OCEANS & ATMOSPHERE
www.csiro.au

Final Report for FAO-CSIRO Project:

Scientific and Technical Support for Indian Ocean Yellowfin and Bigeye Tuna Management Procedure Development: Phase 3

FAO Budget Code TF/FIDTD/TFAA970097099

Dale Kolody (Dale.Kolody@csiro.au)

Paavo Jumppanen

Jemery Day

July 2021

Acknowledgments

This work was jointly funded by the Australian Department of Foreign Affairs and Trade and CSIRO, Australia (administered through FAO-IOTC). Technical oversight and advice was provided by various IOTC Working Groups, and notably the participants of the IOTC MSE Task Force, including Iago Mosqueira, Toshihide Kitakado, Gerald Scott, Gorka Merino, Hilario Murua, Dan Fu and M. Shiham Adam. Operating model conditioning built upon the stock assessment work of Dan Fu, Adam Langley, Agurtzane Urtizberea and others in the IOTC scientific community. The original R-based MSE code from phase 1 was adapted from the Atlantic Bluefin MSE work developed by Tom Carruthers (funded by the ICCAT GBYP project). Summary result graphics were adopted from the Fisheries Library in R (FLR) code provided by Iago Mosqueira.

1 Contents

2 Executive Summary

This report summarizes scientific and technical developments undertaken in support of the Indian Ocean Tuna Commission (IOTC) Management Strategy Evaluation (MSE) process for yellowfin and bigeye tunas from Jan2020 – Jul2021, as defined under the IOTC/FAO-CSIRO Letter of Agreement (FAO Budget Code TF/FIDTD/TFAA970097099). This project was conducted with the oversight of the IOTC Working Party on Methods (WPM), IOTC Working Party on Tropical Tunas (WPTT), the IOTC MSE Task Force (originally an informal sub-group of the WPM and de facto project steering committee, the task force was elevated to a formal subgroup of the WPM in 2021), with feedback from the IOTC Technical Committee on Management Procedures (TCMP). The following points summarize key deliverables from the project:

1. Continued iterative development of bigeye and yellowfin tuna Operating Models. This includes the processes of representing uncertainty, conditioning models (fitting to data), evaluating model plausibility, and updating the MSE projection software to add new structural features in line with feedback from the IOTC working groups, and to improve computational efficiency.
2. Development and evaluation of candidate Management Procedures, to achieve the management objectives, and subject to the constraints, articulated by the TCMP.
3. Presentation of results to the IOTC technical working groups, and participation in the deliberations related to improving stock assessment models and the representation of uncertainty, particularly as related to defining Operating Models.
4. Project outputs include a publicly accessible software archive (https://github.com/pjumppanen/niMSE-IO-BET-YFT/), a series of IOTC working and information papers, updated user manual and technical specification document for the software, quick reference "current state of the Operating Model" documents, and this final report, all of which will be archived in the git repository (along with the documentation of earlier project phases).

Progress toward the above objectives is summarized in an annotated bibliography of 7 IOTC working papers and a critique on the status of the MP development, highlighting ongoing concerns both in the software and methods employed by the authors, and the broader IOTC MP process. The current phase of the project has been completed, however the ongoing need for yellowfin and bigeye MSE scientific and technical support will continue, until either Management Procedures are successfully adopted, or the Commission abandons the effort. Funding has been secured to ensure that CSIRO is able to continue work until June 2023.

3 Acronyms and definitions used in this document

GEF/ABNJ Areas Beyond National Jurisdiction element of the recent Global Environment Facility project.

ABT Atlantic Bluefin Tuna.

BET Bigeye tuna.

CCSBT Commission for the Conservation of Southern Bluefin Tuna.

CPUE Catch per Unit Effort; usually assumed to be standardized into a relative abundance index for fish vulnerable to a particular fishery.

HCR Harvest Control Rule – the numerical algorithm for recommending a management action (e.g. providing a TAC given a biomass estimate). In this document, the term is generally not intended to encompass data collection and analysis or fitting a stock assessment model, as these are considered to be separate components of a complete MP.

ICCAT International Commission for the Conservation of Atlantic Tuna

IOTC Indian Ocean Tuna Commission.

IWC International Whaling Commission.

OM Operating Model – this usually refers to the combination of the generic projection software and suite of model specifications used to simulation test the performance of candidate MPs. We often refer to the OM projection software and OM parameterization separately.

MP Management Procedure – the simulation-tested combination of pre-agreed data collection methods, supporting analysis, and Harvest Control Rule. The term is often used interchangeably with MSE, however the *sensu stricto* MP definition (as used in the IWC and CCSBT) explicitly requires a very high level of pre-specification (i.e. of the data requirements and supporting analyses), to preclude the inherent risk of assessment groups failing to reach consensus during the application of an HCR. MSE is a broader term that does not necessarily imply the same degree of pre-specification.

MSE Management Strategy Evaluation – the process (or final product) of simulation testing a fishery management strategy (see MP).

MSY Maximum Sustainable Yield.

TAC Total Allowable Catch – the catch quota set by an MP (it could be fishery-specific or the aggregate across fisheries, depending on context).

WPM IOTC Working Party on Methods.

WPTT IOTC Working Party on Tropical Tunas

YFT Yellowfin tuna.

4 Introduction

This project report is structured to identify the various objectives and achievements under the FAO-CSIRO Letter of Agreement (FAO Budget Code TF/FIDTD/TFAA970097099), and orient the reader to the appropriate references to follow up on details. There is no synthesis of working papers into a single narrative that re-iterates all of the technical documents and committee deliberations over the past three years. The background below is largely extracted from the phase 1 and 2 final reports (Kolody and Jumppanen 2016, Kolody and Jumppanen 2019). Section 5 itemizes specific objectives and achievements, while section 6 identifies key challenges that remain in the IOTC bigeye and yellowfin Management Strategy Evaluation (MSE) process.

4.1 Background

In pursuing the objectives of achieving conservation and optimum utilization of tuna stocks, the Indian Ocean Tuna Commission (IOTC) committed to pursuing Management Strategy Evaluation (MSE) for the key target species of swordfish and albacore, skipjack, bigeye and yellowfin tunas (IOTC 2011).

MSE is a process in which a fishery system, including the fish population, fishery, and management decisions, are simulated over a medium to long term time horizon, and performance of the management system is evaluated with respect to explicit management objectives (e.g. see Punt et al. 2014 and references therein). A computer simulator (Operating Model, OM) is intended to describe the main uncertainties in the system, including the current state of the fish population and stochastic future events. The Management Procedure (MP) is the algorithm that recommends a unique management action given the data, and is applied at pre-determined intervals. The MP should use feedback control, to change the management action in response to new information about the changing state of the fishery. Simulation-tested MPs offer many advantages over the traditional cycle of stock assessment and ad hoc decision making, including: i) MPs should be robust to the main uncertainties in the system (i.e. provide reasonable management performance regardless of the true underlying dynamics), ii) MP performance is evaluated against multiple explicit management objectives potentially spanning multiple time horizons, and iii) pre-agreement on data collection, analytical methods and harvest control rules pre-empts disagreements about management actions arising from a failure to reach a consensus assessment.

In this project, we have aimed to evaluate MPs using the *sensu stricto* definition, in which the MP explicitly includes the specification of the data collection and analytical methods to be used, in conjunction with a Harvest Control Rule (HCR), e.g. as used in the southern bluefin tuna fishery (Hillary et al. 2016). In some other applications, the MP does not include the internal specification of data collection and analysis. For example, IOTC Resolution 16/02 prescribes an HCR that assumes a sensible skipjack stock assessment will always be available. In the skipjack case, the simulation testing involved simulating stock status outputs with a known degree of accuracy and precision. Given that the former approach requires assumptions about simulated assessment data observation errors, the distinction between the two approaches may appear subtle. But assessment model inferences are often biased in ways that are difficult to anticipate, particularly when there are substantial structural errors in the model. Simulating the assessment process is probably the best way to reliably represent these potential biases (including the time series structure of the errors). However, there is a more important operational distinction between the two approaches that is critical when assessment bodies are unable to reach a consensus view of the stock status (and indeed adoption of the MP approach was motivated by this problem in some international fisheries organizations). The *sensu stricto* MP approach explicitly pre-empts the problem of conflicting assessments, because the MP data and analyses are agreed in advance (and simulation tested to ensure that performance is robust to alternative plausible assessment interpretations). The stock assessments for the IOTC fisheries have undergone substantial changes in recent years, and it would not be surprising if they continue to evolve in the foreseeable future, such that consensus is not inevitable. The *sensu stricto* approach may have the further advantage that resources required for the traditional stock assessment process should not be required every time that the MP is evaluated. If the internal MP "assessment" is a straightforward mechanical calculation, this potentially frees up assessment resources for other strategic research needs (in practice this might be an oversimplification, particularly for contentious stocks with unreliable data).

The MSE process can be partitioned into a series of steps (represented schematically in Figure 1):

1. Identification of management objectives and quantifiable performance measures
2. Development of a range of Operating Models (OMs) to represent the uncertainty in the fishery
3. Development of candidate Management Procedures (MPs)
4. Simulation testing of candidate MPs using the OMs
5. Selection of a preferred MP based of the simulated performance with respect to the management objectives (performance measures)
6. Implementation of the MP

The process is rarely a linear sequence, as individual steps tend to be iteratively revisited as information is exchanged among participants, and decision makers come to understand the possible performance trade-offs. It is useful to think about MSE within the broader context of fishery management as shown in Figure 2. Following adoption of an MP, it should not be expected that the MP will continue to manage the fishery in perpetuity, on "auto-pilot" (i.e. the “meta-rules” in Figure 2 exert a higher level control on the MP). The MP should include periodic performance reviews, to ensure that the MP is meeting the management objectives, and that the management objectives remain appropriate. It would be optimistic to expect that MPs will always perform well, and there should be regular scientific oversight of the fishery to check whether the system has moved into a space that was not encompassed by the original simulation testing. It is probably not possible to anticipate all of the "exceptional circumstances" which could arise, but possible problems include: i) new observations may indicate that the OM understated the uncertainty of the system (e.g. previously unrecognized biological uncertainties become evident), ii) the fisheries data may cease to be as informative as expected (e.g. longline CPUE may no longer be available on the spatial and temporal scales used historically), iii) management actions may not be as effective as expected (e.g. due to IUU fishing), or 4) new data may become available that can improve assessment and management. If exceptional circumstances arise, they should be examined to see if they affect the management recommendation, and the MP may need to be suspended until the issues can be resolved, or a new MP evaluated and implemented.

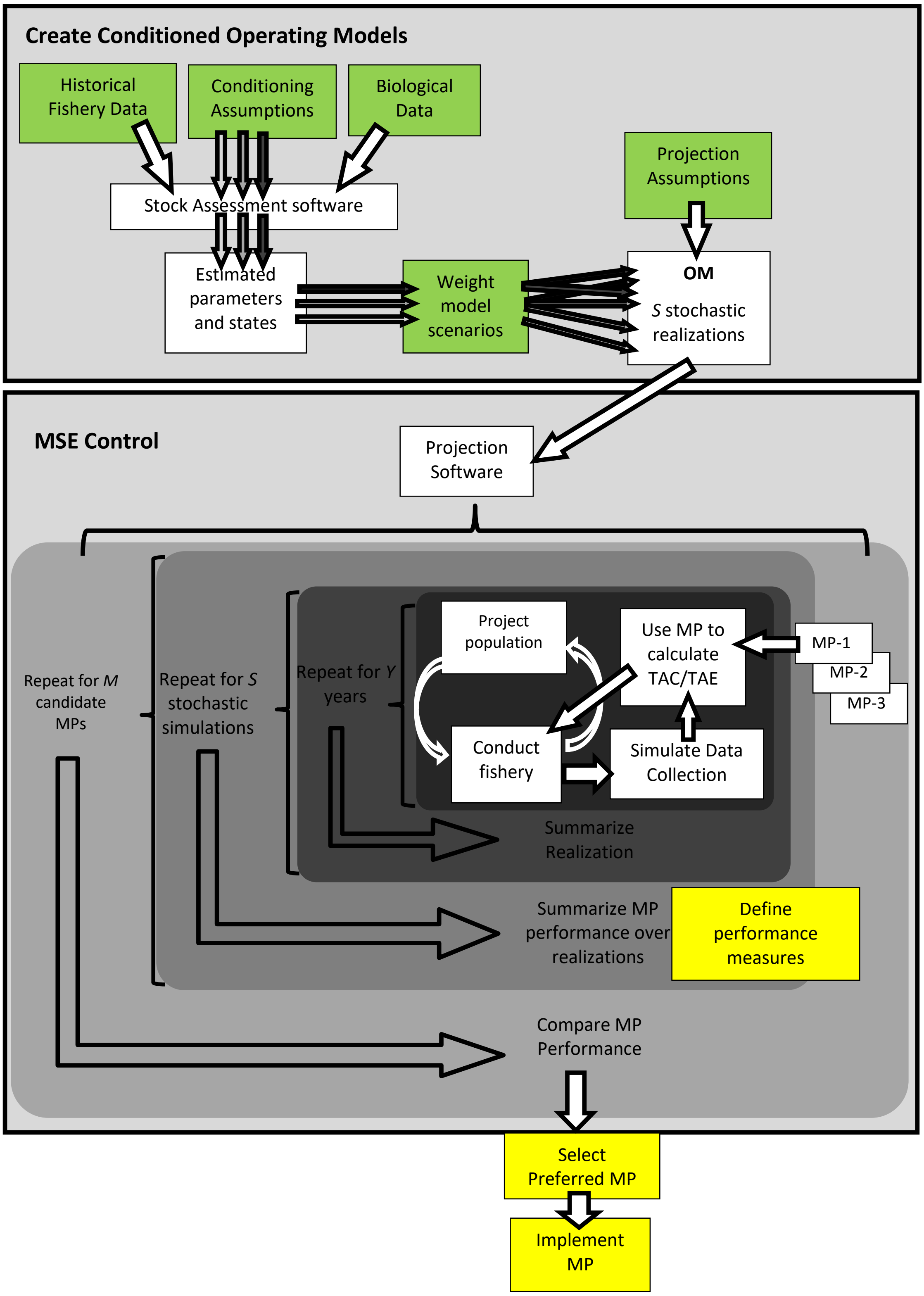

Figure 1. Schematic flowchart of the MSE process, emphasizing the technical elements as implemented in this project. Key points for integrating broader scientific input are highlighted in green, while other stakeholder and manager (Commission) inputs are highlighted in yellow.

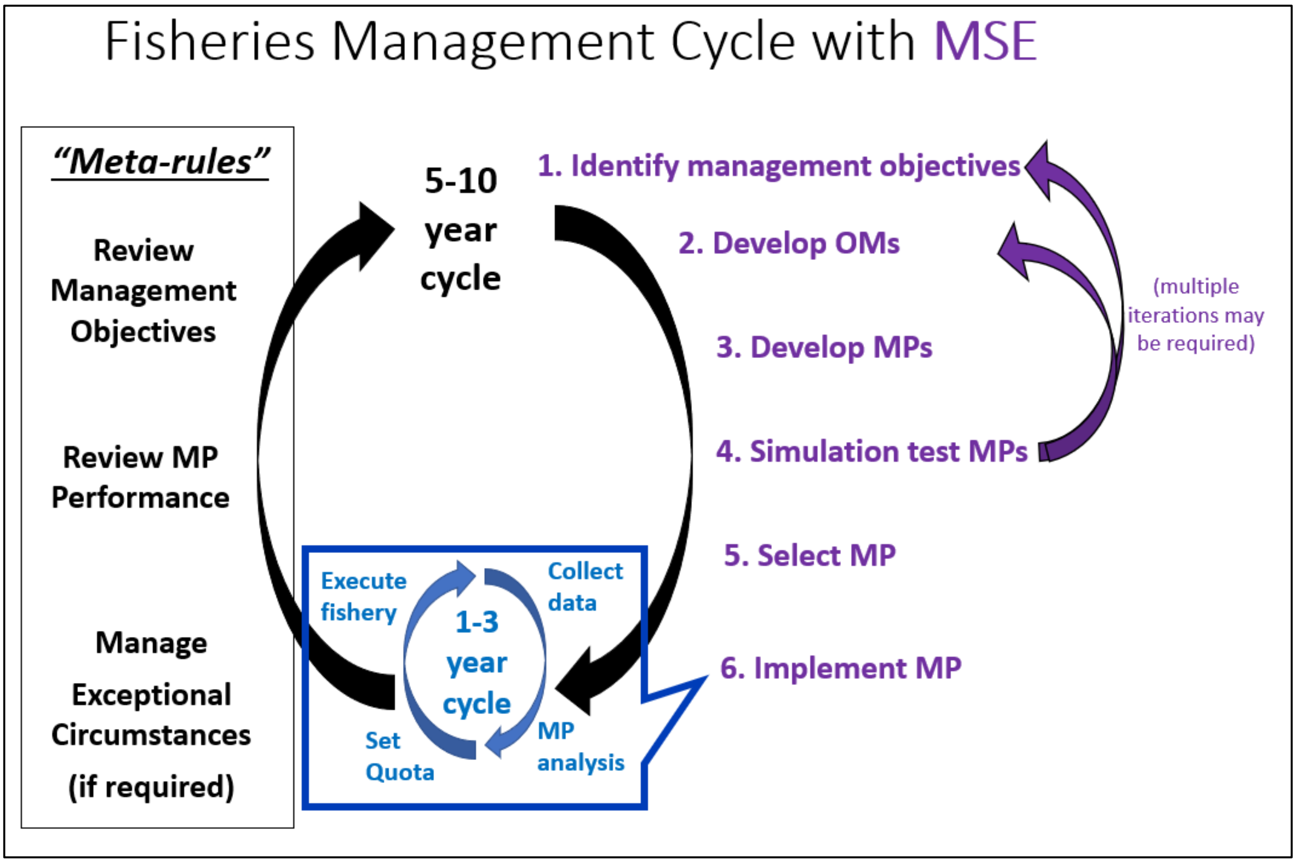

Figure 2. Schematic of a fisheries management cycle in which the MSE process (numbered 1-6) is one part of a larger cycle.

5 IOTC Yellowfin & Bigeye Tuna MSE Phase 3 Objectives and Achievements

The FAO-CSIRO Letter of Agreement identified a number of objectives pertaining to the development of the bigeye and yellowfin Management Procedures. The overarching objectives include:

- Ongoing development of the MP evaluation software.
- Updating of Operating Models in collaboration with the relevant experts at the IOTC WPTT and WPM, and in parallel with the stock assessment processes.
- Development and evaluation of candidate Management Procedures.
- Presentation of results to the IOTC technical working parties (and other scientific bodies as appropriate) for feedback.

5.1 Project reports

Progress is summarized in the series of documents listed below, organized by meeting (archived in the IOTC website, and publicly available from https://github.com/pjumppanen/niMSE-IO-BET-YFT/, and). Documents are cited in reverse chronological order below (results from more recent documents supersede older documents).

Jul 2021 - Phase 3 Final Report (this document)

Jun 2021 - IOTC Technical Committee on Management Procedures

- Two working papers were submitted to the TCMP-04 2021 (Kitakado 2021) describing progress and soliciting feedback:
  - IOTC-2021-TCMP04-08 - *IOTC Bigeye Tuna Management Procedure Evaluation Update June 2021*
    - The working paper presented a contrasting range of tuned MPs, all of which had similar behaviour as determined by the tuning objectives. However, subtle differences in the time series behaviour of the MPs suggested that the model-based MPs are probably more responsive to incoming data and more attractive in the long term. TCMP (2021) endorsed a proposed workplan, timeline, tuning objectives and potential TAC constraints for the continued development of the BET MPs.
  - IOTC-2021-TCMP04-09 – *IOTC Yellowfin Tuna Management Procedure Evaluation Update June 2021*
    - The paper describes the critical failure of the yellowfin OM ensemble (detailed in the March 2021 WPM meeting below), that renders the OM implausible and MP results meaningless at this time.

Mar 2021 – IOTC Working Party on Methods (MSE Task Force) Meeting (WPM 2021)

- Two working papers were presented to this technical meeting:
  - IOTC-2021-WPM12(MSE)-03 – *IOTC Yellowfin Tuna Management Procedure Evaluation Update March 2021*
    - The yellowfin OM was found to be critically flawed. Despite the substantial uncertainty grid, most or all elements of the ensemble could not extract the catches reported in 2019 (2018 and 2019 data were not included in the OM conditioning). The problem appears to be related to the issues with the recent yellowfin stock assessments. The OM ensemble is in general more pessimistic than the assessments, in part because 50% of the OM elements assume a continuous historical increase in longline catchability. This imposes a hyperdepletion effect, making the stock status more pessimistic than the longline CPUE series would otherwise suggest. It was recognized that the OM will need a full reconditioning, in line with 2021 stock

assessment developments. However, there is no guarantee that the assessment will find an easy solution, and the appropriate path for the OM might require a rethinking of the fundamental interpretation of some data sources, and an expansion of the uncertainty grid in new directions.

- IOTC-2021-WPM12(MSE)-04 – *IOTC Bigeye Tuna Management Procedure Evaluation Update March 2021*
  - The bigeye OM appeared to be well-behaved, with MP evaluation results that appear reasonable and similar to previous iterations. The WPM did not identify any need to modify the OM ensemble or tuning objectives (except to clarify target dates in relation to shifting timelines).
  - A promising new MP was presented, which involves i) fitting the joint process and observation error Pella-Tomlinson production model (described in IOTC-2020-WPM11 below), and ii) conducting internal projections to solve for a constant TAC that would hit a pre-specified depletion target at a pre-specified future date. The approach is analogous to how the Commission might interpret a Kobe 2 Strategy Matrix - the MP would be applied, and the TAC updated, with the usual (3 year) frequency, regardless of the projection date adopted within the MP.

Nov 2020 – IOTC Working Party on Methods (WPM 2020) and Working Party on Tropical Tunas (WPTT 2020)

- Three working papers were presented:
  - IOTC-2020-WPM11-11 – *Indian Ocean Bigeye Tuna Management Procedure Evaluation Update March 2020* (*originally prepared for the cancelled 2020 Task Force meeting)*
    - The BET OM was updated with respect to the WPTT/WPM 2019 requests, including adopting the new 2019 assessment for the core data and structural assumptions. Various analyses were undertaken, including investigation of retrospective patterns, the influence of iterative re-weighting on model inferences, and the importance of a non-trivial catch likelihood value (an indicator that at least one time/area/age strata cannot remove the reported catch). Results from a small sample of MPs suggested that tuning objectives were reasonable, and performance qualitatively similar to the previous iteration.
  - IOTC-2020-WPM11-12 - *Indian Ocean Yellowfin Tuna Management Procedure Evaluation Update April 2020.* (*originally prepared for the cancelled 2020 Task Force meeting)*
    - The working paper attempted to address a number of requests arising from the WPTT (2019) and WPM (2019), which involved reconditioning on the basis of new stock assessment work by an international collaboration. It was recognized

that several of the OM requests were problematic for reasons that were not fully appreciated during WPTT (2019), in part due to time constraints and the fqact that the revised assessment was not adopted by the WPTT. Exploration of a range of new structural and parameter assumptions is described, along with efforts to investigate retrospective patterns and iterative re-weighting to improve models. However, the conclusion was that it would be difficult to justify updating the yellowfin reference set OM without further debate in the relevant working parties.

- IOTC-2020-WPM11-13 - *A candidate Management Procedure based on a Joint Process and Observation Error Random Effects Production Model*

    - Working Paper 13 described new developments in model-based MPs. These MPs use a joint process and observation error Pella-Tomlinson production model (implemented with random effects in TMB software) and have proven to be very robust and numerically efficient.

    - Some of the project resources were dedicated to the development of the R package *Buildsys*, which enables sophisticated modern debugging of TMB code (and other C/C++ based applications called through R), described in Jumppanen (in prep.)

2020 TCMP – cancelled due to Covid-19 disruptions

2020 MSE Task Force Mtg – cancelled due to Covid-19 disruptions

5.2 MSE Software developments

All of the source code, compiled executables, and scripts required to reproduce the results described in this report are publicly available for unrestricted use from Github (https://github.com/pjumppanen/niMSE-IO-BET-YFT). Stock Synthesis model fitting results are not publicly archived, because of the file size (but could be circulated if necessary). The OM conditioning process and workflow has not changed substantially since phase 2. However, it is recognized that new methods for evaluating the plausibility of models are desirable, to hopefully avoid the sort of problems observed for yellowfin in recent years (particularly evident in 2021). This might include the suite of diagnostics under investigation for the stock assessment models (e.g. Matsumoto et al 2018).

The OM projection software remains functionally almost the same as the phase 1 product, with most of the new extensions related to utilities for manipulating results and efficiency improvements on the backend of the system. Currently, a simple MP can be tuned (~14 MP evaluations) with a 500 realization OM in less than an hour on a standard laptop PC. The projection software still supports parallel R and C++ projection dynamics. The overhead of maintaining both options will be a burden if substantial changes are required, but to date this proved worthwhile from the perspective of identifying bugs and quantifying structural sensitivities.

In phase 3, considerable effort was taken to improve the model-based MPs. Using TMB software enabled much more numerically-efficient implementation of production models which include joint process and observation error in a random effects state-space configuration.

6 Critique of the IOTC yellowfin and bigeye MP development process

Following the format of the previous final project reports, the following is a compilation of concerns identified by the authors, related to both the MSE scientific and technical development, and issues with the process that may impede the eventual adoption of MPs. These comments are not intended to undermine the MSE process, but rather flag challenges that might benefit from further consideration from the broader scientific community, including reviewers, or the next generation of developers that might carry this work forward in the future. Concerns from the phase 1 and 2 final report are repeated if they remain relevant.

6.1 Operating Model Projections

1) The R-based and C++ based projection sub-routines use different implementations for the catch equations, and a decision will ultimately have to be made to select one or the other for MP evaluation. When compared, the two implementations provided very similar MP evaluation results, except when fishing mortality rates are very high in the yellowfin spatial model. It is conceivable that the choice of sub-routine could influence MP selection in the yellowfin case. However, this conclusion is misleading in the sense that we would question whether either approach is sufficiently realistic when fishing mortality rates are so high that quotas cannot be met, catch rates are uneconomical, and incentives exist for fleets to stop fishing, move among areas and/or switch species targeting. We currently emphasize results from the C++ code because it is more consistent with the SS conditioning model, and is slightly faster. Furthermore, the C++ code offers a mechanism for investigating the plausibility of models in relation to the maximum fishing mortality constraint. e.g. Most or all of the recent conditioned yellowfin OMs could not remove catches that were reported in years (2018-2019) subsequent to the data used in conditioning (2017) with a 20-fold increase in *F* from the recent past. While we are not sure what a plausible increase in effort might be over that short time period, its presumably less than a factor of 2 for the large industrial fisheries. The only downside of the C++ sub-routine is the specialized knowledge (and software dependencies) required to maintain it.

2) The OM was structured to expect exactly one CPUE series from each area. If additional CPUE series are added (as proposed for exploration in the 2021 yellowfin stock assessment), this will require OM code changes. Furthermore, the simulated CPUE output for the candidate MPs is a single relative abundance index (occasional years with missing strata were removed from the MP time series). This aggregate index is compared with the aggregate biomass from the conditioned SS models to ensure that the historical error characteristics can be approximated into the future (i.e. this attempts to account for the potential importance of correlated errors across regions, seasons and years). If additional CPUE series are included in the conditioning, further consideration will be required as to how these additional CPUE series should be simulated for the MPs.

3) Independent reviewers (for multiple IOTC stocks) have suggested that alternative stock-recruit functions (e.g. Ricker) should be considered. We are not aware of evidence from bigeye or yellowfin tuna population to suggest that recruitment initially increased with declining *SB* (as might be predicted by a Ricker function), or that there is a mechanism that would lead one to expect a Ricker function (e.g. high rates of cannibalism or redd superimposition), so we have not considered this to be a priority. Some conditioning tests were undertaken with a Ricker SR function in 2019. The estimated Ricker functions were qualitatively not much different from a Beverton-Holt function (i.e. minimal dome-shape over the observed range of *SB*), such that we would not expect that this recruitment uncertainty would introduce any fundamentally new and insightful challenges to the MPs (at least no more than would be achieved by manipulating steepness, which is already represented in the OMs with multiple fixed values). The status of this issue has not changed since phase 2.

4) Stationarity assumptions – to produce tractable estimators, it is common to assume that many parameters are stationary in population models, even though this is probably often not entirely correct. Non-stationarity in *M*, growth and recruitment processes could have important implications for production dynamics. Non-stationarity in recruitment distribution and/or movement dynamics might introduce the need to reconsider fleet distribution dynamics. The flexibility to add non-stationarity was added to many of the model parameters in phase 2. However, the developers have refrained from open-ended speculation in testing these features. We would argue that such tests should be requested by the relevant experts in the IOTC scientific community, with precise definitions (and perhaps a broad discussion of plausibility, likely importance and priorities).

5) MSY-based reference points depend on fishing selectivity and biology, and are not stationary if the biology changes or selectivity changes (including a change in the relative fishing mortality among fisheries with stationary individual selectivity, as would be observed with fishery-specific quotas allocated with constant relative catches). In the MSE code, the reported reference points are fixed, based on the "recent" effective effort distributions (i.e. estimated fishing mortality). We would suggest that, in the interest of internal consistency, the approach of using "proxy" reference points based on depletion would be preferable for bigeye and yellowfin (as adopted for skipjack). In phase 2, the OM code was modified to allow MSY reference point calculations and reporting based on arbitrary future catch allocations (fixed catch proportions assuming stationary selectivity). TCMP (2019) requested a review of IOTC reference points, and the WPM is currently overseeing the work of that group.

6) There is some evidence to suggest that Indian Ocean yellowfin and bigeye population connectivity might be more complicated than the homogenous mixing assumed in most assessments to date. With this in mind, the projection code partially supports the option of multiple stocks with independent biology, but the implementation has not been tested, and stock-specific population diagnostics are not reported at this time. The recent IOTC stock structure project (e.g. Davies et al. 2020) is the most comprehensive investigation to date, and suggests that yellowfin stock structure probably represents a greater concern than bigeye.

7) Ignoring the multi-species technical interactions in tropical tuna fisheries may limit the utility of MP evaluations (or at least increases the importance of management implementation errors). As noted in the previous point, the projection code partially maintains the feature of multiple independent populations within a species, and it should be reasonably straightforward to extend this feature to a multi-species context. However, it is not a trivial modification, and conditioning would require the data for the different species to be supplied with compatible fishery definitions and spatial structures. Representing the joint uncertainty of multiple species could presumably be easily achieved by independently sampling the species-specific OMs. However, the WPM has expressed the view that trying to address this issue now risks derailing the single species MP evaluation progress that has been achieved to date.

8) Tag dynamics have not been included in the OM projection model. It may prove desirable in the future to simulate the collection and analysis of conventional or genetics-based mark-recapture methods to evaluate candidate MPs with the potential to use this (somewhat) fisheries-independent monitoring capability. This has probably become increasingly relevant, as the WPTT and WPM struggle to produce a satisfactory yellowfin stock assessment, and Close-Kin Mark Recapture appears to be a viable monitoring option for the near future (e.g. ~5 year time horizon from initiation).

6.2 Operating Model Conditioning

1) While the MSE software supports the option of differentially weighting different SS configurations in the OM ensemble, this has not been employed to date. Likelihood weighting provides an obvious candidate for doing this, but this is not very attractive because i) the likelihoods are far from ideal (with many arbitrary manipulations of effective sample sizes and variances etc.), and ii) this form of weighting could only be applied across OM subsets that use the same data in the same way. There is also the qualitative observation that most models actually fit the data reasonably well (or at least, the best fit model often does not appear to be visually much worse than the worst fit model). In practice, there has also been an initial filtering stage, in which some models may be removed (i.e. given weight 0). The exclusion criteria (some of which are somewhat arbitrary) include:

- *Numerical failure – if an individual model failed to reach the gradient convergence threshold it would be excluded (the number of models with this problem was greatly reduced with the adoption of a repeated jittered minimization for every model).*
- *Outlier behaviour – if one or a very small number of models exhibited extremely different behaviour with respect to some stock status or quality of fit indicator it was a candidate for exclusion (not necessary in the final iteration of phase 3 for bigeye)*
- *Catch extraction problems – a non-trivial catch likelihood occurs when Stock Synthesis cannot remove the reported catch for at least one time-age-region strata. This could indicate an overall shortage of fish, or a problem with the spatial distribution. In some bigeye cases, this may reflect a trivial matter of a highly seasonal historical fishery where the model cannot represent seasonality reliably. The problem was more common and serious for yellowfin, and manifested as a fatal projection problem. i.e. Most or all elements of the final yellowfin OM ensemble in 2021 demonstrated a comprehensive failure to remove the catches that were reported after the data that were available for conditioning.*

2) The initiative to bring a more formalized approach to model evaluation diagnostics to the IOTC stock assessment process (e.g. Matsumoto et al. 2016) will presumably have value for filtering and /or weighting the OMs as well. The problem of the unattainable recently-reported catches in the yellowfin context is a good example of evaluating predictive capacity, which is analogous to what the retrospective hindcasting attempts to do. However, we are not convinced that the best approaches have been identified yet, for situations that are more subtle. Ensuring that models are internally consistent is a useful first step, but is not sufficient to guarantee that the models are representative of reality. e.g. There may be multiple different ways to alter a model to minimize a particular diagnostic problem, and unless the specific cause of the problem can be identified, the solution adopted might not be fixing the right part of the model. The southern bluefin tuna assessments prior to 2006 provide an instructive example, in which scientists invested considerable effort representing OM uncertainty in multiple dimensions, but failed to even consider the massive catch under-reporting that was eventually revealed in the market statistics (e.g. Kolody et al. 2008).

3) Ignoring parameter estimation uncertainty (for those parameters not included in the OM grid) might provide an understatement of some key uncertainties. We have attempted to avoid understatement of uncertainty by using the grid-based approach, and adding a number of projection options, including user-defined parameters for some process and observation errors, (CV and auto-correlation), and additional error on the initial numbers-at-age. However, we recognize the arbitrary and subjective nature of this process and other IOTC scientists may have good arguments for alternative assumptions. We expect that this issue would be most influential when inferences are made about MP evaluation outcomes near the tails of distributions.

4) The IWC and CCSBT MP processes found value in distinguishing between "reference set" and "robustness test" OMs (e.g. Punt et al 2014). The reference set OMs were intended to provide a general description of the fishery with a reasonable description of uncertainty evident from the historical data. In contrast, robustness test OMs comprised a series of OMs that were considered to have a low probability, but potentially very negative consequences (e.g. recruitment failure, large IUU catches). Over the years, the technical working groups have proposed many robustness scenarios for both bigeye and yellowfin. These have been addressed in different ways: i) some uncertainty dimensions have been tested and either rejected as implausible (or of minimal influence to behaviour) or were found to be important and elevated to the reference set OM, ii) several standard robustness tests are now presented alongside the reference set MP evaluation results, iii) some of the robustness requests have been referred back to the technical working groups for further clarification and prioritization. In 2021, robustness tests were presented to the TCMP for the first time (for bigeye). To date, all of the robustness tests that have been formally presented have been based on the historical conditioning of the reference set OM, and only differ from the reference set in terms of the projections. Conceivably, some robustness test OMs might require reconditioning. Given the implausible pessimism of the most recent yellowfin OMs, there seems to be reasonable grounds for reconsidering whether the longline increasing catchability trend remains appropriate in the reference set OM.

6.3 State of the Bigeye Reference Set Operating Model

The bigeye reference set OM was updated in relation to the 2019 stock assessment, and no modifications were requested following review of the bigeye OMs in WPM (2020), WPTT (2020) and WPM(MSE) (2021). The inferences from the MP evaluation results presented to TCMP (2021) were qualitatively very similar to the previous iteration presented to TCMP (2019). It appears that the scientific community is satisfied with the status of the OM, and the bigeye MP development process should be aiming to get endorsement of the bigeye OM from the Scientific Committee to move toward the next stage of MP adoption.

6.4 State of the Yellowfin Reference Set Operating Model

The yellowfin OM was found to be critically flawed. Despite the substantial uncertainty grid, most or all elements of the ensemble could not extract the catches reported in 2019 (2018 and 2019 data were not included in the OM conditioning). This is clearly a prediction failure, supported by real data. It was somewhat foreshadowed in the retrospective analyses that have been undertaken in recent years and is presumably shared with the recent yellowfin stock assessments from which the OM is derived. However, the effect is more extreme in the OM ensemble because it is generally more pessimistic than the assessments. This arises in part because 50% of the OM elements assume a continuous historical increase in longline catchability (i.e. this makes the stock status more pessimistic than the standardized longline CPUE series suggests). It was recognized that the OM will need a full reconditioning, in line with 2021 stock assessment developments. However, there is no guarantee that the assessment will find an easy solution, and the appropriate path for the OM might require a rethinking of the fundamental interpretation of some data, and an expansion of the uncertainty grid in new directions. It is possible that the pessimism of the OM (and assessment) arises from a fundamental misinterpretation of the CPUE series (e.g. the WPTT generally accepts that there is CPUE hyper-depletion prior to ~1972, and it is not clear when exactly or why it would have stopped). The increasing catchability trend in 50% of the reference set OMs clearly exacerbates the pessimism, and may be difficult to justify at this time (at least when combined with the other model assumptions in the OM ensemble).

6.5 Candidate Management Procedures

1) The aggregated performance diagnostics suggest that there is relatively little difference in performance expected among the range of tuned MPs tested to date. Presumably this is because there are a limited envelope of options with which to hit the tuning objectives, and all MPs to date are critically reliant on the CPUE series to track relative abundance. We would argue that the joint process and observation error production model, combined with internal projections, provides the most desirable MP behaviour (in terms of not overshooting biomass targets and minimizing catch variability). The CPUE-based MPs appear to offer the worst performance given the same criteria. However, we cannot be certain of the extent to which this is an artefact of not testing enough control parameter combinations. It is notable that the model-based MPs have a very high probability of hitting the (15%) TAC change constraints relative to the CPUE-based MPs. Perhaps the model-based MPs could be made less responsive, or the CPUE-based MPs would benefit from being more responsive. However, it is also possible that the current CPUE-based MPs simply cannot learn to distinguish among OM elements. The CPUE-based MPs always aim for the same CPUE target with the same responsiveness for all OM populations. It makes sense that the model-based MPs might have more capacity to "learn" and differentially respond to each OM population.

2) We have generally not been examining or reporting the specific values of the MP tuning parameter and it is possible that some of these values could be difficult for Commissioners to accept, even if the MP performance appears desirable. e.g. A very high or low target biomass or CPUE could be required within an MP HCR to hit the tuning objective. The MP target could be counter-intuitive (possibly unattainable) and very undesirable if taken out of context. This could arise if a production model is very biased, such that a strange HCR is required to compensate for the bias in the context of the reference set OM. But in the case of yellowfin there is often a disconnect between the catch and TAC, that arises because of the general pessimism of the current yellowfin OM ensemble. Hopefully the implications for the tuning parameters will resolve, with an improved reference set OM (and stock assessment).

3) The potential for implementation errors (catch not equal to TAC) to affect MPs might merit further consideration. There are a number of robustness tests that examined the most obvious potential - reported or unreported over-catch. But there are less intuitive ways in which this could be an issue. The first is the bigeye situation, in which catches appear to be dropping for several years without restrictive quotas, with no obvious incentive for catch reduction. The developers added an additional robustness test (see IOTC-2021-TCMP04-08) to see if several years of under-catching the quota could lead to an information disconnect that would be problematic if quotas were to suddenly start being restrictive. But this issue is also relevant in the context of yellowfin, in which quotas could not be removed (primarily due to the pessimism in the conditioned estimates of abundance, but also movement dynamics to a much lesser degree). We expect this undercatch for yellowfin is an artefact of the current conditioning problem that affects the OM (and assessment). Some facility for relatively minor carry-over of under-catches and over-catches from one year to the next would be worth considering, and provided that the over- and under- catches are small, there would probably be no need to explicitly evaluate this within the MP testing. The MPs examined to date all prescribe quota changes relative to the previous quota regardless of the reported catch – if there is a disconnect, perhaps there is a role for the catch in the MP decision rule.

4) None of the candidate MPs under consideration attempt to use information from the size composition data. Provided that selectivity (and associated assumptions of growth and *M*) and sampling characteristics are stationary, size data could provide useful information about incoming recruitment or spawning biomass. This was the intent of the size composition data in the Prince et al. (2011) empirical MP, however, the species-specific evaluations of that MP (in Kolody et al. 2010) failed to demonstrate that the size composition data improved performance relative to the purely CPUE-based MPs (even under conditions that are much simpler and probably closer to ideal than the IOTC situation). If these approaches are to be pursued, this will require further consideration of how to realistically simulate catch-at-length sampling errors (and non-stationary selectivity). In conjunction with the size composition data, it may be desirable to add some simple age structure to the MP models (e.g. perhaps distinguishing recruits and spawners, in a manner similar to Hillary et al. 2016).

5) Model-based MPs appear to have a performance advantage, but have additional practical problems of numerical stability and efficiency that are not a concern for the empirical MPs. It is important to avoid the awkward situation of creating an MP that offers the best performance subject to flawed model fitting. The model-based MPs implemented using TMB appear to offer markedly better performance than the MPs used in phase 2. Despite a considerable effort to stabilize the numerical characteristics, there is still a small risk of the TMB models failing in particular circumstances. There appears to be a rare and difficult to reproduce bug, related to R-TMB communication, possibly related to a memory leak. The problem can seemingly be minimized by using a PC with a very large surplus of memory, and MP evaluation results can be checked to see if the issue occurred.

6.6 The broader IOTC MSE process

1) MP communication within the TCMP. Despite the ongoing efforts at capacity building (at the TCMP, in the technical working groups, and other initiatives, e.g. GEF/ABNJ sponsored workshops), it seems clear that the majority of TCMP participants are not yet comfortable with their understanding of Management Procedures, the role of the different stakeholders in developing MPs, or how the process of selecting and adopting an MP will work. This is evident from the formal discussion within the TCMP, and the informal (confidential) surveys undertaken at the end of TCMP03-2019. Fortunately, many participants expressed the view that they are learning at the TCMP, and there are several individuals, among multiple delegations, that appear to have a workable understanding of MPs. A reasonable outcome for the IOTC might be achieved if there is a small critical mass of sufficiently-informed individuals spread among the CPCs, that have the respect and trust of other like-minded CPCs (e.g. in terms of representing the general interests of coastal state or distant-water fishing nations).

2) The bigeye MP is at the most advanced stage of development among the IOTC species, and should probably be moving to the next stage of the adoption process. This will require the endorsement of the OM by the Scientific Committee (which should be straightforward given the general acceptance among the technical working groups), and sponsorship of an MP resolution by CPCs for Commission consideration. Peer review should be part of this process. A draft yellowfin MP resolution (IOTC-2019-S23-PropP) was submitted to the 2019 Commission (it was intended primarily for educational purposes and withdrawn for future development and resubmission). Bigeye should probably be given a higher priority for adoption, as the process of developing a plausible yellowfin OM clearly requires further effort.

3) Catch allocations – MP performance will depend to some degree on the catch allocations among fisheries, which are currently being negotiated in a separate process (Technical Committee on Allocation Criteria (TCAC). To date, all of the MP simulations assume that catch proportions will be allocated in line with "recent" historical catches. The 2019 Commission (IOTC 2019) expressed interest in seeing MP evaluation results from alternative quota allocations. While alternative allocations are easily testable, the request was not clearly defined. It should not be the role of the scientists to speculate about (or be seen to be implicitly advocating of behalf of) these sorts of political decisions.

4) Tuning objectives – The TCMP has proposed tuning objectives that encompass a modest range of MP performance, that appear to be sensible and illustrate the important point that the tuning objective is the critical mechanism by which the TCMP has simple and direct control over how an MP performs. However, the probability of being in the Kobe green zone X% of the time from 2030-3034 is not necessarily easy to visualize. In IOTC-2021-TCMP04-08 (and earlier phase 2 working papers), we pointed out that the bigeye tuning

objectives could be redefined into almost equivalent units, that map directly onto some of the standard MP graphical outputs. This might be more convenient for communication purposes. However, since these tuning objectives seem to now be agreed for all IOTC species except yellowfin (and possibly skipjack), it is not clear that attempting to change the tuning units now would be helpful.

7 Concluding Remarks

The phase 3 project has successfully achieved the objectives set out in the FAO-CSIRO Letter of Agreement and the IOTC appears to be moving toward the ultimate goal of adopting Management Procedures for bigeye and yellowfin tunas. At the scientific level, the level of engagement has increased over time, with carefully-considered proposals for alternative modelling assumptions, and a much broader collaborative approach to the stock assessment (for yellowfin in particular) that acknowledges the link to Operating Model development. At the management level, TCMP participants (at least a core group and hopefully a critical mass) appear to have embraced the concept of MP tuning to quantify explicit medium-term management objectives. We note that the CCSBT MP development process was originally targeted to take 2-3 years, but ultimately took around 10 years. Given the far greater complexity of the IOTC fisheries, and the limited scientific capacity among many of the CPCs, it is perhaps not surprising that the IOTC yellowfin and bigeye MP processes are also taking longer than originally expected.

8 References

Davies, C, Marsac, F, Murua, H, Fahmi, Z, Fraile, I. 2020. Summary of population structure of IOTC species from PSTBS-IO project and recommended priorities for future work. IOTC-2020-SC23-11_Rev1.

Hillary, RM, Preece, AL, Davies, CR, Kurota, H, Sakai, O, Itoh, T, Parma, AM, Butterworth, DS, Ianelli, J, Branch, TA. 2016. A scientific alternative to moratoria for rebuilding depleted international tuna stocks. Fish and fisheries 17: 469–482.

IOTC 2011. Report of the fifteenth session of the Indian Ocean Tuna Commission. Colombo, Sri Lanka 18-22 March 2011.

Jumppanen, P. in prep. Introduction to BuildSys: An R Package for Compiling and Debugging R Dynamic Libraries. The R Journal XXX.

Kitakado T. 2021. Presentation of the TCMP-04 (2021) Meeting to the 2021 Commission Meeting. https://www.iotc.org/documents/tcmp04-report-kitakado.

Kolody, D, Polacheck, T, Basson, B, Davies, C. 2008. Salvaged pearls: lessons learned from a floundering attempt to develop a management procedure for southern bluefin tuna. Fisheries Research 94: 339-350.

Kolody, DS, Preece, AL, Davies, CR, Hartog, JR, Dowling, NA. 2010. Integrated evaluation of management strategies for tropical multi-species long-line fisheries. Final report for FRDC project 2007/017. 213p.

Kolody, D, Jumppanen, P 2016. IOTC Yellowfin and Bigeye Tuna Management Strategy Evaluation: Phase 1 Technical Support Project Final Report. Indian Ocean Tuna Commission Working Paper IOTC-2016-WPTT18-32

Kolody, D, Jumppanen, P. 2019. Scientific and Technical Support for Indian Ocean Yellowfin and Bigeye Tuna Management Procedure Development: Phase 2. Final Report for GEF/ABNJ/FAO-CSIRO Project FAO Budget Code TF.FIDFD.TF5G110014024.

Matsumoto, T, Yokoi, H, Satoh, K, Kitakado, T. 2018. Diagnoses for stock synthesis model on yellowfin tuna in the Indian Ocean. IOTC–2018–WPTT20–42_Rev1.

Prince, J, Dowling, N, Davies C, Campbell, R, Kolody, D. 2011. A simple cost-effective and scale-less empirical approach to harvest strategies. ICES J. Mar. Sci. 68 (5): 947-960.

Punt, AE, Butterworth, DS, de Moor, CL, De Oliveira, JAA and Haddon, M. 2014. Management strategy evaluation: best practices. Fish and Fisheries, 17: 303–334. doi: 10.1111/faf.12104.

TCMP. 2019. Report of the 2nd IOTC Technical Committee on Management Procedures. IOTC–2019–TCMP03–R[E].

WPM. 2020. Report of the 11th Session of the IOTC Working Party on Methods. Microsoft Teams Online, 14 -15 October 2020. IOTC–2020–WPM11–R[E]_Rev1.

WPM(MSE). 2021. Report of the 12th Session of the IOTC Working Party on Methods (Management Strategy Evaluation Task Force). Online, 1-4 March 2021. IOTC–2021–WPM12(MSE)–R[E].

WPTT. 2020. Report of the 22nd Session of the IOTC Working Party on Tropical Tunas, Stock Assessment Meeting. Virtual Meeting, 19-23 October 2020.IOTC–2020–WPTT22(AS)–R[E]_Rev1.

CONTACT US

t 1300 363 400
+61 3 9545 2176
e csiroenquiries@csiro.au
w www.csiro.au

FOR FURTHER INFORMATION

CSIRO Oceans & Atmosphere
Dale Kolody
t +61 3 6232 5121
e dale.kolody@csiro.au