SCIENTIFIC COMMITTEE
SIXTEENTH REGULAR SESSION
Online
11–20 August 2020

Developing Management Procedures for WCPO Skipjack: The Estimation Model

---

WCPFC-SC16-2020/MI-IP-09

R. Scott[1], F. Scott, G. M. Pilling and J. Hampton

[1]Oceanic Fisheries Programme, The Pacific Community

Contents

Executive Summary

The estimation method (EM) is an important component of the management procedure (MP). Its role is to provide a reliable estimate of stock status that can be used by the harvest control rule (HCR) to determine future fishing opportunities. The MP can be classed as either empirical, employing an EM based on direct observation of raw data, or model based where, for example, the EM is based on a stock assessment model. WCPFC-SC has recommended that, in the first instance, model based MPs should be investigated for the WCPO skipjack evaluations.

The WCPO skipjack assessment relies on a large quantity of tag release and recapture information which limits the choice of stock assessment model that can be used as the EM. MULTIFAN-CL is perhaps the most effective model to use but, similar to other options, a full assessment can be slow to run and not practical within a simulation framework. For this reason a simplified model is employed that approximates the full assessment. The simplified model must run in a tractable time period but must also provide a reliable estimate of stock status.

We present a series of tests conducted to investigate the performance of the simplified EM specifically with respect to the following three questions.

1. Can the simplified model closely approximate the full assessment model?
2. Can the simplified model provide a reliable estimate of depletion even when the model settings of the OM and EM differ?
3. Does the simplified model perform adequately within the simulations?

These tests included simple comparisons between the EM estimates and those of a full stock assessment which showed that, given the same input data and similar model settings, the EM produces estimates of depletion that are very close to those of the full assessment. Further comparisons between the projected stock status from the Operating Model and estimated stock status from the EM were conducted to illustrate the potential difference in estimates of depletion resulting from model misspecification. For the base case scenario, for which the model settings for the OM and EM are the same, the EM estimates of depletion corresponded well with those of the OM projection. For scenarios where the OM settings differ from the EM settings, the EM estimates, although biased, continue to track the overall trend in depletion and continue to provide a reliable indication of stock status.

To the extent possible, the EM must provide a reliable and relatively unbiased estimate of stock status. Appropriate model settings for the EM should be selected to achieve these basic criteria. The EM investigated here is based on the model settings of the diagnostic case. Although it is quite possible that alternative model settings would also achieve these basic criteria, it is not necessary to test all possible EM scenarios unless one is expected to significantly outperform the others. It is, however, very important that the EM and its settings do not change through time. This is because

the performance of each MP will have been evaluated under the assumption that a specific data collection program, estimation model and HCR will be applied throughout the evaluation time period. Changes to any component of the MP will mean that the MP being applied differs from the one that was tested. The settings of the EM should therefore be clearly documented to ensure that it is consistently applied (see Appendix A).

We invite WCPFC-SC to consider the following points:

- A 3 phase estimation model based on the OM diagnostic case settings provides reliable and consistent estimates of stock status and can be used as part of an MP for WCPO skipjack.
- The EM used for the updated skipjack evaluations (WCPCF-SC16-MI-IP03) is the same as the one described in this paper
- The settings for the EM (see Appendix A) should be formally documented to ensure they are consistently applied and do not change over time.

1 Introduction

The estimation method is an important component of the management procedure (MP). Its role is to provide a reliable estimate of stock status that can be used by the harvest control rule (HCR) to determine future fishing opportunities (see Figure 1). The performance of the HCR will depend on the information provided by the estimation method (EM) which in turn will depend on the information provided by the data collection program. Therefore, when we test the performance of an MP we must consider not only the HCR but also its accompanying EM and data collection program. Although we may test a large number of HCRs, it is most likely that only a small number of EMs and data collection scenarios would be considered.

A MP can be categorised as either empirical or model based depending on the form of the EM. An empirical MP may determine stock status from direct observation of fishery data (e.g. a CPUE index), whereas a model based MP will employ more analytic approaches, such as a stock assessment model. Due to the absence of reliable CPUE information for purse seine fisheries, initial work has focused on the development of model based MPs for the WCPO skipjack harvest strategy evaluations, although empirical approaches have not been ruled out.

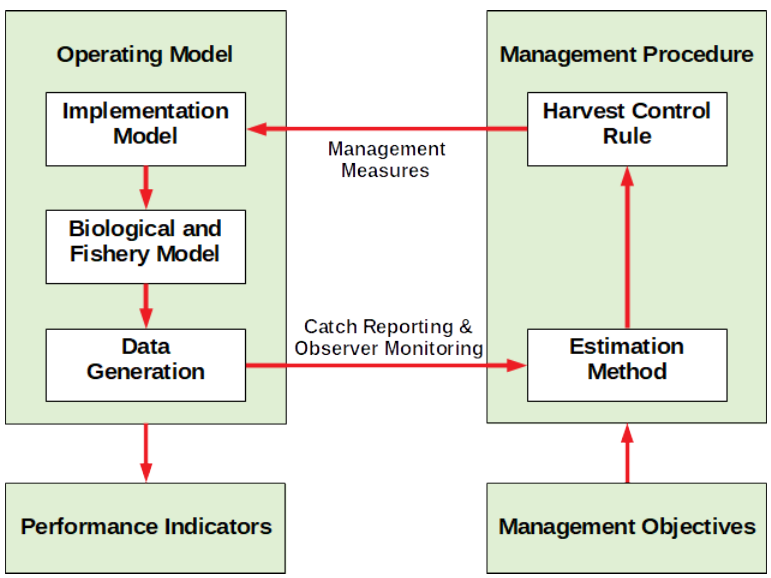


Figure 1: The Management Strategy Evaluation (MSE) framework used for testing management procedures. The estimation model is part of the management procedure and provides information on stock status to the HCR.

In this paper we outline the tested EM for WCPO skipjack and its settings. We present a range of comparative analyses focusing specifically on estimates of depletion, to demonstrate the extent to which it provides a reliable and unbiased estimate of stock status for WCPO skipjack and hence its suitability for use within candidate skipjack MPs. Throughout this paper depletion has been calculated as $\text{SB}_{latest}/\text{SBF0}$.

2 The Estimation Model

The EM tested here is based on the 2019 MULTIFAN-CL diagnostic case assessment for WCPO skipjack. Rather than run a full assessment, which is time prohibitive, an update assessment is conducted in which the EM parameters are initiated with the OM parameter values and then run for a specified number of function evaluations to re-fit the model to the revised data set. This procedure starts the model closer to its expected solution and runs for a shorter and more practical time period but provides only an approximation of the full stock assessment.

For complex stock assessment models, such as MULTIFAN-CL that estimate a large number of parameters, it is often not possible to estimate all of the parameters simultaneously especially when they are not close to their fitted solution values. In such cases the parameters are progressively estimated in phases with additional parameters being freed for estimation at each phase. The 2019 skipjack stock assessment estimated 7642 parameters and ran for 7 phases. The EM estimates the same number of parameters but runs in just 3 phases because the parameters are initialised closer to their expected fitted solution.

2.1 EM settings

The settings of the EM are based closely on the stock assessment diagnostic case model and are shown in bold in Table 1. All other model settings, with the exception of the weighting for the catch likelihood, were set as for the 2019 WCPO diagnostic case skipjack stock assessment (Vincent et al., 2019).

Axis	**Levels**	**Options**		
		0	**1**	**2**
Region structure ∗	2	**8 regions**	5 regions	
Steepness	3	**0.8**	0.65	0.95
Length comp. wtg ∗	3	50	**100**	200
Mixing period (qtr)	2	**1**	2	
Growth	3	Default	Low growth	**High growth**
Hyperstability in CPUE ⋆	2	**0**	-0.5	

Table 1: Skipjack 2019 stock assessment uncertainty grid (Vincent et al., 2019). EM settings are based on the diagnostic case model settings shown in bold. ∗ denotes axes not included in the OM grid. ⋆ denotes axes not included in the stock assessment grid.

When developing the estimation model for WCPO skipjack it was found that model convergence was influenced by the penalty applied to the catch component of the likelihood. The update assessment (EM) was therefore run in 3 phases with the catch penalty weighting being successively increased (100, 10,000, 100,000) in each phase. Phases 1 and 2 were each run for 100 function evaluations and phase 3 for 1,000 function evaluations

2.2 EM testing and validation

A series of tests and comparisons were conducted to investigate the performance of the EM specifically with respect to its ability to estimate depletion. In the first instance a simple comparison was conducted between the EM (3 phase update assessment, as described above) and a full, 7 phase, stock assessment to determine how closely the EM approximates the model estimates of the full assessment (Section 2.2.1).

Subsequent comparisons were made between the projected stock status from the Operating Model and estimated stock status from the EM to illustrate the potential difference in estimates of depletion resulting from model misspecification (Section 2.2.2). Finally we present the EM estimates of depletion that were produced within the evaluation framework (Section 2.2.3).

Comparison	**EM settings**	**Full assess/OM proj settings**
EM vs full assessment	OM model settings	OM model settings
	hist + simulated data	hist + simulated data
	3 phases	7 phases
	OM initialised params	default initialised params
EM vs OM projection	diagnostic case settings	OM model settings
	hist + simulated data	MFCL projection assumptions
	3 phases	MFCL projection
	OM initialised params	MFCL projection assumptions

Table 2: Settings for comparisons between the EM and OM to determine the reliability of EM estimates of depletion. hist + simulated data refers to the full time series of historical observations (1972 to 2018) plus simulated data for the future time period. MFCL projection assumptions refers to the settings assumed when conducting projections e.g. future recruitment taken from SRR, fixed catchability, etc.

2.2.1 Comparison with full assessment

The most recent MULTIFAN-CL stock assessment for WCPO skipjack ran for 7 estimation phases. In each phase of the stock assessment additional model parameters were freed for estimation from their default starting values. In the final phase of the assessment all model parameters are free for estimation. In contrast the EM is run for just 3 phases with all parameters freed for estimation from the outset. However instead of using the default starting values, the parameters are initialised with the values from the OM and therefore closer to their expected fitted solution.

For each comparison the models (EM and full assessment) were fitted to exactly the same data and the same total number of parameters estimated. The data were based on the 2019 stock assessment inputs for the historical time series (1972 to 2018) and on simulated future data (2019 onwards) generated from the OM (the generation of simulated future catch, effort, size frequency and tag recapture data from the OM is described in more detail in WCPFC-SC16-MI-IP10). For both the

EM and the full assessment the model settings (i.e. the parameters that are fixed and not estimated in the model, see Table 1) were also identical.

Comparison of the EM with the full, 7 phase assessment were run for a range of alternative OMs and future time periods. The results (Figure 2) show the EM estimates of depletion to be very similar, and in some cases almost identical, to those of the full assessment. This simple comparison shows that, given the same input data and similar model settings, the EM produces estimates of depletion that are comparable to the full assessment.

2.2.2 Comparison with OM projections

For the first set of comparisons the model settings (Table 1) were held the same for both the full assessment and the EM. This enabled a simple 'like for like' comparison between the EM and full assessment approaches. However, when running the evaluations the settings of the EM will be fixed to specific values (Table 1, bold). These model settings will, in most cases, differ from those of the OM.

A second set of comparisons were conducted to examine the difference between estimates of depletion from projections of the OM and estimates derived from the EM that has fixed model settings. Under this situation the settings for the two models will differ and it is to be expected that the estimates of depletion from them will also differ. The OM may, for example, assume that hyperstability in CPUE is occurring whereas the EM will not. In addition the data for the two models will also differ. The OM projection is based on standard projection assumptions regarding future recruitment, selection patterns, mortality etc. (Pilling et al., 2016), whereas the EM estimates are derived from a model fit to future data that have been simulated from the OM.

Comparisons were made between depletion estimates from status quo 30 year projections of the OM and from EM fits to simulated data generated for the same period under status quo fishing conditions. These comparisons (Figure 3) more accurately represent the performance of the EM within the evaluation framework as they now include both model mis-specification and observation uncertainty.

For the base case scenario (Figure 3a) the model settings for the OM and EM are the same (i.e. there is no model mis-specification) and differences in the estimates of depletion arise solely through observation uncertainty in the input data. The EM estimates of depletion correspond relatively well with those of the OM projection.

Results are also shown for two individual scenarios for which the OM settings differ from the EM (Figure 3b and 3c). These scenarios represent the OMs with the highest and the lowest estimates of depletion. Comparisons with the EM are subject to both model mis-specification and observation uncertainty. The EM estimates, although biased (upwards or downwards) continue to track the overall trend in depletion and therefore continue to provide a reliable indication of stock status to inform the HCR.

For each scenario, the performance of the EM is further examined through a retrospective analysis (Figure 3). The EM was fit to varying time series of simulated data (15, 20, 25 and 30 years). In each case the model estimates are very consistent providing further indication that the EM provides a reliable and consistent estimate of stock status.

A comparison of the ranges of depletion estimates across the OM grid of 24 models for the OM and EM (Figure 4) shows that the EM estimates correspond well with those of the OM, but, as would be expected, have a narrower range of values.

2.2.3 EM performance within the evaluations

Estimates of depletion by the EM for two OM scenarios and one MP (HCR1) for each of the 9 management periods in the evaluations are shown in Figures 5 and 6. Similar to the plots in Figure 3 they show successive estimates of depletion by the EM with increasing quantities of data. They differ from the analysis shown in Figure 3 in that a greater number of retrospective runs are shown (9 runs at 3 year intervals); the random number seeds were different and future fishing is controlled by an HCR rather than assuming status quo conditions.

The EM may be considered to be performing poorly if the terminal estimates of depletion from successive EMs vary substantially, thus providing inconsistent information on stock status to the HCR. Of perhaps greater concern is whether the EM displays a persistent retrospective bias and consistently under- or over-estimates stock status in the terminal years.

Whilst the estimates of depletion show some variation, particularly for the earlier years, there is no evidence of persistent retrospective bias or dramatic variation from one management period to the next. Overall the EM appears to be working well at providing a reliable and consistent estimate of stock status with which to drive the HCR.

3 Discussion

To the extent possible the EM must provide a reliable and relatively unbiased estimate of stock status. Appropriate model settings for the EM should be selected to achieve this basic criteria. The EM investigated here is based on the model settings of the diagnostic case. Although it is quite possible that alternative model settings would also achieve these basic criteria, it is not necessary to test all possible EM scenarios unless one is expected to significantly outperform the others.

It is, however, very important that the EM and its settings do not change through time. This is because the performance of each MP will have been evaluated under the assumption that a specific data collection program, estimation model and HCR will be applied throughout the evaluation time period. Changes to any component of the MP will mean that the MP being applied differs from the one that was tested. This of course does not mean that, once adopted, an MP cannot be changed. However, in such cases where it is considered necessary to change a component of the MP (data collection, EM, HCR), the revised MP should be re-evaluated to ensure that it continues to represent the best performing candidate with regards to achieving management objectives.

Throughout this paper we have noted that the MP comprises a data collection programme, an EM and an HCR. Although we may test a large number of alternative HCRs, it is most likely that only a small number of EMs and data collection scenarios would be considered. This is because the primary role of the EM is to provide a reliable and relatively unbiased estimate of stock status to inform the HCR. Alternative EMs may be considered if they are expected to significantly outperform the current EM with respect to the estimation of stock status. Similarly alternative data collection programmes may be considered if they are believed to be plausible alternatives and are likely to significantly impact the performance of the EM. Ultimately, the performance of the EM is part of the performance of the MP as a whole. If the MP overall is judged to be effective based on the various performance indicators, then the EM that forms part of that MP can also be assumed to perform adequately.

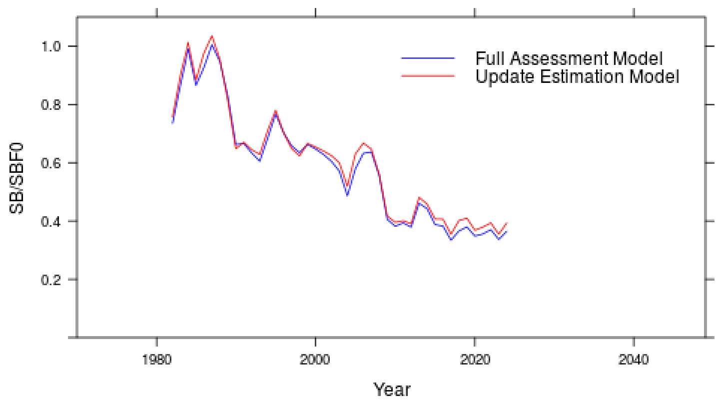
In this analysis we have restricted consideration of model output to the EM estimates of depletion alone, and how they compare to estimates from a full MFCL assessment, estimates from equivalent MFCL projections and to other EM estimates for different time periods and quantities of data (retrospectives). No attempt has been made to interrogate the diagnostics of the EM fit such as residual patterns or likelihood values. In this sense the EM is treated more as an automated algorithm for determining stock status rather than as a stock assessment model.

4 Conclusion

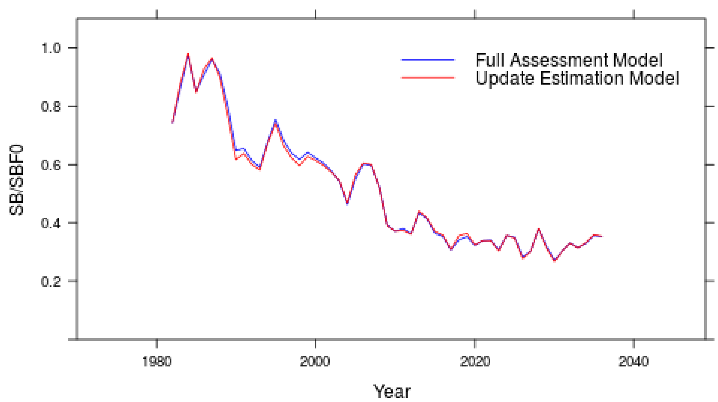
A 3 phase estimation model based on the OM diagnostic case settings, as described above, can provide a reliable and consistent estimate of stock status and can be used as part of an MP for WCPO skipjack.

It is very important that the EM and its settings do not change through time. Changes to the EM, any other component of the MP, will mean that the MP being applied differs from the one that was tested.

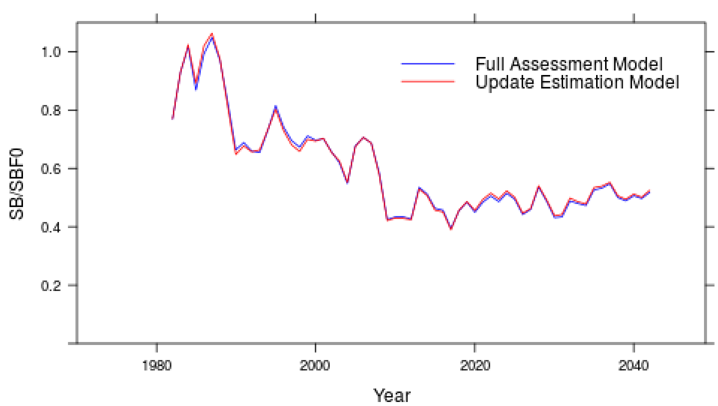
Ultimately, the performance of the EM is part of the performance of the MP as a whole. If the MP overall is judged to be effective based on the various performance indicators, then the EM that forms part of that MP can also be assumed to perform adequately.



(a) A0B0C0D0E0F0: with 6 years of simulated data.

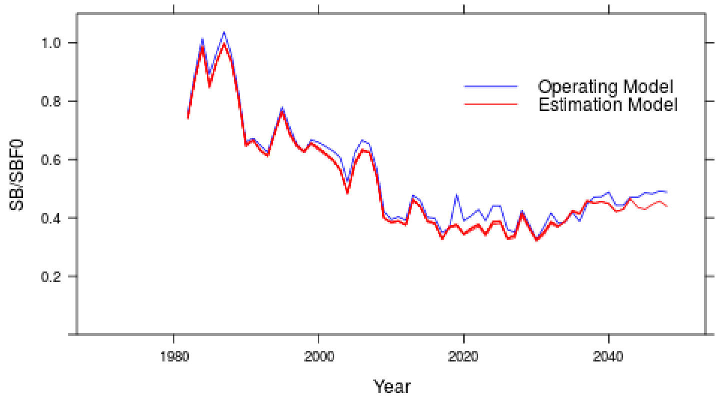


(b) A2B1C1D0E0F0: with 18 years of simulated data

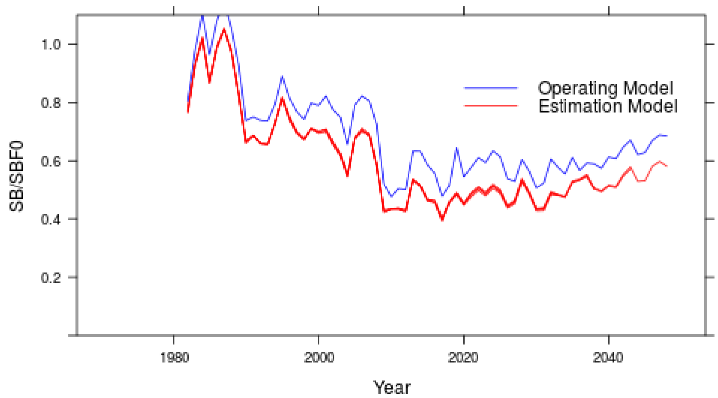


(c) A2B1C1D0E0F0: with 24 years of simulated data

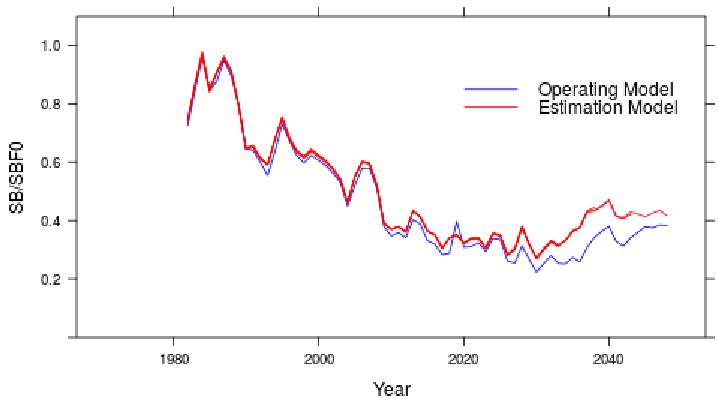
Figure 2: EM comparisons with full 7 phase assessment model run for varying time periods.



(a) A0B0C1D0E0F0.



(b) A2B1C1D0E0F0



(c) A1B0C2D1E1F0

Figure 3: Comparison of EM estimates of depletion with status quo, 30 year OM projections. EM models were run with 15, 20, 25 and 30 years of simulated future data.

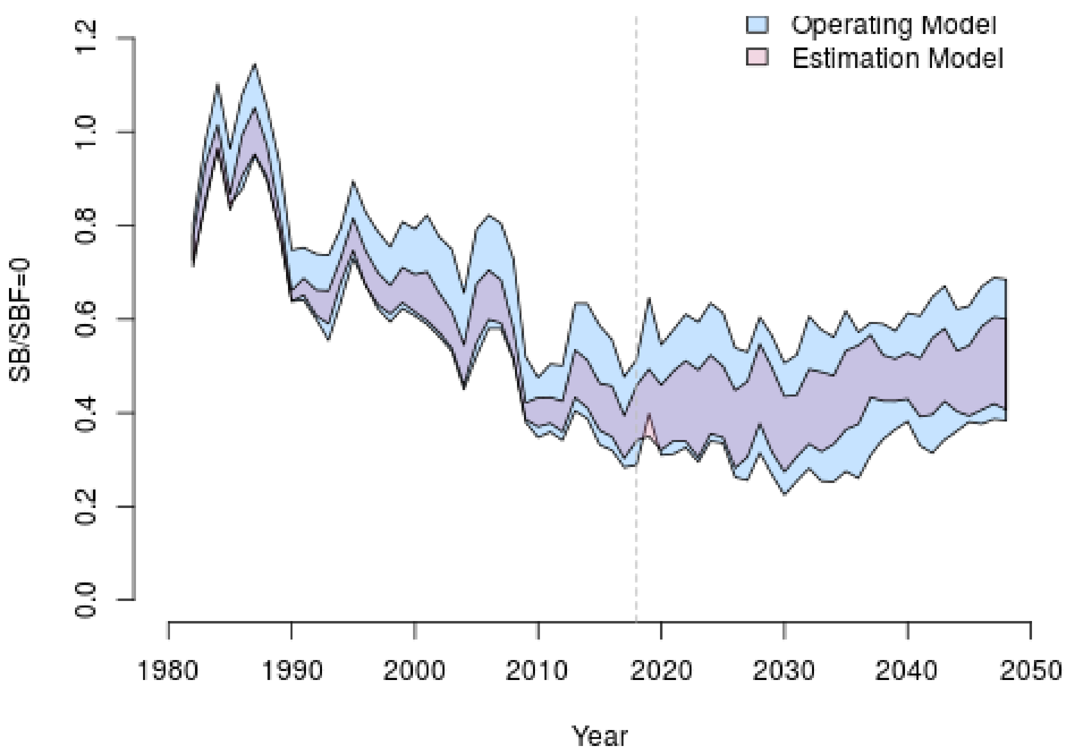


Figure 4: The range of depletion estimates (SB/SB$_{F=0,latest}$) determined from the operating model and the estimation model across the OM uncertainty grid (24 models).

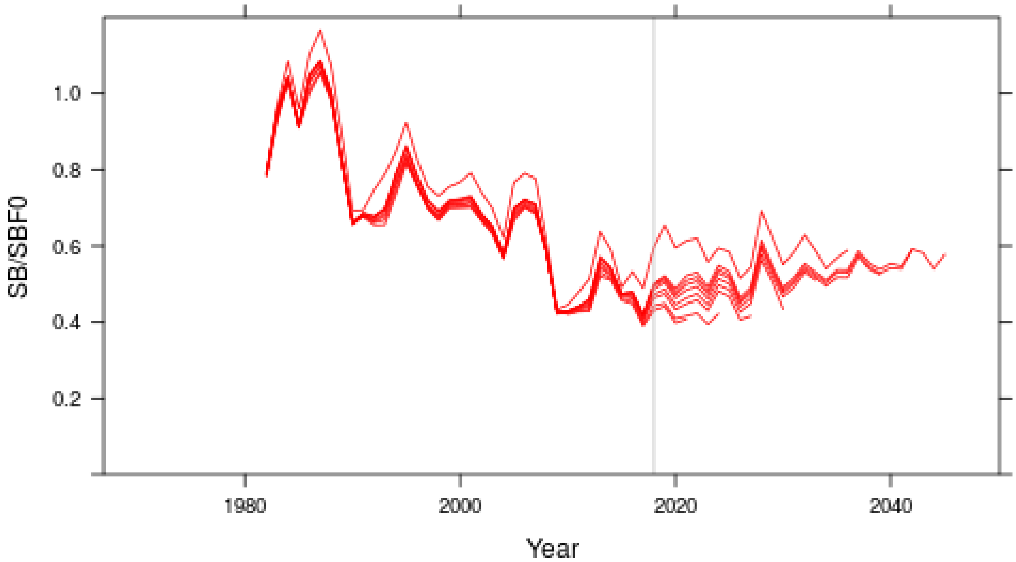


Figure 5: EM estimates of depletion for each management period for OM A0B0C1D0E0 and HCR1. Vertical grey line indicates the beginning of the evaluation period

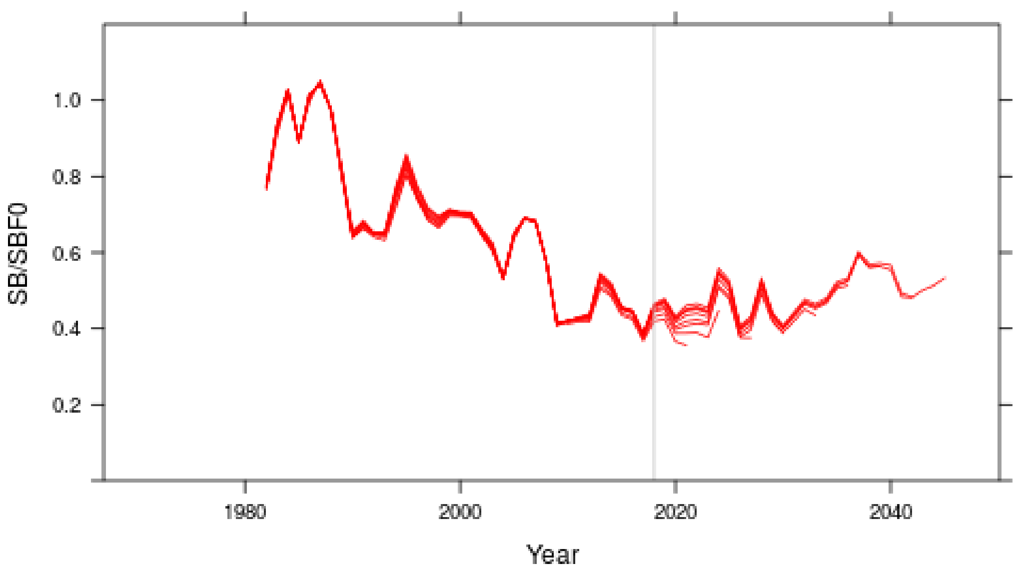


Figure 6: EM estimates of depletion for each management period for OM A1B0C2D1E0 and HCR1. Vertical grey line indicates the beginning of the evaluation period

References

Pilling, G., Scott, R., Davies, N., and Hampton, J. (2016). Approaches used to undertake management projections of WCPO tuna stocks based upon MULTIFAN-CL stock assessments. WCPFC-SC12-2016/MI-IP-04, Bali, Indonesia, 3–11 August 2016.

Vincent, M., Pilling, G. M., and Hampton, J. (2019). Stock assessment of skipjack tuna in the western and central pacific ocean. WCPFC-SC15-2019/SA-WP-05 (rev 2), Pohnpei, Federated States of Micronesia. 12-20 August, 2019.

A EM Settings

The settings of the EM are based closely on the 2019 stock assessment diagnostic case model (Vincent et al., 2019) and are shown in Table 3. All other model settings, with the exception of the growth model, were set as for the 2019 WCPO diagnostic case skipjack stock assessment.

Model Setting	**Value**
Region structure	8 regions
Steepness	0.8
Length comp. wtg	100
Tag mixing period (qtr)	1
Growth	High growth
Hyperstability in CPUE	None

Table 3: Skipjack EM settings.

The EM is run in 3 phases with the catch penalty weighting (age flag 144) being successively increased (100, 10,000, 100,000) in each phase. Phases 1 and 2 were each run for 100 function evaluations and phase 3 for 1,000 function evaluations

The EM is run using MULTIFAN-CL version 2.0.7.x