SCIENTIFIC COMMITTEE
SIXTEENTH REGULAR SESSION

ELECTRONIC MEETING
11-20 August 2020

Report on the THIRD external MSE review: Developments in the South Pacific albacore MSE framework

WCPFC-SC16-2020/MI-IP-11
24/07/2020

Dr Iago Mosqueira
Wageningen Marine Research, Netherlands

WAGENINGEN UNIVERSITY & RESEARCH

P.O. Box 68 | 1970 AB IJmuiden | The Netherlands

Dr Graham Pilling
Deputy-Director (Oceanic Fisheries) of the FAME Division
Pacific Community
CPS - B.P. D5
98848 Noumea, New Caledonia
grahamp@spc.int

Author: Iago MOSQUEIRA

Wageningen Marine Research

DATE
July 23, 2020

SUBJECT
Briefrapportage

OUR REFERENCE
2019663.IM.mnb

POSTAL ADDRESS
P.O. Box 68
1970 AB IJmuiden
The Netherlands

VISITORS' ADDRESS
Haringkade 1
1976 CS IJmuiden

HANDLED BY
Dr. I. Mosqueira Sanchez

TELEPHONE
+31 (0)317 488995

FAX
+31 (0)317 487326

EMAIL
Iago.mosqueira@wur.nl

External expert review of MSE modelling approaches for South Pacific albacore.

Introduction

A review was requested by the Oceanic Fisheries Programme of The Pacific Community (SPC-OFP), under contract CPS20-304, on the ongoing work for the development of Management Strategy Evaluation (MSE) analyses for South Pacific albacore tuna (*Thunnus alalunga*). The present document presents the findings and suggestions of the author with regards to the elements in the MSE approach that are currently being developed, together with some ideas for future work.

The two elements of a simulation of the fishery system constructed for a MSE are the operating model (OM) and the management procedure (MP). The first is a representation of the stock and fishery dynamics, generally as a population and fishery statistical model. The model is usually conditioned on data and a set of assumptions, so as to include both any variability in the system and some quantification of the uncertainties in our knowledge of it. The management procedure to be tested is a combination of the sampling procedure, an estimation method that produces an indication of stock status or exploitation level, and a harvest control rule (HCR), by which a management decision is taken based on this estimate and some reference values. The work under review here includes the general structure of the operating models being conditioned, three harvest control rules (HCR), and work related on the use of the CPUE series as indicators of changes in stock status.

Review terms of reference

The agreed terms of reference for the review read as follows:

The review will focus on the following questions:

1. Consider whether the conditioning and selection of operating models for SP-ALB adequately represent the major sources of uncertainty.
2. Review the approaches used to generate future CPUE indices for long-line fisheries with appropriate variability.
3. Evaluate the overall design, development and testing of empirical HCRs for SP-ALB.

In addition, the researcher experience with the Indian Ocean Tuna Commission will be used to provide some contrasting analysis on:

4. The technical approaches adopted by WCPFC with those of other tRFMOs developing harvest strategies for stocks of albacore tuna.
5. Information as and where relevant on approaches for stakeholder engagement and effective communication that have been successful in other tRFMOs.

Material

The sources used in this review were, primarily, those provided to the author by SPC-OFP. These consist of both published and draft documents covering the three main items in the ToRs, as follows:

- ToR 1 - Conditioning of operating models: Pilling et al. (2018), Scott et al. (2019b) and Tremblay-Boyer et al. (2018).
- ToR 2 - Generation of future CPUE indices: Yao et al. (2020)
- ToR 3 - Design, development and testing of empirical HCRs: Yao et al. (2020b)

In addition, other documents referred to in those above were also inspected when searching for particular pieces of information, or to better understand the context of the work. Those are fully listed in the *References* section of the report.

The review has considered in some detail the similarities and differences with the work carried out by the author for the Indian Ocean albacore stock, under the mandate of the Indian Ocean Tuna Commission (IOTC). The relevant references are also supplied in the final section.

Finally, some demonstration source code, implementing a single step applying a MP that includes a CPUE-based model-free HCR, has also been tested and inspected.

Conditioning of operating models

The OM for South Pacific albacore presented in Scott et al. (2019) is based on a number of runs of the stock assessment model routinely used for management advice. This is a population and fishery model implemented using Multifan-CL (MFCL), in quarterly time steps and stratified in 5 areas and 16 catch fisheries. There are five indices of abundance, derived from catch-per-unit-effort (CPUE) series from various longline fleets.

Structural uncertainty is being considered in the current stock assessment through a model grid that includes alternative values for five fixed quantities or inputs: steepness, natural mortality, growth model, weighting of the size frequency data and CPUE calculation method (Table 1 in Scott et al, 2019). These 72 model runs provide an indication of uncertainty in past stock trajectory and current status, as summarized by Figure 43 in Tremblay-Boyer et al (2018). The grid of model runs used to condition the OM does not include different levels of the weighting factor of the size-frequency data, as it was found to introduce almost no change in the estimated trajectories. This leaves a model grid of 24 runs for the OM reference set.

In comparison, the approach taken by IOTC for the Indian Ocean albacore OM grid was to use a full factorial design on seven variables identified as potential sources of uncertainty, for a total of 1,440 model runs of the stock assessment model, constructed using SS3 (Methot and Wetzel, 2013). The variables and values included are as follows:

- Natural mortality (M): 0.2, 0.3, 0.4, 0.4 at age 0, decreasing to 0.3 at age 5 and older, and 0.4 at age 0, decreasing to 0.2 at age 5 and older.
- Variance of the recruitment deviates (sigmaR): 0.4, 0.6
- Steepness of the Beverton-Holt stock-recruits relationship (h): 0.7, 0.8, 0.9.
- Coefficient of variation of the CPUE series (cpuecv): 0.2, 0.3, 0.4, 0.5.
- Effective Sampling Size of each length data point (ess): 20, 50, 100.
- Catchability trends in the CPUE Longline fleet (LLq): 0% and 2.5% increase per year.

- Form of the selectivity curve for the CPUE fleet (LLsel): logistic or double-normal.

All model runs were considered equally likely and combined without any weighting. No parameter uncertainty was introduced, as the structural uncertainty was shown to be vastly larger. A grid constructed in this manner included of a large number of runs with almost no difference in estimated productivity. But it also provided the material required to explore the relative importance of the variables and levels that IOTC initially decided to include in the grid. An analysis of the impact of each variable and level on the estimated virgin biomass, current status and the stock-yield production function (Mosqueira et al. in prep.) has shown (Figures 1 and 2) that natural mortality (M) of mature ages, and the combination of effective sample size (ESS) and CPUE CV have the greatest effect.

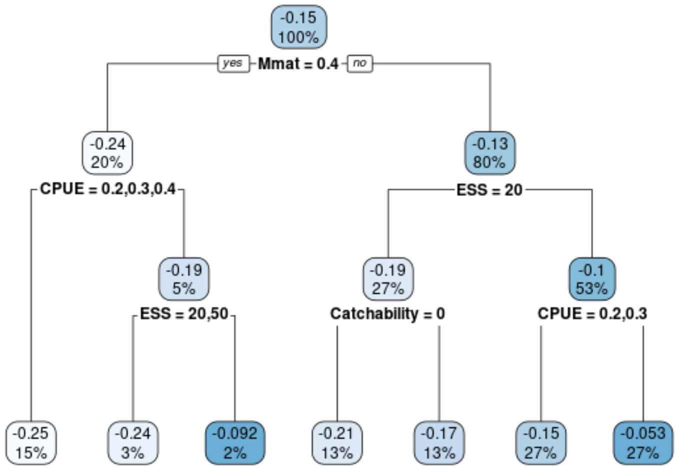

Figure 1. *Correlation tree showing the most significant factors in the Indian Ocean albacore OM grid according to the robustness of the fit as determined by the value of Mohn's rho (Mohn, 1999). Mmat refers to the natural mortality of age 5+ fish, ESS is the effective sample size of the length-frequency data, CPUE is the coefficient of variation of the residuals of the fit to the CPUE series, and catchability refers to the yearly increase in catchability for the CPUE fleet. Presented here as an example use of the method.*

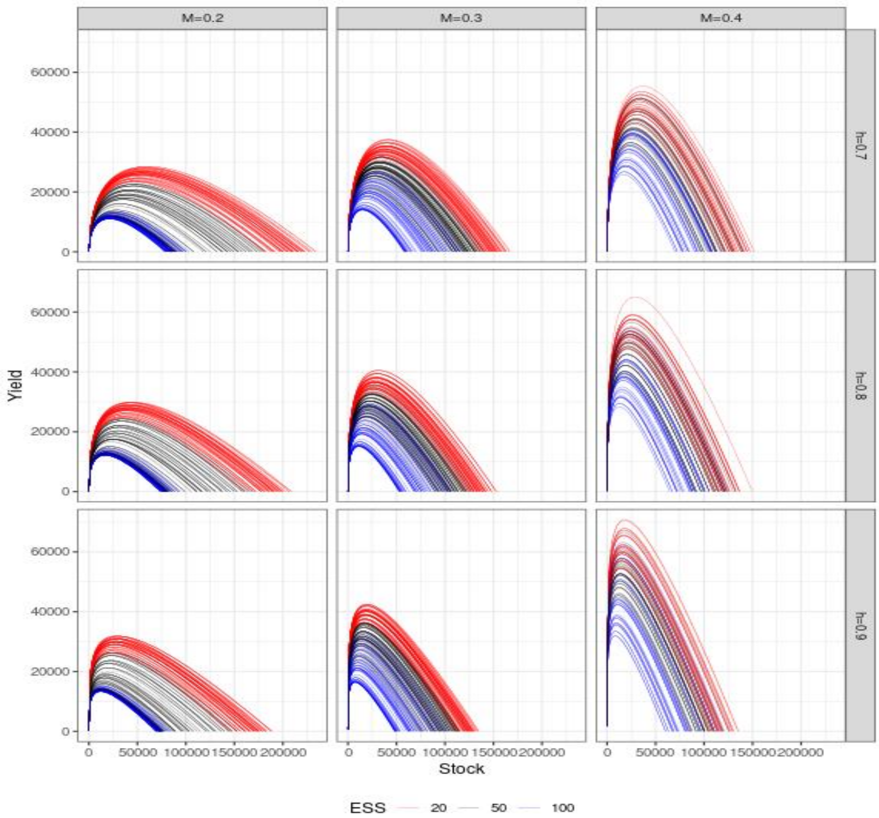

Figure 2. *Production curves obtained from the grid of OM runs for Indian Ocean albacore along four variables: natural mortality of the age 5+ fish (M) in columns, steepness of the stock-recruits relationship (h) in rows, CV of the variance to the CPUE fit (CPUE) as line type, and effective sample size of the length-frequency data (ESS), line colors.*

Although this analysis is still ongoing, and the results might be influenced by particular issues in the structure of the applied model (SS3), the relative weighting of CPUE vs. length data (low ESS combined with high CPUE CV, for example) could be an element to explore for inclusion in the South Pacific albacore OM grid. It might be the case that the contradictory information on the responses to catch levels that the two sources of data appear to have in the Indian Ocean albacore model are not an issue of concern for South Pacific albacore. But the effect of one or other data source dominating in the likelihood might still be worth exploring.

In essence, the proposed OM grid in effect already introduces at different levels the key sources of uncertainty that should always be considered in MSE(Punt et al, 2014): process uncertainty, parameter uncertainty and observation error. Consideration could still be given to the inclusion of extra values in the grid for some of the variables.

Robustness sets

The document accurately indicates the need for robustness tests to be defined and constructed, but this process is yet to start. It might be a pertinent suggestion to agree first on a basic set of scenarios that test the ability of the MP to handle not-so-uncommon short-term negative situations, such as a three to five year series of lower than expected recruitment.

A second set of robustness OM sets could attempt to incorporate the effect of undetected fishery changes on either the past or future dynamics of the fishery. An increase in effective effort or catchability by means of unrecorded technological changes, for example new sonar or satellite technology, could be considered here. Considering alternative relationships between CPUE indices and abundance could supply a number of OMs for this robustness set.

Finally, the definition of long-term scenarios, or those linked to particular events or processes, such as climate change, will undoubtedly require more careful consideration, and the input of relevant experts on ecosystems, oceanography and biology. This process has generally been not too successful in most tuna RFMOs (Sharma et al, 2020).

Generation of future CPUE indices

The candidate MPs being considered in Scott et al (2019) are based on estimates of current relative stock status informed by one or a combination of indices of abundance derived from Catch Per Unit Effort (CPUE) series. There are five series being input to the current stock assessment, and thus used for conditioning of the OM. An analysis is available (Yao et al 2020) to attempt to ascertain the ability of each of those indices to provide information on changes in stock status useful as input to a Harvest Control Rule (HCR). A retrospective analysis has been carried out, five years into the past, where the stock is projected by applying the observed catches by fleet.

The retrospective patterns uncovered are deemed by the authors to be reasonably small and providing support to the use of these CPUE series as estimators of changes in stock status. The estimates of SSB obtained from the retrospective runs (Figure 6 in Scott et al., 2019) show no changes in trends along the five year, but only corrections to the amplitude of change as new data is incorporated. However, the CPUE indices generated by the forecasting step are generally much noisier (Figures 3, 4 and 5 for the aggregated CPUE series), and point to situations in which the HCR could provide a different decision given that trends in abundance move sometimes in opposite directions.

The extend and importance of the retrospective patterns observed in this document should be quantified using one or more of the commonly used metrics, for example Mohn's *rho* (Mohn, 1999). The objective here would be to compare the retrospective deviations across CPUE series rather than as a measure of acceptance of model runs. A particularly useful metric in this case would be the Mean absolute scaled error (MASE, Hyndman, 2006), developed particularly for use in forecasting. The MASE has the desirable properties of scale invariance, predictable behaviour, symmetry, interpretability and asymptotic normality. The one and three step ahead MASE would be computed and compared across CPUEs in the retrospective analyses.

This metric could also be used to directly evaluate the quality of information of each index, by carrying a retrospective forecast (or hindcast) over the recent years and compute the MASE statistic across indices. This exercise was carried out for the IOTC albacore base stock assessment SS3 model, and provided an indication that the CPUE index in area 3 appeared to provide consistent information on stock abundance over the last 20 years (Figure 3).

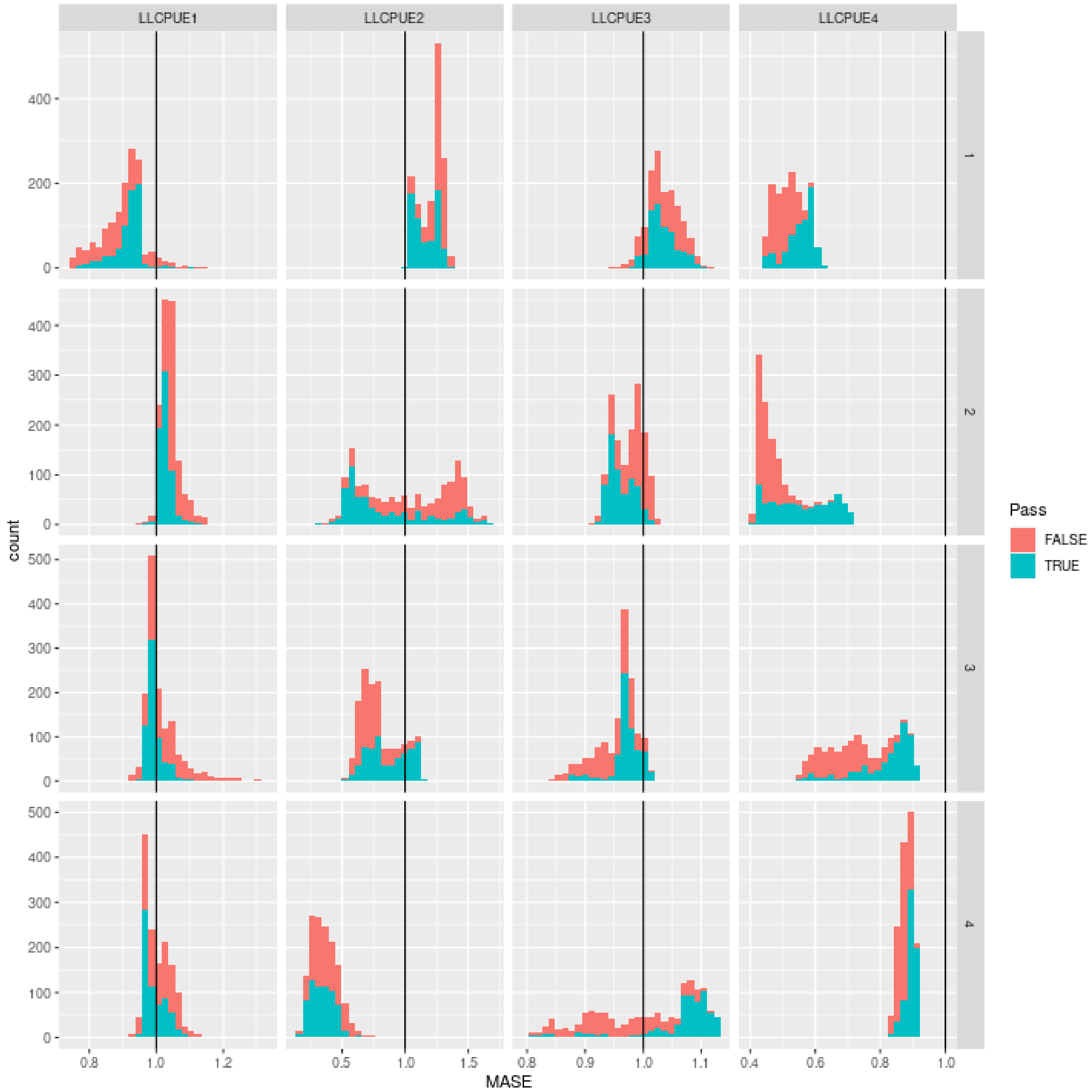

Figure 3. *Distributions of values of the MASE for the Indian Ocean albacore CPUE indices by area (LLCPUE1-LLCPUE4, top panel) sand season (1-4, right hand-side panel). Colors refer to the values of Mohn's rho being larger (TRUE) or smaller (FALSE) than 0.15.*

The procedure would consist on hindcasting, or backtesting, over the recent period, e.g. from the year 2000, using the current deviances in fits to CPUEs and recruitment. The predictive quality of each CPUE could then be computed using MASE so as to better understand their potential use as inputs to the HCR. This procedure would provide a fuller indication of prediction ability that what is obtained from the retrospective analysis. A choice could then be made over what index, or assemblage of them, is chosen to drive the MP. Although Yao et al (2020) employs the CPUE time series of the combined PICT longline fisheries in region 2 to drive the CPUE-based HCR, it is not clear this is the final choice.

Some of the issues previously raised in the past for the albacore MSE work (Scott et al, 2016) appear to be still relevant. For example, the use of an index from a single area or the development of a region-wide index using operational data (as done for IOTC albacore; Hoyle et al, 2019).

Design, development and testing of empirical HCRs

A working document (Yao et al, 2020) presents the description of three harvest control rule (HCR) types, two of them based on trends in CPUE and the third on the observed mean length in the catch. The analysis presents the performance of nine HCRs (three types and three parameter sets each) in terms of their ability to reach and maintain the corresponding target levels.

The precise method employed to generate observations of the distribution of length in future catches should be described, or a suitable reference provided. The observation error mentioned in the document (30% CV) appears to refer to total catches and not specifically to error in length sampling.

No mention could be found in any of the documents reviewed to the use of a tuning procedure to determine the precise values of HCR parameters that would bring the desired performance levels. This might have to do with management objectives not yet being clearly defined and ranked. The comparison of HCR formulations presented here would be better framed in a tuning procedure. A responsive parameter would first need to be identified for each HCR type, for example the gain parameter (*g*) in the CPUE-based HCRs. A search would then be carried out for the value of that parameter under which the MP provides the desired value of a particular performance indicator, linked to the primary management objective. Tuning requires running an MP simulation repeatedly until the performance indicator is within a margin of tolerance of the desired value. The process can be computationally demanding, and might prove too slow to be applied in certain cases.

As mentioned in the document, it would be productive to explore how the relative weighting of both components of the two-phase CPUE HCR affects its behaviour under different circumstances. For example, if an emphasis on CPUE trend vs. absolute level is preferable when levels of variability and error increase.

Technical approaches

The software platform being developed builds on existing code that simplifies access to the inputs and outputs of MFCL, based on classes defined in the FLR toolset (Kell et al, 2007). But only the basic data structures are being used. The example code appears relatively simple at this proof-of-concept stage.

At least one example exists of an MSE analysis being carried out using the tools provided in FLR by the *mse* package in which the operating model was an external model and software, as will be the case here (Pérez-Rodríguez et al, 2019). The Gadget multispecies model (Begley and Howell, 2004) is used as the operating model, while the management procedure is constructed using FLR classes and methods (https://github.com/dgoto2/flr-gadget). The yearly management decision is then applied to the Gadget OM. Choice of software platform is not trivial, and embracing an existing toolset might only be advisable if it contains code, considered both useful and robust, that is best not rewritten. Also if there is interest in contributing to an existing platform with code that others might reuse and extend in the future.

An evaluation of the expected computational needs for a full set of simulations to be executed might be an useful piece of information at this moment. Scaling up to the dimensions of an analysis involving multiple OMs, a large number of iterations, and different MPs might require some design decisions to be made already at this stage.

Example software

The example software contained in file *HCR_2_phase.R* was tested. The required R package, FLR4MFCL, was installed from the latest version of the source code stored at the relevant github repository. The code could be executed without any error using R version 4.0.2, and running on a Linux 64 bit machine.

The code appears to demonstrate that the MFCL OM can be updated every year by applying the catch level coming out of the HCR

A quick attempt was made to increase the number of iterations (*itn*) but the code returned an error in line 259:

```
**************************************
Average years:  5  Gradient:  1.4
Iteration:  2  Counter:  3
**************************************

Error in read.MFCLPseudo(catch = "catch_sim", effort = "effort_sim",
lw_sim = "test_lw_sim",  :
  catch.sim file does not exist
```

Having at this stage a draft design of the software platform, and how it is intended to be implemented, would have been an useful addition for the review. Recent experiences on multiple MSE exercises have lead me to appreciate the value of a complete, even if provisional, design document for software of this kind. Issues like how is the information being passed between the various MSE elements, what inputs, intermediate values and outputs are stored and in which format, or how the software intends to deal with the dimensionality of the analysis (iterations, OMs, MPs), both during execution and for aggregation of outputs, could be presented in a brief design document. For example, if the code is to be executed using some type of parallelization or a high performance computing platform, this should better be considered when the software is designed.

Communication and engagement of stakeholders

Communication and engagement of all stakeholders in the MSE process is not a specific part of the documents under review, but appears to be an issue of concern and interest for WCPFC. It is noteworthy the effort that SPC has placed on developing interactive tools, for example AMPED. The experience of IOTC has shown that dialogue, supported by appropriate tools and engaged and knowledgeable staff, is an essential element for the success of the MP approach. It appears this has also been identified by WCPCF and SPC, and I can only suggest those efforts should continue. In the specific case of the development of MSE analyses, I would like to point to two elements that merit attention, and point to the ways in which IOTC, and other fora in which I am involved, have dealt with them.

The first one is on the presentation of the results of MSE analyses to the plenary of IOTC. As a result of the three meetings of the Management Procedures Dialogue, in which scientists and managers first exchanged their views and opinions about management procedures, objectives, performance indicators and MSE, a document was tabled and agreed that contained the minimum set of outputs that the IOTC plenary would like to see for each MSE analysis (IOTC, 2018, Appendix IV - Presentation of Management Strategy Evaluation Results). This document lists the performance indicators that IOTC uses to evaluate alternative MPs, and the graphical displays and tables used to compare performance and trade-offs across MPs. The document has been updated over time, but has provided a solid basis for discussion at the Technical Committee on Management Procedures to focus on results and content and not on the manner of presentation.

A second element is the accessibility of the analysis and code to any scientist. Development work for IOTC MSEs, for example, has been carried out in open source code repositories and an effort has been made, not always successful, to simplify the accessibility of code and software. Installation of the software employed should be possible for any scientist with a minimum familiarity with the tools or language. Demonstration runs, or scripts that complete an example analysis, should be available and documented. Although it should be expected that a very limited number of scientists will be able and willing to delve into the source code of such an analysis, opening the door to that possibility is a good investment on transparency. In some cases, like a particular RFMO in which the author is currently involved, a number of scientists are in fact ready to install and test the software, and this is considered an essential requirement for the development work.

Final overview

The work carried out so far is a clear progress in the evaluation of candidate management procedures for South Pacific albacore. The technical work is clearly built on a solid understanding of both the system and the model being used. Some software design issues would be better tackled already at this stage, before the complexity of the code increases.

Although not directly as technical issue meant to be under the scope of this review, it would be advisable for WCPFC to provide or agree on an initial set of management objectives under which preliminary evaluations of the MPs can be carried out. A discussion on alternative management objectives is, in my experience, better informed by some initial results. Work on IOTC started based on management objectives that were a direct translation of those in the commission agreement. Thus the aim of conserving stocks at sustainable levels led to an initial target of 50% probability of B being greater or equal to BMSY. From inspection of these results, IOTC suggested exploring other objectives, at 60 and 70% probability. Further dialogue and refinement has taken place, specially for stocks in need of recovery, but this has been much helped by being supported by some preliminary results.

The operating model (OM) includes the key sources of uncertainty in the stock assessment. The relative weight of CPUE and length frequency data proved a significant driver of uncertainty in the IOTC albacore SS3 OM. Their possible role in the MFCL OM could also be explored. It would also be useful to compare the estimates in productivity (e.g. doubling time) and noise (e.g. deviances in the SRR) obtained from each OM in the grid. Those two factors are important determinants of the future behaviour of the OM.

The choice of CPUE series is likely to affect the robustness of an MP informed by it. The current and ongoing analyses on the existing CPUEs could be expanded and complemented so as to best inform the choice of status estimator.

The two harvest control rules (HCR) presented are clear to compute and understand. If a tuning procedure is to be applied, exploring the effect of the HCR parameters on its responsiveness would inform the choice of parameter(s).

The design for the albacore MSE does not mention any attempt at considering the possible changes in the behaviour of the fleets catching this stock to changes in the levels of catch. Fleets can be expected to respond to management measures, especially those of a greater entity. If data and knowledge on the dynamics of these fleets is available, this element could be incorporated, if only tentatively. Changes in targeting dependent on the allowed catch levels, for example, could have a negative effect on the quality of the CPUE series.

References

Begley, J., and Howell, D. 2004. An Overview of Gadget, the Globally Applicable Area-Disaggregated General Ecosystem Toolbox. ICES Document CM 2004/FF: 13. 15 pp.

Hoyle, S.D. and D. Fu, D.N. Kim, S.I. Lee, T. Matsumoto, K. Satoh, S.-P. Wang, T. Kitakado. 2019. Collaborative study of albacore tuna CPUE from multiple Indian Ocean longline fleets in 2019. IOTC-2019-WPTmT07(AS)-10. Working Party on Temperate Tuna (WPTmT).

Hyndman, R.J. and A.B. Koehler. 2006. Another look at measures of forecast accuracy. International Journal of Forecasting, 22(4): 679-688.

IOTC. 2018. Report of the 2nd IOTC Technical Committee on Management Procedures. Bangkok, Thailand, 18–19 May 2018. IOTC–2018–TCMP02–R[E]

Kell, L.T., Mosqueira, I., Grosjean, P., Fromentin, J-M., Garcia, D., Hillary, R., Jardim, E., Mardle, S., Pastoors, M.A., Poos, J.J., Scott, F., Scott, R.D. 2007. FLR: an open-source framework for the evaluation and development of management strategies. ICES Journal of Marine Science, 64: 640-646.

Methot, R.D. and C. R. Wetzel. 2013. Stock synthesis: A biological and statistical framework for fish stock assessment and fishery management. Fisheries Research, 142: 86-99.

Miller, S.K., A. Anganuzzi, D. S. Butterworth, C. R. Davies, G. P. Donovan, A. Nickson, R. A. Rademeyer, V. Restrepo. 2019. Improving communication: the key to more effective MSE processes. CJFAS, 2019, 76:643-656, https://doi.org/10.1139/cjfas-2018-013 Mohn, R. 1999. The retrospective problem in sequential population analysis: An investigation using cod fishery and simulated data, ICES Journal of Marine Science, 56(4): 473–488.

Mosqueira, I. 2018. Update on the development of MSE analysis of candidate Management Procedures for Indian Ocean albacore. IOTC-2018-SC21-17_Rev1

Mosqueira, I. 2019. Indian Ocean Albacore Tuna Management Procedures Evaluation: Status Report. IOTC-2019-TCMP03-09_Rev1

Pérez-Rodríguez, A., D. Howell and I.Umar. 2019. EU SC05 project: "Multispecies Fisheries Assessment for NAFO". Estimation of multispecies based HCRs and use of a multispecies MSE framework to assess the risk of collapse and the fishery-ecological trade-offs. NAFO SCR Doc. 19/017.

Pilling, G.M., R. Scott, F. Scott and J. Hampton. 2018. Technical aspects of a potential South Pacific albacore harvest strategy. WCPFC-SC14-2018/ MI-WP-02.

Scott, R., G. M. Pilling, J. Hampton, C. Reid, N. Davies. 2016. Report of the Expert Consultation Workshop on Management Strategy Evaluation. WCPFC-SC12-2016/MI-WP-05.

Scott, R., F. Scott, G. M. Pilling, J. Hampton and N. Davies. 2018. Selecting and Conditioning the Operating Models for WCPO Skipjack. Selecting and Conditioning the Operating Models for WCPO Skipjack. WCPFC-SC14-2018/MI-WP-03.

Scott, F., R. Scott, N. Yao, R. Hillary, T. Kitakado, N. Davies, G. M. Pilling and J. Hampton. 2019. Report of the Second Expert Consultation Workshop on Management Strategy Evaluation. WCPFC-SC15-2019/MI-IP-03.

Scott, R., N. Yao, F. Scott, G. M. Pilling. 2019. South Pacific albacore management strategy evaluation framework. WCPFC-SC15-2019/MI-WP-08.

Sharma, R, Levontin, P, Kitakado, T, et al. Operating model design in tuna Regional Fishery Management Organizations: Current practice, issues and implications. Fish Fish. 2020; 00: 1– 22. https://doi.org/10.1111/faf.12480

Tremblay-Boyer, L., J. Hampton, S. McKechnie and G. Pilling. 2018. Stock assessment of South Pacific albacore tuna. WCPFC-SC14-2018/ SA-WP-05 Rev. 2.

Yao, N., F. Scott, R. Scott, G. M. Pilling and J. Hampton. 2019. Performance indicators for comparing management procedures for South Pacific albacore using the MSE modelling framework. WCPFC-SC15-2019/MI-WP-03.

Yao, N., R. Scott, F. Scott, G. M. Pilling and J. Hampton. 2019. CPUE analysis for South Pacific albacore. WCPFC-SC15-2019/MI-WP-07.

Yao, N., R. Scott, F. Scott, and J. Singh. 2020. Retrospective CPUE forecasting for the south Pacific albacore. WCPFC-SC15-2020/MI-IP-XX

Yao, N., J. Singh, R. Scott and F. Scott. 2020. HCR design considerations for South Pacific Albacore. WCPFC-SC16-2020/MI-IP-XX.

Justification

Project number: 4311000009-12 (SPC)

The scientific quality of this report has been peer reviewed by a colleague scientist and a member of the Management Team of Wageningen Marine Research.

Approved: Dr Thomas Brunel
Colleague scientist

Signature:

Date: July 22, 2020

Approved: Drs. J. Asjes
Manager Integration

Signature:

Date: July 23, 2020