Western and Central Pacific Fisheries Commission

SCIENTIFIC COMMITTEE
TWENTIETH REGULAR SESSION

Manila, Philippines
14 – 21 August 2024

Summary Report from the 2024 SPC Pre-assessment Workshop

WCPFC-SC20-2024/SA-IP-01
27 July 2024

Prepared by Paul Hamer[1]

[1] Oceanic Fisheries Programme of the Pacific Community

Report from the SPC Pre-assessment Workshop (PAW), March 25-28th 2024

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 sixteenth PAW was held from the 25th – 28th March 2024. The meeting was held under a hybrid format, with 23 external fisheries scientist and consultants travelling to Noumea, 20 SPC staff and a further approximately 25 people joining online. Sixteen organisations were represented, from across at least 12 countries.

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

- Approaches for the 2024 stock assessments of south Pacific albacore, southwest Pacific striped marlin, Western and Central Pacific Ocean (WCPO) silky shark (phase 2) and WCPO oceanic whitetip shark (phase 1),
- Technical developments in Management Strategy Evaluation (MSE), focussing on challenges with development of MSE and an MP for south Pacific albacore,
- Developments to the MULTIFAN-CL modelling framework in 2023 and the 2024 workplan,
- WCPFC Project 123: scoping study for the next stock assessment platform for WCPFC tuna assessments,
- WCPFC Project 122: scoping study on longline effort creep in the WCPO,
- WCPFC Project 113b: developing a stock status and management advice template for consistent reporting of stock assessment outcomes, uncertainties and risk,
- WCPFC project 90: length weight conversions,
- Progress of the Close Kin Mark Recapture (CKMR) project for south Pacific albacore and an overview of the sampling design work and considerations for CKMR work, and;
- Developments in age validation work and SPC's enhanced capacity in fish aging validation using bomb radiocarbon.

The 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. Six external presentations were provided. The meeting operated under the terms of reference provided in Appendix 3.

This report describes the various presentations made, issues discussed by participants, and 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 and ideas from PAW as they develop the assessments and other research activities discussed. The extent to which suggestions can be explored and/or incorporated into the stock assessments prior to WCPFC SC20 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. We also note that the recent political crisis and related unrest in New Caledonia has had an impact on the efficiency of all SPC staff. Ultimately the final decisions on model development, data inputs, and 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 (i.e. the two shark assessments).

The outcomes of this meeting will be reflected in the various papers submitted to WCPFC, SC20. Copies of presentations prepared by SPC or others can be provided on request from paulh@spc.int.

2024 SPC Pre-assessment workshop, SPC Noumea headquarters.

DAY 1 – 2024 south Pacific albacore tuna stock assessment

Day 1 focused on the stock assessment of south Pacific albacore (*Thunnus alalunga)* being led by Thom Teears (OFP-SPC), working closely with John Hampton (OFP-SPC). This assessment covers the entire south Pacific, thus incorporating fisheries and stock components under the jurisdiction of the WCPFC and the IATTC.

The session began with **Paul Hamer (SPC-OFP**) providing a background on south Pacific albacore fisheries, catches and general biology followed by an overview of the 2021 assessment (Castillo Jordan et al. 2021) and the key recommendations from the SC17 review of that assessment. Following this **Jed Macdonald and Giulia Anderson (OFP-SPC)** provided a presentation on the research on spatial population structure based on otolith (shape analysis) and genetic studies. This provided some background context for later discussion on model spatial structure and movement assumptions. The work was supportive of some level of (at least recent, adaptive loci) isolation of the French Polynesian samples from those collected from New Caledonia, however, further work and expanded sampling is required in the EPO region. There are no samples from further east in the EPO region, so additional sampling, particularly off South America, was recommended to improve the understanding of population structure and isolation by distance.

- China indicated as they have vessels fishing throughout the south Pacific and have started port sampling along with tracking methods for locations of individual captured albacore tuna to trace their point of capture. Thus sample collection with accurate position data is viable and they are keen to explore collaboration and training to collect genetic samples. Collection of otolith samples is however problematic since skippers do not want the fish damaged. **SPC and SHOU to follow-up**.

Following this presentation **Inna Senina (OFP-SPC)** provided a presentation on the previous and ongoing work on the SEAPODYM model for Pacific albacore. She covered:

- The Model Configuration: inputs, fixed parameters, data in the likelihood.
- Validation of the quantitative model: parameters, validation against independent studies, validation scores and existing uncertainties.
- Analysis of model outputs: estimated stock structures and stock size, fishing impact, environmental impacts.
- Climate Projections: projected biomass redistributions and changes in abundance.

The current albacore SEAPODYM for the Pacific is fit to data up until 2010. An updated version of the albacore SEAPODYM model is expected by 2025. The updated version will have higher resolution and longer temporal coverage of ocean forcings, with greater opportunities for more regular updates.

In relation to the current assessment, the albacore SEAPODYM model produces estimates of biomass transfer across model regions at monthly resolution and for different life stages. The transfer of biomass is influenced by environmental forcings and habitat/prey fields. The MFCL model cannot use such information to influence inter—region movement estimation, and there is insufficient tag-recapture data for south Pacific albacore to provide the MFCL model with information on movement, particularly between the WCPO and EPO. In the previous assessment, with its more complex region structure, two options were used for movement in the assessment model grid: M1 – MFCL estimated movement, M2 – SEAPODYM movement co-efficients. The model estimates of management quantities were quite sensitive to these two options. It was argued the SEAPODYM option was likely to be more realistic, but this option gave more pessimistic results for stock status and was down weighted by 50% by the SC17. The SEAPODYM presentation indicated that the predicted movement of albacore biomass between the WCPO and the EPO is actually low (<6-13%). Movement is estimated to be highest from the EPO to the WCPO in the larval stage due to east-west transport via the South Pacific Gyre.

Given that there is no tagging data to constrain the model's capacity to transfer fish biomass among regions, and the simpler model spatial structure for the 2024 assessment, it was proposed to use the transfer co-efficient provided from SEAPODYM as prior information for constraining the movement between WCPO and the EPO for the 2024 assessment.

The SEAPDOYM modelling also provides some insights into potential climate influences on recruitment, in particular, the largest historic larval recruitments predicted by SEAPODYM that occurred in 1982/83 and 1997/98 were associated with switches from very strong eastern (cold tongue) type El Ninos to La Ninas. It was note that the large El Nino/La Nina transition that was observed to happen in 2015/16 might be expected to have resulted in high larval numbers, and subsequent prediction of increased adult abundance and longline catch rates in the early 2020's.

SEAPODYM forecasts to 2050 under climate change scenarios for south Pacific albacore suggest (for the unfished state) declines in availability in the western and central Pacific but increases in the eastern Pacific and limited overall change in biomass across the south Pacific as a whole.

The PAW requested that the sensitivity to the SEAPODYM movement coefficients be explored with alternative levels of movement between the EPO and WCPO, in terms of increases and decreases (including zero movement) compared to the proposed SEAPODYM estimates. Preliminary models run during the PAW suggested the model estimates of management quantities are not very sensitive to alternative levels of movement centred around the SEAPODYM levels, including the zero-movement scenario.

Thom Teears (SPC-OFP) then provided a presentation on development of the fishery definitions and spatial structure for the 2024 assessment. This began with a background on what we know about movement of south Pacific albacore and the inconsistencies between the previous internal MFCL model estimates and those predicted from SEAPODYM. It was proposed that given the lack of information to inform the model on movement that a simplified spatial structure be considered, with one region for the WCPFC-CA and one for the EPO and fleets as areas within each of these regions (Figure 1). Simplification of spatial structure was a recommendation from SC17.

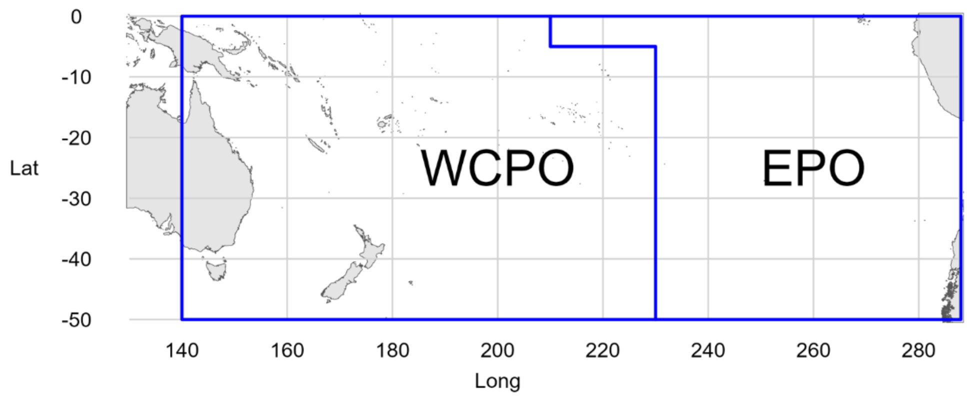

Figure 1. Simplified 2-region spatial structure proposed for the 2024 South Pacific albacore assessment.

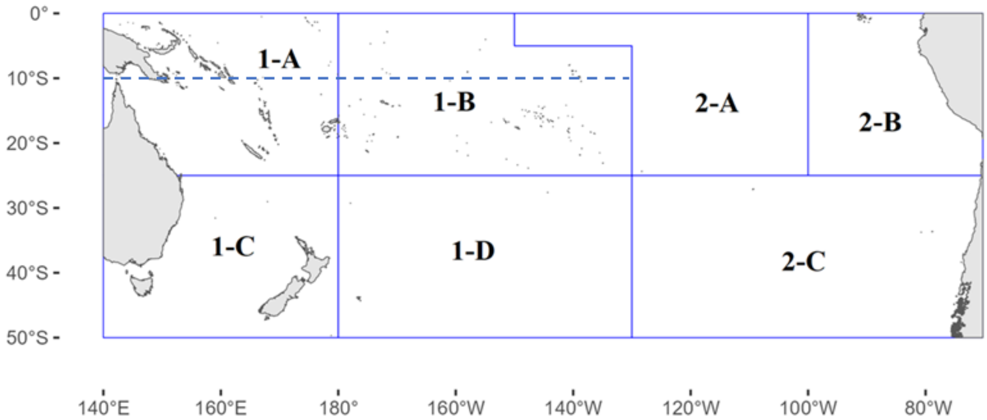

Figure 2. Simplified 2-region spatial structure proposed for the 2024 South Pacific albacore assessment, with sub-divisions for fleet areas, dotted for allocating the tropical longline fisheries.

To inform the fishery structure for the fleet as areas approach, a regression tree analysis of longline length composition data was conducted by **Jo Potts (SPC-OFP)** using previously developed methods that were also applied to the recent yellowfin and bigeye assessments as recommended by the yellowfin peer review [(FishFreqTree Xu 2020)] The results from the analysis were presented and discussed. Based on this analysis a proposed fleet areas structure was presented (Figure 2) along with 14 defined extraction fisheries and 5 potential index fisheries (*although these have been modified since PAW, and a fleet area from 0-10°S has been added to group the tropical longline fisheries that predominantly target yellowfin and bigeye and require separate fisheries for management projections, dotted line in Figure 2*) (Table 1).

Table 1. Proposed fisheries definitions for 2024 South Pacific albacore assessment, refer to Figure 2 for fleet areas/regions.

Fishery Number	Gear	Code	Flag	Fleet Area	Region
1	LL	1.LL.DWFN.1a	DWFN	a	1-A
2	LL	2.LL.PICT.1ab	PICT	a, b	1-AB
3	LL	3.LL.DWFN.1b	DWFN	b	1-B
4	LL	4.LL.AZ.1ab	AU/NZ	a, b	1-AB
5	LL	5.LL.DWFN.1cd	DWFN	c, d	1-CD
6	LL	6.LL.PICT.1cd	PICT	c, d	1-CD
7	LL	7.LL.AZ.1cd	AU/NZ	c, d	1-CD
8	TR	8.TR.ALL.1c	ALL	c	1-C
9	TR	9.TR.ALL.1d	ALL	d	1-D
10	DN	10.DN.ALL.1cd	ALL	c, d	1-CD
11	LL	11.LL.EPO.2a	ALL	a	2-A
12	LL	12.LL.EPO.2b	ALL	b	2-B
13	LL	13.LL.EPO.2c	ALL	c	2-C
14	TR	14.TR.EPO.2abc	ALL	a, b, c	2-ABC
15	LL	15.LL.INDEX.1ab-pre 1990	INDEX	a, b	1-AB
16	LL	16.LL.INDEX.1ab-post 1990	INDEX	a, b	1-AB
17	LL	17.TR.INDEX.1c	INDEX	c	1-C
18	LL	18.LL.INDEX.2abc-pre 1990	INDEX	a, b, c	2-ABC
19	LL	19.LL.INDEX.2abc-post 1990	INDEX	a, b, c	2-ABC

The fleet areas in the WCPO largely reflect a north-south difference in size composition with smaller fish south of 25°S and an east-west separation in the northern area for the DWFN, due to a predominance (unimodal distribution) of larger fish east of 180°, and a bimodal distribution, with a mode of smaller fish to the west of 180°. To south of 25°S, a separate fishery for Australia and NZ is defined to the west of 180°.

For the EPO region data an additional regression tree analysis was conducted by **Haikun Xu (IATTC).** The results were consistent with those from the Pacific wide analysis. The fisheries structure and definitions in Figure 2 and Table 1 thus received general support. There was discussion about the bimodal distributions of the longline size composition data, particularly for DWFNs in the southern region that did not separate spatially upon further splitting (i.e. even at 10 splits) by the regression tree analysis. This bimodality appears related to life history latitudinal differences in occurrence, with small fish occurring in large numbers in the southern region (particularly around NZ and in the Tasman Sea) and larger fish venturing south on a more seasonal basis for feeding, with spawning occurring in more northern sub-tropical/tropical latitudes (Figure 3).

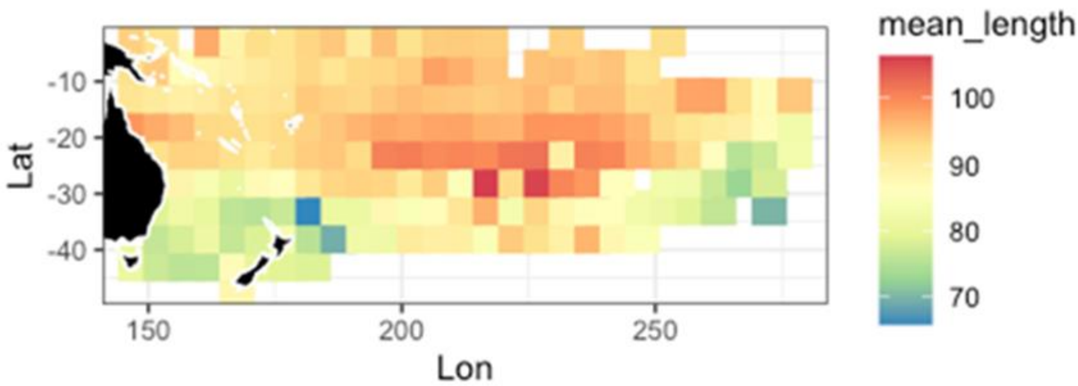

Figure 3. Spatial patterns of mean fork length (cm) for longline caught south Pacific albacore.

There were no objections and general support for the simplified 2-region model and the proposed fleet areas. One question was raised on the alternative fleet areas for the EPO that were suggested in the presentation from Haikun. It was suggested that there would be negligible implications from choosing one or the other, and the choice was to go with the fleet areas that were consistent with the SPC analysis.

Other questions were raised regarding implications of the bimodal size compositions and spatial-temporal dynamics for the development of abundance indices and selectivity. It was noted that the plan was for abundance indices in Region 1 (WCPO) to be structured around the adult spawning region and season – i.e. index of spawner abundance, and the troll fishery, i.e. index of juvenile abundance, with a single adult abundance index for the EPO (all Region 2) (discussed further in CPUE below). In the simplified model there is no north-south movement to estimate so it was noted that seasonal selectivity can be more flexibly applied for the extraction fisheries to account for seasonal difference in availability of different size fish to the longline fisheries in the different area groupings.

There was a query on the use of seasonal partitioning in the regression analysis to see if compositions would split by quarter, however it was replied that this can lead highly complex spatial groups with holes discontinuities that soon become impractical.

The grouping of AU/NZ longline fishery was questioned on the basis of typically higher HBF (i.e. 20-30 HBF) and deeper sets by AU than NZ (i.e. 10-15 HBF). While it was suggested to keep the grouping for simplicity, it was recommended to check the residuals to see if this could be causing problems.

The need to implement a fleet area for longline fishing north of 10°S to the equator (tropical longline) was raised as necessary for Management Strategy Evaluation and projections to allow separate management scenarios for the tropical longline fishery. This was noted to be implemented in the assessment with the assumption of shared selectivity with the LL fleets in areas 1a and 1b

The presentations/discussion on the individual fishery data and size composition data treatment was moved to the next day**, and the next session focussed on the CPUE analyses** to produce the abundance indices. There were five presentations in this session.

The **first presentation by Yi-Jay Chang from the Institute of Oceanography, National Taiwan University**, presented an analysis of the albacore CPUE from the Chinese Taipei distant water fleet (TW-DWF) in the south Pacific. The presentation included discussion on raw catch and effort data, spatial patterns of length frequency and average weight, and a CPUE standardization using a spatio-temporal model (VAST). The catch of south Pacific albacore by the TW-DWF has range from around 2,000 - 10,000 mt since the 1960’s and around 3,000 – 6,000 mt since 2010. The key change observed for this fleet overtime is the reported effort in numbers of hooks, which prior to year 2000 was around 10– 20 million hooks per year, but after 2000 increased sharply and has ranged between around 30-40 million hooks per year since 2010. Most of the TW-DWF catch is now taken from the southern (south of 25°S) region of the WCPO and mostly in quarters 2 and 3. CPUE is highest in the region south of 25°S. Very little data is reported from eastern region of EPO. It was noted that after 2002 there was increased targeting of bigeye and yellowfin with higher HBF, that corresponded to a drop in albacore catches. The quality of the size data prior to 2000 is less certain, however there is a decrease in the proportion of larger albacore and the mean weight since the early 2000s. Consistent with the earlier presentations, mean weight is larger to the north of 25°S. A regression tree analysis (using FishFreqTree Xu 2020) was also conducted on the TW-DWF weight composition data, using average weights (catch weight/number) rather than length distribution as in the previous mentioned SPC analysis, across the entire South Pacific. The results of this analysis were broadly consistent with the proposed fishery area structure presented by SPC earlier, including the split of the southern and northern area at 25°S, and a longitudinal split at around 180°E-170°W. The CPUE standardisation in VAST used data from 2000 for a subset of albacore 'target' data (nominal CPUE 15-30 albacore/1000 hooks). The indices showed the typical strong seasonal variation, particularly in the southern area, with a stable longer-term trend, but suggestion of an overall drop in CPUE from around 2016-17, with some sign of increase in recent years. The indices were generally consistent with the indices used in the previous SPC south Pacific albacore assessment.

Bimodal weight distribution in the southern regions was again discussed by the PAW, noting that in terms of the weight data it was more prevalent to the east of NZ in the high seas, with a more unimodal distribution centred on lower weight fish for the NZ/Tasman Sea/southern AU region. It was suggested that this may indicate differences between flags in the spatial patterns. SPC indicated they could explore this further at flag level, noting the grouping of all distant water flags in the analysis of length comp data. There is a split proposed at 180°E but it is not proposed the split the distant water fleet, more so to separate the AU/NZ fleet in this region. Also, there might be differences that relate to use of length versus weight data, and that there is still bimodal distribution for the region east of NZ in the weight data.

The PAW asked several questions regarding whether the differences in the weight distributions might relate to different gear settings, and also why the TW-DWF doesn't fish much in the southern region in quarters 1 and 4 when the earlier SEAPODYM presentation suggested that adult fish do occur in that region in these quarters. It was replied that as the data is already subset to be albacore target sets, the HBF variation would be minimal, and the differences are not likely driven by gear setting differences. It was suggested to confirm the consistent gear (HBF) settings. The vessels that target south Pacific albacore switch to target albacore in the north Pacific in quarters 1 and 4, and the vessel targeting tropical tuna (bigeye and yellowfin) tend to continuously fish in the tropical region.

It was noted the HBF was only available from 1995, the catch data since 1997 had received additional cross checking, and VMS was available since 2005, so the Chinese Taipei data from 2000 onwards is more certain and has more information.

The second presentation by Hongyu Lin from Shanghai Ocean University, covered an analysis of the Chinese fleet's longline albacore CPUE in the south Pacific. The Chinese data is available from 1990, and Chinese vessels became active in the south Pacific albacore fishery from the early 2000's. The nominal CPUE is around 10-25 albacore/1000 hooks and has gradually increased over the last decade, with generally higher CPUE in the region south of 25°S. Cluster analysis clearly identified albacore targeting sets with mean CPUE of 18.86 albacore/1000 hooks and 88% of catches being albacore, covering most of the longline fishing south of 10°S, and in the EPO region. The CPUE analysis is work in progress and has explored GLM, GLMM with spatial random effects, and a spatiotemporal model in sdmTMB. The results support the continuation with the sdmTMB model exploring temporal autocorrelation, alternative spatial modelling options, barrier mesh structure, and additional covariates including environmental and effort creep related. Standardised CPUE showed slight increasing recent trends for all methods, but with declines in the very recent period.

Questions were raised around the data filtering options, and these were responded, noting around 50% of the data were filtered out, data from 2001 – 2009 was limited, so analysis focussed on data from 2010 onwards.

There was a request from the PAW for more information on the targeting patterns and general operational features of the Chinese fleet, in particular whether they show seasonal shifts in targeting between albacore and tropical tunas, move between the north and south Pacific albacore stocks etc. **Chinese delegates indicted they would follow-up and find out more detail on operational behaviours of their longline fleet in the south Pacific.**

The **third presentation by Phil Neubauer from Dragonfly Data Science covered analysis of the NZ troll fishery size composition and CPUE data.** The troll fishery catches small, predominantly immature, albacore in the waters around NZ, noting that albacore are not a quota managed species in NZ so the fishery has a mix of higher catch trips from targeted troll activity and smaller opportunistic troll catches on trips mainly targeted at other species. NZ accounts for 77% of troll caught albacore in the south Pacific so the data from this fishery has potential to provide information on juvenile abundance/recruitment. The fishery took peak catches up to 6,000 mt in the 1990's and early 2000's, stabilising at around 2,000-3,000 mt over the last decade. Most of the albacore taken in NZ are now by troll. Change to E-Reporting from 2020. Catches is mostly from the northwest coast of the North and South Islands, with highest catches along the northern half of the South Island, and similar CPUE across the North and South Islands. The fishery is seasonal, peaking in January-March (late summer/early autumn). CPUE is fairly consistent from December to April. Previous assessment used the NZ troll fishery catch and size data but not CPUE. The previous assessment appeared to show a missing cohort from 2016, so exploring the troll data might provide some insights into whether this was potentially due to a sampling artifact.

The analysis of size data applied a three category mixture model to try and explain the relative abundance of cohorts based on explanatory variables including sampling areas and times, and fishing trips sampled etc. They also tried a multinomial standardisation. The mixture model could not fit the data that well, but indicated that depending on the fishing year you could observe different cohort patterns in different areas and months, so there was no obvious consistency in space or time. The multinomial standardisation aimed to explore which factors influenced the size composition variation (using BRMS), the full model was best and the most important explanatory variability was still year, but year-month-area interactions and vessel ID effects were also important. The model fitted well across years, and the uncertainty indicated that the sampling was representative of the catches. The interpretation of the relative cohort strengths was however influenced by whether data is displayed as proportions or catches at length. If viewing data as proportions patterns are observed that might suggest weaker and stronger cohorts, and that suggest weaker groups of 2 year old (60 cm) fish in 2015 and 2016 but when these are converted to catch at length the patterns are not as clear. In years with low catches clear age class structure is not evident. However, the groups of 2 years olds in both 2016 and 2015 still appeared similarly with low in catches, which may suggest lower recruitment emanating from spawning in summer 2013/14 and 2012/13.

The presentation by Phil continued with the CPUE analysis using a GLM with and without environmental variables (SST and ENSO). Two analyses were done, one used catch weight per trip data and the other catch numbers per day. Lots of vessels have records of albacore (approx. 1000 vessels) but many have low catches and/or numbers of records and are not likely informative and so were filtered out. The analysis of the catch number per day (1992-2023) predicted a highly variable abundance over time with a stable trend from the early 1990s until around 2010 after which a decline trend is apparent, with 2022 and 2023 being the lowest levels since 1992. The standardisation had minimal effect on the nominal CPUE. Much of the CPUE decline appeared related the northern areas, with more stable CPUE for the main catch area along the west of the South Island. For the trip catch weight model the standardisation had a more notable adjustment of the nominal CPUE, increasing the overall levels prior to 2004 and decreasing them from 2004 onwards. But the standardised and nominal CPUE finished at similar historical low levels in 2023, both declined from 2021 (consistent with the catch numbers analysis), and the overall trend in the standardised CPUE was a decline across the time series. There did not appear to be any influence of SST on the CPUE variation, and while the ENSO index did have an effect in reducing the overall levels, it had limited influence on the trend but reduced the magnitude of the peaks, and with no impact on the CPUE in the final year. Overall, the trip catch weight and daily numbers CPUE series diverged in 2016 due to variable size compositions especially in periods of low CPUE due to absence of strong cohorts of smaller fish and the analysis produced similar results to previous analysis. There remains uncertainty as to how much of the CPUE variation is related to local availability/movement behaviour versus year class strengths, and it is still thought that the CPUE is driven by both year class strength and movement patterns, influenced by environmental/feeding conditions.

In terms of the assessment, it was suggested to consider using the troll index with different levels of weighting, i.e. high weight indicates it represents a strong year class signal, low weight more related to movement.

The PAW noted the low troll CPUE in 2015/16 was consistent with the lowest recruitment estimates from the 2021 stock assessment that did not use the NZ troll CPUE. The influence of the CPUE on either recruitment or migration was discussed. It was noted that the recruitment of albacore in the east part of the spawning region is predicted by the SEAPODYM model to be higher in El Nino, whereas in the western part it is higher in La Nina, so the patterns of recruitment in relation to ENSO are complex. It was noted the trial MFCL models provide reasonable fits to the NZ troll CPUE.

The question was asked about the consideration of lags between ENSO effects on troll CPUE and larval recruitment/spawning success. Ignoring a lag would suggest ENSO effect is more related to availability. It was responded that this could be the case, but in some years the CPUE signal is driven by the 1 year olds. Worth to try further analysis with lags of 1 and 2 years between the CPUE and ENSO. It was also noted that if we accept that there is a stock level abundance signal in the troll CPUE, it is useful to have it in the assessment to provide information on the population scale. Further, by exploring how well the model can fit the troll CPUE and other data (i.e. internal consistency) might provide information on how much the troll data is indicative of local availability versus cohort abundance.

The recent decline in the troll CPUE was noted as a concern but again it is uncertain how much the decline is indicative of a real recruitment decline versus local availability. The PAW asked if the analysis will be submitted to the SC20 given the assessment data will run to 2022 and have limited information on recent recruitment levels. **The NZ troll CPUE and size data analysis paper will be provided to the SC20 by the NZ delegation.** The PAW again emphasised the need to do more exploration to understand what is driving the troll CPUE index, noting the potential importance of the index, for example looking at size data from other parts of the south Pacific albacore fishery to see if smaller fish turned up, for example, in longline compositions in years when there was a lack of fish for the troll fishery around NZ. Phil supported a broader analysis of length composition data from the troll fishery and surrounding fishing areas for longline and observer data. It was noted that a pole and line CPUE juvenile abundance index is used in the north Pacific albacore assessment as a sensitivity. For Pacific blue fin a restricted range juvenile index was actually used in the assessment. The assessment team will explore the use of the troll index in the assessment and look for internal data consistency, and future work should look at data from other fishery components to see if there are indications of movement or spatial variation in occurrence of smaller fish between NZ and adjacent areas that may be influence the troll CPUE dynamics.

The next presentation **by Thom Teears (SPC-OFP) covered development of CPUE indices for the assessment from the multi-fleet data set**. The presentation covered catch, effort, and CPUE spatial summaries, the replication of the 2021 CPUE standardisation and the switch to sdmTMB, an exploration of covariates of vessel ID, targeting cluster, seasonal catchability effects, and the restriction of the WCPO index to a northern region (10-25°C) index to represent spawner abundance. The work on the CPUE analysis was quite advanced.

With regards to the vessel ID being included, this was explored in detail. However, there are over 4,000 vessel IDs in the dataset which is computationally not feasible, so a subsampling method was developed that included consideration of thresholds for sampling coverage, where vessels were subsampled randomly so as new random selections increased the sampling coverage. A threshold for the subsampling was that each 'year-quarter-5 x 5 cell' had at least three observations. This subsampling was conducted 4 times to create 4 data sets that accounted for about 18-20% of the data, each with 800 vessels. The comparisons of the models for each of these data sets indicated that the results and importance of particular covariates were quite sensitive to the particular sample of vessels. This was unsatisfactory and the application of a vessel ID in the standardisation still seems problematic and impractical given the size of the data set.

There was concern that targeting differences and related size selectivity differences may be influencing differences in the observed CPUE patterns among flags. Exploration of spatial size composition data indicated that the size compositions where mostly consistent across flags when they fished in the same areas at the same time, which indicated that at least size selectivity was consistent across flags. The differences in CPUE among flags there likely related other aspects such as seasonally shifts in fishing areas and targeting. This variability in seasonal fishing areas and targeting seems important to consider for developing CPUE indices. Of note was the targeting variability north on 10°S (tropical longline region), where flags will switch between yellowfin/bigeye and albacore – depending on prices and local availability. Whereas in the more southern area the albacore target longline fishery tends to focus most of its fishing in quarters 2 and 3 (as discussed previously for the Chinese Taipei analysis), which creates strong seasonality in the CPUE, partly due to the lower effort, but also the lower availability in the southern areas during quarters 3 and 4 as larger albacore move north to spawn in those months. While CPUE models were run with seasonal effects and HBF (removing the species cluster) that did smooth out the seasonal variation, and CPUE annual variation was positively correlated among seasons, exploration of seasonal patterns in spatial size distributions indicated the strong concentration of larger albacore between 10 – 25°S in quarters 1 and 4. In order to simplify the processes driving the variation in the CPUE, it was proposed to create an abundance index for the spawning component of the population. This spawner index was created by re-running the sdmTMB with HBF but only for the WCPO region and quarter 4 and aggregating the CPUE predictions for the area between 10 – 25°S. This has the implication of changing the annual time period to start from October 1st and end September 30th, with a spawning period at the start of each year between October and December. The resultant annual spawner abundance index was presented and discussed, including some diagnostics for the model performance.

The PAW commented about the issues of confounding of species abundance and species clustering/targeting and suggested running the species cluster analysis in time blocks rather than across all years at once. This was considered a good suggestion but there were still concerns about seasonal patterns of targeting that appear predominant. It is difficult to deal with the issues of targeting, movement and availability with a quarterly index.

The PAW noted a similar spatially constrained LL based spawning index was being applied in the North Pacific albacore assessments, so what is proposed for south Pacific albacore is similar. The spawning index proposed does have a decline at the end which will likely have some impact. The PAW was asked to consider this spawning index proposal. The PAW noted that the north Pacific spawner index dropped to a low level in 2021 but came back to a high level in 2021, which did not seem plausible in relation to stock abundance. One thing to be wary off with a spatially restricted spawning index is that spawning grounds can move around depending on the oceanographic conditions. There is a plan to model the north Pacific albacore spawning index over an expanded area. This could be considered for the south Pacific albacore spawner index.

PAW found the flag differences in LF interesting, noting the three modes in the Japanese data for the region adjacent to New Caledonia. This type of analysis could identify areas/flags to focus on in relation to interpretation of the troll index, i.e. do small fish turn up in other areas/fisheries when troll CPUE drops around NZ for example.

It was noted that some of the unexpected difference between size compositions for Japan and Korea were likely due to low sample sizes. Overall, despite the 100's of size composition comparisons between flags, very few were different so for the most part the LL contact selectivity appeared very similar among flags for south Pacific albacore.

Day 2 South Pacific albacore assessment continued, Western and Central Pacific silky shark assessment

Day 2 started with a recap of day 1 from the chair, and then moved onto discussion of the **pre-treatment of size composition data for the south Pacific albacore assessment, presented by Tom Peatman**. The issue of rounding of size data to 2 and 5 cm bins and subsequent contamination of 1 cm bin data with rounded data was discussed along with the approach to test/detect samples contaminated with rounding (for trip level samples with 100 or more size data). At the trip level, this analysis identified that trips constituting around 12% of the data were contaminated with rounded data. The tests were also performed for size data that had been aggregated. About 20% of samples were contaminated with 2cm rounding (both odd and even rounding occurred) and less than 1% with 5 cm rounding. Some challenges with this type of rounding detection statistical method were noted, in particular, the choice of p-values for indicating the overrepresentation of length bins in the aggregated data, a p-value of 0.01 was recommended as a good compromise to detect clearly contaminated samples while minimising false positives. The presentation ended with a summary of options to deal with the rounding contamination. It was recommended that preparing the data according to 2cm even-odd length bins e.g. 74-75 cm, 76-77 cm etc. (as opposed to 1cm as previous assessment) would retain the most data, and would exclude 2 cm odd-even rounded data, e.g. 73-74 cm, 75-76cm etc.

It was noted by the PAW that this would not be an issue if the data was reported with accurate details of the measurement resolution, i.e. if rounding is conducted this needs to be clearly noted with the data provision. But preferably everyone should aim to measure to the same resolution and or rounding method. The is room for improvement here.

The PAW was interested in how the length data will be used in the model as this could influence the approach to taken for length bins. It was noted that the use of the size data will not be restricted to particular parameters and would be used in a general sense as typically done. It was noted the sensitivity exploration with 1cm and 2cm size bins showed negligible difference, and so 2cm bins is the approach expected to be used. This also allows data that is reported in 2cm bins (as is common) to be included, as this data would have been previously excluded at a 1 cm bin resolution. The recommendation was described again as 2 cm bins with even-odd structure. The PAW did not suggest an alternative and appeared satisfied to continue with this approach.

Tom Peatman then continued to describe the now standard approach for reweighting the size composition data to account for variation in sampling rates and potential biases. This method was applied in the last south Pacific albacore assessment, where bye the extraction fisheries size compositions are reweighted by catch, so that samples from strata with higher catches have more influence on the extraction fishery size compositions used in the assessment. Whereas index fisheries size compositions are reweighted with CPUE so that strata with higher CPUE (i.e. abundance) have more influence on the index fishery size compositions used in the assessment. The technical methods were described in detail.

It was noted that the time blocking of the index fisheries that was suggested earlier was supported by size composition data as the start of the PICT size composition was in the mid-1990s, whereas the DWFN size composition data went back to 1965, with earlier years being dominated by data provided by Japan. There was a suggestion about going to 5cm bins, but this would mean not being able to estimate growth or fit the size compositions of the smaller fish, including the troll. Also, there is very little contamination by the 5cm data (0.2%) so it is considered better to just remove this 5 cm data.

Allen Andrews from SPC-OFP then provided a presentation on age validation of tuna using post-peak bomb radiocarbon dating and the developing capability for otolith research, especially age validation, at SPC. Background on the theory and methods for post-peak bomb radiocarbon dating was provided, with examples for giant trevally around Hawaii, yellowfin tuna in the Gulf of Mexico, and the recent work on yellowfin and bigeye tuna in the western Pacific. The later work involves the use of otoliths for 0-age juveniles where the year of birth is known, and then extracting core (0-age) otolith material by micro mill for older individuals. The 14C values for the 0-age samples were consistent with coral reference datasets and the analyses of the older otolith supported the interpretation of increments in otoliths being deposited annually up to at least 14 years for yellowfin and 13 years for bigeye. The analysis appeared to indicate some under aging of yellowfin for younger ages may be occurring. The method has been applied to albacore in the Atlantic confirming annual ages from 3-17 years. The SPC Sclerochronology Lab is now working on applying the approach to South Pacific albacore, extending the coral records for years beyond the 2010s, work on skipjack, including to confirm ages of large fish samples, work on billfish, and work on deep water snappers. A key point was made that post-peak decline 14C is not as precise as the rise period for determining individual ages, so it not used as an aging method for individual fish, but more so for validation of aging from the central tendency of large samples.

The next session focused on biological aspects, specifically growth and natural mortality, maturity/reproduction and length/weight conversion. The first presentation was from **Donqui Lu from the Shanghai Ocean University, China, on an exploration of spatial variation in growth of albacore across the south Pacific.** The work involved aging of caudal vertebrae samples collected from landings in seven Chinese ports. The samples were caught in three regions or the South Pacific; a western region close to Papua New Guinea and Solomon Islands, a central region around Kiribati and an eastern region in the high seas to the east of French Polynesia. The samples were from fish between 85-102 cm FL. Analysis considered spatial differences in weight at length and length at age, and weight at age. The age range of samples was 2-8 years for the eastern region, but only 4-8 years for the other regions. There were differences in weight at length with heavier fish at length in the western then central then eastern Pacific. There was also a significant different between the growth for western and eastern Pacific samples across the 4-8 years age range, with faster growth of eastern Pacific samples, although the narrow age range and samples size was noted as a caveat to these results and more sampling is required.

The PAW asked about validation of the vertebrae aging method, which requires some more consideration and could be worth comparing to otolith-based growth increment data. **The need for more samples of tissues, otoliths, reproductive biology etc. from the EPO region was raised and the delegates from Shanghai Ocean University expressed their interest to get more involved with sample collection and provision, and that they are building there working relationships with their industry. This is to be followed up with discussions between Shanghai Ocean University (Fan Zang) and SPC staff.**

The next **presentation was from Thom Teears (SPC-OFP), focussing on the consideration of age and growth for the assessment**. The presentation provided a recap of the age-length information available for the previous assessment and the approach to generating the two external growth options that were used in the previous assessment. The two growth curve options (external from otoliths, internally estimated from length data but applied as external input) for the previous assessment were combined with M-at-age to create two growth/M-at-age options for the uncertainty characterisation. This new assessment is aiming to develop a single growth model that fits the troll size modes and the otolith age at length data. I an initial attempt to achieve a growth curve that was more consistent with both the troll length compositions modes and the otoliths a growth model was developed internal to MFCL by fixing L1 and L12 based on the otolith length at age data, allowing the model to estimate K, determine the SD based on the length distribution of the first age class on October 1 (0.75 years, 2.25 cm) and the oldest age class (12 years, ~4 cm).

Fitting an internal growth curve following this approach provided a good fit to the troll length modes, but a degraded fit to the external otolith data. Recollecting that the best model fit to the otolith data was a logistic rather than a VB, but that a logistic growth curve is not currently possible in MFCL. Other options could be further explored such as the VB offsets in MFCL to estimate an internal growth curve that fits both the troll size modes and the external otolith length at age data set. *Note this has since been improved further, see Post PAW updates.*

Discussion by the PAW touched on issues such as temporal variation in growth, particularly the NZ troll size modes that show variability from year to year, which could be growth related or due to the variable time periods during the spawning season that produced recruits sampled off NZ. A single growth relationship might struggle to fit all the data, but it is not possible to have time varying growth in the age structured MFCL model. The use of the otolith data as CAAL data in the model was discussed, noting this was the recommended approach by the yellowfin peer review. I was noted this approach is typically preferred, because external length at age from otolith samples does not take into account fishery selectivity and the length-based sampling process for choosing otoliths to age, such external growth curves can therefore have considerable bias. Inputting the otolith data as CAAL can be explored. The approach of estimating the growth internally from the length data with the fixed L1 and L12, is a kind of compromise, as the selectivity should be accounted for in the length compositions. It was also noted the more flexible growth curves would be useful to explore, i.e. Richards, and perhaps if not available in MFCL perhaps try estimating the growth in a simplified SS3 model. It was noted that MFCL has a Richards growth curve and that this was applied but did not produce any improvement. There was some discussion on factors that might be influence the early growth based on the otoliths and that samples of smaller albacore (< 40 cm FL) would be useful to improve confidence in the early part of the otolith-based growth curve. It was also noted that it appeared from model explorations so far that the first age class is not fully selected by the troll fishery and that the samples of the age class are not likely to represent the real spread of lengths at age 1. Given L1 is subject to selectivity and is fixed to the otoliths data set, this might be a sensitivity. It was responded that we are trying to be conservative with what we try to estimate in model, growth information from length data is only really there for troll size, so we might need to make some fixed assumption about L12. Finally, the PAW asked what the plan would be, would there be a CAAL model approach perhaps with adjusting weighting between fitting troll data well versus fitting CAAL data well. It was responded that the approach is still uncertain, and we need to do a bit more exploration of CAAL.

Post-PAW *we explored applying the VB offsets to early growth phase ages 2, 3, and 4 years in MFCL to approximate a more linear growth in the juvenile period, more consistent with a logistic. The goal is to produce a final growth curve (estimated internally to MFCL with some fixed parameters based on the otolith data) that fits both the troll size modes and the otolith age-length data well. This approach has achieved a satisfactory growth curve using the VB offsets for ages 2, 3 and 4 years, and the otolith data used in the model as CAAL.*

The next **presentation was provided by Thom Teears (SPC-OFP)** which outlined the M-at-age approach used in the previous assessment (Maunder et al.) that depended on growth parameters. As two growth alternatives were used this resulted in two combined growth/M-at-age options. The sensitivity of the previous assessment to these alternatives was noted, but as growth and M are combined the influences of growth versus M could not be separated. The current assessment is considering two options the previous Maunder et al. method with the revised growth curve, and a simpler Lorenzen M-at-age, using a life-history mete-analysis to inform the average M, which could be applied as a prior distribution to use in uncertainty characterisation. For spawning biomass calculation and maturity, the approach for the previous assessment was described and it was noted that there was no new information on reproductive biology that would warrant altering the approach from the last assessment.

Jed Macdonald (SPC-OFP) provided a presentation to update of the progress of WCPFC project 90 – length-weight conversions. He noted that for south Pacific albacore the same L-W conversion has been used since 2005. Since that time many more (several 100 thousand) L-W samples have been collected, and it was time to update the conversion factors. The presentation described the data coverage and presented an updated L-W conversion equation, which was slightly different to the previous equation. The presentation went on to discuss available information for striped marlin, which include L-W and L-L as there are different approaches used to measuring length of striped marlin, and finished with a general update on progress and next steps for project 90, including the collaborative work to improve data on gilled and gutted to wet weight conversions for longline bigeye landed in ports. It was noted that there is still data to come in from recreational sources in Australia and New Zealand for striped marlin that could contribute to updating the conversion factors for that assessment. **Jed noted some more data to come and he would provide the updated conversion factors, and follow up with the Chinese delegates on possible collaboration with port sampling at Chinese ports.**

The next session involved a **presentation from John Hampton (SPC-OFP) on the progress with developing the diagnostic case model for the south Pacific albacore assessment.** The initial work has involved converting the 2021 diagnostic case model to a catch conditioned model with a CPUE likelihood added, adjustment to longline selectivities, collapsing the previous three WCPO regions to a single region (with single EPO region as previous) with fleets as areas approach. The biggest impacts on the depletion from these changes appear related to the addition of the CPUE likelihood as part of the catch conditioned model. The proposed two regions (WCPO and EPO) and fishery definitions/fleet areas was presented. The working diagnostic model at the time of PAW was outlined as below (*but see Post PAW Updates*):

Data

- 14 extraction fisheries defined by area and fleet
- A "spawning biomass" index 10-25S, Oct-Jan from the LL CPUE model
- A NZ troll fishery index (provided by Phil at Dragonfly)
- An EPO longline index
- All indices are annual, LL indices split at 1990 (gear change to mono mainline likely to affect catchability and selectivity)
- Size data for extraction fisheries aggregated with re-weighting based on catch distribution
- Size data for LL index fisheries aggregated with re-weighting based on CPUE distribution

Model structure

- Two regions, WCPFC and EPO, movement by age-class and season fixed at SEAPODYM estimates
- Sex aggregated, 12 annual age classes
- Annual recruitment occurring in October, fixed distribution to regions based on SEAPODYM (82:18)

- Time window is 1954-2022, virgin population assumed in 1954; results compiled for 1965-2022
- Lorenzen natural mortality, asymptotic value of 0.3 per year
- Growth is VB, fixed L1 and L12 parameters, k estimated
- Mean length variance parameters fixed at "sensible" values
- Selectivity is cubic spline (3 nodes), time blocks at 1990 (gear change to mono mainline) for DWFN fleets, seasonal for LL fleets operating year-round in tropical area
- 1 index (WCPFC LL post-1990) assumed to have non-decreasing selectivity with age
- CPUE indices weighted according to their CVs from CPUE model – approx 0.16 for WCPFC LL, 0.70 for EPO LL. NZ troll arbitrarily assigned CV of 0.2
- Size data generally down-weighted, max ESS of 1% of OSS for capture fisheries and EPO indices, and 5% for WCPFC LL index fisheries.

Some key model results were presented; the Lorenzen M-at-age, growth curve, selectivity patterns, recruitment, spawning biomass, spawning biomass depletion, and key diagnostics. The model presented had 198 parameters, 126 selectivity, 69 recruitment deviates, 1 population scaling, 1 SRR and 1 growth, a positive definite Hessian, maximum parameter gradient 9e-07. The only large parameter correlations occurred for selectivity triplets in the 3 node cubic splines. Likelihood profiles showed that CPUE favoured a higher population scale than the size composition. Retrospectives with 5 annual peels showed no major retrospective patterns. Age structured production and catch curve models were also shown along with fits to the CPUE indices which were reasonable. The CPUE residuals showed some trends in residual patterns, most notably for the EPO index pre 1990 with under estimation in the early years. Fits to aggregated size composition showed good fits to the troll size modes (that could not be matched using otolith-based VB parameters), and good fits to the longline fisheries compositions in the north, but not as good in the southern fisheries with the multimodal compositions. Because the longline CPUE indices are annual their selectivity was estimated separately rather than sharing the selectivity with the extraction fisheries that have seasonal selectivity.

The PAW discussed the results on the model development with emphasis on the CPUE and size composition fits and data weighting. The model does not provide adequate fits to the size composition data for the longline fisheries in the southern areas with bimodal size distribution. This requires further work to improve the fits by either considering further refinements to fisheries definitions, selectivity, or data weighting. There were questions regarding the difference in the selectivity pattern between the time blocks, why there is an asymptotic selectivity pattern in the early time block for season 2 but then becomes dome shaped in the latter time block. It was noted that selectivity is being used to explain the lack of larger fish in the catches in the second time block which could be biasing the estimation. It was also noted that the current model was estimating selectivity patterns in a rather unconstrained manner, so this requires some more thought and consider forcing an asymptotic selectivity in the latter time block also. The need for jittering given the correlations between selectivity parameters was raised. The issue of effort creep was also raised in relation to the CPUE trends. Longline effort creep is being considered by a new WCPFC project, and was raised as a potential sensitivity analysis. To a certain extent time blocking the CPUE allows the model to adjust catchability which may account for influences of gear changes.

The discussion then focussed on the characterisation of uncertainty. The areas of uncertainty that are being focussed on are: steepness, CPUE, data weighting, movement, growth and natural mortality.

Note: work since PAW further considered the time blocks, growth and CPUE. As to the time blocks consideration of historical data on the distribution of HBF indicates that there are changes in fishing practices across these three time periods: 1954-1976, 1977-1993, 1994-2022. Hence these periods are being explored as time blocks for the CPUE and selectivity. The CPUE models were rerun with these periods as categorical factors. For growth, a much-improved growth curve was developed using the VB offsets in MFCL and the CAAL data included.

Western and Central Pacific silky shark assessment

Phil Neubauer presented a summary of the previous silky shark assessments, the work done in 2023 (phase 1, Neubauer et al. (2023)) and the plans for the assessment modelling and the associated risk assessment in 2024 (phase 2). All previous silky shark assessments have had very high uncertainty (i.e. Clarke et al. (2018ab)) but have tended to support that overfishing was occurring. One of the key issues with the previous assessment was the catch history and the use of the shark fin trade data, which on more recent analysis is considered unreliable for catch estimates and not recommended. The presentation provided a summary of the data preparation work conducted in 2023, which included catch reconstruction, CPUE and size compositions.

Catch reconstruction methods and results were provided to SC in 2023, a recap on the approach was provided. The approach used a novel multi-model approach that was an enhancement of the approach applied to the recent southwest Pacific mako shark assessment. Observer catches (interactions) were estimated from models developed to predict observer interactions data, then scaled to overall predicted interactions using the effort data in the L-BEST dataset. These estimates were then scaled by estimates of live discards based on observer discard and condition information, as well as by post-release mortality estimates. The approach uses a model-based weighting approach to account for the fact that observer coverage is variable/patchy in space, time and by fleet. A single model cannot adequately predict interactions equally well across space/time/flag strata. This approach applies model stacking using flag, year, latitude and set-type (for purse-seine) as covariates, and produces corresponding predictions of interactions over the full effort data. The model stacking essentially provides higher weight to models that have higher prediction accuracy for particular year/latitude/flag strata. The approach was applied separately for the longline and purse-seine dataset albeit with different model sets.

The new catch reconstruction goes back to 1995 and was reasonably consistent with the previous catch reconstruction by Peatman et al. (2018) for the period of overlap (2003-2018). Interactions in the longline fishery peaked at around 400,000 individuals in 2012 but have declined since then to less than 100,000 individuals per year. An analysis was also run with no year effects in the predictions, which still produced a similar recent downward trend supporting the conclusion that the recent declines were unlikely to have been solely driven by reporting rate changes due to implementing non-retention measures. Silky shark interactions are concentrated in the tropical equatorial region and more so to the west off PNG.

For purse seine, with much more observer data, less spatial-temporal extrapolation was required and there is less uncertainty on the catch predictions. Catches are also highest along the equator and, especially the western equatorial region off PNG. Comparisons with previous reconstructions showed similar catch levels, but the new reconstruction showed a greater increase in catches over the last decade, recently estimated at around 50,000 individuals per year for FAD sets and a similar 50,000 for free-school sets. The recent increasing trend in purse seine catches for both set types contrasts with the decreasing trend for longline.

Catch reconstruction discussion by the PAW noted the positive advancements in the multi-model weighting methods. The biggest decline in LL was from 2013-2104 when the non-retention comes in, but this corresponds to a large reduction in shallow sets that was also noted to have occurred at that time, so the drop in interactions was more likely due to a reduction catchability for the deeper sets.

CPUE was discussed next. Operational data for CPUE from both longline and purse seine is considered too sparse prior to 2012 to be useful, so the decision was to focus on observer CPUE for the abundance indices. For the longline four subsets of observer programmes were considered, all observer data, south Pacific, long running programs, and distant water fleets (DWF). The analysis showed no consistent trends in the standardised CPUE, and essentially flat trends since the late 1990s, with high uncertainty. This result was not surprising as previous analysis also found similar inconsistencies. It is unclear why there are such flat trends and inconsistencies in long CPUE among different observer programs (e.g. low observer coverage etc.) and this could warrant some further exploration, especially in relation to impacts on catch reconstructions.

CPUE for purse seine however did show good consistency among different observer programs, and all showed a consistent increasing trend in CPUE since 2010, when high observer coverage was mandated. While purse seine CPUE for target tuna is problematic as an abundance index, the rate at which silky shark associate with tuna schools is likely more related to abundance, so purse seine CPUE may be more reliable for non-target species such as silky shark.

Length composition data preparation was then discussed. Similar methods were applied as for the NZ albacore troll data discussed earlier using a model-based approach to scale the length compositions to the removals. The approach also provides uncertainty for each length bin. The presence of a mode of larger individuals emerged from 2010 in the purse seine, and in the longline despite the inconsistencies in the CPUE. This would be consistent with the increasing abundance trends observed in the purse seine CPUE.

In summary:

- Reasonable evidence that longline interactions have declined substantially from up to 500k individuals per year, to <100k individuals per year.
- Purse-seine interactions increased since 2012, now higher than longline at ~150k individuals per year.
- Longline CPUE highly variable and inconsistent among observer programs, but purse-seine CPUE shows consistent signs of recent increases since 2012.
- Recent good recruitment appears evident in standardised length-compositions, providing support for recent increases seen in purse-seine CPUE.
- SC19 Recommendation: Alignment of different datasets suggests an integrated stock assessment could work.

The PAW asked about the sharp decline in the interactions for longline in relation to the decreasing shallow set effort around 2013, the stable albeit highly variable longline CPUE and increasing trend in purse seine CPUE – does this all fit together? The longline data it not consistent with the other data, if you look at the example where the year effect is removed, the lack of effect is more or less due to the year effect being flat. The drop in catch is consistent with a change in the fishing practices rather than a change in abundance. **Phil noted he needs to go back and check again the longline to be sure that data hasn't been dropped somewhere in the process.** Despite the variability in the longline indices in previous assessments, ultimately a choice was made to pick an index that was thought to be more representative. The PAW noted the Hawaii data was likely two fisheries one of which changed a lot in mid-2000's.

The trend in cutting off sharks might be part of the issue, noting cutting off is not adequately reported and it is uncertain the level of silky shark cut offs. Need to get better information on cut-offs by species.

The PAW asked about the stock recruitment relationship (SRR), and how or if this was being considered, given that it is a key to stock resilience for sharks but is not well known. Previous assessments used Beverton-Holt (BH) with steepness 0.4, loosely justified. Need more work on the SRR as it has an important influence on the estimated sustainability of these stocks. Tried survivor based SRR for blue sharks but didn't seem to produce reasonable results, so followed ISC back to BH. The PAW questioned if WCPFC might be considering research on the SRR for sharks? This is currently not in the WCPFC Shark Research Plan.

Phil then went on to discuss the preliminary work and ideas for the stock assessment modelling and the risk assessment. In terms of stock structure, electronic tagging does show movement from the eastern tropical Pacific to the WCPO, and silky sharks can move through EEZs, but recent genetic studies indicate there is likely some spatial population structure across the Pacific. Silky shark are mostly found shallower than 100 m. Tagging-recapture studies show a lot of variation in movement patterns and distances, and no spatial patterns in size compositions across the WCPO. There is still no clear understanding of stock structure for silky shark in the Pacific. No evidence of specific pupping latitudes like for blue and mako sharks, the populations are mostly found between 15N-15S. The biology was reviewed to provide information on biological parameters for the models, noting there was variation in biological parameters depending on the studies and locations, e.g. max age 25 – 36 years, conflicting growth curves. Length at maturity 165 cm FL, age 6 years. Post release survival was reasonably consistent at around 80-85% for longline and 85% for purse seine. The previous assessment used the Joung et al. (2008) growth curve based on samples from a restricted area around Chinese Taipei to Okinawa, however, this was questionable and no clear justification. The Oshitani et al. (2003) growth is probably more appropriate as it involves samples from a much wider area in the tropical western and central Pacific. The assessment should try both, the Oshitani et al. growth seems more likely as base case though. Published natural mortality estimates are questionable, perhaps more likely Z estimates. One issue that was raised is the use of fork length and total length with the model and observer data being in FL but the published growth curves being in TL, and when you do the conversion from TL to FL the growth curve changes notably. The previous growth curve used TL for the growth parameters which is not really correct, and had a strong influence on the early growth patterns. There will need to be some work to improve the growth model used as the base case.

The modelled total interactions need to be adjusted to more accurately measure total mortalities, taking account of release rates and associated mortality probability for released individuals. The approach to account for this was described involving a condition at release model, discard model (estimated proportion of released sharks alive) and a post-release morality rate of 15%. Preliminary results were shown, noting the very high longline discard rates in recent years, but also the large uncertainty in discarding and fates of discards due to the patchy/low observer data. For the purse seine there has been increased discarding since around 2015, and with the recent estimates of relatively high post-release survival, this high discarding counteracts the increased trend in interactions.

The model set-up was then discussed, noting the use of the purse seine unassociated and associated set data as the preferred abundance indices, but removing the first data point for the associated sets data that showed an unrealistic increase given the biology of silky shark. The assessment is being done with Stock Synthesis (SS3) and will initially explore using informative priors (rather than external forced parameter values) for M, initial depletion and R_0 as is considered good practice and done in recent assessments (e.g. blue shark). Initial models have placed higher weight on unassociated sets as they catch larger individuals. Francis weights applied for size composition data.

Results from preliminary model runs were discussed, noting the reasonable fits to the purse seine CPUE, selectivity patterns, and size composition fits. It was noted that more work was needed to justify the priors, especially on initial F and R_0. The initial model runs seemed to estimate initial F, M and R_0 with plausible values, but the M seems lower than assumed by the previous assessment which may have been more indicative of Z since it was based on max age from a heavily fished population in the Gulf of Mexico. Not surprisingly in light if large reductions in longline catches F has deceased since 2010 and stock status has improved.

The work program from now will involve:

- Work on demographics and selectivity: - "Get the biology right, or use size-composition data at your own risk" (Minte Vera et al.), also throw in the stock recruitment relationships.
- Priors: push forward checks and better justification for priors on initial F and R0
- Uncertainty axes:
    - Growth
    - Forcing alternative initial F?
    - Alternative catch and discard scenarios
- Integrating across uncertainty:
    - Use approach developed for blue shark to combine models into ensemble outputs
        - Integrate over demographic uncertainty in the model if we can.
        - Assign prior weight to data hypotheses - e.g., catch or discard level.

The PAW asked about the 15% post release mortality, noting that this depends on condition and is possibly higher. It is suggested to confirm the post-release mortality rates from the literature. *(Note: might be some confusion here, the mortality rate of silky sharks that have been hooked, hauled back and handled is more like 60% - but mortality of those released alive is only about 15% - see Francis et al. (2022) study)*. Phil from DragonFly to follow up on this. PAW supported removing the first 2 years of purse seine CPUE as it is acknowledged that there are issues with the reliability of these data as the observer programs were starting, including species ID issues.

The discussion then considered the biological processes, in particular growth and M. Noting the issue of different measuring methods used in the different growth studies available to inform the assessment (e.g. Oshitani et al. (2003) – precaudal length, Joung et al. (2008) – total length), and the previous assessment did not convert the Joung et al. growth to the fork length equivalent, to be consistent with the length measurement used in the assessment (observers consistently measure FL). The model estimates of M will be influence by the growth parameters, this may also be impacting the poor fit to size composition in preliminary model runs. The PAW asked about information on purse seine observer detection rates for silky shark and whether there was information or a way to account for changes in detection rate on the estimation of interactions. It was replied that there is no information to gauge this issue, but could consider higher interaction scenarios for purse seine, similar to longline.

Day 3 Western and Central Pacific silky shark assessment, Oceanic whitetip shark assessment – phase 1, Southwest Pacific striped marlin assessment

Day 3 started with a presentation from Phil Neubauer on the risk assessment methods and options for the silky shark assessment. The presentation provided the context behind the SC19 recommendation to add at least one risk assessment method to each shark assessment, noting:

- Poor reporting of many bycatch species (and target shark catch)
  - Low observer coverage on longline vessels;
  - Catch histories need to be reconstructed - cannot use reported catches alone;
  - Catch histories are uncertain, even before accounting for discarding and handling/release mortality.
- Increasing proportion of sharks are cut free from longlines
  - Potential for biased picture of interaction rates
- Other data inputs (CPUE, length compositions) are similarly uncertain due to low observer coverage on longliners

The presentation provided the theory behind risk assessment methods noting they generally aim to determine a risk ratio which is in essence a ratio of susceptibility to productivity, in the form a ratio of fishing motility to some limit reference point fishing mortality. Typically, risk assessments apply recent data as they are generally applied in situations where there is little historical data. There are various classes of risk assessment: qualitative, semi-quantitative, quantitative. Quantitative methods perform better than semi-quantitative, and are preferred if data is sufficient, which is the case for silky shark. A quantitative method will be preferentially explored. Examples of risk assessments provided for previous WCPFC sharks; bigeye thresher, porbeagle, oceanic whitetip spatial risk assessment using MIST.

The approaches suggested as a focus for the silky shark risk assessment are EASI-Fish and e-SAFE. These approaches were briefly described. EASI-fish and other spatial risk assessments have a core assumption that spatial overlap translates into vulnerability, moderated by a number of potential factors, including contact selectivity, seasonal availability, depth, post-release mortality, catchability. The suggested approach would apply EASI-fish and e-SAFE with extended estimation of gear efficiency as per the previous Oceanic whitetip assessment, and; explore ways to standardise the spatial impact assumptions, compare outcomes under alternative reference points, acknowledging that either approach will be very sensitive to assumptions of catchability ("gear affected area"). A non-spatial alternative could be to use a length based SPR (spawning potential ratio).

The concern remains that there is no way around the need for assumptions when it comes to sharks and that these assumptions will be influential on the outcomes.

The discussion on the risk assessment noted this was a first cut at presenting the combined integrated assessment and risk assessment approaches and there will need to be some patience while the SC and WCPFC develop their understanding and approaches to utilising the outcomes of the alternative methods for deriving management advice. Given this, it was asked should there be a focus on one method for now? Phil indicated he is keen to explore both, noting with EASI-fish there is a point of comparison with the previous application to a range shark species in the EPO and it is relatively easy to set up, but also because the e-SAFE is a more statistical approach, it would be interesting to see what the advantages/difference might be. This might help us to provide a recommendation either way, if we have an integrated assessment, it is useful to compare various methods to the integrated assessment to seem which seems most robust.

Effected fishing area is a key input that is needed, PAW suggested to explore the use of VMS data, but it was noted this might be a significant piece of work that cannot be done in the time available, and the attractive range of longlines is still very uncertain. EASI-fish uses a yield per recruit based reference point that involved growth and M but not reproduction, so this is another reason to consider the e-SAFE also. The PAW asked about CKMR for silky shark, if it was possible or not. Might be difficult to resource this even if it was possible. SPC noted it will also be exploring EASI-fish across various shark species and is developing species distribution models. Further discussion on catchability assumptions ensued.

The PAW supported the approach presented, and noted the reason for the risk assessment is to make sure we can provide some management advice. Finally, the PAW noted the WCPFC work on sharks has three focus areas: to understand fishing interactions with sharks, to mitigate the impacts of the tuna fisheries on sharks, and to understand the stock status. With regards the later, the obligation of WCPFC members is to provide information on sharks related to tuna fishing activities, but in some areas there is still target shark fishing (e.g. Indonesia archipelagic waters/Coral Triangle region) that would not be included in data provided to WCPFC, so there may be substantial shark catches not included in these assessments, what is the way forward to deal with this? No clear answer to this but, yes it should be considered, and options for estimating the effort that are not reported to WCPFC that catch sharks.

The **next presentation was provided by Tyla Hill-Moana of DragonFly on the year 1/phase 1 work for the Western and Central Pacific Ocean oceanic whitetip shark assessment**. The presentation provided a review of the previous oceanic whitetip assessments conducted in 2012 (Rice and Harley 2012), and 2019 (Tremblay-Boyer et al. 2019), both indicated that the oceanic whitetip shark in the WCPO was overfished and undergoing overfishing. More recent work has studied mortality rates associated with longline interactions and how these vary depending on aspects such as wire leaders, trailing gear etc. Using these mortality estimations stock assessment projections were conducted to inform the potential effectiveness of gear regulations on stock status given the associated scenarios of reduced catch related mortalities.

The work to date on the 2024-2025 assessment has only just begun, with review of the biological information, the next focus of the phase 1 work will involve:

- Data characterisation, catch reconstruction
- Biological data to inform stock structure and model parameters
- Length compositions
- CPUE analysis
- Assessment options
- Present work on data inputs at the SC20 meeting

Finally, questions were posed to PAW regarding catch reconstruction work:

- Can we see improved survival from observer data? (Condition and fate models)?
- Does non-retention/cutting free of sharks impact estimated/reconstructed catches?
- Do LF standardisation models provide a more consistent picture of trends in OCS lengths?

The discussion started with the question to PAW on; given it's been 6 years since the last assessment, is there any new research or work in progress that could be useful for the assessment? It was also noted that while there is some confidence that an integrated assessment is possible there will be increased challenges with the recent data due to the non-retention policies. The poor conservation status of oceanic whitetip means there is a lot of interest in the management advice that is provided from this assessment, and the implications of perhaps not producing a successful integrated assessment and relying on the other risk assessment approaches, could be interesting given its CITES and IUCN listings.

The PAW noted the challenges of longline observer coverage in recent years – and it is likely less than 4% from 20N-20S, the prime oceanic whitetip habitat. Is there recent information of the observer coverage in this region and how do you deal with changes over time? The challenges were duly noted, and will have to work with what we have, the signal of depletion was strong in the last assessment with low observer coverage. The prohibition on wire leaders is also influencing catchability strongly in the Hawaii region fishery so this is another challenge to be considered by the assessment scientist (kind of like a negative effort creep situation to be dealt with in the CPUE analyses). PAW noted that C14 age validation work on oceanic whitetip looks promising, but need vertebrate of very large old sharks, which is challenging due to the overexploited state of the oceanic whitetip populations.

Southwest Pacific Ocean striped marlin assessment

The next session focussed on the southwest Pacific Ocean striped marlin assessment, with the **first presentation from Paul Hamer (SPC-OFP) who provided a background on the fisheries, biology and previous assessments.** The assessment covers the WCPFC area south of the equator, including the WCPFC-IATTC overlap area. The presentation discussed fishing methods, catch history, relative importance of striped marlin catches compared to other billfish, the proportions of catch taken by different flags (noting the longline retention ban in NZ since 1987), spatial distribution of catches (noting the spatial distribution of catches has not changed greatly overtime, most catch is from between 20-40°S, 20-25°C SST), and spatial patterns of CPUE over time. It was also noted the patterns of catches by hooks between floats has changed over time, with most catches being on sets with 5-12 HBF prior to the mid 2000's, but since 2010 most catches are from sets with 20-40 HBF. This is due to the shift in gear settings and mainline materials of the tuna fleet to target tuna with deeper sets (300-

400m), but that striped marlin would still be predominantly caught on the shallowest hooks (<100 m). This will have implications for CPUE using hook numbers as an effort metric as a greater proportion of hooks will have low striped marlin catchability since the late 2000s (i.e. negative effort creep). The biology of striped marlin was summarised, noting growth and reproductive biology would be discussed in more detail later by Jess Farley from CSIRO.

- Max length: 250 – 290 cm (lower jaw fork length - LJLF) (possible up to 400 cm and 260 kg)
- Max age: 10 – 15 years, age validation needed
- Age at maturity 2-3 years
- Length at 50% maturity ~ 210 – 220 cm LJFL
- Major southwestern Pacific spawning area in Coral Sea (northeast Australia), also around French Polynesia
- Peak spawning in warmer months in south Pacific (Peak: Nov-Dec, SST 24 - 28ºC)
- Mostly occur shallower than 100m
- Larger/older fish further south, smaller fish in tropical/subtropical the Pacific Islands
- Support from genetics and tagging for largely discrete southwest Pacific stock

Claudio Castillo Jordan (SPC-OFP) provided a summary of the previous assessment conducted in 2019 (Ducharme-Barth et al. 2019). That assessment applied a single region model will the fisheries defined according to 4 subareas. A 2 region spatial model was also explored but was highly sensitive to recruitment distribution and movement and there was not enough information to be confident in these aspects, the model estimates were not plausible. The assessment included an orthogonal uncertainty grid with steepness, growth, M, CPUE, size data weighting and recruitment penalty CV. The assessment indicated high uncertainty, but that the stock was more likely overfished and close to undergoing overfishing. It was noted that the SC15 commented on various technical issues:

- Data conflict – CPUE – size composition, weight of composition data very influential
- Concerns about biological parameters, growth, maturity
- Influence of high early catches and confidence in these data
- Really need for better biological data to inform life-history parameters
- Verify the age methods, more age sampling
- Better estimates of movement/mixing rates – with view to spatial model
- Improved L-W conversions, explore sensitivity to uncertainties in conversion factors

There was clear need for improved biological data for striped marlin in the southwest Pacific raised after the previous assessments, however, since 2018 there have only been 89 new biological samples, including only 7 otoliths and 10 gonads, added to the Pacific Marine Specimen Bank.

The PAW noted concerns from the previous assessment regarding the influence of size data from NZ and the sensitivity of the model results to the CV on the recruitment penalty. It was responded that NZ size data comes largely from the recreational fishery, for which the catches are relatively low, but the size (weight) data have a strong signal of a decline in mean weight. The time series goes back to 1920, although only data from 1952 is used in the assessment as there is no other data prior to 1952, this is around when the declining trend begins. It was noted that this recreational weight data was influential on the estimation of the biomass and depletion trends.

John Holdsworth, noted that the recreational catches around NZ are recently 50-60 tonnes and probably similar levels released, and up until the 1970s pretty much none were released. The weight data for that early period is a good representation of the catch, also they are all green weights on day of capture on certified scales. The question was raised on whether the high striped marlin targeting by the Japanese fleet in the Tasman Sea may have influenced the local abundance around NZ and contributed to the declines in mean weights observed. Also, tagging data indicate the fish caught around NZ migrate from spawning areas both in the Coral Sea and around French Polynesia – the implication might be that the trends in data from around NZ provide indication of trends in adult abundance for the entire southwest Pacific stock.

Dr Julian Pepperell, commented on the recreational data from Australia which goes back to the 1930's, noting in the earlier period of the fishery a lot of the records are from fishing in the southern area off NSW, where larger fish are more prevalent. It was noted that a lot of tag release happens now and weights of released fish are estimated, but interpretation of weight trends in landed fish needs to be wary of the self-imposed weigh in limits for recreational tournaments that have gradually increased from 40-50 to 90 kg since the 1980's. Increasing weight trends are likely biased, and the tag-release weights should be compared.

The PAW asked about the uncertainty grid in the last assessment, noting the orthogonal combinations of steepness and M that would have a strong influence on F_{MSY} reference points and some of the pairings of M and steepness are probably not reasonable. It was agreed this orthogonality of M and steepness is not really a great approach, and the new assessment would be looking to develop a joint prior that incorporates the correlation of M and steepness.

The PAW questioned the issue of the historic Japanese high seas drift net striped marlin catches which influenced the north Pacific striped marlin assessment, but there was suggestion that some of this catch is likely form the south Pacific, so would be good to confirm or better understand if this might be important for the south Pacific striped marlin assessment. **The Japanese indicated they were looking into the location of these historic catches, but that work was still ongoing.**

Claudio Castillo Jordan, then provided a presentation on a proposed fishery structure and fishery data summaries. The fishery structure was proposed to continue as applied in the previous assessment, which partitions the longline fisheries into three subareas and 12 fisheries as per Figure 4, along with two recreational fisheries for AUST and NZ. The main difference to the last assessment is the proposed application of a single index fishery that combines the Japanese and Chinese Taipei fleets, whereas the previous assessment had separated these fleets to create alternative indices that were used in the uncertainty grid.

The PAW commented on the possibility to attempt the regression tree method to see if the size composition data indicates that an alternative/refined fishery structure might be more appropriate. Post- PAW analysis of spatial size composition data using the regression tree approach (see SC paper), did not provide any strong support to diverge from the previous fishery structure.

The fishery data summaries were presented by fishery. The Japanese fleet, despite reporting striped marlin in reasonable numbers from the Tasman Sea region, the data has no weight data over the last two decades. **It was noted that the lack of Japanese weight data may be because the measurement**

method switched from weights to length measurements, so there is need to check if length data might be available.

The Australian data was questioned as to whether it was whole weights or processed weights.

The NZ longline data has very little length data, but despite the retention ban since the late 1980 there are still catches. The NZ representative noted the observer record of catches when matched to logbook records are quite different, so catches in numbers are estimated for the observer records scaled to logbook effort, and to estimate the catch weight, recreational length data I used. There are concerns on accuracy due to uneven distribution of observer effort. Also, the issue of post release mortality was raised and it was noted that in 2021, 78% of striped marlin from the NZ longline fishery were released alive. Post release mortality should be considered and there are studies with satellite pop-up tags to provide estimates.

The PAW also noted that 95% of recreational striped marlin are released but need to consider the tag/released fish numbers also. For the NZ data it was suggested there may be some additional length data from other research work. The PAW noted the previous assessment assumed that all the commercial NZ longline catch died, but that tag and released recreational fish survived, and only the landed catch is used in the assessment.

There was a discrepancy in the Fishery 14 (French Polynesia) size data the needs checking, i.e. period of small lengths in the 1990s early 2000s. May have been related to issue with species ID.

Release mortality and numbers of released fish expanded from observer records, may require some further consideration to more accurately reflect the mortality for the NZ longline fishery.

There was little further feedback on the fishery size data.

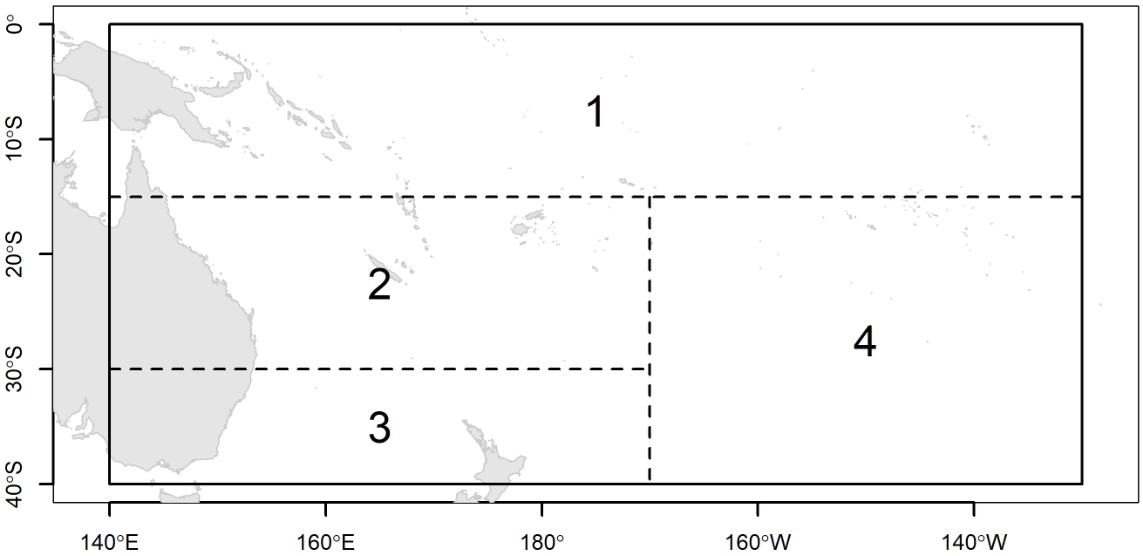

Figure 4. Spatial domain of the southwest Pacific striped marlin assessment, with sub-divisions for fleet areas.

Table 3. Southwest Pacific striped marlin assessment fisheries definitions from the 2019 assessment.

Fishery	Nationality	Gear	Sub.region	Catch
1: JP 1 LL	JP	Longline	1	Number
2: JP 2 LL	JP	Longline	2	Number
3: JP 3 LL	JP	Longline	3	Number
4: JP 4 LL	JP	Longline	4	Number
5: TW 4 LL	TW	Longline	4	Number
6: AU 2 LL	AU	Longline	2	Number
7: AU 3 LL	AU	Longline	3	Number
8: NZ 3 LL	NZ	Longline	3	Number
9: AU 3 REC	AU	Recreational	3	Number
10: NZ 3 REC	NZ	Recreational	3	Number
11: OTHER 1 LL	DWFN/PICT	Longline	1	Number
12: OTHER 2 LL	DWFN/PICT	Longline	2	Number
13: OTHER 3 LL	DWFN/PICT	Longline	3	Number
14: OTHER 4 LL	DWFN/PICT	Longline	4	Number

The previous assessment did not conduct a reweighting of size composition data to account for sampling biases with respect to catch and CPUE as has become standard in recent tuna and billfish assessments. The 2024 assessment will introduce the reweighting procedure. Tom Peatman provided a brief outline of the approach.

There was some consideration by the PAW on the need for reweighting of the recreational fishery data, but it was noted that reweighting in terms of it's implications for extraction selectivity is not likely to have any influence. What is more concerning is the issue that the self-imposed high grading (non-retention on smaller individuals) by the recreational fisheries might mean that the weight data could be biased to higher weights in recent years, so the decline in mean weight might be underrepresented noting the last assessment was influenced by the recreational weight data.

The PAW was also asked about alternative weighting of different fisheries to improve fits to size composition, as it was noted that in the last assessment some fisheries size data with significant catches were not well fit. i.e. region 1 Japanese longline has multimodality due the catch of small fish near the equator. It was suggested that it could be worth considering seasonal selectivity for this fishery.

Laura Tremblay-Boyer then provided a presentation on the work on the CPUE standardisation for the Australian Eastern Tuna and Billfish fishery. This is a multi-species fishery that evolves and changes year to year due to economics, regulations and species availability. It developed from the 1980s and expanded in the mid to late 1990s. It has focussed more on tuna (including bluefin) and swordfish, striped marlin is a byproduct species. A key feature of this fishery is the diversity of gear settings applied depending on the species targeting in time and space. Another complication is the changes in spatial fishing patterns over time influenced by targeting preferences (i.e. swordfish) and regulations. The nominal CPUE shows and a long-term downward trend from the late 1990s, but a big recent spike in 2022 which was confirmed by industry anecdote and may be a local abundance pulse. The catches are seasonal with highest CPUE in quarter 4 (late spring/summer). Highest CPUE the area 20-35°S, is partly influenced by regulations. Weight trends are stable over time at an average of 70-80 kg whole weight and 3-5 years age.

The standardisation is based on a simulation tested method (Zhou et al. 2019) and reviewed/updated annually. It is an annual index but includes quarter effects. Results for 1998-2022 were presented. A core area, subdivided into subareas with assumed constant abundance, is defined that represents a region of consistent effort over time. Used a GAM to do the standardisation, lots of covariates in the model as it is applied to various species, has a strong negative influence in early years and positive in later which flattens the long-term trend. The most influential covariates were targeting cluster, HBF, hook density per Km of mainline (which increased from 2005). Decline in nominal CPUE in recent years was due to increased HBF (to target swordfish) and was standardised up.

The PAW asked about a TAC on striped marlin – there is one in Australia – around 400 t but has not been met recently. The PAW asked about how the CPUE modelling approach deals with spatial variation of the data coverage, and importantly the issue of including a species cluster covariate along with the many other gear covariates for a non-target species. Specifically, by including the cluster variable as a catchability covariate, it could remove some of the abundance signal. It was responded that this is likely an issue, so moving to an alternative approach that is more focussed on the species is more appropriate. It was also noted that the negative hook density effect on catchability and the increasing trend in hook density (i.e. standardises up the CPUE), could be counteracting a real declining abundance trend that is indicated in other data, but this is difficult to disentangle. It was noted that the plan is to transition to a more continuous spatial approach for the CPUE analysis. Given some of the concerns on the Australian CPUE modelling it is unclear that it would be considered as an alternative indicator for the assessment.

Claudio then presented the work so far on the longline CPUE spatio-temporal modelling for the current assessment. The presentation noted the last assessment transitioned from a GLM approach to a spatio-temporal model in VAST, and developed multiple indices for the uncertainty grid, Japan LL in subregion 2, Chinese Taipei LL in subregion 4 and the Australian LL subregion 2. The Japanese index for subregion 2 was used for the diagnostic model. Japanese data started in 1952. The difficulty of replicating previous analyses was noted in relation to the restrictions on operational data sharing agreements with DWFNs, i.e. maintaining raw data sets from previous analysis is problematic. The plan for the CPUE model is to start from 1967 to use the best data and combine Japan and Chinese Taipei into one index fishery, and apply an annual model implemented in sdmTMB. Some early results were presented, compared to the previous method and they show a similar trend but less extreme short-term variability, particularly in the early period. The PAW was asked for any advice, noting this was early stages of development. The sparsity of data was noted at the annual level, particularly when modelling data for one season (i.e. quarter 4). The PAW confirmed that the data for the annual index was indeed restricted to one quarter. MFCL would need the index to be specified to a particular month, which could be the middle month of the quarter. The PAW suggested that given the amount of unsampled area if data is restricted to one quarter, it could be better to use the full year data instead, add a month or quarter effects and allocate the index to the middle month if the year. The PAW discussed alternative ways to create CPUE indices, such as using certain fleets with more data early in the time period and transition to other fleets later.

The striped marlin discussion then moved on to biology with a **presentation from Jess Farley of CSIRO on updated work on growth and reproductive biology in 2021** (WCPFC-SC17-SA-IP11,) after the previous assessment. There were concerns that the previous growth curves based on back calculated lengths at age from fin spin ages were biased towards slower growth rates, and that the previously used reproductive ogive was biased to larger sizes at maturity due to misclassification of smaller spent fish as immature. The study used otoliths that were available for the same fish that were aged from fin spines and re-examined the histology slides of gonads from the previous study.

It was found that fin spines can lead to significant underestimation of age for fish over 3 years age compared to otoliths. New growth curves were developed from the otoliths, and showed the oldest age of the samples were 15 years for one male, and females have higher Linf. It was recommended that the new otolith-based growth curves were used in the assessment but noting the small sample size and lack of age validation for the otolith readings, more research on striped marlin age and growth was strongly encouraged.

The re-reading of the histology slides indicated that only three fish were likely misclassified, but this had an effect on the shape of the ogive, while not having much affect on the L50 size. There were also differences in the models used to fit the maturity data and size binning. Ultimately the updated ogive was recommended to be used for the assessment. Again, the appalling lack of samples and data was emphasised as well as the need to explore age validation. **Something to add to the research needs for the WCPFC Billfish Research Plan.**

The PAW talked about what sample needs for further studies, noting the biggest fish are important, but indeed all sizes are needed. Difficulty dealing with the big heads and tiny otoliths, need to develop approaches/protocols for collecting and storing samples, noting both tissues and otoliths are important. The sampling program for the north Pacific was noted and that some of the samples from this program might be available and useful. **It was suggested that a small paper on what would be required to look at age validation studies would be useful to bring into the billfish research plan (Allen Andrews to provide some guidance).**

Claudio continued on to discuss other biology aspects for the assessment, providing a recap of the previous externally input biological parameters, again noting apart from the updates to the growth and reproductive ogives based on the previous collected samples there was very little new information or samples since the last assessment. The PAW noted the updated max age which will have implications for M and translation of length-based inputs to age based.

The PAW was asked if there was support to update the growth curves and reproductive ogives, and this was supported, noting the M should be updated to account for/be consistent with the growth. The PAW suggested that if looking to update the Lorenzen, explore the length-based version. The PAW sought clarification on some of the differences in the new growth curves including the faster growth of the smaller fish, and it was noted the new growth is based on the observed data rather than back calculated. The PAW asked about internal growth estimation using CAAL, but it was noted there are only 60 annual age otolith samples. Also noting the peer review recommendations on applying CAAL if possible, but again the lack of samples was noted. It can be tried, but unlikely to be useful given the lack of samples. In terms of two sex model, indicated that the biological information is insufficient and the catch conditioned approach in MFCL is still to be thoroughly tested for two sex models.

Claudio then went on to discuss the initial model development, starting with a recap of the previous assessment approach. Preliminary results of stepwise model changes were presented, included a comparison with an SS3 model. The models run so far were estimating a more depleted stock status than the previous diagnostic model, with the largest change related to the new growth parameters with lower Linf, noting the M was not yet updated to account for the higher max age and the reproductive ogive has not yet been updated. The PAW noted that there are some very large fish in some of the size samples and depending on the weight applied these might be having a disproportionate influence on the model estimation, also these larger fish are well above the Linf from the growth curve. This comes back to the limited number of samples in the otolith data.

There was discussion around the issue of the early very high catches reported by Japan in the second year of the catch time series in relation the very rapid decline in stock status (depletion) at the start of the model time series. This big drop seems somewhat unrealistic, and another model was run starting in 1966, but a similar large initial drop in stock status was also estimated. Various thoughts on this were discussed, but this will require further exploration.

The PAW raised the consideration of effort creep given the time series of CPUE was very long. It was suggested to apply some fixed rate of effort adjustment to the CPUE, but it was also noted that for striped marlin (a non-target species), the effort creep would likely be negative as they are surface orientated and the trend in longline fishing targeting tuna has been to deploy deeper sets which would likely reduce striped marlin catchability. Effort creep scenarios are a case-by-case consideration depending on the species, so any levels applied in a grid used for management would need some reasonable justification.

The PAW commented on the differences that can occur depending on the method used to estimate the weights for the composition data in the likelihood. It is unclear what weighting method will be used at this stage.

Overall, the model development was still very preliminary, but seemed to show that there was very minimal influence of the change from catch errors to the catch conditioned approach, but that updating the growth to the otolith-based growth parameters could have a notable influence on stock status estimates.

Day 4 Multifan-CL updates and workplan, next generation tuna model, south Pacific albacore MSE development, developing a standardized reporting template for stock assessment and management advice, Close Kin Mark Recapture update, longline effort creep project, revisit the south Pacific albacore assessment uncertainty grid and sensitivities.

Day 4 started with a presentation **from Nick Davies on the MFCL developments over the past year and the focus of work for next year**. The presentation covered:

- Current and development versions - status
- Benchmark testing
- Update – items since PAW 2023 reported to SC19
- Update – items since SC19
- Work plan 2023-24

The presentation outlined a large amount of work that has occurred since the 2023 PAW (MFCL version 2.2.3.0), including:

- Optimised parameter scaling applied as default setting
- Lorenzen functional form for natural mortality – allowing for fixed and supplied initial start values
- Catch-conditioned models with estimation of the *Fishing mortality_effort_relationship* regression – revised the coefficient bounds
- Variable penalty weight for grouped non-concentrated CPUE likelihood, allowing time varying CV on CPUE
- Corrections to concentrated CPUE likelihood formulation with assumed sigma (grouped) and applies normalised lambda
- von Bertalanffy stdev (length-at-age) - correction to variance formulation
- Correction to likelihood components report - Dirichlet multinomial term for LF and WF data
- Abbreviated variance calculations of dependent variables – can be applied to only selected quantities of management interest (SBF=0 and MSY-related) for grid models

Benchmark testing was conducted with the release of the version 2.2.5.0 around the time of SC19. Development since then has involved:

- Simulation of stochastic CPUE data - catch-conditioned model
- Simulation of projected effort - catch-conditioned fisheries
- Consolidation of total likelihood components report
- Constraint on regional recruitment distribution by the orthogonal polynomials parameterization
- Parameter listing output file

Each of these recent developments were described in detail by Nick. The development version at the time of PAW was 2.2.5.1, and had involved the following updates and correction:

- duplication and obsolete use of parest_flags(173)
- maximum tag recapture iterations in pooled group increased
- selectivity temporal structures in respect of month_1
- xinit.rpt indexing fixed for grouped selectivity parameters
- SSMULT_noRE correction to likelihood calculation
- non-decreasing penalty for time-block selectivity estimation

The work plan for 2024 was outlined noted a stronger focus on the developments required to facilitate the MSE frameworks and associated projection aspects required for the catch conditioned model. Focal areas included:

- Independent variables report – extend to full range; add st-devs
- Updated and writing new sections for the MFCL manual:
    - Catch-conditioned model
    - Stochastic simulation CPUE pseudo-observations
- MSE Team support:

    - improving stochastic projection efficiency (estimator model evaluation)
    - stochastic projection functionality for: terminal numbers, recruitments with autocorrelation, selectivity deviates
- Catch-conditioned model:
    - enable estimation of selectivity deviate coefficients
    - review the operation of existing control phase routines
    - apply fml_effort_rltnshp estimation conditional on a fish_flags(fi)
    - testing with multi-sex example
    - generation of simulation pseudo-observations of tagging data
- Extend von Bertalanffy st.dev correction to Richards curve, and multi-species/sex instances
- Correct discrepancy between Frecent/FMSY in variance report and Fmult in plot.rep.

Nick also noted work that remains on the list from the recent yellowfin assessment peer review, including:

- Extend MULTIFAN-CL so that variability in weight-at-length can be taken into account.
- Extend MULTIFAN-CL so that it is possible to specify the number of spline knots when defining selectivity and where they are located with respect to age (length) as the current approach means that the selectivity for some knots is constrained to zero.
- Extend MULTIFAN-CL so that account can be taken of age-reading error when fitting to conditional age-at-length data.
- Add the ability to specify overdispersion in CPUE as an additive rather than multiplicative factor.
- Integrate the calculation of M-at-age from the sex-ratio data into MULTIFAN-CL unless a sex-specific assessment is used.

(Note: since PAW a new benchmark tested version of MFCL has been released Vers. 2.2.6.0)

Arni Magnusson then presented an overview of the WCPFC Project 123: Scoping the next stock assessment platform. He covered:

- Project P123: objectives, background, terms of reference
- Software Platforms: operational, current and future development
- Road Ahead: assessments, workshops, collaboration, adaptive plan
- Possible Outcomes: level of funding, partnerships

Arni emphasized the need to start working towards a successor for MFCL given the recent retirement of Dave Fournier and the retirements of John Hampton and Nick Davies expected in the not too distant future. Project 123 was supported by the WCPFC to the extent of 50,000 USD per year (plus a 10,000 USD additional contribution for ISSF), with a 3 year window. The funds will provide support for SPC staff to start driving a process of transition which will require building collaborations, workshops, and scoping of what will be required from future tuna stock assessment software, and trialing alternative software platforms on tuna data. Also noting the ADMB project is now finished, and models are moving to TMB.

Arni outlined the workplan:

2024

1. Review and identify important model features for tuna assessments
2. Identify existing platforms that have these features or can be extended
3. Reach out to and initiate collaboration with model developers
4. Conduct two workshops in 2024, one online and one in person

2025-2026

5. Conduct simulation studies
6. Determine which platforms can be considered viable candidates
7. If a viable platform has been identified, plan transition
8. If no viable platform is identified, extend a platform or create a new one

A number of existing software packages were discussed, along with several that would be considered in as in ongoing development.

- Stock Synthesis (nearing end of life)
- Casal2 (still being developed)
- Gadget (still being developed)
- SAM fitted to length comps (Colin Millar, Anders Nielsen) (still being developed)
- WHAM fitted to length comps (Giancarlo Correa, Tim Miller) (still being developed)
- ALSCL (Fan Zhang, Noel Cadigan) (research model – could be developed further)
- CCSBT (D'Arcy Webber, Rich Hillary) (still being developed)
- FIMS (NOAA project) (major model development project to superseded Stock Synthesis)

The push for developing State Space Models with length and age structure, spatial partitioning and ultimately the capacity to model tagging data and incorporate CKMR data was noted.

Options to consider how to progress this transition were discussed: in particular, should we just roll the handle on the next MFCL assessments and devote resources to implementing the assessments in other software. It was noted that this transitioning and testing of alternative software is also not a 'hobby' project and will require full-time dedicated staff resources. A strawman was proposed to guide the transitional workflow (Figure 5).

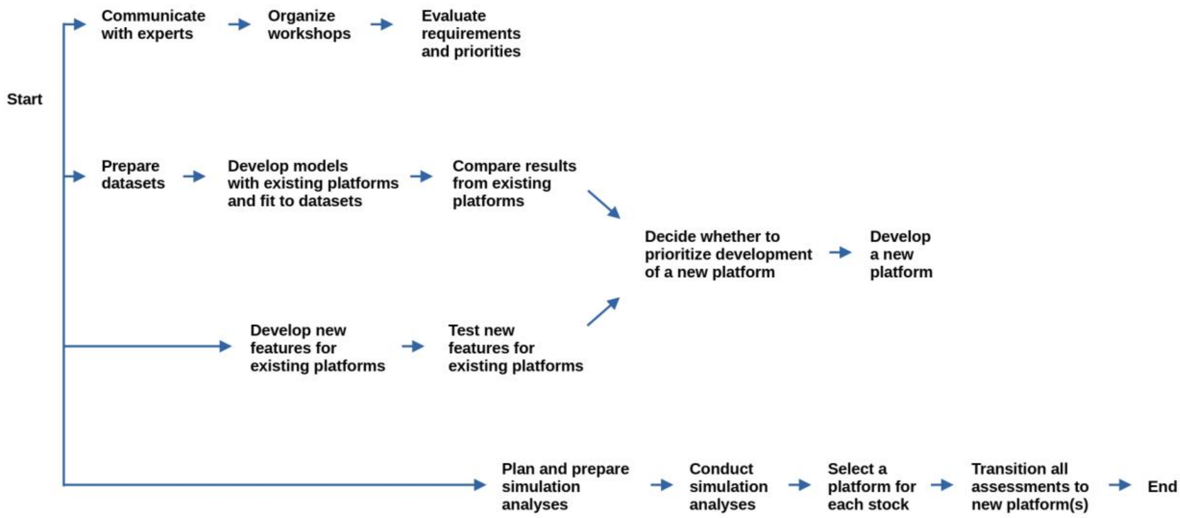

Figure 5. Workflow strawman pathways for the transition from MFCL to a successor stock assessment platform.

Arni noted the risk that continuing with a small allocation of funds each year will likely lead to limited real progress as the work will be a side project. We will continue to make smaller improvements to MFCL, try other software that will not do everything we need, have a workshop and repeat this, rather than devoting a fulltime resource to the work, have a clear plan with a target to be implementing an MCFL successor within X years. For the latter, we would be likely requiring a fulltime two-person team for 5 years. The importance of collaboration was emphasized, this is not something that SPC would want to go alone on.

The PAW noted the need to reach out to other RFMOs for collaboration and the need for a well formulated proposal if planning on asking for greater contributions from WCPFC members through SC/WCPFC contributions.

The PAW noted the step involving identifying what features and capabilities are desired is a really key step. Noting that the list of software mentioned did not include SEAPODYM, and that we need to be looking at these types of models and fully exploring the features and capabilities, including the incorporation of the environmental data. Develop a comprehensive wish list and take it to the groups actively developed stock assessment software and see if they can build in these features. SEAPODYM was never really considered a stock assessment tool, but with its ongoing development, perhaps it can be added to the stock assessment list and considered is an option in this context.

The PAW noted that the CAPAM next generation stock assessment model meeting in 2019 came to a point where there was general support that developing a collaborative general modeling framework or suite of models that can serve needs across very many users and contexts was the way forward. However, this has not eventuated, and different groups have tended to go back to working on their own models, the global collaboration has not happened. This project may end up being just another solo effort, which seems an inefficient use of time and resources. The PAW emphasized to not go down a SILO path, but to make strong effort to build the broader collaborations.

The PAW noted the benefits of having a large/moderate user base for whatever succeeds MFCL – both for resilience to staff turnover but also detecting bugs. Also, the SC really needs to have a conversation on where and what roles they see that stock assessments will play in the future as MSE and management procedures are becoming the main drivers of management decisions. This may factor into refining the focus of these types of model development projects, and this discussion probably needs to happen before going to the commission for large amounts of funding. The PAW noted that collaboration will be promoted if all the RFMOs chipped in resources so that their scientists have an incentive and are supported to collaborate and contribute work.

The next presentation was provided by **Phil Neubauer on WCPFC Project 113b: Standardised stock assessment reporting framework.** The project is contracted directly from the WCPFC secretariat to DragonFly data science. The presentation outlined the terms of reference and scope of the work. The first part of the work will be to conduct a survey of managers across the WCPFC membership, with feedback expected by May and then follow-up with stock assessment scientists. A draft report would be expected to be shared with CCM heads of delegations in June, and then revised and updated for consideration at SC20.

Rob Scott from (SPC-OFP) then provided an update on the south Pacific albacore harvest strategy development, with a focus on the MSE development. Rob noted the WCPFC workplan aims to adopt a management procedure for south Pacific albacore in 2024, and to achieve that, there is a lot of technical work on MSE that will be required this year. Rob noted the implication of the new assessment, and the interim TRP adopted by WCPFC 2023. The presentation outlined the history of the albacore MSE work, and the approaches to the stock status estimation methods, noting simpler empirical approaches failed and have moved to model based estimation methods, exploring various forms or surplus production models. So far, the MSE framework, operating models etc. have been based on the 2021 assessment, but considerable pressure to update everything to the 2024 assessment, which places some serious workflow and time challenges if an MP is to be adopted in 2024.

In terms of the HCR component of the MP to start testing, the form used for the skipjack MP is being used as the basis. The OM grid currently includes steepness, movement, size data weighting, growth, recruitment distribution, CPUE hyperstability, effort creep, and is considering options to dealing with the recent big dip in stock status (depends on new assessment), scenarios regarding options for how to manage the EPO stock component and climate change. Noting these would all be reviewed pending the 2024 assessment. Rob noted the retrospective issues with the previous assessment models and the improvements presented in the new assessment models, plus the problematic issue of the big dipper for the MSE projections, and that moving to the 2024 model might be desirable, especially if it is an improvement on the 2021 model with respect to these issues. Rob discussed the issue of the bigger dipper and the evidence in the size composition that recent recruitment failure and continued high catches was driving the recent dip.

Rob showed some comparisons of the overlap in the depletion trajectories of the 2018, 2021 and some preliminary 2024 models – indicating the early signs are that the 2024 assessment would be relatively consistent (overlap the ranges) with previous assessments. Rob discussed the things that would need to be worked through to update to the 2024 model spatial and fishery (areas and fleets) structures. He also noted the simplified spatial structure for the 2024 assessment may limit the capacity to explore climate change influences on migration and distributional changes. He discussed the potential climate related movement scenarios that might be explored.

The current working assumptions for the MSE work were noted:

- 3 year management cycle
- MP applies to whole WCPFC-CA south of the equator – EPO encouraged to adopt compatible measures.
- Options for EPO fishing levels included in robustness scenarios – but let's see now, more discussion imminent.
- Request that both catch and effort management approaches be investigated
- MP applies to all fisheries (including troll fisheries – but ongoing discussions)
- Candidate MPs designed around
  - Interim TRP (0.96 SB2017-2019/SBF=0)
  - Additional range of candidate TRP values (SB/SBF=0 0.42 - 0.56 or preferably equivalent levels - but defined in terms of a reference period).
  - Meta-rules to limit change in catch/effort between management periods.
- Performance indicators
  - Depletion relative to iTRP
  - Risk of falling below the LRP
  - CPUE relative to baseline levels
  - Catch relative to baseline levels.
  - Catch stability

In terms of the imminent work plan, a key requirement at SC20 will be to achieve agreement of the OMs for the MSE work, review the TRP options and evaluation of candidate MPs. The SMD02 in September 10-12th will provide opportunity to solidify these options and define additional MSE work to do before WCPFC. There is also a desire from stakeholders to explore both catch and effort controls, so will need to consider how to go about this.

The PAW asked about the basis for the OMs, commenting that the use of the 2024 assessment as the basis for the OMs would likely be the preferred approach given the differences to the 2021 assessment. But it was noted that updating OMs for each new assessment is not necessary and should not be expected, the main thing is the assessments, and the OM grids capture a similar range of variation/uncertainty. There was discussion on the CPUE as a performance indicator, and it was noted that this is proxied by the vulnerable biomass. The question was posed that an overall CPUE/vulnerability type performance indicator might not be relevant to a particular fleet, and is it possible to consider fishery or area specific indicators? The response was that for now this was not being included, although it has been thought about. The preference from the point of the technical work is not to put forward fleet or area specific performance indicators as this is creeping into the area of allocation trade-offs. All fleets are assumed to be affected equally by the MP, and in terms of differential impacts or performance of different fleets, that really is for WCPFC to work through. The aim of the MSE is to provide a range of options and performance indicators at the stock scale and up to stakeholders to make their own decisions.

The PAW further accepted that it is not necessary to update the OMs for the new assessment, but seems likely that this would be preferred, with some additional factors in the OM grid or as robustness tests. The question was asked whether if the OMs are to be updated should there be any communication of this to the SC/PAW? **The response was that if a switch to the new assessment as the basis for the OM grid, it would be preferred if this was communicated more broadly to SC and at least back to the PAW.**

The PAW suggested, if time, it would be useful to understand what has changed between the assessments to be confident that the better assessment was the basis for the OM. **Nicholas offered to do some explorations on the previous assessment.**

The PAW asked about what explorations were being done on varying the MP, is it both the estimation method and the HCR? It was responded that we are trying to implement a fixed estimation method and tried various options leading us to the age structure production model, so for much of what is being evaluated among MPs, it is really the HCR component that is differing.

The PAW discussed the concept of exceptional circumstances and in this case if the 2021 and 2024 assessments show notable differences, this could justify moving to the new assessment.

The PAW noted that the SC has previously made the call on what the OMs are based on, so leaving this to SPC might be a risk even if the new assessment is showing notable improvements.

It was followed up that while SPC can put forward their recommendations it will still be up to SC to make the call on what is used as the OMs for the evaluations. PAW suggested to move forward with 2024, as 2021 has been done. Work towards the new 2024 grid options and park work on 2021 for now, but the communication of this beforehand is noted. It was also noted that it is important that WCPFC members understand the options, and importantly that they are aware that there are options.

The Chair of the WCPFC South Pacific Albacore IWG noted that this meeting was also now planned to occur before SC so this can be discussed in the IWG.

Finlay Scott (SPC-OFP) then provided a presentation on the work on the estimation method. He noted some key qualities to aim for in an estimation method:

- Should work well for a range of stock statuses
- Detect if stock is approaching or is below LRP
- Robust to uncertainty
- Performs well under a wide variety of scenarios
- Simple as possible
- Simpler than a stock assessment model
- But need to balance simplicity with performance
- Remembering that the estimation method is not a stock assessment
- Only provides estimate of current stock status for HCR

Finlay then when onto the discuss the work so far on developing model-based EMs,

- SC19 MI-IP-02 looked at two candidate EMs - both model based
- Continuous time surplus production model - SPicT

- Age Structured Production Model (implemented with Multifan-CL)
- Simulation tested across a broad range of uncertainties, stock statuses and future scenarios to compare performance.

The discussion then focussed on the MFCL Age Structured Production Model (ASPM) as it performed better than SPicT in the initial explorations. The tests indicated that the ASPM produced stock status estimates that where similar to the OM estimates (based on 2021 assessment model) for both historical and projected periods. The notable exception was that the ASPM tended to provide overly optimistic stock status ($SB/SB_{F=0}$) compared to the OM when the stock was heavily depleted. This was a concern; however, it was noted that a good HCR should maintain the stock status well above the levels at which that bias seems to become prevalent and can be considered also in the design of an HCR. The next step is to test the full EM and HCRs together in the MSE framework, and see if we can improve it, also the move to a catch conditioned approach will need to be explored.

The PAW suggested the EM is like trying to get a simpler model emulating a more complicated model, and another option is to consider non-parametric emulation of models.

The PAW noted the value of delving into the EM performance in the way presented but acknowledged that the best evaluation will be within the full MP using MSE. It was asked if it was possible to test the MPs in situations where the stock is, or becomes, heavily depleted, like a robustness test. This could possibly be explored.

The next presentations from Laura Tremblay-Boyer (CSIRO) and Giulia Anderson (SPC-OFP) provided an overview and update on the Close Kin Mark Recapture work for south Pacific albacore. The presentation focussed on the sampling design aspects and basic theory of CKMR methods.

The sampling objective of 25-30,000 samples over several years was noted and the presentation discussed how this sampling objective was derived. A key decision to make in determining the sampling design is the CV for the estimated quantities, to improve precision you just need to detect more kins, and to do this requires more sampling. The design is also influenced by whether the objective if is to estimate population size or mortality. If population size, you need to sample adults and offspring, because detecting parent offspring pairs is critical, if the interest is mortality, then adults are not required, and focus is on detecting half-sibling pairs. A key requirement for designing the sampling program is that you need a rough indication of the population magnitude and age structure, so you need do this modelling first. For south Pacific albacore the sampling design was developed by Mark Bravington in 2021 (SC17) using a model developed based on the 2018 assessment. Challenges with south Pacific albacore include the issue that we can only get small juveniles in large numbers from around NZ, so the sampling design had to consider the implications of this and account for this by spreading samples for adults across the south Pacific.

Sampling designs for yellowfin, bigeye and swordfish in the Pacific are now being developed. The PAW noted that the SPC sampling program was trying to get a wider spread of samples by sampling from US and Canadian fleets fishing in the south Pacific. The PAW asked about the progress on epigenetic ageing, and sex markers. The sex markers are done, the epigenetic ageing is ongoing. However, the sampling widget designed by CSIRO currently needs to have a larger tissue sample capacity to allow the analysis of the various genetic markers from the one sample. It was pointed out that the likelihood of the juveniles around New Zealand coming predominantly from spawning immediately north is highly unlikely considering oceanographic and larval transport in the south Pacific gyre.

Giulia Anderson then presented on the progress with the sampling program. She noted the target for kin pairs of either type is 50 minimum, and this means around 30,000 samples are required distributed from across the south Pacific. Sample locations have been established at eight ports across the Pacific, NZ ports for juveniles, and chasing addition ports for distant water vessels. Training and development of port samplers and protocols has been conducted in various locations and sampling systems must be adapted depending on the situations at each port.

- CKMR training events in Fiji, Tonga and Marshalls—plus Samoa this week and Solomons next month
- 28 collaborators from 10 countries fully trained
- Protocols specialisation at 6 ports
- Sampling kits for 26,000 fish distributed to 8 countries
- 7,300 samples collected from 3 ports in 2023
- 2,000+ samples collected so far in 2024
- Countries will start sampling independently once current training events are finished
- The ONSHORE port sampling app. now has a CKMR component

SPC staff are based some of the countries and these staff are critical in quality control.

In terms of availability of CKMR for the 2027 assessment, this remains uncertain. There are many logistical constraints to moving samples and risk that as the sampling program is still developing, we might expect issues with sample quality that means additional samples are required.

The potential for support for sampling in American Samoa was raised as there may be staff available back at American Samoan ports in the future, although most of the US troll vessels go back to Oregon.

The option of having gene extraction stations in different countries could improve the logistics of moving samples around so this is something being explored.

Jemery Day of the (SPC-OFP) presented the next presentation on WCPFC project 122: Scoping Study on Longline Effort Creep in the WCPO. The presentation outlined the scope of the project and noted that some previous WCPFC assessments had included effort creep scenarios on CPUE (i.e. 0.5 – 2% p.a.) as sensitivities, but that it had not be applied in models used for management advice. Jemery discussed various ideas on ways to think about effort creep, and thoughts of others. The project is in the early stages and the PAW was asked for input on their experiences and advice for determining levels of effort creep and how to account for it in stock assessment.

The PAW provided examples – NZ rock lobster where they tried to statistically determine effort creep which in the first attempt provided a plausible estimate of 1% per annum, but in follow up analysis produced implausible rates. Ultimately, they conducted a review of literature and came up with a number to apply and included as an axis of uncertainty, with a linear increase over time. This seems to be the approach that is often applied, but there is also the need to consider the functional form and timing of changes in catchability due to effort creep – e.g. step/sigmoid changes, linear, intermittent etc.. The next phase of the project will involve reviewing literature and having discussions with experts in this area and others in the WCPFC delegations with strong interests. Ultimately, the aim will be to provide guidance to the SC on an approach or framework to consider longline effort creep and its implications and need for incorporation into WCPFC stock assessments and some options for this.

The meeting circled back to the south Pacific albacore assessment to revisit the assessment plan, the model sensitivity analysis and potential uncertainty grid to form the basis of management advice. This discussion was led by Thom Teears (SPC-OFP). Thom went back to some further considerations of the spatial patterns of CPUE in relation to the proposed spawner index (i.e. CPUE for quarter 1 (October – January for the new assessment) 10-25°S), he showed how at the south Pacific wide and annual scale the spatial patterns of CPUE vary substantially, however looking at the restricted spawner region and season, the CPUE spatial patterns are more consistent across decades. This provided further support that the spawner index is measure a specific component of the stock that is not subject to major differences in spatial occurrence. Thom delved further into the bimodal size distribution, noting that further splitting in the regression tree cannot remove the bimodality in the southern central Pacific region. Thom also looked at flag and quarter specific data for Chinese Taipei and the bimodality was still there for the southern region.

Thom showed some updated diagnostics for a model fitted to the CAAL data and with the VB offsets. He also showed a model where the length data were not used to estimate growth, and this showed that while the prediction of CAAL data looked better the fits to the troll data size modes was poor, so estimation of growth in the model benefits from the size composition data.

Thom then noted the various sensitivities and further modelling exploration they would seek to explore including, truncating the CPUE time series to start from the mid-1960s – 1970, placing higher CVs on the early years of CPUE, movement estimation in MFCL – noting they had run that model and the results didn't seem plausible – everything ended up going to the east, movement sensitivities around the SEAPODYM estimates, further consider the use of the Chinese or Chinese Taipei CPUE indices, further develop the stock synthesis model if time allows, think about possibility of sex specificity (yeh-nah), try a global CPUE index for the entire northern sub-area, exploration M-at-age methods Lorenzen/Maunder/various levels of asymptotic M from meta-analysis, run models for split CPUE indices (i.e. pre and post early 1990's).

The components of the potential uncertainty grid were then discussed, including; steepness, movement, size data weighting, CPUE, growth, M-at-age. It was also suggested that the assessment was aiming to explore a more ensemble Monte-Carlo approach using priors/joint prior distributions for uncertainties that this is amenable to rather than orthogonal/fixed grid values.

Overall, the PAW was supportive of the uncertainties to be considered, while noting that sensitivities would need to be conducted before final decisions on what to include in the uncertainty characterisation to put forward as the basis for management advice. The PAW noted the M-at-age with alternative growth coupling, and suggested that a max age M-based prior option could be considered, also that if one of data weighting methods could be chosen over the others, the data weighting axis could be removed. The PAW was keen to see the CPUE axis with the global northern region CPUE included (0-25°S) and asked for at least one option other than zero could be included for longline effort creep. Recruitment distribution was also suggested as a sensitivity that could be added.

The aim to move towards a prior distribution-based approach for drawing values for external inputs was encouraged and can help bring more objectivity into the models included in the uncertainty characterisation.

The PAW was concluded by the chair who thanked all the participants who travelled to Noumea, those joining online, presenters and SPC staff. The chair noted that more time for discussion would have been desirable, and an extra day or half day could be useful for next PAW, along with consideration of some changes to the format, including potential workshopping discussions on key problematic areas of assessments.

Appendices

Appendix 1. Post PAW – pre-SC communiques

1. PAW Group Communique 12/07/2024

2024 South Pacific albacore stock assessment

This is a brief update on the progress of the WCPFC 2024 south Pacific albacore assessments. This communique is aimed to provide key information on the progress and important technical elements and changes incorporated into the 2024 assessment, and to elicit any feedback that could be useful for us to consider prior to SC, noting much of the core assessment is now more or less finalised.

Progress summary

As of writing, the south Pacific albacore assessment is in good shape. Assuming no last-minute curve balls the computational work is mostly completed for all the models, including sensitivities, and we are now working on finishing some diagnostics and the report compilations for both the stock assessment paper and the background analysis/inputs paper. We hope to deliver the papers close to the deadline, but to forewarn, given the implications of the New Caledonia crisis a slight delay might be expected. We will keep you informed closer to paper submission due date of 27th July.

1. Core structure of the 2024 assessment model

Considerable changes have been made to the core structure of the assessment, largely in response to SC comments on the 2021 assessment, and comments by the PAW earlier this year. The main changes are summarised in Table 1 and the spatial configuration of the assessment is shown in Figure 1. Note the move to a major simplification of the assessment structure (recommended by SC) meant that the diagnostic model is quite different to the previous diagnostic model which was no longer a suitable starting point for a typical stepwise development/progression approach applied in previous assessments.

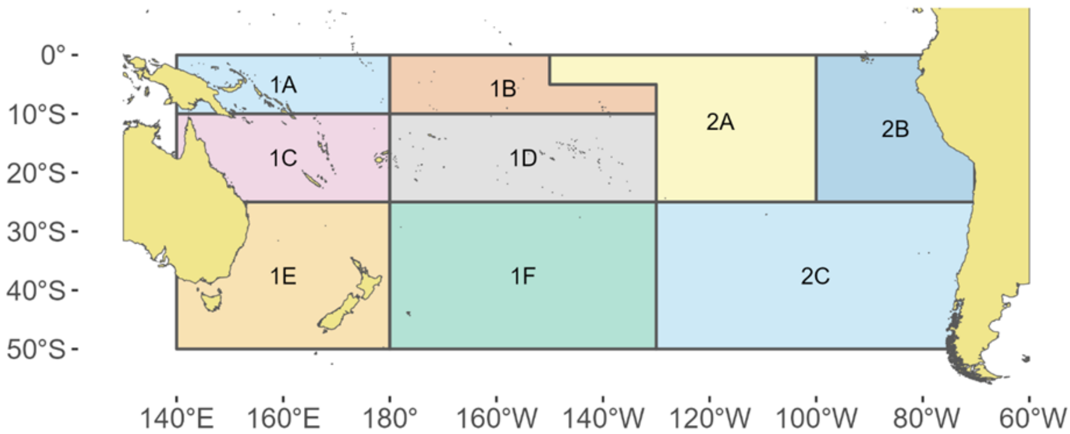

Figure 1. Definition of regions and sub-regions in the assessment. Numbers 1 and 2 indicate explicit regions in the assessment model. The letters (A, B, etc) indicate sub-regions used for the definition of fisheries.

Key assessment features and changes

Table 1. Summary of model features for the 2021 assessment and those for the 2024 assessment.

Feature	2021	2024	Comments
Spatial structure	4 regions	2 regions, WCPFC-CA, EPO	The simplification of the spatial structure was suggested by SC. The two explicit regions correspond to 1. WCPFC-CA, and 2. EPO. The "overlap" region is included in region 1.
Fisheries	21 capture, 4 index	17 capture, 3 index	LL fisheries defined by sub-region (see Figure 1) and fleet category. Fishery structures informed by spatial analysis of size composition data (regression tree approach).
Time window	1960-2019	Jan 1954- Oct 2022	Extension back to 1954 to meet the assumption of an unexploited initial population.
CPUE indices	LL, one per region	a) WCPO LL "spawning" – 10-25°S, Oct-Jan, NZ troll, EPO LL full year b) WCPO North – 0-25°S, full year, with, NZ troll, EPO LL full year (Note: the full year indices have month as a covariate in the sdmTMB model, and are applied as annual indices for month 8 in MFCL)	The alternative WCPO LL indices are used as an equally-weighted dimension (branch) in the model uncertainty ensemble. • Sensitivity analysis is conducted with and without the troll CPUE index.
Length data	1-cm bins, unfiltered, arbitrary likelihood weighting	2-cm bins, 50 minimum sample size per fishing incident, Francis weighting	LF data were initially screened to remove data obviously measured at 5-cm precision but reported as 1-cm precision. Estimation of LF data weighting by fishery applied the Francis's method (as implemented in SS3).
Catch estimation	Catch errors	Catch conditioned (no error)	

Selectivity	Time invariant,	Quarterly or semestral for some LL	Seasonal variation in selectivity is explored as a proxy for regular size-related movement within each model region. We are using seasonal selectivity in the main model. Sensitivities on which fishery(s) are used as asymptotic. *Sensitivity* • 3 selectivity time blocks for DWFN LL fisheries were trialled in a sensitivity analysis.
Catchability	Explicit, effort devs estimated	Implicit (NOTE: For projections using effort for fisheries with effort data the fishing mortality-effort-relationship predictions of catchabilities are now estimated as a second step on the converged models but with fixing all other parameters in the converged models to retain the original model estimated quantities intact.)	*Sensitivity* • Breaking LL CPUE indices into 3 x time blocks trialled in a sensitivity analysis. • Various levels of LL effort creep trialled in sensitivity analysis.
Growth	VB fixed, based on otolith age	a) L1 fixed, age-dependent variance fixed, estimate generic var, k, L12 and VB offsets for ages 2, 3 and 4 years. Include conditional age-at-length (CAL) data in the assessment model. b) Standard VB, all parameters estimated except age-dependent variance. CAL data included.	The two growth structures were explored, option a) is preferred due to superior fits to key data components. *Sensitivity* • Growth option b) is included as a sensitivity • Alternative weightings for CAL data.
Length – weight a and b parameters	a=6.9587e-06 b=3.2351 e	Updated to: a = 1.708e-05 b = 3.0483	Length -weight parameters updated with addition of approximately 6,000 new samples added to data set.
Maturity at length ogive	Farley et al. (2014), smoothed via a logistic curve.	No change	

Movement	Estimated, or provided by SEAPODYM	SEAPODYM	Movement was specified as age and season specific. *Sensitivity* • No movement and higher movement between the WCPO and EPO regions was trialled in sensitivity analysis.
Natural mortality	Specified age-specific pattern (Maunder et al. method)	Lorenzen, average M sampled from prior.	100 values of average M were sampled from a lognormal prior of mean 0.36 yr^{-1} and a CV (SD on the log scale) of 0.2. The prior mean was based on the Hamel and Cope Amax method, with Amax assumed to be 15 yr based the oldest fish aged using otoliths. Average M was converted to Lorenzen M-at-age and used as fixed parameters in each 100 model branch of the model ensemble (see below).
Recruitment	Quarterly, σ=0.7. Proportions fixed (SEAPODYM)	Annual (Oct), σ=0.7. Proportions fixed (SEAPODYM).	*Sensitivity* • The effect of assuming different fixed proportions was investigated in a sensitivity analysis.
Stock-recruitment	B&H, steepness 0.65, 0.8 or 0.95	B&H, steepness sampled from prior.	100 values of h were sampled from a censored (0.2-1.0) beta prior whose characteristics were informed by life history considerations and previous specifications based on meta-analysis. The 100 values of h were randomly combined with the 100 values of average M in each 100 model branch of the ensemble (see below).
Parameters estimated	4,637	166 or 164	

2. Diagnostics

The performance of the core model will be evaluated by applying a suite of diagnostics, including:

- Convergence diagnostics – maximum parameter gradient, positive definite hessian, minimal parameter correlations, parameters within bounds, sensitivity to starting values (jittering)
- Fits to CPUE, LF and conditional age at length data
- Annual age class mode interpretation for NZ troll LF
- Parameter correlations, within bounds
- Retrospective analysis
- Likelihood profile
- ASPM, catch curve analysis

3. Sensitivity analyses

A range of sensitivity analyses were conducted as follows:

A. Data weighting – troll CPUE
    - Run model with (i.e. diagnostic case) and without troll CPUE. Troll CPUE is down-weighted to a trivial level so that we can still see the fit to the data.
    - Additionally, test down-weighting of TR capture fishery LF data.

B. Data weighting – conditional age-at-length data
    - Run model with down-weighting coefficients of 0.5, 0.75 (diagnostic case) and upweighting to 1.0.

C. Movement
    - Sensitivities would be (a) zero movement and (b) movement probabilities that are double those indicated by SEAPODYM.
    - Additional sensitivity on running a 'WCPO only' single region model.

D. Effort creep
    - Assume annual cumulative (additive) catchability increase of 0.5%, 1.0% and 1.5% per year from 1954 to present (applied as adjustments to CPUE indices).

E. Recruitment distribution
    - Diagnostic case assumption is that recruitment is distributed 82:18 on average in regions 1 and 2, respectively (from SEAPODYM).
    - Test two contrasting proportions – 90:10 and 70:30.
    - Time-series variability is currently assumed to be the same in both regions.

F. Number of age classes
    - 12 age classes are assumed in the model, sensitivities with 10 and 15 age classes.

G. Non-decreasing selectivity
    - Currently the two LL index fisheries are assumed to have non-decreasing selectivity.
    - Change the non-decreasing selectivity to the 2 capture fisheries (one per region) that catch the largest fish.

H. Time-blocked selectivity/catchability
    - The analysis of HBF distributions (see attached) indicate major changes in about 1977 and 1994. Analyse the impact of time-blocking the selectivity of DWFN capture and index fisheries at these years as a sensitivity.
    - We can also split the indices at these years to test for catchability change across the time blocks.

I. Growth
    - Two alternative growth formulations (standard VB and VB with offsets) were explored. Initial trials indicated that better fits to LF and CAL data could be obtained

by allowing some deviation from von Bertalanffy growth for juvenile age classes. In this model we estimated "VB offsets" for age classes 2, 3 and 4, while also estimating L12, k and the generic variance of length at age. L1 and the parameter describing the age dependency of the variance of length at age were held fixed at values deemed to be reasonable. The second growth model assumed a more traditional standard von Bertalanffy approach (no offsets), estimating L1, L12, k and the generic variance parameter. This second growth approach was not preferred due to its poorer performance but is included in the uncertainty ensemble.

4. Ensemble model uncertainty framework

Previous assessments have characterised uncertainty in the stock assessment results using a factorial grid of discrete structural or fixed parameter settings for processes or data considered to have significant uncertainty. For this assessment we have adopted a model ensemble approach similar to the type described by Ducharme-Barth and Vincent (2022) and recommended by the SC19 paper of Neubauer at al. (2023). The estimation uncertainty for each ensemble model is included for the key management quantities as done in the 2023 yellowfin and bigeye assessments. The different uncertainty elements included in the ensemble approach applied here are:

Natural mortality – 100 values of average M sampled from a log-normal prior, converted to Lorenzen M-at-age (Figure 2a). In constructing this prior, we considered the views reached in recent CAPAM meetings on this topic. The report of the 2023 "tuna best practice" CAPAM workshop isn't published yet but the recording of the session discussing M best practice can be found at https://www.youtube.com/watch?v=_CFhMETAhDw&list=PLKeH-azh54PVgOjmJ1Gw4gmaCBQ0PDrz3&index=15 starting around the 27 min mark. The consensus seemed to be to (a) recognise the age-specific pattern in M by using the Lorenzen inverse-mean-length-at-age method; and (b) scale the average M using the Amax approach (at least when not possible to estimate internally, which we can't do for albacore). There seemed to be limited support for using other life-history info so the approach we took was to use Amax, as follows:

- We assumed an Amax of 15 yr based mainly on the observations of the oldest fish in the otolith aging data set – oldest 15.16 yr, with 2 others >14 yr. Allen Andrews also advised during the PAW that something "mid-teens" is appropriate for ALB, based on bomb radiocarbon observations in other oceans. That provides an expected value of average M of 0.36.
- In sampling average M's from the prior, we initially used the Hamel recommended CV of 0.31, but found that this produced a significant number of low <0.25 and high >0.7 values that we felt were outside the range of feasible albacore M. So, we opted to reduce the CV to 0.2 to focus the replicates on a more feasible range of about 0.25-0.55.
- Having sampled an average M from the prior, we did an estimation of what M12 would be assuming that the sampled M represents the simple average over ages 1-12 years. The way the Lorenzen is parameterised in MFCL is that one specifies LN(M12), so it was necessary to do this first. The model then used M-at-age provided by the Lorenzen assuming inverse proportionality with mean length at age. So, the Lorenzen pattern used in any particular model replicate was linked to the growth parameters internally estimated for that replicate. We trialled actually using the M as a real prior and estimating M internally conditioned on that. We found that the estimate drifted off on the high side, to the limit of what was allowed

by the prior, which didn't seem reasonable. Similar was found in the last North Pacific albacore assessment.

Steepness – 100 values sampled from censored beta prior, and randomly paired with the sampled M values (Figure 2b). We initially considered using the approach of Brodziak et al. (2011), that uses various life history parameters, including M-at-age, in characterising a steepness prior. With Jon's assistance, we applied this approach using SP albacore reproductive parameters, and M-at-age. An example of the steepness distributions that were obtained from this procedure is shown in Figure 3. The mode of the distribution is at 0.99 and initial runs of the Monte-Carlo approach were generating h values of >0.99 in around 20% of the replicates. This distribution also implies a non-trivial probability of very low steepness. Obviously, this represented a big change from the previous approach of using 0.65, 0.80 and 0.95 as discrete values in a factorial grid with equal probability. We found that MFCL was having trouble adequately fitting the SRR to the SB and recruitment estimates when using the very high values of h. And there were several replicates for which h had been sampled to be very low, <0.5, which were producing what we judged to be unreasonable estimates of population dynamics in the MFCL models. To proceed, we would have had to reject around 30% of the models on these grounds, which would have been effectively truncating the steepness prior on the low and very high ends. Our conclusion was that, on balance, it was better to modify the prior for steepness so that such high values >0.99 and low values <0.5 had much lower probability than as indicated in the distribution below. The approach that we therefore adopted was similar to one used by Ducharme-Barth and Vincent (2022). First, we wanted to recognise that the application of Jon Brodziak's approach using life-history criteria was indicating that steepness on average was likely to be considerably higher than what we had been assuming in previous assessments. Secondly, we also wanted to respect the previous thinking, including meta-analyses, that indicated a reasonable range of steepness was likely to be around 0.65-0.95, which had led us to our previous approach in the factorial grid. We therefore ultimately landed on using the prior shown in Figure 2b that takes these considerations into account.

CPUE indices – Two alternative WCPO LL CPUE indices, taken to be equally weighted. One index was based on LL fisheries data for the area 0-25°S over the entire year. The second index had more restricted spatial (10-25°S) and temporal (Oct-Jan) to focus the index on the SP albacore spawning area and season.

The CPUE alternatives provide two equally weighted core model branches for the ensemble. The same 100 sets of M-at-age and steepness values sampled from their priors were used with each of these two core models thus making up 200 potential models in the ensemble. We opted to use the same M and h couplets for each of the core models to be able to more clearly determine their respective impacts on the stock assessment results.

The 200 potential models in the ensemble were screened for inclusion in the final ensemble using the following criteria:

1. Models have a maximum parameter gradient of no more than 1e-04.
2. Models have a positive definitive hessian.
3. Models have a minimum acceptable fit to the CPUE indices.
4. Models fit the NZ troll fishery LF data in such a way that the annual age-class interpretation of clear length modes, where they occur, is preserved.

5. Models have an acceptable fit to the overall LF data in aggregate, bearing in mind the degree of down-weighting dictated by the Francis weighting method.
6. Models have a minimum acceptable fit to the conditional age at length data in aggregate.

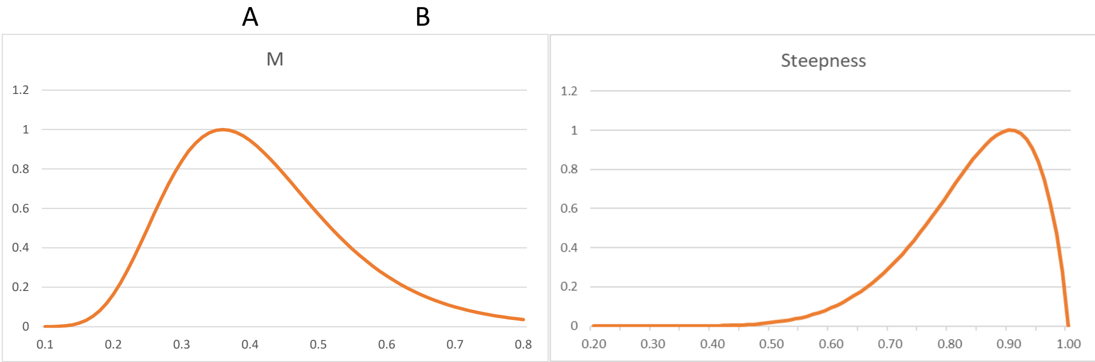

Figure 1. Prior distributions for a) average M; and b) steepness.

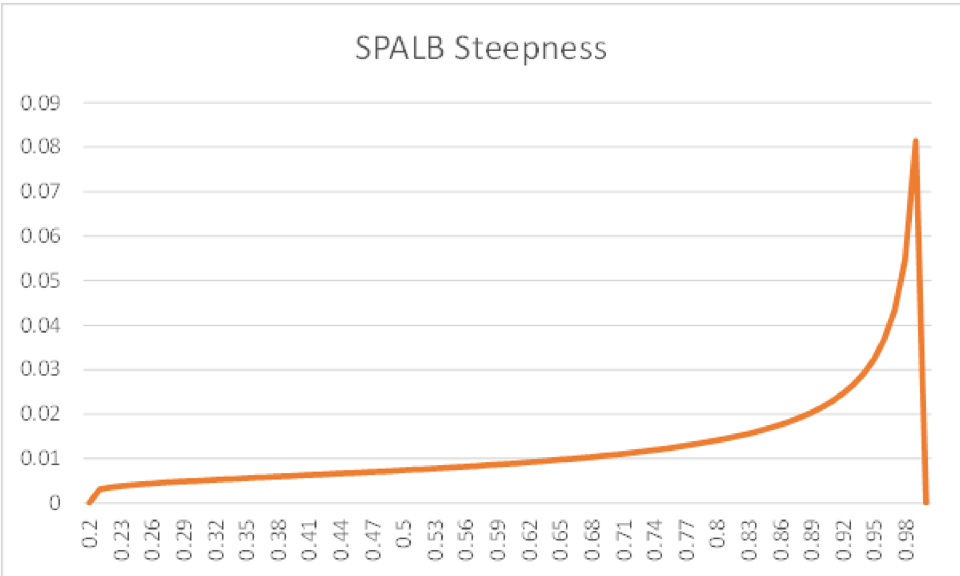

Figure 2. Prior for steepness derived using the procedure of Brodziak et al. (2011).

2. PAW Group Communique 18/07/2024

2024 southwest Pacific striped marlin stock assessment

This is a brief update on the progress of the WCPFC 2024 southwest Pacific striped marlin (MLS) assessment. This communique is aimed to provide information on the progress and important technical elements and changes incorporated into the 2024 assessment, and to elicit any feedback that could be useful for us to consider prior to SC, noting much of the core assessment is now more or less finalised.

Progress summary

As of writing, the southwest Pacific MLS assessment is close to completing all the analysis, and we are now working on finishing some diagnostics and the report compilations for both the stock assessment paper and the background analysis/inputs paper. Note that **we do not expect that these papers will be available by the due date, and we will be focussing on the assessment paper as the priority**. The crisis in New Caledonia has disrupted the MLS assessment work, as staff on this assessment have moved between countries and been mostly working remotely and separated from each other for the critical period over the last few months. We will keep you informed closer to the paper submission due date of 27th July. We are aligning approaches as closely as we can with the albacore assessment so that methods for generating the key outputs for albacore – such as the uncertainty ensemble can be applied to MLS.

5. Core structure of the 2024 assessment model

The core structure of the 2024 assessment is consistent with the previous assessment. Although analysis of longline size composition was conducted using the regression tree method (see SC20-SA-IP-03), the results did not provide a convincing argument to change the spatial/fishery structure from that used in the 2019 assessment, and the spatial-temporal coverage of data was poor or even non-existent for many spatial cells in the analysis.

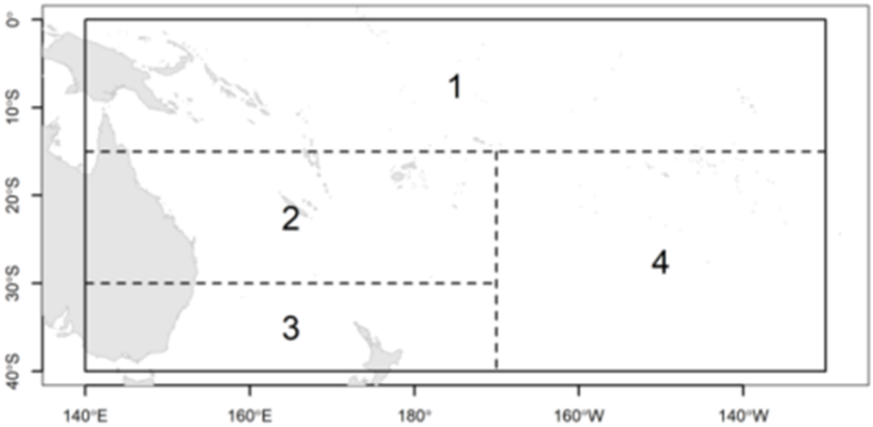

Figure 1. Spatial structure of the previous and current stock assessment for MLS. Note this represent a spatial structure for defining fleets as areas, the model is a single region model with no movement.

Key assessment features and changes

Table 2. Summary of model features for the 2019 assessment and those for the 2024 assessment.

Feature	2019	2024	Comments
Spatial structure	Single region, 4 sub-regions for fleets as areas	No change	LL size composition did not provide strong case to modify fleet structure of move to discrete spatial regions
Fisheries	14 capture	No change	Rationale for fishery structure in previous assessment was sound.
Time window	1960-2019	Jan 1979- Oct 2022	This is a significant change in the 2024 assessment. While we embarked a model staring in 1952 to try to include all the historic fishing data, the CPUE in the early period from Japan was not considered reliable and there were no size data, plus a very large single year spike in catch, which continues to be questionable and problematic. The full model (start in 1952) predicted very large implausible recruitment spikes and high early biomass which dropped implausibly rapidly. We considered maintaining the early period catches – as a kind of burn-in phase – and over time the model estimates did settle to be consistent with the previous assessment model. However, we were ultimately not comfortable with including these clearly unreliable early estimations. We therefore explored a model starting in 1979 when CPUE and size data were available/and or considered more reliable. • We explored the sensitivity of the model with the shorter time period (1979) to alternative starting conditions of total morality centred around a level that provided a starting depletion similar to the previous diagnostic model and the full time series model in 1979.
CPUE indices	LL, one per sub-region based on different DWFN fleets, applied as quarterly in MFCL, VAST	1 combined JP/TW LL for entire model region starting in	• Sensitivity analysis with AUS LL index, although, discussion on the derivation of the index during the PAW seemed to indicate that it was based on a generic standardization model, and that it really needed a standalone alone analysis to have confidence to apply for an

		1979, applied as annual in MFCL, sdmTMB	assessment. Based on this it is not included in the ensemble for management advice.
Length data	6 cm bins, unfiltered, raw (neither catch weighted, nor CPUE weighted) Arbitrary data weighting grid axis,	5 cm bins, minimum sample size of 10 (per fishing incident), sample sizes reweighted using Francis method (as implemented in r4SS) and no size composition weighting axis used in the uncertainty grid.	Size composition data for each fishery was weighted using the Francis method (as implemented in r4ss)
Catch estimation	Catch errors	Catch conditioned (no error)	Key change – all SPC tuna/billfish assessments moving to simpler catch conditioned models.
Selectivity	Time invariant, LL asymptotic for NZ recreational fishery	No change, possible sensitivities on seasonal selectivity	
Catchability	Explicit, effort devs estimated	Implicit	
Growth	External, VB from Kopf et al. 2011 based on spines, plus alternative combined otolith/spines growth curve	External, updated to a full otolith-based VB from Farley et al. (2021) as recommended at PAW. No alternative growth is available for south Pacific.	The new growth is a notable change and involves a higher k and Linf.
Length – weight a and b parameters	a = 4.4989964037849e-07, b = 3.61648435378567	Updated to one based on direct EFL v weight data (see adjacent comment). Very minor change.	Explored data set from Kopf for n = 114 coupled EFL and WW records that were measured directly to EFL on MLS captured in NZ and Australian waters between 2005 and 2008. a = 5.399420e-07 (95% CIs: 2.338e-07, 1.247e-06) b = 3.583776 (95% CIs: 3.425, 3.743)
Maturity at length ogive	Based on Kopf et al. (2012)	Updated, from Farley et al. (2021) as recommended at PAW.	The updated maturity at length ogive moved the L50 to a slightly larger size and had a steeper transition from immature to fully mature.

Natural mortality	External age specific with age deviate based on Piner and Lee (2011), average M 0.3, 0.4, 0.5, diagnostic model was 0.4.	Lorenzen – using an M prior based on A_{max} (Cope and Hamel 2022), max age 15 years (Farley et al. 2021)	M cannot be estimated internally for the striped marlin assessments. Applied approach as per albacore. 100 values of average M were sampled from a lognormal prior of mean 0.36 yr^{-1} and a CV (SD on the log scale) of 0.2. The prior mean was based on the Hamel and Cope Amax method, with Amax assumed to be 15 yr based on the oldest fish aged using otoliths. Average Ms drawn from the prior were converted to Lorenzen M-at-age matching the prior M to the average of ages 2-10 years (age at 50% maturity age of plus group).
Recruitment	1 per year, Recruitment penalty 0.2, 0.5 and 2.2 included in uncertainty	1 per year Diagnostic 0.2	*Sensitivity* Will run sensitivities on this as per previous assessment.
Stock-recruitment	B&H, steepness 0.65, 0.80, 0.95 in uncertainty grid	B&H, steepness sampled from prior (modified based on Brodziak (2015) method, incorporates M distribution from A_{max})	100 values of h were sampled from a censored (0.2-1.0) beta prior whose characteristics were informed by life history considerations and previous specifications based on meta-analysis. The 100 values of h were randomly combined with the 100 values of average M in each 100 models used in the uncertainty ensemble
Parameters estimated	2648	76	Big reduction in parameters – due shorter time period, switch to catch conditioned model

6. Diagnostics

The performance of the core model will be evaluated by applying a suite of diagnostics, including:

- Convergence diagnostics – maximum parameter gradient, positive definite hessian, minimal parameter correlations, parameters within bounds, sensitivity to starting values (jittering)
- Fits to CPUE and size composition data
- Parameter correlations, within bounds
- Retrospective analysis
- Likelihood profile
- ASPM, catch curve analysis

7. Sensitivity analyses

A range of sensitivity analyses were conducted as follows:

- AUS CPUE
- Recruitment penalty
- Assumed starting stock conditions for the 1979 model

8. Ensemble model uncertainty framework is consistent with south Pacific albacore approach

Natural mortality – 100 values of average M sampled from a log-normal prior, converted to Lorenzen M-at-age. Average M derived as per south Pacific albacore using the Amax method, as follows:

- Amax of 15 yr based mainly on the observations of the oldest fish in the otolith aging data set That provides an expected value of average M of 0.36, same as albacore.
- Reduce the CV to 0.2 to focus the replicates on a more feasible range of about 0.25-0.55, as per albacore.
- Having sampled an average M from the prior, we did an estimation of what M10 (MLS plus group age) would be assuming that the sampled M represents the simple average over ages 2-10 years (based on recent feedback on the approach described in the previous communique for albacore, apply the sampled M from the age at maturity). The way the Lorenzen is parameterised in MFCL is that one specifies LN(M10), so it was necessary to do this first. The model then used M-at-age provided by the Lorenzen assuming inverse proportionality with mean length at age. The growth is fixed as an external input for the MLS assessment.

Steepness – as per albacore. We initially had a go at the Brodziak et al. (2011) prior estimation, but as per the albacore, it didn't seem satisfactory and needs more time than we had to explore in depth. It estimated a slightly lower expected *h* than albacore, we created a prior along the lines of that described for the albacore in the previous communique, with some slight modifications to reflect the indication of lower expected *h*.

NOTE: Given the desire to move to an ensemble approach (Ducharme Barth and Vincent 2022, Neubauer et al. 2023) we suggest that further work is required, perhaps as projects under the various tuna, shark and billfish research plans to develop guidance for creating these prior distributions – especially for steepness and M.

Appendix 2: PAW Agenda

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

2024 SPC Pre-assessment Workshop Agenda (original version) 25th- 28th March,

Monday 25th March (Sun 24th US)	**DAY 1: 2024 south Pacific albacore assessment**	**Presenter initials and presentation number**
09:00 – 09:10	**Introduction** • Reminder of TOR and objectives for the SPC preparatory workshop • Agenda and meeting format/procedures • Any other introductory comments	PH
09:10 – 10:30 *Session 1* *(80 mins)*	**Background** • Background south Pacific albacore fishery and stocks (10 mins) • Previous south Pacific albacore assessment summary and SC17 recommendations (15) mins) • Discussion (10 mins) **Spatial structure and movement** • Research on connectivity between the WCPO and EPO (15 mins) • Albacore and SEAPODYM (20 mins) • Discussion (10 mins)	• PH/TT **(P1)** • JM/GA **(P2 - online)** • IS **(P3)**
10:30 – 11.00	BREAK	
11.00-11.40 *Session 2* *(40 mins)*	**Spatial structure and movement** • Revising the spatial and fisheries structure (15 mins) • Discussion (20 mins)	• TT/JP **(P4)**

11.40 -13.00 *Session 3* *(80 mins)*	**Data Inputs** ● Data summaries by defined fisheries (20 mins) ● Size composition data issues and pre-treatment (20 mins) ● Discussion (15 mins)	● TT **(P5)** ● TP **(P6)**
	● NZ troll fishery size comp analysis (15 mins) ● Discussion (10 mins)	● PN **(P7)**
13.00-14.00	Lunch BREAK (60 mins)	
14.00-16.30 *Session 4* *(140 mins)*	**CPUE indices (part 1) (80 mins)** ● NZ troll CPUE (15 mins) ● Chinese longline CPUE (15 mins) ● Chinese Taipei longline CPUE (15 mins) ● Other flags longline summary (15 mins) ● Discussion (20 mins) **CPUE indices (part 2) (60 mins)** ● Developing the multi-fleet abundance indices (30 mins) ● Discussion (30 mins)	● PN **(P8)** ● Hongyu Lin **(P9)** ● Yi-Jay Chang **(P10 - online)** ● TT **(P11)** ● TT **(P12)**
16.30-17.00 (30 mins)	**Buffer time**	
17.00	**Conclude/wrap-up day 1**	PH
Tuesday 26th March (Mon 25th US)	**DAY 2: 2024 south Pacific albacore assessment (continued) and WCPO silky and oceanic white tip shark assessments**	
8.30-10.00 *Session 5* *(90 mins)*	**Biology south Pacific albacore** ● Age and growth data summary, alternative growth model (20 mins) ● Exploring spatially distinct growth patterns of albacore across the Southern Pacific (15 mins) ● SPC's new otolith lab R and D program and some insights on albacore age (15 mins) ● Growth discussion (20 mins) …………………………………………………………	● TT/JH (**P13**) ● Dongqi Lu **(P14)** ● AA **(P15)**

	• Natural mortality approach (15 mins)	• **TT (P16)**
10:00 – 10.30	BREAK	
10.30-11.10 *Session 5 continued* *(50 mins)*	• Discussion – natural mortality (15 mins) ……………………………………………… • Maturity/reproductive biology (10 mins) • Length-weight conversions (project 90 update) (10 mins) • Discussion (10 mins)	• TT **(P17)** • JM **(P18 - online)**
11.10 – 13.00 *Session 6* *(110 mins)*	**Model development** • Key changes from 2021 • Preliminary model exploration and results • Proposed next model development steps • Data weighting • Sensitivity tests and uncertainty characterization • Diagnostics Including discussion	• TT/JH **(P19)** ALL

Appendix 3: List of participants

Name	Affiliation	In person (IP) / Online (O)
Leyla Knittweis	Ministry for Primary Industries, NZ	IP
D'Arcy Webber	Quantifish, NZ	IP
Philipp Neubauer	Dragonfly Data Science	IP
Kath Large	Dragonfly Data Science	O
Kyuhan Kim	Dragonfly Data Science	IP
Tyla Hill-Moana	Dragonfly Data Science	IP
Steven Brouwer	Sagittas LTD	O
Nick Davies	Takina LTD, consultant	IP
Jessica Farley	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	O
Adele Dutilloy	FFA	IP
Keith Bigelow	NOAA (Pacific Islands Fisheries Science Centre), US	O
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
Moses Mataika	Fiji	IP
Sera Waqa	Fiji	IP
Mickael Lercari	New Caledonia	IP
Tom Peatman	Independent Consultant	IP
Fayakun Satria	Indonesia	IP
Lilis Sadiyah	Indonesia	IP
Adrian Hordyk	Blue Matters	O
Eric Chang	National Sun Yat-sen University, TW	O
Yi-Jay Chang	National Sun Yat-sen University, TW	O
Wei-Chuan Chiang	TW	O
Jia-Rong Wu	TW	O
Atsushi Tawa	Japan Fisheries Research and Education Agency	IP
Hidetada Kiyofuji	Japan Fisheries Research and Education Agency	IP
Yoshinori Aoki	Japan Fisheries Research and Education Agency	IP
Haewon Lee	Korea	O
Heewon Park	Korea	O
Jung-Hyun lim	Korea	O
Youjung Kwon	Korea	O
Domingo Ochavillo	American Samoa	O
Berry Muller	Marshall Islands	O
Francisco 'Curro' Abascal	EU	IP
Fan Zhang	Shanghai Ocean University	IP
Hongyu Lin	Shanghai Ocean University	IP

Donqui Lu	Shanghai Ocean University	IP
SungKwon Soh	WCPFC Secretariat	IP
Elain Garvilles	WCPFC Secretariat	O
Glen Holmes	PEW	O
Allen Andrews	SPC	IP
Claudio Castillo Jordon	SPC	IP
John Hampton	SPC	IP
Graham Pilling	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	O
Steven Hare	SPC	IP
Joanne Potts	SPC	O
Inna Senina	SPC	IP
Lucas Bonnin	SPC	IP
Marino Wichman	SPC	IP
Jennyfer Mourot	SPC	IP
Giulia Anderson	SPC	IP
Caroline Ton	SPC	IP

Appendix 4: 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.

References

Brodziak, J., Mangel, M. and Sun C. 2025. Stock-recruitment resilience of North Pacific striped marlin based on reproductive ecology. Fisheries Research 166, 140-150.

Castillo Jordan, C., Hampton, J., Ducharme-Barth, N., Xu, H., Vidal, T., Williams, P. Scott, F., Pilling, G. and Hamer, P. 2021. Stock assessment of South Pacific albacore tuna. WCPFC-SC17-SA-WP-02.

Ducharme-Barth, N., Pilling, G. and Hampton, J. 2019. Stock assessment of SW Pacific striped marlin in the WCPO. WCPFC-SC15-SA-WP-07

Ducharme-Barth, N., and M. T. Vincent. 2022. Focusing on the front end: A framework incorporating uncertainty in biological parameters in model ensembles of integrated stock assessments. Fisheries Research: 255, 106452. https://doi.org/10.1016/j.fishres.2022.106452

Clarke, S., Langley, A., Lennert-Cody, C., Aures-de-Silva, A. and Maunder, M. 2018a. Addendum to Pacific-wide silky shark (*Carcharhinus falciformis*) stock status assessment. WCPFC-SC14-SA-WP-08a

Clarke, S. Langley, A., Lennert-Cody, C., Aures-de-Silva, A., and Maunder, M. 2018b. Pacific-wide silky shark (*Carcharhinus falciformis*) stock status assessment. WCPFC-SC14-SA-WP-08b

Cope, JM. and Hamel, OS. 2022. Upgrading from M version 0.2: An application-based method for practical estimation, evaluation and uncertainty characterization of natural mortality. Fisheries Research 256 https://doi.org/10.1016/j.fishres.2022.106493

Farley, J. Krusic-Golub, K. and Eveson, P. 2021a. Updating age and growth parameters for South Pacific albacore (Project 106). WCPFC-SC17-SA-IP-10.

Farley, J., Eveson, P., Krusic-Golub, K. and Kopf, K. 2021b. Southwest Pacific striped marlin population biology (Project 99). WCPFC-SC17-SA-IP-11.

Francis et al. (2022) Post-release survival of shortfin mako (*Isurus oxyrinchus*) and silky (*Carcharhinus falciformis*) sharks released from pelagic tuna longlines in the Pacific Ocean. Aquatic Conservation: Mar Freshwater Ecosystems, 33, 366–378.

Joung, S. J., Chen, C. T., Lee, H. H., and Liu, K. M. 2008. Age, growth, and reproduction of silky sharks, *Carcharhinus falciformis*, in northeastern Taiwan waters. Fisheries Research, 90, 78–85. doi:10.1016/j.fishres.2007.09.025

Neubauer, P., Kim, K., Large, K. and Brouwer, S. 2023. Analysing potential inputs to the 2024 stock assessment of Western and Central Pacific silky shark (*Carcharhinus falciformis*). WCPFC-SC19-SA-WP-10-Rev 1.

Neubauer, P., K. Kim, T. A'mar, and K. Large. 2023. Addressing uncertainty in WCPFC stock assessments: Review and recommendations from WCPFC Project 113. WCPFC/SC19-SA-WP-10. https://meetings.wcpfc.int/file/13043/download

Oshitani, S., Nakano, H., and Tanaka, S. 2003. Age and growth of the silky shark *Carcharhinus falciformis* from the Pacific Ocean. Fisheries Science, 69, 456–464.

Peatman, T., Bell, L., Allain, V., Caillot, S., Williams, P., Tuiloma, I., Panizza, A., TremblayBoyer, L., Fukofuka, S., and Smith, N. 2018. Summary of longline fishery bycatch at a regional scale, 2003-2017. WCPFC-SC14--WP-03.

Rice, J. and Harley, S. 2012. Stock assessment of oceanic whitetip sharks in the western and central Pacific Ocean. WCPFC-SC8-SA-WP-06 Rev 1.

Tremblay-Boyer, L., Carvalho, F., Neubauer, P., and Pilling, G. (2019). Stock assessment for oceanic whitetip shark in the Western and Central Pacic Ocean. WCPFC-SC15-SA-WP-06.