SCIENTIFIC COMMITTEE

EIGHTEENTH REGULAR SESSION

Online meeting

11-18 August 2022

Report from the SPC Pre-assessment Workshop - March 2022

WCPFC-SC18-2022/SA-IP-02

6 July 2022

Report compiled by Paul Hamer

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

Report from the SPC Pre-assessment Workshop (PAW), March 2022

Oceanic Fisheries Programme, Pacific Community (OFP, SPC)

Introduction

To help guide stock assessment and related modelling work 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 and representatives from regional fisheries organisations through the SPC pre-assessment workshop (PAW) process. The fourteenth PAW was held from the 29th March – 31st March 2022. Due to the COVID-19 pandemic, the meeting was held electronically.

Day 1 focused on data inputs and modelling approaches for the assessment of skipjack tuna (*Katsuwonus pelamis*) in the WCPFC – Convention Area (WCPFC-CA) lead by SPC-OFP (Paul Hamer, Claudio Castillo Jordan, John Hampton and Thom Teears), with additional contributions from Tom Peatman (independent consultant), and Aoki Yoshinori (Japan Fisheries Research and Education Agency (FRA)). A background paper to the skipjack assessment was circulated prior to the PAW and is provided as appendix 4 to this paper. Joe Scutt Phillips (SPC-OFP) provided a presentation on a novel simulation approach for estimating tag mixing periods. Day 2 focussed on data inputs and modelling approaches for the assessment of shortfin mako shark in the southwest Pacific Ocean (WCPFC-CA), along with follow-up work to refine the model grid used to characterise uncertainty and provide management advice for the 2021 stock assessment of blue shark in the southwest Pacific Ocean. Mako shark presentations lead by Dragonfly Data Science (Kath Large) and Sagittus (Steve Brouwer) and blueshark by Dragonfly Data science (Philipp Neubauer) and NOAA (Felipe Carvalho). Day 3 included an overview of recent developments and consolidation work on the MULTIFAN-CL (MFCL) stock assessment software presented by Nick Davies (Takina), a background and overview of the process for the peer review of the 2020 yellowfin tuna assessment in the WCPFC-CA (SPC-OFP, Arni Magnusson), presentation of approaches for providing open, repeatable and transparent stock assessments (SPC-OFP, Arni Magnusson), and a session on progress with MSE modelling and harvest strategy development for the four target tuna stocks in the WCPFC-CA (SPC-OFP, Rob Scott, Finlay Scott). The meeting was chaired by Paul Hamer (SPC-OFP). The meeting operated under the terms of reference provided in Appendix 3.

This report briefly summarises the various presentations and related comments and discussions with a focus on the important issues and suggestions made. The report does not attribute comments to countries or individuals except where the comment is related to the agreement to provide data or to undertake particular analyses. The stock assessment scientists will address or respond to the recommendations provided in this report to the extent possible. It must be noted that the extent to which suggestions can be incorporated into the stock assessments or other related work prior to the Scientific Committee (SC18) will be constrained by the ability of the models to converge under the assumptions required, and the data, time and staff resources available. It should also be noted that the PAW is conducted in the early stages of the development of the assessments, and issues will often come up that are not raised or discussed in the PAW and may change the direction and approaches for the assessment. The stock assessment team may consult with particular individuals and countries as required to provide clarity on specific issues, particularly those related to data inputs. Ultimately the final decisions on model development, data inclusion, 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.

The comments and discussion related to the various presentations are noted in the document, with the comment/suggestions from the PAW participants plain text followed by the responses in italics. References in text are hyperlinked.

The outcomes of this meeting will be reflected in the papers submitted to WCPFC-SC18. Copies of presentations prepared for the workshop can be provided on request to Paul Hamer (paulh@spc.int).

Workshop agenda is in appendix 1.

DAY 1 – 2022 WCPFC-CA skipjack tuna stock assessment

Stock assessment overview presentation

Paul Hamer (SPC-OFP) provided a background presentation for the 2022 skipjack stock assessment. The presentation provided an overview of the skipjack fishery, catch history by gears and areas, fishing methods, selectivity patterns, and the most recent 2019 assessment (final data year 2018, Vincent et al. 2019a). Key points relating to the outcomes of the 2019 assessment were that the skipjack stock in the WCPFC-CA was not considered to be overfished or undergoing overfishing according to the reference points used by the WCPFC. The stock was estimated to have declined notably from the mid-2000s, but had remained relatively stable over recent years.

The 2019 assessment applied two spatial structures: a 5-region (as per 2016 assessment) and a new 8-region structure (Fig. 1). However, at SC15, the 8-region model structure was preferred for provision of management advice. **Scientific Committee (SC) 15 "*agreed to use the 8 region model to describe the stock status of skipjack tuna because SC15 considers that it better captures the biology of skipjack tuna than the existing 5 region structure*"** *(SC15 Summary Report).*

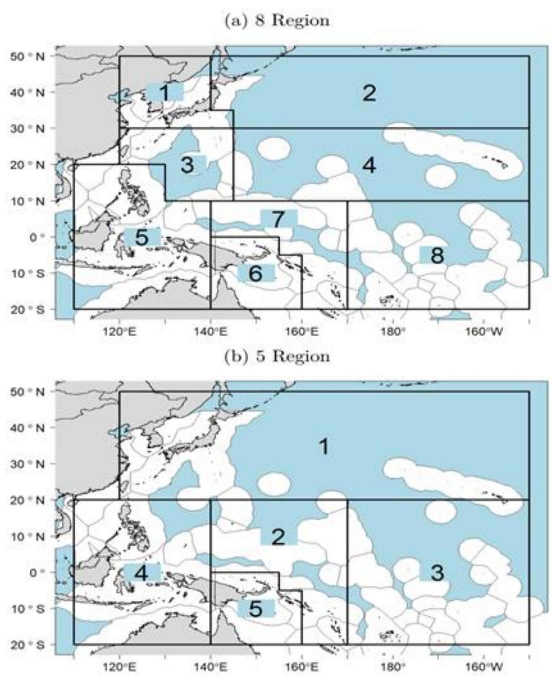

Figure 1. a) 8-region and b) 5-region model structures used in the 2019 WCPFC-CA skipjack stock assessment.

The 2019 assessment highlighted some areas of concern/deficiency that warranted further consideration. These included:

- Conflict among data sources, with tagging data favouring more optimistic stock status than size composition and CPUE.
- Increasing trends in estimated recruitment overtime and potential buffering by large model regions (i.e., region 4) with low fishing pressure but limited information to inform movements to and from.
- Hyperstability in CPUE indices (i.e., Japan pole and line CPUE) and spatial contraction of the Japanese pole and line fishery that provided the main CPUE indices of abundance.

The assessment included uncertainty with respect to growth, tag mixing period, steepness of the stock recruitment relationship, and weighting of the size composition data. Of these, the most influential were the tag mixing period and the choice of growth model. As such the assessment recommended further work to improve the understanding of tag mixing periods and growth parameters. Other practical suggestions included working with fisheries agencies on the efficiency of sending recaptured tags to SPC for validation to maximise the tagging data available for each assessment, making adjustments to tag input file so that time at liberty for individual recaptures is better represented in relation to the tag mixing periods that are specified at quarterly resolution for MFCL (i.e., accounting for tags released at the end of one quarter and recaptured early in the adjacent quarter, tags released at the start of one quarter and recaptured at the end of that quarter). The assessment also recommended further exploration and analysis of CPUE for the Japanese pole and line fishery and the equatorial purse seine fishery, with the view that purse seine CPUE may provide an alternative abundance index to pole and line CPUE in the equatorial model regions where pole and line effort has become negligible.

Following the background presentation, information was present of the proposed diagnostic model regional and fishery structures. The 8-region structure, preferred by SC15, was proposed for the diagnostic model development, with the same fishery structure for the *extraction* fisheries (31 fisheries) as applied in the 2019 assessment. Various options to be considered to provide CPUE indices of relative abundance (referred to herein as 'survey fisheries'), including the Japan pole and line and equatorial purse seine, and Philippines purse seine for region 5. Summaries of catch and length composition data were provided for each of the 31 extraction fisheries. An issue with the pole and line fishery data for region 4 (PL ALL 2) was evident in the data summary where it appeared that data for the Japanese distant water fishery (JP DW) may have been missing since 2004. *Follow up with Japanese scientists and Peter Williams indicated that from 2004 the catches from both fisheries were combined, so the overall catches are correct*.

Discussion

Discussion on the background and spatial/fishery structure presentation generally focussed on the topics of spatial structure and biological uncertainty, specifically growth.

The question was raised as to whether or not there would be a discussion on biological information for this skipjack assessment. *It was noted that there is no new information on skipjack growth, reproductive* *biology, or population structure in the WCPO since the last assessment, so there was nothing new to discuss or present to PAW 2022. Options for modelling growth to be considered later in the day.*

It was noted that work is currently underway in the Eastern Pacific Ocean (IATTC) analysing tag-recapture growth increment data from purse seine fishery recaptures. This work appears to suggest that growth rates are relatively constant across sizes of tagged and recaptured fish in the study. However, the study lacks the larger size fish that are caught in the long line fisheries. *OFP indicated they will also analysis tag-recapture growth increments to see how they compare with growth estimated by MFCL. OFP staff will make contact with the IATTC staff to compare methods and growth estimates. Another study on skipjack growth in the Indian ocean has stalled, and new information on reproductive biology from a Japanese sampling programme is not expected to be available for the current assessment.*

It was recommended that to provide more confidence in the growth estimates from MFCL, that the assessment team go back to the literature and pull together the existing information on skipjack growth to support the development of growth models perhaps as priors and/or for comparison with estimates from tagging data and the MFCL based estimations. *OFP to do a meta-analysis on skipjack growth from the literature to provide alternative growth option(s).*

It was noted that selectivity assumptions are very important, and the assessment was unstable without inclusion of the longline fishery size composition that includes the large sized fish and with the assumption of asymptotic selectivity improves fits to other gear selectivity's and model stability. There are still some issues around estimating growth inside the model to be discussed later, and further explored in this assessment.

Several alternative thoughts on spatial structure were expressed. One suggested that simplification of the spatial structure would be good to explore because of the convergence issues (lack of positive definite Hessian solution) with the 8-region model. It was suggested to try a model where regions 1-4 of the 8-region model are collapsed into a single region. This structure may be less sensitive to the lack of movement information for region 4 and the related biomass accumulation in this region. However, concerns were expressed regarding collapsing model regions, particularly in relation to how the model uses the tagging data and the very influential assumptions related to tag mixing. Simpler regional structures will have implication for tag mixing assumptions that will need to be considered but are very important given the sensitivity of the model to these assumptions. *OFP indicated exploring a 4-region spatial structure (collapsing region 1-4) could be done as a sensitivity, with key results and diagnostics added as an appendix in the stock assessment paper.*

Composition data

Tom Peatman provided a presentation on the treatment of size composition data for the skipjack assessment. The presentation outlined how fishery-dependent size composition data can be biased due to uneven sampling in relation to the spatial and temporal distribution of the catches and the different gears that select for different size fish (for extractions fisheries), or the relative abundance patterns across space and time (survey fisheries). Statistical reweighting of size composition data is therefore important to improve how the size composition data used in the assessment represents the composition of the harvested fish (extraction fisheries) or the relative abundance (survey fisheries). The reweighting procedure presented for the 2022 assessment was not applied in the 2019 assessment. The approach for reweighting purse seine composition is consistent with that applied to the 2020 bigeye and yellowfin tuna assessments (Peatman et al. 2020). For the pole and line fishery, the approach used is consistent with that applied to the long line fishery in the 2021 south Pacific albacore assessment (Vidal et al. 2021). At this stage of the assessment work, the reweighting has been conducted for the extraction fisheries as the survey fishery CPUE indices are not finalised. The reweighted compositions will be re-generated when the complete data become available in late April. The reweighting generates length compositions for both the robust normal and self-scaling likelihood approaches. A 50% reduction in sample sizes is also applied to account for use of samples in both extraction and survey fisheries where this occurs.

Discussion

Discussion on size composition analysis considered how selectivity was being modelled in relation to the extraction and survey fisheries and any variation in compositional data over time, also considering the implications for effective sample sizes (i.e., estimating time-varying selectivity can result in very good fits to compositional data and as a result higher effective sample sizes*). It was noted than in past skipjack assessments selectivity has typically been modelled as time invariant, and that this is generally suitable for the more consistent equatorial fisheries that take most of the catch. However, there is seasonal dynamics in the north Pacific that appear to relate to availability rather than selectivity, which may be better modelled with seasonal fisheries.* It was also noted that a key concern is to try to avoid a bad lack of fit to size composition data as this can lead to data conflicts and issue with biomass scaling. This may mean that excluding or down weighting composition data for periods where strange patterns in the composition data or poor/uncertain data occur is worth considering. *We do also note that lack of fit alone is probably not sufficient for excluding data but more so where lack of fit corresponds to composition data that are limited and/or of poor/unreliable quality.*

Region 3 pole and line composition data shows high variation, particularly for periods where sampling was relatively high (i.e., 1970s and post-2000). This was suggested to be related to the periodic occurrence of larger spawning fish as well as smaller recruits. While seasonal selectivity could be tried, this variation is not obviously seasonal and is due to variation in availability rather than selectivity. In this case we may have to accept a poorer fit to composition data for region 3 pole and line, rather than increasing parameterisation.

Tagging data

Thom Teears provided an overview of the tagging data preparation for the skipjack assessment, noting that the preparation would follow similar steps and methods as conducted in the 2019 assessment. The first step in the process to generate the tag data input file involves identifying the recaptures that are unusable, that is, tag recaptures that do not have associated meta-data on recovery date (year and month), location (latitude and longitude), fleet (gear and flag), and tag number. These unusable recaptures need to be accounted for in the tag release data. Essentially, the adjustments for unusable recaptures preserves the recovery rates of tags from the individual tag groups that would otherwise be biased low when an often significant proportion of recaptures cannot be assigned to a recapture category in the assessment. The adjustments to the tag releases are conducted by length bins of releases (2 cm bins for skipjack).

Once the usability adjustments are made, further adjustments to tag release numbers are made to account for tagging mortality, tagger effects (i.e., various effects related to the experience of the tagger, tagging station on a vessel, release condition of the fish etc. that influence the probability of the tagged fish being recaptured), and tag shedding rates. Further analyses or assumptions from literature are required to inform these latter adjustments. The analysis used for informing tagger effects adjustment is discussed in more detail below.

In the previous skipjack assessment, a base tagging mortality rate of 0.07 was applied. This value originates from the paper by Hoyle (2011). Further research is however required to better understand the tag-related mortality rate. The 0.07 tag-related mortality rate is assumed as applying to expert taggers, and the value is adjusted to account for tagger effects (i.e., some taggers/tagging events induce higher tag-related mortality than others, discussed below). Instantaneous tag shedding rate (0.0697) was estimated previously from double tagging experiments (Vincent et al. 2019b) and this value will be applied as the base tag shedding rate. While MFCL now has the ability to include chronic tag shedding, we do not have information on this, and for the short times at liberty of most skipjack recaptures it may be less of an issue that for other species/assessments. Chronic tag shedding will not be added into this assessment.

For the current assessment it is expected to use tagging data from release events up to and including the releases made during the 2019 western Pacific tagging cruise (WP5). This cruise included releases in three model regions, greater proportions of free-school releases and larger fish (>50 cm) than previous cruises.

There were no comments/discussion on the tagging data initial treatment steps.

Tagger-effects

Tom Peatman provided a presentation describing the analysis used to inform the adjustments of tag releases to account for tagger effects. Tagger effects occur on top of the base levels of tag shedding and tagging-related morality mentioned previously, and can be due the condition of the fish, placement/quality of the tag insertion, identity of the tagger and tagging station on the vessel. These factors can affect the tag retention and mortality of the fish and therefore its probability of becoming a recapture. The tagger effects analysis aims to estimate correction factors for tag release events that account for these factors. The approach used is described in Berger et al. (2014). Review of the approach was conducted in a 2020 SPC convened tagger effects workshop, which recommended further analyses to explore the modelling approach. Other recommendations included: using separate models to estimate correction factors for central and western Pacific cruises, undertake simulation studies to explore the ability of the modelling approach to separate tagging event effects from tagger and station effects, and explore the use of random effects (e.g., for tagging event, tagger). The new work conducted by Tom Peatman and described in the presentation focussed on the simulation studies to test whether the approach developed by Berger et al. (2014) was robust to issues related to imbalances with respect to tagging event and taggers and correlations between tagger experience levels and tagging station. The results indicated that the estimation of tagger effects using the Berger et al. (2014) methods, were not systematically unbiased, despite the imbalanced nature of the tagging dataset.

Tagger effects were modelled using the Berger et al. approach and the results presented for the RTTP and the PTTP (noting that tagger effects corrections can only be modelled for these tagging programmes due to lack of operational data from the other programmes). Both programmes had median release event corrections of around 0.8, ranging from around 0.4 to 1, however for the RTTP there were some events that had correction factors >1, and it was suggested that these few outliers are not applied and should receive a correction factor of 1.

Uncertainty in correction factors was also estimated using random draws of parameters sets from a multivariate normal distribution defined by parameter means and variance covariance matrix to account for correlation between parameters.

Random effects were also explored but made no difference to fixed effects structure.

The presentation concluded with some discussion points, including:

How to correct the Central Pacific PTTP given the recommendation to separate out these releases, but very few skipjack released (814 released with 36 recoveries). Either apply the median correction factors from the western Pacific cruises as is done for the earlier SSAP and the JPTP or model the central Pacific cruises separately. Finally, the issue of imprecision in estimation of correction factors was noted and how sensitivity of the assessment to this imprecision could be considered.

Discussion

The simulation study was positively received, with a suggestion to consider further work where the operating model is designed to replicate the imbalances in the tagging and the estimation models then apply the Berger et al. model, rather than estimation and operating models being specified the same. It was also recommended to do some sensitivity models to explore the implications of imprecision in the correction factors for the MFCL estimation of management quantities. Initially run a few scenarios, median, high and low scenarios, noting that the parameter sets drawn influenced the correction factors similarly across release events. It was also suggested that structural uncertainty in the tagger effects model is a more genuine uncertainty to explore. There was some discussion on this in respect to inclusion of tagger experience, for example, which has been tried and did not produce sensible results, tagger identity seems to have a stronger effect. *Due to time constraints, inclusion of structural uncertainty in the tagger effects model would not be possible for this assessment, but the assessment will undertake to explore the implications of imprecision in the tagger effects corrections.*

Simulating tag mixing of WCPO skipjack tuna

Joe Scutt Phillips provided a presentation on a new Lagrangian simulation approach to estimate tag mixing periods. The study was developed to provide a more informed approach for specifying tag mixing periods to the MFCL model, noting that tag mixing period is an influential uncertainty in the previous skipjack assessments. The approach aimed to simulate mixing periods specifically for each release event taking into account the unique locational and temporal (environmental, fishing effort) contexts of each release event that may result in different rates of mixing of released fish. It applies an individual based Lagrangian model (Ikamoana) (Scutt Phillips et al. 2018) to track movement of individual fish (particles) and quantify the fishing pressure that individuals experience across their dispersal trajectories. Ikamoana uses the forcings and parameters such as fishing mortality, growth, natural mortality that have been estimated from real data by the Eulerian model SEAPODYM (Lehodey et al. 2008). The individual based modelling approach simulates post-tagging movement and probability of capture for all individual tagged fish in a release group while also simulating the corresponding, 'idealised' tagging and mortality of the broader population as the simulated 'truth' to compare to the tagged groups. The key results from the individual simulations are trajectories of survival and capture probabilities that can be compared between tagged fish and the fully mixed population. By comparing distributions of capture probabilities for the tagged and untagged fish it possible to use criteria to estimate at what period after release, fish from particular tagging events are experiencing the same fishing mortality as the untagged fish for particular model regions, and therefore can be considered to be mixed with the untagged population.

The presentation provided and overview of the methods and preliminary results for several tag release events. Feedback was requested on the statistical criteria for indicating that tag release groups were mixed, utility of the approach and mixing periods to explore (i.e., 0-3 quarters or more).

Discussion

Request for clarification on how the movements of the simulated fish are driven, and how to account for the tendencies of some fish to stay in the one place or show different behaviours? *The movement is driven by the Lagrangian formulations of the advection/diffusion equations and environmental forcings from SEAPODYM. It was suggested that in theory this should capture locational 'stickiness' of tagged fish because the tagging data informed the SEAPODYM parameterisation, also the geography is influential and should be captured by SEAPODYM.*

It was pointed out that variability in individual fish behaviours might not be appropriately captured by the SEAPODYMN driven movement, and more work is needed to include this in the modelling. Including different behaviours into the individual fish is a work in progress. A version of this model is being developed that can be used to explore implications of varying individual fish behaviours. More studies will be important to improve understanding of what influences individual behaviours, school cohesion etc., to inform modelling experiments (i.e., paired release/recaptures, electronic tagging, stable isotopes). Further exploration of SEAPODYM results could be useful to see if the influence on geography is captured. Acknowledged that it is early days with this work to explore its usefulness to inform tag mixing, important step forward to identifying differences between release groups.

It was noted that smaller scale influences are likely critical to this analysis, so can you run sensitivities to explore some of the uncertainties in movement modelling. *This is a good idea and could be explored in future work. It was also discussed that the simulation estimates of mixing periods can be applied in the*

MFCL modelling with plus and minus one quarter to include uncertainty, while at least providing information on variation among the release events. This is an improvement on the unrealistic assumption of the same mixing period for all release groups and just having several fixed levels in the assessment uncertainty grid.

Will be interesting to see how the assessment outcomes depend on the assumptions of variable or fixed mixing periods across release groups. Will the other issues in relation to considering the use of tagging data be discussed, especially tag recapture reporting rates. *Chair replied that we can discuss the tag reporting rates issue in a later session, noting that there is no new information to inform updated analysis of reporting rates compared to what was used in the most recent assessments, and no new tag seeding data. Reporting rates will be re-estimated for this assessment by Tom Peatman.*

It was suggested that it would be worthwhile to compare to the mixing rates suggested by the individual based models of mixing to what is indicated by the goodness-of-fit-tests (CUSTARD and TART) using the tag-recapture data described in the study by Kolody and Hoyle (2015). It was suggested to redo the analysis by Kolody and Hoyle with the updated data and model regional structure and compare the results to the previous regional structure and those from simulation work. It was acknowledged that this is a reasonable idea and would be nice to do, some uncertainty on the data requirements. Requires paired analysis and tag density analysis, but should be applicable. Question as to whether that analysis provides results similar to what the individual based simulations provide? The Kolody and Hoyle approach is more about indicating whether individual release groups are mixed, or not, under the mixing period assumptions, so could be used to indicate which groups are mixed under the predictions of the individual based simulations, and which groups do not appear to be mixed – could be uses to to support decisions on whether or not to exclude certain tag release groups. *Not sure how feasible this analysis is for the current assessment, but worthwhile to consider.*

It was noted that clumping occurs due to fish keying in on certain features – anticyclonic eddies are important for tuna clumping, so there are more than bathymetric influences in mixing.

It was also suggested to look at the tag recapture distributions by quarter/region across years to look at area/quarter consistency of recaptures across time and compare to predictions of the simulation modelling. *It was noted that this may not be possible due the lack of continuity of tagging effort in space and time, also the central Pacific tagging cruises do not tag many skipjack. But might be possible to look at this with the dispersal models. Could compare the observed tag-recaptures with those predicted by the simulation models. Also followed up that the simulation modelling can be useful to inform planning of tagging work.*

Finally, the point was made that SEAPODYM integrates behavioural ecology with environmental variables to create the habitat fields to drive the movement, and the tag mixing simulation just turns the Eulerian approach to a Lagrangian approach. Many of the things we have been talking about, SEAPODYM is already providing a fit to and with diagnostics to show how well the model predicts the catch data, noting the fishery dependent nature of this data. So we need to careful about circularity, when thinking about the *ways to interrogate the results of the tag mixing simulations in relation to how they predict the movement behaviour.*

CPUE analyses

The session on CPUE analysis started with a presentation on the Japanese pole and fishery by Yoshinori Aoki from the Japan Fisheries Research and Education Agency (FRA). The presentation provided background on the Japan pole and line fishery operations, noting that there are typically 10-30 poles active in a fishing period and that fish can be caught at a rate of one every 5 secs. It's a surface fishery targeting mostly skipjack and albacore. Temporal/spatial shifts in catch distributions were described showing the contraction in the catch and effort distribution in relation to the skipjack assessment model regions. Catches have contracted to be more focussed closer to Japan overtime. Notably, region 8 of the model and the western halves of regions 2 and 4 have produced very low catches in recent decades. Regions 5 and 6 have shown very low catches by Japanese pole and line over most of the assessment history. Up until the late 1980s the pole and line catch was taken across most model regions, in the last 5 years the catch is mostly taken from region 2 and 4 (8 region model). Effort in the Japanese pole and line fishery is now at less than 15% of its peak in the late 1970s. Big drops in effort during the 1980s. Spatiotemporal models for predicting CPUE are used to account for the variation in the spatial coverage of fishing overtime, aim to provide a better index of abundance.

Most of the catch is taken during quarters 2 and 3, which is spring and summer at Japan's latitude. The next most important quarter is quarter 1. For model region 2 (northern most region), catches only occur in quarters 2 and 3, corresponding with the northward movements of skipjack when the warmer waters occur. In quarter 1 catches are mostly in regions 1, 3 and 4. In quarter 4 catches are low across all regions, and most catch is from region 3 and 4. The seasonal changes in skipjack distribution and abundance off Japan due to temperature variation have a major effect of the dynamics of the pole and line fishery.

The preparation of the catch and effort data was outlined, noting that it does not vary to that applied to the 2019 assessment. The logbook data has been updated to include 2018-2020. The main issue to consider is the implications of changes in the spatial coverage of the pole and line fleet's catch and effort overtime.

Clarification was sorted on filtering step 5 that removes vessels that had operated for less than 5 years in the fishery and less than 10 days per year. *This was modified for the last 5 years to only remove vessels with less than 10 days per year.*

Thom Teears from SPC-OFP followed up with an overview of the approach applied to the standardisation of the Japanese pole and line CPUE.

A spatiotemporal delta-lognormal GLMM implemented using the VAST software will be used as per the preferred approach for the 2019 assessment. The initial phase of the analysis was to replicate as close as possible the index times series applied in the 2019 assessment, using the same data. This is not a straightforward process, as new analysts are involved, and the data are not made available by Japan to the SPC assessment team. The SPC analyst, who was also new to SPC, took the scripts from the previous analysis and modified/simplified them as required due to some customized functions not being compatible with updates to R packages etc., and then worked closely with Japan scientists from FRA to provide training on how to run the VAST models using R. Together, they were able to closely replicate the previous standardised pole and line CPUE indices and the comparison was presented. The analysis was then conducted with the updated Japanese pole and line data until the end of 2021. Results were presented and discussed across all model regions, noting that the pole and line CPUE was not applied in regions 5 and 6 for the 2019 assessment due to the very low data coverage in these regions.

The final proposed model was the same as that applied in the previous assessment, and modelled encounter probability, catch, knot (280 knots distributed uniformly), year-quarter, spatial random effect, spatiotemporal random effect, catchability covariates (fishery class (OS - offshore, DW – distant water), vessel size, number of poles), habitat covariate of SST, and a random effect of vessel ID. Various alternative model configurations were compared, that incrementally removed model covariates from the full model, and confirmed the full model was the best performing model. It was noted that a density-based knot distribution was very similar to a uniform distribution. A uniform distribution was used for the final model. Further work will be done to include diagnostics for the CPUE analysis including Dharma residuals, PIT residuals, covariate plots, spatial effects and uncertainty for the indices. These will be provided in the usual assessment inputs paper.

Discussion

Discussion regarding the reduction in data coverage by the Japanese pole and line fishery overtime and how to deal with this in terms of including particular time periods/truncating of the indices. Recent simulation paper Ducharme-Barth et al. (2022) discusses how well the spatio-temporal model can do in predicting these indices depending on the sampling coverage. It was recommended to look at plots of the proportions of individual regions sampled over specific time periods and make some judgements regarding the time periods for which a viable index could be estimated for each region. Maximum gradient seemed a bit large, perhaps try some more steps to improve the fits. It was suggested that a uniform knot placement is recommended. *The uniform knot distribution will be used*. *Options for truncating indices will be considered.*

There was a query as to why individual gear covariates (i.e., bird radar, sonar etc.) were not included, as an earlier GLM analysis (2010) showed these to be most influential. *The individual gear covariates are only available for the DW component of the fishery, to include both OS and DW fisheries in the analysis and maximise the data available (spatial-temporal coverage) it meant not being able to include the individual gear covariates. Analysis of gear covariates in the 2019 VAST analysis indicated they did not have strong effects and a vessel effect was included instead.* Follow-up question as to why not important? *The geospatial VAST and standard GLM approaches are different and the lack of influence of inclusion of individual gear covariates may have been due to the variability due to gear effects being soaked up by the spatio-temporal random effects of the VAST model. This is a potential downside of the geospatial VAST* *models in that they may model the spatio-temporal variability well, but the specific effects due to gear etc. can then be difficult to detect. If the data are available, suggested to repeat the analysis for DW fishery with individual gear covariates to see if it is consistent with updated data.*

It was mentioned that an important consideration of applying the VAST CPUE indices is whether or not to allow them to inform spatial scaling of the biomass in MFCL. This was not done previously, but if it can be done it provide significant information to the model.

Following the PAW discussion, the assessment team has further considered the strong seasonality displayed by the pole and line indices. This strong seasonal component may be somewhat problematic for the MFCL model to fit. Further exploration with the catch condition modelling approach indicated that for the survey fisheries such seasonal dynamics would be accounted for by seasonal population process such as recruitment or growth, as opposed to 'catchability' which would be estimated for the extraction fisheries. Adding a seasonal factor to the VAST models is perhaps worth exploring but will require some thought as to how the seasonality is modelled as the seasonality may be both due to availability and catchability, and we would only want to remove the catchability influence from the abundance indices. This remains a work in progress, and a seasonal model may or may not be feasible in time but will be explored.

Developing a purse seine CPUE index

Thom Teears followed with a presentation on the exploration of a CPUE index from analysis of purse seine data, focussed on the equatorial model regions 6, 7, 8. The 2019 assessment applied a purse index for region 6 only using a GLM approach. For the 2022 assessment, considering the issue of the decline in Japanese pole and line fishery effort in the equatorial region, further analysis of purse seine catch and effort information is being conducted in an attempt to create suitable ongoing abundance indices for region 6, 7, and 8. These investigations are applying the geostatistical VAST approach. The investigations so far were presented considering both the observer and logbook databases, with the initial analyses focused on the post 2010 period so that indices based on both data sources could be compared. Additional analyses were presented for logbook only data going back to 1990. The models discussed included only unassociated (free school) sets only, and both unassociated and associated (FAD) sets. Two different effort metrics were explored, 'set' and 'time between sets'. It was considered that 'set' as an effort metric could lead to hyperstability in the index and less confidence that it would be sensitive to changes in abundance. On the other hand, 'time between sets' might better capture the effort component that is more likely influenced by abundance, i.e., search time required to find suitably sized tuna schools. The introduction of satellite FAD buoys with echosounders however is an issue for using time between sets as the effort metric for the associated component of the fishery due to the increases in information over time to inform and reduce the travel time between sets (i.e., effort creep). Therefore, combined set type models used set as the effort metric for associated sets and time between sets (excluding night-time, as searching is assumed not to occur at night) for unassociated. Models using only unassociated sets applied time between sets. Indices based on unassociated fishing intuitively should provide more representative indices of abundance than those involving associated sets, and should be less subject to effort creep – at least since 2010 when the observer data begins and much of the current technology (bird radars, modern echosounders, better communications etc.) was likely already in use.

The covariates explored for the combined set type models were: set type, vessel length, vessel ID, species cluster, and ENSO (El Nino-Southern Oscillation), and for the unassociated set type models: vessel length, vessel ID, species cluster, and ENSO. Fifty spatial knots were used, data were initially filtered by: removing skunk sets <1 mt, includes vessel active for ≥ 20% of time series, include vessel between 50 and 80 m length, and remove data with no spatial co-ordinates and/or >99th quantile for catch.

Differences between the indices developed from the observer and logbook data where greater for the unassociated sets analysis, although difference were still notable for the model that contained both set types. For the analyses of observer data, the best models for the unassociated sets analysis included; species cluster, vessel length, ENSO and vessel ID, and for the combined set types, these same covariates with addition of set type. For the logbook analysis, the best models included species cluster, vessel length and vessel ID for the unassociated analysis, and for the combined set types; species cluster, vessel length, ENSO, vessel ID and set type. Results were also presented for the logbook models going back to 1990.

Discussion

Overall, the differences between the indices generated by the logbook and observer data were not trivial. This was especially the case for the unassociated sets analysis, even when using the same effort metric of time between sets. It was noted that issues such as species composition are likely important in differences between the logbook and observer data, as the observers use a more systematic approach to estimating the species composition than the skippers. The skippers may tend to make general estimates that are more similar across time and may not capture the variation as well as the observer records. The observer data is a more reliable dataset and it would perhaps be better to focus on using those data, even though a shorter time period is available.

There was discussion on considering the inclusion of more specific gear covariates in the analysis, as was also suggested for the pole and line, rather than assuming all the gear effects are captured by the vessel length and vessel effects. Suggested to go back and see what other gear covariates could be include from the observer data. *It was pointed out that previous extensive analysis of purse seine CPUE used vessel length, net length, net depth, well capacity, gross tonnage, skiff horsepower, species cluster. The best model simply included just the vessel length covariate as most of the other covariates correlated with vessel length or had negligible effects. In any case the covariates in general did not have much effect. We can go back and have another look at what is in the observer records to see in anything else might be useful.*

It was also mentioned that using FAD sets within the CPUE analysis should be treated with caution due to the technological advances and associated difficulties in capturing effort appropriately and factors influencing increased effectiveness of FAD fishing. Further work to explore a purse seine CPUE index based only on unassociated sets is worthwhile. It will be important however, to consider the implications of the FAD closure period, as during this period there will be vessels that are not as skilled in free school fishing conducting more unassociated sets. This may have the effect to depress the index during quarters when

FAD fishing is mostly banned. Consider approaches to filter vessels so that vessels that mostly specialise in FAD fishing are excluded from an index based on unassociated sets. *SPC will further explore the derivation of an abundance index based on unassociated sets only.*

The discussion on CPUE indices then focussed on the issue of effort creep, which may not be accounted for in standardisations, but could be biasing CPUE trends and creating hyperstability in the indices. There is an assumption that effort creep must be notable, particularly, in the purse fishery due to the introduction of technologies such as satellite and sonar equipped FAD buoys, better communications and fishery analytics etc. over the last 20 years. Effort creep was also raised as a concern for the pole and line fishery. However, despite significant work, effort creep has been difficult to clearly detect and quantify in the purse seine fishery with the available data, at least for the more detailed analysis since the late 2000s when more operational data became available from observers and much of the new technology was implemented by the fishery. Effort creep would also not be a consistent effect overtime and would be expected to occur in phases and at different rates depending on when new technologies or other gear improvements were introduced and their trajectory of uptake by individual vessels.

Suggestion of approaches to explore effort creep and providing some evidence to support applying certain levels of effort creep to adjust the CPUE indices overtime were mentioned as worthy of further exploration, in particular the approach by Kolody applied to estimating skipjack catchability trends in the Indian ocean (IUTC-2018-WPTT20-32). It was suggested that given whatever CPUE is used it will be subject to considerable uncertainty, this may need to be included in the uncertainty grid.

Further comments that effort creep issue are relevant to pole and line also focussed on this active fishing method and that search time is expected to be affected by technology uptake etc.. Suggested that a study to explore implications of adding different levels of effort creep to the pole and line indices and seeing how this influence aspects such as increasing recruitment trends overtime, tag reporting rates on upper bounds etc. *Agreed that this would be an interesting exercise, and we will need to consider effort creep scenarios to apply, likely not feasible before this assessment, but will consider adding an effort creep sensitivity. Effort creep studies on the Japanese pole and line fishery would also need to be led by the Japanese as SPC does not have direct access to the relevant data. (Note that concerns related to including effort creep scenarios in the 2022 assessment are discussed further below).*

Keith Bigelow of NOAA provide a short presentation describing the approach to generate the CPUE index for the high seas pocket 1, representing model region 5, using the Philippines domestic purse seine fishery data provided by BFAR. The method applies a GLM with negative binomial, with year:quarter, area, vessel and log effort. The year:quarter and vessel effects are most important. The approach used in 2022 will be the same as in 2019. Data is available from 2005 to 2021.

Model development

The next session discussed model development for the stock assessment. Claudio Castillo Jordan of SPC-OFP provided an overview of the model development so far and potential options for the diagnostic case model and the uncertainty grid.

Noting that a new version of MFCL will be used with new features, most notably the catch-conditioned approach. Other features that will be explored, that were not available/applied to the previous assessment, but are now available in MFCL include; orthogonal polynomial recruitment, a new approach to movement estimation and several self-scaling approaches to estimation of effective samples sizes for size composition data (i.e. self-scaling multinomial, Dirichlet).

A recap on the fishery and regional structure was provided before handing over to Nick Davies for an overview of the catch conditioned approach to estimating fishing morality. The catch conditioned approach is also discussed in the section on MFCL developments update (day 3). In summary this new catch conditioned approach has passed through testing and is ready to apply in a formal assessment. Consistent with the objective to reduce the complexity and parameterisation of the skipjack assessment the catch condition is expected to be applied in the 2022 assessment. It is a simpler approach in terms of the function minimiser rather than a major simplification of the process being modelled. It is hoped that it will improve stability of the model estimations and help achieve the positive definite Hessian solution.

Model development so far has focused on using the 2019 data inputs. Updating to the new MFCL version 2.0.8.4 applied as it was in the previous assessment produced virtually identical results. The next phase of model development was to explore the catch conditioned approach. Implementing this on the 2019 data showed some notable differences that required further investigation, and implementation of this approach is still in its early stages. It was noted the large reduction in estimated parameters when applying the catch conditioned approach, and other new features.

The approach for estimating growth was also discussed, with options being considered including estimation by MFCL with high weight placed on the size composition data and then applying the estimated growth externally, estimate growth externally using other information such a tag-recapture growth increments and literature review.

Discussion of potential uncertainty components to include in the structural uncertainty grid followed. The possible grid components include, Tag Mixing, Length Scalar (or remove thus axis by using a self-scaling option), Growth and M, and Steepness. Other options to consider in relation to structural uncertainty include the regional structure, movement including the possibility to provide movement coefficients derived from the SEAPODYM analysis.

Discussion

It was noted that a new feature in MFCL to be discussed on day 3 is the ability to deal with chronic tag loss. This could be considered if there is thought to be an issue with the skipjack tag retention overtime. Also, as to calculation of the Hessian, it is very early days for considering this and getting a few negative eigenvalues is not worth getting caught up in during the model development phase. As you get closer to a model where you are settling on inputs, assumption etc. you can use tools developed for MFCL to explore the parameters responsible for negative eigenvalues and work on improving those.

Other comments related to growth estimation: estimating growth internally with length frequency data that is noisy and variable over time, whatever comes out is likely to represent the dominant data or some kind of average. Might be better to identify particular fisheries with good data overtime, fit to a number of these, and this may provide of an appreciation of variation in the MFCL estimated growth, that can then be considered as uncertainty. Comment on grid explosion, but don't have to run all combinations, can use a confounded design to reduce the number of models in the grid as demonstrated in an earlier assessment.

Comment made that we are short of time and lots to talk about in relation assumptions an uncertainty, need to revisit this topic again on day 2.

It was noted the value of working to get to a positive definite Hessian solution so that estimation uncertainty can be incorporated. There are other uncertainties that were talked about including; tagger effects, CPUE, factoring in effort creep and if so how much, regional structure is also relevant in terms of the estimability of the model. This discussion is important and warrants more time to discuss. Information on growth from publications can be used to constrain growth estimation if you use an internal estimation. Regional structure approach, noted the SC15 recommendation but issues still exist with movement parameters etc., could pull in SEAPODYM, but still movement is an issue. Combining regions 1-4 seems a viable way forward and may help with some of the seasonal and buffering concerns etc.

The orthogonal polynomial recruitment was questioned, need to think carefully about its inclusion. It was suggested that for effort creep, just because you don't know what it is, doesn't mean you don't include it in the grid. The one quarter mixing period could potentially be removed based on outcomes of goodness of fit tests if you can run these. For length frequency data weighting these should be reconsidered in light of the work by Tom Peatman to reweight the composition data. Consider also adjusting down the maximum effective sample size. *Clarification was made that the devisor is applied after the maximum sample size.* Combining the northern regions could be problematic for use of the Japanese tagging data. Alternative model region structure is an uncertainty and therefore could be considered for inclusion in the grid.

The chair agreed to continue this discussion in an additional session on day 2.

Continuation of the discussion on model uncertainty on day 2:

The discussion of uncertainty characterisation was continued at the end of day two with some initial recap, and the preface that while there are many uncertainties that could be considered for inclusion in the characterisation of uncertainty, the uncertainty grid presented to SC18 should avoid having arbitrary options for the levels of each uncertainty axis. While it might be acknowledged that a particular uncertainty is important to consider, if there is a lack of reliable information to inform the ranges of uncertainty to include for a particular factor it is difficult for SC to consider and can lead to even more arbitrary weighting of that axis. In this context, it's important to consider what is included in the uncertainty grid used for management advice and what is acknowledges as a likely influential uncertainty but is provided as a sensitivity analysis with less informed ranges, pending more reliable information to bound the uncertainty.

Options for the uncertainty grid we then discussed.

The key proposed axis were: **tag mixing period, growth, and steepness**. It is proposed to estimate natural mortality internally, similar to the previous assessment so this would not be a grid axis. Further, it is hoped that the axis for the length composition sample size divisor can be removed by applying the Dirichlet multinomial with no random effects, which has recently shown good performance in other models. This will reduce the grid to the key uncertainties that were influential in the previous assessment and removes the arbitrary decisions on weighting divisors for the length composition data.

The proposed options for the tag mixing axis were: 2 quarters, 1 quarter (but questions over whether 1 quarters should remain), variable mixing be release group based on the tag mixing simulation study.

The proposed options for the growth axis were: estimated by MFCL, external growth options from literature and or tag-recapture growth analysis

The proposed options for the steepness axis are: 0,65, 0.8, 0.95

Other uncertainties that could be considered as sensitivities included: estimation uncertainty for tagger effects (i.e., apply median and plus or minus some error range), effort creep scenarios on CPUE indices (i.e., 0, 1, or 2%), regional structure (collapsing regions 1-4), CPUE (pole and line truncation, purse seine options), using movement from SEAPODYM, revision of reporting rate priors. If the estimated management quantities are notably sensitive to these uncertainties they could be included in the model grid for management advice.

Follow-up discussion on treatment of uncertainty

Effort creep scenarios – this needs more work, but evidence not necessarily available to include in the grid. Very useful to include as a sensitivity analysis and include in the assessment report somehow to raise the issue and put it to the industry to provide their perspectives.

Follow-up on the complex temporal dynamics of effort creep, it's not a continuous compounding process. But needs information to inform how to structure the scenarios to be modelled. Suggestion of industry survey of Japanese pole and line fishery could very useful, we also do not have access to operational data to do any in depth analysis, surveys and analysis would need to be run by the Japanese scientists.

Comment on separating out effort creep scenarios for pole and line and purse seine, perhaps consider 1-4% range for purse seine, perhaps lower for pole and line. But over what periods, compounding rates lead to large effects that are unrealistic over long time periods.

Regarding effort in the Japanese pole and line, Japanese scientist from FRA suggest they are looking into this but for now do not have a quantitative measure of effort creep. In future they will look to getting perspectives of the fishery. Effort creep may have occurred overtime related to bird radar, sonar, SST from satellite, which could have improved the search times. Japan suggested that effort creep should be taken into account for this assessment for all index fisheries. Further, discussion will be required on the levels of effort creep but propose it should be captured in the assessment uncertainty.

Followed up with the question as to whether to include effort creep in the grid – Japan confirmed they would like this. *If so, Japanese scientists would need to conduct an analysis or work with SPC to do an analysis to provide some defendable values for the pole and line fishery. There are conflicting views on this, one is to include sensitivity analysis or effort creep, the other is to try to include it in the uncertainty grid for management advice.*

Inclusion of effort creep in the uncertainty grid was supported by another attendee but would need to be supported by further analysis to determine the effort creep scenarios to include, perhaps based in the work done in the Indian Ocean (Kolody method), maybe also by comparing pole and line versus purse seine. *In response, it was suggested the ratio methods suggested might not be that applicable to the pole and line fishery which is 90% skipjack. If formal inclusion of effort creep in the grid was to occur, it gives it a level of authority that is not yet supported by the data available, in particular for pole and line.*

Clarification that previous assessment seemed to indicate different rates of change in fishing power between pole and line purse seine, also acknowledge using ratios for rates for different species to explore effort creep in the pole and line fishery may not be feasible. *Previous assessments have estimated catchability changes for purse seine against pole and line which is assumed to have constant catchability. The catchability trends for purse seine in the previous assessment could provide some basis for informing effort creep levels to explore for purse seine.*

At this stage it will be difficult to come up with values of effort creep, support was expressed that a sensitivity or small suite of sensitivities informed by the best available information be included as an appendix to the stock assessment report.

The issue of effort creep uncertainty will not go away, as a collective we need to think about what to do about this problem and express this need to the SC. Either start collecting the data needed to measure this or move to fishery independent indices. Either way a strong push to SC to support work to deal with the effort creep issue is warranted.

The final proposal is to include effort creep as a sensitivity analyses, with scenarios for both pole and line and purse seine.

Query on indices being used in the assessment in relation to the mix of the pole and line and possible use of a purse seine index for equatorial regions. *The likely approach will use the pole and line index for regions 1, 2, 3, 4, 7 and 8, the Philippines purse seine index for region 5, and a newly generated purse seine index using the VAST approach for regions 6, 7 and 8. In relation to the purse seine index for regions 6, 7, and 8, we are uncertain over what period this would apply at this stage (mostly likely the purse seine index would be based on observer data, so starting from about 2010), the thought is that as the pole and line effort drops away over time in regions 6, 7 and 8 that the pole and line index would be replaced by a purse seine index for these regions, so the historic period will continue to be informed by the pole and line CPUE. Still in early stages on the purse seine CPUE analysis. Also, to note that the previous assessment applied a GLM purse seine index for regions 6 (PNG area) based on logbook data from 1997, so switching to the use of observer data will reduce the length of the time series and replace the GLM region 6 purse seine index used in the previous assessment.*

The simplification of the skipjack model was commended and suggested that simplification of the uncertainty grid could involving removing the 1 quarter mixing options based on evidence discussed by others and adding the variable mixing estimates derived from the simulation studies. As to the regional structure issue, recommend that this be included in the report with results from a 5 region model that combines region 1-4 included.

The suggestion that removing a mixing period of 1 was not justified, and not supported by others who argued we don't have a broad justification for this yet. Need to further consider the results of the simulation work before considering this. This was followed up with support to keep the 1 quarter in the grid, but important to include the heterogeneity that would be expected based on the simulation results.

Possible that the simulated mixing periods could be added as a sensitivity rather than as a grid axis, pending SC consideration of the mixing simulation study.

The previous approach of Kolody was raised again as a way to see whether any groups are likely mixed by 1 quarter after release. Suggested that this approach should be revisited and applied to the larger data set now available. This could also be useful for providing confidence for the SC to endorse the simulation modelling approach for estimating tag mixing periods to input to the MFCL model.

Query on movement priors: reflecting that these were set to zero in the previous assessment, and for region 4 of the 8 region model the estimation was at the prior. Would be interesting if the 8 region model is retained that changes to this prior are explored to see if the region 4 movement estimation simply follows the prior and then if so, what are the implications of this for the management quantities. *Trying a different set of priors would be a good thing to do, and we could look at getting the transfer coefficients from the SEAPODYM model and try using these as the priors rather than specifying rather arbitrary low values.* Further support for this idea and that estimating movement in the model is difficult given the narrow spread of tagging relative to the area covered by the assessment.

The comment was made to consider implications of the COVID pandemic on the tag reporting rates, and that priors on this might need to somehow be freed up in the final years of the model period. It was also noted to be careful with assumptions related to relationships between growth and steepness when considering plausibility of particular grid combinations (i.e., high steepness is inconsistent with low growth), recent studies suggests that these assumptions are questionable.

Tag reporting rates were discussed in relation to the issue of model convergence and achieving a positive definite Hessian solution. As some reporting rates were on the lower and upper bounds, it is worth looking into this issue further. For lower bounds, some may have no recaptures, in this case just turn the reporting rate estimation off, for those on upper bounds, increase the upper bound and see if the estimates follow the bound or settle on a value below the increased upper bound. *Agreed these are good ideas to explore.*

A comment was made regarding the implications of COVID in relation to this assessment. The fishery has continued to operate well, but observer coverage has dropped since early mid-2020. This may have implications for the precision of species composition estimates, as for the purse seine the observer records are important for adjusting catch estimates for species composition. Other aspects such as trip length may have changed due to different port/unloading aspect arrangements, but uncertain on any implications of this. Finally, the amount of observed fishing activity will be lower meaning reduced coverage for CPUE analysis using observer data.

Summary of proposed uncertainty grid

Uncertainty	**Option 1**	**Option 2**	**Option 3**
Tag mixing	1 quarter	2 quarter	Variable (simulation)
Growth*	**MFCL (Dirichlet)**	Tagging and/or literature based	
Steepness of stock recruitment relationship	0.65	**0.8**	0.95

**Modelling since PAW shows the Dirichlet multinomial distribution has performed well and can be used in place of the rather arbitrary sample size divisors used in the previous assessment to provide an uncertainty axis for alternative weighting of for the size composition data. The tests show the growth estimated by applying the Dirichlet multinomial is virtually the same used in the previous assessment for the diagnostic model, but with a more realistic range of uncertainty that increases with age. The alternative options for growth will depend on the outcomes of the analysis of tag-recapture growth increments and the literature review. The findings from this work will be compared to the model estimated growth to inform if, and what, alternative growth functions will be applied within the uncertainty grid, but we expect at least one other growth option besides the MFCL estimate.*

We are currently unsure on which mixing period option will be used as the diagnostic case - pending consideration of the results of the simulation work and an exploration of the diagnostics of models run using these alternative mixing approaches.

Steepness will be the standard options used for the WCPFC tuna assessments.

Summary of proposed sensitivities

Several sensitivity analyses are proposed to be applied to the final diagnostic model:

- Apply a 5 region model structure with region 1-4 combined into one region
- Apply effort creep scenarios to the CPUE indices. More work is required with Japanese scientists to inform effort creep scenarios that are consistent with their data and other information on uptake of gear and technological advances. As this will not be possible in time available, we would apply a 1% annual rate of effort creep on the pole and line CPUE indices. As to the effort creep to apply on purse seine indices, the observer data is only available since 2010. We suggest that effort creep has been far less over this period than in the 1990s and 2000s when technologies were being taken up by the purse seine fishery. We propose to apply a modest scenario of 1% to the equatorial purse seine indices since 2010.

- Apply the transfer coefficients from the SEAPODYM as the priors or as fixed coefficients for the movement parameters.

DAY 2 – 2022 Southwest Pacific mako shark assessment and follow-up work on 2021 southwest Pacific blue shark assessment

Southwest Pacific mako shark assessment

Data overview

Stephen Brouwer of Saggitus provided an overview of data available for mako shark in the southwest Pacific focussed on the longline fishery. Purse seine was not discussed but will be presented in the data characterisation paper for SC18.

The presentation summarised the data available for the logbook and observer records, along with information on biological parameters, drawing from summaries available in the most recent Shark Research Plan (Brouwer and Hamer 2020), and general data on longline fishery. Records of mako shark have increased over time due to improved reporting but have dropped in recent years which is likely related to increased discarding and reduced observer coverage. A significant amount of length data is available for the longline fishery from about 2000, with over 10,000 length samples, and mostly recorded by sex. Very few data are available for the purse seine fishery. Overall, the recent Shark Research Plan indicated that a data rich integrated assessment could be considered, with the caveat that much depends on the ability to achieve a suitable catch reconstruction. Some age and growth data are available from the south Pacific (New Zealand) and the east Pacific (Mexico).

A key issue is that while there is a reasonable amount of data on blue shark catches for logbooks, they are all reported as the western Pacific (WX), with no distinction between the north and south Pacific. Reporting to a more specific catch location should be improved and/or adjusted retrospectively if possible. There is a lot of catch data provided by flag, but again they are not separated by north and south Pacific. Shark catch reporting, in general, was low up until the mid-2000s when CMM 2006-05 was introduced that included reporting requirements for sharks, but much of the shark catches were not reported to species. CMM 2009-04 was introduced in 2009 that required species reporting, from 2010 increases in mako reporting was detected. Issue was noted on porbeagle reporting around Fiji which seems too far north of their typical range, suggested to assume porbeagles reported north of 25°S be reclassified as short fin mako.

Data on fate and condition on capture and release is available from 2000, which is important for incorporating discard mortality rates into catch estimates. The proportion of mako shark released as “dead” has decreased over time, and from 2015 most makos are reported as released alive.

More detailed information on length composition was discussed, including the amount of length data by flag over time, noting that the length data is truncated around 300 cm, and mako can grow larger than this, so large older individuals are not captured. Mean length has been quite stable since the late 1990s. The best data coverage over time is for Australia, New Zealand and Fiji, followed by Chinese Taipei. Patterns of length composition by latitude and quarter were also presented, suggesting that small sharks appear more common at southerly latitudes and move north as they grow. Catch by hook number (per basket) was also shown indicating that mako shark are mostly caught on hooks closer to the surface. Catch proportions were presented showing the high proportions of mako and other shark catches in shallower hooks. Observer reported catch proportions by shortfin and longfin mako and porbeagle were discussed to describe the issues with patterns of species misclassification. The key point was the dramatic reduction in the proportion of porbeagle shark reported from 2000 onwards when it appears species classification dramatically improved.

Summaries of observer coverage in terms of hooks set and bait use were provided, followed by summaries of nominal CPUE variation by year and fleet along with numbers of observed sets with CPUE, noting that flags with most data are New Zealand, Australian and Fiji. Also noted that for some flags, i.e., Australia, the introduction of electronic monitoring (EM) has likely influenced recent trends in logbook reported CPUE (getting better) and CPUE data from observer is not available for the recent years due to other process required to analyses and deal with EM data. Also note to be wary of variation in CPUE due to changes in reporting requirements over time.

Overall concluding points were:

- Mako sharks are wide ranging across the south Pacific Ocean, with size and seasonal movement patterns.
- There are reasonable amounts of data from 1990-2020, but the data by fleet are fragmented.
- Mako shark aggregated data should be reported by ocean area not simply WCPO, and where possible these data should be retrospective corrected.
- Catch reporting has improved across all fleets over time.
- For most fleets after 2015 most mako sharks are released (cut free), and a high proportion of releases are alive and healthy at release.
- Some length data are available but not for all fleets.
- Mako sharks are landed in both shallow and deep sets, but most frequently caught in the shallow hooks and comprise a higher proportion of the catch in shallow sets. But the number of deep sets is much larger and therefore contains the most data.
- Both observed and reported data are available for CPUE standardisation. Gear attributes (e.g. HBF) more likely to be informative than specified targeting information as targeting is poorly reported.
- Past management interventions may complicate the CPUE standardisations as in recent years most flags require sharks to be released.
- There appears to be a species identification issue in the logsheets with some mako sharks probably recorded as porbeagle. We may need some rules to deal with this e.g., all POR recorded north of 25°S re-classified as SMA.

Biology and assessment planning

The mako shark data characterisation was followed by an overview of plans for the assessment presented by Kath Large of Dragonfly. The presentation provided background on shortfin mako shark biology and stock structure and an overview of the planning for the assessment.

Life history and biological parameters for shortfin mako shark vary slightly depending on the studies. Shortfin mako shark are moderately long live; it is thought they can reach at least 30 years of age. Age at maturity is around 7-9 years (180-185 cm TL) for males and 20 years (275-285 cm TL) for females. Fecundity is low, with an average of 12.5 pups per litter (range 4–25), but with a reproductive cycle that is thought to be two or three years depending on the study region. Natural mortality is suggested to be low (i.e., 0.128-0.15). Growth is typically fast during the first few years and slows down from about 5 years age. Some studies suggest growth rates for males are lower compared to females from around the age at maturity, although more recent studies cast some doubt on this observation and suggest that growth differences might not appear until beyond 16 years age. The issue of growth uncertainty between sexes is important for deciding if the assessment should attempt to be sex specific or not. Advice was requested on if or not to incorporate biological parameters used in the north Pacific mako shark assessment.

Southwest Pacific shortfin mako shark are genetically distinct from those in the southeast Pacific and the north Pacific. This genetic separation is further supported by tagging studies that show regular movements around the southwest Pacific but only one known movement of a southwest Pacific tagged shark to the north Pacific, and none to the southeast Pacific. The locations of large females are poorly known. There is reasonable confidence that the southwest Pacific is an appropriate scale for the stock assessment.

The planned approach for the assessment will be similar to that applied to the 2021 assessment of southwest Pacific blue shark (Neubauer et al. 2021), which will be complimented by a comparable dynamic Bayesian surplus production model. By comparable this means a model that can use the same inputs as the Stock Synthesis model. The surplus production model could also provide a back-up should the Stock Synthesis model prove problematic for this first assessment. Discard mortality will be included, in combination with the uncertainty in catches based on catch reconstructions.

Likely uncertainties to be included in the assessment (with the caveat that they may change depending on the catch reconstruction work) include: Initial F/catch (combinations of discard mortality and post-release mortality), CPUE series, growth and fecundity, stock recruit assumption, natural mortality, recruitment deviations, prior on initial population size and productivity (SP-models). The approach used for characterising uncertainty/selecting models/model weighting for management advice will be guided by the outcomes of the concurrent work on the 2021 southwest Pacific blue shark assessment grid. Outputs will be similar to the blue shark assessment with addition of dynamic depletion, $B/B_{F=0}$.

Reference related to biology noted in the makos shark presentations:

Bishop, S.D.H.; Francis, M.P.; Duffy, C.; Montgomery, J.C. (2006). Age, growth, maturity, longevity and natural mortality of the shortfin mako (*Isurus oxyrinchus*) in New Zealand waters. Marine and Freshwater Research 57: 143–154.

Carreón-Zapiain, M. T., Favela-Lara, S., González-Pérez, J. O., Tavares, R., Leija-Tristán, A., Mercado-Hernández, R., Compeán-Jiménez, G. A. (2018). Size, Age, and Spatial–Temporal Distribution of Shortfin Mako in the Mexican Pacific Ocean. Marine and Coastal Fisheries: Dynamics, Management, and Ecosystem Science 10:402–410.

Clarke, S.; Coelho, R.; Francis, M.; Kai, M.; Kohin, S.; Liu, K.-M.; Simpfendorfer, C.; Tovar-Avila, J.; Rigby, C.; Smart, J. (2015). Report of the Pacific Shark Life History Expert Panel Workshop, 28-30 April 2015. Western Central Pacific Fisheries

Commission Scientific Committee eleventh regular session WCPFC-SC11-2015/EB-IP-13. 111 p.

Francis, M.P. (2016). Size, maturity and age composition of mako sharks observed in New Zealand tuna longline fisheries. New Zealand Fisheries Assessment Report 2016/22. 34 p.

Holdsworth, J.; Saul, P. (2014). New Zealand billfish and gamefish tagging, 2012–13. New Zealand Fisheries Assessment Report 2014/11. 26 p.

ISC Shark Working Group (2018). Stock assessment of shortfin mako shark in the North Pacific Ocean through 2016. WCPFC-SC14-2018/ SA-WP-1. WCPFC.

Kai, M.; Thorson, J.T.; Piner, K. R.; maunder, M.N. (2017). Spatiotemporal variation in size-structured populations using fishery data: an application to shortfin mako (*Isurus oxyrinchus*) in the Pacific Ocean. Can. J. Fish. Aquat. Sci. 74: 1765–1780.

Mollet, H.F.; Cliff, G.; Pratt, H.L.; Stevens, J.D. (2000). Reproductive biology of the female shortfin mako, *Isurus oxyrinchus Rafinesque*, 1810, with comments on the embryonic development of lamnoids. Fishery Bulletin 98: 299–318.

Neubauer, P.; Large, K.; Brouwer, S. (2021). Stock assessment for Southwest Pacific blue shark in the Western and Central Pacific Ocean. WCPFC-SC17-2021/SA-WP-03). WCPFC.

Neubauer, P.; Richards, Y., & Tremblay-Boyer, L. (2019). Alternative assessment methods for oceanic white-tip shark (tech. rep No. WCPFC-SC15-2019/SA-IP-13). WCPFC.

Takahashi N.; Kai M.; Semba Y. (2016) A potential method for meta-analysis of growth curve for shortfin mako shark in the North Pacific. ISC/16/SHARKWG-1/08. Working document submitted to the ISC Shark Working Group Workshop, 14–21 November 2016, Busan, Korea.

Discussion

Question was raised regarding the plan for CPUE, as whatever is decided will have implications for both proposed assessment approaches. *Acknowledged* t*here are a number of challenges posed by the earlier presentations for CPUE analysis. These will be followed up in next presentation*.

Release mortality is currently being studied in NZ and a draft paper can be provided (refer to Simon Hoyle, NIWA). How to deal with natural morality, are you considering age dependency in this, and or different base levels? Noting age at reproduction is around 18-20 for females and the oldest observed mako have been 29 years (32 years in Atlantic), and with a 2-3 year reproductive cycle this provides a short window for reproduction, so it may be likely the maximum age is well underestimated. How might this be dealt with. *The uncertainty in longevity of the species was acknowledged, but unclear how to resolve or deal with this in the assessment at this stage.*

Follow up comment on age and longevity – big mako sharks are very hard to find and large females don't turn up in fishery catches or data, very little is known about them. Furthermore, ageing methods are not validated, bands maybe be 1 or 2 years or even vary in periodicity.

The issue of sex specific assessment was raised again, with the recommendation that if there is enough data the assessment should be sex specific. However, the issue of determining catch by species, which could be problematic for a sex specific assessment, was raised, noting that it is common for spatial distributions of males and female sharks to differ so the sex ratio might be expected to differ spatio-temporally. Question was asked on how much information there is on spatial distribution of sex ratios. *At this stage we have not explored the spatial patterns in sex specific information, starting to look at the length compositions, so will consider sex ratio also. We have the sex ratios from observer data that are used to do the catch reconstruction so it could be relatively straight forward to proportion the reconstructed catches by sex.*

Length frequency data and catch reconstruction

Kath Large provided a presentation on the length frequency data from observer records and the catch reconstruction approach. She noted that the analysis of CPUE for the purposes of abundance indices had not begun and would not be discussed.

The summary of the length composition data noted that observer sampling coverage is biased towards different flags overtime; Japan and Australia in the mid-1990s, New Zealand and Fiji in the later 1990s, Australia in the latter 2000s, and for 2010 to 2021; Fiji, Chinese Taipei and China. Unscaled length frequencies were presented by year, north and south of 35°S, sex and fleets. Spatial coverage appears reasonable, but number of annual samples is low. The pattern of smaller individuals further south was not as clear as with blue shark. Mako sharks are endothermic and aren't restricted by environments like some other sharks. Larger mako likely do occur in colder waters, smaller individuals reported further south, could be including porbeagle.

The catch reconstruction approach will follow the method applied to 2021 southwest Pacific blue shark assessment (Neubauer et al. 2021), where overall catches are estimated from observer catch rates using GLMM, estimated within the general Bayesian framework "brms" using negative-binomial models. Models of CPUE are built based on observed sets and covariates, and then used to predict catches based on total reported longline effort (hooks) by fleets across the assessment region. The response variable is formulated as CPUE aggregated to the resolution of the L-Best dataset (i.e., 5◦ x 5◦, flag, month).

A preliminary example of a catch reconstruction model was presented and the issue of variable observer coverage overtime depending on fleets was important to consider in model fitting. The influence of species ID codes (SMA and SMA+MAK) was also discussed, and it was decided that the analysis would use the dataset for SMA+MAK. The preliminary catch reconstruction models predicted captures (posterior mean) at similar magnitude and trend to the previous study by Peatman et al. (2018), with a stronger decline in mid-2000s, and continuing decline from 2010 to 2020, but with large uncertainty associated with the entire series. It was suggested that the recent decline could be influenced by the increased rates of cut-offs.

Fleet definitions were discussed, with a proposed fleet definition stratified by two area: NZ/Tasman, High-latitudes and High-seas, Low-latitudes, to be further explored.

The immediate focus of work will be on the catch reconstruction aspects, with consideration on the following issues:

- Explore species identification/mis-identification (SMA, LMA, MAK, POR)
- Investigate proportions cut free
- Consider to issue of extrapolating into spaces where there are no SMA, or not at a plausible catch rates?
- Implications of variable observer coverage over time and space
- Are we predicting CPUE and catch with bias? (i.e., related to observer coverage biases)
- Are we adequately capturing uncertainty of predicting into spaces with low information
- How do these affect the catch and CPUE trajectories, and what effect will this have on assessment outputs?

Discussion

Initial questions were posed by the chair in relation to choice of datasets (MAK v SMA+MAK) for catch reconstruction, specifically in relation to the reporting changes, and the length of time series required to do an assessment on a moderately long lived, late maturing species. The point was also made that this is first assessment and we should have realistic expectations given the data.

Comments made on 1990s Australian observer data that it maybe from the Japanese fleet, and to follow-up with SPC perhaps. With using observer data to reconstruct catches we assume that the catch rates recorded by observers are representative, but with low/patchy coverage they could be biased. How are you taking this into account? *This issue was acknowledged and some of the covariates should help deal with this, but more exploration of the observed data and how best to deal with this issue of spatial-temporal biases and low samples sizes is required. The next phase of work will spend a considerable amount of time on this issue.*

It was suggested to consider comparing observer catch rates with logbook catch rates as you would expect logbook catch rates to be less than observer catch rates if the observer catch rates are a good representation of the catch rates in an area. If this is not the case, it may indicate biases due to low samples sizes.

It was noted that the assessment team will collaborate with Japanese scientists (Kai-san) on the analysis of CPUE from the Japanese longline fleet.

Comments on biological parameters between north and south Pacific shortfin mako, specifically the growth parameters which have been problematic for assessments of north Pacific mako, where a meta-analysis was used to construct sex specific growth parameters. The growth parameters link to other parameters such as maturity at age, natural mortality and steepness. Suggest to start with north Pacific mako shark biological parameters and one by one go through and replace them with more appropriate values for south Pacific shortfin mako if available.

Question was raised as to the quite notable difference in the catch reconstruction using the SMA versus the SMA+MAK datasets when the CPUE variation was similar for both and the it would good to compare the reconstructed catches between fleets for the SMA versus the SMA+MAK datasets to see which fleets data are influencing the differences.

Point was made that the way the plots were presented makes the differences look greater than they are.

In relation to how long of a time series of data is reasonable to do the assessment the point was made that for these assessments we are fixing productivity parameters and exploring alternatives for these. Not trying to estimate the productivity in these assessments due to the short-time frame of the data compared to the longevity of the species. However, simpler surplus production models will be attempted that will try to estimate productivity parameters, albeit with informed priors. Hopeful that using these simpler model might help to inform some of the fixed parameter settings required for the more complex integrated assessment approach, that due to data limitations require a lot more fixed parameter assumptions.

In relation to the patchy observer coverage data, this is something to spend some time on, particularly in relation to model selection for catch reconstruction, which has perhaps been overly simplistic. Might want to look at more strategic cross validation across spatial, temporal and fleet strata to better identify models with good prediction on the fleets that take most catches as opposed to those that might have a lot of observer data but don't catch much. This is an important part of this work and more so for mako given the patchiness of the data.

Follow-up discussions also pointed to the importance of dedicated work on catch reconstruction outside of an assessment. More time is required to explore catch reconstruction methodology across a range of shark and other species with patchy catch or observer data history. A dedicated programme of work on catch reconstruction, that might include simulation studies, could be proposed to SC for funding support.

Question raised around the data used to generate catch for unsampled areas by observers. *L-best provides the effort from the logbooks and the observer data provides the CPUE, but further consideration required around how to treat areas with no CPUE but that have effort.*

Question on the sex specificity of the assessment – are there certain areas where the males and females are more vulnerable. *There are data that might help explore this, i.e. the sex specific length data, but we will need to look at that more closely.*

There appears to be some evidence in the spatial size composition figures of larger sharks being taken by distant water fleets to the east in the high seas. Larger mako sharks are typically also not found close to land. Could be evidence of some sex segregation, but pattern might also be influence by low sample sizes in that region.

In relation to catch reconstruction – noting that more mako sharks are caught in shallow sets – can this be factored into the analysis. Not responded to.

Question raised on how you can predict catches where you have no observer CPUE for a fleet and given the fishing practices can vary greatly among fleets. *We can only really predict the catches based on the information available with the effort data. Catch composition cluster helps to account for the different fishing/targeting practices. Flags are included where observer data is available, where there is no observer data, the flag effects are integrated across flags to estimate the unsampled strata, but this is generally more uncertain. Trying to be more thorough and doing more cross validation to explore how well we are extrapolating across poorly sampled fleet strata.*

Suggestion to perhaps explore the important variables using the fleets with good observer coverage and see if any of these variables are ones that cannot be included in the more expansive model with the logsheet data. *Acknowledged that this is a good idea. Ultimately aiming to capture the true trends and uncertainties and this could provide another option to explore along with the cross validations planned.*

Southwest Pacific Ocean 2021 blue shark assessment follow-up work on diagnostics and grid weighting

In the final session of days 2 Philipp Neubauer discussed approaches being explored to refine the model grid for management advice from the 2021 southwest Pacific blue shark assessment. The work relates to the TOR for WCPFC Project 107b: Towards Providing Scientific Advice for Southwest Pacific Blue Shark (*Prionace glauca*).

Background on the problem was provided noting that the 2021 uncertainty axes covered input data (Catch/F; discard scenarios; CPUE series) and biology (M, Growth, Sigma R, Stock-recruit) resulting in 3888 models used to characterise the uncertainty in management quantities. This large number of models was not acceptable to the SC17 who requested further work to refine down the model grid to provide management advice. The TOR provided suggestions to consider evaluating models in the uncertainty axes in terms of convergence tests, model fit and retrospective/predictive pattern that may allow for elimination of poor or implausible models or provide for statistical model weighting. As part of the work, calculation of an alternation dynamic depletion reference point of $B/B_{F=0}$ was requested to replace B/B_0.

The update on progress indicated that the $B/B_{F=0}$ reference point was now evaluated, and various model convergence tests had been conducted (i.e., final gradients for grid models, Jitter, postive definite Hessian checked with functionality from SS3 adnuts package). The grid of models was run with additional SS3 diagnostics from SS3 diags package, including retrospective analysis: Mohn's ρ for SSB and F and MASE predictive skill for CPUE.

Detailed descriptions of results of the various convergence and diagnostic tests were provided.

Overall, convergence tests did not suggest convergence failure for any models. Positive retrospectives were mainly associated with a single growth model and low recruitment deviation. Predictive skill was difficult to evaluate and further work is required on weighting of MASE for integrated assessments. None of these diagnostics provided a straightforward way to "weight" models. Taken together, results so far do suggest that the Joung et al. growth model is less compatible with other assumptions: higher final gradients, positive retrospective bias and as such could be dropped from the grid axes. The Manning and Francis growth seems better. Dropping the growth axis doesn't really reduce the spread of uncertainty but reduces the number of models by half. Low M also leads to very high, perhaps unlikely, estimates of $B/B_{F=0}$ but unsure if this is enough reason to drop this axis. Different combinations of CPUE inputs and weightings also makes it more difficult to apply model fit diagnostics across the grid, and the uncertain catch and CPUE time series make it difficult to formalise the model fit. Issues with CPUE conflicts was discussed and why this causes MASE > 1 for all CPUE series. Hard to look at MASE in absolute terms when there are multiple CPUE series.

An attempt at model weighting was conducted after dropping the grid axes that are consistently worse in some diagnostics across the whole grid: Joung et al. growth model and Low M scenario. A three step procedure was trialled, Step 1: Remove consistently implausible models, Step 2: Weight the input data, Step 3: Weight by predictive performance. The performance of the approach was described in detail for $B/B_{F=0}$, with most of the effects on the distribution of this reference point being attributed to steps 1 and 2. Issue with weighting based on MASE where discussed, in particular that while it provide an absolute score that can be turned into relative weights, it does not really triage amongst the models, and you can end up with two similarly performing models retained in the grid. Suggested to explore an approach referred to as 'stacking' that looks for a set of models that best describes all the features of the data. It works to try to maximise the predictive ability of a set of models rather the predictive ability of individual models and thus can remove redundant models. However, more work is required. The plan is to try alternative approaches:

1. Weighting for BSH according to above mentioned three-step procedure

2. Compare 'stacking' and MASE based weighting

3. Generate outputs based on weighted ensemble using $B/B_{F=0}$ reference point

Discussion

Question on the three step process was raised regarding how the weights in the second step were done as it still looks to be fairly subjective. *For the catch weighting this was based on the confidence intervals, so is more principled and less subjective, but for the CPUE time series it is more subjective and no clear way to get around this.*

Suggestion to look at other reference points F/F_{ref}. *Agree – good idea to look at performance of the approach for other reference points.*

Some outliers are still present – should explore these combinations of parameters on the extremes and see what can be done with these models. *Yes, this is a good idea, haven't looked at these in great detail yet.*

Question on the CPUE trends presented with the prediction skill (MASE) analysis. Why are these trend different? *It was explained how this related to the smaller sharks in the NZ fishery and therefore the increase in CPUE starts earlier than for the EU and Japan data. Also, because the increasing trends are greater in the EU and NZ data, the model tends to overestimate the increasing trend in the Japan data, trying to fit all the data sets but as a results does poorly on each due to conflicts in trends.*

It was suggested that the SPM (surplus production model) doesn't work, so it is difficult to use the SPM in the model checks, can you explain the reason further. *Two aspects were noted: what diagnostics to use to compares the ASPM value outcomes, as you will get different trajectories, but how do you compare them across 1000's of models in a systematic/consistent way. The other point is we are not estimating much in terms of the production function, the model is using the recruitment to explain the population increase.*

Question was raised regarding the number of models that you might end up with given model weighting is not removing models and SC will be placing a certain focus in the number of models in the grid. *Removing the growth and M axis gets down to around 700 models, but that is still a lot. Some hope that the stacking procedure will be able to remove redundant models to reduce this further.*

Point was made about redundant models potentially biasing management advice derived from the model grids the way they are currently used.

Support was provided to try the stacking, but still some problematic aspects. When this gets presented to SC will also need to be clear that is very difficult, probably impossible, to remove all subjectivity from the process. Also how do you plan to include estimation uncertainty. *Doubtful that this process can ever be totally objective, decision will need to be made by experts or committee. Using stacking can include estimation uncertainty and we plan to look into it.*

Further support and encouragement for the approach being followed was provided.

DAY 3 – Recent development work for MFCL, yellow fin peer review, improving openness and transparency of fisheries science, and progress in harvest strategy and MSE work.

Recent developments in MFCL

Day 3 began with a presentation from Nick Davies of Takin LTD on the extensive development work that has occurred over the last year for the MFCL stock assessment package. The presentation provided thorough descriptions of a number of features that are now incorporated into the latest release version (2.0.8.5) of MFCL, including:

1. Catch-conditioned model
2. Survey Fishery CPUE likelihood
3. Long-term tag loss

The presentation went through the technical description and application of these features on example assessments and noted areas that still required further refinements and testing. Some further work on the catch conditioned approach will involve post-minimization operations such as fishing impact analysis and projections. However, these features will be available for the current skipjack assessment. The catch conditioned feature will be used in the assessment.

The workplan for 2022 was outlined, starting with news that the lead developer of MFCL, Dr Dave Fournier, has retired as of December 2021. This will have implications for ongoing development of MFCL. Until a longer-term plan for MFCL is decided, no significant new developments are planned, and the workplan will focus on consolidation of the recent new features. Beyond that, work in 2022 will involve tidying of the MFCL code; completing a backlog of bug fixes; some outstanding bigeye tuna review panel recommendation; and some "small-scale" requests in the tasks list. Support for MSE requirements will also be important in 2022 along with documentation required to update the MFCL Manual. A number of specific activities for immediate focus were indicated including benchmark testing to ensure integrity of the most recent changes to the development version 2.0.8.5.

Note: MFCL 2.0.8.5 is now released and is the version used for the skipjack assessment.

Discussion

Recognition and appreciation of the contribution of Dave Fournier to the development of MFCL and tuna assessments in the Pacific, and his global contribution to the development of modern fisheries stock assessment. Best wishes to Dave for what comes next.

Someone asked the question of Dave maybe coming out or retirement – unlikely.

2020 Yellowfin tuna assessment peer review

The next presentation provided by Arni Magnusson gave an overview of the planning and progress of the independent peer review of the 2020 yellowfin tuna assessment.

An overview of the 2020 yellowfin assessment was provided as background, pointing out the concerns that the assessment outcomes appeared to be overly optimistic on stock status and presented a major shift from the previous assessment. The review was required to better understand the major difference between the last two assessments and provide recommendations for the upcoming assessment in 2023. Because the structure of the yellowfin assessment is very similar to the bigeye assessment the review has relevance for the 2023 bigeye assessment also.

The review panel selected according to the WCPFC selection process consists of Dr Mark Maunder (IATTC), Dr Andre Punt (University of Washington) and Dr Jim Ianelli (NOAA). These three reviewers also conducted the previous review of the WCPFC bigeye assessment. The review process will consist of periodic communications with the panel while preliminary work is conducted before a week long intensive in-person modelling workshop in Noumea over the period 7-13th September. The review panel chair (Andre Punt) will provide a report that may be provided to a special SC session and/or SC19, and will be presented for discussion at the 2023 PAW.

Preliminary work has involved the new assessment scientist becoming familiar the 2020 yellowfin assessment, training in MFCL and setting up an efficient working system and repository on GitHub. The review process and model development can be followed on GitHub: [PacificCommunity/ofp-sam-yft-review]

Initially modelling work will involve conducting one-off sensitivities to the explore the individual influences of changes made from the 2017 to the 2020 assessment, and then explore alternative structural assumptions such as regional structures. New features of MFCL such as the catch condition approach would be explored later in the process, with the initial aim being to better understand the changes from the 2017 to 2020 diagnostic models.

A summary paper on progress will be provide to SC18.

Discussion

Region 9 was questioned as still being required or not, as it was introduced in the last review process for bigeye to account for the localised tagging data. *May not be necessary for yellowfin at least*.

Data conflicts were raised as an issue that will need to be considered. *Acknowledged.*

A question was raised on how the phases of the preliminary modelling work will link. *For the one-off sensitivities of the stepwise changes from 2017 to 2020 we would use the same model configurations and data, once this phase of work is done we would look at testing the new MFCL features and likely explore the alternative spatial structure with the implementation of new features such as the catch conditioned approach.*

A single region model was attempted at the last minute for the 2020 assessment but did not work for some reason. It was suggested to look again at a single region model. *We will explore simpler region structure.*

Selectivity was an issue and poor fit to length composition was raised as a particularly problematic area in the 2020 assessment. Asked if this would be explored and if so it was recommended to do this earlier on, rather that after implementing new MFCL features. *Yes this will be explored*.

Open and reproducible fisheries science

Arni Magnusson provided a presentation on the topic of open and reproducible fisheries science and plans for the OFP to improve their systems for supporting collaborative work, better repository approach and accessibility, repeatability and transparency. The presentation described what is meant by open and reproducible. Open pertaining to scripts being available to run analysis, data being available for the analysis and any software required being available free. Reproducible is more challenging, it should mean that the analysis can be run on any computer, by different people, on different operating systems and in different software environments. Analysis should be able to be run now and into the future, and can be modified and rerun with different data, different data preparation and even different model options. This is important because it all provides for:

1. Repeatability
2. Institutional memory
3. Reviewability
4. Scientific method
5. Interregional research
6. Dissemination
7. Collaboration
8. Traceability
9. Credibility

The presentation went on to describe approaches to make analysis reproducible and provided an introduction to GitHub as a platform for facilitating open and reducible stock assessments. Following this the Transparent Assessment Framework (TAF) (ICES TAF page: https://taf.ices.dk) developed for ICES assessments was described.

The presentation finished with mention of the plan for OFP to work on improved systems, especially reproducibility of assessment including supporting analysis. This is critical for a group such as OFP with high turnover of assessment scientists and new scientist having to pick previous assessments and preparatory analyses done by others. This is an acknowledged major bottle neck in the workflow efficiency that needs to be improved. GitHub provides the platform but there needs to be a structure and framework for how it is used, and importantly, that is consistently applied by all assessment staff.

There was no discussion from the floor on this presentation.

Progress on harvest strategy work

Update on WCPFC harvest strategy workplan

Rob Scott for SPC-OFP started the harvest strategy session of with an update on progress of the WCPFC harvest strategy workplan and priority areas for 2022. For both skipjack and south Pacific albacore there is an expectation under the revised workplan that the WCPFC in 2022 will review and adopt management procedures. This will require the SC18 agree to the operating models in the MSE frameworks, provide advice of the performance of candidate management procedures, provide advice on relevant elements of the monitoring strategy. TCC (Technical Compliance Committee) will also need to consider the implications of candidate management procedures. For bigeye and yellowfin the work will continue to develop the mixed fishery MSE framework and development of candidate management procedures for bigeye, depending on progress with the technical work SC18 may be required to provide advice on potential bigeye management procedures. For bigeye and yellowfin the current WCPFC workplan aims to adopt management procedures in 2024.

The Science Management Dialogue (SMD) meeting that will occur shortly after SC18 was then introduced. This will be the first SMD on the harvest strategy work and is being referred to as trial meeting. The first SMD will bring together managers from WCPFC CCMs with scientists, with the aim to achieve consistent understanding amongst members of the harvest strategy approach, through sessions of capacity building and general discussion and dialogue. The SMD hopes to prioritise/identify (a subset of) preferred management procedures (focussed on skipjack and south Pacific albacore) for consideration by WCPFC19 and to develop/consider future processes/pathways (short and long-term) to facilitate Commission decision making on management procedures by determining how the Commission can be best supported by SC; TCC; and future SMDs.

Under the Pacific Tuna MSE project (funded by NZ Ministry of Foreign Affairs and Trade) 11 papers are planned for SC18 and the Management Theme will be dominated by MSE/Harvest strategy papers and discussion.

Following the update on the WCPFC workplan, updates on progress of the technical work and MSE frameworks were provided for skipjack, south Pacific albacore and the mixed fishery approach.

Skipjack MSE framework

- All components of the Skipjack MSE model were updated for the 2019 assessment, tested and suitable for progressing formal evaluation studies
- Results for 9 alternative management procedures available on PIMPLE.
- Proposed **monitoring strategy** needs refining including any **CCM specific performance indicators** that cannot be derived directly from the MSE framework.
- Ongoing work to refine elements of the **robustness set**.

Details can be found in SC17-MI-WP-04

Basis of the skipjack Management Procedure is outlined in Appendix A
Assumptions regarding the implementation of archipelagic waters detailed in Appendix B
Evaluation software and input data repository described in Appendix C

The details of the skipjack uncertainty/operating model grid (below) were discussed.

Axis	Levels	Options		
	Reference	0	1	2
Recruitment variability	2	1982-2014	2005-2014	
Catch & effort	1	20%		
Size composition	1	Estimated ESS		
Tag recaptures	1	Status quo		
Steepness	3	0.8	0.65	0.95
Mixing period	2	1 qtr	2 qtr	
Growth	2		Low	High
Movement	1	Estimated		
Hyper-stability in CPUE	2	0	-0.5	
Effort creep	2	0%	2%	

South Pacific albacore MSE framework

The South Pacific albacore MSE framework is based on the 2018 (5 region) assessment models. The initial focus was on Empirical MPs (using CPUE), but after a good go at this approach in the end it was not that successful, so currently investigating model based approaches for the MP using the surplus production model 'SPiCT' (Stochastic Production model in Continuous Time). SPiCT accounts for observation error in catches and biomass indices and autocorrelation in time series of catches and biomass indices. The proposed monitoring strategy needs refining including any CCM specific performance indicators that cannot be derived directly from the MSE framework and more work is required to refine elements of the robustness set. It is expected that the model based MP will be presented to the SC18 and provide the basis further MP evaluations.

The details of the south Pacific albacore uncertainty/operating model grid (below) were discussed.

Axis	Levels	Options		
	Reference	0	1	2
Recruitment variability	1	1960-2014		
Catch & effort	1	20%		
Size composition	-			
Tag recaptures	-			
Steepness	3	0.8	0.65	0.95
Natural Mortality	2	0.3	0.4	
Growth	2	Est. von B	Chen-Wells	
Size Comp. Weighting	3	20	50	80
CPUE	2	Geo-stats	Traditional	
Movement	-			
Effort creep	2	0%	2%	

The PAW was reminded on the R Shiny apps. available for stakeholders to explore the performance of alternative harvest control rules:

Skipjack: PIMPLE https://ofp-sam.shinyapps.io/pimple/

South Pacific albacore: SPAMPLE https://ofp-sam.shinyapps.io/spample/ (note currently uses the CPUE based MPs)

Also, the Moodle learning repository: https://spc.learnbook.com.au/login/index.php (login as guest using password '2cool4school').

Development of bigeye tuna operating models (OMs) for the mixed fishery framework

Progress on the develop of bigeye OMs was presented. The initial work has been to do retrospectives on the 24 grid models from the 2020 bigeeye assessment. The result look reasonable except for some models where the size composition was severe downweighted (divisor 500), runs of these models struggled to converge and with poor retrospective performance. Considering removing the models with the severely downweighted composition data from the OM set.

Hindcast analysis has also been connected to test predictive skill, with the result also looking reasonable. A surplus production model in SPiCT was also run in comparison to MFCL and looks to work quite well as a potential estimation method for the MPs.

The details of the bigeye uncertainty/operating model grid (below) were discussed.

Axis	Levels	Options			
	Reference	0	1	2	3
Recruitment variability	2	1982-2018	2005-2018		
Catch & effort	1	20%	??		
Size composition	1	-			
Tag recaptures	1	-			
Steepness	3	0.8	0.65	0.95	
Tag mixing period	1	1qtr			
Size Composition Wt	4	10	60	200	500
Growth	2	Diagnostic	M-high		
Movement	1	Estimated			
Hyper-stability in CPUE	-	-			
Effort creep	2	0%	2%		

Question was raised by the presented as to whether hyperstability and effort creep scenarios should apply to longline fisheries.

Mixed Fishery harvest strategies

Finlay Scott of SPC-OFP provided a presentation on the mixed fishery MSE approach being developed. The approaches integrates the singles stock management procedures developed for skipjack, south Pacific albacore and bigeye, yellowfin does not require is a management procedures and is managed under the skipjack, albacore and bigeye MPs. The table below indicates under which MPs will the various tuna stocks and major fishery grouping be controlled under. i.e., yellowfin tuna in the tropical purse seine fishery would be managed under the skipjack MP.

	Skipjack	Yellowfin	Bigeye	South Pacific Albacore
Tropical PS	**SKJ MP**	SKJ MP	SKJ MP	
Northern PS	SKJ MP	SKJ MP	SKJ MP	
Tropical LL		BET MP	**BET MP**	ALB MP / BET MP
Northern LL		BET MP	BET MP	
Southern LL		ALB MP	ALB MP	**ALB MP**
PL	SKJ MP	SKJ MP	SKJ MP	
ID / PH / VN (non-AW)	SKJ MP	SKJ MP	SKJ MP	
Southern Troll				ALB MP
Archipelagic Waters and Territorial Seas	Aligned to SKJ MP, national plan or local MP	Aligned to SKJ / BET MP, national plan or local MP	Aligned to SKJ / BET / ALB MP, national plan or local MP	Aligned to ALB MP, national plan or local MP

A detailed description of the interactions among the three MPs and ow they control fishing of the four stock across the different fishery groupings was provided. Given that south Pacific albacore is mostly caught in the southern long line fishery that also does not catch much yellowfin, bigeye or skipjack, a specific example was provided where albacore could be excluded for the mixed fishery framework to further simplify it. This is worth considering.

Mixed fishery indicators are also being developed to support evaluation of combined MPs operating with the mixed fishery MSE framework.

A demonstration 'proof of concept' evaluation of several skipjack HCRs showing implications for bigeye and skipjack was presented.

The next phase of work on the mixed fishery MSE framework will involve developing the BET and YFT OMs, noting the current evaluations are based on most recent assessment (2020) and the next assessments are in 2023. Also noted that the peer review of the yellowfin assessment is occurring and may have further implication for which OMs used in the mixed fishery MSE. Work on develop MPs for BET, start with a SPiCT model-based approach for the estimation model.

Discussion

Regarding the OM grid for south Pacific albacore the question was asked around what recruitment is included give the recent recruitment patterns detected in the last assessment. *Currently basing this OM grid on the 2018 assessment with recruitment variability up to 2014. Recruitment variability is one of the biggest sources of uncertainty particularly in the projections for south Pacific albacore. Projections from the 2018 assessment were however quite consistent with the estimates from the 2021 assessment. Uncertainty should be included in the current OM and the recruitment variability in these should represent the expected variation.* Temporal correlations in poor or strong recruitments – there is potential to simulate extended periods of reduced recruitment in SP Albacore. *Should be possible to include some additional recruitment scenarios into the albacore OMs to include temporal correlation in poor recruitment.*

Question asked to remind in the skipjack OM grid the difference between the hyperstable axis and the effort creep axis. *The effort creep is a scalar applied to the effort, it is either turned off, or has a value- i.e., 2%, it means a fixed year on year 2% increase in effort on 2012 levels is applied through the projections. Hyperstability in CPUE is a scalar applied to catchability depending on biomass of the stock. For effort creep it is just a forward projection, hyperstability is a model fit so includes the historical catchability, may increase the sensitivity to the catchability in the robustness set.*

Question on dip at end of bigeye hindcasts. *Likely related to recruitment, but early stages more peel to do.*

Point made on the need for SC to agree on OMs. Any thoughts on criteria to guide SC on formal adoption of OMs. *We'll talk more on this after lunch, hopefully discussion today will help with this.*

Point on effort creep made that might be useful to compile a list of the changes that have occurred in the purse seine and longline fisheries that might be able to guide discussions on appropriate levels of effort creep to include in the projections.

Chair respond that this would require work by DWFN, PICTs and industry involvement and would take a bit of work to bring together – worthwhile to do the compilation but the team at OFP would be unable to do it by SC and would need collaboration by DWFN scientists.

Work so far was commended. Comment on the skipjack OM uncertainty grid, seems that density dependent effects and spatial and temporal process variation is missing. The OMs and estimation models are very similar so what about putting in some temporal variation of processes into the OMs, include axes that represent impacts of environmental variation/uncertainty, such as ENSO.

Density dependent effects are captured to some extent by the hyperstability axis and recruitment estimations but other processes are assumed stationary. Currently difficult to have time varying process (i.e., growth). The projections go for thirty years and in reality, MPs will be reviewed over much shorter time frames, so they can be adapted as things change, also there is the monitoring strategy that can also indicate if things are no longer working as expected and need review. Further, noted this is the first attempt to build working MSE frameworks, but non-stationarity is clearly a very important thing to have foremost in minds as we continue to improve these modelling systems.

Skipjack operating models

Rob Scott continued with a presentation on the skipjack operating models focussing on the need to from agreement on the OMs and the potential implications of new assessments for OM grids. He showed the comparison between the skipjack OM grids conditioned on the 2016 and 2019 assessments and the comparisons of the south Pacific albacore stock assessment grids for 2019 and 2021. Making the point in both cases that although the assessments change quite lot, their ranges of uncertainty in the key management output ($SB/SB_{F=0}$) still overlapped. It is not necessary to update the OMs in an MSE framework when a new assessment is conducted, the OMs in the MSE framework should capture a sufficient range of uncertainty that will encompass the range of estimates provided by new assessments. The traditional stock assessments are then used as a monitoring tool that can change overtime as new information comes to light and assumption might get changed. If the assessment outcomes (considered the best scientific information on stack status) start to move outside the range of uncertainty predicted the MSE then this can trigger a review and the OMs may then need to be updated and MPs re-evaluated.

The skipjack OM grid was revisited and R Shiny app. Hierophant: https://ofp-sam.shinyapps.io/hierophant/ was presented. This app. allows exploration of model diagnostics across all models in the OM grid. This app was produced to allow people to explore and compare each model in the OM grid to perhaps assist with decisions on the composition of the models retained in the final agreed OM grid.

In relation to reaching agreement on an OM grid at SC18 – there are perhaps two key questions to ask:

- Have all important sources of uncertainty been considered?
- Do the model diagnostics indicate acceptable fits to the data?

The chair raised the point we really need to move forward on agreement or otherwise on the OMs, especially for skipjack, and that the PAW has the mix of people who can help advise on what might help to do this.

Comments were and on outputs provided and how to improve them for the end users:

1. For the ribbon overlay plots include the medians to show differences/similarity between new and old assessments and existing OM grid in terms of stock depletion.
2. Also, would be good to show projections from the 2019 assessment OM grid with the estimates for the most recent (2022) assessment so people can see how similar the projections from the 2019 OM grid are to the new assessment for recent years.

In relation to environmental effects, most of the major effects are not predicted to occur for 15-20 years based on SEAPODYM, so for incorporating environmental effects, might want to consider these effects coming into play somehow in 15 years' time.

Comment regarding the inclusion of some of the tag related uncertainties, depending on how the sensitivities come out in the current assessment.

Perhaps there are some simple options to include periodic environmental effects on recruitment dynamics, stock recruitment etc.. Do some analysis to look for evidence of these effects in the data, comparing model predictions of recruitment etc. to binary classifications of ENSO events for example. *Consider for future work.*

Further comments regarding effort creep for skipjack. Effort creep could be considered in the historical period as an aspect of uncertainty, by including another OM grid axis for effort creep. This would be expected to increase the range of uncertainty in the OM models to lower levels of depletion (based on results of the preliminary analysis by NDB). *We will consider this.*

Re-emphasised the key question is to capture the most important and most consequential uncertainties. There are many uncertainties, but they may not all result in broadening the uncertainty.

Question on alternative catch levels in archipelagic waters. The MPs can't control these catches. Need sensitivity analyses for increased future catches particularly in region 5 archipelagic water. *Alternative future archipelagic catch trajectories can easily be included and could form either part of the robustness set or the reference set. (need further discussion with WCPFC members)*

Comment on the different catch outcomes for purse seine depending on catch or effort controls in relation to changes in density and species composition of schools. Under the current MPs pole and line is controlled by catch and purse seine by effort, if population stays stable the catch outcomes are similar for both. However, when the stock changes significantly the catch responses can be different for the effort and catch controlled fisheries. For example, if the stock declines rapidly the allowed catches will stay the same for the catch managed fisheries (they will need more effort to catch it but their effort is not constrained by the HCR), but for the effort managed fisheries although effort is the same they will catch less due to the lower available stock and they cannot increase effort as it is constrained by the HCR.

Clarified that as the density of species changes the school compositions can also changes which could change how the projected effort for purse seine sets impacts the different stocks. For, example if skipjack declines perhaps purse seine sets are more compromised of yellowfin. *I guess we were unclear on this question but might be getting at that the model makes some assumptions of the impacts of a set on the F for different species, but that this might change as densities change due to school composition effects. Not clear how we could deal with that.*

Revisiting the queries around the recruitment time series. Noting that the range of recruitment deviation was less in the 2021 assessment compared to the 2018 assessment. Does this mean that you should change the recruitment deviates used in the projection? *Can't mix the recruitment series across models but we can add temporal autocorrelation into the recruitment deviates and creates different scenarios for the projections.*

Support the need to develop a system with a set of criteria to tick or score to guide agreement of the OMs.

Commenting on the lack of responses to feedback on OM grids. Recommendation of a paper from this workshop to request for comments around the OM grids. *Will consider this with the MSE team*.

As to deciding on OMs – can consider the diagnostics, as to the ranges of uncertainties – seems reasonably broad. For the SC perhaps they should focus on that the uncertainty captures the next 5 or so years. Also, perhaps better to focus on the uncertainty coverage than the diagnostics which can be problematic as to how to use them to choose among OM models. There will be review processes. Also need to consider further the exceptional circumstances. It's more about doing something sensible, appropriate and practical now rather than trying to fully account for everything at this stage.

Skipjack Management Procedure Dry Run

The final presentation outlined a paper planned for SC18 to demonstrate how the combination of a harvest strategy and the traditional stock assessment work together when the harvest strategy is implemented. The presentation covered background on harvest control rules (HCRs), and the management procedure, including the estimation model. It then showed how the traditional stock assessment fits alongside the harvest strategy as part of the 'monitoring strategy'. The harvest strategy components, specifically the MP (management procedure) and its HCR and estimation model should remain fixed for as long as the MP is continuing to meet the objectives. That is are we meeting the performance indicators such as the target reference point for stock biomass etc.. The way we check we are meeting objectives is through the monitoring strategy. The traditional stock assessment is the key part of the monitoring strategy and can and does change as the goal of the process is to provide the best available science to estimate the current status of the stock, so as new credible information comes available, it should be used to better inform the stock assessment. The key point is that the original MSE framework used to test and select the MP should capture a large enough range of uncertainty so that the new assessment estimates sit within the range of uncertainty estimated by the models used to evaluate the MP. If overtime the traditional stock assessment estimates that stock status moves outside the range expected from the evaluations of the chosen MP, then the models used for the MP evaluations might need to be updated and the MPs re-evaluated.

Point was made in the discussion that it needs to re-emphasised that even though the MP evaluations might show the average expected performance, the reality is that stock will fluctuate around the target due to random process such as recruitment variation. The MPs HCR responds to the estimated stock status to continually try to adjust fishing (up or down) to push the stock towards the target objective.

Meeting closed

The meeting was closed by the chair, thanking everyone involved, especially the presenters, OFP stock assessment team, and those that were joining outside of hours, noting that 50-60 attendees logged on each day.

APPENDIX 1: Agenda

Tuesday 29th March (Monday 28th US)	**DAY 1: Preparatory Workshop for 2022: skipjack tuna assessment** E-meeting (Zoom) Link: https://spc.zoom.us/j/99881478003?pwd=b0h1bjg5V2wreFB6RWVoTE9sSnJLZz09 Meeting ID: 998 8147 8003 Passcode: 910567 **Times are New Caledonia time**	**Presenter initials and presentation numbers**
09:00 – 09:15	• Introduction • Reminder of ToR and objectives for the SPC preparatory workshop • Agenda and meeting format/procedures • Any other introductory comments	PH
09:15 – 10:30 *Session 1* *(75 mins)*	• Previous skipjack tuna assessment summary (10 mins) • Proposed model spatial and fishery structures (20 mins) • Fisheries overviews and data inputs (30 mins) (60 mins)	• CCJ/PH **(D1-P1)**
10:30 – 10.50	BREAK	
10.50-12.10 *Session 2* *(80 mins)*	• Composition data – reweighting procedure (15 mins) • Tagging data - background (10 mins) • Tagging data pre-treatment overview (20 mins) • Tagger effects (20 mins) (65 mins)	• TP **(D1-P2)** • PH/JSP **(D1-P3)** • TT **(D1-P4)** • TP **(D1-P5)**

12.10-13.00	BREAK	

13:00 – 14:30
 	
Session 3
 	
(90 mins)
	• Tag mixing period simulations (40 mins)
	• CPUE analyses for 2022 assessment (40 mins)
	    ○ Pole and line
	    ○ Purse seine
	    ○ Philippines purse seine
	
	(80 mins)
		• JSP **(D1-P6)**
		• TT/AY/KB **(D1-P7)**

14.30-15.30 *Session 4* *(60 mins)*	• Diagnostic model development and new MFCL features (focussing on catch conditioned and CPUE treatment) (30 mins) • Preliminary model results (15 mins) • Key uncertainties (15 mins) • Other (60 mins)	• CCJ/ND **(D1-P8)**
15.30-15.40	Discussion and wrap up day 1	PH
Wednesday 30th March (Tuesday 29th US)	**DAY 2: Preparatory Workshop for 2022: Southwest Pacific mako shark assessment and southwest Pacific blue shark 2021 assessment follow-up work on model selection**	
09:00 – 9.10	Get online and intro to day 2	PH
9.10-10.30 *Session 5* *(80 mins)*	• Review of data inputs for southwest Pacific mako shark (30 mins) • Biology and assessment plan (30 mins) (60 mins)	• SB **(D2-P1)** • KL/PN **(D2-P2)**

10:30 – 10.50	BREAK (20 mins)	
10.50-12.20 *Session 6* *(90 mins)*	Data inputs and preparation including preliminary catch reconstruction (70 mins)	• KL/PN **(D2-P2)**
12.20-13.10	BREAK (50 mins)	
13:10 – 14:40 *Session 7* *(90 mins)*	• Southwest Pacific blue shark assessment follow-up work: background • Approaches to model selection for management advice • General discussion on the proposed approach (as required) (60 mins)	• PN/FC **(D2-P3)**
14.40-15.30 (40 mins)	Skipjack uncertainty characterization cont.	All
Thursday 31st March (Wed 30th US)	**DAY 3: Preparatory Workshop for 2022: Yellowfin tuna peer review, MFCL developments, Management Strategy Evaluation/Harvest strategies**	
09.00-9.10	Get online and intro to day 3	PH

9.10-10.00 Session 8 (50 mins)	• Recent developments and future work for Multifan-CL (35 mins) (35 mins)	• ND **(D3-P1)**
10.00-10.40 Session 9 (40 mins)	**Yellowfin tuna peer review** • Yellowfin tuna assessment peer review (20 mins) (35 mins)	• AM/JD **(D3-P2)**
10:40 – 11.00	**BREAK**	
11.00-11.30 Session 9 (cont) (30 mins)	• Making things open and reproducible (20 mins)	• AM **(D3-P3)**
Session 10 11.30-12.30 (60 mins)	**<u>MSE sessions</u>** • Update on WCPFC harvest strategy work plan (10 mins) • Update on MSE frameworks: skipjack, south Pacific albacore, mixed fishery (30 mins) (40 mins)	• RS/FS **(D3-P4)**
12.30-13.20	**BREAK**	

13.20-14.30 Session 11 (70 mins)	• Skipjack operating models (OMs): what is required from SC18 for agreeing on the OMs? (30 mins) • Skipjack management procedure trial run with 2022 assessment (10 mins) (60 mins)	
14.30 Wrap up and Follow-up	**WRAP UP** • Workshop recommendations/key points circulated • Recommendations agreed • Meeting draft paper circulated for comments • Comments received • Meeting paper finalized for SC16 submission	PH

PH Paul Hamer, AM Arni Magnusson, CCJ Claudio Castillo Jordan, ND Nick Davies, JD Jemery Day, TT Thom Teears, RS Rob Scott, FS Finlay Scott, TP Tom Peatman, AY Aoki Yoshinori, NM Naoto Matsubara, NY Nan Yao, FP Felipe Carvalho, SB Steve Brouwer, KL Kath Large, PN Philipp Neubauer, JSP Joe Scutt Phillips, KB Keith Bigelow

APPENDIX 2: List of participants

(NB – there may be some inaccuracy due to the online format, multiple people joining from the same link and people dropping in an out of the meeting for shorter periods)

Name	Affiliation
John Annala	Ministry for Primary Industries, NZ
Leyla Knittweis	Ministry for Primary Industries, NZ
Brad Moore	NIWA, NZ
Simon Hoyle	NIWA, NZ
Brian Kumasi	PNAO
Philipp Neubauer	Dragonfly Data Science
Kath Large	Dragonfly Data Science
Steven Brouwer	Sagittas LTD
Nick Davies	Takina LTD SPC consultant
Rob Campbell	CSIRO, AU
Jessica Farley	CSIRO, AU
Ashley Williams	CSIRO, AU
Rich Hillary	CSIRO, AU
Laura Tremblay Boyer	CSIRO, AU
James Larcombe	Department Agriculture Water and the Environment, AU
Reuben Sulu	FFA Secretariat
Lianos Triantafillos	FFA Secretariat
Adele Dutilloy	FFA Secretariat
Keith Bigelow	NOAA (Pacific Islands Fisheries Science Centre), US
Filipe Carvalho	NOAA (Pacific Islands Fisheries Science Centre), US
Jon Brodziak	NOAA (Pacific Islands Fisheries Science Centre), US
Michelle Sculley	NOAA (Pacific Islands Fisheries Science Centre), US
Matthew Vincent	NOAA (Beaufort Lab), US
Nicholas Ducharme-Barth	NOAA (Pacific Islands Fisheries Science Centre), US
Mark Fitchett	Western Pacific Regional Fishery Management Council (US)
Haikun Xu	IATTC
Carolina Minte-Vera	IATTC
Eric Chang	National Sun Yat-sen University, TW
Yi-Jay Chang	National Sun Yat-sen University, TW
Kwang-Ming Liu	Deep Sea Fisheries Division, TW
Ren-Fen WU	Overseas Fisheries Development Council, TW
Keisuke Satoh	Japan Fisheries Research and Education Agency
Hiroki Yokoi	Japan Fisheries Research and Education Agency
Kei Okamoto	Japan Fisheries Research and Education Agency
Yuichi TSUDA	Japan Fisheries Research and Education Agency
Aoki Yoshinori	Japan Fisheries Research and Education Agency
Naoto Matsubara	Japan Fisheries Research and Education Agency
Mikihiko Kai	Japan Fisheries Research and Education Agency
Yasuko Semba	Japan Fisheries Research and Education Agency
Jun Matsubayashi	Japan Fisheries Research and Education Agency
Zhe Geng	Shanghai Ocean University
Jiangfeng Zhu	Shanghai Ocean University
Fan Zhang	Shanghai Ocean University
Jingtao Wang	Shanghai Ocean University

Yang Wang	Shanghai Ocean University
Yuchen Huang	Shanghai Ocean University
Mi Kyung Lee	National Institute of Fisheries Science, KR
Youjung Kwon	National Institute of Fisheries Science, KR
Glen Holmes	PEW
SungKwon Soh	WCPFC Secretariat
Elain Garvilles	WCPFC Secretariat
Mickael Lercari	Service du parc de la mer de Corail et de la pêche – SPNMCP), NC
Tim Adams	Gonedau
Claudio Castillo Jordon	SPC
John Hampton	SPC
Graham Pilling	SPC
Sam McKechnie	SPC
Lauriane Escalle	SPC
Rob Scott	SPC
Finlay Scott	SPC
Nan Yao	SPC
Jemery Day	SPC
Arni Magnusson	SPC
Thom Teears	SPC
Paul Hamer	SPC
Joe Scutt Phillips	SPC
Simon Nicol	SPC
Jed McDonald	SPC
Steven Hare	SPC
Tom Peatman	Independent Consultant
Shana Miller	Ocean Foundation

APPENDIX 3: Terms of Reference

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

Terms of Reference

- Review the most recent completed assessments, in particular, any concerns, suggestions and/or recommendations raised by the Scientific Committee, the Commission, research providers, individual CCMs, or any independent reviews;
- Review preliminary work undertaken by the service provider relating to the stock assessments, including any proposed:
    - revisions to biological parameters
    - revisions to historical data
    - changes to structural assumptions in the model
    - methodological issues, e.g. characterization of uncertainty
    - standardized CPUE analysis
    - incorporation of tagging data or other auxiliary data

- Provides guidance to the OFP on:
    - the suitability of any proposed changes and any suggested additional work
    - a minimum set model runs to be undertaken, in particular the range of key sensitivity analyses
    - desired model diagnostics to be presented
    - alternative modelling approaches that could be considered

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

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

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

APPENDIX 4: Background paper for the skipjack assessment

2022 Western and Central Pacific Ocean Skipjack tuna assessment: background paper for the Pre-assessment Workshop

Oceanic Fisheries Programme, SPC

Introduction

Stock assessments of skipjack tuna (*Katsuwonus pelamis*) in the Western and Central Pacific Fisheries Commission convention area (WCPFC-CA) (Fig. 1), have been conducted at three-year intervals since 2000 (Bigelow et al. 2000). The most recent assessment was conducted in 2019 (Vincent et al. 2019). That assessment is described in more detail below. The assessments apply the integrated assessment approach using the stock assessment framework Multifan-CL. The development of the 2022 assessment is based off the 2019 assessment and where possible aims to improve on the data inputs and modelling approaches applied. Major changes to structural assumptions may occur if reliable new information comes to light to support such changes. The Pre-assessment Workshop (PAW) discussions provide an important forum for information exchange, advice and constructive feedback to assist with the improvement and development of the new assessment. SPC scientists make the final decisions on the data inputs, preparatory analyses, structural and other assumptions required during model development, taking into consideration the advice and recommendations from the PAW and the available time and resources.

This background paper is aimed to provide a general introduction to skipjack fishery in the WCPFC-CA, the previous WCPFC-CA skipjack assessment, and the key aspects where changes/improvements in the assessment approach are being considered. Hopefully it provides useful background for your preparation to participate in the skipjack discussions at the PAW.

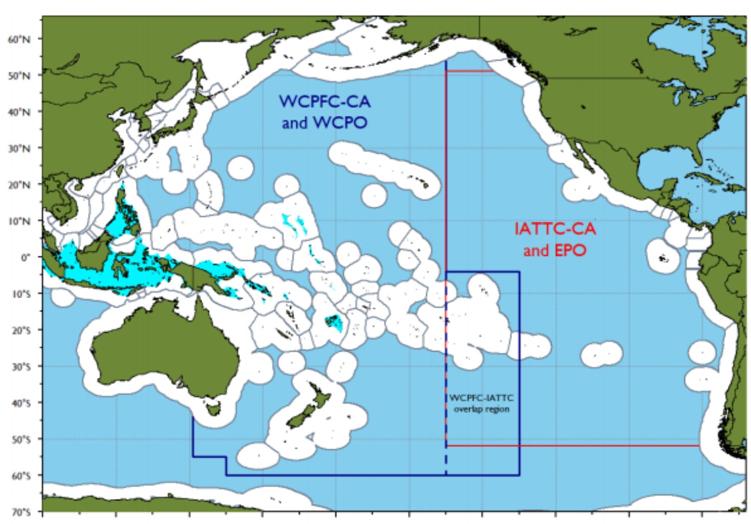

Figure 1. Important national, regional and management zones in the Pacific. The WCPFC Convention Area (WCPFC-CA) is outlined in dark blue, the IATTC Convention Area (IATTC-CA) area is outlined in red. The western and central Pacific Ocean (WCPO) includes all of the WCPFC-CA, minus the overlap with the IATTC-CA; the eastern Pacific Ocean (EPO) is coincident with the IATTC-CA. Pacific nation EEZs are outlined in grey and archipelagic waters are shaded turquoise (from Hare et al. (2021)).

WCPO skipjack fishery background

The skipjack tuna fishery in the WCPFC-CA accounts for approximately 35% of the global tuna catch, and around 60% of the global skipjack catch. The annual catch has ranged from around 1.6 to 2.04 million tonnes since 2015, with the largest catch recorded in 2019, estimated at approximately 2.042 million tonnes (Fig. 2) (Hare et al. 2021). The catches are dominated by the purse seine fishery, that has accounted for 77-83 % of the catch since 2015. This compares to the early 1980's when catches were in the order of 400 – 700 thousand tonnes and the pole and line fishery accounted for around 50-70 % of the catch. The majority of the WCPFC-CA skipjack catches in recent years are from the equatorial region between 10°N and 10°S (Fig. 3).

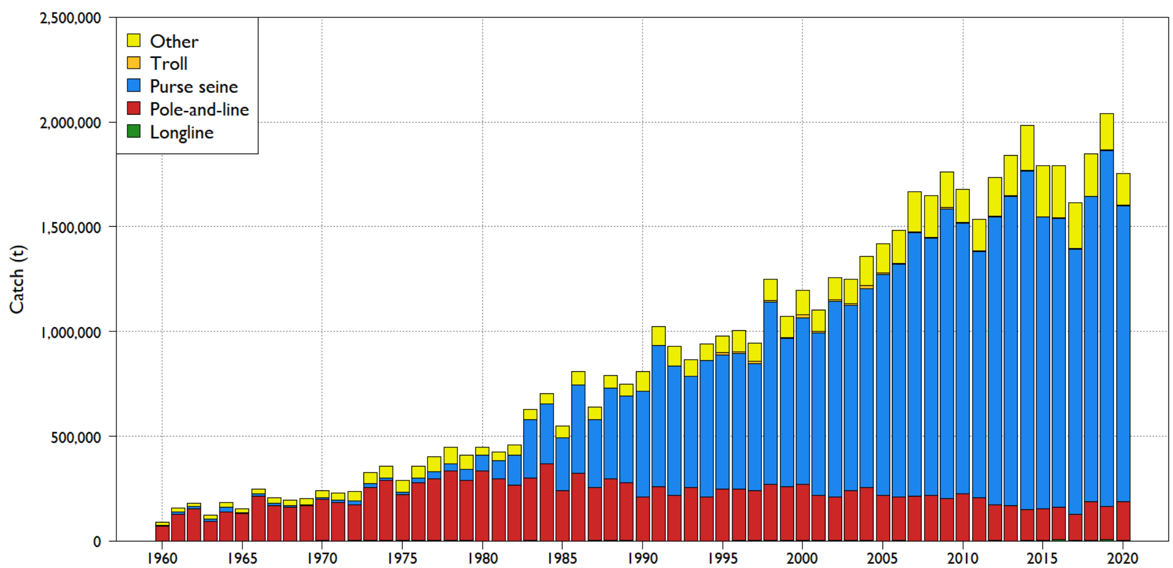

Figure 2. Skipjack tuna catch by gear in the WCPFC-CA since 1960.

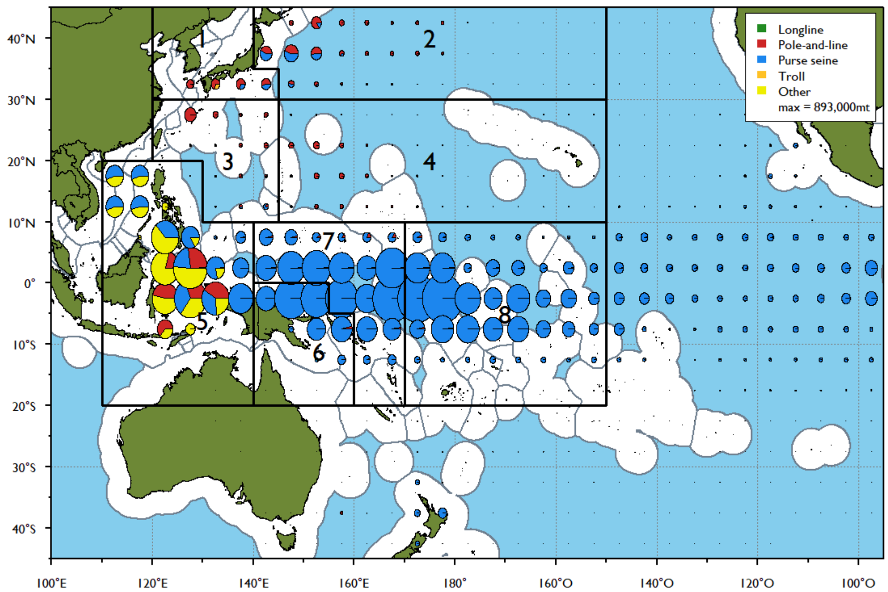

Figure 3. Skipjack tuna catch distribution by gear in the Pacific Ocean for 2016-2020. The 8 regions used for management advice in the 2019 assessment are overlayed.

The growth of the purse seine fishery took off from the mid-1980's (Fig. 2). There are two main modes of purse seine fishing: 'unassociated' (or free school) which refers to targeting free ranging tuna schools, and 'associated', which refers to purse seine sets targeting tuna schools aggregated around floating objects such a natural logs or purpose built anchored or drifting fish aggregation devices (FADs). Overtime the 'associated' fishing mode has become dominated by constructed drifting FADs (dFADs). This has been facilitated by the availability of satellite tracking buoys attached to the FAD rafts, and more recently, acoustic sensors added to the buoys that transmit information on the presence and biomass of tuna associated with individual dFADs. Despite the introduction of these FAD buoy technologies, the proportion of the purse seine skipjack catch taken by 'associated' and 'unassociated' fishing modes has remained relatively similar over the last decade.

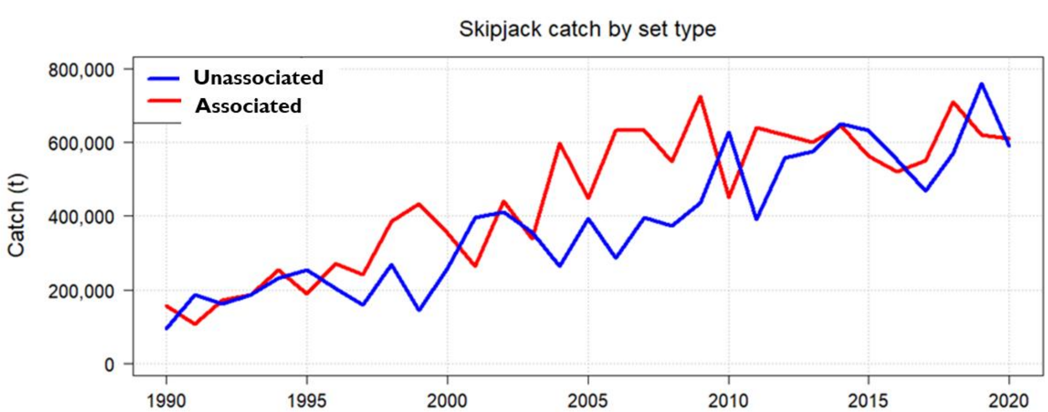

Figure 4. Comparison of annual skipjack catches by set type in the WCPO.

Small amounts of skipjack catch are taken by longline and troll fisheries (< 5,000 tonnes/year for each gear). 'Other' gears, which mostly refers to artisanal fishing gears in the Indonesia, Philippines and Vietnamese fisheries now account for a similar percentage, approximately 10%, of the catch as the pole and line fishery, although these fisheries catch higher numbers of fish due to their selectivity for smaller skipjack (Fig. 5).

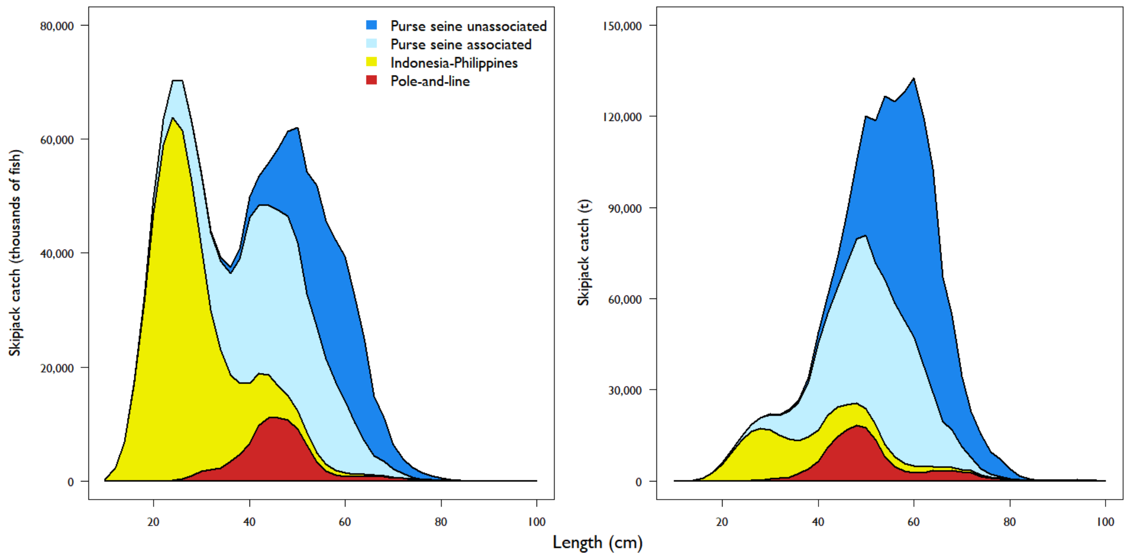

Figure 5. Size composition of skipjack tuna caught by various fishery components – by numbers (left) and tonnes (right).

The revenue derived from the skipjack fishery is incredibly important to Pacific Island countries. For some, the tuna resources within their exclusive economic zones (EEZs) represent their only significant renewable resource and their best opportunity for economic development. Pacific Island countries who are Parties to the Nauru Agreement (PNA) (Federated States of Micronesia, Kiribati, Marshall Islands, Nauru, Palau, Papua New Guinea, Solomon Islands and Tuvalu, plus Tokelau) control most of skipjack catch from the WCPFC-CA and obtain the most financial benefit from the resource through the Vessel Day Scheme (VDS) and other arrangement such as chartering. The VDS is a system of effort management that limits the number of purse seine fishing days in the PNA country EEZs, with days being sold at or above a minimum rate agreed by The Parties. The main controls on skipjack fishing overall are through limits on the number of purse seine fishing days either by the VDS or flag specific day limits on purse seine fishing on the high seas stipulated under the Western and Central Pacific Commission's tropical tuna conservation and management measure (CMM 2021-01) (although some flag based catch limits also apply). This CMM also provides the overarching stock conservation objectives for the management of the skipjack stock, as part of the broader tropical tuna fishery, including yellowfin and bigeye.

While CMM 2021-01 provides the current high-level guidance, objectives and management limits for the tropical tuna fisheries, including skipjack, this CMM is viewed as a bridging measure while a harvest strategy is developed and implemented. The development of the skipjack harvest strategy has progressed to the point where a fully functional MSE (Management Strategy Evaluation) framework is available and is being used for testing of candidate management procedures (including the harvest control rules). The WCPFC, under its current harvest strategy workplan (WCPFC Harvest Strategy

Workplan Progress), is scheduled to adopt a management procedure for the WCPFC-CA skipjack stock by the end of 2022. The operating models in the skipjack MSE framework are based on the 2019 assessment. However, it is not necessary to update MSE operating models for every new assessment, so long as they continue to adequately incorporate the range of uncertainty in stock assessment predictions of stock status. The regular stock assessments (i.e., the new 2022 assessment) become a key part of the ongoing monitoring programme that tracks the performance of the harvest strategy overtime, and this will be discussed in sessions 10 and 11 on day 3 of the PAW.

Summary of the 2019 assessment

The 2019 WCPFC-CA skipjack assessment (Vincent et al. 2019) involved two spatial structures, a 5-region (similar to the 2016 assessment) and an 8-region structure (Fig. 6). While there were merits to both spatial structures, the 8-region structure was used as the diagnostic (base case) model and the 5-region model as a sensitivity. However, for the provision of management advice the **Scientific Committee (SC) 15 "*agreed to use the 8 region model to describe the stock status of skipjack tuna because SC15 considers that it better captures the biology of skipjack tuna than the existing 5 region structure*"** *(SC15 Summary Report).*

The 8-region model included 31 fisheries, across pole and line, purse seine, longline, troll and other methods. The structural uncertainty grid used to represent uncertainty in model derived management quantities included axis for:

- Regional structure (5 and 8 regions - noting the 5-region model was dropped by SC15)
- Steepness (0.65, 0.80 as the diagnostic model, 0.95)
- Tag mixing period (1 quarter as the diagnostic, 2 quarters)
- Growth estimated internally from length composition modal structure (Default, low growth, and high growth)
- Length composition weighting factor (i.e., sample size devisor) (50, 100 as the diagnostic model, 200)

The assessment involved 108 models of which the spatial structure options were considered separately resulting in two sets of 54 models.

The main outcomes of the assessment (summarised in Figs. 7 and 8) were that:

- Total biomass and spawning potential remained relatively stable, with fluctuations, until the mid-2000s, after which it declined. Estimated recruitment showed an increasing trend from 1980 to the recent period.
- Average fishing mortality rates for juvenile and adult age-classes increase throughout the period of the assessment.
- The 8-region model structure provided slightly more optimistic estimates of stock status when compared to the 5-region model structure. In both cases, the stock was assessed to be above the adopted LRP, and fished at rates below F_{MSY}, with 100% probability. It was concluded the skipjack stock was not overfished, nor subject to overfishing.
- Overall median depletion over the recent period (2015-2018; $SB_{recent}/SB_{F=0}$) was 0.44 (80 percentile range 0.36-0.52) for the 8-region model, and 0.40 (80 percentile range 0.30-0.50) for the 5-region model.

- Results from both regional structures indicate a stock status currently on average below the interim TRP for skipjack at that time.
    - For the 8-region grid, 47 of the 54 models (85%) estimated $SB_{recent}/SB_{F=0}$ to be less than the TRP (50% $SB_{F=0}$).
    - For the 5-region grid, 48 of the 54 models (87%) estimated $SB_{recent}/SB_{F=0}$ to be less than the TRP (50% $SB_{F=0}$).
- Recent median fishing mortality (2014-2017; F_{recent}/F_{MSY}) was 0.44 (80 percentile range 0.34-0.61) for the 8-region model, and 0.48 (80 percentile range 0.35-0.66) for the 5 region model.

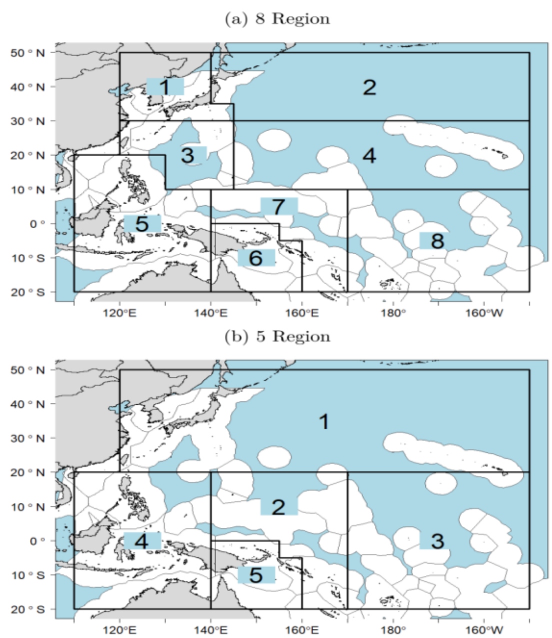

Figure 6. a) 8-region and b) 5-region model structures used in the 2019 WCPFC-CA skipjack stock assessment.

As noted above the SC15 favoured the 8-region model structure. For the provision of management advice, they also chose to down weight (x 0.8) some aspects of the uncertainty grid, including the higher and lower steepness values, and the length composition scalar of 50. The down weighting of these grid components had very minor influence (Fig. 7d).

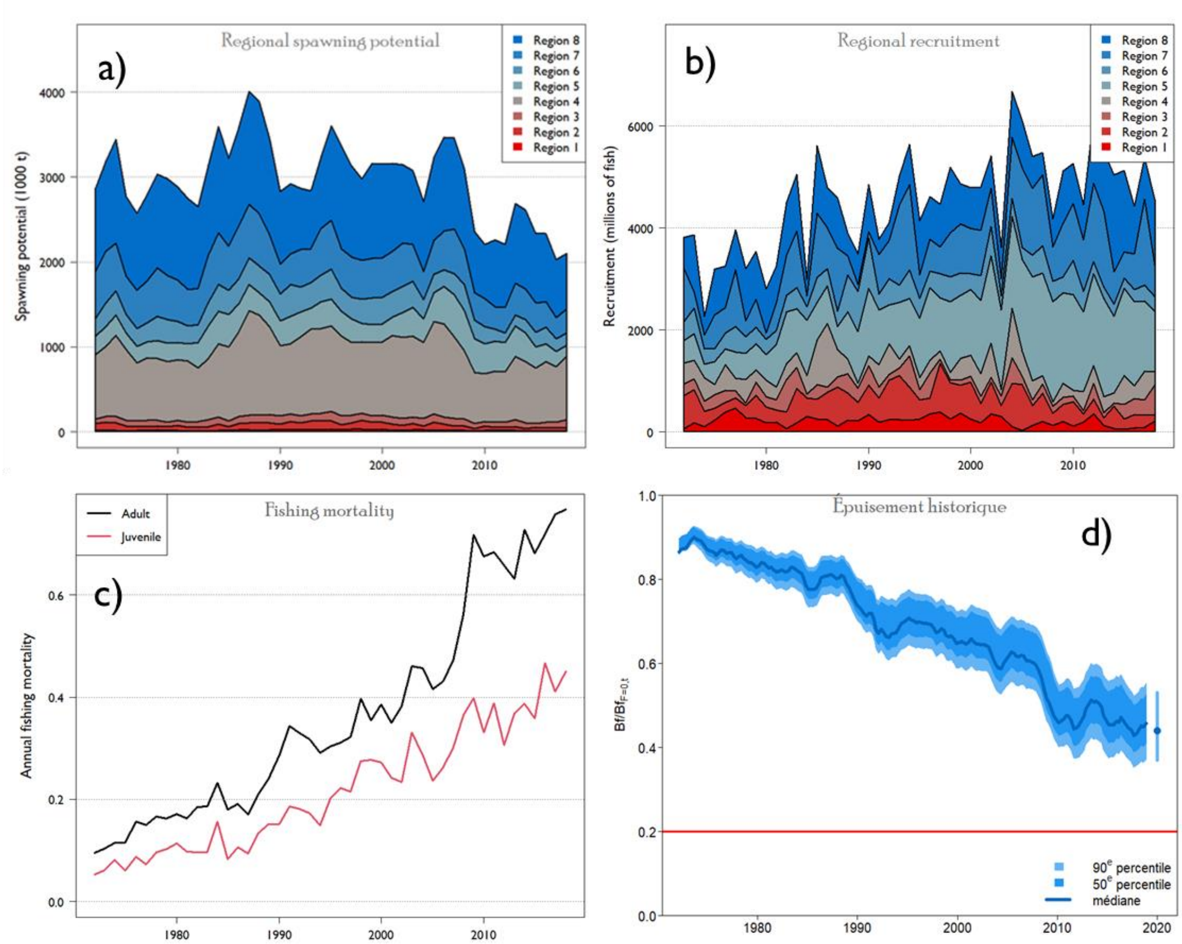

Figure 7. Estimated spawning potential time series by model region (a), recruitment by model region (b), fishing mortality for adults and juveniles from the skipjack diagnostic case model (c), and (d) spawning potential depletion across the uncertainty grid of 54 models (8-region structure), with the SC15 weighted median value illustrated by the large blue point.

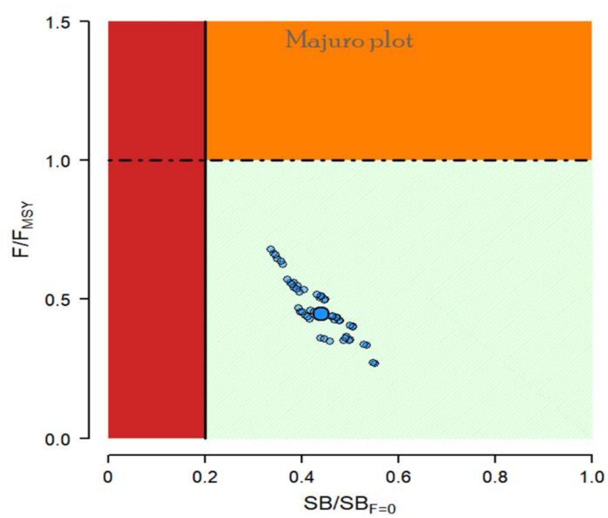

Figure 8. Majuro plot of the skipjack stock status for the 2019 assessment (last data year 2018) displayed as the end points from the uncertainty grid of 54 models (8-region structure).

2022 assessment planning summary

Spatial structure

An early step in the development of the assessment is to consider the spatial structure. This provides the basis for stratification of much of the data inputs and the assessment outcomes used for management advice. SC15 recommended that the 8-region model is preferrable over the 5-region model. As there is no new information on population structure of skipjack that could provide a strong basis for changing the 8-region model, to our knowledge, we propose to continue to use the 8-region model (Fig. 6a) for the 2022 diagnostic model. If PAW participants have strong concerns about the 8-region structure, they should raise these, and their reasoning for any alternatives. Sensitivity models on alternative structures can be attempted if there is a good reasoning for this, noting that running alternative spatial structures that do not simply involve combining (collapsing) regions of the 8-region structure involves much more work and may not be possible in the time frame available.

Fishery structure

Applying the 8-region structure will mean that changes to the general fisheries structure from 2019 are not necessary. The previous 8-region model had 31 fisheries across six gears/methods; pole and line (PL), long line (LL), purse seine 'associated' (SA), purse seine 'unassociated' (SU), purse seine not specified (S) and miscellaneous/artisanal gears ('other' gear) (i.e., small vessels in Philippines, Indonesia and Vietnam).

For the 2022 assessment we plan to apply an alternative approach to estimating fishing mortality, referred to as the 'catch-conditioned method'. The 'catch-conditioned' approach was initially considered for development in 2007 (Hampton et al. 2007) and is now available in the most recent version of MFCL (Vers. 2.0.8.4). Previous assessments have applied what is referred to as the 'catch-errors method'. The details of how these alternative methods differ will be presented in the PAW, session 4, day 1.

The catch-conditioned approach results in a large reduction in the numbers of parameters estimated. It assumes the catch data are correct (with no observation error) and solves the Baranov catch equation for fishing mortality. This method does not require estimation of the (often thousands of) additional parameters required for the catch-errors method and is thus computationally more efficient and less complex. The catch-conditioned method however adds a likelihood component for 'survey' indices of abundance that can provide information on relative abundance across space and time. This is particularly useful if the assumption of constant catchability can be made across fisheries operating in different model regions at the same time (something to discuss in the PAW). In WCPFC tuna assessments the indices of abundance are typically based on standardised fishery dependent CPUE time series. As such the catch- conditioned model requires specification of the additional 'survey' fisheries to provide the abundance indices.

For this assessment we propose to explore the application of the following 'survey' fisheries:

- The Japanese DW (distant water) and OS (offshore) pole and line fishery (regions 1- 8) (*using the VAST method, similar to previous assessment*)
- The equatorial purse seine fishery (regions 6, 7 and 8) (*previous assessment only included region 6 with GLMM, this assessment aims to apply a VAST model*)
- The Philippines purse seine fishery (region 5) (*as per previous assessment*)

The development of these indices will be discussed in session 3 of day 1. There are various ways in which these 'survey' fisheries could be used in the assessment depending on assumptions of shared catchability and how reliable they are thought to be as indicators of skipjack relative abundance trends across time and space. There are also some concerns around hyperstability of fishery dependent CPUE indices from skipjack fisheries that target schools and lack detailed information on effort related variables. These will be important topics to discuss in the PAW.

Model development

To develop the 2022 diagnostic model, we will follow a series of steps, often referred to as a 'stepwise' model development or 'bridging' analysis. The first step in this process is to run the 2019 diagnostic model with the most recent version of MFCL (2.0.8.4), and no other changes. This has been completed and has confirmed virtually identical results compared to the 2019 diagnostic model. The next step is to run the 2019 MFCL updated diagnostic model with a catch-conditioned approach. This has been the focus of recent work and will be discussed at the PAW. Once we are satisfied with a catch-conditioned version of the 2019 diagnostic model, we will conduct a series of one-off sensitivity tests to explore sensitivities to changes being considered for the new assessment, including data inputs, biological and structural uncertainties etc. The outcomes of these one-off sensitivities will be used to inform the structure of the final stepwise model development to establish the new diagnostic model. These one off-sensitivities will be conducted on the catch-conditioned 2019 diagnostic model, allowing us time to explore sensitivities earlier in the assessment time schedule, before the final data updates for the 2021 terminal year are completed in May. The PAW is an important forum for considering sensitivities to explore and discuss proposed changes from the 2019 to 2022 assessment. The final stepwise model development phase will occur when the new data inputs are finalized. Hopefully this will lead to fewer last-minute surprises.

Uncertainty treatment

Treatment of uncertainty in the assessment outcomes used for management advice is an important consideration of any assessment. Previous skipjack assessments have characterised uncertainty using a factorial grid approach (i.e., Table 1), that considers combinations of various components of structural (or model) uncertainty. These components and their respective levels/values in the uncertainty grid have been an important topic of discussion at the PAW, with the goal of developing a grid that incorporates the most important structural uncertainties, while limiting the total number of models required to be run. This group of models is used for formulating management advice, providing a range of potentially plausible states of the stock and fishing impacts against which managers can consider risk.

The model grid used in the 2019 assessment is shown in Table 1. The most influential uncertainties in the 2019 assessment were the mixing period, growth and steepness assumptions. The steepness assumptions applied are standard options applied to tropical tuna assessments and we propose to continue with the range of steepness values applied in 2019. PAW attendees may be aware of new work that could be important in refining the choice/range of skipjack tuna steepness assumptions, if so, please raise this.

Between each assessment, research should be ongoing to reduce uncertainties. In the case of skipjack, the choice of the growth relationship remains problematic due to the difficulty of obtaining accurate age estimates from hard parts such as otoliths. We are not aware of any new growth studies in the western and central Pacific which could be applied to the 2022 assessment. In terms of growth estimation, we plan to further explore the application of MFCL to estimate growth from length composition modal structures (to be discussed at the PAW) by initially placing high weight on the length composition data, as was done in the previous assessment, and applying the resulting estimated growth form externally to the final models, where the size composition receives lower weight. This may include attempting shorter time steps and selecting data that have clearer modal structures. We also plan to analyse tag-recapture growth increment data as a potential source of growth information that may lend support to the estimates derived from MFCL.

Natural mortality (M) at age was not included as an axis in the previous assessment but was estimated internally, meaning that for each model of the uncertainty grid a different M at age was estimated. Rather than allowing M at age to vary in relation to changes in other assumptions, an alternative is to determine M at age from an external analysis and apply these values as fixed inputs. We propose to explore this option for the 2022 assessment with a similar approach as applied to the 2021 South Pacific albacore assessment. The approach involves estimating growth models using MFCL (similar to the previous assessment), and then apply these growth parameters to estimate M at age using an approach that is based on that developed by Maunder and Wong (2009) and Maunder et al. (2011). This results in a combined growth and M at age axis for the uncertainty grid (see Castillo Jordan et al. (2021) appendix 1).

The most influential uncertainty in the 2019 assessment was the tag mixing period. While the general approaches to preparing the tagging data and incorporating tagging mortality, tag shedding, tagger effects and priors/penalties for tag reporting rates will not differ notably from the 2019 assessment, tag mixing has received targeted research focus. To include tag recaptures in the model estimation, it is assumed that tagged fish are fully mixed with the untagged population and hence are exposed to similar levels of fishing pressure as the untagged population. The time period that is required for this mixing to be complete is uncertain and is likely to vary among different tag release events. Previous assessments have simply applied assumptions of either one or two quarters for the mixing period to all release events. Recent work that will be presented in session 3, day 1, has developed a new tag mixing simulation approach. This approach incorporates environmental forcing on dispersal/movement of individual tagged fish (i.e., modelled particles) and those of untagged fished to estimate when simulated tagged fish have mixed to extent that they experience similar probabilities of being captured as untagged fish. It is possible that this simulation modelling can be applied to estimate tag mixing periods specific to each release event, rather than fixing the mixing period at one or two quarters for all release events. It is too early to decide whether this alternative to specifying tag mixing periods should be included in the assessment, either as the diagnostic case or as an alternative tag mixing axis in the uncertainty grid. This will be a key topic for discussion at the PAW.

Table 1. Structural uncertainty grid used in the 2019 WCPFC-CA skipjack assessment

Uncertainty	Option 1	Option 2	Option 3
Tag Mixing Period	1 quarter	2 quarter	
Length Scalar	50	**100**	200
Growth Model	Low	**Diagnostic**	High
Steepness (H)	0.65	**0.8**	0.95

Previous iterations of the skipjack assessment have not included parameter estimation uncertainty in the characterisation of uncertainty. Further, applying all of the grid axis/level combinations for management advice can result in the inclusion of combinations that are unlikely in reality, for example, low growth and high steepness, high growth and low M etc. The inclusion of grid combinations that have low plausibility can result in artificially expanding the range of uncertainty in management advice. In the past certain combinations of grid axis/levels have been down weighted for management advice during the SC meetings. This has been done in a fairly subjective manner, which is not desirable. Recent work in the area of ensemble model selection/filtering (i.e., 2021 Southwest Pacific swordfish assessment, Ducharme Barth et al. (2021) *) and objective weighting approaches for model ensembles using various diagnostics provide some good direction for how to more appropriately characterise uncertainty in management quantities. This subject will also be discussed in relation to the follow-up work on model selection/weighting for the 2021 Southwest Pacific blue shark assessment in PAW session 7 of day 2. This is an active area of research and no general guidelines on approaches have been developed for WCPFC assessments. Discussions of approaches to characterising uncertainty that are feasible to apply to skipjack in the available time frame (models take approximately 7 hours to run) or as focus for further development are welcome.*

References that are not hyperlinked:

Maunder, M.N., Wong, R.A., 2011. Approaches for estimating natural mortality: Application to summer flounder (Paralichthys dentatus) in the U.S. mid-Atlantic. Fish. Res. 111, 92–99.

Maunder, M.N., Aires-da-Silva, A. Deriso, R.B., Schaefer, K. and Fuller, D., 2009. Preliminary estimation of age- and sex-specific natural mortality for bigeye tuna in the eastern Pacific Ocean by applying cohort analysis with auxiliary information to tagging data. Inter-Amer. Trop. Tuna Comm., Stock Assessment Report, 10: 253-278.