SCIENTIFIC COMMITTEE
TWENTY-FIRST REGULAR SESSION

Nuku'alofa, Tonga
13–21August 2025

SUMMARY OF SC21 ONLINE DISCUSSION FORUM

WCPFC-SC21-2025/ODF-01

INTRODUCTION

1. The twenty-first meeting of the Scientific Committee (SC21) was held in Nuku'alofa, Tonga. SC21 made use of an online discussion forum (ODF) to facilitate consideration of discussions on 2025 SC projects and other items. The ODF was closed on 17 August 2025, during SC21, to allow the outcomes of the ODF discussions to be considered by CCMs at SC21.

2. For reference during the SC21 Work Programme and Budget discussions, the table below summarizes the input provided by SC participants on ODF Topics related to WCPFC projects. The full comments are presented in this paper under each Topic.

SUMMARY OF INPUT FROM SC21 ON SC PROJECTS

Topic #	Subject	Comments
1	**Project 127:** Review and reconciliation of size data collected in the WCPFC-CA for stock assessment purposes	**Chinese Taipei:** correction to Table 12
2	**Project 114:** Progress in improving coverage of cannery receipts data for WCPFC scientific work	**ISSF:** support for SPC's efforts to obtain a more complete set of cannery receipts and for identifying networks to improve tag reporting
5	**Project 123:** Scoping the next generation of tuna stock assessment	**USA:** supportive of continued work on Project 123, with high priority on continuing collaborations with DTU on tagging data analysis, and collaboration with IATTC on development of a new modelling platform. In view of SPC's workload, support outsourcing the work from SPC when developing the proposal for phase 2.
10	**Project 124:** Stock Assessment of Oceanic Whitetip Shark in the WCPO	**Japan:** appreciated work to improve the stock assessment of this regulated species, noting it is remarkable that the positive effect of management measures was evaluated and confirmed quantitatively. Additional detailed technical questions.
21	**Project 121:** Ecosystem and climate indicators	**USA**: agree that this is a reasonable collection of indicators. Suggest adding the estimated median phytoplankton size as either a fishery or ecosystem indicator. Requested (and received in comments) update from SPC on the anticipated next steps towards implementation.

TOPIC 1. Project 127 – Review and reconciliation of size data collected in the WCPFC-CA for stock assessment purposes

Relevant Documents: SC21-ST-WP-02. P. Hamer, E. Schneiter, T. Vidal, P. Williams. Review and reconciliation of size data collected in the WCPFC-CA for stock assessment purposes (WCPFC Project: 127)

Questions and Comments

Chinese Taipei

Just a correction on weight data for ALB (Table12, page41). "TWWT: …Weights are processed weights…" should be changed to " Weights are whole weights in both fresh and frozen forms".

Reply

SPC: Many thanks for the correction, Ren-Fen, I will correct it in the paper.

TOPIC 2. Progress in improving coverage of cannery receipts data for WCPFC scientific work (Project 114)

Relevant Documents: SC21-ST-WP-04 T. Vidal, S. Gislard, J. Scutt Phillips, and T. Peatman. Project 114 Update: Progress in improving Cannery Receipt Data for WCPFC scientific work

Questions and Comments

ISSF

It is good to see that SPC continues to endeavor to obtain a more complete set of cannery receipts, which could be beneficial to stock assessments. The idea of identifying networks to improve tag reporting also seems novel and worthwhile.

TOPIC 3. Project 90 update 2025 – Better data on fish weights and lengths for scientific analyses

No questions or comments

TOPIC 4. Project 109 – Training observers for elasmobranch biological sampling

No questions or comments

TOPIC 5. Scoping the next generation of tuna stock assessment software (Project 123) Update

Relevant documents: SC21-SA-WP-01. A. Magnusson, N. Davies, G. Pilling, and P. Hamer. Project 123: Scoping the next generation of tuna stock assessment software

Questions and Comments

USA

The USA is supportive of continued work on Project 123. We see real value in having interested members and the SC continuing to have the opportunity to guide prioritization and scope of the project. We look forward to working with other CCMs during the small group meeting to further this project.

In the meantime we place a high priority on continuing collaborations with DTU on tagging data analysis either as an input to a new modeling approach, or as a dedicated spatiotemporal modelling platform. We also prioritize further collaboration with the IATTC on the development of a new modelling platform. Lastly, while we consider the development of a single-region MFCL model to be an interesting exercise, we emphasize the importance of simulation testing modelling approaches relative to a known "truth" rather than only comparing to alternative modelling approaches fit to the same data. In the absence of a suitable operating/simulation model, the IOTC YFT simulated data from the spatial simulation workshop (Goethel et al., 2024; https://onlinelibrary.wiley.com/doi/10.1111/faf.12819) could be used to evaluate performance.

We place a low-priority on developing alternative modelling approaches (e.g., GADGET) outside of a simulation framework.

As the SC considers the next phase of Project 123 which ends next year at SC22, we expect that model development (including simulation testing) will likely entail a substantial workload. We also recognize the ever-increasing workload the SSP is asked to complete each year in support of existing Commission obligations. Therefore, we would like to suggest that when developing the proposal for phase 2, it includes plans to outsource the work from the SSP.

TOPIC 6. Skipjack Tuna Stock assessment

Relevant Documents: SC21-SA-WP-02 T Teears. WCPO skipjack tuna stock assessment

Questions and Comments

New Zealand

We acknowledge the substantial work undertaken in preparing the 2025 SKJ assessment, and the significant improvements, particularly the removal of Skipjack Survey and Assessment Programme (SSAP) tagging data and inclusion of pole and line effort creep corrections. However, several technical concerns remain that could affect the reliability of stock status estimates. New Zealand recommends addressing these issues before the next assessment.

1. Recruitment trends (Comment)

SC18 noted that the steadily increasing recruitment in the SKJ assessment may be a model artefact. In response, the 2025 Pre-assessment Workshop (PAW) suggested that three modelling changes might address the recruitment trend in the 2022 assessment: removing conflict between tagging programs, removing some anomalous size data, and including pole and line effort creep. The SSP has integrated some of these changes into the 2025 SKJ assessment, including removing the SSAP tagging data, which we consider has improved model realism.

2. Model scaling from tag data (Question)

The SKJ assessment is built around data from Tagging Programmes (TPs). The tagging data, and the assumptions that go into modelling tags, strongly determine the biomass estimated during each TP period. The effect of tagging data is evident in the model development series (figure 26): the biomass trend from the 1970s to 1990 is fixed until the SSAP is removed. Once the SSAP data is removed the early biomass trend becomes more responsive, such as when pole and line effort creep is included.

However, the biomass trend after 1990 stays largely the same, even when effort creep is added (figure 26). This may be due to the constraining effects of the Regional Tuna Tagging Project (RTTP) in the early 1990s, and subsequently the Pacific Tuna Tagging Programme (PTTP) - from 2005 to the present.

The PAW recommended that the model should be run with one TP at a time to test the effect on biomass trends (p. 37 of SC21-SA-IP-01), because of the conflicts between all the TPs. These scenarios have not been reported in the current assessment document. Could the SSP please comment on the results of these runs? If they were not conducted, we recommend including them in future to reduce internal conflict, and to help interpret each programme's influence on biomass scaling.

> **2. Reply (SPC)**
>
> Runs excluding, in turn, the RTTP, PTTP and JPTP tagging programs were not undertaken due to time constraints. There is however a run reported in the Diagnostics section of the report where all tagging data are excluded (Fig. 61). When time permits, the three additional runs, excluding the TP's one at a time, can be undertaken and the results shared.

3. Apparent conflict with CPUE data (Question)
The pole and line CPUE indices have a declining residual trend from 1990 to 2005 (figure 46), indicating that the RTTP and PTTP data conflict with the CPUE indices. The CPUE indicates that biomass declined from 1990 to 2005, but the TPs indicate stable biomass. The increasing recruitment trend from 1990 to 2005 (figure 34) is another diagnostic indicator of this conflict. The declining CPUE is consistent with the fact that the catch approximately doubles from 1990 to 2010 (figure 3), while the TP data indicate stable biomass across this period. Could the SSP please comment on this apparent conflict indicated by the model diagnostics?

3. Reply (SPC)
The scale of the CPUE residuals is generally very small for those fisheries showing a slight decline over this period. Overall, we feel that the residuals pattern is as good as one could expect in a complex assessment integrating several data types. Regarding the increase in recruitment from 1990, we agree that this could be a result of the model accommodating increasing catch over this period, as well as fitting simultaneously the tagging and CPUE data. However, overall scaling seems reasonably consistently informed by the CPUE and tagging data (Fig. 62).

4. Tag mixing and Ikamoana (Comment)
It is well accepted that tuna tags mix slowly, and that unmixed tags cause bias in model parameter estimates. Ikamoana has been used to identify release-specific mixing periods, by simulating mixing between tag releases and untagged fish. This approach appears preferable to assuming uniform mixing. However, the simplifying assumption that there is no entrainment by coastlines and FADs may affect the mixing estimates from Ikamoana. We suggest providing diagnostics to compare Ikamoana's predicted levels of mixing with observed levels of mixing, such as via aggregated tag density plots.

More importantly, plots of tag densities for the tags included in the stock assessment are an important diagnostic to determine how much bias may be caused by unmixed tags; plots of this type were provided for YFT and BET in WCPFC-SC19-2023/SA-WP-03 (figs 48-49, Appendix 1). However, more aggregated plots can better monitor for the density trends characteristic of poor mixing, such as plots with densities combined across time and release groups (e.g. WCPFC-SC9-2013/SA-IP-06, figure 37).

These would be useful standard diagnostics for a stock assessment that assumes tags are mixed and would help members to evaluate whether mixing assumptions could affect scaling.

4. Reply (SPC)
The phenomena of poor mixing in tropical tuna tagging experiments and it's potential to bias demographic parameters has been well discussed. The use of the most recent SEAPODYM solution to simulate tag releases in Ikamoana is an attempt to characterise the potential variability for different release groups. As SEAPODYM includes tagging data in its estimation of movement parameters, in essence this approach leverages information from many tags over time to permit this characterisation, even for tag release groups with very few releases and/or recaptures.

Entrainment of fish around islands or other bathymetric features is considered to the degree that such features influence the oceanographic forcing at the scale that they are included in SEAPODYM (1 degree). The effect of attraction/entrainment by FADs, vessel level behaviour, or any other local scale dynamics hypothesised to influence mixing, is not included in the simulation. Indeed, we are unaware of any model of such dynamics that is quantitatively estimated on data.

In response to discussion at the 2025 Pre-assessment Workshop, a comparison of observed recapture rates and simulated recapture probability distributions was made, and discussed in

SC21-SA-IP-3 (table 3.). Tag recapture density plots can be calculated as was done for the 2023 yellowfin and bigeye, although aside from a qualitative and visual examination of uniform tag recapture over time, it is unclear how this would support selection of an appropriate mixing period for stock assessment.

5. Fits to size data (Comment)

Models that downweight tagging data or apply stricter tag mixing criteria suggest substantially higher biomass, based on the influence of size data (see figures 61, 62). However, size data are difficult to fit well in the skipjack model (figures 48-55), and they may be receiving too much statistical weight. Multiple factors can contribute to poor size data fits: skipjack growth is poorly known and assumed to be uniform throughout the WCPO, and all fisheries assume constant selectivity (even though realized selectivity tends to change seasonally and through time).

Size data that fit poorly may not provide reliable information about biomass scale, as seen with the 2024 and 2025 striped marlin assessments. We note the benefits of the work to improve size data fits in the swordfish assessment and we appreciate the efforts already made to improve the skipjack size data. However, we believe this issue is very challenging and will require additional attention for the next skipjack assessment.

5. Reply (SPC)

The fits to the aggregate size data are easily the best ever seen in a WCPO skipjack assessment. Some fisheries that have small observed sample sizes are naturally not fitted as well, but the fits to the large-catch fisheries in particular – tropical purse seine and the domestic fisheries in Indonesia and Philippines – are fitted quite well in aggregate. We note that the data weight of the size data is estimated in the model and the extent of estimated down-weighting is shown in Fig. 21. We prefer this approach to arbitrary down-weighting.

6. Management parameters (Question)

After increasing linearly until 2010, catch has varied between 1.5 and 2 million tonnes for 15 years, as though it has reached a plateau. Nevertheless, the model estimates biomass to be between 2 to 5 times BMSY, and the Fmultiplier is estimated to be between 2 and 5, which suggests the potential to substantially increase the exploitation rate. There are various possible explanations for this, but it is an inconsistency that needs to be addressed. Could the SSP please comment on this?

6. Reply (SPC)

We note that WCPFC (and PNA) effort limits on the purse seine fishery have been in place since 2010. This, related access restrictions, and economic factors in the purse seine fishery, will have been a major factor in the 'plateauing' of the catch. We see no particular inconsistency in the management quantities, with the median fmult = 2.85, Frecent/Fmsy (1/fmult) = 0.35, and median SBrecent/SBmsy = 3.9. Aside from the constraints mentioned above, there is technically some 'potential' to increase the exploitation rate (if this refers to fishing mortality or related metric). The estimated MSY is about 20% higher than the 2024 catch, but since we are reaching the dome-shaped top of the yield curve, it would require an estimated 2.85 times the recent total effort across all components of the fishery to achieve that. This does not make a lot of sense from a fisheries management perspective. And note that the level of SB at MSY is estimated to be 14% of the unexploited SB, which is less than WCPFC's LRP.

7. Model diagnostics (Comment)

The three main sources of scaling information are tagging data, size data, and the relationship between catch and CPUE, but all of the diagnostic alternative models include either tagging or size data. We would be interested to see a standard ASPM model, with CPUE and catch data but no tagging or size data, as an

additional diagnostic for this assessment. A BSPM model would also be a useful sensitivity / diagnostic approach to consider in future assessments.

7. Reply (SPC)

We can quickly do this and include in the presentation. Presumably you would like the estimated movement parameters to be fixed, as are the selectivities.

In respect of NZ point 7, we re-ran the ASPM removing the tagging data in addition to removing the size data. In so doing, we fixed the movement coefficients at their values estimated in the diagnostic case. As for the original ASPM, we did two runs, one excluding the estimation of recruitment variability and one including it. Below is the equivalent plot to Fig. 59 in the SKJ assessment report.

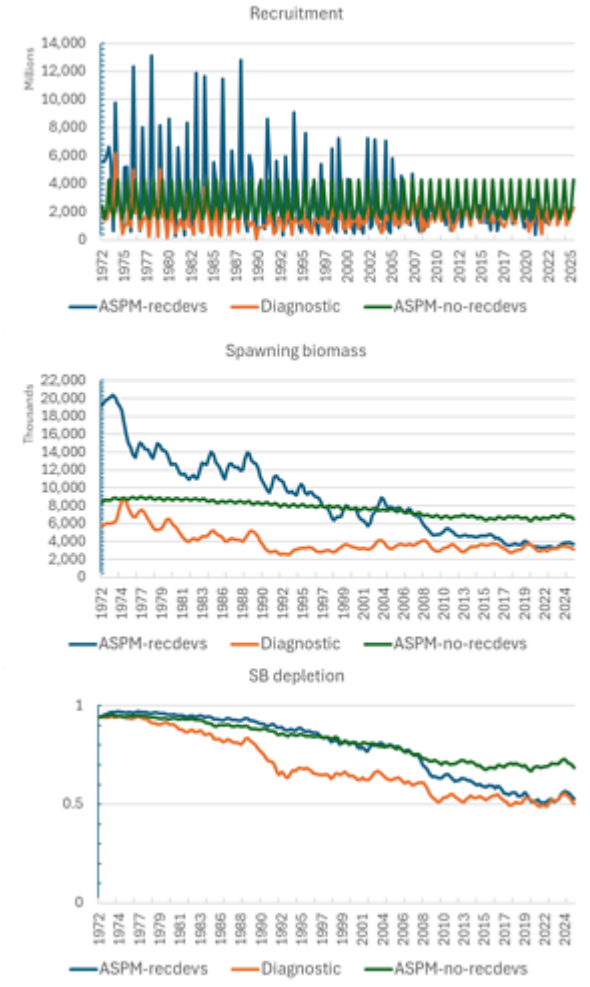

image907×1549 126 KB

We see that when the tagging data are omitted, the recruitment and SB are up-scaled considerably, and SB depletion is more modest for most of the time series, particularly when recruitment variation is disabled. When recruitment variation is allowed, the most recent estimates of depletion converge with the diagnostic case, but there is considerable divergence in the earlier estimates.

8. Review process (Comment)
We recognise that the stock assessment team are under huge pressure to deliver on time and that this year there were issues with data preparation which caused delays. That said, this year the members received this very complex SKJ assessment document with one week to review it, which is simply not enough time for a full review. This is further complicated by the fact that we only have reports, not the assessment input and output files which tend to be published well after SC is finished. [We note the SWO assessment files have been provided in a GitHub repository, which is very helpful.]

Given the importance of the skipjack stock to the Pacific region, we think the current review process is not working well, and we would like to suggest a full external peer review of this assessment in the period leading up to the next assessment. This could follow a similar approach to the 2024 review of the northwest Pacific striped marlin stock assessment. Terms of Reference for the review could be determined by SC22.

> **8. Reply (SPC)**
> All relevant files relating to the 2025 skipjack stock assessment will be placed on a GitHub repository as soon as possible. SPC has a long history of making assessment input and output files available for interested parties to download. But perhaps WCPFC should establish a formal policy for such transparency, noting that this is not routinely provided for stock assessments conducted for WCPFC by other organisations, or by most other tRFMO's.
>
> While we would welcome a peer review for the WCPO skipjack assessment, SPC's experience with the sort of peer review being suggested is that it requires a time commitment from the SPC stock assessment and data team approximately equivalent to undertaking a full assessment. If such a review is to be conducted during the next 3 years, it would likely require modification of the stock assessment schedule.

USA
The US would like to recognize the efforts made by the SSP in addressing many of the concerns raised by SC18. We consider this to be an improved assessment in many aspects. However, despite these positive steps forward we do have a number of concerns, questions and/or technical comments which we will raise here in the ODF.

The key area for discussion with this assessment is the high conflict in the different data sources as evidenced by the likelihood profile, jitter analysis, and leave data out analyses (ASPM, CCA, and no tag model). It is worth noting that the expanded diagnostics provided by the assessment team facilitate this identification, and we are grateful for their provision. The overall assessment result is one that appears to split the difference between these components and as a result doesn't fit any of them as well as desired. As noted by New Zealand, there is evidence of mis-fit to certain data components: indices, mean length over time, and JPTP tagging data. The mis-fit to data components (particularly the large length observations for many fisheries in recent years) and the impact on the likelihood could limit confidence that the model accurately captures the population dynamics.

Given this concern, and the importance of skipjack to the region and the world, we are supportive of New Zealand's suggestion of an external review of the skipjack assessment and for SC22 to help determine the Terms of Reference for that review.

There are also a number of additional comments/questions related to modeling choices, and estimated outputs that we wish to raise. Some issues are interrelated. These are briefly summarized below (in no particular order):

Model convergence (Comment)
This assessment is a step forward with respect to model convergence. Most ensemble models meet minimum convergence criteria (gradient and PDH). Though a jittering analysis was conducted only 30 jitters were conducted which may or may not be a large enough sample size to identify if the model has converged to a minimum. Furthermore, though no jittered model had a better total likelihood, there appeared to be very high variance in terms of minima with respect to fits to different data components. This tells the same story as the likelihood profiles, that the data are in conflict and that the total likelihood is achieved by trading off fits to the different data components. Given the apparent conflict and sensitivity to starting conditions, this does not appear to be a well-determined solution.

Reply (SPC)
Mostly agree. The final run of jittering was the 3rd or 4th such set of jitters. In the earlier runs, we found one or more somewhat better solutions (but with for all intents and purposes identical assessment results). We repeated the jitter on each occasion using the best model as the basis for the new jitter run. We agree that this sort of sensitivity is not ideal and there is a need to review model complexity and, to the extent possible, conduct additional data quality assessment.

Regarding data conflict, we note that overall the CPUE data provided similar information on overall scaling to the tagging data. The main conflict in fact is between the size data and the priors for tag reporting rates. This motivated us to test for the influence of the tag reporting priors by setting all of them to be uninformative. The results of this run are shown in Fig. 67, where we see the expected result that the population scaling is shifted upwards and the depletion ratio increased by several percentage points. This occurs because the reporting rate penalties are much reduced (but many of the estimates now go to their upper bound), allowing the size data to have more influence on scaling. The size data weight given by the Dirichlet Multinomial likelihood is estimated internally and the down-weighting is quite substantial, as indicated by Fig. 21. One indication of the effect of further down-weighting of the size data is provided in the ASPM that includes recruitment variability (Fig. 59). In this test, the size data are effectively down-weighted to zero, and the impact on scaling is positive, but relatively minor. That is a somewhat curious result, and we might have expected a negative impact on scaling. But of course, many parameters that were estimated in the full model (selectivity, growth, etc) were held fixed at the estimated values in the ASPM.

Model complexity (Comment)
The model reduced active parameter counts by switching to the OPR recruitment parametrization. While this is a step in the right direction, the model remains highly complex and flexible with quarterly movement, and quarterly regional recruitment parameters able to move biomass around as needed in order to reconcile clear conflicts in the data. This results in some interesting regional biomass and recruitment patterns. However, the current documentation and figures are not sufficient for determining if these regional patterns are reasonable (does movement follow tags or another data component; does biomass accumulate in lightly fished regions; does recruitment occur in temperate regions in quarters where it would not be expected). While management is done at the aggregate level, the details at the regional level are important for developing confidence in the model (is it consistent with empirical observations or other model outputs, e.g., SEAPODYM). Future assessments of skipjack (and other species) should re-evaluate the modelling approach used in order to "right-size" the assessment approach to match model complexity to data limitations and stock dynamics.

Reply (SPC)
We agree that more information on the regional allocation of recruitment and biomass could have been given in a more digestible format and we apologise for this oversight. We were generally

comfortable in how these regional estimates looked, and can offer the following additional results (from the diagnostic case) to support this:

The overall proportional allocation of recruitment by region was as follows (region 1-8 in sequence): 0.017 0.018 0.027 0.148 0.283 0.097 0.284 0.126

For SB, the regional distribution is shown in the following plot:

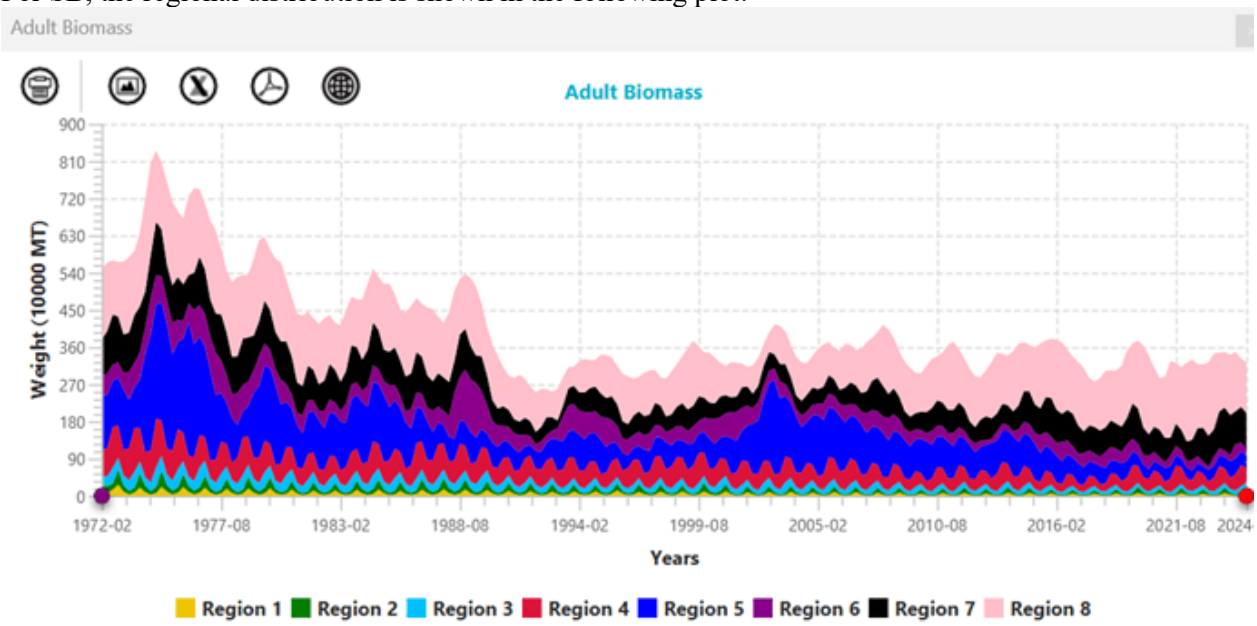

image940×470 169 KB

In both cases, the regional distributions obtained were not surprising, with the tropical regions dominating.

Model diagnostics (Comment)

While there does appear to be a greater focus on model diagnostics, some conventional diagnostics such as reporting of goodness of fit, length composition bubble plots, and hindcast cross-validation are lacking. Some figures are also difficult to interpret (for example the lack of observation error bars on CPUE fit makes it difficult to evaluate the quality of the fit, it is also very challenging to interpret the movement estimated by the model relative to what the tags indicate). The likelihood profiles by data component and fleet were useful for the index and length composition. It would be useful to do something similar for the tagging programs/release groups.

Reply (SPC)

Yes, there are many ways that additional graphics could contribute to interpretation of results – everyone has their favourites! We will take this comment on board for future assessments.

Model fit to data (Comment)

It is difficult to fully evaluate fits to the indices without also seeing the observation error associated with them (see the above comment). There does appear to be some trend in the PL fit residuals as pointed out by NZ. The PH PS index is best fit, unsurprisingly given the tight CV. However, the free-school PS indices do not appear to be well fit at all. The seasonal dynamics and long term dynamics are poorly captured. Future assessment should consider re-balancing the CVs used to fit the indices rather than using the raw CVs as output from the different standardization analyses.

Aggregate length composition data fits appear reasonable for many fisheries, however the fit to the trend over time is poor. In particular large observations are noted at the end of the time series for many fisheries, and these are poorly fit. Given the dome-shaped nature of selectivity, and the time invariance of selectivity, these observations likely have high leverage on model outcomes. As seen by the CCA the comp data indicates an increasing biomass trend, possibly driven by these large observations. In general, it is not preferred for composition data to drive either scale or trend estimates. It is recommended that composition data be scrutinized to exclude problematic observations, and that selectivity be made time-varying to better fit the size data if it represents a real shift and to reduce leverage on model outcomes.

Fits to the JPTP tagging data are poor. Fit to the PTTP and RTTP appear ok, except for RTTP region 8. It was recommended at the PAW to only fit one program at a time.

This is a sensible recommendation raised by New Zealand, and should be investigated. We note that SPC provided a response that said that models excluding one tagging program at a time were explored but this differs from the New Zealand request. The New Zealand request would help alleviate conflict between tagging programs and also with the indices.

> **Reply (SPC)**
> We note the comments on the PS CPUE. They are certainly not as well fit as some of the other indices, but we have seen many worse fits in other assessments. We would be interested in receiving more information in how to "re-balance" CVs in an objective way.
>
> We do not agree with the comment that the fit to the size data over time is generally poor. The time trends for LF data sets that have continuously high sample sizes over time (e.g. the tropical PS fisheries responsible for the vast majority of the catch) are generally reasonably well fitted it seems to us. Of course, some of the LF data sets for some fisheries are very sporadic, and these will be not as well fitted. But please point out which fisheries are causing concern.
>
> Regarding the lack of fit to the some of the larger fish sizes in the last year or two, we acknowledge this and point out that this is likely a result of constraints on recruitment variability that we applied at the end of the time series. We had to constrain recruitment in the last 3 years of the assessment to a common level, because earlier runs when this was not done produced enormous spikes in recruitment that we felt were unreasonable. Constraining recruitment in this way inevitably results in some lack of fit to the size data. We have heard in presentations today that very large catches of SKJ occurred in 2024 in both the WCPO and EPO. Therefore strong recruitment in the last year or two of the assessment may in fact have occurred; however we did not feel comfortable in relying on such huge estimates of recruitment in the last couple of years that are based on limited information (at this stage) and because of the way such estimates would impact short-term projections and potentially stock status indicators.

Potential hyperstability in CPUE (Comment)
While hyperstability was addressed for the PL, it remains a potential issue with the free school purse seine as well. This was raised at the PAW but not investigated or addressed in the report.

> **Reply (SPC)**
> We were less concerned with effort creep in the PS CPUE indices partly because of the way in which the data were filtered (free-school specialist vessels) and the use of VMS-based searching distance as the basis of the effort metric. When we thought about possibly imposing some effort creep scenarios on the PS CPUE indices, we realised that we would also have to consider negative effort creep due to the possibly of increasing dFAD density since 2010 negatively impacting the availability of SKJ in free schools. And in either the positive or negative effort creep scenario, we

would have to have made arbitrary assumptions, unlike the PL CPUE effort creep scenarios that were based on actual estimates by Japanese colleagues.

Biomass scaling (Question)
It is unclear from the model what is now driving the regional estimates of scale as catchability is no longer linked between index fisheries. Regional biomass scale is now no-longer constrained and can be whatever is needed to fit the data. Given concerns with data fits, and noisiness of some data streams, along with a lack of clear visuals it is difficult to understand if the estimated patterns are legitimate or if it is the model simply chasing noise. Could the authors comment on what is now informing regional estimates of scale now that index catchabilities are no longer shared?

We note that the regional scaling was an important uncertainty in previous assessments as the allocation of biomass to more lightly fished regions helped buffer overall stock status (e.g., 5 region model that was presented in 2019 and conducted as a sensitivity in 2022).

Reply (SPC)
CPUE indices were originally grouped (PS and PL indices separately) for catchability, but this resulted in some persistent lack of fit with some indices. It was clear that there was other information in the model regarding regional scaling – possibly from the tagging data – that was causing this. Additionally, the shared selectivity across indices within gear type was causing additional lack of fit to the index size data. Relaxing this grouping occurred in one of the later phases of the estimation. We were pleasantly surprised when the regional scaling of recruitment and biomass remained within expected bounds. With the considerable improvement in both CPUE and size data fits that resulted, we could see no reason to not retain the ungrouped index catchability and selectivity.

Movement estimation (Comment)
It is difficult to determine which data components (should be tags) drive the movement estimation and if this is reasonable. We request that a more useful plot be developed (or consider reverting to the plots shown in previous assessments) which more clearly show how model estimates of movement differ from what is indicated by the tagging observations.

Reply (SPC)
Well noted and agreed.

Regional recruitment estimation (Question)
Regional recruitment estimation is noted to be problematic on occasion, and the considerable flexibility of the model to estimate regional recruitment can lead to some counterintuitive results such as the very high recruitments predicted in quarter 1 for region 1 & 2, but no recruitment in quarter 2 for these regions. Based on biology and environmental preferences the reverse would be expected and we welcome comments from the authors if available.

Recruitment is predicted to decrease in regions 1-5, and increase in regions 7-8. At the regional scale there appears to be some correlation between recruitment and the evolution of catches. Regions with increasing catches appear to show increasing recruitment, regions with decreasing catches appear to show decreasing recruitment, and regions where catch is relatively stable show relatively stable regional recruitment estimates.

Reply (SPC)
The model will want to put recruitment in a region in a way that allows the seasonality of the catch in that region to be accurately predicted. In the case of R1 and R2, the majority of the catch

occurs in Q2. Looking at the size data, most of the catch is small fish in the 2nd quarterly age class. So the model places fish in the 1st quarterly age class (i.e. recruitment) in R1 and R2 in Q1 so that they are available at the right size for capture in Q2. Note that "recruitment" in R1 and R2 might not necessarily be based on spawning in those regions 1 quarter prior. There is the possibility (even likelihood) of transport of larvae in the Kuroshio Current that originates in the more tropical Philippines region for example.
While the use of the orthogonal polynomial recruitment formulation has simplified the recruitment parameterisation compared to the previous recruitment devs approach, there may be a case to simplify further, as the fully saturated year effect and season-region interaction effect still gives the model considerable (perhaps too much) flexibility to put fish where it needs to best fit the available data. We accept that further simplification of the recruitment parameterisation may be required, even if it is inevitably at the expense of poorer fits to data.

Mixing period assumptions (Comment)
The tagging data are unlikely to meet the mixing period assumptions for this large box model, we recommend more appropriately modelling the tagging data externally using the approach being developed in collaboration with DTU.

Reply (SPC)
We agree. Also, we think there is potential to further adapt the SEAPODYM model (e.g. to model tagging data in a release conditioned rather than recapture conditioned mode) for this purpose.

Growth /M estimation (Question)
We note the tradeoff identified in needing a high k to fit the tagging data given the Lorenzen mortality shape (needing high attrition of young ages), request comment from the authors if this is potentially indicative of model mis-specification and warrants a review of how the tagging data is used within the model (is the inclusion of potentially un-mixed tags causing problems?).

Reply (SPC)
This is a possibility than needs investigation. However we note that estimation of high (>0.4 per quarter) was also occurring for the tag mixing scenario (K=0.1) that specified longer mixing periods.

Initial conditions (Question)
The justification for initial depletion of 98% appears inconsistent from previous estimates where initial conditions were estimated (noting that those previous estimates could be wrong). Could the authors provide more justification about the initial condition assumptions and confirm that results were not sensitive to the initial condition assumptions?

Reply (SPC)
The 2022 assessment assumed an unexploited population that was unrealistic given the presence of significant catches prior to 1972. We made the change to assuming a small amount of fishing mortality in the initial equilibrium population to recognise this prior exploitation. The level of 0.02 x M to represent the initial fishing mortality was chosen so as to provide a constant small level of depletion over the first few years of the model. We noted that this specification impacted model estimates only over the first few years and subsequent results were completely insensitive to the setting. Also, note that the previous approach of setting F for the initial population to be at the average of the level estimated for the initial few years was done for the catch-errors version of the model. This approach is not technically feasible for the catch-conditioned approach due to the timing of the availability of the estimates.

Chinese Taipei

During plenary discussion Chinese Taipei asked if SPC could provide a ratio of the estimated spawning potential depletion for 2019 relative to the 2012 based on this year's diagnostic case model to enable comparison of the historical stock trajectory.

Reply (SPC)

In response to a request from Chinese Taipei during plenary, we calculated the median SBrecent/SBF0 and the median SB2012/SBF0 across the model ensemble and then calculated the ratio between these median estimates. The results indicated the median of SBrecent/SBF0 = 0.5113 and the median SB2012/SBF0 = 0.5115, and the resulting ratio of 0.999.

Response to SPC (Chinese Taipei)

Thanks Thom!

This makes it easier for my delegation to interpret stock status—using the ratio of recent depletion to the 2012 iTRP depletion level rather than an absolute value.

SBrecent(2015-2018)/SBF=0 / (SB2012/SBF0) = 1.04 (from 2019 assessment) (0.44/0.42)
SBrecent(2018-2021)/SBF=0 / (SB2012/SBF0) = 0.85 (from 2022 assessment) (0.51/0.60)
SBrecent(2021-2024)/SBF=0 / (SB2012/SBF0) = 0.99 (from 2025 assessment) (0.51/0.51)

Reply 2 (SPC)

Just to flag that this calculation provides a comparison of SBrecent/SBF=0 with the estimated depletion in 2012, estimated within the three recent skipjack stock assessments. But we note that the adopted iTRP for skipjack, as defined in CMM 2022-01, is calculated as the average of two spawning potential depletion values:

• the estimated average depletion of the skipjack tuna stock over the period 2018-2021 (SB2018-2021/SBF=0).
• the long-term median equilibrium stock depletion that would be reached under the agreed baseline fishing conditions for skipjack tuna (purse seine effort at 2012 levels, pole and line effort at average 2001-04 levels, and the domestic fisheries in assessment region 5 at average 2016-18 levels).

Further details of that calculation are provided in CMM 2022-01. We recalibrate this iTRP value using the 2025 assessment in Section 11.6.1 of SC21-SA-WP-02 (REV3), and provide the recent status of the stock relative to that recalibrated iTRP within the paper.

TOPIC 7. Southwest Pacific Swordfish Stock Assessment

Relevant Documents: SC21-SA-WP-05 J. Day. Stock assessment of Southwest Pacific swordfish

Questions and Comments

EU

Thank you SPC for the work. We are aware of the challenges and difficulties this kind of assessement may imply. We would like to make a request for its potential inclusion in the management advice section: As we all know, there is a significant proportion of SWO catches that take place as bycatch in tuna longline fisheries north of 20S, which are not currently covered by the CMM. This is a major challenge for the management of the stock without an easy solution. You provide in the assessment report some figures that illustrate this point, but we would like to kindly request the numbers of the estimated catches both north and

south of 20S (as an example, in the latest 5 years). Similarly, if you can produce a time series with the proportions, it can be very helpful at a later stage

Reply (SPC)

Thanks for the question - which is clearly important for management. The answer to this question requires some modelling, rather than simply summarising the input data, as catch is reported in a mixture of: numbers of fish caught (for most fisheries); and in t (for the EU fisheries). It makes no sense to compare coconuts with mangoes - so using the catch converted to a common unit (t) by Stock Synthesis, I have calculated the following proportions of the total catch (caught north of 20S), for the last 1, 5 and 10 years.

Period	% N	Catch N ('000 t)	Catch S ('000 t)
2023	0.333	1976.57	3953.38
2019-2023	0.379	10005.19	16411.36
2014-2023	0.441	28710.15	36396.16

A full time series could be produced - but we would need to know exactly what is required. Do you want a time series of annual proportions, or a time series of rolling averages (and over how many years), or of total catches, or catches (or proportions) split into .1N and.2N, either combined or separately, or some other combination. Is this something that will need to be regularly and consistently reported in future swordfish assessments?

Incidentally, all of this data is in the Stick Synthesis output file, Report.sso and can be found in the following section:

CATCH report:15
in the column labelled:
dead_bio
listed by fishery, and by quarter.

This file is publicly available at the swordfish 2025 GitHub repository: GitHub - PacificCommunity/ofp-sam-swo-2025-diagnostic: Swordfish 2025 diagnostic model

The fisheries north of 20S are labelled with either .1N or .2N on the end of the fishery name, or are fisheries numbered 1,4,8,10,15 and 18.

This raw data enables anyone to construct any time series they desire, either as catch (in t), or as a proportion of the total catch, or catch by fishery, by simply summing the appropriate figures. That may be more useful, as it allows any required option to be calculated by any interested party as required?

USA

The US would like to recognize the efforts made by the SSP in improving many aspects of the SWPO swordfish assessment. We consider this to be an improved assessment in many aspects. However, we do have a number of concerns and/or technical comments and suggestions which we will raise here in the ODF.

Fit to the CPUE

There appears to be a substantial lack of fit to the CPUE indices included in the model, with the Australian index slightly better than the New Zealand index. Both indices indicate patterns in the residuals where there

are persistent negative residuals through 2010, positive residuals from 2010-2017, and then negative residuals again after 2017. Fitting to CPUE indices should be prioritized over fitting to size data.

There is also a misfit in the CPUE indices in the last few years of the model where both observed CPUE time series are increasing after 2019 but the fitted index declines after 2020. This is most likely driven by the decrease in the mean weight of the Australian 1C fishery after a peak in 2015. As the likelihood profile indicates the CPUE indices are in conflict with the weight data, it is clear that this conflict needs to be better understood in future assessments. Exploration of the decrease in the mean size of the Australian fishery 1C should be fully explored to determine if it is a change in the fishery or a true indicator of a change in abundance.

Alternative CPUE indices

We would also like to suggest that future assessments consider exploring a CPUE index that uses the DW LL fleets such as Korea, Japan, and/or Chinese Taipei. These fleets, while often not targeting swordfish, do catch substantial amounts of swordfish and have a footprint that is much larger than the Australian or New Zealand fleets which may be more representative of the relative abundance of the entire stock. Work in the north Pacific has developed models to identify fleets based upon CPUE and size data, which has resulted in being able to identify fleets which provide information on recruitment, and fleets which catch adult swordfish and can be used as an index of abundance.

Treatment of Size Data

Generally, it is clear from the mis-fit to the CPUE data and the results of the ASPM that the size data are an important driver of the population trends. In fact, the fit to the CPUE in the ASPM model is a straight line, which indicates that it is providing no information to the population dynamics. Stock assessments should be driven by the indices of relative abundance not the size composition data. However, even within the size data from the same fleets operating in different areas (i.e. Australia 1C and 1N) indicate opposite estimates of population size. The sharp increase in mean size from 2015-2019 compared to the years before and after in the AUS 1C fleet is what is driving the decline in population in the last few years of the assessment, as it is attempting to fit the decline from 2016-2023 while the CPUE indices suggest an increase in population. It would be useful to understand why those data are so different than those before and after and if it is a change in fishery operations or operating area, time varying selectivity would likely reduce the influence of those data.

While we appreciate the removal of sparse or unreliable data, we would recommend a slightly different treatment of the selectivity estimates for those fleets fixed based upon an initial model run, as this is a very strong assumption and removes the ability of the model to adjust the selectivity of the model to fit the other data components. Aggregating the data into the superyears option in the size composition module of Stock Synthesis would allow for estimation of the selectivity parameters of fleets that have a reasonable aggregate composition pattern and allow the model to adjust the selectivity of those fleets based upon the other input data. For other fleets with highly uninformative size data, it might be useful to determine if any of the fleets fishing in the same area have a similar operation which would allow for mirroring of selectivity parameters. We also suggest including plots of the residuals of the size data by year in the report in the future, as it can indicate periods when the size data are not well fit and may need time-varying selectivity. This is especially useful for fleets that are not targeting swordfish and may change their operations over time as their species of target changes.

Sensitivity Runs

We also suggest that for the sensitivity runs including the alternative CPUE indices, a likelihood profile be run to evaluate how well the alternative indices agree with those selected in the diagnostic model.

We would like to note the very large catch of swordfish that occurs in the central tropical pacific just outside and along the boundary of the stock assessment. In previous assessment of swordfish stocks, there was a lot

of discussion with IATTC and ISC about which stock these catches belong, with the general agreement that these are most likely a mixed stock area where swordfish are feeding and that each stock would include some or all of these catch in their assessments as a sensitivity run. This was done in the 2023 NPO SWO assessment, the 2021 SEPO SWO assessment and the 2021 SWPO SWO assessment. We would recommend that this continue to be the practice for this stock until better information is available to appropriately allocate the catch in that area to the correct Pacific swordfish stock.

TOPIC 8. Revised 2024 stock assessment of striped marlin in the southwestern Pacific Ocean

No questions or comments

TOPIC 9. 2025 Stock Assessment of Striped Marlin in the Southwest Pacific Ocean

No questions or comments

TOPIC 10. Stock Assessment of Oceanic Whitetip Shark in the Western and Central Pacific Ocean 2025 (Project 124)

Relevant documents: SC21-SA-WP-08 P. Neubauer and K. Large. Stock Assessment of Oceanic Whitetip Shark in the Western and Central Pacific Ocean 2025 (Project 124)

Questions and Comments

Japan

We appreciate the work for the improvement of stock assessment of this regulated species. It is remarkable that the positive effect of management measures was evaluated and confirmed quantitatively also in this assessment. We have several questions as follows;

Question1: Treatment(filtering) of HBF

In the current assessment, re-estimation of catch by reconsidering the treatment of HBF with zero or missing data is significant progress. It is reasonable to remove data without HBF and zero data. Given that this process is influential, we have one question on the treatment of HBF. In Figure2, there are several large HBFs (such as >50) in some flags. Is it realistic gear configuration? Those high HBFs may be the number of floats (not HBF) and we would like to know your thoughts on the treatment of such large HBFs.

Reply (SPC)

HBF: While very high HBF numbers may not reflect real gear configurations, it is difficult to determine an effective cutoff point. However, at high HBF, the change in the effect is very flat (see Figure 1), so the exact number is probably less important unless the "true" number would have been low (e.g., <10). Given the small number of occurrences of high HBF in the dataset (compared with zeros or missing HBF), we suspect that any alternative treatment of high HBF records would have a very minor influence - however, this could be investigated in the future to ensure that this assumption is correct.

Response to SPC (Japan)

For Q1, it was well understood that the number of high HBF is small and no threshold is available. Misreporting of HBF and basket number is sometimes observed and high HBF (misreporting: correctly it is basket number) tends to be corrected to low HBF after the correction. This question was based on our experience of such observation.

Question 2: Growth parameter

In the previous stock assessment, both high growth by Seki et al. (1998) and slow growth by Joung et al. (2016) were used. Enhancement method of growth band pair, the period and area of sampling area and sample size are very different between studies, and the discrepancy of growth may not be attributed to regional difference in growth with the current assumption of stock structure.

In the current assessment, both estimates were considered in the informal prior for M, but their means were lower than M used in previous assessment. We would like to know the reason for this difference. In addition, given the several differences between studies, the slower growth was used in the diagnostic model in the stock synthesis. It would be appreciated if you could show the rationale of selecting the slower growth and whether the higher growth was taken into consideration in the model ensemble of stock synthesis.

Reply (SPC)
We maintained the two growth assumptions from the previous assessment, and maintained the same diagnostic case assumption for growth; largely for continuity reasons as opposed to reflecting plausibility. Both growth assumptions received equal weight in the SS ensemble, so the diagnostic choice is purely for illustrative purposes.
The M priors were derived based on available life-history based estimators; while the mean and median of the generated prior densities are lower than previous (fixed) M, this reflects a broader considerations of potential ways to estimate M than the previous assessment, which was based on a single estimate derived from a relatively ad-hoc approach to deriving the M estimates (see methods in Cortes (2002) - Incorporating Uncertainty into Demographic Modeling: Application to Shark Populations and Their Conservation). The actual estimate used in previous assessments was a relatively arbitrary choice within the range reported in Cortes (2002). Our updated priors do incorporate the range of survivorship estimates from Cortes (2002) within the 95% confidence bounds of the prior, and given the wide prior distribution, the prior mean should not be taken as a point estimate of M; the actual M in the model is estimated using this wide prior and informed by data (albeit with some conflict), as shown by the likelihood profiles - but it suggests a somewhat lower M than previously assumed (although the upper end of the posterior distribution is 0.17; which is close to the value of 0.18 previously assumed).

Response to SPC (Japan)
Q2. it was understood that two growth curve was treated equally and the text was just for the illustrative purpose. I also understood that your approach of estimating M with broad prior is more reasonable approach to consider the uncertainty of M rather than fixed and arbitral M. I appreciate your explanation.

Question 3: Treatment of the impact by regulation
The effect of introduction of management measures is well noted in this report in that the quality of CPUE and size data of OCS has been degraded. In addition, historical change of the leader material (if any) would be influential in the CPUE standardization. In this point, we would like to hear your thoughts on the treatment for this change of catchability or data availability in the stock assessment (such as the introduction of time block in the stock synthesis).

Reply (SPC)
We agree that there is a likely impact of regulation on CPUE - but this impact is difficult to treat. Ideally, these changes are standardised for in the CPUE standardisation; however, the observer data available for standardisation are relatively sparse already, and our experience is that operational parameters are inconsistently recorded and their inclusion usually leads to dropping large amounts of data, especially in early years.

An alternative, as noted in the question, would be to split the CPUE series or the catchability parameter in SS. We have not attempted this but agree that this would be a good sensitivity to explore in future shark models (not just for OCS); it has the potential to remove significant signal from the model, as much of the signal likely comes from the stock response to the non-retention measure (i.e., the rate of increase); splitting the time-series at the introduction of non-retention

measures could remove this signal - but it is worth exploring within the model to understand the impact in more detail.

Response to SPC (Japan)
Q3, I agree with the difficulty in treating the change of data quality, but I think it is very influential in the results. So it is future work to consider how we can remove artificial impact on the data from the data in regulated species, as well as continuation of data collection. As you pointed out in the presentation, there may be bias of size data sampling on board (e.g., smaller shark may tends to be hauled and measured rather than larger shark).

TOPIC 11. Progress towards a Close-Kin-Mark-Recapture application to South Pacific Albacore (Project 100c)

No questions or comments

TOPIC 12. Project 125: Billfish Biological Data Collection

No questions or comments

TOPIC 13. Project 126: Shark Biological Data Collection

No questions or comments

TOPIC 14. Project 128: Connectivity study on key tuna species in the WPEA region

No questions or comments

TOPIC 15. Projects 117 and 118: Biological sampling plans for tuna and billfish

No questions or comments

TOPIC 16. Project 120: Progress report on Reproductive Biology of WCPO Yellowfin Tuna

No questions or comments

TOPIC 17. Project 122: Progress report on Longline CPUE

No questions or comments

TOPIC 18. Climate Change Vulnerability Assessment (Consultancy report)

Relevant Documents: SC21-EB-WP-01 K. Robertson and M. Baird. WCPFC CMM Climate Change Vulnerability Assessment

Questions and Comments

Authors (K. Robertson and M. Baird)

Thank you very much for the opportunity to share our progress on the WCPFC CMM Climate Change vulnerability framework. It has certainly been an interesting exercise.

We have provided all deliverables to the Scientific Committee. It is a lot of information to work through. It includes an SC paper, as well as the CCVA excel workbook and .doc guidance document, as well as 5 CMM assessments, We will present on the main findings at SC on the Thursday session.

Our consultancy runs through to December so we are happy to take on your feedback. The key information we are looking for, consistent with our paper recommendations are:

1. Review and comment on the scientific approach and methodology of the draft framework
2. Provide input on the indicator system and data requirements

3. Identify potential data sources and quality considerations within the SC's expertise
4. Advise on integration with existing scientific processes and assessments
5. Recommend capacity building requirements for successful implementation
6. Support pilot implementation of the framework for 2025 CMM assessments.

We remain at your disposal for any questions or feedback.

FFA

Thank you very much Kerrie for the amazing work. The FFA has some technical comments re the assessment:

- Suggest an inverse color coding for Adaptive Capacity, considering low AC contributes to a high vulnerability, while high AC contributes to low V (inverse relationship that with Sensitivity),
- The framework considers the lack of information as a "high" score contributing to risk, and although the rationale behind this makes sense (as it contributes to the precautionary approach for decision making), it is not possible to see in the overall outcome whether a "high risk" is due to a majority of "unknown" score, or to actual high risk score. It would be useful to have those details shown in the final outcome, which would save the decision makers a lot of time having to scroll through the individual indicators to understand whether the CMM need more science data, or needs more management action,
- To ease the identification of which exact aspects contribute to higher grading each element of risk (exposure, sensitivity and adaptive capacity), and hence have a quicker understanding of where action could be needed, it would be helpful to have some automated feature in the excel that provides this outcome for each of the elements. In this way, when looking at the climate risk results sheet, it could be clearly identified what is contributing most to the scores. We welcome views.

TOPIC 19. Project 62: Updates on the 2019 SEAPODYM Review

No questions or comments

TOPIC 20. Project 110 and 110a: Progress report on Non-entangling and biodegradable FAD trial in the WCPO

No questions or comments

TOPIC 21. Ecosystem and Climate Indicators (Project 121)

Relevant Documents: SC21-EB-IP-01 SPC-OFP. Project 121 Update: Ecosystem and Climate Indicators

Questions and Comments

USA

The US appreciates the work that the SPC has done, including hosting the workshop last November, and we agree that this is a reasonable collection of indicators. One suggestion would be to add estimated median phytoplankton size as either a fishery or ecosystem indicator. It can be derived from satellite remotely sensed SST and chl-a data, and may have a more direct mechanistic relationship to fish size and other ecosystem properties than either SST or chl-a on their own.

Reply (SPC)

We have started to commence exploratory analyses along similar lines of thought. We look forward to working with the USA to fully test this indicator

Question

We note that there is no updated workplan included in SC21-EB-IP01. The previous workplan from SC20 (SC20-EB-WP-01) indicated that a WCPFC member workshop to refine the indicators would tentatively be held in 2026 and that the indicators would be "adopted" in 2026. We would appreciate an update from SPC on the anticipated next steps towards implementation.

Reply (SPC)

Workshops are planned for 2-6 March 2026 (Ecosystem Indicators) and 9-13 March 2026 (Climate Indicators) in Noumea, New Caledonia. We are waiting confirmation of technical expert participation. Once we have that we will ask the WCPFC Secretariat to send a circular notifying CCMs of the two workshops. Note that participation is not budgeted and will be at CCMs own expense.