SCIENTIFIC COMMITTEE
TWENTIETH REGULAR SESSION
Manila, Philippines
14–21 August 2024

Testing and developing an estimation method for South Pacific albacore

WCPFC-SC20-2024/MI-WP-05

9th August 2024

F. Scott[1], R. Scott and N. Yao

[1]Oceanic Fisheries Programme, The Pacific Community

Executive Summary

An Age-Structured Production Model (ASPM) implemented using Multifan-CL was tested as a candidate estimation methods (EM) for South Pacific albacore (SPA).

The model was fit to data generated by performing 800 stochastic projections to 2052 across the proposed SPA operating model grid, excluding effort creep. The future catches in the projections were set to a sine wave to provide contrast in the results. Observation error was included on both the historical and projected catch and CPUE data, and used as input data to the candidate EM. Of the 800 projections, 60 failed due to stock collapse from over-fishing. The input data for the ASPM were truncated to ten different final years, from time series ending in 2007 to 2052 in increments of five years. This gave over 7000 input data sets of increasing length for the candidate EM, across a range of true $SB/SB_{F=0}$ values. The ASPM fitted well across all of the projected input data sets with a high level of convergence and a low failure rate (about 6%).

The ASPM showed a linear relationship between the estimated and true $SB/SB_{F=0}$ across a range of true $SB/SB_{F=0}$ values for four different examined metrics of $SB/SB_{F=0}$:

- Terminal year $SB/SB_{F=0}$
- Mean $SB/SB_{F=0}$ of the last three years
- Mean $SB/SB_{F=0}$ of the last three years relative to the mean $SB/SB_{F=0}$ in 2000-2002
- Mean $SB/SB_{F=0}$ of the last three years relative to the mean $SB/SB_{F=0}$ in 2019-2022

For each of the metrics, two calculations of $SB/SB_{F=0}$ are used: instantaneous ($SB_y/SB_{F=0,y}$) and latest ($SB_{latest}/SB_{F=0}$, i.e. SB in year y relative to the average $SB_{F=0}$ in years y-10 to y-1), giving eight possible metrics that could be considered as an input to the harvest control rule (HCR).

Uncertainty in natural mortality has a strong impact on the estimated stock status. The ASPM uses a fixed level of natural mortality and this can result in high levels of bias when using estimates of absolute stock status such as $SB/SB_{F=0}$, impacting the performance of the ASPM as an EM.

As the trend in stock status was generally well estimated and unaffected by the assumed level of natural mortality in the ASPM, the impact of uncertainty in natural mortality could be reduced by using an estimate of stock status relative to some period in the past, for example, $SB/SB_{F=0}$ relative to the mean $SB/SB_{F=0}$ in 2017-2019, the period used in calculating the interim Target Reference Point for SPA. This is different to the inputs of previous example SPA HCRs and the HCR of the adopted SKJ management procedure (MP), and will need careful consideration and testing.

All candidate MPs will be tested using the Manangement Strategy Evaluation framework and their relative performance compared using performance indicators.

Next steps:

- Develop candidate management procedures (MPs) using estimation methods based on both

absolute and relative measures of stock status.

- The shapes of the harvest control rules in the candidate MPs can be adapted from those considered at WCPFC20.
- Test the candidate MPs using an updated Management Strategy Evaluation framework, using the updated operating model grid.
- Calculate and present performance indicators so that members may evaluate the relative performance of the candidate MPs and provide feedback.

SC20 is invited to:

- Consider the results of the estimation method trials for the South Pacific albacore management procedure.
- Endorse the use of the proposed ASPM as the estimation method for candidate management procedures.
- Endorse the use of relative and absolute depletion metrics for input into candidate South Pacific albacore harvest control rules.

1 Introduction

Initial investigations in 2023 into an estimation method (EM) for use as part of a management procedure (MP) for South Pacific albacore (SPA) suggested that an age-structured production model (ASPM), based on the 2021 stock assessment for SPA and implemented using Multifan-CL, might be appropriate (Scott et al., 2023).

The stock assessment upon which the ASPM was based had several features that received additional scrutiny from members, for example, the presence of a 'big dip' in stock status in the first few years of a projection based on that assessment.

Advancements have been made to the most recent stock assessment for SPA, most notably by moving to a fully catch-conditioned model, which have led to much improved diagnostics (Teears et al., 2024). These advancements have been adopted by the updated proposed operating models (OMs) for management strategy evaluation (MSE) (Scott et al., 2024). This prompted a re-evaluation of the EM presented to SC19.

This report presents the performance of that updated ASPM. The same approach as described in WCPFC-SC19-2023/MI-IP-02 to investigate its performance is used here, where the updated EM is fitted to simulated future data. Its performance is then evaluated by investigating the strength of the linear relationship between the estimated and 'true' stock status; although any non-linearity or bias can potentially be accounted for by the harvest control rule (HCR) design.

As always, it is worth noting that when using a model-based approach for the EM it should not be treated as a full stock assessment. Instead, it should be considered as an algorithm that generates the input signal to the HCR. The input to the HCR is the current stock status, which means that the EM does not need to effectively estimate the historical stock status. Therefore, the focus of the EM results is on the most recently estimated value of stock status and the estimate of the full time series is not considered.

2 Generating test data

Test data were generated in a similar way to Scott et al. (2023). Stochastic projections based on the proposed 2024 SPA OM grid were used to generate time series data which were used to fit the candidate EM and evaluate performance (Scott et al., 2024). The 30 year projections start in 2023 and continue until 2052. Effort creep is not considered in the analysis at this stage but will be investigated in the future.

The projections were conditioned on fixed levels of future catches for each model fishery. During the first three years (2023-2025) of the projection the future catch for each fishery was set as a mean of the 2020-2022 catches. To introduce contrast in the projected catch values, from 2026 the catches were set as a sine wave, with a mean value of the mean 2020-2022 catches, an amplitude of 25% of the mean value and a wavelength of 27 years, i.e. one full period was completed over the projection period (Figure 1).

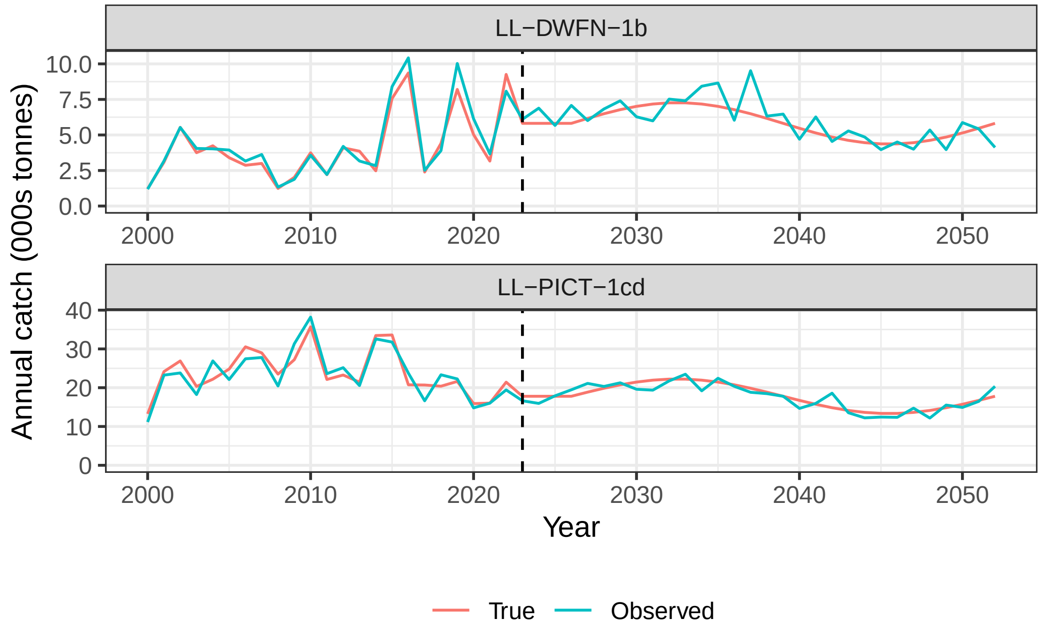

Figure 1: Observed and 'true' annual catches of two example fisheries, DWFN longline in model region 1b and PICT longline in model region 1cd, from a single projection. Each projection has different observed catches due to observation error. Future catches in the first three years (2022-2025) are at the 2020-2022 average, and then from 2026 as a sine wave, with a mean value equal to the 2020-2022 average and an amplitude of 25% of the mean value. For illustration purposes, only data from 2000 are shown. The dashed vertical line is at 2023, the first year of the projection.

400 projection models were generated where each had a different sample of natural mortality (M) and steepness, sampled from the distributions used to generate the OMs (Scott et al., 2024). Recruitment variability was included in the projections by applying recruitment residuals sampled from the historical period to the recruitment predicted by the stock-recruitment relationship. Two levels of future recruitment variability were included, with residuals sampled from either the period 1972-2020 or 2000-2020, noting that residuals earlier in the time series were more negative leading to a more pessimistic stock status. Each of the 400 projection models was projected twice with different randomly sampled recruitment residuals, giving a total of 800 projections.

Observation error was included for the historical and projected periods for the catches of each extraction fishery and for the catch-per-unit effort (CPUE) of each index fishery, with a coefficient of variation equal to 0.20. The observed catch and CPUE were used to generate input data for the candidate EM.

Of the 800 projections, 60 failed due to the stock biomass falling too low, leaving 740 projections to be used as test data for the candidate EM.

3 Multifan-CL Age-Structured Production Model

An Age-Structured Production Model (ASPM) does not use catch size distribution data, only catch and effort data. The Multifan-CL ASPM used here is based on the diagnostic 2024 stock assessment for South Pacific albacore (SPA). The regional and fishery structure of the ASPM is the same as the 2024 assessment, i.e. two model regions, 17 extraction fisheries and three index fisheries. Various parameters are fixed to the same value as the diagnostic case: steepness of the stock-recruitment relationship, movement rates, growth, natural mortality, selectivity and recruitment distribution (Teears et al., 2024).

The candidate EM was fitted to each of the 740 projections described above. The time series of input data were truncated to different final years, similar to a retrospective analysis. The final years ranged from 2007 to 2052 in increments of five years, giving 10 final years per projection: four in the historical period (2007 to 2022) and six in the projected period (2027 to 2052). The candidate EM was therefore fitted 7400 times, each fit giving an estimate of stock status in the final year.

The model was fitted in several phases, with each phase fixing more parameters, similar to fitting a Multifan-CL stock assessment. The final phase was run for a maximum of 3000 evaluations. This was sufficient to reach a maximum gradient of less than 1e-7 for all successful model fits, i.e. all models were considered to have converged. This is an improvement on the ASPM presented to SC19, for which out of the 6060 model fits, only 1000 finished with maximum gradients less than 0.01

Out of the 7400 ASPM fits, 480 models failed to fit, a failure rate of 6.5%. These were from projections that produced unrealistically low CPUE for the index fisheries at some point in the projected time series (hundreds of times smaller than seen in the historical period) due to low biomass. This left 6920 model fits.

4 Impact of uncertainty in natural mortality

Estimating natural mortality (M) can be difficult and uncertainty in M is an important component of the uncertainty grid for both the stock assessment and operating models (Teears et al., 2024; Scott et al., 2024). The assumed value of M has a big impact on the estimates of stock status and depletion. For example, Figure 11 in SC20-MI-WP-04 shows that a stock becomes less depleted as M increases.

This feature is also present in the ASPM presented here. The ASPM assumes M is fixed at the median level of the distribution used in the 2024 stock assessment and operating model grids. The selection of M effectively scales the stock status but does not affect the trend. This means that there is the possibility that estimates of stock status by the ASPM may be biased by the selection of the fixed value of M. This bias is on top of any other bias that may occur from using the ASPM.

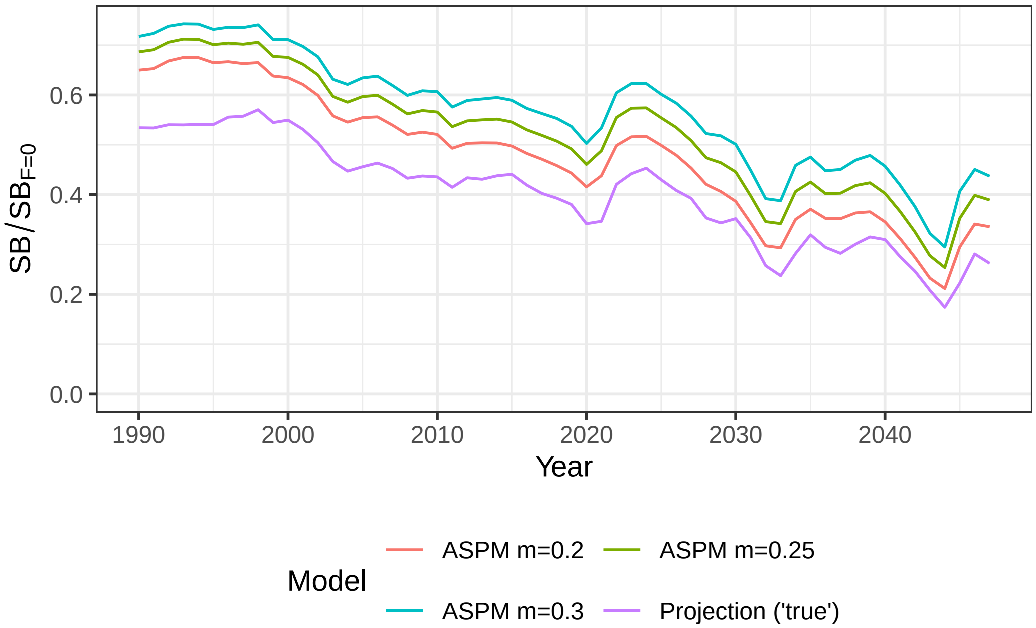

Figure 2: The 'true' $SB/SB_{F=0}$ from a single stochastic projection with recruitment variability. The true value of m for the projection model is 0.23. The estimated $SB/SB_{F=0}$ from three fits of the ASPM with different fixed values of natural mortality (m): 0.2, 0.25 and 0.3 are shown. No observation error was included to help emphasise the impact of the selection of m for the ASPM. The fits are performed on data up to 2047. For illustration purposes, only data from 1990 are shown.

For example, Figure 2 shows the $SB/SB_{F=0}$ from a single stochastic projection from a model which has a 'true' value of M of 0.23. The resulting estimated $SB/SB_{F=0}$ from three different ASPMs, each with a different fixed value of M (0.2, 0.25 and 0.3), are shown. No observation error was included to help emphasise the impact of the selection of M for the ASPM. Each of the estimates show a similar trend to that of the projection model, but the bias increases with the fixed value of M. For these fits, the estimated stock status is always overestimated.

As only the bias, and not the trend, appears affected by the value of M in the ASPM, to reduce the impact of the selection of the value of M, one approach is to calculate the stock status relative to some reference period in the past (Figure 3). This reduces the bias, making the selection of the value of M for the ASPM less important. The input to the harvest control rule (HCR) would therefore be $SB/SB_{F=0}$ relative to some period in the past, instead of an absolute measure of $SB/SB_{F=0}$.

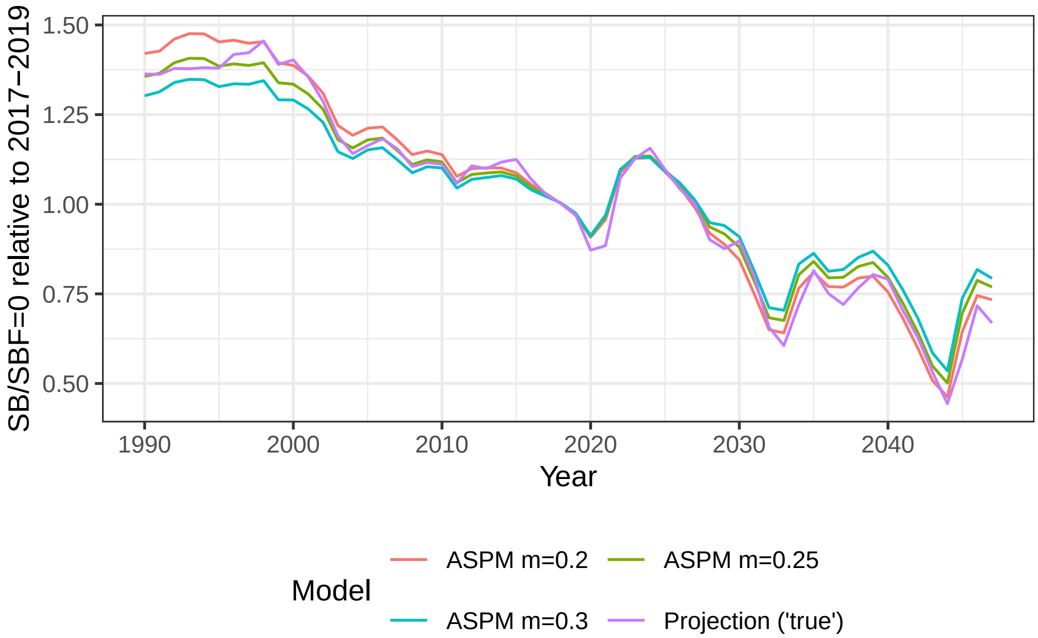

Figure 3: The SB/$SB_{F=0}$ relative to the mean SB/SBF=0 in 2017-2019 for all four models shown in the previous figure. The impact of M on the performance of the ASPM has been reduced.

An appropriate reference period is 2017-2019, i.e. the period used to calculate the interim TRP for SPA (WCPFCP20 summary report, paragraph 238). However, in this report the performance of the ASPM using only historical data is also of interest, i.e. earlier than 2017. Therefore, as an illustration of this approach, an example alternative reference period is chosen as 2000-2002. Results for 2017-2019 are also included but these only include mostly projected values of stock status.

5 Methods

As noted above, the EM should not be considered as a stock assessment, but rather an algorithm that calculates an input to be used by the HCR. Only the terminal estimated values of stock status are therefore used.

Four metrics of ASPM estimated stock status are calculated and compared to the 'true' value from the projection to evaluate the performance of the ASPM as an estimation method:

- Terminal year SB/$SB_{F=0}$
- Mean SB/$SB_{F=0}$ of the last three years
- Mean SB/$SB_{F=0}$ of the last three years relative to the mean SB/$SB_{F=0}$ in 2000-2002
- Mean SB/$SB_{F=0}$ of the last three years relative to the mean SB/$SB_{F=0}$ in 2019-2022

The last two metrics are relative measures of SB/$SB_{F=0}$ and are designed to reduce the bias resulting from uncertainty in natural mortality described above.

For each of the metrics, two calculations of $SB/SB_{F=0}$ are used: instantaneous (where depletion is calculated as the spawning biomass in year y relative to the unfished spawning biomass estimated for that year y: $SB_y/SB_{F=0,y}$) and latest ($SB_{latest}/SB_{F=0}$, i.e. SB in year y relative to the average $SB_{F=0}$ in years y-10 to y-1), giving eight possible measures of stock status. Each of these measures could potentially be used an an input to the HCR.

To examine whether the ASPM estimated stock status reflected the true value, the estimated and corresponding true values were plotted and a linear relationship fitted through the points. The performances of the ASPM and the alternative measures of stock status calculations were examined through their R-squared values.

Only the stock status in model region 1 (WCPFC-CA) is considered, ignoring model region 2, the rest of the Eastern Pacific Ocean.

6 Results

Summarising the results of fitting linear regression models to the ASPM estimated $SB/SB_{F=0}$ and the true $SB/SB_{F=0}$ from the projection results, using a relative $SB/SB_{F=0}$ results in better performance (Table 1). The R-squared is higher when relative to 2017-2019, than to 2000-2002. Using 2017-2019 excludes most of the historic data and this potentially suggests that the projected data is 'better behaved' than the historical data.

Table 1: Results of fitting linear regression models to the ASPM estimated $SB/SB_{F=0}$ and the true $SB/SB_{F=0}$ from the projection results for all eight $SB/SB_{F=0}$ metrics.

Metric	R-squared	Intercept	Slope
Final inst. SB/SBF=0	0.59	-0.13	1.01
Mean inst. SB/SBF=0	0.63	-0.13	1.02
Final latest SB/SBF=0	0.55	-0.01	0.81
Mean latest SB/SBF=0	0.67	-0.04	0.87
Mean inst. SB/SBF=0 rel. 2000-2002	0.92	-0.15	1.12
Mean inst. SB/SBF=0 rel. 2017-2019	0.97	-0.16	1.14
Mean latest SB/SBF=0 rel. 2000-2002	0.83	-0.03	0.92
Mean latest SB/SBF=0 rel. 2017-2019	0.93	0.03	0.89

6.1 Absolute $SB/SB_{F=0}$ metrics

The results of the non-relative metrics are shown in Figure 4. The 'stripes' for the historical estimates are the result of the projection models having different values of natural mortality but the same historical input data of catch and CPUE. The ASPM has a fixed value for natural mortality and therefore estimates simular values of $SB/SB_{F=0}$ for each projection model, giving rise to the 'stripes' (the difference between them arises from the observation error on the input data). This further illustrates the impact of the uncertainty in natural mortality on the performance of the ASPM.

The solid line shows the results of a linear regression and the diagonal dashed line has a slope of 1 and an intercept of 0. The dashed line is always higher than the solid line, indicating that the ASPM tends to overestimate the stock status. This may be a particular problem as the ASPM may not always detect when the stock is near the Limit Reference Point (LRP).

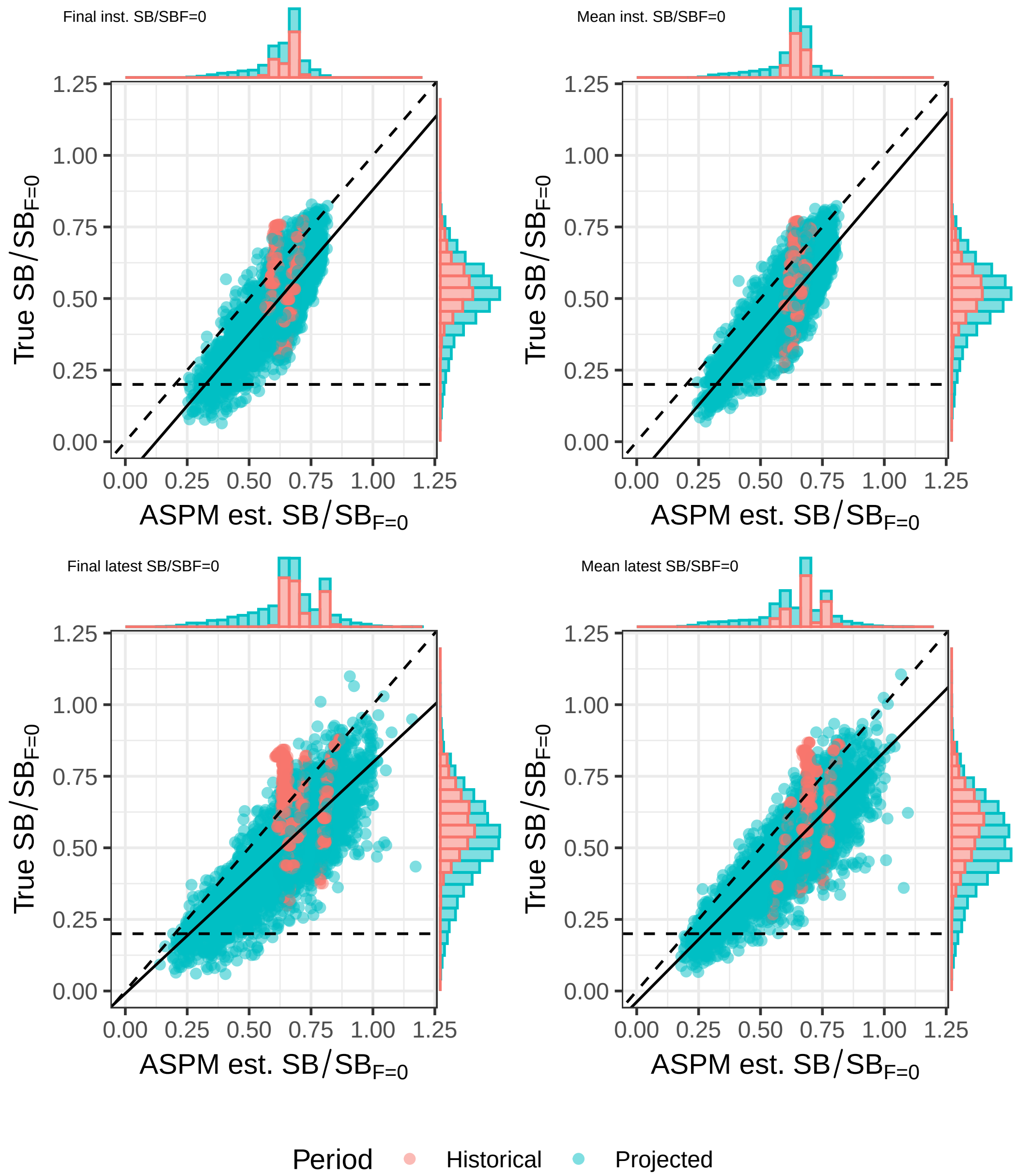

Figure 4: True SB/SB$_{F=0}$ from the projections against the ASPM estimated SB/SB$_{F=0}$ for two metrics: the final year value and the mean estimated value of the last three years and for two calculations of SB/SB$_{F=0}$: instantaneous and latest. Each panel shows the final estimate of the 6920 fits across the 740 projections. The points are coloured by time period (historical - up to 2022; projected - from 2023). The solid straight line is a linear regression. The dashed diagonal line has a slope of 1 and an intercept of 0 (i.e. perfect agreement between ASPM estimated values and the truth). The horizontal dashed line shows the Limit Reference Point of 0.2. Marginal histograms are shown for each model.

Investigating the residuals of the linear regression by true SB/SB$_{F=0}$ grouping (low, medium or high stock status) shows that using the mean value of SB/SB$_{F=0}$ rather than the final value performs slightly better for lower values (0 - 0.3) but the difference between the four metrics is small (Figure 5). The fits do not appear to be worse when the true $SB/SB_{F=0}$ is low, an important characteristic of an EM.

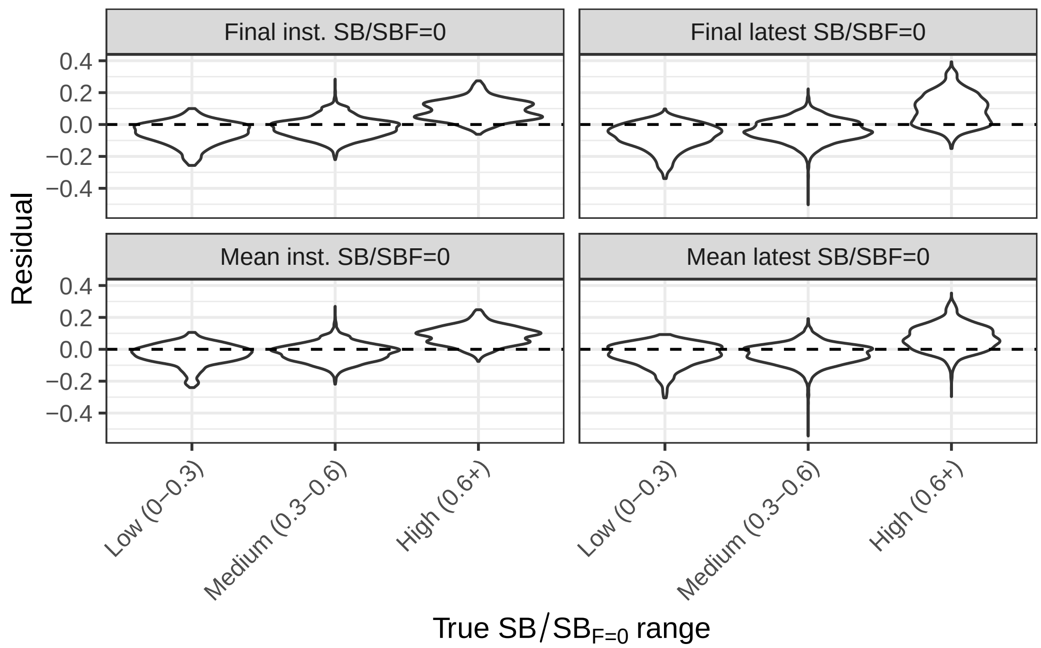

Figure 5: Violin plots of the residuals from fitting a linear regression of the ASPM estimated $SB/SB_{F=0}$ against the true $SB/SB_{F=0}$ for four different metrics of $SB/SB_{F=0}$ (terminal year & mean of last three years; instantaneous and latest). The residuals are placed into groups based on the range of true $SB/SB_{F=0}$ values.

6.2 Relative $SB/SB_{F=0}$ metrics

The ASPM performs better as an EM when using relative measures of $SB/SB_{F=0}$ (Figure 6). There is much less scatter than when using the absolute measures, and less bias. This demonstrates that using a relative measure may reduce the impact of uncertainty about natural mortality on the performance.

Two relative time periods are analysed here. A more recent time period, such as 2017-2019 may be preferred, but it is not possible to fully evaluate the EM's performance for the historical period. An example alternative reference period is chosen as 2000-2002 to illustrate the approach.

As 2017-2019 is the basis for the current interim SPA TRP it may be appropriate to use this period. The input to the HCR would effectively become relative to the interim TRP (by also including the 0.96 multiplier) which may make designing the HCR easier than using an alternative period.

The mean latest $SB/SB_{F=0}$ relative to 2017-2019 gives the least biased results, and for this reason may be preferred to the other metrics.

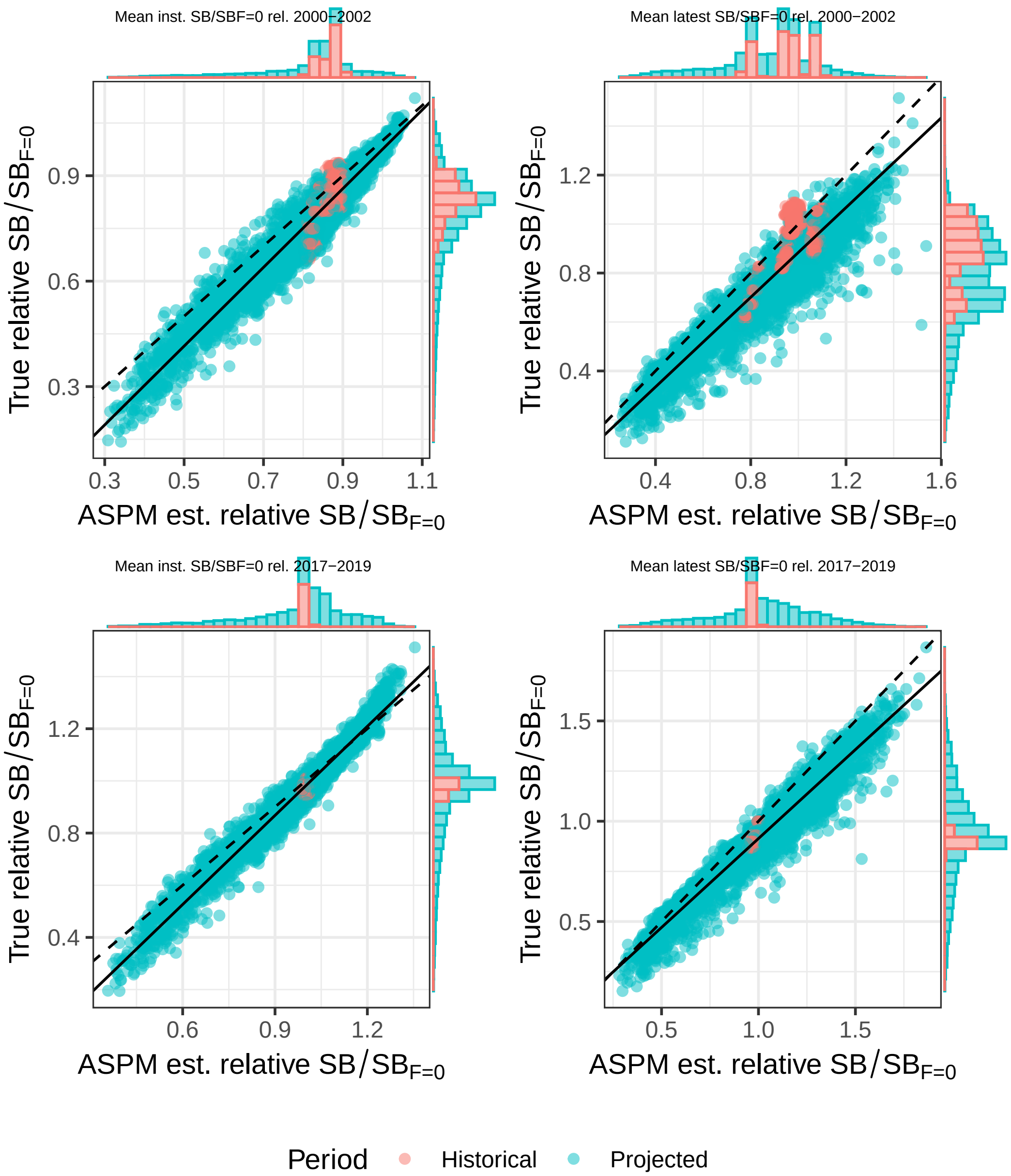

Figure 6: True relative $SB/SB_{F=0}$ from the projections against the ASPM estimated relative $SB/SB_{F=0}$ for two metrics: the mean estimated value in the last three years relative to the mean value in 2000-2002 and 2017-2019 for two calculations of $SB/SB_{F=0}$: instantaneous and latest. Each panel shows the final estimate of the 6920 fits across the 740 projections. The points are coloured by time period (historical - up to 2022; projected - from 2023). Note that there are fewer historical points when relative to 2017-2019. The solid straight line is a linear regression. The dashed diagonal line has a slope of 1 and an intercept of 0 (i.e. perfect agreement between ASPM estimated values and the truth).

Investigating the residuals of the linear regression by true $SB/SB_{F=0}$ grouping shows little difference between the different metrics (Figure 7). As with the absolute metrics, the fits do not appear to be significantly worse when the true $SB/SB_{F=0}$ is low, an important characteristic of an EM.

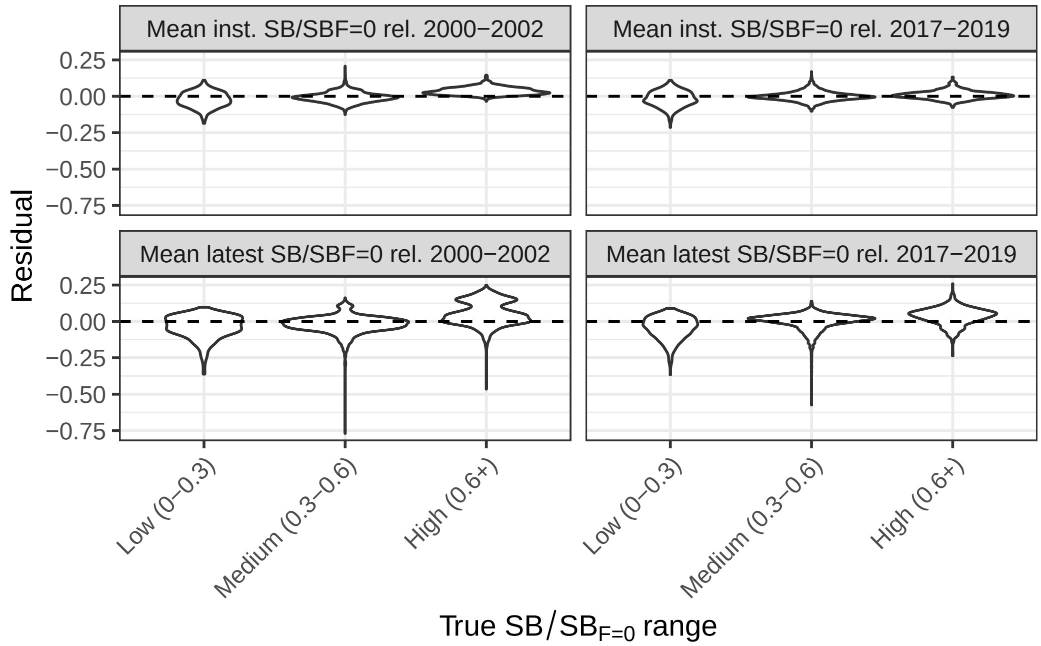

Figure 7: Violin plots of the residuals from fitting a linear regression of the ASPM estimated $SB/SB_{F=0}$ against the true $SB/SB_{F=0}$ for the relative metrics of $SB/SB_{F=0}$ (relative to 2000-2002 and 2017-2019; instantaneous and latest). The residuals are placed into groups based on the range of true mean $SB/SB_{F=0}$ of the last three years.

7 Discussion

The ASPM fitted well across all of the projected data sets, with a low failure rate of 6% and a high level of convergence.

Uncertainty in natural mortality has a strong impact on the estimated stock status. The ASPM uses a fixed level of natural mortality and this can result in high levels of bias when using estimates of absolute stock status such as $SB/SB_{F=0}$. While the trend in stock status over time appeared to be captured, estimated depletion was scaled by the value of M assumed. This impacts the performance of the ASPM as an EM.

The impact of natural mortality can be greatly reduced by estimating $SB/SB_{F=0}$ relative to some period in the past, for example $SB/SB_{F=0}$ relative to the $SB/SB_{F=0}$ in 2017-2019, the period used as a basis for the interim SPA TRP.

This is different to how previous example SPA HCRs and the HCR in the adopted SKJ MP work, and will need careful consideration and testing. If a relative measure of stock status is preferred, then the relative period will need to be decided.

If an absolute measure of stock status is preferred, then thorough testing of candidate MPs will be necessary to ensure their performance is not adversely affected by the limitations of the EM described here. Using the mean estimated stock status of the last few years may be preferred to using the stock status in the terminal year as the R-squared from the linear regression is higher.

All candidate MPs will be tested using the Manangement Strategy Evaluation framework and their relative performance compared using performance indicators.

Next steps:

- Develop candidate management procedures (MPs) using estimation methods based on both absolute and relative measures of stock status.
- The shapes of the harvest control rules in the candidate MPs can be adapted from those considered at WCPFC20.
- Test the candidate MPs using an updated Management Strategy Evaluation framework, using the updated operating model grid.
- Calculate and present performance indicators so that members may evaluate the relative performance of the candidate MPs and provide feedback.

8 Recommendations

SC20 is invited to:

- Consider the results of the estimation method trials for the South Pacific albacore management procedure.
- Endorse the use of the proposed ASPM as the estimation method for candidate management procedures.
- Endorse the use of relative and absolute depletion metrics for input into candidate South Pacific albacore harvest control rules.

Acknowledgments

We gratefully acknowledge funding for this work from the New Zealand Ministry of Foreign Affairs and Trade (MFAT). In addition we thank New Zealand eScience Infrastructure (NESI) for providing access to their computing resources.

References

Scott, F., Scott, R., Yao, N., and Natadra, R. (2023). Testing and developing estimation methods for South Pacific albacore. Technical Report WCPFC-SC19-2023/MI-IP-02, Koror, Palau, 16–24 August 2023.

Scott, R., Scott, F., Yao, N., Hamer, P., and Pilling, G. (2024). Selecting and Conditioning Operating Models for South Pacific Albacore. Technical Report WCPFC-SC20-2024/MI-WP-04, Manila, Philippines, 14–21 August 2024.

Teears, T., Peatman, T., Castillo-Jordan, C., Davies, N., Day, J., Hampton, J., Magnusson, A., Pilling, G., Xu, H., Vidal, T., Williams, P., and Hamer, P. (2024). Stock Assessment of South Pacific Albacore: 2024 - Rev.02. Technical Report WCPFC-SC20-2024/SA-WP-02, Manila, Philippines, 14–21 August 2024.