SCIENTIFIC COMMITTEE
NINETEENTH REGULAR SESSION
Koror, Palau
16–24 August 2023

Testing and developing estimation methods for South Pacific albacore

WCPFC-SC19-2023/MI-IP-02
26th July 2023

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

[1]Oceanic Fisheries Programme, The Pacific Community

Executive Summary

Two different candidate estimation methods for South Pacific albacore were investigated: a continuous-time surplus production model (SPiCT) and an Age-Structured Production Model (ASPM) implemented using Multifan-CL. These models were fit to data generated by performing stochastic projections to 2049 across the 72 models in the South Pacific albacore 2021 assessment grid. The future catches in the projections were set to a sine wave to provide contrast in the results. Recruitment variability was included in the projections. Observation error was included on both the historical and projected catch and effort data, and used as input data to the candidate estimation methods. Ten projections were run per stock assessment model, giving 720 projections. Of these, 114 projections failed due to stock collapse from over-fishing.

The input data were truncated to ten different final values, from 2004 to 2049 in increments of five years to give over 6000 fits per candidate estimation method, across a range of true $SB/SB_{F=0}$ values.

Six different variants of the SPiCT model were explored. However, the results suggest that none of them are as effective an estimation method as the ASPM. In particular, the relationship between the estimated biomass relative to the carrying capacity (estimated biomass / K) and the true $SB/SB_{F=0}$ from the projections showed a high degree of scatter. Additionally, the models' performances became noticeably worse at low values of true $SB/SB_{F=0}$.

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 three different metrics of $SB/SB_{F=0}$. The performance of the three metrics of $SB/SB_{F=0}$ was largely the same. The relationship was largely unaffected by the historical period but performed slightly worse for lower levels of true $SB/SB_{F=0}$. Additional investigation may be necessary to understand why the maximum gradient for some fits is larger than expected, despite the the estimated $SB/SB_{F=0}$ being close to the true value.

The results suggest that the ASPM performs better than SPiCT within these evaluations and may be an effective estimation method. More work may be needed to explore how the performance can be improved, and which metric of $SB/SB_{F=0}$ is preferred. At this stage it is worth using the ASPM as part of candidate management procedures to be tested.

1 Introduction

Under the harvest strategy approach, a management procedure has three essential elements: data collection, an estimation method and a harvest control rule (HCR). The job of the estimation method is to provide an estimate of current stock status to the HCR to define future fishing levels.

There are several important qualities of an effective estimation method that must be considered. The estimation method should work well for a range stock statuses. For example, an estimation method that only provides an accurate measure of stock status when the stock is healthy and around the target reference point (TRP) is of only limited use. It is particularly important that it is able to detect when the stock is declining or approaching the limit reference point (LRP) so that appropriate management action can be taken in a timely manner. Additionally, it is necessary to know if the stock status is higher than desired so that fishing opportunities can potentially be increased.

The estimation method should be robust to different sources of uncertainty. This includes being robust to sources of process error, such as variations in stock-recruitment processes, and observation uncertainty, such as in the collected catch and effort data.

Ideally, the estimation method should be as simple as possible, e.g. if a stock assessment model is being used, it should be simpler than that used to regularly assess the stock as part of the monitoring. However, very simple models may not adequately capture important components of stock dynamics, for example age or length structure of the stock, and spatial distribution and movement, which can result in poor performance. The complexity of the estimation method is therefore a balance that will vary between stocks and fisheries.

Estimation methods can be broadly split into two types: empirical, which may use direct observations thought to imply current stock status, such as catch per unit of effort (CPUE) and model-based, which use the collected data to estimate the current stock status through estimation, such as a stock-assessment model. Initial work to develop the management procedures for SPA focused on developing empirical procedures that used longline catch per unit effort (CPUE) as the primary indicator of stock status, consistent with the noted objectives. Results of those preliminary analyses highlighted the difficulty of using CPUE as the primary measure of stock status through empirical methods. SC17 supported the continued investigation of simple model-based alternatives that could address some of these problems. While these model-based approaches also depend on CPUE to estimate stock status, they typically provide more reliable and more stable estimates through the use of additional information (e.g. catches) (Scott et al., 2019; Yao et al., 2019).

In this report we explore two model-based approaches as candidate estimation methods:

- SPiCT (a continuous-time surplus production model)
- An Age-Structured Production Model (ASPM) implemented using Multifan-CL

WCPO tuna stocks are managed relative to the key levels of stock status of the Limit Reference

Point (LRP) and the Target Reference Point (TRP), that are typically defined using $SB/SB_{F=0}$. An MP is designed to keep the stock status around the TRP and have a high probability of avoiding the LRP. One way of measuring how well an estimation method is performing is to compare the estimated stock status to the true $SB/SB_{F=0}$. Ideally, while the estimated stock status does not need to exactly match the 'true' $SB/SB_{F=0}$, the estimation method should be a good predictor of it and there should be a strong linear relationship between them.

In this report the two candidate estimation methods are 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 HCR design, it does makes the HCR design process less clear.

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

2 Generating test data

Stochastic projections based on the 2021 South Pacific albacore stock assessment grid were used to generate time series data on which to fit the candidate estimation models and evaluate their performance (Castillo Jordan et al., 2021). The 30 year projections start in 2020 and continue until 2049.

During the first three years of the projection the future catches for each fishery were set as a mean of the 2017-2019 catches. Although it would have been possible to set the catches from 2020-2022 to their observed values, it would not affect the performance of the candidate estimation method. To introduce contrast in the projected values, from 2023 the catches were set as a sine wave, with a mean value equal to the 2017-2019 average, 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).

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.

Observation error was included on the catch and effort for both the historical and projected periods, with a coefficient of variation equal to 0.25 for each fishery (Appendix B). The observed catch and effort were used to generate input data for the candidate estimation models (CPUE data for the index fisheries, catches for the non-index fisheries).

Each of the 72 models in the grid was projected ten times giving 720 projections. Of these, 116 projections (16%) failed due to the stock biomass falling too low, leaving 606 projections.

The candidate estimation methods were fitted to each of the 606 projections. The time series of input data were truncated to different final years, from 2004 to 2049 in increments of five years, giving ten final years: four in the historical period (2004 to 2019) and six in the projected period (2024 to 2049). Each estimation method was therefore fitted 606 x 10 = 6060 times, giving 6060 estimates of biomass in the final year.

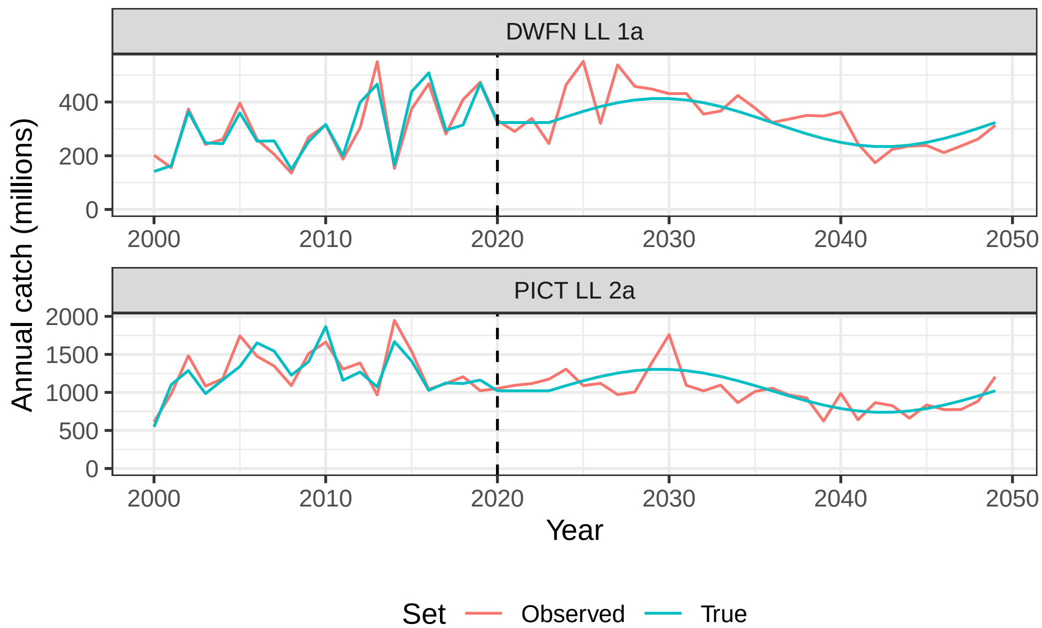

Figure 1: Observed and ‘true’ annual catches of two example fisheries, DWFN longline in model region 1 and PICT longline in model region 2, from a single projection. Each projection has different observed catches. Future catches in the first three years are at the 2017-2019 average, and then from 2023 as a sine wave, with a mean value equal to the 2017-2019 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 2020, the first year of the projection.

3 Candidate estimation models

3.1 SPiCT

SPiCT is a surplus production model in continuous time implemented in R (Pedersen and Berg, 2017). Surplus production models are relatively simple in that they consider only the total biomass of the stock and do not consider the age or size structure. Additionally, there is no regional or spatial structure.

The input data for SPiCT is the total catch in weight, plus one or more indices of catch per unit of effort (CPUE). Here, the input data was annualised. Model region 4 (the Eastern Pacific Ocean) in the projections is not considered and only model regions 1 to 3 are included. Only the total catches in regions 1 to 3, including observation error, are used. The observed CPUE, calculated from the catch and effort with observation error, of the index fisheries in regions 1 or 2 of the projection model are used to generate the CPUE for the SPiCT model.

The outputs from SPiCT include the total biomass through time, as well as estimated parameters. It is not possible to calculate $SB/SB_{F=0}$ from the output as there is no age or maturity data meaning it is not possible to directly compare the outputs of SPiCT to the TRP and LRP. Here, the resulting total biomass is scaled by the estimated carrying capacity of the stock, K, to give an equivalent total biomass relative to the total biomass with no fishing ($TB/TB_{F=0}$). This is considered as a potential input to the HCR.

Six alternative SPiCT models were tested, with different index fisheries providing the CPUE (either from model region 1 or 2) and options for fixing certain parameters (Table 1). Up to two parameters are fixed:

- The shape of the production curve, n, which when fixed is set at 2 gives a Schaefer production model;
- The carrying capacity, K, which when fixed is set at the mean total biomass with no fishing ($TB_{F=0}$) as estimated by the diagnostic stock assessment model.

All other model parameters are freely estimated.

Table 1: Estimation method options for SPiCT.

Model	CPUE index	n	K
Model 1	Index fishery 2	Freely estimated	Freely estimated
Model 2	Index fishery 2	Fixed at 2	Freely estimated
Model 3	Index fishery 2	Fixed at 2	Fixed at mean $TB_{F=0}$ from the diagnostic case
Model 4	Index fishery 1	Freely estimated	Freely estimated
Model 5	Index fishery 1	Fixed at 2	Freely estimated
Model 6	Index fishery 1	Fixed at 2	Fixed at mean $TB_{F=0}$ from the diagnostic case

Each of the six SPiCT models was fitted ten times to each of the 606 projections, with each fit having a different final year. Of these 6060 fits, 62 of them failed to converge and were dropped from the results.

As mentioned above, although SPiCT estimates a full time series of biomass, we are only interested in generating an input to the HCR. Therefore, only the final estimated value of biomass is considered here. By comparing the final estimates of biomass estimated by SPiCT, scaled by K, to the 'true' $SB/SB_{F=0}$ in model regions 1 to 3 in the projections it is possible to evaluate how well the SPiCT models perform as estimation methods.

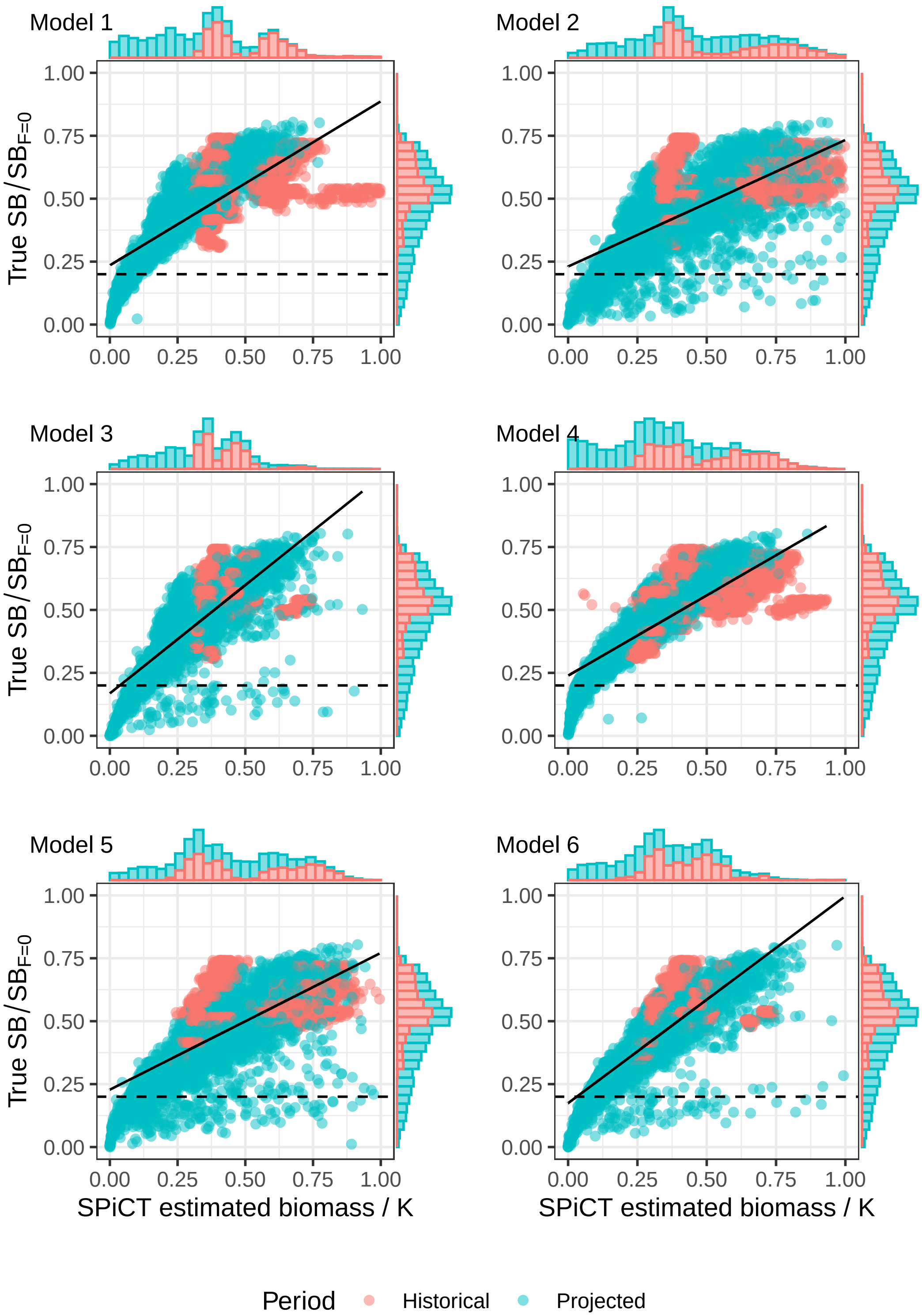

Figure 2: True SB/SB$_{\mathrm{F=0}}$ from the projections against the SPiCT estimated biomass / K for each of the six SPiCT models. Each panel shows the final estimate of the 6060 fits across the 606 projections. The points are coloured by time period (historical - up to 2019; projected - from 2024). The solid straight line is a linear regression. The horizontal dashed line shows the Limit Reference Point of 0.2. Marginal histograms are shown for each model.

As mentioned above, to be an effective estimation method the SPiCT estimate of biomass / K should be a good predictor of the true $SB/SB_{F=0}$. Each of the SPiCT models show high levels of scatter, particularly for the historical period, suggesting that none of the SPiCT models tested here would be an effective estimation method (Figure 2).

It is expected that the model fits for the projected period are better than for the historical period, as the input data is model generated and possibly 'better behaved'. However, it is important that the estimation method works well for the historical data as this is expected to be more representative of what the real future data will be like.

As mentioned above, ideally the relationship between the SPiCT estimate of biomass / K and the true $SB/SB_{F=0}$ should be linear. To explore this a linear regression was fit to the results for each model (Table 2). All of the models have R-squared values below 0.7. Models 2 and 5 (fixing n but not fixing K) have the lowest R-squared value and noticeable scatter can be seen in Figure 2. Models 1, 3, 4 and 6 have better R-squared values but a show a strong degree of non-linearity for lower values of true $SB/SB_{F=0}$.

The residuals of each linear regression can be explored by groupings of different ranges of true $SB/SB_{F=0}$ (split into three groups: 0-0.3, 0.3-0.6, 0.6+) (Figure 3). The residuals are noticeably worse for the lower range of true $SB/SB_{F=0}$, particularly for models 3 and 6. This is a potential problem as this is when when you most want your estimation method to perform well and have a clear relationship between the estimated and the true stock status. Models 1 and 4 perform reasonably well but it is difficult to recommend either of them as an estimation method.

Although a model-based estimation method should be as simple as possible, SPiCT has no spatial, age or length structure. These simplifications may be too much for it to be an effective estimation method for South Pacific albacore, particularly given the spatial structure of the stock.

Table 2: Results of fitting linear regression of the SPiCT estimated biomass / K against the true $SB/SB_{F=0}$ from the projection results.

SPiCT model	Intercept	Slope	R-squared
Model 1	0.24	0.65	0.63
Model 2	0.23	0.50	0.49
Model 3	0.17	0.86	0.62
Model 4	0.24	0.64	0.69
Model 5	0.23	0.54	0.51
Model 6	0.17	0.82	0.68

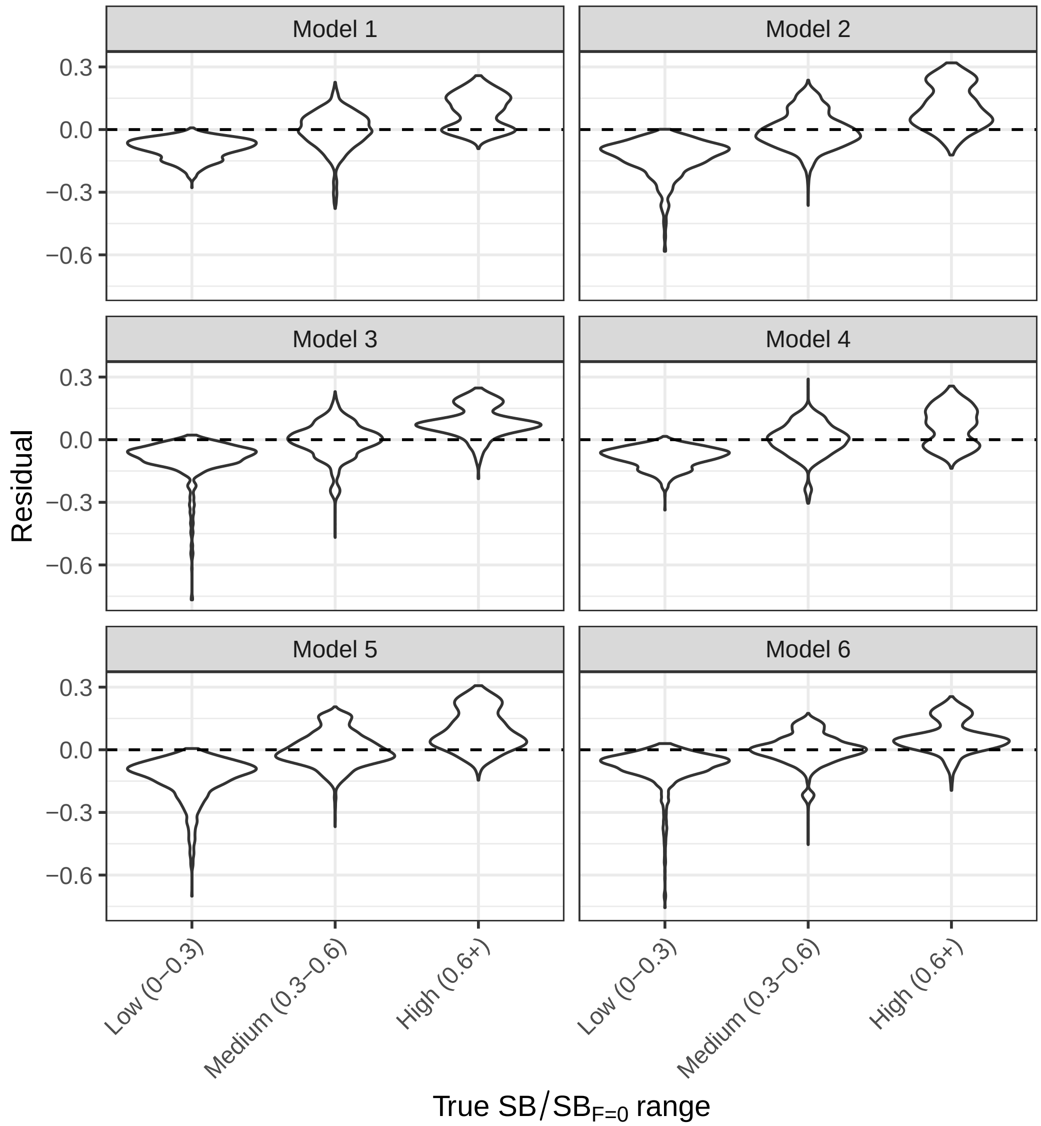

Figure 3: Violin plots of the residuals from fitting a linear regression of the SPiCT estimated biomass / K against the true SB/SB$_{F=0}$. The residuals are placed into groups based on the range of true SB/SB$_{F=0}$ values.

3.2 Multifan-CL Age-Structured Production Model

An Age-Structured Production Model (ASPM) includes the age-structure of the population instead of just the accumulated biomass (as with a surplus production model like SPiCT). As such, they can include greater complexity in the stock and fishery dynamics, such as maturity at age and selectivity. ASPMs can come in many different flavours, with differing levels of complexity. The Multifan-CL ASPM is based on the diagnostic 2021 stock assessment for South Pacific albacore. An ASPM does not use catch size distribution data, only catch and effort data. The regional and fishery structure of the ASPM is the same as the 2021 assessment, i.e. four model regions, 21 fisheries and four index fisheries. Various parameters are fixed to the same value as the diagnostic case: steepness of the stock-recruitment relationship (fixed at 0.8), movement rates, growth and natural mortality, selectivity and recruitment distribution (Castillo Jordan et al., 2021).

The model was fitted in several phases, with each phase fixing more parameters, similar to fitting a Multifan-CL stock assessment. The final phase is run for 20,000 evaluations or until the maximum gradient is less than 0.001. Estimated outputs values include the spawning biomass (SB), the spawning biomass in the absence of fishing ($SB_{F=0}$) and the ratio between them ($SB/SB_{F=0}$).

As with the testing of the SPiCT model, the ASPM was fitted to each of the 606 projections. The time series of input data was truncated to 2004 to 2049 in increments of five years, giving ten final years: four in the historical period (2004 to 2019) and six in the projected period (2024 to 2049). The same level of observation error on the catches (for the non-index fisheries) and CPUE (for the index fisheries) was used as with the SPiCT model (Appendix B). The ASPM was therefore fitted 606 x 10 = 6060 times. The final estimates of $SB/SB_{F=0}$ estimated by ASPM were compared to the 'true' $SB/SB_{F=0}$ values to evaluate how well the ASPM performs as an estimation method.

Three different metrics of $SB/SB_{F=0}$ were explored as inputs to the HCR: instantaneous ($SB_y/SB_{F=0,y}$), 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) and the mean instantaneous $SB/SB_{F=0}$ of the last three years. The ASPM estimated $SB/SB_{F=0}$ was compared to the equivalent 'true' value in the projections for each of the different metrics. Here, only the $SB/SB_{F=0}$ in model regions 1 to 3 are considered, ignoring model region 4, the Eastern Pacific Ocean.

Out of the 6060 model fits, only 1000 finished with maximum gradients less than 0.01, and 280 with maximum gradients greater than 100, suggesting that many of the models had not converged sufficiently. Increasing the number of evaluations in the final phase to 50,000 evaluations improved the convergence, with 3000 fits having maximum gradients of less than 0.01, and 100 greater than 100. However, the estimated $SB/SB_{F=0}$ from the 20,000 and 50,000 evaluation fits were almost identical suggesting that simply increasing the number of evaluations and further reducing the maximum gradients does not necessarily change the estimated value.

Appendix A compares two fits of the ASPM from the same underlying projection model, with contrasting maximum gradients of 133 and 0.000564. Despite the difference in maximum gradients, each fit estimates $SB/SB_{F=0}$ equally well, suggesting that the maximum gradient alone is not a good guide to how well the model has fit. This is an area of work that needs to be explored further.

The relationships between the ASPM estimated $SB/SB_{F=0}$ and the 'true' projected $SB/SB_{F=0}$ for the three different $SB/SB_{F=0}$ metrics are linear. Some scatter is evident, particularly for higher values of $SB/SB_{F=0}$ (Figure 4). The performance of the three $SB/SB_{F=0}$ metrics is broadly the same.

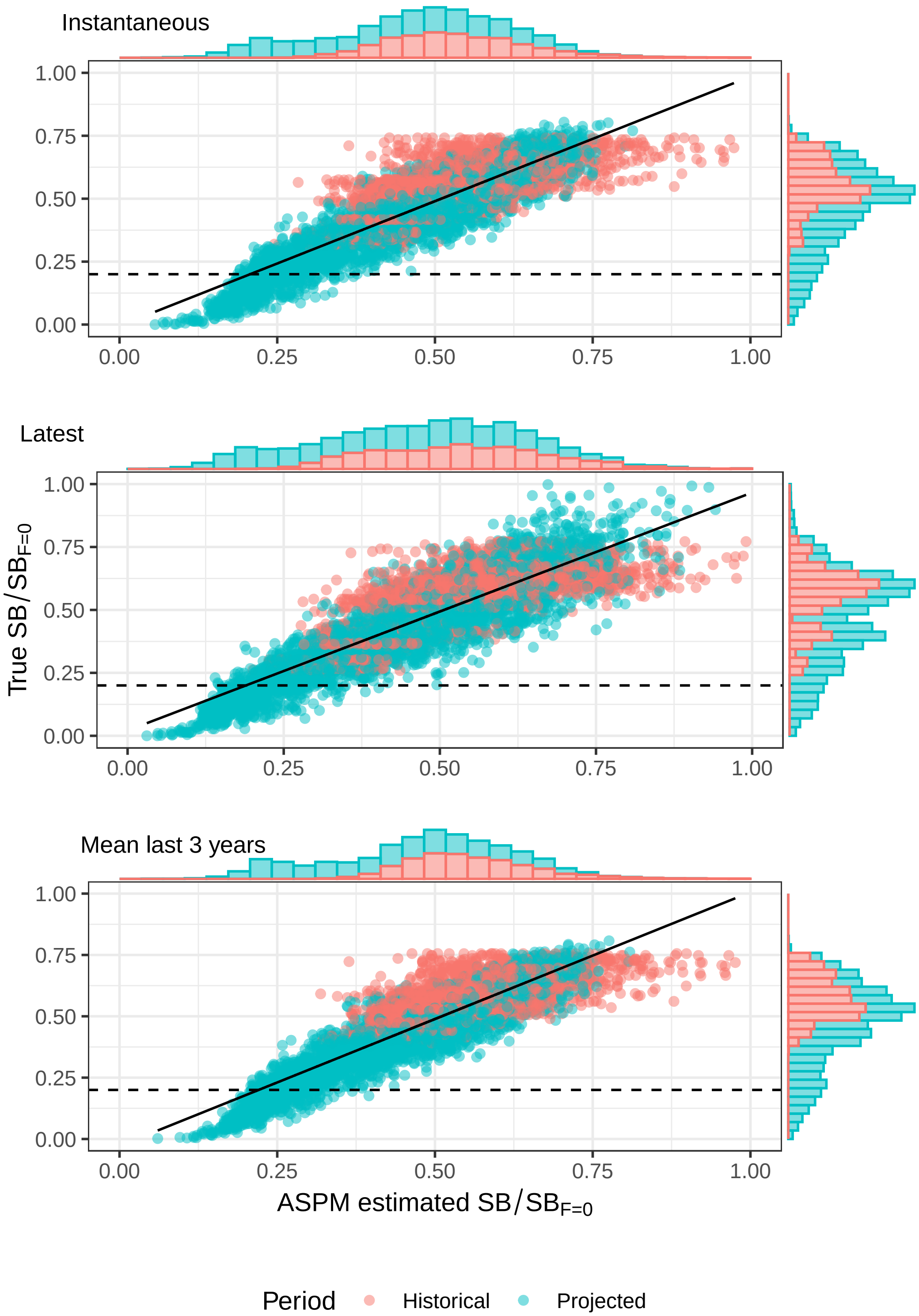

Figure 4: True SB/SB$_{F=0}$ from the projections against the ASPM estimated SB/SB$_{F=0}$. Three different metrics for SB/SB$_{F=0}$ are tested: instantaneous, latest and the mean of the last three years. The plot shows the final estimate of the 6060 fits across the 606 projections. The points are coloured by time period (historical - up to 2019; projected - from 2024). The solid straight line is a linear regression. The horizontal dashed line shows the Limit Reference Point of 0.2. Marginal histograms are shown.

Fitting a linear regression model to the estimated and 'true' values of $SB/SB_{F=0}$ results in R-squared values close to 0.8, an intercept close to 0 and a slope close to 1 for each metric (Table 3).

Table 3: Results of fitting linear regression model to the ASPM estimated $SB/SB_{F=0}$ and the true $SB/SB_{F=0}$ from the projection results for the three different $SB/SB_{F=0}$ metrics.

SB/SBF=0 metric	Intercept	Slope	R-squared
Instantaneous	-0.01	0.99	0.77
Latest	0.02	0.95	0.77
Mean last 3 years	-0.03	1.03	0.79

Investigating the residuals of the linear regression by true $SB/SB_{F=0}$ grouping shows that the estimation method performs slightly worse for low values of $SB/SB_{F=0}$ for all three metrics of $SB/SB_{F=0}$ (Figure 5). Very low values of true $SB/SB_{F=0}$, i.e. those less than 0.125, tend to be overestimated by the ASPM model (Figure 4). Although this is undesirable behaviour for an estimation model, the bias in the residuals is small and can be accounted for in the design of the HCR. Additionally, if the management procedure is doing its job properly then the true stock size would have only a small probability of falling to low values of $SB/SB_{F=0}$.

Investigating the residuals by time period (either historical or projected) shows that the ASPM model performs equally well across time periods (Figure 6). As mentioned above, it is important that the estimation method works well for the historical data as this is expected to be more representative of what the real future data will be like.

These results suggest that the ASPM may be an effective estimation method and is worth testing further in candidate management procedures (Scott et al., 2023). Additionally, any of the three metrics of $SB/SB_{F=0}$ could be considered as an effective input to the HCR.

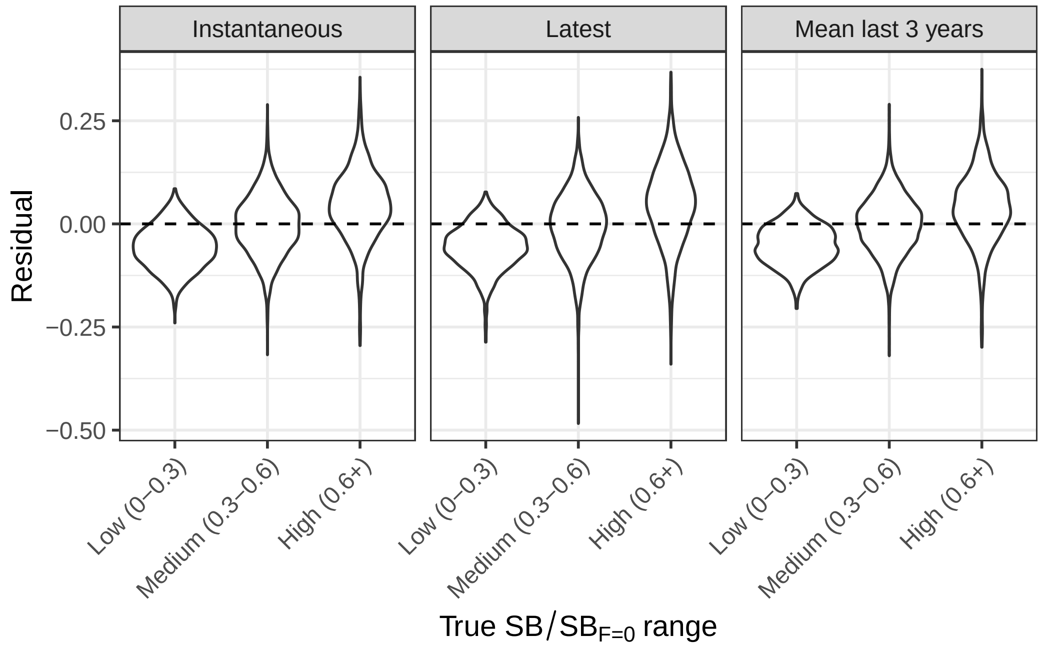

Figure 5: Violin plots of the residuals from fitting a linear regression of the ASPM estimated $SB/SB_{F=0}$ against the true $SB/SB_{F=0}$ for three different metrics of $SB/SB_{F=0}$. The residuals are placed into groups based on the range of true $SB/SB_{F=0}$ values.

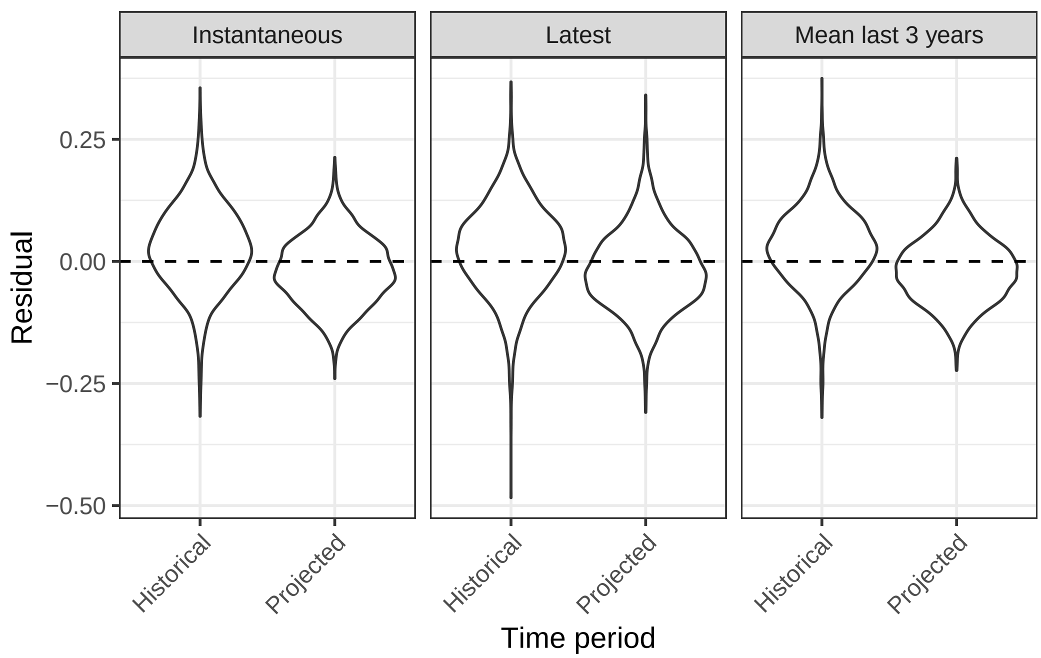

Figure 6: Violin plots of the residuals from fitting a linear regression of the ASPM estimated $SB/SB_{F=0}$ against the true $SB/SB_{F=0}$ for three different metrics of $SB/SB_{F=0}$. The residuals are placed into groups based on the time period of the final year in the input data (historical 2004-2019; projected 2020-2049).

Acknowledgments

We gratefully acknowledge funding for this work from the New Zealand Ministry of Foreign Affairs and Trade (MFAT) funded project "Pacific Tuna Management Strategy Evaluation". In addition we thank the Center for High Throughput Computing (CHTC UW-Madison) for generously providing access to their computing resources.

A ASPM Exploratory diagnostics

This appendix explores the results of two fits, using two iterations (6 and 10) from the same projections model (S1M1D1R1G1), truncated to end in 2019, i.e. only considering the historical data. The only difference between the projection results in historical period is from the observation error. The ASPM is fitted to both iterations. The maximum gradient for the two iterations are 133.3 and 0.000564 respectively, i.e. according to the maximum gradient iteration 6 has fitted poorly.

The two iterations appear to estimate the projected 'true' $SB/SB_{F=0}$ equally well (Figure 7) so the maximum gradient does not appear to be a good guide as to how well the ASPM estimates $SB/SB_{F=0}$.

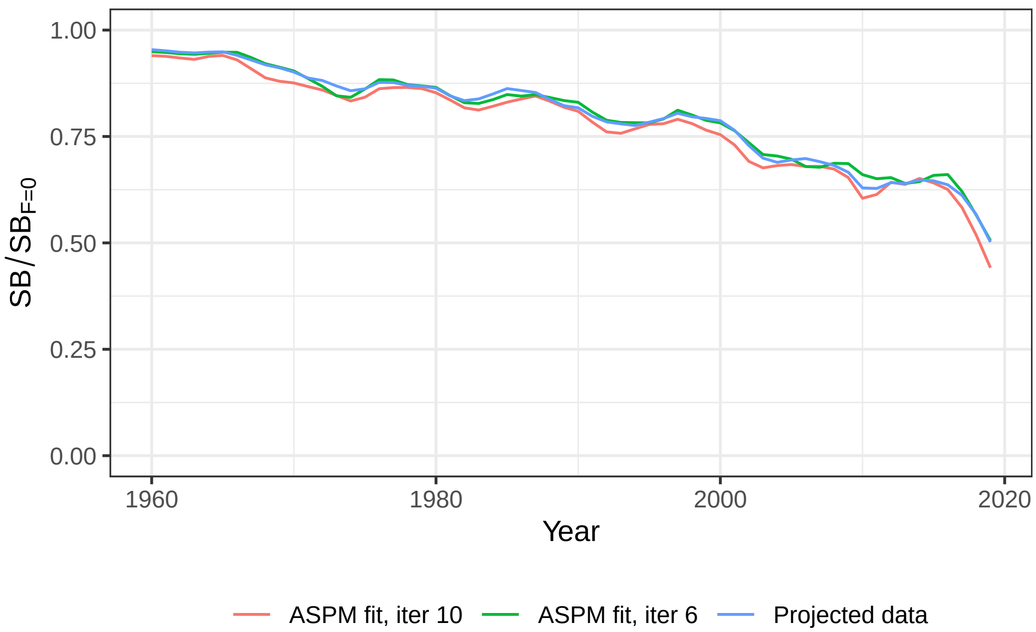

Figure 7: The projected 'true' $SB/SB_{F=0}$ and two iterations of $SB/SB_{F=0}$ as estimated by the ASPM. The difference between the iterations is only observation error applied to the input catch and effort data. Iteration 6 has a maximum gradient of 133; iteration 10 has a maximum gradient of 0.000564.

The effort devs of the index fisheries of each fit show several outlying values for each fishery in iteration 6, but only one outlying value in regions 2, 3 and 4 in iteration 10 (Figure 8). This is probably what is driving the poor maximum gradient for iteration 6 despite the $SB/SB_{F=0}$ being well estimated.

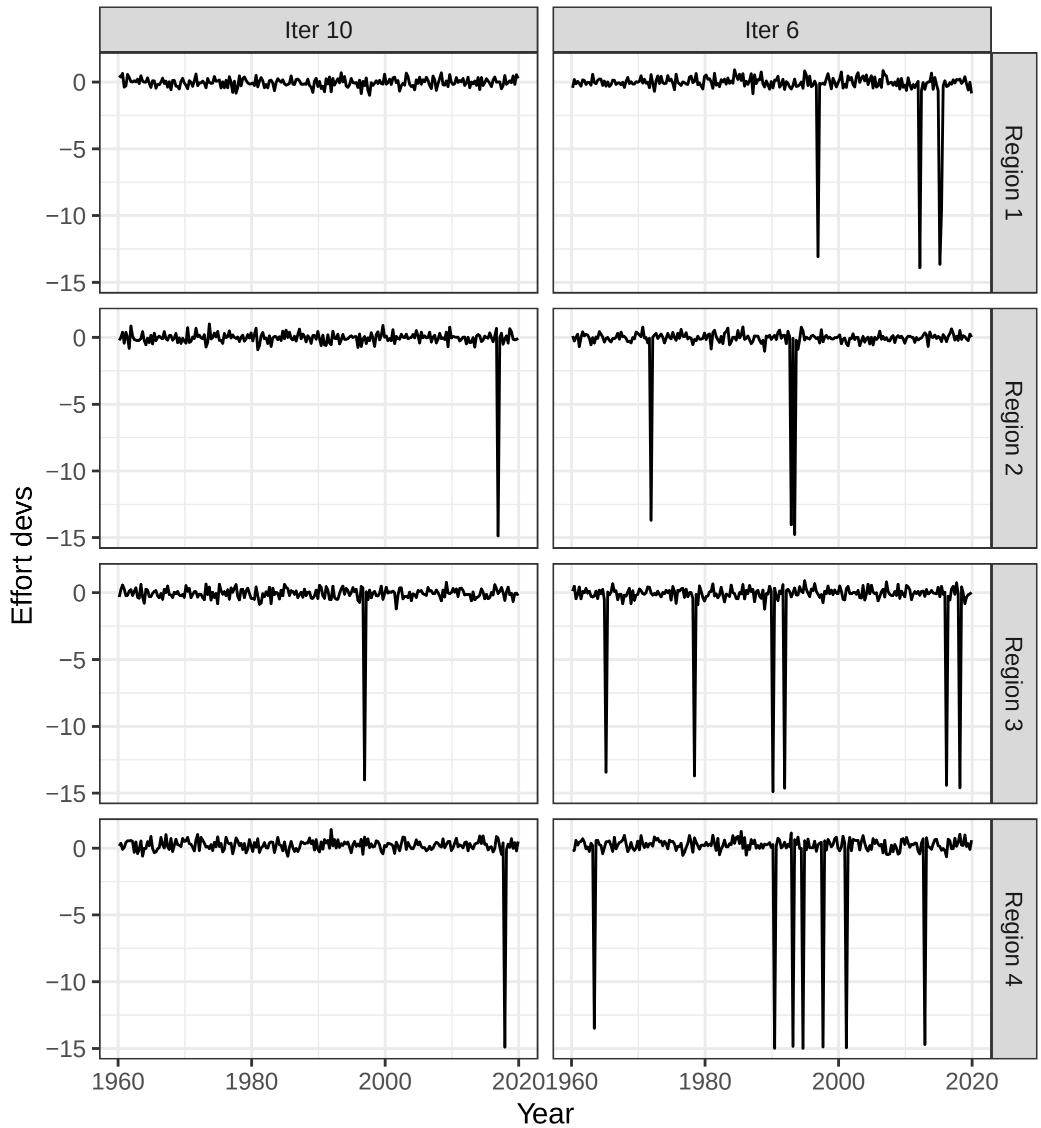

Figure 8: Effort devs of the index fisheries in model regions 1-4 for two ASPM iterations from the same projection. The difference between the iterations is only observation error applied to the input catch and effort data. Iteration 6 has a maximum gradient of 133; iteration 10 has a maximum gradient of 0.000564.

B Estimating observation error

This appendix discusses the level of observation error to be applied to catch and effort of the fisheries in the projection model for the purposes of generating input data for the candidate estimation methods (CPUE for the index fisheries, catches from the non-index fisheries).

Observation error is and an inevitable component of the data collection process, It can be characterised as the difference between the 'true' value of the data and that which was recorded and used for analysis, including the difference between 'true' catches taken and fishing effort of a fishing operation, and that which was reported. It is important that a management procedure (MP), including the the estimation method, is robust to observation error.

When performing projections for Management Strategy Evaluation (MSE), observation error should be simulated and included in the data used by the estimation method. The level of observation error in the simulations should correspond to that seen in the real world. However, an appropriate level of observation error is hard to estimate. It is not necessary for the exact level of error to be known but it should be sufficient to generate confidence that the candidate MPs have been adequately tested against this source of uncertainty.

In the projections, observation error is included on the fishery specific quarterly catch and effort through the application of log-normal error based on a user defined coefficient of variation (c.v.) (Davies et al., 2018). Separate coefficients can be applied for either catch or effort but cannot currently be applied to individual fisheries. In other words the catch c.v. applies to all catch projected fisheries and the effort c.v. to all effort projected fisheries. It is possible to specify different c.v.s for catch and effort. However, for simplicity, here the c.v.s are kept the same.

For projections, ideally the observation error applied in the future should be similar to that seen in the past. To estimate an appropriate level of observation error stochastic projections were performed using the diagnostic 2021 South Albacore stock assessment model. The projections were run until 2049 with future catches fixed at 3/4 of the 2017-2019 level and included recruitment variability and observation error in the future.

The historical observation error was taken to be the difference between the historical catch and effort data used to fit the model (taken to be the observed data) and those estimated internally by the model (taken to be the 'true' data). The future observation error was calculated as the difference between the future catch and effort specified to run the projection (taken to be the true data), and the simulated observed catch and effort.

The difference between the observed and 'true' catch per unit of effort (CPUE) of the index fisheries was calculated, i.e. the residuals in both the historical and projected periods. The c.v. of the residuals in the recent historical period (1990 - 2019) was compared to that in the projected period (2020 onward) for different levels of observation error. If the c.v. of the residuals in the projected and historical periods are similar, then it can be assumed that the level of observation error is the same in both periods.

Different c.v.s of future catch and effort error were tested. The projections were run 100 times for each candidate level of observation error (Figure 9). The recruitment variability differed between iterations, but was kept the same for each level of observation error meaning that the results of different levels of observation error are directly comparable. The historical observed and true data were the same for each iteration and level of observation error.

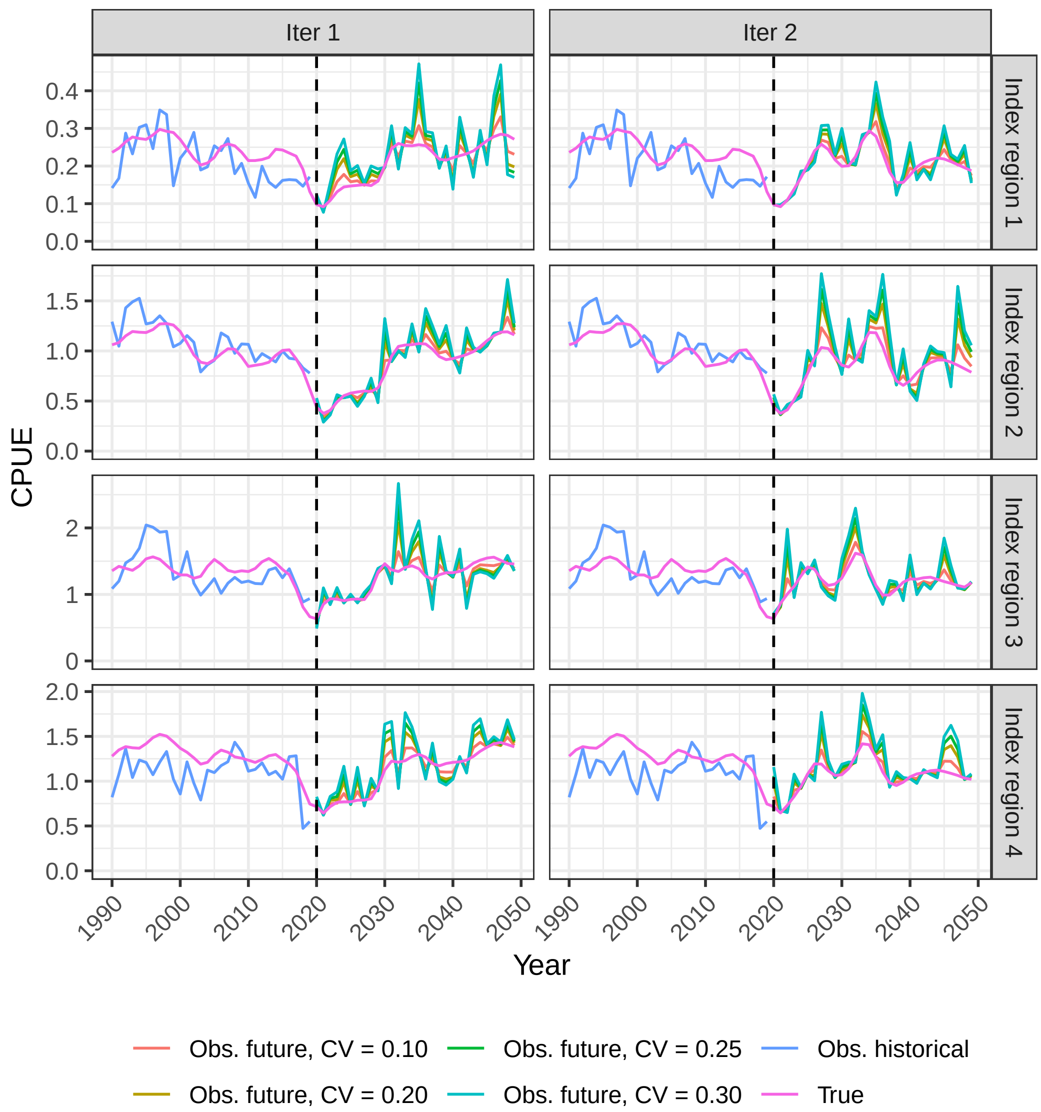

Figure 9: Catch per unit effort (CPUE) from two example projections for the four index fisherires. The dashed vertical line seperates the projected and historical periods. Observed CPUE in the future period has four different levels of observation error. The difference between the true data and the observed data is the observation error.

The c.v. of the residuals in the projected period was calculated for each of the 100 projections and for each level of observation error. The distribution of the c.v. of the projected residuals was compared to that in the past (Figure 10). An observation error c.v. of 0.25 results in a c.v. of the residuals in the projected period being the same or slightly above that seen in the historical period for index fisheries in regions 2, 3 and 4, but slightly lower for index fishery in region 1. This level of observation error for the catch and effort in the projections is therefore thought to give equivalent levels of future observation error as seen in the past. The other tested levels of observation error result in the c.v. of the residuals in the projected period being either too low (0.10 and 0.20) or too high (0.30),

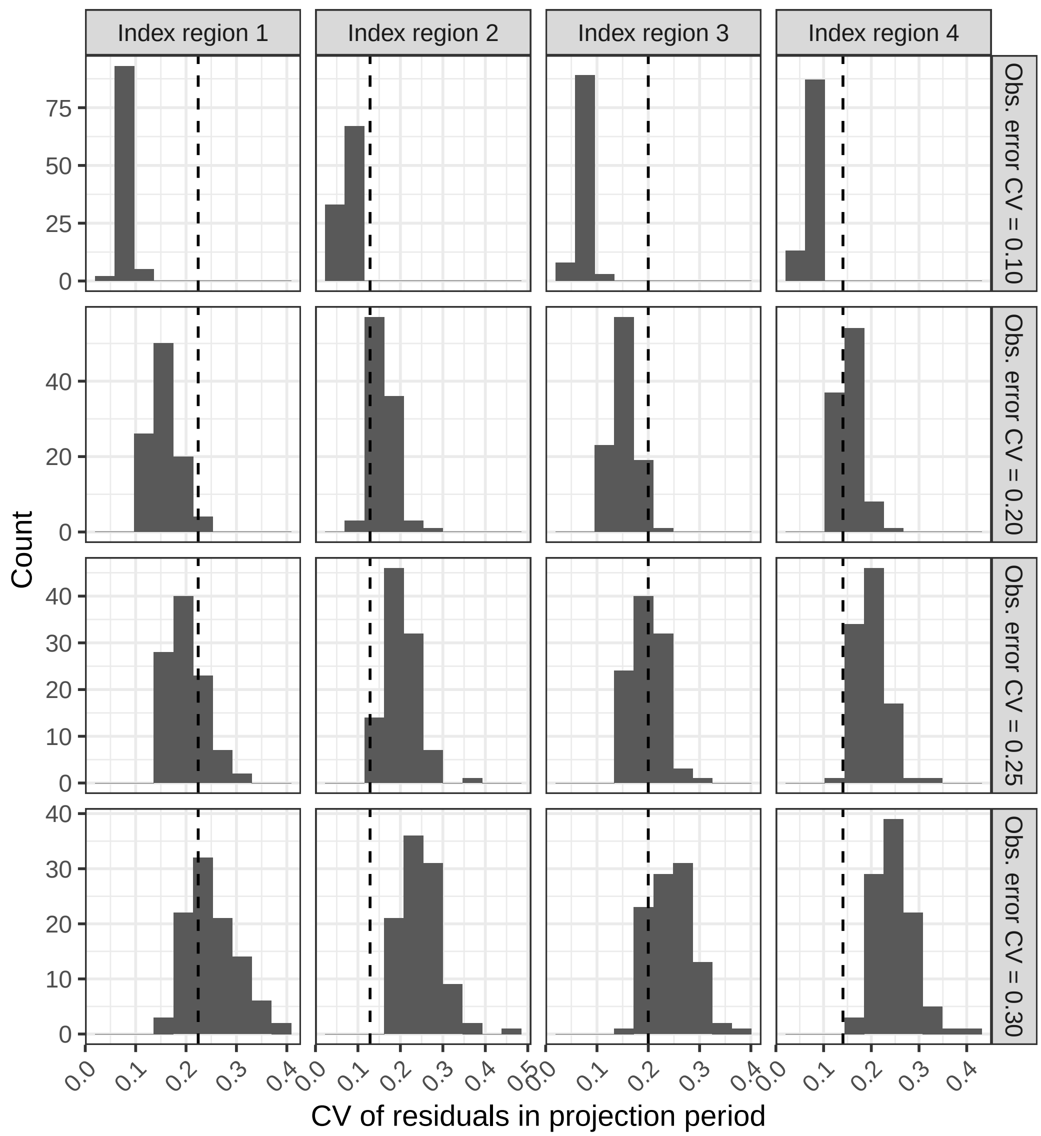

Figure 10: Histograms of the coefficient of variation (CV) of the residuals in the projected period for each index fishery and level of observation error. The dashed vertical line is the equivalent CV of the residuals in the recent historical period.

References

Castillo Jordan, C., Hampton, S., Ducharme-Barth, N. Xu, H., Vidal, T., Williams, P., Scott, F., Pilling, G., and Hamer, P. (2021). Stock assessment of South Pacific albacore tuna. Technical Report WCPFC-SC17-2021/SA-WP-02, 11–19 August 2021.

Davies, N., Fournier, D., Takeuchi, Y., Bouye, F., and Hampton, S. (2018). Developments in the MULTIFAN-CL Software 2017-18. Technical Report WCPFC-SC14-2018/SA-IP-02, Busan, South Korea, 5–13 August 2018.

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

Scott, F., Scott, R., Yao, N., and Natadra, R. (2023). Evaluation of candidate management procedures for South Pacific albacore. Technical Report WCPFC-SC19-2023/MI-WP-06, Koror, Palau, 16–24 August 2022.

Scott, R., Yao, N., Scott, F., and Pilling, G. (2019). South Pacific albacore management strategy evaluation framework. Technical Report WCPFC-SC15-2019/MI-WP-08, Pohnpei, Federated States of Micronesia, 12–20 August 2019.

Yao, N., Scott, F., Scott, R., Pilling, G., and Hampton, S. (2019). CPUE analyses for South Pacific albacore. Technical Report WCPFC-SC15-2019/MI-WP-07, Pohnpei, Federated States of Micronesia, 12–20 August 2019.