SCIENTIFIC COMMITTEE
FIFTEENTH REGULAR SESSION
Pohnpei, Federated States of Micronesia
12–20 August 2019

Report of the Second Expert Consultation Workshop on Management Strategy Evaluation.

WCPFC-SC15-2019/MI-IP-03

F. Scott[1], R. Scott, N. Yao, R. Hillary [2], T. Kitakado [3] N. Davies[4], G. M. Pilling and J. Hampton

[1]Oceanic Fisheries Programme, The Pacific Community
[2]CSIRO
[3]Tokyo University of Marine Science and Technology
[4]Te Takina Ltd.

Contents

Executive Summary

The Western and Central Pacific Fisheries Commission (WCPFC) is developing a harvest strategy framework for the management of skipjack, bigeye, yellowfin and South Pacific albacore. As part of this process it is proposed to test the performance of candidate management procedures using Management Strategy Evaluation (MSE). In order to plan the work ahead and to prioritise the issues that will need to be addressed, the Pacific Community (SPC) held a second expert consultation workshop with external MSE experts. The workshop aims were to review and critically appraise the work conducted to date by SPC in support of the WCPFC harvest strategy approach, in particular the technical development of the MSE framework, to provide recommendations for future work and to identify priority work areas.

The main summary points and recommendations from the workshop are presented below. Additional points are raised within the main body of the text. Points are broken down by general subject area.

General

- The Implementation Model, through which the output of the harvest control rule (HCR) is applied to the fishery, should be as simple as possible. In particular, it should not focus on allocation scenarios.
- Although implementation error may be important, if the output of the management procedure is not respected and implemented correctly then that management procedure cannot be considered as having been tested.
- To reduce stakeholder confusion, it is a good idea not to run the full stock assessment in the same year that the output from the management procedure is being evaluated.

Operating Model

- When performing the simulations, operating models (OMs) should be drawn at random from the grid. There is no need to weight the OMs in the grid when performing those draws.
- To investigate the potential impacts of ENSO on the biological population the future movement matrices inside MULTIFAN-CL can be set in time blocks, e.g. every five years.
- Effort creep scenarios within the skipjack OM grid could be modelled through catchability changes.
- There is no need to condition the OMs too frequently (e.g. with every new assessment), unless the realised performance of the management procedure as identified by those assessments falls within the bounds of the agreed exceptional circumstances.
- The general approach for modelling the mixed fishery interactions for bigeye and yellowfin as a single OM with parallel stock models was agreed to be worth investigating.

Management Procedure

- Data collection (the type, frequency and quality of observed data) should be considered as an integral part of the management procedure. For example, a management procedure that includes an Estimation Model (EM) that uses tagging data (such as MULTIFAN-CL) will rely on tagging programmes being continued in the future in order to operate correctly.
- It is possible to use MSE to evaluate the value of data collection. For example, it could be possible to evaluate the decrease in performance of a management procedure if the scope of future tagging programmes were reduced.
- Developing pseudo data generation routines for effort and selectivity deviates within MULTIFAN-CL is important for the application of an empirical based management procedure, i.e. one that is based on CPUE.
- Time varying future selectivities could be considered if shown to be important to the performance of the management procedure.
- As minimising computational time of the simulations is important, this may be helped by truncating the time period over which pseudo data are generated.
- The structural similarities between MULTIFAN-CL based OMs and candidate skipjack EMs are only of concern if the EM has access to information that it should not have. To this end, the same EM settings should be used within the MP tested across all OMs in the grid.
- Before deciding on whether to use a model or empirical based management procedure, the relative performance of both should be evaluated where possible.
- If a real world fishery has been identified as being useful for generating a CPUE series for an empirical based management procedure it may be possible to include that fishery in the OM as an index fishery.
- The "tuning" of HCRs to maximise performance against a specific performance indicator (PI) considered to be of key importance can be computationally intensive. Therefore, it may be necessary to agree on the most important PIs before tuning is possible.
- It is possible to add additional meta-rules to the output of the HCR. For example, the maximum change in output could be limited to +- 20%.

Selecting the MP

- The number of PIs should be as small as possible.
- The proposed tool for allowing stakeholders to interactively explore the PIs and other evaluation outputs should continue to be developed (Scott et al., 2019b).

This document makes extensive use of a number of acronyms. Some of these acronyms will be familiar to regular readers of papers submitted to WCPFC-SC, others may be less familiar. The use of such acronyms is unavoidable when describing and recounting the technical discussions that took place during the workshop. For the benefit of the reader, we provide here a brief summary of some of the key acronyms used throughout this paper.

Acronym	Description
MSE	Management Strategy Evaluaion
OM	Operating Model (see figure 1)
MP	Management Procedure (see figure 1)
HCR	Harvest Control Rule (see figure 1)
IM	Implementation Model (see figure 1)
EM	Estimation Model (see figure 1)
PI	Performance Indicator
TRP	Target Reference Point
ENSO	El Niño Southern Oscillation

Table 1: Key acronyms used throughout this paper

We invite WCPFC-SC to consider the main outcomes of the second technical expert consultation workshop on MSE, and in particular to note the technical outcomes of the workshop, the general support for the approaches being taken, and the suggestions for future development.

1 Introduction

The Western and Central Pacific Fisheries Commission (WCPFC) is developing a harvest strategy framework for the management of skipjack, bigeye, yellowfin and South Pacific albacore (WCPFC, 2014). As part of this process it is proposed to test the performance of candidate management procedures (MPs) using Management Strategy Evaluation (MSE). A first expert consultation workshop on MSE was held in 2016 which focussed on the general MSE approach (Scott et al., 2016). In order to plan the technical work ahead and to prioritise the issues that will need to be addressed, the Pacific Community (SPC) held a second expert consultation workshop with external MSE experts. The workshop aims were to review and critically appraise the work conducted to date by SPC in support of the WCPFC harvest strategy approach, to provide recommendations for future work and to identify priority work areas.

The discussions were largely based on the skipjack evaluations as this is currently the most advanced. However, the discussions were also relevant to the evaluations of the other species.

This report presents the main discussions and recommendations from the workshop.

2 Management Strategy Evaluation framework

The workshop discussed the most recent structural and technical developments of the WCPFC MSE modelling framework (Scott et al., 2018a, 2017). This provided an overview on which to base the more detailed discussions later in the workshop agenda. Items presented and discussed included: general modelling framework; data availability and data pre-processing; file structures and supporting software (HTCondor, FLR4MFCL, MULTIFAN-CL, etc.).

Most of the framework is implemented using R (R Core Team, 2019) with the operating model (OM) and elements of the management procedure (MP) implemented using MULTIFAN-CL (Davies et al., 2017, 2018, 2019) (see Figure 1). There was general agreement that the overall MSE modelling framework was satisfactory. In particular, it was noted that MULTIFAN-CL should continue to be used as OM for skipjack and South Pacific albacore.

The main issues arising were that data collection (the type, frequency and quality of observed data used by the MP) should be considered as an integral part of the MP. This is because the performance of a particular MP is dependent on the data that is collected. An adopted MP is therefore agreed with the stipulation that the required data will continue to be collected in the future. For example, if an adopted MP includes an Estimation model (EM) that uses tagging data (such as MULTIFAN-CL), appropriate tagging programmes will need to continue to be run in the future. That is, the required future tagging programmes become part of the adopted MP.

It was noted that it is possible to use MSE to evaluate the value of data collection. For example,

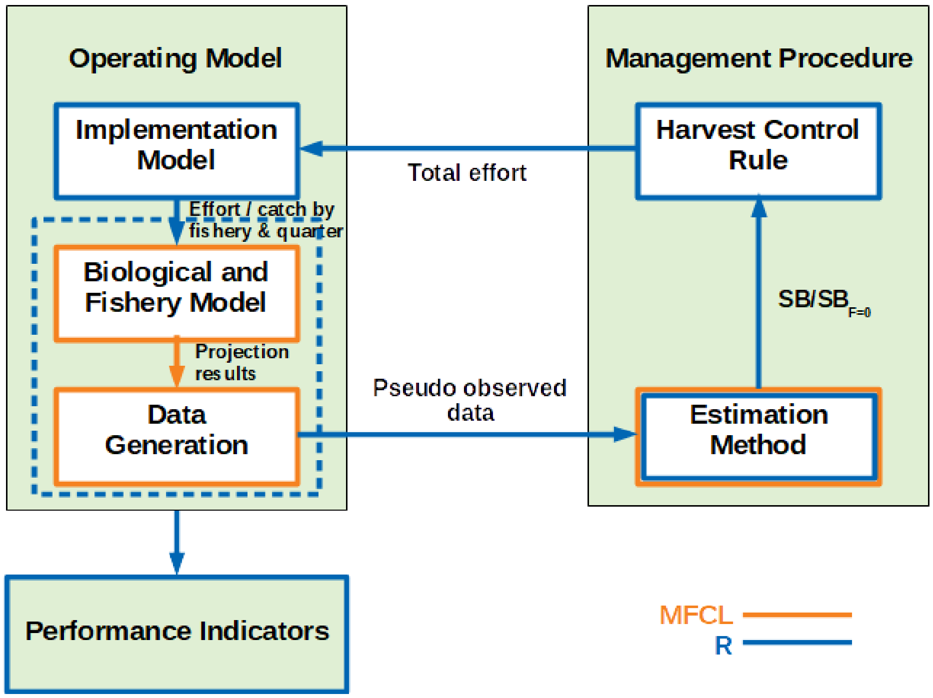

Figure 1: Proposed implementation of the MSE framework for skipjack.

it could be possible to evaluate the change in performance of an MP if future tagging programmes were reduced or increased. As a further step, an adaptive sampling procedure could be considered in the future whereby the data collection requirements in each management cycle are determined as part of the MP. However, this approach is not ready to be considered at this stage.

It was also noted that although the skipjack evaluations have so far focussed on a model based MP, it may still be worth examining an empirical based MP based on a CPUE index, as proposed as the primary approach for South Pacific albacore. This would require both the identification of a suitable CPUE index and, for the purpose of the evaluations, the generation of a pseudo CPUE series based on that index. The generation of such CPUE series will be assisted by the ability to resample effort deviates from the MULTIFAN-CL OM. This feature is not currently available in MULTIFAN-CL but is expected to be added in the near future.

The importance of repeatability of the simulations was discussed. This can be organised through recording the seeds of the random number generators used in the evaluations.

Summary of key points and recommendations:

- Data collection (the type, frequency and quality of observed data) should be considered as part of the MP. For example, an MP that includes an EM that uses tagging data (such as MULTIFAN-CL) will rely on tagging data being collected in the future to operate correctly.
- It is possible to use MSE to evaluate the value of data collection. For example, it could be possible to evaluate the decrease in performance of an MP if future tagging programmes were

reduced.

- An adaptive sampling procedure could be considered in the future whereby the data collection requirements are determined as part of the MP. However, this approach is not ready to be considered at this stage.

3 Operating Model

The OM is a mathematical representation of the "true" system. It simulates the real world by attempting to capture all existing knowledge and data processes for the exploited populations and associated fisheries. The main components of the OM is the the biological and fishery model (see Figure 1). An additional Implementation Model (IM) may also be included.

As noted above, the workshop discussions were largely based on the skipjack evaluations as these are the most advanced. However, the discussions were also relevant to the evaluations of other species.

3.1 Implementation Model

The IM translates the management decision made by the MP into management actions affecting the biological and fishery model.

During the workshop it was agreed that the IM should be kept as simple as possible. In particular, it should not focus on allocation scenarios (as discussed previously, see Scott et al. (2018a)). However, for the purpose of running the evaluations, it will be necessary to make basic assumptions about how changes in catch or effort, as determined by the harvest control rule (HCR), will be implemented by the fishery. Currently, the basic assumption for the skipjack evaluations is that the HCR will apply to all fisheries (with the possible exception of archipelagic waters); that each fishery will be subject to an equal proportional increase or decrease in catch or effort and that the overall catch or effort will be consistent with that specified by the HCR.

The workshop acknowledged that, in reality, the distribution of future catch and effort amongst fisheries may differ from that assumed in the evaluations as a consequence of ongoing negotiations on the allocation process. It was considered that fine scale changes such as different allocations between fisheries within an assessment region (e.g. EEZ vs high seas allocation) would likely have only a small impact on the performance of the HCR. However, differences in effort allocation at the broader assessment region level may have greater impact. It was recommended that a small number of alternative scencarios be investigated to determine the sensitivity of the results to such changes but that these scenarios be restricted to the robustness set of OMs (Scott et al., 2018d, 2019d).

When testing MPs, the assumption has been made that the output of the MP is respected and implemented perfectly. Although "implementation error" (when the management actions specified by the MP are not followed precisely by the fishery) has been considered as part of the MSE framework, if the output of the MP is not respected and implemented correctly then the MP is effectively different to that which was originally tested. Scenarios such as "what happens if the actual effort that is implemented is 25% higher than that set by the MP?" should not therefore be included within the reference set of OMs. However, such scenarios could be considered as part of the robustness set of OMs to illustrate the impact of implementation error (Scott et al., 2019d).

Effort creep is an issue that is known to be important to stakeholders and can have a detrimental impact on the performance of MPs. However, this is an issue that can be considered in the conditioning of the OMs rather than as something that is dealt with by the IM (Scott et al., 2019a).

Summary of key points and recommendations:

- The IM should be as simple as possible. In particular, it should not focus on allocation scenarios.
- Although implementation error may be important, if the output of the MP is not respected and implemented correctly then the MP is effectively different to that which was originally tested.

3.2 Conditioning Operating Models

The conditioning of the OMs is an important element of the evaluations. The range of OMs in the grid of options should cover a broad range of plausible uncertainties. Additionally, the OMs can be broadly split into two categories: the reference set and the robustness set.

For the skipjack evaluations, the stock assessment uncertainty grid for WCPO skipjack has been largely adopted for the MSE reference set although some factors have been dropped and a number of additional factors have been added (e.g. hyperstability, recruitment period) (Scott et al., 2018d). A new stock assessment for WCPO skipjack will be presented to SC in 2019. At the time of writing the results of this assessment are not known. The workshop noted that the conditioning of the OMs was part of the monitoring strategy, to check that the most important sources of uncertainty continue to be appropriately represented in the modelling framework. The conditioning process would therefore be repeated periodically but there is no need to condition the OMs too frequently (e.g. with every new assessment).

Additional consideration of the conditioning of the OM will be required in the event that the management procedure falls outside the bounds of the range of uncertainty assumed for the evaluations. An example might be if future recruitment were to fall to very low levels that were not considered during the initial conditioning process. Such events were not part of the original testing of the

HCR, they are considered to be exceptional circumstances and may require additonal management action that deviates from the HCR. The workshop noted that exceptional circumstances should not be specified too rigidly and that a relatively loose description of such events was preferrable.

The workshop discussed sampling OMs from the grid for the evaluations. The MSE simulations will require many replicates to be run for each candidate MP (> 1000). Each replicate will use one of the OMs in the grid and should be drawn at random.

The relative plausibility of each OM can be crudely evaluated using the likelihood of the fitted OM. This approach can be used to identify OMs that should potentially not be included in the grid. However, weighting the plausibility of each OM based on its likelihood value has some important drawbacks. Previous studies have found that this approach can disproportionately allocate significant weighting to just a select few model scenarios. In addition, some options in the grid only become important in combination with other grid options. Although it is possible to weight the models in the grid, the process to agree the weighting can be very long and the basis for the weighting can change over time. As each OM in the reference set is considered to be plausible, its individual weighting is not considered to be important. As such it is not recommended to weight the OMs when sampling from the grid.

As mentioned above, effort creep scenarios can be included as part of the grid. This may be important as the presence of effort creep can have a detrimental effect on the performance of an MP. Effort creep can be simulated through a range of alternative future catchability scenarios. It may be challenging to agree appropriate effort creep scenarios. An envelope approach (no effort creep plus maximum plausible effort creep) was discussed as a viable start point. It may be necessary to additionally monitor the estimated level of effort creep in the monitoring strategy to check that it does not exceed the tested range.

To investigate the potential impacts of short-term climate effects on the biological population, the future movement matrices inside MULTIFAN-CL can be set to represent those of specific El Niño Southern Oscillation (ENSO) conditions (El Niño, La Niña or neutral). This can become more sophisticated when environmental covariates are added to the movement matrix. However, this addition to MULTIFAN-CL is not expected in the near future. Time-blocks of movement, to approximate environmental impacts on movement, could be specified within the OM.

For skipjack, the HCR in the MP is likely to be driven by estimated $SB/SB_{F=0}$. It was noted that one method to identify important options to include in the conditioning grid is to look at the effect of the likelihood profile $SB/SB_{F=0}$ with respect to other options (i.e. the marginal likelihood). For example, it may be possible that the assumptions about the tagging data have the largest impact on the estimated $SB/SB_{F=0}$, in which case those assumptions on tagging should be included in the grid. However, this marginal likelihood may be difficult to calculate in practice.

Summary of key points and recommendations:

- For the replicate simulations, OMs should be drawn at random from the grid. There is no need to weight the OMs in the grid.
- It may be interesting to examine the likelihood profile of depletion ($SB/SB_{F=0}$) to further identify important options to include in the grid.
- To investigate the potential impacts of ENSO on the biological population the future movement matrices inside MULTIFAN-CL can be set in time blocks, e.g. every five years.
- Catchability scenarios could be considered as part of the skipjack OM grid to model effort creep.
- There is no need to condition the OMs too frequently (e.g. with every new assessment), unless the realised performance of the MP identified by that assessment falls outside the bounds of the variability assumed when testing the MP.

4 Management Procedure

The MP can be described as the formally specified combination of data collection, analysis method (or EM, which may be a stock assessment) and HCR (decision rule) that are used to set the fishing opportunities in each management cycle. MPs are tested by simulation and chosen for their performance in meeting the specified management objectives and robustness to a range of uncertainties. It is the whole MP that is ultimately adopted as part of the harvest strategy framework. This means that a HCR can only be agreed with the condition that the data collection and EM used in the simulations as part of the MP continue into the future.

As an example, for the simulations it is necessary to make assumptions about the future tagging events, including the frequency and length composition of the releases. This assumption becomes part of the MP so that the chosen MP is accepted with the condition that the simulated future tagging events are actually realised. Within the work pursued so far, an initial scenario was considered to be that the future tagging events would be similar to the recent historical tagging events.

As well as the application of an MP, the tuna stocks will continue to have a full stock assessment every three years. It was noted that it is not advisable to run the full stock assessment in the same year that the output from the MP is being evaluated.

4.1 Generation of pseudo data

As mentioned above, data collection is an integral part of the MP. In the MSE simulations the data collection process is modelled through the generation of pseudo data. In the MSE framework presented above, the actual generation of the pseudo data is carried out by MULTIFAN-CL during the OM simulations. However, it is really part of the MP.

Presently, pseudo data can be generated for the length and weight catch compositions, catch and effort by fishery and tag recapture data, in both the historical and future periods (Davies et al., 2017; Scott et al., 2018c). The repeatability of the pseudo data generation is governed by the use of seeds for the random number generators. It is possible to generate the pseudo data without error to enable validation checks to be made.

One of the key discussion areas for the workshop was how close the pseudo data needs to be to real data for it to be considered appropriate for the purpose of the evaluations. After comparing generated pseudo data from preliminary simulations and real data, the workshop agreed that the pseudo data generation processes were generally sufficient for the evaluations. However, there were some issues that might be need to be considered further to improve the realism of these data.

The length and weight composition data are generated from a multinomial distribution using an estimated sample size (ESS). The future is ESS taken from the estimated ESS as determined from an MFCL fit using the self scaling multinomial without random effects minimisation option. Although only length composition data are used in the skipjack evaluations (for the EM), evaluations for the other tuna species will use a mix of length and weight composition. Comparing the length compositions of the pseudo data to the historical data showed that the pseudo data may be too well behaved, particularly regarding the presence of modes in the data. It is not yet understood how important this might be. To overcome this it might be possible to include time varying selectivity in the future period so it better matches the properties of the historical period.

Presently, the variability for the pseudo catch and effort observations is not by fishery. This could be added into future versions of MULTIFAN-CL. It was noted that although pseudo data for both catch and effort can be simulated, it may only be necessary to simulate pseudo data for either catch or effort for each fishery, depending on what data are actually required by the MP. Initial investigations into simulating appropriate catch and effort data are detailed in Scott et al. (2018b). Further work on the generation of pseudo catch and effort data will be aided by future developments in MULTIFAN-CL to enable re-sampling of effort deviates and catchability deviates. These developments may be particularly important when constructing a simulated pseudo CPUE index for an empirical based MP.

The pseudo data generation routines can be applied to both the historical and future periods. Including variability in the historical period allows variability in the starting position of the projection to be included. This could be further enhanced through the addition of variability in the terminal numbers at age, a feature that has been planned for MULTIFAN-CL. Also, it was noted that with an effort controlled management regime, including variability on the observed error through the generation of pseudo data was equivalent to including implementation error in the IM.

Further methods for comparing pseudo generated and real data were suggested including examining the proportion of zeros in the data, given that there is a known zero inflation feature in the data. Additionally, retrospective cross validation could also be used.

It was suggested that it may be possible to speed up the computational time of the simulations by truncating the time period over which pseudo data are generated.

Summary of key points and recommendations:

- Developing pseudo data generation routines for effort and selectivity deviates within MULTIFAN-CL is important for the application for an empirical based MP, i.e. one that is based on CPUE.
- Retrospective cross validation could be employed to investigate the pseudo data generation processes.
- Time varying future selectivities could be considered if shown to be important to the performance of the MP.
- It may be possible to speed up the computational time of the simulations by truncating the time period over which pseudo data are generated.

4.2 Estimation Model

The EM is used within the MP to assesses the change in stock status and provide an appopriate signal to the HCR. The choice of EM depends on whether an empirical or model-based approach is being used for the MP. The choice of EM has an impact on the MP performance as well as practical implications for running the MSE simulations.

One of the key discussions at the workshop was about the relationship between the EM and OM. The skipjack evaluations use MULTIFAN-CL for both the OM and also potentially the EM. Therefore, there was a potential problem of the EM being too close to the OM. However, the workshop agreed that this was not a problem. In the MSE simulations, the same EM (model settings, parameters etc) must be used for all OMs in the grid. For example, if MULTIFAN-CL is being used as the EM, then the same steepness assumption in the EM must be used for all the OMs, even though the steepness in the OMs can be different. So long as the EM does not have access to information that it is not supposed to have (a basic tenant of MSE simulations), using MULTIFAN-CL for the EM is acceptable.

It is expected that the MP for skipjack will be model based. However, consistent with SC advice, it was suggested that it is still worth running a limited set of tests with an empirical based MP to see how well it performs.

Regarding a model based MP for skipjack, several methods for the EM have been tried including the A4A model (Jardim et al., 2015) and a stripped-down MULTIFAN-CL model (Scott et al., 2017). The A4A model requires a CPUE input. However, for skipjack it is assumed that the CPUE index is not guaranteed to be informative. Although a surplus production model has not been tried as an EM it is widely assumed that it will also not perform well, for similar reasons. It was noted that it is important that the EM is able to make use of the type of data that is currently being collected, for example tagging data. The workshop therefore agreed that the preferred option was to continue developing an EM using MULTIFAN-CL.

One of the disadvantages of using MULTIFAN-CL as the EM is the length of time required to run the MSE (current skipjack simulations take about 10 hours for one replicate). The workshop discussed the issue of convergence of the EM and how to reduce the number of model evaluations. As the main signal to the HCR from the EM is likely to be $SB/SB_{F=0}$, it is not necessary to estimate $SB/SB_{F=0}$ to a high level of precision, i.e. more than 2 or 3 significant figures. This raises the possibility of reducing the number of model evaluations required before an acceptable estimate of $SB/SB_{F=0}$ is obtained.

With regards to empirical-based MPs, it was noted that if a real world fishery has been identified as being useful for generating a CPUE series for that MP it may be possible to include that fishery specifically in the OM.

Preliminary results from the skipjack EM using MULTIFAN-CL were presented. It was noted that although the OM had constant catchability in the future, the estimated catchability varied in time, and that this may have an impact if managing the fishery through effort. These impacts may become stronger under an effort creep scenario.

Summary of key points and recommendations:

- The structural similarities between MULTIFAN-CL based OMs and EMs are only of concern if the EM has access to information that it should not have.
- The same EM settings should be used for all of the OM in the conditioning grid.
- Before deciding on whether to use a model or empirical based MP, the relative performance of both should be evaluated where possible.
- If a real world fishery has been identified as being useful for generating a CPUE series for an empirical based MP it may be possible to include that fishery in the OM as an index fishery.

4.3 Harvest Control Rules

An HCR is an agreed rule (algorithm) that describes how fishing opportunities are intended to be controlled by management in relation to the state of some indicator of stock status (Scott et al., 2016). Within the MSE framework the HCR takes the output of the EM and defines overall future fishing opportunities. This output is then passed to the IM component in the OM. For example, the HCR can output overall future catch or effort limits based on the estimated depletion from an EM.

The issue of tuning HCRs was discussed at the workshop, whereby a set of HCR parameters are found that maximise an agreed sub-set of most important performance indicators (PIs). Other RFMOs have used a competitive style of tuning, where participants are invited to submit alternative HCR parameterisations to see which one "wins". However, this may not be appropriate for the WCPO skipjack evaluations given the length of time it takes to run a single evaluation replicate (currently about 10 hours). Instead an alternative approach will need to be found, perhaps using an initial grid of HCRs that is subsequently thinned out and developed.

A key issue with tuning is that it requires stakeholders to be very clear about which are the key management objectives they want to achieve with their fishery so that appropriate indicators can be developed. In the case of the WCPO tuna stocks, the main indicator is thought to be the proximity to the target reference point (TRP). However, other indicators, such as catch stability, are also important.

When considering an empirical based HCR (as is expected to be used for albacore), it may be possible to use multiple input signals. For example, multiple CPUEs from different fisheries with different selectivities can be integrated in the HCR (Hillary et al., 2013).

It was emphasised that even after the adoption of an MP, there is still the possibility of continuing to tune the adopted HCR as more data become available in the future.

It was considered that step changes in the HCR output (e.g. from a step shaped HCR) were undesirable as these can mean that small changes in the input signal (e.g. $SB/SB_{F=0}$) lead to large decreases in catch or effort. This can lead to disagreements about rerunning the EM until a slightly different input signal is achieved that does not lead a large decrease in catch or effort. It was also noted that it is possible to add additional meta-rules to the output of the HCR. For example, the maximum change in output could be limited to +- 20%.

Summary of key points and recommendations:

- The tuning of HCRs to maximise performance can be computationally intensive. It may be necessary to agree on the most important PI(s) before tuning is possible.
- There is the possibility of continuing to tune the adopted HCR as more data become available in the future.
- It is possible to add additional meta-rules to the output of the HCR. For example, the maximum change in output could be limited to +- 20%.

4.4 Performance Indicators

PIs are used to evaluate how well a candidate MP is expected to perform in relation to the fishery management objectives and enable the selection of a preferred MP option from a range of candidates. A suite of 11 PIs have been proposed for the skipjack evaluation which cover a range of economic, biological, social and ecosystem objectives (Scott et al., 2018b). However, not all of the proposed PIs can be calculated from the MSE outputs and alternatives should be found, where possible.

Different stakeholders can have different objectives for a fishery. This means that a wide range of PIs can be proposed as part of the MP selection process. However, the workshop agreed that the number of PIs should be kept as small as possible. The WCPO tuna stocks have a TRP which is considered to be the "best" place to be regarding economic and biological considerations. Tuning the HCR, as discussed above, could therefore be performed by focussing on a single PI that quantifies the ability of an MP to maintain the stock at the TRP (for example, PI 8 in Scott et al. (2018b)). However, such an indicator may not be able to capture all of the information required on the performance of an MP that is required by stakeholders. For example, an indicator that measures the proximity of the stock to the TRP does not consider stability of catches etc.

Communicating the results of the evaluations to stakeholders is a key element of the HS process. It was suggested at SC 14 (2018) that the PIs are calculated over three time periods: short-, medium- and long-term (Scott et al., 2018b). Additionally, most of the PIs are calculated as distributions of values, rather than point values. This means that as the number of candidate MPs increases, a potentially large amount of information will need to be communicated to stakeholders to enable them to make an informed decision about preferred MPs. A proposed option for communicating the results is to use an interactive tool that allows users to explore the evaluation results and PIs using a range of different graph types. A prototype version of this tool was demonstrated at the workshop and it was agreed that it should continue to be developed (Scott et al., 2019b).

It was noted that it may be possible to develop automatic rejection criteria for MPs based on the PIs. For example, as agreed by WCPFC for the purporses of harvest strategy development, if the probability of $SB/SB_{F=0}$ falling below the limit reference point (LRP) is greater than 20%, the MP could be thought as too risky and not considered further.

Summary of key points and recommendations:

- The number of PIs should be as small as possible.
- The proposed tool for allowing stakeholders to interactively explore the PIs and other evaluation outputs should continue to be developed (Scott et al., 2019b).

4.5 Mixed fishery interactions

Skipjack, South Pacific albacore, yellowfin and bigeye tuna are all caught using an overlapping mix of different fishing gears in overlapping regions, leading to a mixed fishery with complex species and gear interactions. This is particularly true for yellowfin and bigeye tuna which are caught together by a mix of purse seine, longline and other gears. The workshop discussed the possibility of performing the harvest strategy evaluations for yellowfin and bigeye together (Scott et al., 2019c). It was agreed that the general approach of using an OM based on the two stocks, each being modelled by MULTIFAN-CL in parallel, was worth considering further. Additionally, MPs based on the estimated status of both stocks could be developed. This is discussed further in Scott et al. (2019c).

Summary of key points and recommendations:

- The general approach for modelling the mixed fishery interactions for bigeye and yellowfin as a single OM with parallel stock models was agreed to be worth investigating.

4.6 Summary

This report summarises the outcomes of the second technical expert workshop on MSE. The two external experts provided scientific and technical expertise on MSE, reviewed the work to date, and provided guidance on the next steps in the development of the WCPO MSE framework.

We invite SC15 to note the technical outcomes of the workshop, the general support for the approaches being taken, and the suggestions for future development.

Acknowledgments

We gratefully acknowledge funding for this work and the workshop from the New Zealand Ministry of Foreign Affairs and Trade (MFAT) funded project "Pacific Tuna Management Strategy Evaluation"; We thank Rich Hillary and Toshi Kitakado for generously giving their time to attend the workshop.

References

Davies, N., Fournier, D. A., Takeuchi, Y., Bouye, F., and Hampton, J. (2017). Developments in the MULTIFAN-CL software 2016-2017. WCPFC-SC13-2017/SA-IP-05, Rarotonga, Cook Islands, 9–17 August 2017.

Davies, N., Fournier, D. A., Takeuchi, Y., Bouye, F., and Hampton, J. (2018). Developments in the MULTIFAN-CL software 2017-2018. WCPFC-SC14-2018/SA-IP-02, Busan, South Korea, 9–17 August 2018.

Davies, N., Fournier, D. A., Takeuchi, Y., Bouye, F., and Hampton, J. (2019). Developments in the MULTIFAN-CL software 2018-2019. WCPFC-SC15-2019/SA-IP-02, Pohnpei, Federated States of Micronesia, 12–20 August 2019.

Hillary, R., Preece, A., and Davies, C. (2013). Mp estimation performance relative to current input cpue and aerial survey data. CCSBT-ESC/1309/19, CSIRO, Canberra, Australia, 2-7 September 2013.

Jardim, E., Millar, C. P., Mosqueira, I., Scott, F., Chato Osio, G., Ferretti, M., Alzorriz, N., and A., O. (2015). What if stock assessment is as simple as a liner model? The a4a initiative. *ICES Journal of Marine Science*, 72(1):232–236.

R Core Team (2019). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.

Scott, F., Scott, R. D., Davies, N., Pilling, G., and Hampton, S. (2018a). Technical developments in the MSE modelling framework. WCPFC-SC14-2018/MI-IP-02, Busan, South Korea, 5–13 August 2018.

Scott, F., Scott, R. D., Pilling, G., and Hampton, S. (2018b). Performance indicators for comparing management procedures using the MSE modelling framework. WCPFC-SC14-2018/MI-WP-04, Busan, South Korea, 5–13 August 2018.

Scott, F., Scott, R. D., Pilling, G., and Hampton, S. (2019a). Skipjack MSE framework. WCPFC-SC15-2019/MI-IP-02, Pohnpei, Federated States of Micronesia, 12–20 August 2019.

Scott, F., Scott, R. D., Pilling, G., Yao, N., and Hampton, S. (2019b). Harvest strategy engagement tools. WCPFC-SC15-2019/MI-WP-09, Pohnpei, Federated States of Micronesia, 12–20 August 2019.

Scott, F., Scott, R. D., Pilling, G., Yao, N., and Hampton, S. (2019c). Mixed fishery and multi-species issues. WCPFC-SC15-2019/MI-WP-04, Pohnpei, Federated States of Micronesia, 12–20 August 2019.

Scott, F., Scott, R. D., Pilling, G., Yao, N., and Hampton, S. (2019d). Modelling key uncertainties in the skipjack MSE. WCPFC-SC15-2019/MI-WP-06, Pohnpei, Federated States of Micronesia, 12–20 August 2019.

Scott, R., Pilling, G. M., Hampton, J., Reid, C., and Davies, N. (2016). Report of the expert consultation workshop on management strategy evaluation. WCPFC-SC12-2016/MI-WP-05, Bali, Indonesia, 3–11 August 2016.

Scott, R. D., Davies, N., Pilling, G., and J., H. (2017). Developments in the MSE modelling framework. WCPFC-SC13-2017/MI-WP-04, Rarotonga, Cook Islands, 9–17 August 2017.

Scott, R. D., Scott, F., Davies, N., Pilling, G., and Hampton, S. (2018c). Generating pseudo data in MULTIFAN-CL. WCPFC-SC14-2018/MI-IP-03, Busan, South Korea, 5–13 August 2018.

Scott, R. D., Scott, F., Davies, N., Pilling, G., and Hampton, S. (2018d). Selecting and conditioning the operating models for WCPO skipjack. WCPFC-SC14-2018/MI-WP-03, Busan, South Korea, 5–13 August 2018.

WCPFC (2014). Conservation and Management Measure on establishing a harvest strategy for key fisheries and stocks in the Western and Central Pacific Ocean. Conservation and Management Measure 2014–06, Eleventh Regular Session of the WCPF Commission, Apia, Samoa, 1–5 December 2014.

A Workshop Participants

Name	Institute	
Nick Davies	Te Takina Limited, New Zealand	Invited expert
John Hampton	OFP-SPC	SPC
Richard Hilary	CSIRO, Australia	Invited expert
Toshihide Kitakado	Tokyo University of Marine Science and Technology, Japan	Invited expert
Graham Pilling	OFP-SPC	SPC
Finlay Scott	OFP-SPC	SPC
Robert Scott	OFP-SPC	SPC
Nan Yao	OFP-SPC	SPC