Western and Central Pacific Fisheries Commission

SCIENTIFIC COMMITTEE
TWENTY-FIRST REGULAR SESSION

Nuku'alofa, Tonga
13–21 August 2025

Bigeye tuna management procedure design considerations

WCPFC-SC21-2025/MI-WP-06
28th July 2025

M. Wickens[1], F. Scott, R. Scott, N. Yao

[1]Oceanic Fisheries Programme of the Pacific Community

Contents

Executive Summary

The WCPFC harvest strategy workplan schedules the consideration and refinement of bigeye tuna management procedures (MPs) in 2025, with the adoption of the bigeye tuna MP by the end of 2026 (WCPFC, 2024). This report presents design considerations for candidate MPs and identifies a candidate estimation method (EM). The companion report (Scott et al., 2025a) presents performance evaluations of candidate MPs.

Two different candidate models were investigated for the bigeye tuna EM: an Age-Structured Production Model (ASPM) implemented using Stock Synthesis 3 (SS3) (Methot Jr., 2013) and a continuous-time surplus production model (SPiCT) (Pedersen and Berg, 2017). The models were fitted to data generated using stochastic projections from 2021 to 2051 across the proposed operating model (OM) grid. The future catches in the projections were set to four contrasting levels: 50%, 75%, 100% and 125% of status quo, where status quo was the mean catch from 2019 to 2021. Observation error was included on both the historical and projected catch data. EM performance was judged by comparing the 'true' stock status from OM simulations to the stock status estimated by the EM in the terminal year (the final modelled year).

The best performing EM was the ASPM implemented in SS3 with a single area, thirty-two extraction fisheries, one 'consensus' index fishery, observation error on catch, and two types of estimated parameters: a population scalar and recruitment deviations.

To evaluate the performance of EMs, four metrics of stock status were compared: the spawning biomass in the terminal year relative to the spawning biomass in the initial year (SB/SB_{Initial}), the spawning biomass in the terminal year relative to the spawning biomass in the terminal year if no fishing had occured $SB/SB_{F=0}$, the $SB/SB_{F=0}$ in the terminal year relative to the average $SB/SB_{F=0}$ in the reference period from 2012 to 2015 (SB/SB_{2012}), and the $SB/SB_{F=0}$ in the terminal year relative to the average $SB/SB_{F=0}$ in the reference period from 2017 to 2019 (SB/SB_{2017}). Of the four metrics assessed, SB/SB_{2012} produced the most accurate estimates of true stock status.

The EM implemented in SS3 sufficiently estimated stock status for the range of uncertainties simulated for this report. The SB/SB_{2012} metric was unbiased but produced less accurate estimates of stock status when the simulated stock was low. The EM exhibited some sensitivity to initial conditions and some models failed to converge. Further work may be needed to explore whether performance can be improved, and if there are more appropriate metrics for the estimation of stock status.

In addition to the EM, other operational considerations relevant to MP design were identified and discussed. It is essential that data collection, the shape of the HCR, the output type (catch- or effort-based), the spatial jurisdiction, and the frequency of MP implementation function cohesively to achieve management objectives.

SC21 is invited to:

- Consider the results of the estimation method trials for the bigeye tuna management procedure,
- Endorse the continued development of an estimation method implemented in Stock Synthesis 3 for the bigeye tuna management procedure, and
- Provide guidance on the operational considerations of MP design, including the harvest control rule, meta-rules and output type (catch- or effort-based).

1 Introduction

In accordance with the work plan for the adoption of harvest strategies under CMM 2022-03, SC21 is scheduled to consider and refine a candidate set of management procedures (MPs) for bigeye tuna in 2025 (WCPFC, 2024). This report presents design considerations for candidate MPs and the companion report (Scott et al., 2025a) presents evaluations of the candidate MPs.

An MP comprises three main components; a data collection program that ensures appropriate and sufficient information is available to monitor the stock and determine its status; an estimation method (EM)[2] that follows a pre-specified and fixed procedure to estimate the current status of the stock; and a harvest control rule (HCR) that sets the management action to be applied given the estimate of stock status from the EM. All three components should be considered as a whole when designing, evaluating and implementing an MP.

A significant portion of this report is dedicated to the EM, which is the most complex component of MP design. The EM must accurately estimate the stock status over a wide range of uncertainties and produce a useful metric for the HCR. Two models were considered and their design and evaluations are presented in Section 2.

It is important to note that the EM serves a different purpose to the stock assessment. The EM is a single model that provides an estimate of current stock status to the HCR to define future fishing levels and it need not accurately estimate historical population changes. In contrast, the stock assessment is used for monitoring purposes under the harvest strategy approach and is typically an ensemble of models that aim to estimate historical and current stock statuses.

Section 3 discusses the design considerations for data collection, the HCR, and other operational decisions within the MP.

2 Estimation method

Two candidate EMs for bigeye tuna were developed and tested: an Age-Structured Production Model (ASPM) implemented using Stock Synthesis 3 (SS3) and a continuous-time surplus production model (SPiCT) (Methot Jr., 2013; Pedersen and Berg, 2017). SS3 is a flexible and advanced framework that accomodates complex dynamics such as selectivity and spatial structure. SPiCT is a simple biomass model that does not include length, age or spatial variation.

The purpose of the EM is to produce an indication of stock status which, through the HCR, determines the necessary management to achieve management objectives. Often the EM produces a measure of relative stock depletion but this does not always need to be the case. EMs should be judged by their ability to produce an informative indicator over a wide range of potential future scenarios, and be robust to different sources of uncertainty.

[2] In this context, EM refers to estimation method, not electronic monitoring.

The EMs were designed with the principles from Rademeyer et al. (2007) in mind. The models used in the candidate EMs were simpler than the stock assessment for two main reasons: clear interpretation and reduced computational cost. The EM model structure was parsimonious and many parameter values were fixed to avoid extensive multiparameter estimation.

A grid of operating models (OMs) as described in Scott et al. (2025b) produced simulated population trajectories and generated test data for input into the EM. The performance of SPiCT and SS3 was evaluated by comparing the estimated current stock status to a known stock status derived from the OM simulation. The OMs represent the 'true' stock status and are based upon six models from the model grid in the 2023 stock assessment of bigeye tuna (Day et al., 2023). The estimated stock status metric was calculated differently depending on the candidate EM with the aim of measuring the relative biomass depletion.

Observation error in model inputs must be considered in EM design. An adopted EM should be robust to observation error as it is an inevitable component of the data collection process. Initial testing of the candidate EMs did not include observation error in order to evaluate their performance under idealised conditions. Even without observation error, the SPiCT model did not perform well, so further testing was not carried out. For the SS3 model, initial testing without observation error suggested that it could be considered as a candidate EM. Therefore, further testing with observation error was conducted.

In the proposed EMs, catch and CPUE are assumed to be available from all fisheries described in Table 1 of the most recent stock assessment (Day et al., 2023).

2.1 Generating test data

To evaluate the performance of the EMs, test data were generated from stochastic projections. The projections were generated using a grid of OMs, developed from the 2023 stock assessment implemented in Multifan-CL (Day et al., 2023; Fournier et al., 1998) and discussed further in Scott et al. (2025b). Stochasticity was included in future recruitment through the application of randomly selected historical recruitment deviates. Note that the OM grid utilised in this analysis may not necessarily exactly match the OMs described in Scott et al. (2025b). Nevertheless, the range of depletion in the simulated projections provided sufficient coverage to evaluate the performance of the EM.

For EM testing, the OM projections were run for 30 years from the final year of the stock assessment, 2021 to 2051. The EM was fit to 2400 simulations which covered the uncertainties outlined in Table 1.

Each combination of the uncertainty axes was simulated 5 times with different recruitment residuals, resulting in a total of 2400 simulations. The simulations covered the four main types of error pertinent to fishery models: model error, process error, implementation error and observation error. These settings produced a wide range of future population trajectories, which provided sufficient

Type	Axis	Levels	Options
Model error	Steepness	2	0.65, 0.8 and 0.95
	Tag mixing period	2	1 quarter and 2 quarters
Process Error	Recruitment period	2	long term and short term
Implementation Error	Effort creep	2	0% and 2% PS, 1% for LL
Observation Error	Catch variance	1	20%
Future scenario	Catch target	4	50%, 75%, 100% and 125% of SQ
	Terminal year	5	2011, 2021, 2031, 2041, 2051

Table 1: Axes of uncertainty in the OM projections used for EM testing. Effort creep is a measure of annual percentage increase in effort in the purse-seine (PS) and longline (LL) fisheries. Catch target is a scaling of the status quo (SQ) catch, where SQ was the mean catch from 2019 to 2021.

variation in stock status to evaluate the performance of the EMs.

Settings for 'steepness', 'tag mixing period', 'recruitment period', 'effort creep' and 'catch variance' axes are discussed in Scott et al. (2025b).

Two more axes were considered for EM testing. The 'terminal year' axis was added to examine model performance over different time scales. EMs with the terminal year prior to 2021 used 'true' catch and 'true' index data from the historical period and were not subjected to the data generation process. The 'catch target' axis was included to drive the stock to different depletion levels. Four different levels were used: status quo, 125% of status quo, 75% of status quo, and 50% of status quo, where status quo was the mean catch from 2019 to 2021.

CPUE was calculated as the ratio of total catch to effort for each index fishery. No observation errors were added to the CPUE time series. The fishery-specific indices used for the 2023 bigeye stock assessment (see Table 1 of Day et al. (2023)) were summed to produce a 'consensus' index to avoid the difficulties of reconciling differences in multiple indices within an assessment (Hoyle et al., 2024). The geostatistical model used to generate indices for the bigeye tuna fishery produced CPUE with implicit regional scaling so the 'consensus' index represents an area-weighted CPUE. Further development of the EM could explore alternative treatment of index fisheries. Future catchability was assumed to remain at the same level as the final three years in the OM so the mean from 2019 to 2021 was used, with seasonal catchability used for non-purse seine fleets in regions 1 and 2.

In these models, catch from the historical period and projection period were treated differently. Historical catch from 1952 to 2021 was provided directly from the OM to the EM and assumed to already include observation error. Observation error was added to the catch for the projection period, 2022 to 2051, by sampling from a normal distribution with a coefficient of variation equal to 0.2 and mean equal to the 'true' catch observation as shown in Figure 1 (Scott et al., 2025b). This approach aimed to mimic the error accumulated through the real-life data collection process.

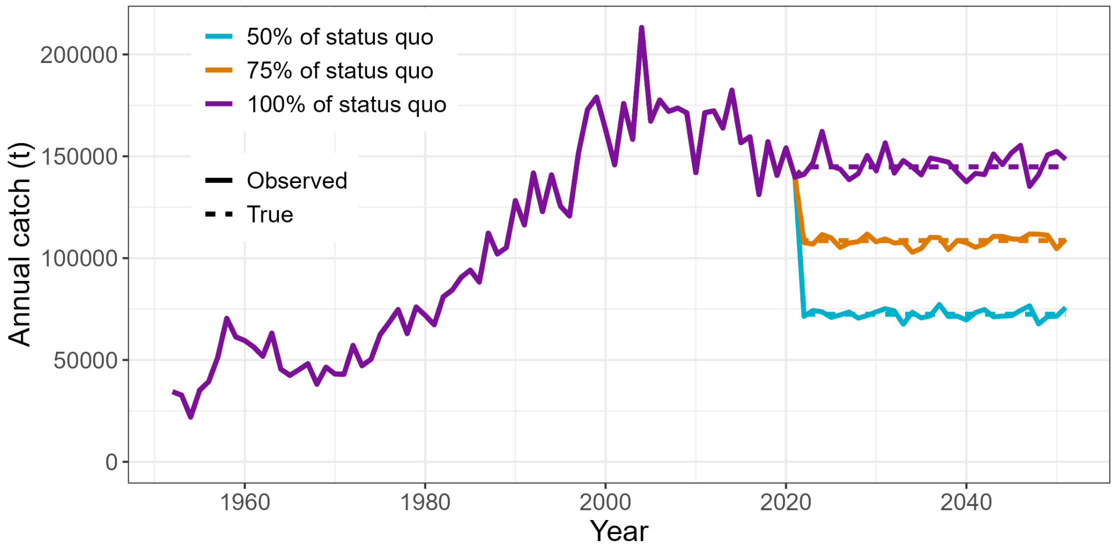

Figure 1: Three illustrative future catch scenarios for the bigeye tuna operating model. Future catches are set to 50%, 75% and 100% of the 2019-2021 average catch (dashed line) but the catches are observed with error (solid line) when provided to the estimation method. Some simulations could not sustain 125% of 2019-2021 average catch during the projection, and are therefore omitted to avoid obscuring the plot.

Neither candidate EM used tag data nor size composition data, so the simulation of tag data and size composition data was not required.

2.2 Stock Synthesis age-structured production model

The age-structured production model (ASPM) implemented in Stock Synthesis 3 (SS3) was run on an annual timestep with four seasons with thirty-two extraction fisheries as per the 2023 stock assessment but only one index fishery and no spatial structure. Previous work in 2023 and 2024 to develop an estimation method for South Pacific albacore highlighted the capability of ASPMs for estimation of stock status (Scott et al., 2024).

In fishery population models, structural parameters such as selectivity and growth are difficult to estimate without tagging and size composition data so selectivity and growth were fixed based on the diagnostic model from the 2023 stock assessment. Productivity parameters such as steepness and natural mortality are also difficult to estimate so for simplicty, these were also fixed based on the diagnostic model from the 2023 stock assessment. Movement and the distribution of recruits were obsolete without spatial structure so they are not considered in the tested SS3 EM.

Initial population size has a strong correlation with depletion so the parameter that dictates initial population size (R_0) must be estimated in the EM. Annual deviations from the stock recruitment curve called 'recruitment deviations' were also estimated in SS3 to allow for a better fit to CPUE data and increase flexibility in the model.

Parameter estimation was performed using maximum likelihood estimation. The convergence of the model was determined by the maximum gradient component. The acceptable maximum gradient component suggested by SS3 was 0.001, so this value was used to determine if the model was sufficiently converged.

Of the 2400 SS3 simulations, 23 failed for unknown reasons (possibly exceeding the compute limit) and 912 finished with a maximum gradient component greater than 0.001. There were 2242 converged simulations resulting in a 61% convergence rate. However, most non-converged models produced good estimates so all non-converged models are included in plots and summaries. The simulations that did not converge were mostly from higher catch targets and later terminal years.

During SS3 parameter estimation, it was discovered that the model exhibited sensitivity to the initial value of R_0. Despite the small number of estimated parameters, some models failed to converge when the initial value of R_0 was too high. Further development on this model should address this behaviour.

Details of the SS3 model can be found in Appendix A.

2.3 SPiCT surplus production model

An alternative candidate EM was investigated using Surplus Production in Continuous Time (SPiCT) (Pedersen and Berg, 2017). SPiCT is a simple biomass model with growth, natural mortality and recruitment all contained in a single parameter (r). Additionally, features such as selectivity, spatial dynamics and time-varying processes are not available. The model relies heavily on indices of abundance so requires a strong assumption that bigeye tuna CPUE timeseries are sufficiently informative indices of abundance and are representative of the population.

As a starting point, SPiCT was run with perfect observations of catch and CPUE in future projections. Nine CPUE timeseries from the index fisheries in Day et al. (2023) were provided to SPiCT. The population scale parameter (K) and growth parameter (r) were estimated. These preliminary investigations revealed SPiCT could not reliably estimate the stock status despite perfect observations of catch and CPUE. SPiCT has previously been explored for South Pacific albacore but dismissed as a candidate EM due to the high variance in stock status estimates (Scott et al., 2023b). Therefore, SPiCT was not considered further as a potential EM for bigeye tuna (see Figure 7).

2.4 Estimation method output metric

As noted above, the EM should not be considered as a stock assessment, but rather an algorithm that provides an input to the HCR. A relative measure of stock status is often used due to the ease of interpretation and reduction of bias in comparison to absolute measures. The HCR takes a single representative quantity so the EM does not need to recover the entire historical population accurately, as long as the estimated trend provides insight into the 'true' population.

To evaluate the performance of the EMs, we compared the estimated value from the EM to the 'true' value from the OM projection. We investigated the performance of four different metrics of the SS3 estimated stock status:

- the spawning biomass in the terminal year relative to the spawning biomass in the terminal if no fishing had occured ($SB/SB_{F=0}$), and
- the spawning biomass in the terminal year relative to the spawning biomass in the initial year (SB/SB_{Initial}), and
- the $SB/SB_{F=0}$ in the terminal year relative to the average $SB/SB_{F=0}$ in the reference period from 2012 to 2015 (SB/SB_{2012}), and
- the $SB/SB_{F=0}$ in the terminal year relative to the average $SB/SB_{F=0}$ in the reference period from 2017 to 2019 (SB/SB_{2017}).

The $SB/SB_{F=0}$ metric in SS3 is also commonly referred to as "dynamic B_0" which is sometimes a more appropriate method to measure relative stock status for management decisions (Bessell-Browne et al., 2022) and generally insensitive to the time period selected for the calculation (Berger et al., 2013).

The year range considered for the calculation of stock status metrics represent the reference period for the bigeye tuna target reference point (TRP) (2015 - 2015) and the recent time period (2017 - 2019), as used for South Pacific albacore (Pilling et al., 2024a,b).

SPiCT does not calculate the unfished biomass and also does not separate the spawning biomass from the total biomass so the metrics used for SPiCT were slightly different:

- the total biomass in the terminal year relative to the carrying capacity (TB/K), and
- the total biomass in the terminal year relative to the total biomass in the initial year (TB/TB_{Initial}).

2.5 Results

An estimation method should be able to determine the 'true' stock status measure in most plausible situations, including low biomass, high biomass, during the historical period and into the projected period. Ideally, the relationship between the EM estimate of stock status and the 'true' stock status should be linear. In other words, the stock status estimated by the estimation method should be consistent with the 'true' value from the operating model. To explore this, a linear regression was fitted to the results for each model and the findings are presented in Table 2.

The linear regression revealed that SPiCT resulted in no significant linear relationship between estimated and 'true' stock status. As can be seen in Figure 7, SPiCT fails to detect changes in the true relative total biomass (TB/TB_{Initial}). This initial investigation led to the rejection of SPiCT as a candidate EM.

In the case of SS3, the SB/SB_{2012} and SB/SB_{2017} metrics resulted in a strong linear relationship ($R^2 > 0.78$) between the estimated and 'true' stock status. Whereas, the SB/SB_{Initial} and $SB/SB_{\text{F=0}}$ metrics were slightly weaker ($R^2 > 0.62$) but still informative. Due to the high R^2 value and relevance to the proposed bigeye tuna reference period, SB/SB_{2012} was the preferred metric for the EM output.

Model	Output	Intercept	Slope	R^2
SPiCT	SB/SB_{Initial}	0.42	0.01	0
	SB/K	0.68	0.21	0.07
SS3	SB/SB_{Initial}	0	0.51	0.72
	$SB/SB_{F=0}$	-0.04	0.4	0.62
	SB/SB_{2012}	-0.34	1.31	0.78
	SB/SB_{2017}	-0.38	0.53	0.78

Table 2: Results of fitting linear regression of the estimated stock status against the 'true' stock status.

It is important that the EM is robust to all axes of uncertainty, including different depletion levels and time scales. The terminal depletion level in the simulations was mostly dictated by the projected catch levels, as seen in Figure 3. The SS3 EM was able to correctly detect changes in stock status between the tested range of 0.05 to 0.76 for $SB/SB_{F=0}$ and range of 0.12 to 2.72 for SB/SB_{2012}. However, accuracy declined for lower values of true stock status, indicating slightly worse performance in the lower range (see Figure 5).

EM performance should be invariant to the terminal year and consistent across the historical period and the projection period. The projection period may result in better fits due to "better behaved" data generated from the simulation. Data from the historical period exhibit characteristics of real future data and should therefore also support successful EM model fits. The EM results grouped by terminal year are shown in Figure 6 and Table 3. All groups showed consistent estimation of 'stock status' indicating the EM performed equivalently well in the historical and projected period. Table 3 contains diagnostics for the linear regression broken down by terminal year excluding 2011 and 2021 which have little variation in true stock status. Only the preferred metric (SB/SB_{2012}) and the widely-used metric ($SB/SB_{F=0}$) are shown for simplicity. The EM performs consistently well irrespective of terminal year.

A further point of interest in these results was the systematic tendency for the EM to underestimate the 'true' stock status. There was a directional bias which caused the estimated stock status to be 16% lower for SB/SB_{Initial} and 28% lower for $SB/SB_{F=0}$. It was for this reason that two more metrics were added which used a comparison to a reference period. Overall, the SB/SB_{2012} metric produced less biased results (1%). Although it is possible for the HCR to account for bias, addressing bias in the EM can be easier to interpret.

In summary, the SS3 model that included observation error performed well. All EM output met-

Model	Output	Terminal year	Intercept	Slope	R^2
SS3	$SB/SB_{F=0}$	2031	-0.01	0.31	0.92
		2041	-0.01	0.38	0.88
		2051	0.03	0.36	0.54
	SB/SB_{2012}	2031	-0.18	1	0.94
		2041	-0.18	1.26	0.91
		2051	-0.09	1.36	0.82

Table 3: Results of fitting linear regression of the estimated stock status against the 'true' stock status for different terminal years in the projection.

rics had high R^2 values from the linear regression. The SB/SB_{2012} metric sufficiently estimated the 'true' stock status, exhibited less biased results and led to the highest R^2 value from linear regression.

3 Management procedure design considerations

The management procedure contains three primary components: data collection, estimation method and harvest control rule. The candidate estimation methods are discussed in Section 2 and data collection and harvest control rules are discussed here. This section also discusses the following operational considerations: the frequency of running the MP, the harvest control rule output most practical for implementation and the spatial jurisdiction of the MP. Evaluations of candidate MPs are discussed in Scott et al. (2025a).

3.1 Data collection

To support the future implementation of an MP, the data produced by the fishery must meet the requirements of the MP by providing the estimation method with the necessary information. The candidate estimation methods implemented in SS3 relied on total catch estimates for 32 extraction fisheries and relative abundance indices from 9 index fisheries which are combined to generate a whole-fishery index fleet. This fishery structure was based on the 2023 stock assessment (Day et al., 2023). Future data collection must continue to produce catch and effort timeseries suitable for anuual catch by weight totals and standardised CPUE indices. Neither of the candidate EMs tested here required tagging data nor size composition data so the collection of these data is not necessary for the candidate EM.

3.2 Harvest control rule

An HCR is a pre-agreed decision rule that is used to set fishing opportunities in the future. It takes an estimate of stock status from the EM and prescribes the necessary level of fishing in alignment with management objectives. The purpose of the HCR is to achieve fishery objectives by adjusting catch or effort. For example, the HCR should ensure the stock has a high probability of avoiding

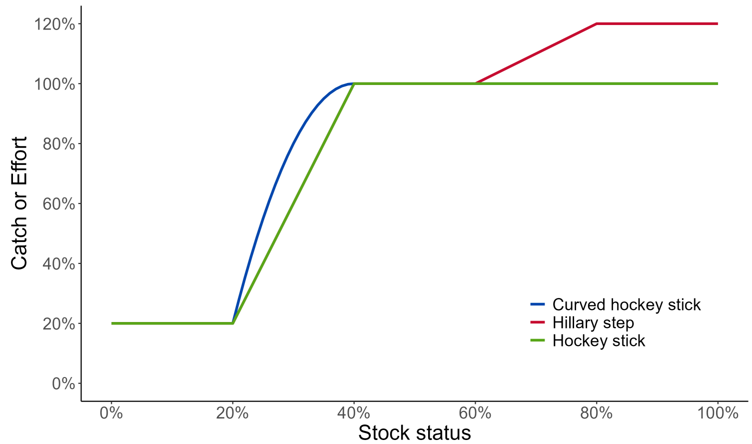

Figure 2: Example HCR shapes.

the limit reference point (LRP) while keeping the stock around the target reference point (TRP) on average. This section discusses HCR design.

The HCR should reduce fishing pressure if the stock status is approaching or below the LRP and maintain fishing pressure if the stock status is near the target reference point. The most common HCR shape, referred to as the 'hockey stick', is determined by four parameters: B_{lim}, B_{elbow}, C_{min} and C_{max} (see Figure 2). B_{lim} refers to the leftmost bend and B_{elbow} refers to the rightmost bend in the 'hockey stick' shape. When the stock status is less than B_{lim} the catch limit is set at C_{min}. When the estimated stock status is greater than B_{elbow} the catch limit is set at C_{max}. When the stock status is between B_{lim} and B_{elbow}, the catch limit is set according to the slope.

An alternative to the hockey stick shape is to replace the straight sloping line with a curving line (see Figure 2). The curving line allows for more gradual adjustments in fishing pressure if the estimated stock status drops slightly below the TRP. This avoids making potentially large reductions in fishing pressure when the population is only marginally below the desired level. However, if the biomass continues to decline, the HCR calls for steeper cuts to address the worsening situation and prevent further depletion.

An additional feature, referred to as the 'Hillary step', allows for increased fishing pressure if the stock is above the target reference point (see Figure 2). When the estimated stock status aligns with the Hillary step, the HCR maintains the baseline fishing effort or catch limits, even if stock status increases or decreases by a limited amount. Beyond the step, the fishing pressure is allowed to increase. The HCR adopted for skipjack at WCPFPC19 in 2022 and some candidate HCRs for South Pacific albacore use a Hillary step shape (WCPFC, 2022; Scott et al., 2023a).

3.3 Meta-rules

The HCR can be further restricted with additional meta-rules which limit the percentage change in catch or effort between management periods. Meta-rules prevent large fluctuations in fishing levels which may be difficult for the fishing industry to adapt to. However, the inclusion of meta-rules may slow the response of necessary changes in fishing pressure to reach the TRP. The adopted skipjack MP contains a meta-rule whereby the output of the HCR cannot change by more than 10% between management periods (WCPFC, 2022).

3.4 Operational considerations

Operational considerations of MP design include the frequency of MP evaluation, the type of control (catch or effort), and the spatial jurisdiction of the management procedure.

The MP must be run frequently enough to detect changes in stock status and allow management to react in a timely manner. However, the frequency may be limited by the time taken to collect data, perform EM analyses and run the MP. The previously developed skipjack and South Pacific albacore MPs are run every three years (Scott et al., 2022, 2023b) and the proposed revision of the harvest strategy schedule indicates a three year cycle for bigeye tuna as well (Larcombe and Pilling, 2025).

Fishery management advice may be expressed in terms of either catch or effort, with conversions used to translate between the two. For example, if the harvest control rule prescribes catch adjustments to regulate fishing pressure, a fishery that operates under effort constraints may apply a conversion to determine the equivalent level of effort.

Under the proposed mixed fishery framework, the spatial jurisdiction of the bigeye tuna MP is restricted to longline fisheries operating between 20N and 10S. Purse-seine fisheries are managed through the skipjack MP and fisheries south of 10S are managed through South Pacific albacore MP. The current proposal is for only the longline fisheries operating between 20N and 10S to be controlled by the bigeye tuna MP.

4 Conclusion

All aspects of the MP must function cohesively to achieve management objectives. We propose an EM implemented in SS3, with stock status measured as the $SB/SB_{F=0}$ in the terminal year relative to the average $SB/SB_{F=0}$ in the reference period from 2012 to 2015 (SB/SB_{2012}). There may be a need to re-evaluate the performance of the EM following any significant developments with the OMs.

SC21 is invited to:

- Consider the results of the estimation method trials for the bigeye tuna management procedure,
- Endorse the continued development of an estimation method implemented in Stock Synthesis 3 for the bigeye tuna management procedure, and

- Provide guidance on the operational considerations of MP design, including the harvest control rule, meta-rules and output type (catch- or effort-based).

Acknowledgements

We gratefully acknowledge funding for this work from the New Zealand Ministry of Foreign Affairs and Trade (MFAT) funded project "Sustainable Pacific Fisheries".

5 References

Berger, A., Pilling, G., Kirchmer, C., and Harley, S. (2013). Determination of appropriate time-windows for calculation of depletion-based limit reference points. WCPFC-SC9-2013/MI-WP-02, Pohnpei, Federated States of Micronesia, 6-14 August 2013.

Bessell-Browne, P., Punt, A. E., Tuck, G. N., Day, J., Klaer, N., and Penney, A. (2022). The effects of implementing a 'dynamic B0' harvest control rule in Australia's southern and eastern scalefish and shark fishery. *Fisheries Research*, 252:106306.

Day, J., Magnusson, A., Teears, T., Hampton, J., Davies, N., Castillo-Jordan, C., Peatman, T., Scott, R., Scutt Phillips, J., McKechnie, S., Scott, F., Yao, N., Natadra, R., Pilling, G., Williams, P., and Hamer, P. (2023). Stock assessment of bigeye tuna in the western and central pacific ocean: 2023. WCPFC-SC19-2023/SA-WP-05 (Rev. 2), Pohnpei, Federated States of Micronesia.

Fournier, D., Hampton, J., and Sibert, J. (1998). MULTIFAN-CL: a length-based, age-structured model for fisheries stock assessment, with application to South Pacific albacore, *Thunnus alalunga*. *Canadian Journal of Fisheries and Aquatic Sciences*, 55:2105–2116.

Hoyle, S. D., Campbell, R. A., Ducharme-Barth, N. D., Gruss, A., Moore, B. R., Thorson, J. T., Tremblay-Boyer, L., Winker, H., Zhou, S., and Maunder, M. N. (2024). Catch per unit effort modelling for stock assessment: A summary of good practices. *Fisheries Research*, 269:106860.

Larcombe, J. and Pilling, G. (2025). Wider issues for consideration within the harvest strategy workplan review. WCPFC-SC21-2025/MI-WP-05, Nuku'alofa, Tonga, 13–21 August 2025.

Methot Jr., R. D. Wetzel, C. R. (2013). Stock synthesis: A biological and statistical framework for fish stock assessment and fishery manegement. *Fisheries Research*, 142:86–99.

Pedersen, M. W. and Berg, C. W. (2017). A stochastic surplus production model in continuous time. *Fish and Fisheries*, 18(2):226–243.

Pilling, G., Scott, R., Hamer, P., and Scott, F. (2024a). Analyses to inform discussions on candidate bigeye and yellowfin target reference points. WCPFC-SC20-2024/MI-WP-07, Manila, Philippines. 14 Aug - 21 Aug 2024.

Pilling, G., Scott, R., Hamer, P., Scott, F., and Nan, Y. (2024b). Recalibration of the adopted south pacific albacore interim target reference point and review of WCPFC20 requested options. WCPFC-SC20-2024/MI-WP-03REV1, Manila, Philippines. 14 Aug - 21 Aug 2024.

Rademeyer, R., Plaganyi, E., and Butterworth, D. (2007). Tips and tricks in designing management procedures. *ICES Journal of Marine Science*, 64:618–625.

Scott, F., Scott, R., and Yao, N. (2024). Testing and developing an estimation method for South Pacific albacore. WCPFC-SC20-2024/MI-WP-05, Manila, Philippines, 14-21 August 2024.

Scott, F., Scott, R., Yao, N., and Natadra, R. (2023a). Evaluation of candidate management procedures for south Pacific albacore. WCPFC-SC19-2023/MI-WP-06, Koror, Palau, 16-24 August 2023.

Scott, F., Scott, R., Yao, N., and Natadra, R. (2023b). Testing and developing estimation models for south pacific albacore. WCPFC-2023/SC19-MI-IP-02, Koror, Palau, 16-24 August 2023.

Scott, F., Scott, R., Yao, N., and Wickens, M. (2025a). Evaluation of bigeye management procedures. WCPFC-SC21-2025/MI-WP-05, Nuku'alofa, Tonga, 13–21 August 2025.

Scott, R., Scott, F., Yao, N., Hamer, P., and Pilling, G. (2022). Evaluations of candidate management procedures for skipjack tuna in the WCPO. WCPFC-SC18-2022/MI-WP-02, online, 10-18 August 2022.

Scott, R., Scott, F., Yao, N., and Wickens, M. (2025b). Candidate bigeye operating models. WCPFC-SC21-2025/MI-WP-05, Nuku'alofa, Tonga, 13–21 August 2025.

WCPFC (2022). Conservation and management measure on a management procedure for WCPO skipjack tuna. Conservation and Management Measure 2022-01, Nineteenth Regular Session of the WCPF Commission, Da Nang City, Vietnam, 28–3 December 2022.

WCPFC (2024). Indicative work plan for the adoption of harvest strategies under cmm 2022-03. WCPFC21-2024-HS Workplan, Twenty-First Regular Session of the WCPF Commission, Suva, Fiji, 28–3 December 2024.

A Appendix - SS3 model description

Growth

Growth in the SS3 EMs used the single-sex von Bertalanffy length-at-age growth model, based on the 2023 stock assessment. For the stock assessment, the von Bertalanffy growth curve was fitted to length and age data in the MFCL model. The estimated parameters in MFCL were the mean length of the first age class (L_1), the mean length of the oldest age class (L_A) and the von Bertalanffy growth coefficient (κ). Across the stock assessment grid, the estimated growth parameters resulted in similar growth curves so the mean value of each parameter was provided to SS3. SS3 has a slightly different parameterisation to MFCL but by setting the first age class to 0.25 and the oldest age class to 10, the parameters have similar intepretations.

The von Bertalanffy growth model in SS3 takes the form:

$$L_a = L_\infty + (L_1 - L_\infty)\exp\left(-\kappa(a - A_1)\right) \tag{1}$$

where

$$L_\infty = L_1 + \frac{L_2 - L_1}{\exp\left(-\kappa(A_2 - A_1)\right)}. \tag{2}$$

The parameters were set up as follows: the mean length of the first age class $L_1 = 24.06$ cm, the mean length of the oldest age class $L_2 = 151.28$ cm, the growth coefficient $\kappa = 0.43$, the minimum age $A_1 = 0.25$ and the maximum age was $A_2 = 10$. The coefficient of variation was set to 0.1 for both young and old ages.

The weight-length relationship in SS3 was parameterised as:

$$W = \alpha L^\beta \tag{3}$$

with $\alpha = 0.000031$ and $\beta = 2.93$ as per the diagnostic model of the 2023 stock assessment.

Maturity and fecundity

Maturity and fecundity in the EM was assumed known and fixed at the maturity-at-age values from the diagnostic model of the 2023 stock assessment. It was assumed that fecundity was already accounted for in the MFCL maturity-at-age vector.

Selectivity

Selectivity in the EM was assumed to be known and fixed at similar values to the diagnostic model of the 2023 stock assessment model by grouping similar fisheries and mirroring the selectivity functions. The logistic and double normal selectivity options in SS3 were most appropriate for the bigeye tuna fisheries due to the simple implementation and similarity to selectivity shapes in the operating model. The age structure in SS3 used only 10 'nodes', making the selectivity function coarse compared to MFCL which used quarterly ages for a smoother selectivity function.

Fisheries 6 (LL-OS-7) and 18 (2-ID-PH-7) used a logistic selectivity to prevent the model from storing cryptic biomass in older age classes. The inflection and width of the logisitic curve were both set to 2.

Longline fisheries (1, 2, 3, 4, 5, 7, 8, 9, 10, 11, 12, 27, 29, 33, 34, 35, 36, 37, 38, 39, 40, and 41) used a double normal selectivity with peak at age 3.5 and non-zero asymptote at 0.53 (the mean selectivity at maximum age from the 2023 stock assessment grid).

Purse seine and domestic fisheries (13, 14, 15, 16, 17, 19, 20, 21, 22, 23, 24, 25, 26, 28, 30, 31 and 32) also used a double normal selectivity with peak at age 0.5 and a zero asymptote, typical for non-longline fisheries (see Day et al. (2023) for fishery descriptions).

Natural mortality and steepness

Natural mortality and steepness were assumed known and set to the same value used in the diagnostic model of the 2023 stock assessment, $M = 0.11$ at age 10 and $h = 0.8$ respectively. The Lorenzen form was used to scale natural mortality by the length of the fish.

Recruitment deviations

Recruitment deviations were estimated by the EM to improve the fit to catch-rate data. The variability in recruitment, σ_R, was set to 0.2 to reduce large deviations from the stock-recruitment relationship in the context of limited data.

B Appendix - Figures

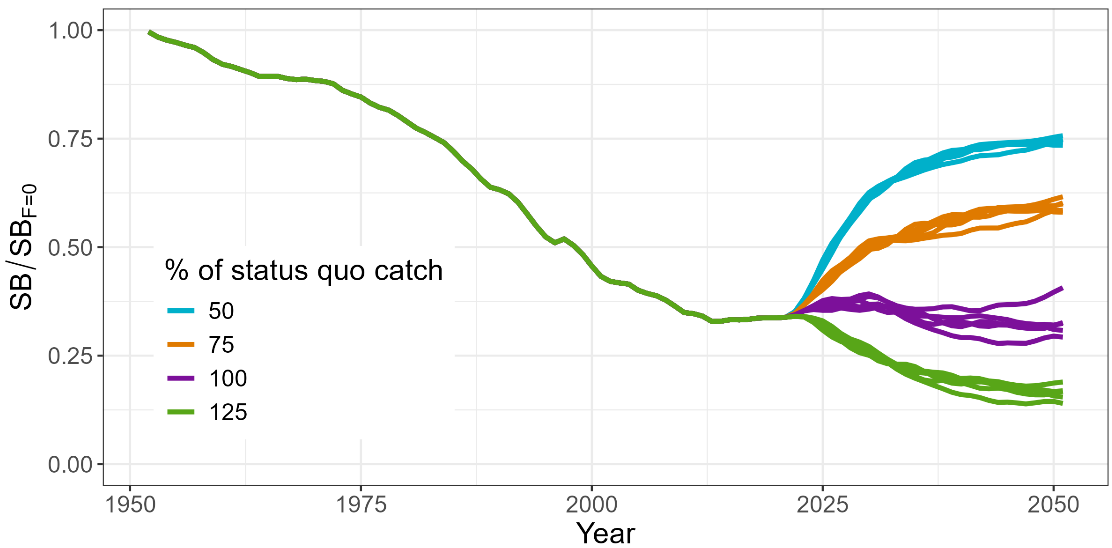

Figure 3: Example population trajectories under the four tested catch target levels with five illustrative simulations for each level.

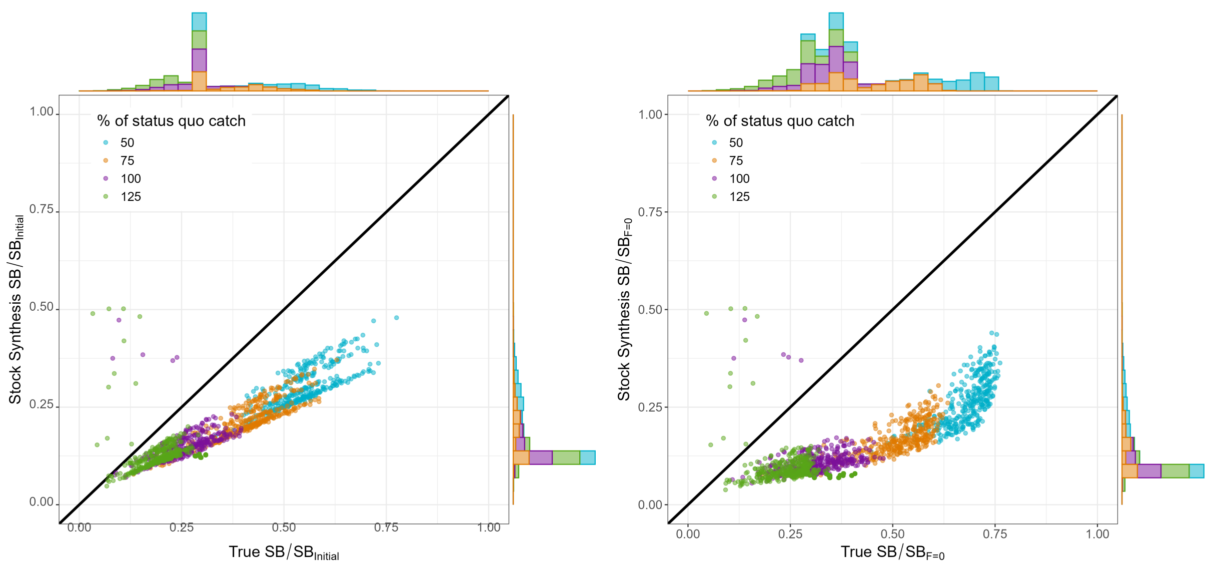

Figure 4: Results of SS3 coloured by projected catch level. Comparison of 'true' operating model stock status and estimate of stock status from SS3. Estimated stock status was measured using the spawning biomass in the terminal year relative to the spawning biomass in the initial year (left) and the spawning biomass in the terminal year relative to the unfished spawning biomass in the terminal year (right).

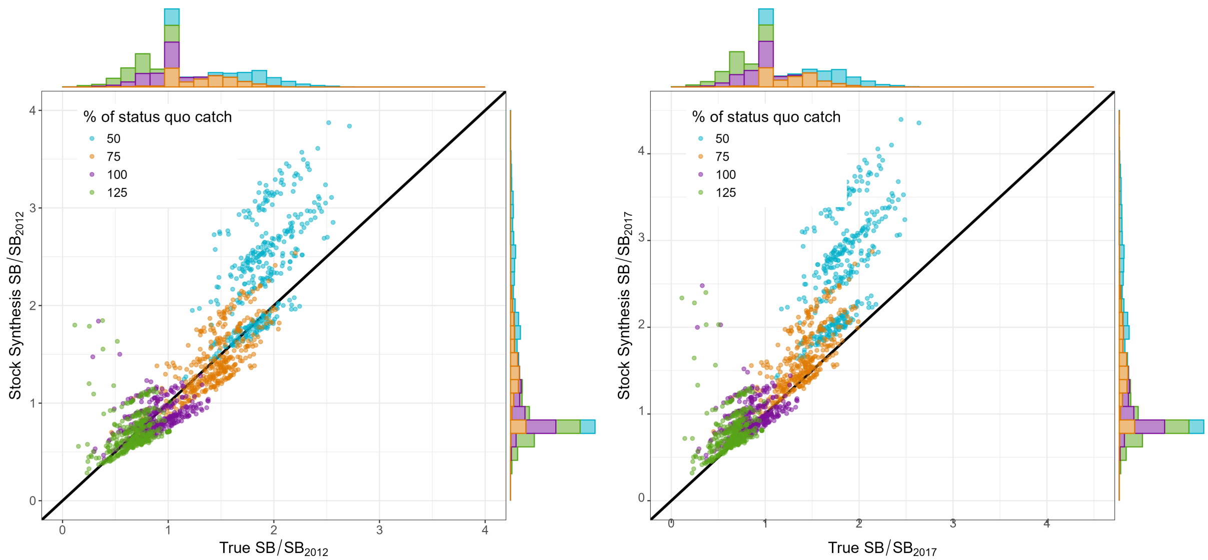

Figure 5: Results of SS3 coloured by projected catch level. Comparison of ‘true’ operating model stock status and estimate of stock status from SS3. Estimated stock status was measured using the $SB/SB_{F=0}$ in the terminal year relative to the average $SB/SB_{F=0}$ in the reference period from 2012 to 2015 (SB/SB_{2012}) (left) and the $SB/SB_{F=0}$ in the terminal year relative to the average $SB/SB_{F=0}$ in the reference period from 2017 to 2019 (SB/SB_{2017}) (right).

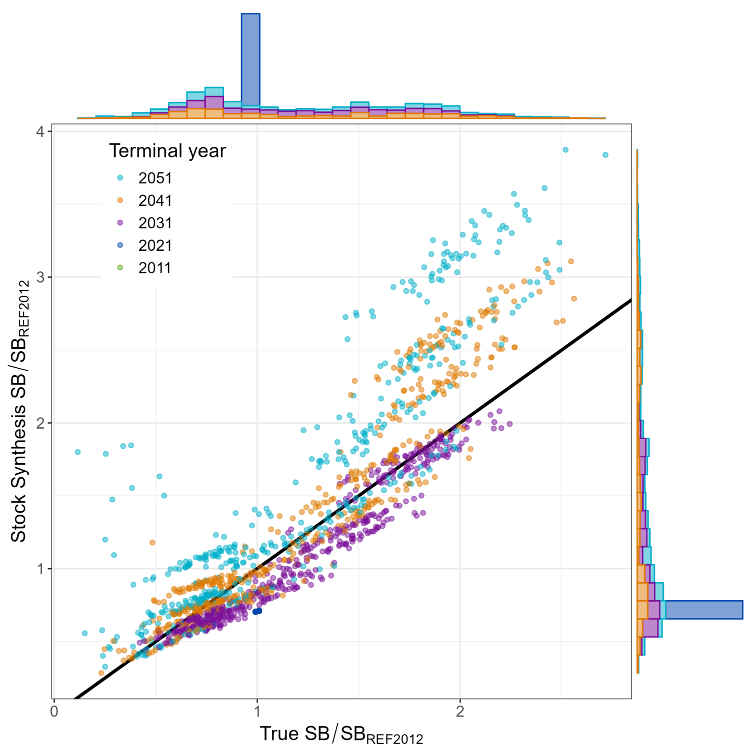

Figure 6: Results of SS3 coloured by terminal year. Comparison of ‘true’ operating model stock status and estimate of stock status from Stock Synthesis using the $SB/SB_{F=0}$ in the terminal year relative to the average $SB/SB_{F=0}$ in the reference period from 2012 to 2015 (SB/SB_{2012}). Note, the SB/SB_{2012} metric can not be calculated for the simulations with a terminal year of 2011.

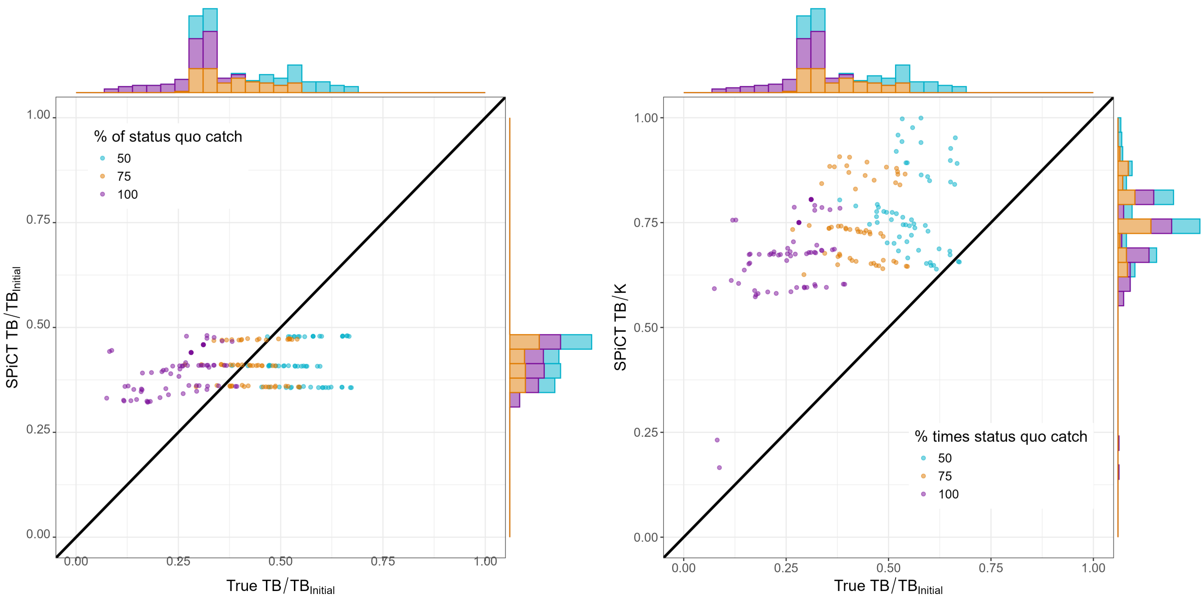

Figure 7: Results of SPiCT coloured by projected catch level. Comparison of 'true' operating model stock status and estimate of stock status from SPiCT without observation error. Estimated stock status was measured using total biomass relative to initial total biomass (left) and total biomass relative to carrying capacity (right).