The implementation of a management procedure approach to set catch limits for the Southern Bluefin Tuna and the Atlantic Bluefin Tuna

Melissa Jacobs

A thesis submitted in fulfillment of the requirements for the degree of

MASTER OF SCIENCE

in the Department of Mathematics and Applied Mathematics

University of Cape Town

Supervisor: Prof Doug S. Butterworth

July 2020

DECLARATION

I, Melissa Jacobs, know the meaning of plagiarism and declare that all of the work in the thesis, save for that which is properly acknowledged, is my own; some of the work was conducted in collaboration with others members of the MARAM group, as will be evident from the referencing in the thesis, but in particular the overall context and summarization of what is a large body of work and which the thesis provides is my work alone. I declare that this thesis has not been submitted to this or any other university for a degree. All the sources I have used or quoted have been indicated and acknowledged as references. I hereby grant the University of Cape Town free licence to reproduce for the purpose of research either the whole or any portion of the contents in any manner whatsoever of the above thesis.

Signed by candidate

Melissa Jacobs

July 2020

ACKNOWLEDGEMENTS

This thesis would not have been possible without the support and guidance of a number of people whose contributions I thankfully acknowledge. First and foremost, I would like to thank the Lord Jesus for graciously being my source of strength and guidance throughout my research. I would like to pay special regards to my supervisor, Prof Doug Butterworth, for granting me the opportunity to do this research, and for his patient mentorship and extensive expertise during my studies. This is sincerely appreciated. I wish to thank Ms Di Loureiro (Administrator for the MARAM Group) for her excellent co-ordination of tasks for this group and her general assistance in such a pleasant manner. A special thanks to members of the MARAM group, especially Dr Rebecca Rademeyer, Dr Susan Holloway and Dr Anabela Brandao, for their guidance, support and for proofreading my chapters. I also wish to thank Dr Mitsuyo Miyagawa for her kind assistance, useful insight and ever-willingness to help me, especially with R and ADMB code. I would like to thank the University of Cape Town for awarding a merit bursary to me. Furthermore, I wish to thank all those who were involved in the development of code and software packages used, specifically Dr R. Hillary, Dr A. Parma, Dr D. Webber, Dr J. Ianelli and Dr N., Takahashi for SBT and Dr T. Carruthers for ABFT. I would also like to thank the two anonymous examiners for their time and effort in reading, assessing and commenting upon an earlier version of this thesis, which was very helpful. Lastly, I wish to thank my family and friends for their love, support, encouragement and prayers.

ABSTRACT

The aim of this thesis is to develop, test and evaluate candidate management procedures (CMPs) for the implementation of the Management Procedure (MP) approach to set catch limits for two bluefin tuna species: Southern bluefin tuna (SBT - *Thunnus maccoyii*) and Atlantic bluefin tuna (ABFT - *Thunnus thynnus*). The thesis reviews possibilities and advances proposals for the implementation of this approach in two international tuna Commissions: the Commission for the Conservation of Southern Bluefin Tuna (CCSBT), and the International Commission for the Conservation of Atlantic Tunas (ICCAT) in relation to stocks of bluefin tuna. These approaches aim to provide a simple and widely agreed way to calculate annual catch limits that will lead to meeting fishery management objectives such as attaining or maintaining reasonably large catches in the future, while at the same time avoiding reduction of the resource's abundance to a level which would put the stock and its future productivity at risk.

The SBT component leads to the development, simulation testing and comparative analysis of five final candidate management procedures. The first three, called DMRMCPUE, DMRMGT and DMRMCKMR, each use only CPUE, gene-tagging (GT) and close-kin mark-recaptures (CKMR) indices of abundance respectively. These are followed by DMRcomb1 and DMRcomb2, which are weighted combinations of the first three CMPs. Each CMP is tuned to two different recovery objectives set by the CCSBT: to achieve a median spawning stock biomass (SSB) which is either 30% of its pristine value by 2035, or 35% of this value by 2040. This must be achieved over a weighted set of different Operating Models (OMs) for the resource, which serve as a reference set. Each CMP is applied to the reference set (base18), a CPUE variable squares robustness test (cpuew0) which reflects a resource of worse current status, and a low recruitment robustness test (reclow5) which includes an extended period of poor recruitment in the future. Simulation testing of DMRMCPUE, DMRMGT and DMRMCKMR, together with an equally weighted combination MP DMRcomb1, indicated the need to focus on improved levels for the lower percentile for SSB depletion, and that this can be achieved by placing a higher weight on the GT component in a combined CMP. The DMRcomb2 then does this by placing a 60% weight on the GT component of the MP, and 20% weight on the CPUE- and CKMR-components each. Even though this results in lower stock risk, it also leads to lower catch limits which is a common trade-off.

Subsequent to these analyses, a final MP for SBT was chosen. This CMP was selected on a broader (international) basis where four final CMPs were evaluated and compared, with DMRcomb2 being one of these. These four final CMPs are compared briefly. Although the performance of DMRcomb2 compared favourably with the others, eventually a Management Procedure (MP) called RH13 developed by an Australian scientist was preferred. The primary basis on which RH13 was considered to outperform DMRcomb2 was that it provided a greater probability of catches and SSB continuing to increase after the 2035 recovery target had been attained.

The ABFT component reports on work to an intermediate stage of an MP development process yet to be completed. This process is complex because Atlantic bluefin consist of Western and Eastern origin stocks, which mix in much of the North Atlantic. Results for two CMPs, FXP_1 and FXP_2, are presented. The purpose of these CMPs is to provide direction on where further exploration and discussion needs to focus to improve the trade-off between high catches and resource conservation. These CMPs are applied to a reference set of OMs and robustness tests, OM1 to OM15, from ICCAT's OM package version 5.2.3. The results from this showed that ROM14, a primary robustness test reflecting lower current Western stock abundance, heavily dominates CMP performance because it is difficult to avoid resource extinction under this scenario. Such avoidance leads to low catches and underutilisation for many of the other OMs. Suggestions are made about how to further refine these CMPs and advance the overall process.

Contents

Chapter 1

Introduction

1.1 Introduction

Populations (stocks) of many marine species in the world are overfished, and some are even at risk of extinction. Stocks become overfished when fish are caught at a faster rate than the stocks can replenish themselves by natural regeneration. This emphasizes the necessity for fishery management because stocks that have been managed generally achieve successful rebuilding, while stocks that are not managed are generally declining (Hilborn and Ovando, 2014). Many people rely on fish as a source of protein, and fish accounts for 17% of all the animal protein consumed in the world (Kituyi and Thomson, 2018). Furthermore, the ocean provides an important source of income because the fisheries sector offers approximately 200 million jobs to people worldwide (Kituyi and Thomson, 2018). Therefore, there is an urgent need for widespread sustainable fisheries management.

Today, fisheries management is moving slowly towards a new approach which is called management strategy evaluation (MSE) based harvest strategies, or alternatively the management procedure (MP) approach. This approach will be explored in this paper as a basis to set catch limits for the Southern Bluefin Tuna (SBT) and the Atlantic Bluefin Tuna (ABFT). The purpose of this work is to provide guidance for future analyses for and management of SBT and ABFT, by developing and investigating the properties of some Candidate Management Procedures (CMPs). The basis for this work is simulation testing of these Management Procedures (MPs) with the aim of recommending a sustainable catch, given uncertainties about the status and dynamics of the resource.

The Southern and Atlantic Bluefin Tuna stocks provide a highly priced valuable food source that is much sought after by the lucrative Japanese sushi and sashimi markets. These fish are so valuable to these markets in Japan and some other places around the world, that the effort to catch them is almost unmatched worldwide. This is the main reason they are under heavy fishing pressure.

SBT and ABFT are on the IUCN's[1] redlist of threatened species (IUCN, 2020). Both are listed as endangered, but SBT is categorized as critically endangered. Therefore, both SBT and ABFT are currently in rebuilding phases. Because they occur mainly outside national Exclusive Economic Zones (EEZs), they are managed internationally by the Regional Fishery Management Organizations (RFMOs) for these species, mainly the Commission for the Conservation of Southern Bluefin Tuna (CCSBT) and the International Commission for the Conservation of Atlantic Tunas (ICCAT), respectively.

There is a Glossary in Appendix A that defines terms that although not used in the thesis are useful for context in the overall field. This Glossary is specifically for terms used for harvest strategies, management procedures and management strategy evaluation (Joint Tuna RFMO Management Glossary, 2018). These originate, in the main, from a Glossary developed by the 2018 Joint Tuna RFMO Management Strategy Evaluation Working Group Meeting, with some embellishments and a few additions.

1.1.1 About SBT and the role CCSBT

SBT are large pelagic fish that are found primarily in the Southern Ocean and a known breeding ground off Java, Indonesia in the Indian Ocean. SBT have a maximum length of about 2.5m, can weigh up to 200 kg and have a life span of 40 years. Despite the widespread and migratory nature of SBT, they are considered to be, and are managed as, one breeding stock. Figure 1.1 shows the statistical areas used to stratify information for the fishery for SBT.

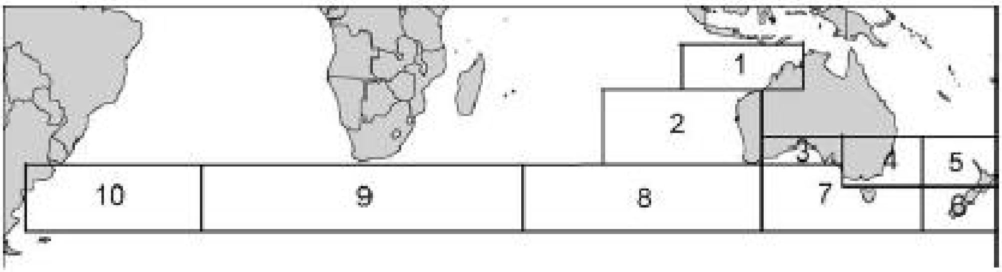

Figure 1.1: The statistical areas used to record information for the fishery for SBT (Campbell, 2004).

The management of SBTs is conducted by the CCSBT, which is an inter-governmental organization that is responsible for sustainable management and conservation of SBT throughout their distribution. The members of the Commission (strictly the Extended Commission when Taiwan is included) are: Australia, the European Union, the Fishing Entity of Taiwan, Indonesia, Japan, the Republic of Korea, New Zealand and South Africa (CCSBT, 2020).

The CCSBT uses the MP approach to address the aim of rebuilding the SBT stock. An MP called the Bali Procedure has been used to provide Total Allowable Catch (TAC) recommendations from 2012

[1]The International Union for Conservation of Nature is a global organization that is primarily focussed on sustainability of natural resources and necessary measures to protect them. The IUCN's red list of threatened species serves as a critical indicator of the health status of different species (IUCN, 2020).

until 2020. The 2017 stock assessment indicated that the stock is at a low level of about 13% of its initial (or pristine) spawning stock biomass (SSB_0) before harvesting commenced, according to the CCSBT's ESC23[2] report (CCSBT, 2018b). Even though this level is low, there has been improvement compared to previous stock assessments, which indicated that stock was at 5.5% SSB0 in 2011 and 9% in 2014 (CCSBT, 2018b).

There is a new MP, called the Cape Town Procedure, that will be implemented from the year 2021. There are rebuilding targets for this MP which will be discussed further in Chapter 4; these are set with the intent of reaching and maintaining an optimal level of SBT abundance in the future.

1.1.2 About ABFT and the role of ICCAT

ABFT is the largest tuna species, reaching a maximum length of about 4m and weighing up to 900 kg. ABFT are found in the western and eastern Atlantic Ocean, as well as its adjacent seas such as the Mediterranean Sea. They have a wide geographical distribution, with spawning typically occurring in the warmer Mediterranean Sea and Gulf of Mexico regions from where the fish move to Atlantic waters to feed. The spawning grounds are indicated in yellow and the migration routes of ABFT are indicated by black lines and arrows in Figure 1.2. This unique extensive migratory nature of ABFT, and the indication of mixing of stocks occurring in different proportions, causes the management of this species to be particularly challenging. This is why they are currently managed as two separate stocks, taken to be separated by the dotted red line in Figure 1.2: the Western stock (originating from spawning in the Gulf of Mexico) and the Eastern stock (originating from spawning in the Mediterranean Sea), though in reality each stock moves across this mid-Atlantic boundary.

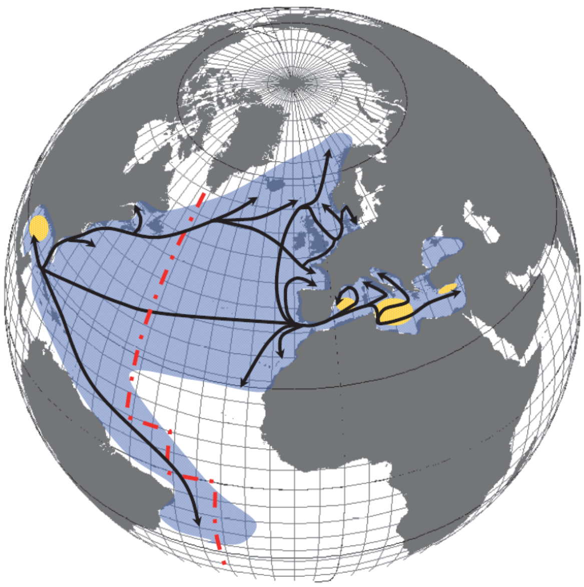

Figure 1.2: Map of the special distribution of ABFT (blue) (Fromentin, 2010 modified from Fromentin and Powers, 2005).

[2]ESC23 refers to the Report of the Extended Scientific Committee for the Twenty Third Meeting of the Scientific Committee (CCSBT, 2018b).

The management of ABFTs is conducted by ICCAT, which is an international organization responsible for the management of tuna-like species in the Atlantic Ocean and its adjacent seas (ICCAT, 2020a). The ABFT is one of many species managed by ICCAT. SBT also occurs in the area for which ICCAT is responsible; however, as noted above the CCSBT is the body primarily responsible for the management of SBT. ICCAT develops management advice for its member nations based on scientific data. Research conducted by ICCAT involves the collection and analysis of fishery statistics relating to trends in the tuna-like resources in the region for which it carries responsibility (Fromentin, 2010).

The implementation of an MP for ABFT currently constitutes work still in progress, with MP development at an intermediate stage. The purpose of the work presented in this thesis is to develop CMPs that will result in reasonable catches for acceptable fishery performance whilst also maintaining a low risk of unintended depletion of either of the Western and the Eastern origin stocks.

Chapter 2

Background and History

2.1 What is the traditional approach to fisheries management

Prior to the development of the MP approach, the "traditional" approach (TA) was widely used to provide scientific recommendations for setting catch limits. The TA is a mathematical evaluation/process which basically uses population monitoring data to estimate resource abundance and productivity, typically called a "stock assessment" of the resource. Essentially, the estimates are obtained from statistical analyses that are dependent on the fishery dynamics and which require extensive population data and expert analysis (Hilborn and Ovando, 2014).

The TA then combines these estimates with a "harvest control rule" (HCR) to calculate a catch limit recommendation, formally referred to as the total allowable catch (TAC). This approach involves the development of a "best assessment" of a resource based upon an associated "best set of assumptions" regarding the factors that impact the resource's dynamics (Butterworth, 2007).

The methodology of the MP approach can be described very briefly as formulating, testing and selecting the most suitable approach overall to set catch limits for a resource, whereas the methodology of the TA would be limited only to formulating and selecting what seems to be the best model to set catch limits. Thus, the TA is like the MP approach but, arguably, the most important aspect is excluded, which is simulation testing to check that the approach will "work" (provide desirable performance[1]) if applied over a period of time.

[1]Performance is used here to refer to an evaluation of the results expected after the implementation of an MP, with the hope of reaching pre-specified objectives.

2.2 What are the problems with the traditional approach

One of the main aims for this area of fisheries management is to obtain a sustainable catch from the resource given current and future uncertainties. Therefore, a method is needed that will be able to account realistically for the behaviour of a harvested resource in terms of aspects which are known, as well as inevitable uncertainties pertaining to resource and fishery dynamics. The TA does not have a process in place to account extensively for uncertainties by some form of testing performance before implementation. Therefore, there are a range of problems that may arise from the use of the TA, which stem from its failure to have used a method of evaluation to first test the performance of the management process to follow.

Simulation testing is the ideal way to consider and evaluate long-term trade-offs realistically. The concept of simulation testing will be explained in Section 2.3, as it is fundamental to conform fully to the MP approach. Since no simulation testing is conducted for the TA, this does not conform fully to the precautionary approach because it does not allow longer-term trade-offs to be evaluated properly, by also accounting for future data that will become available. The only way to check whether the TA is actually appropriate for real use would be to implement it anyway, whilst risking the possibility that the "best assessment" could be wrong.

Regardless of incorporating quantitative data pertaining to uncertainty, there is not a safety-net of assurance that management objectives will be met with the TA (Punt and Donovan, 2007). Two main problems with the TA are related to data: the effect of data variability on TAC recommendations from the "best assessment", and the development of a "best assessment" that might incorrectly depict the resource dynamics.

The selection of a specific TA can have a large effect on the catch recommendation output. Every year new data become available, the TA fits to these extra data, which results in a (yearly) change in output from the "best assessment". This is fine because catch limit advice is reacting to the further information by making use of the signal provided by these new data. However, if the scientists involved in the TA process decide to select a different assessment method for the following year, this change in the methodology will change the catch recommendation output. The problem with this is that the change may not be due to the new data indicating a real change in the resource, but instead could simply reflect the changed selection of the "best assessment" approach. In the long term then, a changed "best assessment" selection may add unnecessary extra variability to the catch limit outputs over time; therefore, it is probably better to adhere to one method instead of (possibly) annually updating it.

Since there is no way to make proper allowance for uncertainties in the TA, the risk from basing decisions on a "best assessment" is always there. In this case the TA might try to take precaution into account by basing a TAC recommendation on, say, the lower 5%-ile value for the TAC estimated (instead of on the median value), which could result in an unnecessarily low catch.

2.3 What is the MP approach (broadly)

The MP approach was first developed by the Scientific Committee of the International Whaling Commission (IWC) during the late 1980s (Butterworth, 2007). Since then, there have been developments and improvements to the approach dependent on the management goals of an organisation. The essence of the MP approach to set catch limits for a resource involves first setting up a candidate management procedure (CMP) which is a formula that, like the TA, inputs population monitoring data with the aim of suggesting a catch limit. This is to provide advice on the total allowable catch (TAC). In this regard, this approach is similar to the TA. However, extensive analysis using simulation testing (using a range of trials) takes place involving the application of the CMP to a range of operating models (OMs). An operating model is a model that describes the dynamics of the resource. A range of different scenarios is developed; these scenarios are represented by different OMs, which need to be considered due to the uncertainty about resource and fishery dynamics. Chapter 3 discusses this more thoroughly.

A CMP is applied to an OM, and in turn the testing software, in this case written using ADMB[2] and R[3], is used. This then outputs performance statistics from simulations (such as average catch, resource depletion and average annual catch variation (AAV[4])) based on the specified OM. This concept is summarized visually in Figure 2.1 below. In addition to analysing the effects of the OMs on CMP performance, this approach takes matters one step further by also examining the effect of possible less likely scenarios which are referred to as robustness tests. CMPs are applied to robustness tests, which also describe the dynamics of the resource but for more extreme or rather unlikely situations, and in turn the performance statistics outputs are evaluated. Taking the outputs from the robustness tests into account serves as a precautionary approach by allowing for the occurrence of situations that could affect the resource negatively, such as an unusual environmental condition which hinders successful reproduction.

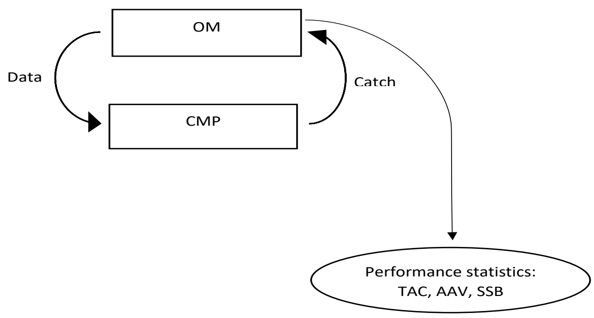

Figure 2.1: An illustration of the use of an OM to evaluate the performance of a CMP.

[2]ADMB is the software AD Model Builder that is used to compile and run CMPs for SBT.

[3]R is the statistical software used to compile and run CMPs used for ABFT with the use of a package called ABTMSE.R (Carruthers, 2019b). R is also used to produce graphics for both SBT and ABFT.

[4]$AAV = \frac{1}{y_2-y_1+1}\sum_{y=y_1}^{y_2}\frac{|C_y-C_{y-1}|}{C_{y-1}}$, where C_y is the catch (or projected TAC) for year y (Carruthers, 2019c).

2.4 The advantages of a MP approach

The use of an MP approach has a high possibility to reduce the risk of overfishing, the extent of depletion of a resource, and a fishery "crashing", because of its robustness to uncertainty due to the simulation testing that takes place prior to making a TAC recommendation. Since there will be multiple OMs and robustness tests within any MP approach, a catch recommendation approach will be tested for each one in turn to ensure that the stock does not drop beyond repair regardless of the details of the resource and fishery dynamics. The consideration of uncertainty allows for a catch formula to be chosen that will still provide reasonable performance, regardless of how the resource behaves. This is advantageous to everyone, and gives industry a greater sense of security.

The MP has a structured nature, which eliminates the possibility to argue back and forth about considering changes to the TA and its catch limit output each year. It doesn't allow for *ad hoc* decisions because all the "rules" concerning catch and making a TAC recommendation are not only discussed but also decided on beforehand. This is beneficial because it reduces the time spent haggling over catch recommendations (Butterworth, 2007).

Not only is the MP approach structured, but it incorporates the concept of acting according to the status of the stock, i.e., if the status of the stock improves then more catch may be taken, but if the status of the stock worsens then less catch should be taken than previously recommended. The fishing industry also needs stability, which can be ensured only by placing limitations on the rate at which TAC's may be increased and reduced. But the only way one can ensure that this will not lead to undue harm to the resource is by using simulation testing. The primary advantage of the MP approach is that catch limits are forced to be based primarily on stock status, and not on the motives of the different parties involved in the catch limit decision process.

2.5 How to apply the MP approach in practice

Applying an MP in practice requires interdisciplinary input. Scientific researchers, stakeholders, computer specialists and technicians, government compliance officers, industry executives, and fishers are all involved in the process of implementing an MP.

Scientists formulate models to use and procedures to follow for setting TACs. The best practice manner in which MPs should be developed is laid out in an overview article by Punt *et al.* (2016). This process requires computer simulations in which computer specialists are involved to use computer code to represent the MP process and develop the simulation tests. This process to first develop the MP approach to set catch limits is evaluated in this thesis from a research perspective. It is the responsibility of scientists and researchers to present different MP options or harvest strategies resultant from this process, as well as the implications of each MP, to stakeholders, managers and government compliance officers. In turn, it is the government compliance officers and managers responsibility to abide by the rules stipulated by the MP approach for the MP which they chose.

If an MP approach leads to setting a catch limit of, for example, 70 000 mt of fish that may be caught, it will be the responsibility of fishers to abide by and not exceed the limit. Additionally, it is also the responsibility of government compliance officers to ensure that industry laws and government regulations are respected.

Meetings and conferences are scheduled as part of the process. There will be meetings with the appropriate participants to discuss a specific subject within the process leading to making a catch limit recommendation. Scientists, researchers and computer specialists will need meetings to discuss and agree on a reference set of operating models and different MP options to present to the other parties. Following that, there are meetings where different candidate MPs will be presented for managers to choose the most suitable of these candidates.

Each organisation will apply an MP appropriate for its objectives, goals and situation dependent on the species being fished. The CCSBT and ICCAT are inter-governmental organisations; therefore, there are delegates from each member country that participate in the process.

CCSBT set up a work plan each year to follow a process with meetings that, in summary, aim to develop, discuss and decide on an MP approach to stipulate an agreed TAC. These involve Operating Model and Management Procedure Technical Meetings, Extended Scientific Committee Meetings and Extended Commission Meetings, amongst others. Table 5 in Appendix B2 shows the CCSBT work plan from 2018 to 2020, and provides a description of what each meeting entails to be able to achieve an agreed TAC to be implemented for 2021 to 2023. Once the MP is adopted, a schedule for implementation is initiated.

ICCAT is responsible for broad range of species, so that there are different meetings taking place during the year. To mention some, there are meetings of the Standing Committee on Research and

Statistics (SCRS) that is responsible for updating statistics concerning fishing activities and making final scientific recommendations about catch limits. There are also SCRS species group meetings that are responsible for MP work and stock assessment updates, and meetings of the compliance committee which is generally responsible for ensuring that ICCAT management measures are in line with compliance regulations (ICCAT, 2020a).

Chapter 3

Methods

3.1 Introduction

The purpose of this research is to select the best candidate management procedure to implement for each of the two bluefin tuna species investigated, Southern Bluefin Tuna (*Thunnus maccoyii*) and Atlantic Bluefin Tuna (*Thunnus thynnus*). Extensive analyses are conducted to investigate the properties of the suggested CMPs developed, with the aim of determining which CMP provides the best basis for future scientific recommendations for TACs.

The process of formulating, testing and selecting the most suitable MP approaches to set catch limits for these bluefin tuna species will be discussed here; it follows the general approach set out in Section 2.3. The results of the MP tests are produced computationally, using and appropriately modifying code developed by R. Hillary, A. Parma and D. Webber for SBT and T. Carruthers for ABFT. The selection process depends on the performance statistics chosen to evaluate the MPs on the basis of a chosen set of OMs; these models reflect alternative hypotheses for the dynamics of the resource and the fishery concerned.

3.2 Southern Bluefin Tuna

3.2.1 Background

In 2019 the CCSBT's Extended Scientific Committee (ESC) recommended a new MP, called the Cape Town Procedure, which is intended to be used to set the global TAC for the management of SBT across all waters from the east coast of South America across the South Atlantic and Indian Oceans to the east of New Zealand in the Pacific from the year 2021. Following this, the Extended Commission (EC) decided to adopt this MP (CCSBT, 2019b). Prior to this, an MP called the Bali Procedure had been developed and will continue to be implemented until the year 2020. The development of CMPs for SBT occured over a period of nine years, from 2002 to 2011, before the Bali Procedure was finally adopted in 2011 and implemented in 2012.

The implementation of SBT MPs to provide catch limit advice applies to periods of three years, which are referred to as TAC blocks, where the second year in each block is used to determine the TAC to be implemented for the next block starting after the following two years. Figure 3.1 shows visually how the process has operated since the Bali MP was adopted in 2011, as well as the times when TAC calculations occurred within the TAC blocks. TAC Block 4 represents the first implementation period after the new MP, the Cape Town Procedure, was adopted in 2019.

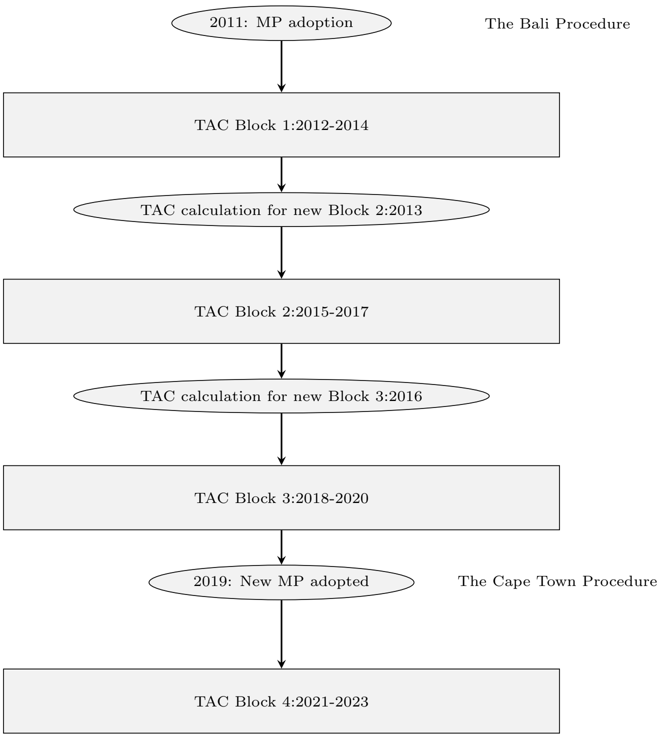

Figure 3.1: The SBT MP timeline from 2011 to 2023.

CPUE and Aerial Survey (AS) indices are the two main data inputs to the Bali Procedure. A key difference between the Cape Town Procedure and the Bali Procedure is that the former uses gene-tagging (GT) estimates of the abundance of 2-year-olds and information on spawning stock estimates from close-kin mark-recapture (CKMR) in place of the AS indice that the latter had used (both MPs use CPUE indices from Japanese longliners). One of the two main reasons for the development of the new MP is that collection of one of the primary data inputs, the AS indice, had to be discontinued (for various reasons, as set out below).

The conduct of an AS is quite expensive, and for financial reasons there were no AS data collected in the year 2015. However, the absence of AS data for that year had essentially very little to no effect on the performance of the implementation of the Bali Procedure in terms of achieving management objectives for the sustainability of the stock (Takahashi *et al.*, 2015a).

Before an MP is implemented, the newly observed values of the indices for each data type are compared to predictions made at the time the MP was adopted. In 2016, the most recent observed CPUE index value lay within the predicted 95% probability range when the Bali Procedure was applied to the reference (base) case OM. However, the AS index values for the years 2012, 2014 and 2016 did not lie within the corresponding predicted probability range. In addition to this, the AS index value for the year 2016 was far higher than the upper 5th percentile, even when the Bali Procedure was applied to the "high aerial cv" robustness test. This robustness test assumes a larger variability of AS indices (Takahashi *et al.*, 2016).

Regardless of the absence of the 2015 AS data and the high result for the 2016 AS index, the decision was still made to use the Bali Procedure to make TAC recommendations for the 2018 to 2020 fishing seasons because the concerns mentioned above did not lead to increased stock conservation risk. The reason for this is because the 2016 AS index value was higher (rather than lower) than the range predicted for the value of that index. The reason the AS was later discontinued was because experienced spotter pilots had retired and the data collection was too expensive to continue.

3.2.2 The data available for CMPs and the operating model

3.2.2.1 Data available for the CMPs

The CCSBT ESC agreed in 2017 to the development of a revised MP that would use two other (additional) sources of data in the place of the aerial survey data which would no longer be available. These are the GT data from a gene-tagging monitoring programme and close-kin mark recapture (CKMR) data from a CKMR project (CCSBT 2017b). Thus, the three data inputs that are used in the development and evaluation of candidate MPs are catch per unit effort (CPUE), gene-tagging (GT) and close-kin (CK) data. These three data types/inputs used are explained briefly below.

Catch per unit effort (CPUE):

The CPUE index data that are projected for use have similar behaviour and statistical properties to those shown by the standardized Japanese longline CPUE used in the conditioning process (CCSBT, 2010). They are generated from the projected future abundance (numbers-at-age) of the resource.

The CPUE index that is used as an input to the CMPs developed below is defined as follows, where J is the index:

$$J_y^{CPUE} = \frac{\frac{1}{5} \cdot (CPUE_{y-2} + CPUE_{y-3} + CPUE_{y-4} + CPUE_{y-5} + CPUE_{y-6})}{\frac{1}{5} \cdot (CPUE_{2016} + CPUE_{2015} + CPUE_{2014} + CPUE_{2013} + CPUE_{2012)}}; \tag{3.1}$$

where J_y^{CPUE} is a relative CPUE index

The reason for averaging over five years is to reduce the impact of random errors on the resource abundance signal provided by these data. The index is normalised to the average of the values over the most recent five years at the time these analyses commenced; all indices are normalised in a similar fashion to make their values and changes comparable in terms of their size.

Gene-tagging (GT):

Gene-tagging is essentially genetic fingerprinting. A juvenile fish is caught, and a DNA sample or genetic fingerprint is collected from this juvenile fish which is then used as a tag in order to monitor that specific fish with its unique genetic signature. Gene-tagging is a data collection method that is simple and environmentally friendly because genetic samples are taken from a number of fish caught, following which they are released back into the water with no damage or harmful marking to the fish. After a set period, the fish are recaptured to determine which fish have DNA corresponding to those originally genetically sampled (or "marked"), and hence what proportion of the fish that were marked have been caught. This information then provides estimates of absolute abundance and fishing mortality relatively straightforwardly, and can also be used to provide estimates of natural mortality for the tagged cohort(s) (Cramer Fish Sciences, 2015).

The incorporation of juvenile abundance estimates into the SBT CMPs can be a valuable source of information for testing those CMPs. The CCSBT have previously used aerial survey indices as the primary source of information for estimates of juvenile abundance. However, under the CCBST Scientific Research Program for 2014-2018, a gene-tagging programme was developed which provides annual estimates of the number of 2-year-old tuna entering the population (Preece *et al.*, 2015).

The GT index is an estimate of the number of two year old tuna that is used as an input to the CMPs developed below is defined as follows, where J is the index:

$$J_y^{GT} = \frac{\frac{1}{5} \cdot (J_{y-2}^{GT} + J_{y-3}^{GT} + J_{y-4}^{GT} + J_{y-5}^{GT} + J_{y-6}^{GT})}{J_{2016}^{GT}}; \tag{3.2}$$

where J_y^{GT} is a relative GT index averaged over five years, with the average taken over 5 years for the same reason as for the CPUE index. Note that only the 2016 value is used to normalise the index, as that was the first year the programme started and the only result available at the time these analyses commenced.

Close-kin mark recapture (CKMR):

The CKMR index of adult abundance used in CMPs is another index that will be available each year in the future. CKMR is a method to monitor fish populations that is similar in concept to GT because it also relies on a 'mark and recapture' method. However, this method relies specifically on matching adults to offspring, where adults are sampled off Indonesia where they spawn, and the offspring are sampled in the Great Australian Bight when they are only a few years of age at most. A parent-offspring pair (POP) refers to two fish that have matching DNA; a half-sibling pair (HSP) refers to two offspring which share one but not both parents (Bravington, 2017). This information is used to enable the SBT OM (that will be discussed in Section 3.2.2.4) to generate parent-offspring and half-sibling pairs, but the index that is used as an input to CMPs is not these numbers of parent-offspring and half-sibling pairs, but instead (for simplicity) an index of adult abundance for each year using a method developed by Hillary (Hillary et al., 2018).

The CKMR index that is used as an input to the CMPs developed below is defined as follows, where S_y is the annual index:

$$J_y^{CKMR} = \frac{\frac{1}{2} \cdot (S_{y-5} + S_{y-6})}{\frac{1}{2} \cdot (S_{2013} + S_{2012})}; \tag{3.3}$$

where J_y^{CKMR} is a relative CKMR index averaged over two years and S_y is the estimated SSB value obtained from analyzing the CKMR-related data using code provided by Hillary (Hillary et al., 2018). S_{2012} and S_{2013} depend on the year y for which J_y^{CKMR} is evaluated and not "directly observed" indices such as the CPUE (which is essentially catch divided by effort) and GT (which is essentially numbers of genetically tagged fish multiplied by numbers caught and divided by number of such tagged fish in the catch) indices above (Butterworth *et al.*, 2018b).

3.2.2.2 Introduction to the OM

This section describes the 2017 CCSBT OM which was reconditioned in 2017 before the MP testing started in 2018. The purpose is to use an OM that has been satisfactorily fit (or conditioned) to past data to be able to then provide future predicted values for data. The structural changes to the OM, after the reconditioning (i.e. an update to earlier conditioning) had taken place, were reviewed and accepted by the 8th Operating Model Management Procedure (OMMP) meeting in June 2017 (CCSBT, 2017a).

This research covers work that began with the testing of an MP in 2018, which was carried out based on the 2017 reconditioned OM and selected robustness tests. Thereafter ended at the ESC24[1] in September 2019 (CCSBT, 2019b), soon after which the EC[2] selected and adopted a new MP, namely, the Cape Town Procedure (CCSBT, 2019a). Subsequent to the start of this work, a few further generally minor changes were made to the OM, such as the addition of unaccounted mortality (UAM1) and the update is listed and described in Appendix B3.

Prior to 2017, the last time that the CCSBT OM had been reconditioned was in 2014. The OM had to be reconditioned anew, there would be two new data inputs available, namely GT and CKMR. In addition to the new data sources to be taken into account, there were also data types used previously that needed to be updated. The code used to simulate SBT projections then had to be adjusted to account for these changes to the OM.

3.2.2.3 The population model

The parameters of the CCSBT OM relate to a population model. Fish populations can be depicted by a model in terms of various characteristics such as age, area, size, sex, stock, natural mortality and somatic growth. There are various population models that can be used to reflect the dynamics of a fish population, which are dependent on the species itself and/or its behaviour (Hilborn and Walters, 1992). The basic population model used for SBT is age-structured with a Beverton-Holt stock recruitment relationship.
The OM is described by equations which are shown in detail in Appendix B4. There are some parameter values in the OM that are assumed and fixed on input, but there are also other parameter values that are estimated when conditioning the OM to historical catch and other resource monitoring data. The process for estimating the latter set of parameter values for all the OMs in the reference set (or "grid") in Table 3.1 is based on the principle of maximum likelihood. The estimation of the parameter values depends on the goodness of fit of the historical data to the population model. The fitting is conducted by maximising the likelihood using a software package such as ADMB (Note that this likelihood is referred to as an objective function - whose value is minimized in the fit - because some penalty terms are added to the negative log-likelihood).

[1]ESC24 refers to the Report of the Extended Scientific Committee for the Twenty Fourth Meeting of the Scientific Committee (CCSBT, 2019b).

[2]EC refers to the CCSBT's Extended Committee whose tasks include deciding upon a total allowable catch and its allocation among the members of the CCSBT.

The historical data include catch, CPUE indices, catch length distribution data, age-at-length data and data on tag returns. The historical catch data represents the total catch estimated each year from the various countries that have participated in the fishery, and the CPUE series used are for the Japanese longline fishery. The data used in the OMs also include results from the aerial survey, and more recently from gene-tagging and close-kin genetics.

3.2.2.4 The reconditioned OM and parameter description

Table 3.1 provides the grid configuration for the agreed reference set of OMs for the 2017 conditioning process, which is a slightly simplified version of Table B.2 in Appendix B3 which is an update for the conditioned OM for 2019. This Table reports on the models included in the OM grid, where each row represents an OM parameter that is explained below. The first column lists these parameters, the second column indicates the corresponding values chosen for each parameter, and the third column "cumulative N" indicates the total number of models there are in the grid as each parameter is added, since the number of models is a combination of all possible choices listed (separated by commas) in the parameter values column.

These combinations lead to an eventual total of 432 different models. The fourth column represents the weighting given to each parameter value, where M_0 and M_{10} have a likelihood-based weighting (related to the quality of the fit of the OM to the data) and the remainder of the weightings were uniform or based on the expert judgement of the CCSBT ESC members. However, each of these 432 different models are not all are given equal weighting. Some are weighted more than others, depending on the weight given to each of the parameter values to which that OM corresponds; this is done by multiplying together the respective weights for each of these parameters. Therefore, as the (control parameters of the) CMPs are tuned to the median target value for recovery, this is not to the 432 models weighted equally, but instead to this individual weighting of each of the 432 models. The results for the control parameter values that are tuned to achieve a median SSB recovery target for the reconditioned OMs as shown in Table 4.1 on page 45, and described in Chapter 4.

In addition to the different models in the OM grid, there are more scenarios that can be used to test the performance of an MP which are called robustness tests or sensitivity tests. Robustness tests generally relate to (OM) scenarios which are less likely to occur, and they are used to test the robustness or flexibility of an MP in the event of such extreme or drastic eventualities. All the robustness tests considered are listed in Appendix B1 with their descriptions as well their priorities in an MP testing context.

Parameter descriptions:

h — The parameter h is steepness, which is related to the shape of the stock-recruitment function. A key assumption of the population model is that there are circumstances where the number of recruits born each year is the same as the number of tuna that die each year, i.e. the stock's abundance will be kept steady - this is when it is at carrying capacity (K). There is a corresponding recruitment in

Table 3.1: Grid configuration for the agreed reference set of OMs for the 2017 stock assessment.

Parameter	Value	Cumulative N	Weighting
h	$0.6, 0.7, 0.8$	3	0.33 : 0.33 : 0.33
M_0	$0.35, 0.4, 0.45, 0.5$	12	Likelihood-based
M_{10}	$0.05, 0.085, 0.12$	36	Likelihood-based
Omega (ω)	1	36	1
CPUE series	w0.5, w0.8	72	0.5 : 0.5
CPUE age range	$4-18, 8-12$	144	2 : 1
Psi (ψ)	$1.5, 1.75, 2$	432	$1, 2, 1$

this situation, $R(K_{SB})$, where K_{SB} is the carrying capacity expressed in terms of spawning biomass. When the spawning biomass is $0.2K_{SB}$, the corresponding recruitment is $hR(0.2K_{SP})$, which then defines the steepness parameter.

M The parameter M is the natural mortality rate of the fish. The mortality of fish (for each age a) in a population is assumed to be proportional to the number of those fish in the population. Natural mortality is mainly the result of fish being eaten by natural predators. The natural mortality rate at age a in a closed population is defined (in the absence of any mortality caused by fishing) by:

$$M_a = \frac{dN_a}{dt}/N_a; \tag{3.4}$$

where N_a is the number of fish of age a.

M_0 The parameter M_0 is the natural mortality rate for small 0 year old fish.

M_{10} The parameter M_{10} is the natural mortality rate for larger 10 years and older fish.

Omega(ω) This is the power parameter in the CPUE-exploitable biomass relationship: $CPUE = kB^{\omega}$.

CPUE series There are 10 statistical areas covered by the SBT fishery; these are the 10 rectangular blocks shown in Figure 3.2. The Japanese longline CPUE series considered relates to areas 5 to 9 only. These areas cover a wide expanse from South Africa to New Zealand, and are fished over the Southern Hemisphere winter months. GLM standardization is carried out to account for factors other than fish density on which CPUE depends, such as latitude and longitude, and essentially accounts for the non-random nature of the collection of the data. For the conditioning process, the population model is fit to the standardized CPUE.

The CPUE abundance series is developed by each year multiplying the CPUE in each area (which is an index of the fish density there) by the size of that area. Either "Variable squares" or "Constant squares" series are created, depending on how these area-sizes are defined. "Variable squares" assumes that the density of fish in a small region that wasn't fished in a particular year is zero, whereas "Constant squares" is more optimistic in assuming the densities in those unfished regions are the same as the density in nearby regions where fishing did take place that year; all regions that have been fished at least once during the history of the fishery are included in the summation concerned. Two different weighting of these series are then considered, w0.5 and w0.8, where w0.5 takes the average of Constant squares and Variable squares series (a 50:50 weight) and w0.8 puts more weight on Constant squares because it is considered to be closer to the real situation, and a weighting ratio of 80:20 is used (CCSBT, 2009).

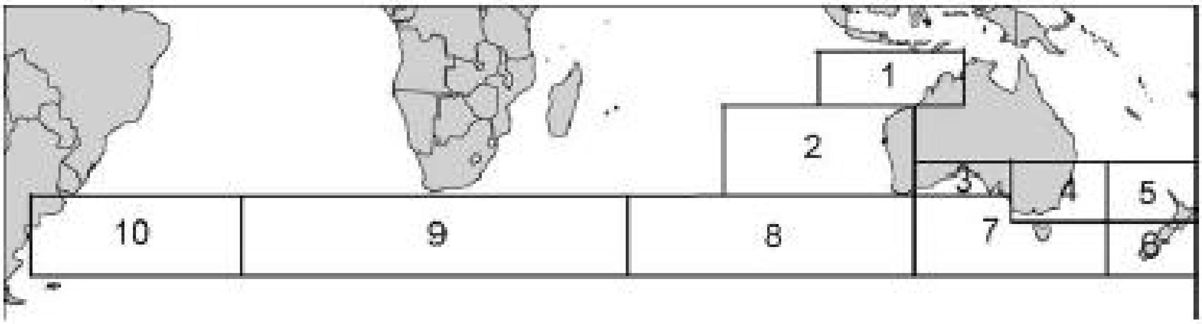

Figure 3.2: The statistical areas used to record information for the fishery for SBT (Campbell, 2004).

CPUE age range — The Japanese longline CPUE series used relates to fishing of tuna of intermediate ages, which range from approximately age 4 to age 18 years. Since CPUE is assumed to be proportional to the biomass available to the fishery, it is necessary to consider how to weight the ages in this range to the fishery to best reflect the contributions from this range. There are two options given in Table 3.1, which are either to average over the whole range (age 4 to age 18) or to concentrate more towards the middle of that range (age 8 to age 12).

Psi (ψ) — This parameter relates to the CK data. Once fish are mature, they may spawn from not at all to multiple times each year, with the frequency increasing with age. The parameter ψ is the parameter of the associated non-linear relationship which is modelled by a power law. There are three alternative values for ψ that are considered.

3.2.3 The SBT CMPs

3.2.3.1 The development of a CMP

This section will focus on the development of five "final" CMPs. The results of the five CMP's will be discussed extensively in Chapter 4. These CMPs have been based on a series of CMPs that all initiated from the CMPs DMM2 to DMM6 which are described below.

An interim objective set by the CCSBT was to avoid the resource from failing to achieve a certain level of recovery by a specified time. The specific aim was to attain at least a 70% probability of achieving an interim target recovery of 20% of pristine spawning stock biomass by the year 2035 (CCSBT, 2018a). The CMPs tested were all tuned (by varying the values of their control parameters) to ensure that this objective, as well as other more specific tuning criteria, were met.

All the CMPs had been tested to meet this requirement, before proceeding (in June 2018) with the agreed objectives stipulated by OMMP9[3] for the initial and final rounds of CMP development and testing. These were to attain a 50% probability of achieving various recovery objectives of either 25%, 30%, 35% or 40% of unfished spawning stock biomass SSB_0 (or initial total reproductive output(TRO_0)[4]) by the tuning year 2035, and further post-2035 performance should also be evaluated by the year 2045 (CCSBT, 2018a).

There are three types of data input used for the CMPs, as discussed in section 3.2.2.1. The final objective is to see how best the three data types could be used in combination to provide a catch limit (TAC) recommendation. Therefore, CMPs were initially developed for each data type. Thereafter, combinations of either two or three data types were used/included to evaluate the results, and the implications for the TAC and other aspects of MP.

3.2.3.2 Initial round of CMP development

A simple CMP, **DMM1**, that was the first CMP considered in the development process (Butterworth *et al.* 2018a):

This CMP set the TAC very simply as a multiple of the J_y^{CPUE} value at the time, where J is an aggregate index of abundance. This is a simple constant proportion CMP. It intends a constant fishing mortality strategy, with the TAC being moved up or down proportional to the abundance which is indexed by J_y^{CPUE} (an index based on CPUE data):

$$TAC_{y+1} = \alpha \times J_y^{CPUE} \times \frac{TAC_{2016}}{J_{2016}^{*CPUE}} \tag{3.5}$$

where:

TAC_{y+1} is the TAC that is determined for the following year,

[3]The Ninth Operating Model and Management Procedure Technical Meeting (CCSBT, 2017b)

[4]Strictly TRO was always used, but for convenience and understandability it is referenced as SSB.

J_y^{CPUE} is the relative CPUE index that is observed (actual data), and is then averaged over a period of the last five years; this is to lessen the unwanted impact of noise in these data which is unrelated to the underlying abundance, this index is shown in equation 3.1,

TAC_{2016} which is the actual TAC for the year 2016 which was 14647 mt,

$J_{hist,x}^{CPUE}$ is the historical CPUE index for year x,

J_{2016}^{*CPUE} averages the historical CPUE indices over three years as follows: $\frac{1}{3} \cdot (J_{hist,2014}^{CPUE} + J_{hist,2013}^{CPUE} + J_{hist,2012}^{CPUE})$

α is a control paramter that is tuned to achieve a desirable outcome when the CMP is applied to the OM.

DMM2 is a target-based empirical CMP that makes use of only CPUE index data and incorporates a target parameter called J_{targ}^{CPUE}. It is based on the following formula, where β and J_{targ}^{CPUE} are tuning parameters (Butterworth et al. 2018a):

$$TAC_{y+1} = TAC_y \times (1 + \beta \cdot (J_y^{CPUE} - J_{targ}^{CPUE})) \tag{3.6}$$

where:

β is a control parameter (which is related to the CPUE index data) that is tuned to achieve a desirable outcome when the CMP is applied to the OM, and

J_{targ}^{CPUE} is the parameter whose value is adjusted to achieve the recovery SSB level set for a target year (usually 2035).

DMM1 is a constant proportion CMP, whilst DMM2 is a target-based CMP. At this stage of development, the target approach was preferred and all the CMPs following DMM1 and DMM2 are either similiar to DMM2 though with different data input(s) or an improvement to DMM2 with respect to the TAC formula itself. The reason for choosing the target approach over the constant proportion approach is due to the output performance statistics for DMM2 being judged to be better than those for DMM1.

DMM3 is a target-based empirical CMP, and is an extension to DMM2 that makes use of CPUE index data as well as GT data, and their respective target values (Butterworth *et al.* 2018a).

$$TAC_{y+1} = TAC_y \times (1 + \beta(J_y^{CPUE} - J_{targ}^{CPUE}) + \gamma(J_y^{GT} - J_{targ}^{GT})) \tag{3.7}$$

where:

γ is a control parameter (which is related to the GT data) that is tuned to achieve a desirable outcome when the CMP is applied to the OM, and

J_{targ}^{GT} is the parameter whose value is adjusted to achieve the recovery SSB level set for a target year (usually 2035).

DMM4 is a modification to DMM3 that incorporates the tuning parameters in an improved (more readily understood) way as follows (Butterworth *et al.* 2018b):

$$TAC^1_{y+1} = TAC^1_y \times (1 + \beta(J^{CPUE}_y - J^{CPUE}_{targ})) \tag{3.8}$$

$$TAC^2_{y+1} = TAC^2_y \times (1 + \gamma(J^{GT}_y - J^{GT}_{targ})) \tag{3.9}$$

$$TAC_{y+1} = \omega \cdot TAC^1_{y+1} + (1 - \omega) \cdot TAC^2_{y+1} \tag{3.10}$$

Here there is the addition of an extra paramater ω, where ω provides a relative weight to the CPUE index data compared to the GT data, such that $\omega + (1 - \omega) = 1$. The approach used to select parameter values for DMM4 was first to tune γ and J^{GT}_{targ} when setting $\omega = 0$ to provide a TAC if GT data only were used. Thereafter, the next step in the process was to test the (remaining) combinations of β, J^{CPUE}_{targ} and γ and to select the parameter values based primarily on obtaining less TAC variability during the projection period.

DMM5 is a CMP that makes use of only the CKMR index[5] generated from the SSB estimation method developed by Hillary (Hillary *et al*, 2018):

$$TAC_{y+1} = TAC_y \times (1 + \kappa(J^{CKMR}_y - J^{CKMR}_{targ})) \tag{3.11}$$

where J^{CKMR}_{targ} and the control parameters are defined below:

κ is a control parameter (which is related to the CKMR index data) that is tuned to achieve a desirable outcome when the CMP is applied to the OM, and

$$J^{CKMR}_{targ} = \begin{cases} \frac{T2-T1}{y2-y1} \cdot (y - y_1) + T1, & y1 \leq y \leq y2 \\ T2, & y2 < y \end{cases}$$

with $T1 = 0.5, T2 = 2.0, y1 = 2021, y2 = 2035$ as these led to reasonable performance.

where:

J^{CKMR}_y is a relative CKMR index averaged over two years.

S_y is the estimated SSB value obtained from analyzing the CKMR-related data using the code provided by R. Hillary, and

S_{2012} and S_{2013} depend on the year y which J^{CKMR}_y is evaluated and not "directly observed" indices (Butterworth *et al.*, 2018b).

There is a dependence on year for the J^{CKMR}_{targ} value because S_{2012} and S_{2013} depend on the year y for which J^{CKMR}_y is evaluated and not on the "directly observed" indices which are detailed in section 3.2.2.1; thus when simulation tests were conducted, the SSB reflected a decrease in later projection

[5] The (combined) CKMR index generated from the SSB estimation method referred to here is used in the code provided by R. Hillary for the purpose of SBT CMP development (Butterworth *et al*, 2018b).

years after reaching the SSB tuning requirement for 2035. The solution to this was to first have a smaller J_{targ}^{CKMR} value, achieved by setting $T2 > T1$, which allows the TAC to increase initially, and thereafter a bigger J_{targ}^{CKMR} value to prevent the TAC from continuously increasing and resulting in a decline of SSB after the target has been reached at 2035.

DMM6 is a target-based empirical CMP that makes use of all three data inputs, i.e. the combination of CPUE index data, GT data and CKMR index data (Butterworth *et al.* 2018b):

$$TAC_{y+1}^{1} = TAC_{y}^{1} \times (1 + \beta(J_{y}^{CPUE} - J_{targ}^{CPUE}))....... \quad (3.8)$$

$$TAC_{y+1}^{2} = TAC_{y}^{2} \times (1 + \gamma(J_{y}^{GT} - J_{targ}^{GT})).............. \quad (3.9)$$

$$TAC_{y+1}^{3} = TAC_{y}^{3} \times (1 + \kappa(J_{y}^{CKMR} - J_{targ}^{CKMR}))....... \quad (3.11)$$

$$TAC_{y+1} = \nu \cdot TAC_{y+1}^{3} + (1-\nu)[\omega \cdot TAC_{y+1}^{1} + (1-\omega) \cdot TAC_{y+1}^{2}] \quad (3.12)$$

This CMP combines the CMPs DMM4 and DMM5 with the addition of an extra paramater ν, where ν provides a relative weight for the CKMR index data compared to the CPUE and GT data, such that $\nu + (1-\nu) = 1$.

3.2.3.3 Final round of CMP development

DMRCPUE is a revised version of DMM2, equation 3.6, which makes use of CPUE index data only, and allows for the differential adjustment of the TAC based on whether the observed CPUE index is above or below a specific target level J_{targ}^{CPUE}. The reason for this is to get relatively smaller increases when the observed CPUE index is above the target and bigger decreases when the observed CPUE index is below the target (Butterworth *et al*, 2019d, e).

This can be achieved by adjusting control parameter β, such that:

$$\beta = \begin{cases} \beta_{up} = \beta_1 - \beta_2 \\ \beta_{down} = \beta_1 + \beta_2 \end{cases}$$

where: $\beta_1 > \beta_2$

$$\text{If } J_{y}^{CPUE} \geq J_{targ}^{CPUE} : \ TAC_{y+1}^{CPUE} = TAC_{y}^{CPUE} \times (1 + \beta_{up} \cdot (J_{y}^{CPUE} - J_{targ}^{CPUE})) \quad (3.13)$$

$$\text{If } J_{y}^{CPUE} < J_{targ}^{CPUE} : \ TAC_{y+1}^{CPUE} = TAC_{y}^{CPUE} \times (1 + \beta_{down} \cdot (J_{y}^{CPUE} - J_{targ}^{CPUE})) \quad (3.14)$$

Constraints on DMRCPUE:

1) If $TAC_{y+1}^{CPUE} > 28000$, then $TAC_{y+1}^{CPUE} = 28000$

2) For y >y_t : $J^{CPUE}_{targ,y}$ is replaced by $J^{CPUE}_{targ}(1 + \lambda^{CPUE}(y - y_t))$

where λ^{CPUE} is a parameter that stipulates the per annum (p.a.)
increase in J^{CPUE}_{targ} after tuning year y_t which is either
2035 or 2040.

The first constraint was a bound of 28000 mt. Without this bound, TACs could increase too much and consequently cause later problems with resource status. The second was intended to reduce the increase in TAC after year y_t, so as to try to avoid a subsequent drop in abundance.

DMRGT is a revised version of equation (3.9); it is a CMP that uses GT data only, based on the formulae below. The rationale for the form of these various formulae and constraints are the same as for the two constraints for CPUE (Butterworth *et al*, 2019d, e).

The control parameter γ is adjusted similarly as β above:

$$\gamma = \begin{cases} \gamma_{up} = \gamma_1 - \gamma_2 \\ \gamma_{down} = \gamma_1 + \gamma_2 \end{cases}$$

where: $\gamma_1 > \gamma_2$

$$\text{If } J^{GT}_y \geq J^{GT}_{targ}: \ TAC^{GT}_{y+1} = TAC^{GT}_y \times (1 + \gamma_{up} \cdot (J^{GT}_y - J^{GT}_{targ})) \tag{3.15}$$

$$\text{If } J^{GT}_y < J^{GT}_{targ}: \ TAC^{GT}_{y+1} = TAC^{GT}_y \times (1 + \gamma_{down} \cdot (J^{GT}_y - J^{GT}_{targ})) \tag{3.16}$$

Constraints on DMRGT:

1) If If $TAC^{GT}_{y+1} > 28000$, then $TAC^{GT}_{y+1} = 28000$

2) For y >y_t : $J^{GT}_{targ,y}$ is replaced by $J^{GT}_{targ}(1 + \lambda^{GT}(y - y_t)))$

where $\lambda^{GT}$is a parameter that stipulates the p.a.
increase in J^{GT}_{targ} after tuning year y_t which is either
2035 or 2040.

The rationale for these constraints is similar to that stated above for DMRCPUE.

DMRMCKMR is a revised version of equation (3.11), the CMP that uses CKMR summary data only, based on the formulae below. The rationale for the form of these various formulae is the same as for the two constraints for CPUE and GT above, but the rationale for the constraints is more complicated as described below (Butterworth *et al*, 2019d, e).

The control parameter κ is adjusted similarly as β and γ above:

$$\kappa = \begin{cases} \kappa_{up} = \kappa_1 - \kappa_2 \\ \kappa_{down} = \kappa_1 + \kappa_2 \end{cases}$$

$$\text{where: } \kappa_1 > \kappa_2$$

$$\text{If } J_y^{CKMR} \geq J_{targ}^{CKMR}: \; TAC_{y+1}^{CKMR} = TAC_y^{CKMR} \times (1 + \kappa_{up} \cdot (J_y^{CKMR} - J_{targ}^{CKMR})) \tag{3.17}$$

$$\text{If } J_y^{CKMR} < J_{targ}^{CKMR}: \; TAC_{y+1}^{CKMR} = TAC_y^{CKMR} \times (1 + \kappa_{down} \cdot (J_y^{CKMR} - J_{targ}^{CKMR})) \tag{3.18}$$

Constraints on DMRCKMR:

$$1) \text{ If If } TAC_{y+1}^{CKMR} > 28000, \text{then } TAC_{y+1}^{CKMR} = 28000$$

$$2) \;\; \text{For } \;\; \mathrm{y}_1 \leq y \leq y_2: \;\; J_{targ}^{CKMR} = (\tfrac{T_2 - T_1}{y_2 - y_1}) \cdot (y - y_1) + T_1,$$

$$\text{For } \;\; \mathrm{y}_2 < y: \;\; J_{targ}^{CKMR} = T_2,$$

$$\mathrm{J}_{targ,y}^{CKMR} = J_{targ}^{CKMR}(1 + \lambda^{CKMR}(y - y_t))$$

where λ^{CKMR} is a parameter that stipulates the p.a. increase in J_{targ}^{CKMR} after tuning year y_t which is either 2035 or 2040.

The constraints for DMRCKMR required a more complicated form than was needed for CPUE and GT. The first constraint was a bound of 28000 mt tuna, which is the same constraint that was applied for CPUE and GT. The same form of the second constraint was set up for the CKMR projections, as shown in equation 3.11, to achieve the 2035 target level; but this resulted in a TAC that first decreased and later increased rapidly. Such a trend in the TAC is undesirable for industry stability.

DMRcomb is a CMP that uses a weighted combination of DMRCPUE, DMRGT and DMRCKMR from equations 3.13 - 3.18 based on the following formula (Butterworth *et al*, 2019d, e):

$$TAC^{COMB}_{y+1} = w_{CPUE} \cdot TAC^{CPUE}_{y+1} + w_{GT} \cdot TAC^{GT}_{y+1} + w_{CKMR} \cdot TAC^{CKMR}_{y+1} \tag{3.19}$$

$$\text{where } w_{CPUE} + w_{GT} + w_{CKMR} = 1, \text{i.e. the weights sum to one.}$$

All the CMPs described above are tested with the following CCSBT SC-agreed operational constraints:

- TACs are set in 3-year blocks.
- TAC is restricted to a maximum change of 3 000 mt (up or down).
- The minimum change limit is 100t, hence: $100 \leq |TAC_{y+1} - TAC_y| \leq 3000$ in years when there is a TAC change.
- The maximum TAC for all the CMPs considered is 28 000 mt.

The first three constraints listed above were defaults set by CCSBT, while the last constraint was by choice because this restriction on TAC allowed reasonable SSB future projections. For each of these operational constraints, alternative options were tested in earlier versions of the development process. This was suggested by the CCSBT SC with details described in Appendix B2. Those options explored did not result in a any difference of note so that the constraints above were retained. The alternative options explored for each constraint are listed in the Table below.

Table 3.2: Alternative options to operational constraints explored.

Operational constraint	Alternative option(s)
TACs are set in 3-year blocks.	TACs are set in 2-year blocks. TACs are set in 4-year blocks.
TAC is restricted to a maximum change of 3 000 mt (up or down).	TAC is restricted to a maximum change of 2 000 mt (up or down). TAC is restricted to a maximum change of 4 000 mt (up or down).
The minimum change limit is 100 mt, $100 \leq \lvert TAC_{y+1} - TAC_y \rvert \leq 3000$ in years when there is a TAC change.	
The maximum TAC for all the CMPs considered is 28 000 mt.	The maximum TAC for all the CMPs considered is 32 000 mt.

3.3 Atlantic Bluefin Tuna

3.3.1 Introduction

This section describes the methodology followed in moving towards the development of CMPs to set TACs for Atlantic Bluefin Tuna (ABFT) in the West and East areas of the Atlantic, where the division of the Atlantic into these two areas is shown in Figure 1.1 in Chapter 1. There is a continuing series of deliberations in the process of developing an MP for ABFT to provide improved catch limit recommendations. The CMP approaches developed in this thesis have been part of multiple rounds of these development phases over the period from early 2018 until the end of September 2019, and thus do not describe the final selection by ICCAT (which is still to come). However, for the purpose of this thesis, the "final round" of testing refers to work until the end of September 2019. All documents and updates relate to this period during which the work reported was conducted.

Unlike for SBT, the ABFT OMs will not be covered here in detail because they are currently a work in progress and much more complicated. Therefore, for the purpose of this thesis, the main concentration is on the SBT work and its OMs in relation to the development of that CMP.

Figure 3.3 displays the specific spatial strata defined for the OMs developed for the ABFT MSE, where strata 1 to 3 are in the West area and strata 4 to 7 in the East area. These strata are used in the development of stock mixing aspects incorporated into the OMs (Carruthers, 2019a). Data from different countries and researchers have been assigned to these strata, and incorporated into the ABTMSE R package to be used to provide projections of abundance indices per stratum. These are summarized in Table 3.3.

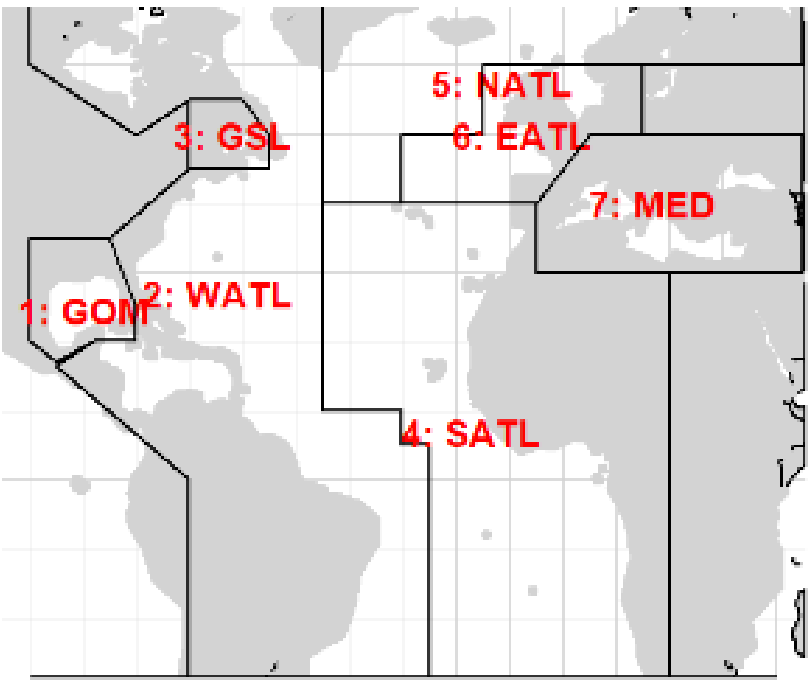

Figure 3.3: The seven spatial strata used for the ABFT OMs (Carruthers, 2019a).

Table 3.3: Details of the indices specified in the ABTMSE.R package version 5.2.3 with the corresponding spatial strata.

Number	Name	Description	Stock	Corresponding spatial strata	Period data available
1	JPN_LL_NEAtl2	Japanese Longline in the North East Atlantic	East	5	2010 - 2017
2	FR_AER_SUV	French Aerial Survey	East	7	2009 - 2017
3	MED_LAR_SUV	Mediterranean Aerial Survey	East	7	2001 - 2016
4	GBYP_AER_SUV	GBYP Programme Aerial Survey in the western Mediterranean	East	7	2010 - 2018
5	JPN_LL_West2	Japanese Longline in the western Atlantic	West	2	2010 - 2017
6	US_RR_66_114	US Rod and Reel 66cm – 114cm West Atlantic	West	2	1993 - 2017
7	GOM_LAR_SUV	Gulf of Mexico Larval Survey	West	1	1977 - 2017
8	CAN_ACO_SUV	A survey in the Gulf of St Lawrence based on acoustics	West	3	1994 - 2017

The migratory nature of ABFT impacts the OMs for the MSE in an important way through these having to account for the mixing of the Western and Eastern origin stocks. This means that the catch recommendations will be provided for each (West or East) area, but reference to population status (in terms of SSB) will be per stock, i.e. SSB of Western Stock or SSB or Eastern Stock. The reason for this is because the mixing of the stocks affects both areas, and consequently the development and selection of a CMP for the West area will directly affect the development and selection of a CMP for the East area, and *vice versa*.

A solution to the ABFT stock mixing situation is to develop a coupled CMP that takes account of mixing considerations in such a way that a CMP for the West area and a CMP for the East area are tested simultaneously. This is structured computationally within an ABFT MSE R package (ABTMSE.R) provided by T. Carruthers under the support of the ICCAT Grand Bluefin Tuna Year Programme (GBYP). The R package is accompanied by a Candidate Management Procedure Developers Guide that advises users on how to install, design and test CMPs in R using the package, an ABFT MSE package user guide which is a more extensive guide, and a trial specifications document (TSD) which gives full details of the MSE trials for ABFT. These two user guides and the TSD are comprehensively used (and referenced) in this research.

3.3.2 The operating model and robustness tests

In 2014 the GBYP proposed an MSE project for ABFT to produce robust management advice for the resource that is consistent with the precautionary approach (Carruthers, 2019a). To achieve this, an extensive OM is necessary to be able to examine the effect of catch recommendations on ABFT realistically. Like the SBT OM, the ICCAT ABFT OM is also based on reference set of trials which are all specified in the TSD which is kept updated by T. Carruthers (Carruthers, 2019b). Recruitment, spawning fraction, natural mortality rate and the mixing of both Western and Eastern Stock are the main uncertainty factors that are covered by these OMs, as shown in Table 3.4. Table 3.5 specifies the factorial design and labelling of the reference set OMs. Table 3.6 lists three primary robustness that tests have been examined, but there are 35 specified robustness tests which are listed and described in Appendix C1 (Butterworth *et al.* 2020). The deterministic results for depletion and average catch for the CMPs described in Section 3.3.3 which are applied to the corresponding ROM_1 to ROM_30 are also provided in Appendix C1. ROM13, ROM14 and ROM15 in Table 3.6 correspond to ROM_1, ROM_2 and ROM_3 from the priority robustness tests in Table C.1 in Appendix C1, and these are used to explore an extreme scenario related to Western stock abundance.

Table 3.4: Uncertainty factors and uncertainty levels (within each factor) that constitute the Operating Models reference set and robustness tests.

Recruitment		
	Western stock	**Eastern stock**
1	B-H with h=0.6 ("high R0") switches to h = 0.9 ("low R0") starting from 1975	1950-1987 B-H h=0.98 switches to 1988+ B-H h=0.98
2	B-H with h=0.6 fixed, high R0	B-H with h=0.7 fixed, high R0
3	Historically as in Level 1. In projections, "low R0" switches back to "high R0" after 10 years	Historically as in Level 1. In projections, 1988+ B-H with h=0.98 switches back to 1950-1987 B-H with h=0.98 after 10 years.
	Spawning fraction both stocks	**Natural Mortality rate both stocks**
A	Younger (W and E same)	High
B	Older (W and E older but different for the 2 stocks)	Low
	Mixing	
I	Best estimates	
II	Four times increase in weight of likelihood component for electronic tags (increased Eastern stock in West, decreased Western stock in East)	

Table 3.5: The factorial design and labelling of the reference set operating models (Carruthers, 2020).

Mixing	I		II	
Spawning fraction/Mortality	A	B	A	B
Recruitment 1	OM_1	OM_4	OM_7	OM_10
Recruitment 2	OM_2	OM_5	OM_8	OM_11
Recruitment 3	OM_3	OM_6	OM_9	OM_12

Table 3.6: Priority Robustness Tests (Carruthers, 2020)

	One factor deviation from OM		
	OM_4: 1wBI	OM_5: 2wBI	OM_6: 3wBI
Western Contrast. Increased precision (CV of 15%) of the GOM_LAR_SUV index to create greater contrast in current Western stock status	ROM 13	ROM 14	ROM 15

The notation of the robustness tests in Table 3.6 is to link to the factors and levels specified in Table 3.4. There have been ongoing discussions about issues related to uncertainty and the associated factors and levels, where each axis reflect a particular uncertainty factor; these remain in discussion and under consideration for revision, since research on them amongst ICCAT scientists is still in progress. There were modifications to Table 3.4 as work conducted progressed from 2018 to 2019. A broad summary of the key uncertainty factors and levels as they stood in September 2019 is provided below.

Recruitment:

The recruitment factor accounts for the argued of regime shifts in the past and consequently also possible regime shifts in the future; should such occur the Beverton-Holt (B-H) stock-recruitment relationship parameters change.

Level 1: Both the Western and Eastern stock undergo regime shifts, where the B-H stock-recruitment relationship parameters change as follows (Carruthers, 2020):
Western stock: B-H with steepness fixed at 0.6 until 1974 with a high estimated R0 value switches to B-H with steepness h fixed at 0.9 with a low estimated R0 as of 1975.
Eastern stock: B-H with steepness fixed at 0.98 with an initial estimated R0 value from 1950 to 1987, then switches to Beverton-Holt (B-H) with steepness still fixed to 0.98 but now a different estimated R0 value as from 1988.

Level 2: No regime shift for both the Western and Eastern stock in the past, nor future. The B-H stock-recruitment relationship for each stock remains unchanged:
Western stock:: Beverton-Holt (B-H) with steepness fixed at 0.6 with a single estimated value R0.

Eastern stock: Beverton-Holt (B-H) with steepness fixed at 0.7 with a single estimated value R0.

Level 3: Since there may have been regime shifts in the past as specified in level 1, there is then the possibility of a regime shift in the future. Both the Western and Eastern stock undergo these.
Western stock: Historically as for level 1. Future projections reflect a change in the Beverton-Holt (B-H) stock-recruitment relationship where steepness changes from 0.9 back to 0.6 and low R0 switches back to high R0 after 10 years.
Eastern stock: Historically as for level 1. Future projections reflect a change in the Beverton-Holt (B-H) stock-recruitment relationship where steepness remains the same but R0 switches back to the initial R0 after 10 years.

Spawning fraction and natural mortality:

The Western stock spawns in the Gulf of Mexico and Eastern stock spawn in the Mediterranean. In the Mediterranean there is evidence that fish spawn from about age three, but in the Gulf of Mexico only large fish from about age 10 move there where they spawn. This then raises the question: how is it possible that the same fish exhibit such different behaviour? Similarly, there is ongoing discussion about how long the mature fish (of spawning age) live. This is an unresolved issue, so that there are options termed Level A and Level B applicable to the Western and Eastern stocks. **Level A** reflects the assumption that fish that mature young also have a higher natural mortality, whereas **Level B** assumes that fish that mature older have a lower natural mortality.

Mixing:

The extent to which the Western and Eastern stock mix is still unclear, as archival tagging as well as microchemistry and genetic data indicate that there are a fair number of Eastern stock fish in the West area and that there are some Western stock fish in the East area. There is still ongoing discussion concerning the extent of this mixing; thus there are two levels as options for mixing: best estimates are represented by **Level I** and a higher level represented by **Level II**.

3.3.2.1 Description of the fixed parameters of the OM.

Table 3.7 lists the fixed parameters (user specified) used in the OM, and the numerical values for the base case and alternative options are shown in Appendix C3. The TSD (version 19-4) provides more details of other estimated parameters and the conditioning of the OMs (Carruthers, 2020).

Table 3.7: The parameters of the OM that are fixed.

Parameter	Symbol
Steepness	h
Maximum length	L_∞
Growth rate	K
Age at length zero	t_0
Natural mortality rate at age	M
Selectivity of at least one fleet	θ
Maturity at age	mat

Steepness (h):
The parameter h is steepness, which is related to the shape of the stock-recruitment function. There is a description of this parameter in section 3.2.2.4 for SBT. Recruitment is calculated from a Beverton-Holt stock recruitment relationship with steepness h, where Table 3.4 specifies the different steepness values used for the OMs.

Maximum length, growth rate and age at length (L_∞, K and t_0):
These are parameters of the von Bertalanffy growth curve used for eastern stock; for the western stock the Richards growth curve was eventually used - the associated parameters are given in Appendix C3.

Natural mortality rate of fish (M):
There is a description of this parameter in section 3.2.2.4 for SBT, and details of these values used for natural mortality are given in Appendix C3.

Selectivity of at least one fleet (θ):
Gear selectivity for each baseline fleet type is estimated through use of the ICCAT catch-at-size dataset in conditioning the OMs. There are 17 fishing fleets that are incorporated into the OMs as described in Table C.4 in Appendix C2. The selectivity of the two fleets in question, selected for the base case and for the alternative option OMs, are Fleet13 'TPnew' and Fleet14 'CAN RR' as shown in Appendix C2; selectivity for these fleets is assumed to be logistic, whereas more flexibility is allowed for the other fleets (Carruthers, 2019c).

Maturity at age (mat):
As with SBT, maturity changes with age but the OM equations and the values differ for the Western and Eastern stock with details given in Appendix C3.

3.3.3 The ABFT CMPs

This section will focus on the development of a coupled CMP which consists of two CMPs that are tuned to indirectly take account of the mixing of the Western and Eastern ABFT stocks. The CMPs are applied to various OMs and robustness tests as shown in section 3.3.2. However, prior to these most recent (as of September 2019) OMs and robustness tests, there were three earlier ABTMSE package updates over the period from 2018 to 2019, where the OMs and robustness tests differed from the TSD - Draft annex Version 19-4 considered here. The values of the control parameters of the CMPs undergo a process of tuning and testing to achieve CMPs that result in reasonable trade-offs between the conservation and utilisation of ABFT by taking substantial but low risk catches from the Western and Eastern stocks; hence these values had to be adjusted each time the package was updated.

The CMPs considered were based on composite abundance indices that aggregated over those which will be available in the future in the manner described below.

First, in an effort to attain comparable sizes at the present time per area, combine across all indices to be considered in an area. This was done by normalising each index relative to its mean over a recent period of years for which the index had been roughly stable. This was done by use of equation 3.20 which will result in the 2019 values of normalised index I_y^{i*} being close to 1.

$$I_y{}^{i*} = \frac{I_y^i}{\text{Average of historical } I_y^i} \tag{3.20}$$

Following this, the variance for each normalised index in each area over the period used for normalisation is calculated to determine the weight to give to each index when combining across the different indices. This is done as follows, taking the weight to be inversely proportional to the variance in the equation below:

$$w_i = \frac{1}{(\sigma^i)^2} \tag{3.21}$$

A composite index is then calculated as a weighted average over all the indices available for each of the West and East areas, where n is the number of indices per area:

$$J_y = \frac{\sum_i^n w_i \times I_y^{i*}}{\sum_i^n w_i} \tag{3.22}$$

The final form of this composite index which is used in the CMPs is the average of each index (from equation 3.22) for the last three years for which data are available, where these averages are taken to smooth over the noise in the series:

$$J_{av,y} = \frac{1}{3}(J_y + J_{y-1} + J_{y-2}) \tag{3.23}$$

Then the TAC for the West and East areas are calculated as:

$$TAC_{W,y} = (\frac{TAC_{W,2018}}{J_{W,2016}}) \cdot \beta \cdot J^W_{av,y-2} \tag{3.24}$$

$$\text{If } TAC_{W,y} \geq 1.2 * TAC_{W,y-1} \text{ then } TAC_{W,y} = 1.2 * TAC_{W,y-1}$$

$$\text{If } \mathrm{TAC}_{W,y} \leq 0.8 * TAC_{W,y-1} \text{ then } TAC_{W,y} = 0.8 * TAC_{W,y-1}$$

$$TAC_{E,y} = (\frac{TAC_{E,2018}}{J_{E,2016}}) \cdot \alpha \cdot J^{E}_{av,y-2} \tag{3.25}$$

$$\text{If } TAC_{E,y} \geq 1.2 * TAC_{E,y-1} \text{ then } TAC_{E,y} = 1.2 * TAC_{E,y-1}$$

$$\text{If } \mathrm{TAC}_{E,y} \leq 0.8 * TAC_{E,y-1} \text{ then } TAC_{E,y} = 0.8 * TAC_{E,y-1}$$

CMP specifications:

- These TACs are set in two year blocks.
- Each TAC is subject to a 20% minimum/maximum change for each of the two areas, as indicated by the equations above.

Equations 3.20 to 3.25 represent the core equations used in the development of the CMPs. However, several subsequent changes were made to these CMPs by adjusting the equations above, and consequently also retuning of the values of the control parameters that ultimately determined the TACs, as is described below. Similarly, to the initial constant proportion CMP used for the development of SBT CMPs in equation 3.5, here reference is made to "fixed proportion" CMPs where the control parameters β and α are multiples of the J_{av} value for the West and East areas at the time. These parameters are used to tune to achieve a CMP that reflects an acceptable trade-off between catch and stock recovery.

A summary of CMP modifications and OM package revisions is described briefly below. Further investigations of CMPs applied to OMs were carried out to improve results that initially reflected large declines in SSB values or extremely low catches within a 30 year projection period. The OMs were also revised throughout the testing process to be able to improve the realism of the scenarios. The results of CMPs applied to OMs highlight the crucial effect that OM plausibility has on the range of stock abundances that need to be considered, in particular as regards low Western stock abundance.

It is important to note that each set of OM revisions corresponded to a specific ABTMSE R package at the time. Here, details reported pertain only to the "final round" OM package during the period of this work (which started in 2018 and ended in September 2019). Table 3.8 displays a list of the OM package versions, the different CMPs tested specified by their parameter values and the number of OMs specified in the OM package version. The results of the final CMPs developed for the West and East areas will be presented and discussed in Chapter 4.

The ABTMSE package includes deterministic and stochastic OMs. The CMPs applied to deterministic OMs were set to the selection of the perfect observation model and with no implementation error, and the CMPs applied to stochastic OMs were set to the selection of the good observation model and with no implementation error (Butterworth *et al.*, 2020). The perfect observation error model tests MPs with with no observation error, which accounts for no difference between what the model predicts and what is observed. Additionally, the recruitment follows the stock recruitment relationship exactly, so that there is no variation there. The stochastic OMs are the ones that introduce variations in both those respects.

Only the stochastic OM is realistic but applying CMPs to stochastic OMs is time consuming since multiple simulations are projected and averaged, and there is more variation in observation error and about the stock recruitment relationship. However, the reason CMPs are also applied to deterministic OMs is because it is time effective since there is only one projection, and if a CMP does not reflect desirable performance when applied to deterministic OM, it will not either for a stochastic one. In that way, deterministic runs are conveniently used for initial CMP testing.

Table 3.8: List of the OM package versions, CMPs names with corresponding β and α parameter values, the number of OMs specified in the OM package version and the year CMP testing occurred.

	CMP			OMs	Year
	CMP name	β	α		
Initial CMPs applied to OM package v2.8.0	CMP1.1	1.1	1.1	8	2018
	CMP1.5	1.5	1.5		
	CMP1.7	1.7	1.7		
	CMP1.5Jthresh	1.5	1.5	1 (OM1)	
Second round CMP testing applied to OM package v3.3.0	CMP0.5	0.5	0.5	8	2018
	CMP1.0_0.5	1.0	0.5		
	CMP0.5_1.0	0.5	1.0		
	CMP1.0	1.0	1.0		
Third round CMP testing applied to OM package v4.2.15	FXPx	0.75	1.5	16	2019
	M_FXP	0.5	1.75		
Final round CMP testing applied to OM package v5.2.3	FXP_1	0.5	0.5	12	2019
	FXP_2	1.0	1.0		

Summary of CMP modifications:

Initial CMPs applied to OM package v2.8

A key concern that arose from applying CMPs CMP1.1, CMP1.5 and CMP1.7 initially to the OM package v2.8.0 was that future biomass projections were shown to drop quite low. Specifically, the lower 10 percentile of B/B_{MSY} at the end of the 30 year projection was very low (Butterworth *et al.*, 2019a). To resolve this issue, a threshold value, $J^{W/E}_{av,thres}$, was introduced to allow the values of β and α to drop linearly to zero as J_{av} decreases as follows:

$$\text{If } J^W_{av,y} < J^W_{av,thres}, TAC_{W,y} = (\frac{TAC_{W,2018}}{J_{W,2016}}) \cdot (\frac{J^W_{av,y}}{J^W_{av,thres}}) \cdot \beta \cdot J^W_{av,y}, \tag{3.26}$$

$$\text{If } J^E_{av,y} < J^E_{av,thres}, TAC_{E,y} = (\frac{TAC_{E,2018}}{J_{E,2016}}) \cdot (\frac{J^E_{av,y}}{J^E_{av,thres}}) \cdot \alpha \cdot J^E_{av,y}, \tag{3.27}$$

where the values $J^W_{av,thres} = 0.405$ and $J^E_{av,thres} = 0.463$ were selected to get improved conservation performance without impacting catches too much in situations where the resource was not seriously depleted.

Second round CMP testing applied to OM package v3.3.0

There was an update to the OM for the second round of CMP testing. Simple fixed proportion CMPs were again applied to the 8 conditioned OMs in OM Package version 3.3.0, where equations 3.20 to 3.25 above are again used as the CMPs. There were no modifications to the CMP equations at this stage, but there was an update to the OMs to provide more realistic models of the Western stock. As noted in Table 3.8, four CMPs with ranges for two control parameters are selected (Butterworth, *et al.*, 2019b).

Third round CMP testing applied to OM package v4.2.15

The above was followed by a third round of CMP testing applied to OM package v4.2.15. Two CMPs, FXPx and M_FXP, were used, with M_FXP being a more conservative variant of FXPx. FXPx is described by equations 3.20 to 3.25 above, but M_FXP incorporates a restriction that allows for the reduction of TACs in the West area to be faster if the resource indices there are reflecting low values. Instead of four indices used to determine the TAC in the West area, only the Gulf of Mexico larval survey index is used instead to calculate J_y (from equation 3.22) for reasons which are detailed below. The adjustment was that if $J^W_{av,y-2} > 1$, $TAC_{W,y}$ then equation 3.24 still provided the TAC in the West area for M_FXP, but if $J^W_{av,y-2} \leq 1$, the following applied:

$$TAC_{W,y} = (\frac{TAC_{W,2018}}{J_{W,2016}}) \cdot \beta \cdot \frac{J^W_{av,y-2} - J2}{1 - J2} \quad \text{if } J2 \leq J^W_{av,y-2} \leq 1 \tag{3.28}$$

$$TAC_{W,y} = 0 \quad \text{if } J^W_{av,y-2} < J2 \tag{3.29}$$

$$\text{and if } TAC_{W,y} \geq 1.2 * TAC_{W,y-1} \text{ then } TAC_{W,y} = 1.2 * TAC_{W,y-1}$$

Hence, if $J^{W}_{av,y-2} < J2$, the TAC for the West area can decrease to as low as zero, instead of the initial 20% minimum restriction for the CMP specification (see equation 3.24). $J2$ was set to 0.7 to obtain reasonable performance (Butterworth *et al.*, 2019c).

Final round CMP testing applied to OM package v5.2.3

The final round of CMP testing explored the results of two updated CMPs, FXP_1 and FXP_2, that both reflect conservative approaches. Table 3.8 shows the differences in the alpha and beta values for FXP_1, for which $(\alpha, \beta) = (0.5, 0.5)$ and for FXP_2, for which $(\alpha, \beta) = (1.0, 1.0)$. These CMPs were applied to the conditioned OMs in OM Package version 5.2.3. Importantly, there were two key changes to both these CMPs from the CMPs developed earlier: (1) the TACs in the West and East areas were restricted to be below specified values for FXP_1 and FXP_2, and (2) a modification was made to the TAC for the West area in circumstances where indices related to the west stock abundance dropped too low for FXP_1 and FXP_2 to provide adequate conservation performance (Butterworth, *et al.*, 2020). Hence for these new CMPs:

1) There are TAC (maximum) restrictions for the West and East areas to prevent stock abundance from dropping excessively where regime shifts to lower productivity levels occur. These restrictions are additional to the 20% minimum/maximum TAC restrictions placed on all CMPs throughout the development process, as indicated in equations 3.24 and 3.25.

$$\text{If } TAC_{W/E} \geq TAC_{W/E,max} \text{ then } TAC_{W/E,y} = TAC_{W/E,max} \tag{3.30}$$

2) CMP deterministic runs reflected that in some cases the Western stock abundance dropped too low because the TAC for West area (equation 3.24) did not reduce the TAC fast enough to promote stock recovery. The Gulf of Mexico larval survey index may reflect low stock abundance in the West more reliably because it is the only index in the West area where the tuna abundance is not influenced by the presence of eastern origin fish. This is addressed by altering the TAC for West area set through equation 3.22, and consequently equation 3.30 as well to now use Gulf of Mexico larval survey index only, as opposed to all four Western indices.

$$TAC^{*}_{W,y} = X_{W,y} TAC_{W,y} \tag{3.31}$$

$$X_{W,y} = \begin{cases} 1 & = \text{if } I^{*smooth}_{y-2} \geq T \\ \\ I^{*smooth}_{y-2} & = \text{if } I^{*smooth}_{y-2} < T \end{cases}$$

where T is a control parameter specifying the index value at which the formula changes (see Fig. 3.4) and and $I^{*smooth}_{y}$ is a smooth normalized index based on the Gulf of Mexico larval survey index I_y:

$$I^{*}_{y} = I_y / \left(\frac{1}{y_2 - y_1 + 1} \sum_{y'=y1}^{y_2} I_{y'}\right) \tag{3.32}$$

This normalized index is then smoothed:

$$I_y^{*smooth} = \frac{1}{3}(I_y^* + I_{y-1}^* + I_{y-2}^*) \tag{3.33}$$

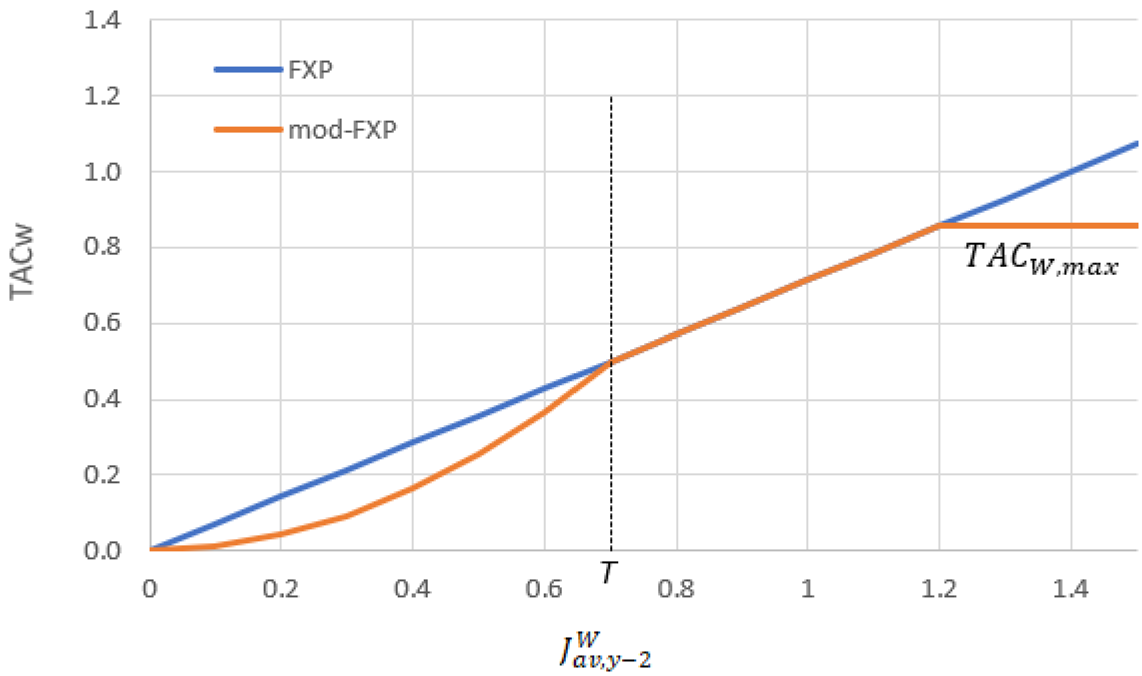

Figure 3.4: Illustrative relationship between the TAC set for West area and the index used for West area for FXP and its modified CMP denoted mod-FXP.

Figure 3.4 shows the relationship between the TAC set for West area ($TAC_{W,y}$) and the index used for West area (J^w). The TAC formula for both FXP_1 and FXP_2 are represented by FXP and mod-FXP, depending on how the Western stock behaves. The blue line represents the TAC formula set for West area in equation 3.24 using the index for the West area described in equations 3.20 to 3.24, denoted FXP. The orange curve represents the TAC formula set for West area subject to modifications depicted by equations 3.31 to 3.33, denoted mod-FXP. Since only the Gulf of Mexico larval survey (GOM_LAR_SUV) is used for FXP_2, $J_{av,y}^W$ from equation 3.23 can be compared to $I_y^{*smooth}$ in equation 3.32, where $I_{y-2}^{*smooth}$ can be referred to as $J_{av,y-2}^W$ (even though it's not an average) for FXP_2 for simplicity in the plot.

Chapter 4

Results and Discussion

4.1 Southern Bluefin Tuna

4.1.1 Results

Table 4.1 lists the values determined by tuning for the various control parameters used in the five CMPs for the final round of CMP development detailed in section 3.2.3.3. DMRMCPUE, DMRMGT and DMRMCKMR each use only CPUE, GT or CKMR indices of abundance respectively, while DMRcomb1 and DMRcomb2 are weighted combinations of these three. The CMPs are tuned (using the control parameters) to achieve a median SSB which is (1) 30% of its pristine value by 2035 or (2) 35% of its pristine value by 2040, for the reference set (or grid, termed base18). These are the recovery targets specified by the Commission for SBT (i.e. the CCSBT). Since both DMRcomb1 and DMRcomb2 are weighted combinations of the first three CMPs, and the weight parameters for these CMPs (denoted w_{CPUE}, w_{GT} and w_{CKMR}) sum to 1, these combination CMPs did not (in practice) require further tuning.

Table 4.2 and Table 4.3 list the stochastic results for the performance statistics for the CMPs for the base18 reference set, the CPUE variable squares robustness test (cpuew0) and a future low recruitment robustness test (reclow5) for the two tunings. Median values are shown with 90% PIs except for P(2up/1down). This last statistic is the probability that after increases in the first two TAC changes, there is a decrease; the reason for its inclusion is explained in the section below. The stochastic simulations involve 2000 repetitions. Figure 4.1 and Figure 4.9 display guitar plots of the results listed in Tables 4.2 and Table 4.3, respectively.

Figures 4.2 to 4.7 show the TAC and SSB trajectory plots for CMPs tuned to achieve a median SSB which is 30% of its pristine value by 2035, for the base18 reference set, the CPUE variable squares robustness test (cpuew0) and a low recruitment robustness test (reclow5), respectively. Figures 4.10 to 4.15 show these same trajectory plots for CMPs tuned instead to achieve a median SSB which is 35% of its pristine value by 2040 for the same three scenarios.

Figures 4.8 and Figure 4.16 show guitar plots of the performance statistics for the DMRcomb2

CMPs which achieve median SSBs of 30% by 2035 and 35% by 2040 of the pristine SSB value for the base18 reference set and when these are applied to the ten robustness tests, respectively.

4.1.2 Discussion

4.1.2.1 Performance statistics

Before discussing these results, the performance statistics are explained in the context of what would constitute optimal CMP performance. These statistics are listed in Tables 4.2 and 4.3 and many are displayed as guitar plots in Figures 4.1, 4.8, 4.9 and 4.16.

$\boldsymbol{Mean\ TAC(2021-2035)}$ and $\boldsymbol{Mean\ TAC(2036-2050)}$ are the average annual total allowable catches over the specified periods. It is generally desirable for the TAC to be high, but this performance statistic must be assessed in conjunction with SSB recovery. The reason is the inevitable trade-off that a lower recovery objective leads to a high TAC but to perhaps unsatisifactorily poor 5%-iles for resource recovery at the lower end of the resultant distributions, while a higher recovery objective reflects the danger of TAC dropping too low to maintain a healthy industry at the expense of high resource recovery.

$\boldsymbol{SSB_{2035}/SSB_0}$ and $\boldsymbol{SSB_{2040}/SSB_0}$ are the SSB ratios at the end of the projection period relative to the value for a pristine (unfished) resource. Stock conservation is a main priority for this work, so that the selection of a CMP that ensures stock recovery with a value of this statistic typically close to that corresponding to MSY is very important.

$\boldsymbol{SSB_{2035}/SSB_{2018}}$ and $\boldsymbol{SSB_{2050}/SSB_{2018}}$ are the SSBs at the end of these projection periods relative to those in the year 2018 when the projections start. Note that management under the new MP starts from 2021, as TACs as determined by the Bali MP for 2018 to 2020 are already in place, as shown in Figure 3.1. These fixed TACs early in the projection period are taken into account in the code.

$\boldsymbol{\%AAV(2021-2035)}$ and $\boldsymbol{\%AAV(2036-2050)}$ are the average percentage annual catch variation values, which provide an indication of industrial stability. The intent is to keep AAV low to promote such stability. Footnote 4 in Chapter 2 (page 7) shows the formula used for AAV.

$\boldsymbol{P(2up/1down)}$ is the probability of initial TAC increases for the first two TAC changes, followed by a decrease in the third. The intent is to keep P(2up/1down) low, again for industrial stability. However, a relatively high value of P(2up/1down) for certain robustness tests involving a future period of low recruitments, such as reclow5, could be desirable as such a drop in the TAC may be needed to reduce the risk of resource depletion once the abundance indices indicate that this is occurring.

$\boldsymbol{P(SSB_{2030} > 0.3SSB_0)}$ is the probability that the SSB in 2030 is higher than 30% of the pristine SSB; this performance statistic indicates the probability that the 2035 tuning target level has already been reached five years before the first tuning target date.

4.1.2.2 Comparative CMP performance of DMRMCPUE, DMRMGT and DMRMCKMR

The control parameters were selected to achieve desirable CMP performance that reflects what are considered to be reasonable trade-offs among the performance statistics. The control parameters β, γ and κ used in DMRMCPUE, DMRMGT and DMRMCKMR respectively, are referred to as "gain parameters". They indicate the extent to which the TAC is to be adjusted in relation to the corresponding stock level at the time relative to some specified value (another control parameter). The formula for adjusting the TAC is set up in such a way that the gain parameter increases the TAC when the index is above the specified value, or decreases it when the index is below this value.

In Figure 4.1, the SSB_{2035}/SSB_0 guitar plots show that the GT based CMP generally has the highest value for the lower 5%-ile for this statistic for the base18 and reclow5 scenarios (but DMRCPUE is slightly higher for cpuew0), and the CKMR based CMP has the lowest, so that the GT based CMP is the best performing in this respect. However, the CKMR based CMP clearly exhibits the lowest AAV values over 2021 to 2035, while the GT based CMP reflects the highest, so that the CKMR based CMP is best in those terms. Even though the CPUE based CMP is intermediate in regard to the trade-off between avoiding low SSB and securing low AAV, they have the lowest 5%-ile for the mean TAC for the period 2036 to 2050. Poor performance in that regard would risk the TAC dropping unnecessarily low in the long run because median SSB for the CPUE based CMPs remains well above its 2018 level by 2050 as shown in Figure 4.1. Similar trends and trade-offs are evident in Figure 4.9 for the CMPs tuned to recover to 35% of pristine SSB by 2040.

The TAC trends for the different CMPs are more readily evident from Figures 4.2, 4.4 and 4.6. These show that (in median terms) the TAC for the CKMR based CMP initially increases steeply, whereas the CPUE and GT based CMPs show more gradual increases. However, the ordering of these rates of increase reverses later before the tuning year is reached.

4.1.2.3 The key differences in performance for the two tunings for base18 from Figure 4.1 and Figure 4.9

The 5%-ile of the mean TAC (from 2021 to 2050) is generally lower for the second tuning (for 2040) than for the first (for 2035). This means that there is a higher risk of the catch dropping to low levels for the second tuning. There is no particular difference in AAV between the two tunings, with the CPUE based CMP generally exhibiting the highest AAV for both. The SSB ratio statistics generally exhibit similar depletion patterns across CMPs, with CKMR having the lowest 5%-iles. The medians for SSB_{2035}/SSB_0 and SSB_{2040}/SSB_0 for all the CMPs are above the 13% abundance level at the start of 2016. There are no major differences between $P(SSB2030 > 0.3SSB0)$ for the two tunings, but the second tuning has slightly higher probabilities. $P(2up/1down)$ is reasonably low across CMPs for both tunings, but the CKMR based CMPs show distinctly higher $P(2up/1down)$ statistics than the other CMPs for the second tuning.

4.1.2.4 Robustness test analyses for cpuew0 and reclow5

The worst depletion performance occurs under the cpuew0 scenario for which the 5%-ile SSB_{2035}/SSB_0 and SSB_{2040}/SSB_0 values for all the CMPs are well below the 13% abundance level at the start of 2016. For all the SSB performance statistics, the worst stock recovery is evident for this robustness test for both tuning levels. Though the GT based CMP generally shows the lowest risk of undesirable stock depletion (lower 5%-ile SSB_{2035}/SSB_0 and SSB_{2040}/SSB_0) for base18 and reclow5, the CPUE based CMP shows better depletion performance for cpuew0. This can be seen in Table 4.2 and 4.3, as well as in Figures 4.1 and 4.9.

Reclow5 is a more plausible robustness test scenario than cpuew0, and it was rated as a high priority robustness test by the CCSBT ESC compared to cpuew0 which they rated as low priority. These two were selected for more attention in this thesis as the first analyses conducted indicated that they led to especially large impacts on the results. Figure 4.1 and Figure 4.9 show that the GT based CMP shows the best trade-off providing the largest TAC in the long run (from 2036-2050) and the lowest SSB depletion for reclow5.

The weighted combination CMPs address these problem areas of CMP performance by assigning appropriate weights to the more basic CMPs which are being combined. The CPUE based CMP addresses the SSB depletion best over all for all the robustness tests combined, but the GT based CMP is best for the reclow5 robustness test. DMRcomb1 assigns equal weight to each indices component as a basis for evaluating the performance for the "simplest" form of combination CMP. For DMRcomb2, the decision was made to assign more weight to the GT based CMP compared to the other two because not only does it provide the best performance for reclow5, but it also generally shows lower risk for stock depletion except for cpuew0.

4.1.2.5 Key differences in performance for the two tunings of DMRcomb2 for base18 and the ten robustness tests

The 5%-ile for the mean TAC (from 2021 to 2050) is generally lower for the second tuning than for the first; this means that there is a higher risk of catch decreasing further for the second tuning of DMRcomb2. For both tunings, there is clearly a higher risk of the catch dropping to low levels for the variable squares robustness tests (cpuew0) and the robustness test that combines variable squares and low recruitment (cpuew0reclow5). Generally, AAV is lower for the first tuning, with cpuew0 and cpuew0reclow5 exhibiting the highest AAVs for both. The 5%-iles for the SSB ratio indices are the lowest for cpuew0 and cpuew0reclow5 for both tunings. From Figure 4.8, the 5%-ile SSB in 2035 is below the 2016 level for those two robustness tests.

For the first tuning, cpuew0, reclow5, as2016reclow5 and cpuew0reclow5 reflect high $P(2up/1down)$ values. However, the $P(2up/1down)$ is relatively lower for the second tuning than for the first for all the robustness tests, even for reclow5 and as2016reclow5 which have the highest $P(2up/1down)$ for the second tuning. There are no major differences for $P(SSB2030 > 0.3SSB0)$ for both tunings, but the second tuning has slightly higher probability under robustness test fis20.

In general, the worst performance for both tunings for DMRcomb2 occurs for the robustness tests cpuew0 and cpuew0reclow5.

4.1.2.6 Comparison of the four CMPs submitted to the CCSBT ESC meeting in September 2019

At the end of this CCSBT MP revision process, the “best” CMPs developed by four different groups, together with the results from their simulation tests, were submitted to the CCSBT ESC meeting held in Cape Town in September 2019. These four included DMRcomb2; the other three were named RH13, AAA and NT4. Note that all the CMPs reported upon below were applied to the base18 reference set and its associated robustness tests, and that unless otherwise indicated, references below to trends refer to medians.

A general overview of the comparative performance of these CMPs is provided below. All four reached both tuning levels without needing short-term decreases in the TAC, all achieved the specified recovery objectives and exceeded the interim rebuilding target of attaining at least a 70% probability of reaching 20% of SSB_0 by 2035 (see Appendix B2), and all generally showed satisfactory performance under the base18 reference set.

For the first tuning to 30% of pristine SSB by 2035 in median terms

Figure 4.17 shows that there is a general increase in TAC for all four CMPs over the first tuning period to reach 30% of pristine SSB by 2035. There is a steady TAC increase for DMRcomb2 and AAA throughout this period, with DMRcomb2 projecting the largest increase by 2035. The increasing trends in TAC for RH13 and NT4 are initially slightly steeper than those for DMRcomb2 and AAA, but then start to flatten towards a more gradual steady increase. Figure 4.18 shows comparative stochastic results for 50 iterations of all four CMPs for this tuning period. There is a maximum TAC constraint for DMRcomb2 of 28 000 mt which restricts the TAC increase from 2035, but for RH13, NT4 and AAA the TACs continue to increase to 2045 (though still remaining below 28 000 mt in median terms by that time).

Figure 4.19 (deterministic) and Figure 4.20 (stochastic) shows that all four CMPs are above the interim target level (of 20% of pristine SSB) by 2035. There is a general gradual increase in SSB for all the CMPs from the start of the projection until 2035, but after that the SSB projection for RH13 increases slightly, decreases slightly for DMRcomb2 and is broadly steady for NT4 and AAA. Even though DMRcomb2 leads to the highest TAC, viewed in a broader context it rises too high because this fast initial increase to 2035, even though followed by a period where it is steady, leads a problem because this results in the SSB decreasing from 2035. In contrast, RH13 initially increased TACs slower overall to 2035, allowing for SSB to increase after 2035 which was an important reason why this CMP was preferred over the others.

For the second tuning to 35% of pristine SSB by 2040 in median terms

Figure 4.21 shows that there is a general gradual increase in the TAC projections that is similar across all four CMPs over the second tuning period for which they reach 35% of SSB by 2040. This tuning reflects a more conservative approach, which is why these TAC increases occur more gradually through this period than for the first tuning. DMRcomb2 shows the largest TAC increase by 2040. Figure 4.22 shows comparative stochastic results for 50 iterations of all four CMPs for this tuning level. Even though DMRcomb2 provides high TACs, its TAC projections also show the lowest 5%-ile projections which constitute a large risk for the fishery. The possibility of the TAC dropping rather low does lead to a corresponding lower stock risk for DMRcomb2, as shown in Figure 4.24. The possible drops in TAC for RH13, NT4 and AAA are not as low relative to their medians.

Figure 4.24 shows that all four CMPs are above the interim recovery level (20% of pristine SSB) by 2040. The 5%-ile of SSB for DMRcomb2 remains above this interim level, but for RH13, NT4 and AAA this falls slightly below by 2040. The median projections show that RH13 once again shows a slightly higher increase in SSB (Figure 4.23).

At the end of this September 2019 meeting, the ESC recommended the acceptance of the RH13 MP and the tuning to 30% SSB_0 by 2035, based on its performance overall. This recommendation was subsequently accepted by the Commission.

The ESC summarised the reasons for its decision as follows[1]:

"The ESC commended the cooperative, open nature of the MP development and testing process and that this had resulted in considerable sharing of knowledge, data, code and learning. This had improved the performance of all MPs and the understanding of Members. All CMPs perform well, each with their own positive features, making the task of recommending a MP to the EC a challenging one, because generally the differences in performance statistics were quite small. There are, nevertheless, some important differences, and some CMPs perform better over a wider range of criteria and robustness tests than others.

The ESC therefore considered the CMP performance across a broad range of attributes: (i) Risk to SSB; (ii) Short term level of TAC; (iii) Probability of two increases in TAC followed by a TAC drop; (iv) Longer term performance beyond 2035; (v) Nature of the TAC trajectory; (vi) Certainty of future TACs; and (vii) Incorporation of available data sources.

The ESC noted that there are important trade-offs between these attributes, which imply that they need to be considered simultaneously when evaluating the CMPs. The most important trade-off was between the degree of certainty about future catches and the degree of responsiveness and robustness to different uncertainties. The CMPs that resulted in higher certainty about future TACs (narrower range in future catches), also had higher risks to the stock and lower robustness over the range of scenarios evaluated." (CCSBT, 2019b)

[1] This is from the Report of the Twenty Fourth Meeting of the Scientific Committee (ESC24) (CCSBT, 2019b)

Table 4.1: The values of the control parameters for the five CMPs considered.

	Tuned to achieve a median SSB which is 30% of its pristine value by 2035 for the RC (base18) OM grid.					Tuned to achieve a median SSB which is 35% of its pristine value by 2040 for the RC (base18) OM grid.				
	DMRMCPUE	**DMRMGT**	**DMRMCKMR**	**DMRcomb1**	**DMRcomb2**	**DMRMCPUE**	**DMRMGT**	**DMRMCKMR**	**DMRcomb1**	**DMRcomb2**
CPUE:										
β_{up}	0.13			0.13	0.13	0.13			0.13	0.13
β_{down}	0.7			0.7	0.7	0.7			0.7	0.7
J_{targ}^{CPUE}	0.9			0.9	0.9	1.152			1.152	1.152
λ^{CPUE}	0.03			0.03	0.03	0			0	0
GT:										
γ_{up}		0.25		0.25	0.25		0.25		0.25	0.25
γ_{down}		1.25		1.25	1.25		1.25		1.25	1.25
J_{targ}^{GT}		0.47		0.47	0.47		0.65		0.65	0.65
λ^{GT}		0.08		0.08	0.08		0		0	0
CKMR:										
κ_{up}			0.17	0.17	0.17			0.17	0.17	0.17
κ_{down}			0.17	0.17	0.17			0.17	0.17	0.17
$T1$			0.4	0.4	0.4			0.595	0.595	0.595
$T2$			1.5	1.5	1.5			1.695	1.695	1.695
$y1$			2021	2021	2021			2021	2021	2021
$y2$			2030	2030	2030			2030	2030	2030
λ^{CKMR}			0.03	0.03	0.03			0.03	0.03	0.03
Comb:										
w_{CPUE}				1/3	1/5				1/3	1/5
w_{GT}				1/3	3/5				1/3	3/5
w_{CKMR}				1/3	1/5				1/3	1/5

Table 4.2: Performance statistics for the CMPs listed for the reference set (or grid, called base18) and two robustness tests, cpuew0 and reclow5. Median values are shown with 90% PIs except for P(2up/1down). Each CMP is tuned to achieve a median SSB which is 30% of its pristine value by 2035 for the base18 reference set.

MP	run	Mean TAC(2021-2035)	Mean TAC (2036-2050)	% AAV (2021-2035)	% AAV (2035-2050)	SSB_{2035}/SSB_0	SSB_{2040}/SSB_0	P(2up/1down)
DMRMCKMR		21276(20701, 21828)	24670(19554, 27453)	3.9(2.9, 5.5)	3.5(1.2, 7.5)	0.301(0.173, 0.494)	0.316(0.140, 0.588)	0
DMRMCPUE		21286(18978, 23344)	26878(14160, 28000)	7.6(4.2, 11.1)	3.3 (0.000, 13.1)	0.303(0.184, 0.488)	0.315(0.166, 0.566)	0
DMRMGT	base18	22056(19164, 24398)	27679(21921, 28000)	9.4(3.9, 11.2)	1.5(0.000, 9.1)	0.301(0.191, 0.479)	0.301(0.158, 0.548)	0.012
DMRcomb1		21743(19823, 23561)	27130(17636, 28000)	7.3(3.3, 10.1)	2.4(0.000, 10.4)	0.301(0.184, 0.483)	0.305(0.156, 0.553)	0.005
DMRcomb2		21997(19625, 24043)	27400(19375, 28000)	8.6(3.2, 10.6)	1.9(0.000, 9.2)	0.300(0.187, 0.477)	0.301(0.157, 0.547)	0.013
DMRMCKMR		20311(19654, 20930)	15701(11537, 21809)	4.4(3.0, 6.0)	9.3(1.8, 22.3)	0.154(0.050, 0.339)	0.169(0.014, 0.453)	0.048
DMRMCPUE		17087(14825, 19487)	12439(3929, 26171)	7.4(1.9, 14.8)	10.0(2.7, 35.6)	0.186(0.086, 0.364)	0.226(0.084, 0.487)	0.01
DMRMGT	cpuew0	20180(15737, 23563)	24762(9521, 28000)	6.8(2.3, 11.0)	6.3(0.000, 25.8)	0.165(0.083, 0.328)	0.168(0.062, 0.409)	0.058
DMRcomb1		19019(16544, 21379)	16199(5845, 27079)	4.7(2.2, 9.9)	7.6(1.6, 34.6)	0.170(0.077, 0.342)	0.194(0.055, 0.445)	0.31
DMRcomb2		19474(16219, 22594)	19341(6936, 27733)	4.6(1.7, 10.0)	7.0(0.719, 34.2)	0.168(0.080, 0.334)	0.183(0.067, 0.424)	0.152
DMRMCKMR		21259(20698, 21832)	22148(17631, 27052)	3.9(2.9, 5.5)	4.3(1.3, 12.4)	0.224(0.125, 0.368)	0.229(0.086, 0.456)	0
DMRMCPUE		20130(17827, 22555)	20685(9456, 27840)	8.0(4.3, 11.2)	6.9 (0.674, 17.3)	0.234(0.142, 0.371)	0.249(0.124, 0.459)	0.001
DMRMGT	reclow5	20001(16006, 23311)	26435(15419, 28000)	7.5(3.4, 10.8)	4.7(0.000, 12.3)	0.242(0.157, 0.372)	0.247(0.136, 0.459)	0.254
DMRcomb1		20480(18001, 22431)	23794(12176, 27867)	4.9(2.6, 8.5)	4.2(0.540, 14.5)	0.234(0.143, 0.369)	0.240(0.123, 0.448)	0.188
DMRcomb2		20338(17158, 22847)	25407(13496, 27960)	5.5(2.4, 9.4)	4.2(0.117, 13.4)	0.237(0.150, 0.367)	0.243(0.128, 0.454)	0.275

Table 4.3: Performance statistics for five CMPs listed for the base18 reference set and two robustness tests, cpuew0 and reclow5. Median values are shown with 90% PIs except for P(2up/1down). Each CMP is tuned to achieve a median SSB which is 35% of its pristine value by 2040 for the base18 reference set.

MP	run	Mean TAC(2021-2035)	Mean TAC (2036-2050)	% AAV (2021-2035)	% AAV (2035-2050)	SSB_{2035}/SSB_0	SSB_{2040}/SSB_0	P(2up/1down)
DMRMCKMR	base18	19336(18789, 19847)	20948(16511, 25217)	2.4(1.6, 3.8)	4.7(1.3, 9.4)	0.318(0.189, 0.513)	0.350(0.170, 0.620)	0.19
DMRMCPUE		19301(16487, 21732)	24458(10606, 28000)	6.4(2.6, 11.1)	5.8 (0.000, 14.6)	0.322(0.202, 0.507)	0.349(0.200, 0.601)	0
DMRMGT		19429(14051, 23162)	26835(14644, 28000)	8.5(3.3, 13.1)	3.9(0.000, 12.3)	0.329(0.219, 0.497)	0.347(0.215, 0.586)	0.018
DMRcomb1		19456(16388, 21782)	25132(13545, 27920)	4.6(1.8, 9.7)	4.4(0.304, 9.8)	0.322(0.207, 0.502)	0.347(0.201, 0.591)	0.032
DMRcomb2		19508(15262, 22661)	26137(14126, 28000)	6.1(2.3, 10.9)	4.0(0.000, 10.7)	0.325(0.213, 0.498)	0.346(0.208, 0.586)	0.022
DMRMCKMR	cpuew0	18488(17909, 19040)	13389(9789, 18439)	5.0(3.1, 6.9)	8.1(1.6, 20.9)	0.172(0.066, 0.359)	0.202(0.041, 0.491)	0.746
DMRMCPUE		13890(12340, 17195)	8122(2438, 21751)	14.4(3.5, 20.4)	10.9(3.3, 27.0)	0.214(0.106, 0.393)	0.280(0.120, 0.546)	0
DMRMGT		16131(11540, 21582)	18914(4865, 28000)	9.1(3.1, 20.8)	7.2(0.000, 16.5)	0.202(0.119, 0.351)	0.240(0.133, 0.457)	0.011
DMRcomb1		15699(12905, 18922)	11864(3931, 25042)	8.3(1.9, 18.8)	6.6(1.8, 18.4)	0.200(0.106, 0.367)	0.251(0.119, 0.496)	0.026
DMRcomb2		15859(11972, 19940)	14271(4307, 27042)	7.3(1.8, 20.4)	6.4(1.5, 16.5)	0.202(0.114, 0.357)	0.247(0.130, 0.482)	0.029
DMRMCKMR	reclow5	19325(18799, 19808)	18787(15025, 23539)	2.4(1.7, 3.8)	4.3(1.2, 11.0)	0.241(0.140, 0.387)	0.262(0.113, 0.498)	0.172
DMRMCPUE		17759(15836, 20512)	15983(7462, 26786)	9.5(4.0, 11.9)	9.4 (2.9, 17.2)	0.254(0.157, 0.392)	0.288(0.155, 0.502)	0.001
DMRMGT		15827(12214, 20814)	20830(9494, 27848)	10.3(4.3, 17.7)	8.2(0.701, 15.6)	0.276(0.189, 0.403)	0.315(0.195, 0.516)	0.104
DMRcomb1		17208(14393, 20192)	18051(8563, 27118)	5.6(2.1, 13.3)	5.9(1.5, 12.8)	0.261(0.171, 0.395)	0.295(0.174, 0.507)	0.239
DMRcomb2		16484(13033, 20516)	19251(8762, 27551)	7.1(2.6, 15.4)	6.7(1.4, 14.4)	0.270(0.182, 0.400)	0.307(0.188, 0.508)	0.16

Results representative of achieving a median SSB which is 30% of its pristine value by 2035.

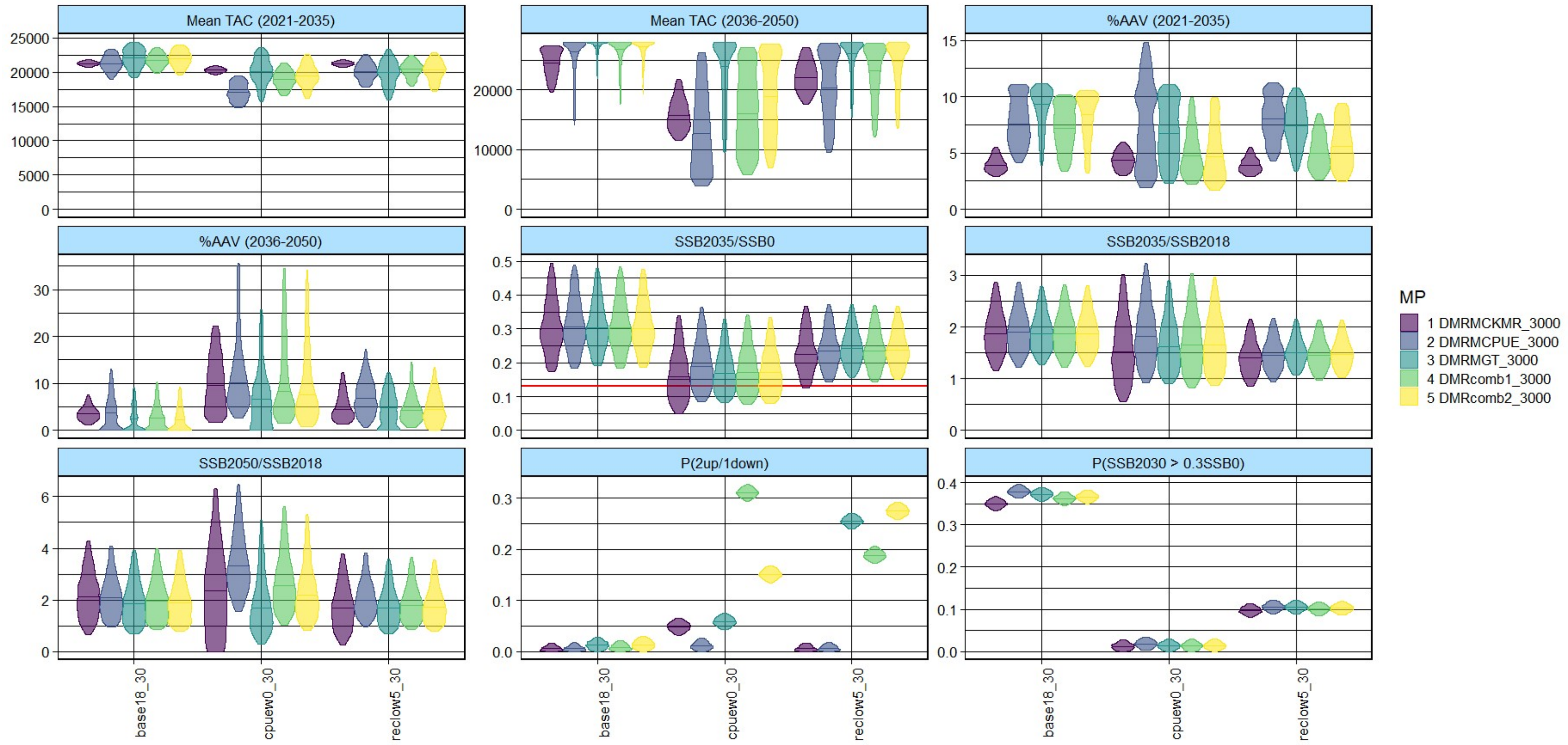

Figure 4.1: Guitar plots for performance statistics for the five CMPs which achieve a median SSB which is 30% of its pristine value by 2035 for the base18 reference set. The red horizonal line on the SSB_{2035}/SSB_0 indicates an abundance level of 13%, estimated in 2017 for 2016, when these projection specifications were agreed by the CCSBT ESC. Guitar plots are mirror image probability density functions shown vertically; the central horizontal line for each is its median.

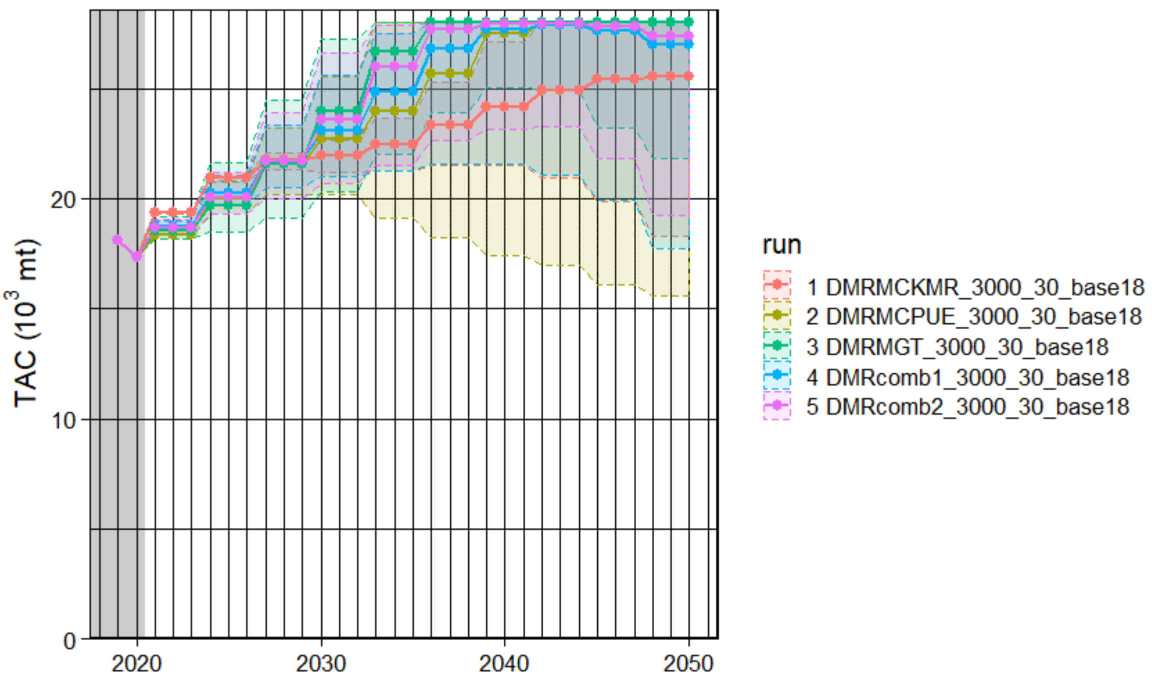

Figure 4.2: TAC plots for the five CMPs investigated tuned to achieve a median SSB which is 30% of its pristine value by 2035 for the base18 reference set. Medians and 90% probability envelopes are shown. In this and the following Figures, probability envelopes retain the same colours as the corresponding medians, but become grey when they overlap.

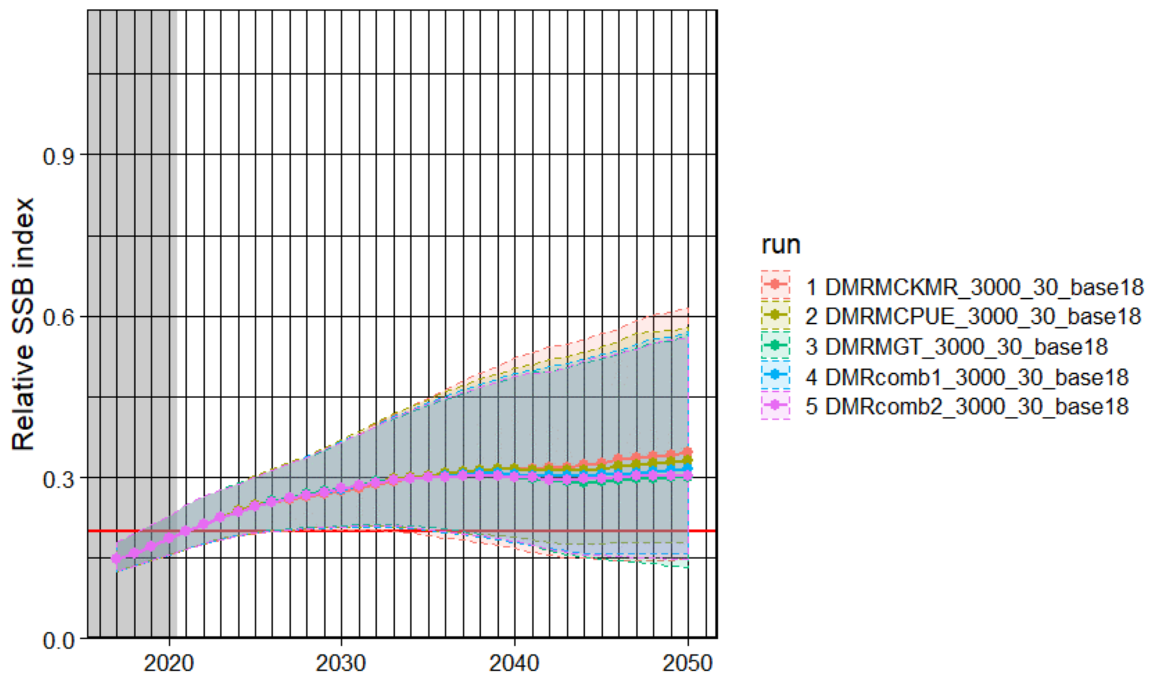

Figure 4.3: SSB trajectory plots for five CMPs investigated tuned to achieve a median SSB which is 30% of its pristine value by 2035 for the base18 reference set. Medians and 90% probability envelopes are shown. The red horizontal lines here indicates 20% SSB_0. This reflects the interim rebuilding target for which the CCSBT specified that any CMP should achieve at least 70% probability of reaching this interim target by 2035.

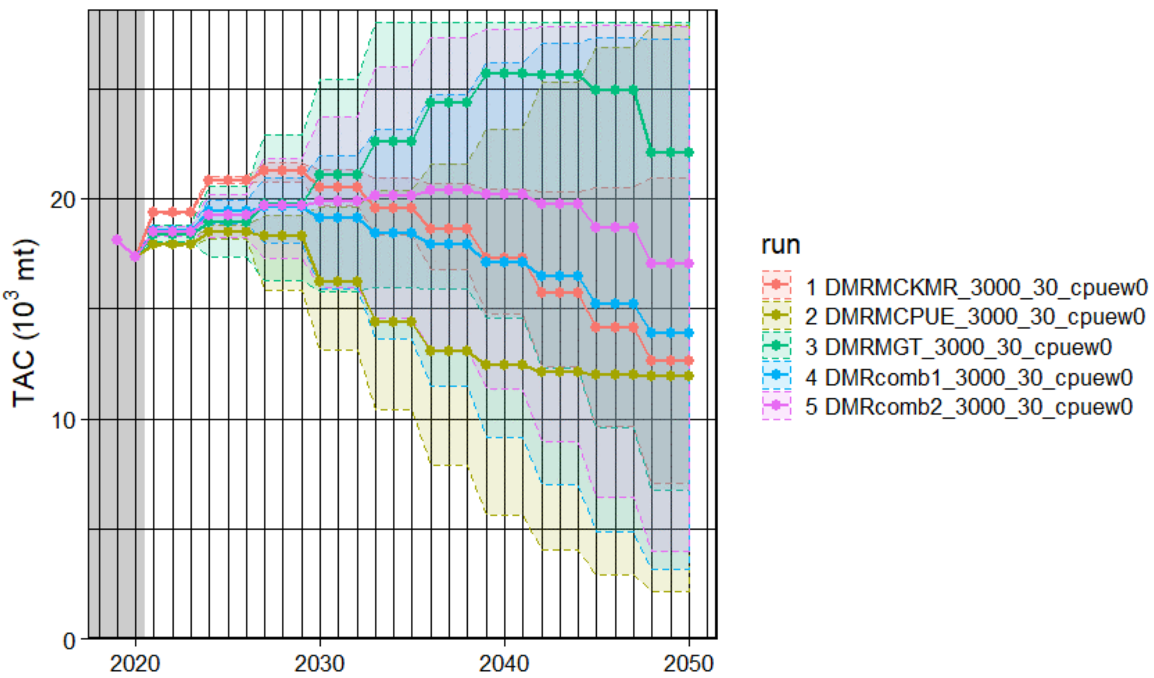

Figure 4.4: TAC plots for the five CMPs investigated tuned to achieve a median SSB which is 30% of its pristine value by 2035 for the base18 reference set when these are applied to robustness test cpuew0. Medians and 90% probability envelopes are shown.

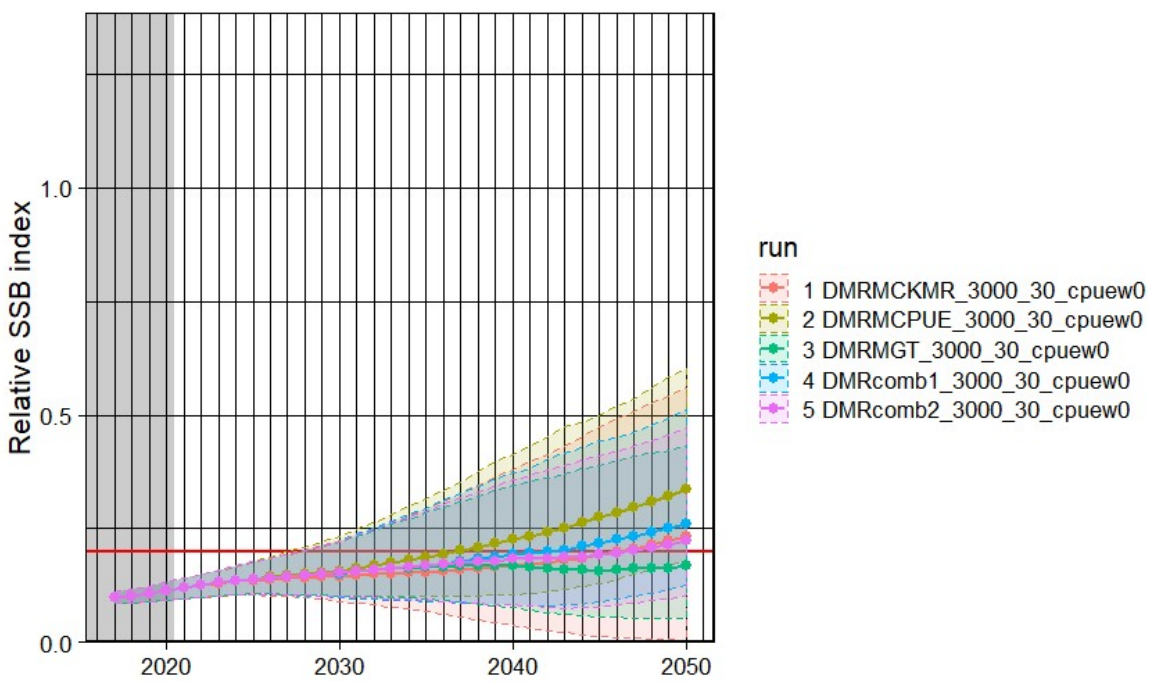

Figure 4.5: SSB trajectory plots for the five CMPs investigated tuned to achieve a median SSB which is 30% of its pristine value by 2035 for the base18 reference set when these are applied to robustness test cpuew0. Medians and 90% probability envelopes are shown. The red horizontal line here indicates 20% SSB_0. This reflects the interim rebuilding target for which the CCSBT specified that any CMP should achieve at least 70% probability of reaching this interim target by 2035.

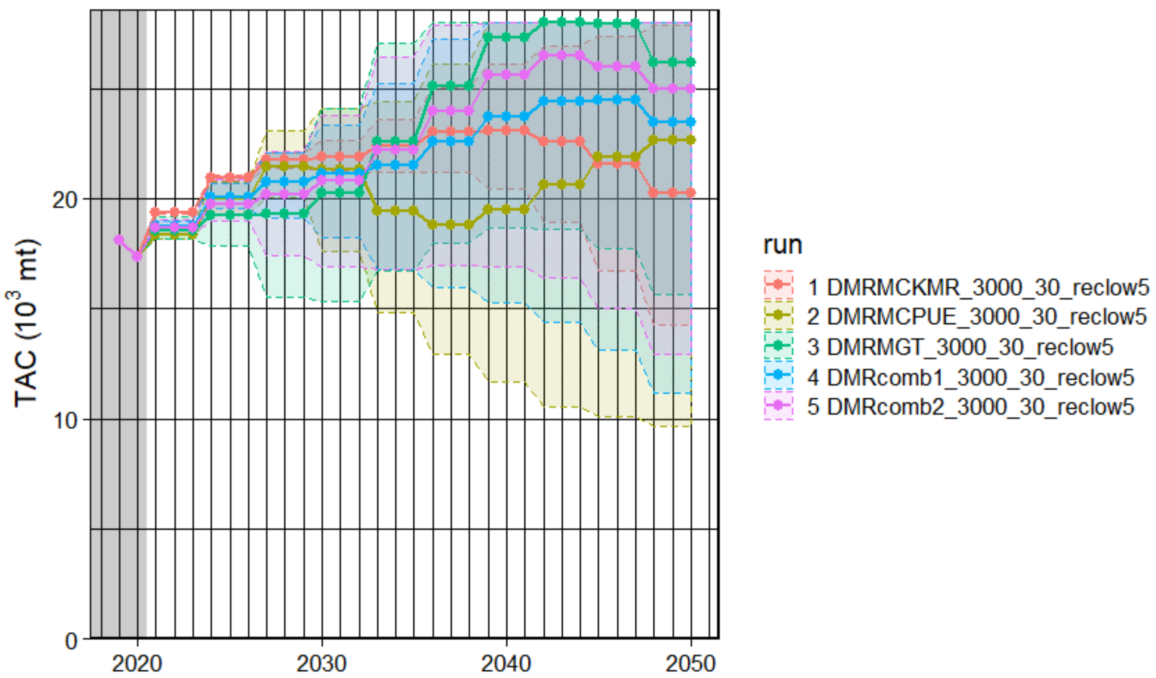

Figure 4.6: TAC plots for the five CMPs investigated tuned to achieve a median SSB which is 30% of its pristine value by 2035 for the base18 reference set when these are applied to robustness test reclow5. Medians and 90% probability envelopes are shown.

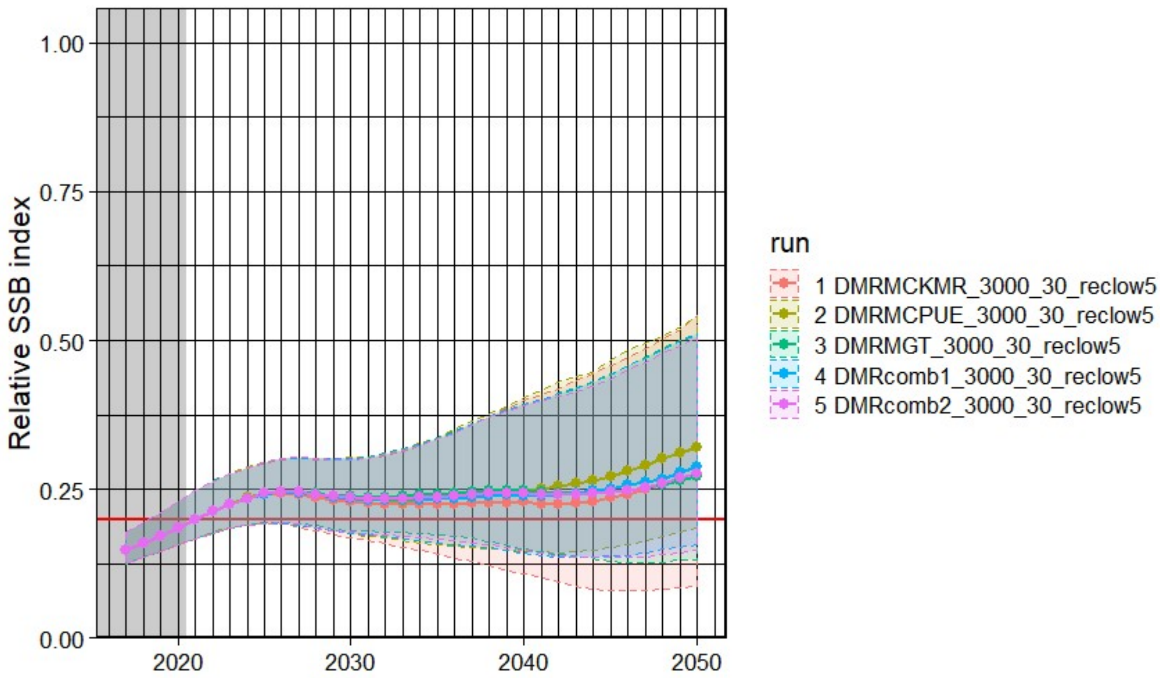

Figure 4.7: SSB trajectory plots for the five CMPs investigated tuned to achieve a median SSB which is 30% of its pristine value by 2035 for the base18 reference set when these are applied to robustness test reclow5. Medians and 90% probability envelopes are shown. The red horizontal line here indicates 20% SSB_0. This reflects the interim rebuilding target for which the CCSBT specified that any CMP should achieve at least 70% probability of reaching this interim target by 2035.

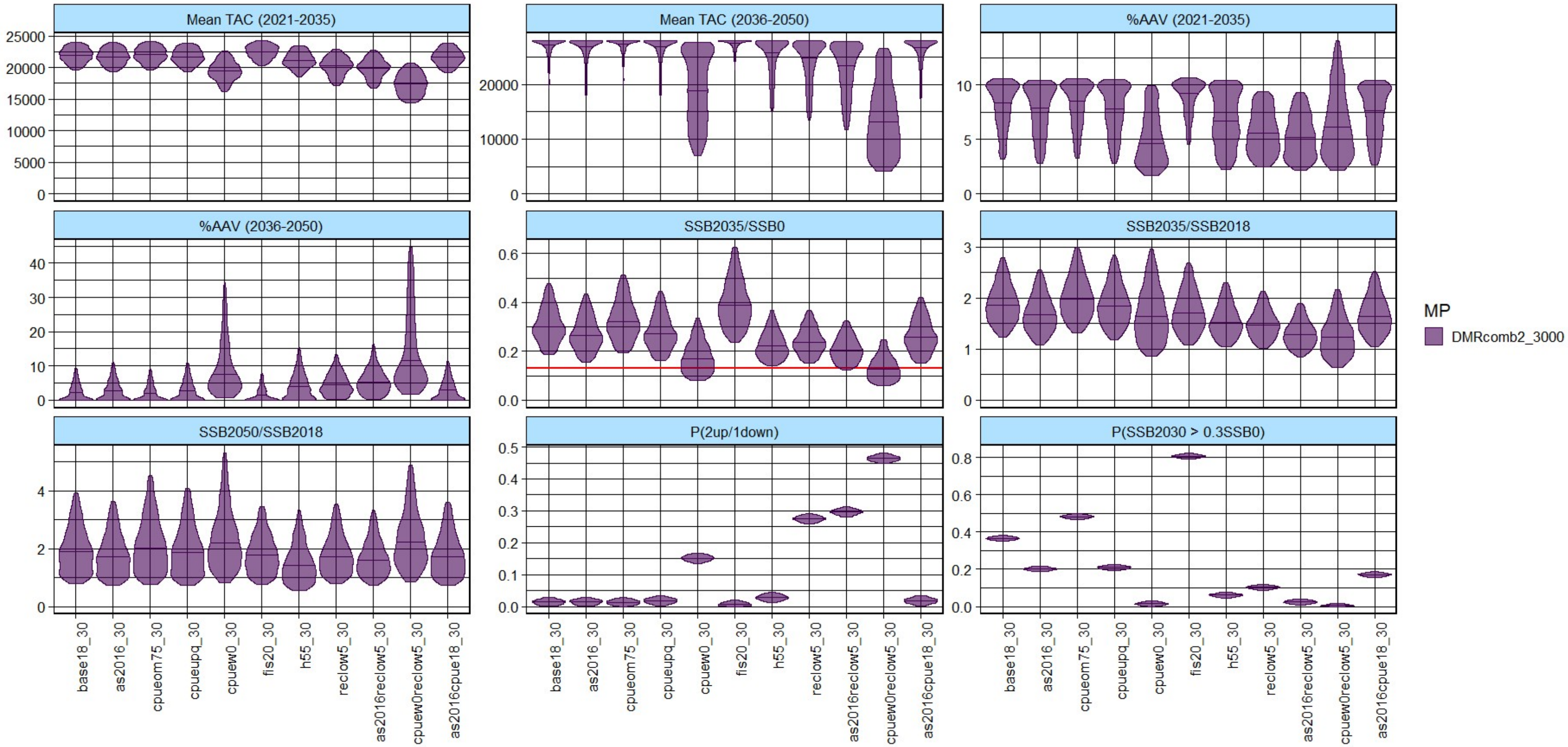

Figure 4.8: Guitar plots of performance statistics of the CMP DMRcomb2 representative of achieving a median SSB which is 30% of its pristine value by 2035 for the base18 reference set when applied to base18 and to ten robustness tests. The robustness tests are all described in Appendix B1. The red horizontal line on the SSB_{2035}/SSB_0 indicates an abundance level of 13%, estimated in 2017 for 2016, when these projection specifications were agreed by the CCSBT ESC. Here "cpue18" means remove the 2018 cpue point; thus the robustness test as2016cpue18 means remove both (CCSBT, 2019a).

Results representative of achieving a median SSB which is 35% of its pristine value by 2040.

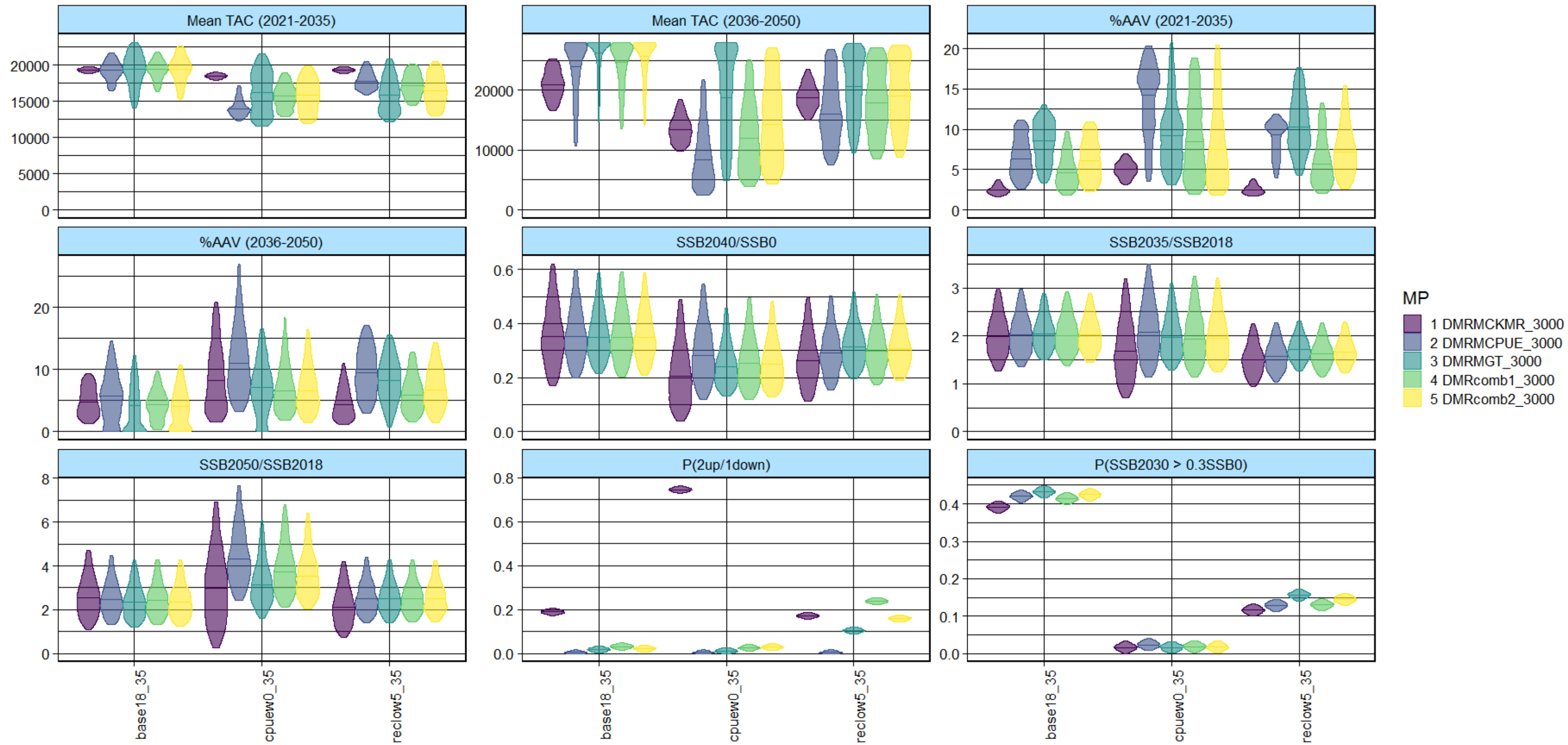

Figure 4.9: Guitar plots of 5 CMPs of performance statistics representative of achieving a median SSB which is 35% of its pristine value by 2040 for the base18 reference set. The central block is tuned to give depletion results for the tuning year 2040, instead of 2035.

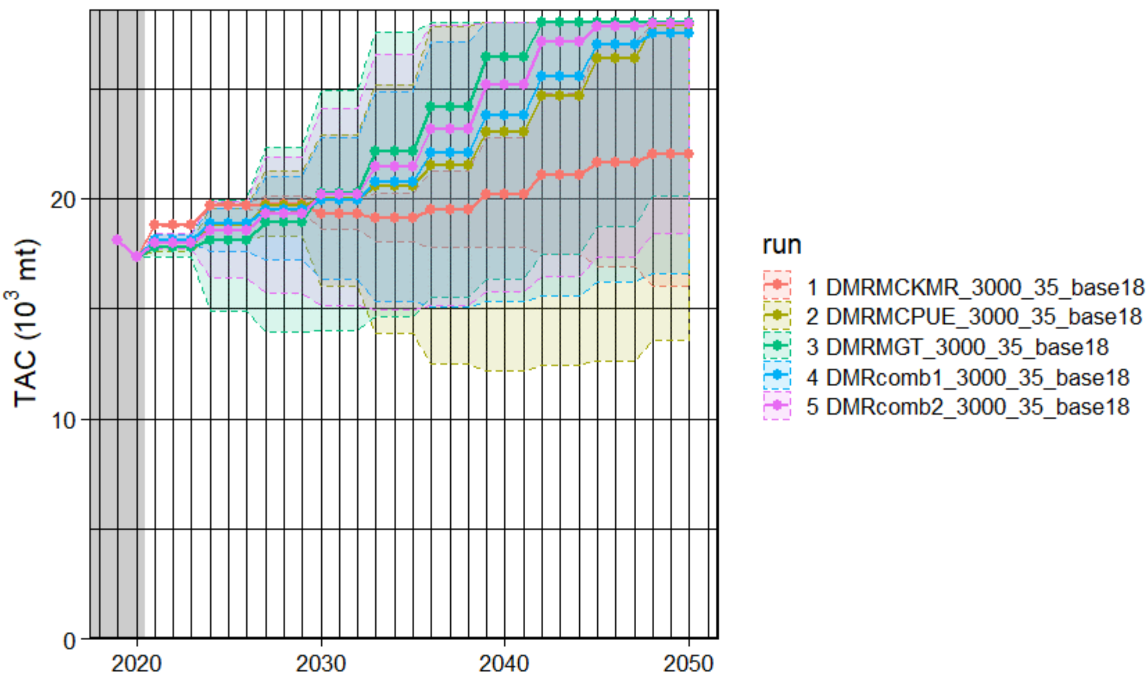

Figure 4.10: TAC plots for the five CMPs investigated tuned to achieve a median SSB which is 35% of its pristine value by 2040 for the base18 reference set. Medians and 90% probability envelopes are shown.

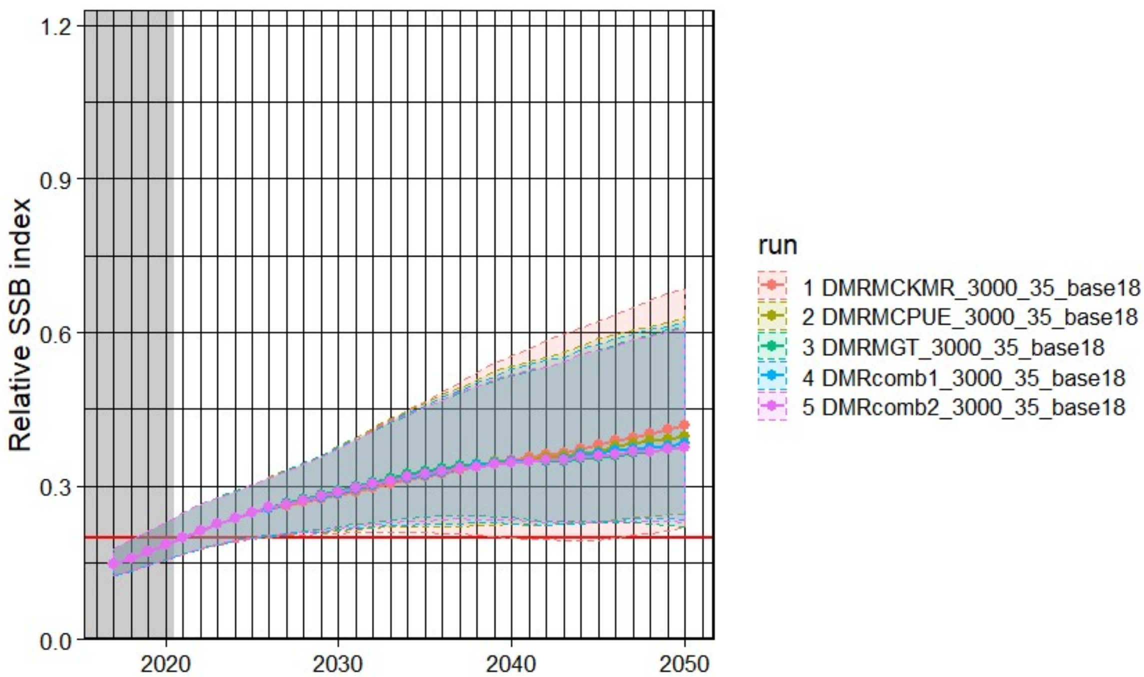

Figure 4.11: SSB trajectory plots for the five CMPs investigated tuned to achieve a median SSB which is 35% of its pristine value by 2040 for the base18 reference set. Medians and 90% probability envelopes are shown. The red horizontal line here indicates 20% SSB_0. This reflects the interim rebuilding target for which the CCSBT specified that any CMP should achieve at least 70% probability of reaching this interim target by 2035.

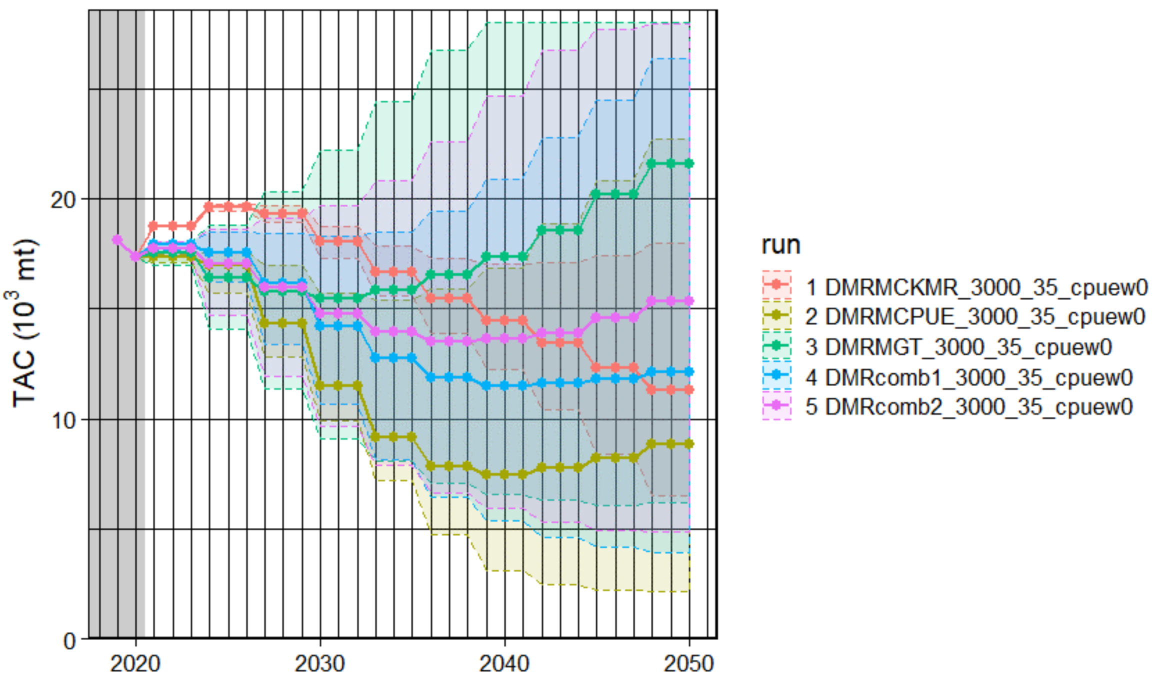

Figure 4.12: TAC plots for the five CMPs investigated tuned to achieve a median SSB which is 35% of its pristine value by 2040 for the base18 reference set when these are applied to robustness test cpuew0. Medians and 90% probability envelopes are shown.

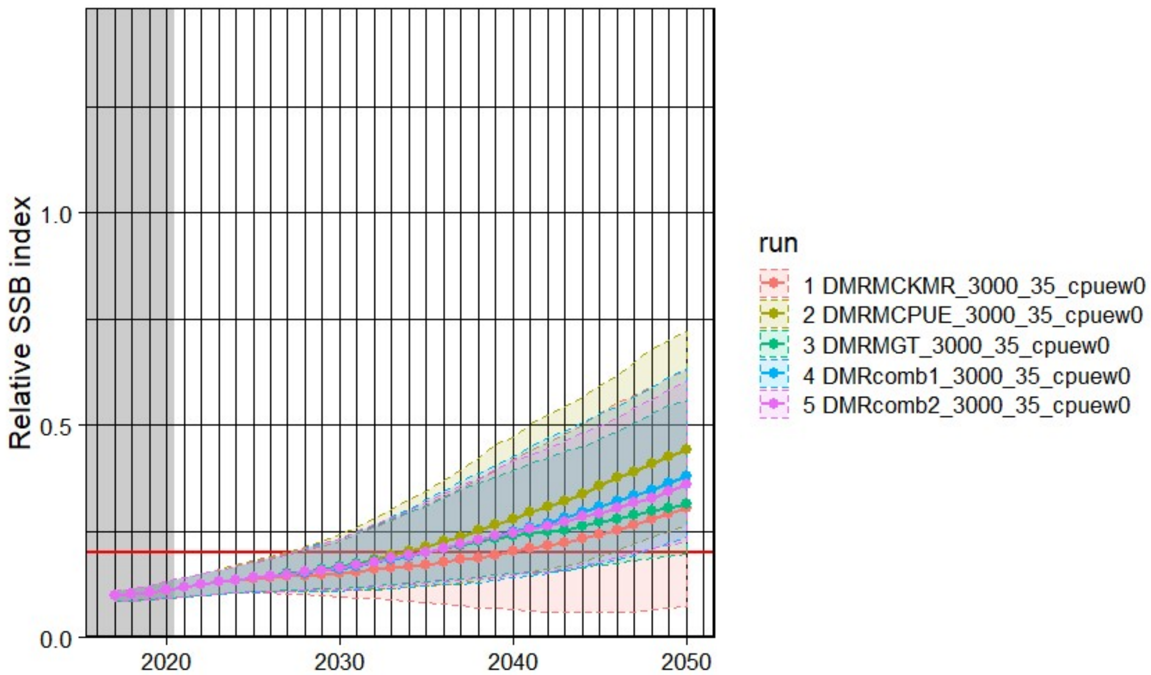

Figure 4.13: SSB trajectory plots for the five CMPs investigated tuned to achieve a median SSB which is 35% of its pristine value by 2040 for the base18 reference set when these are applied to robustness test cpuew0. Medians and 90% probability envelopes are shown. The red horizontal line here indicates 20% SSB_0. This reflects the interim rebuilding target for which the CCSBT specified that any CMP should achieve at least 70% probability of reaching this interim target by 2035.

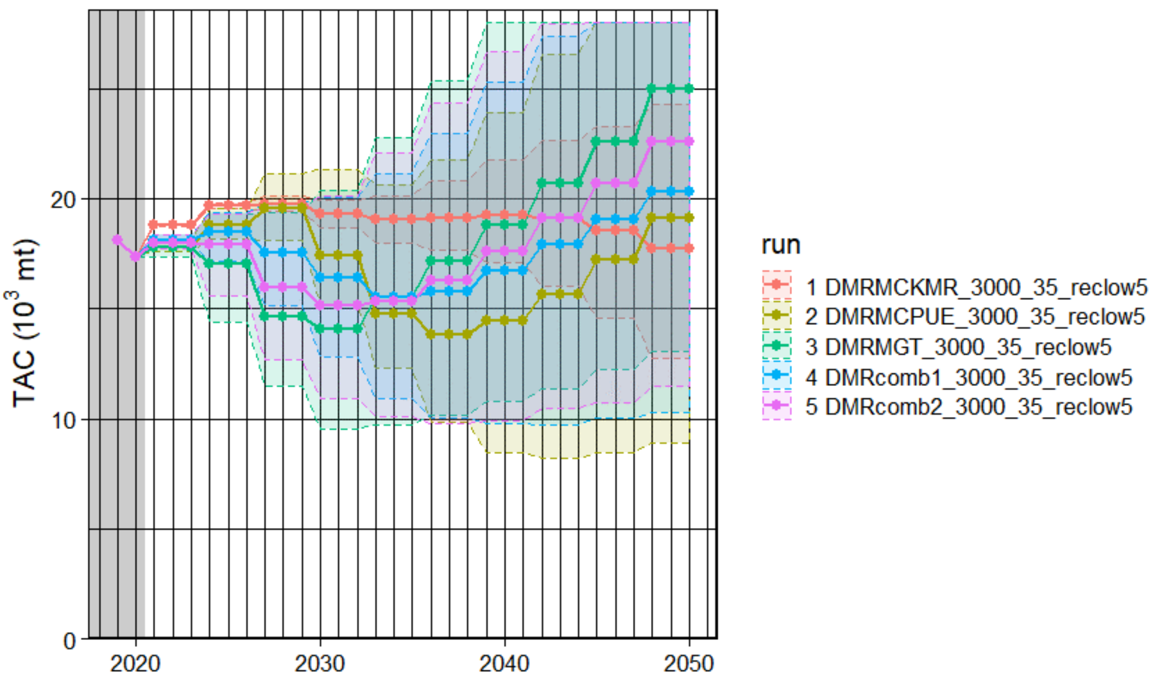

Figure 4.14: TAC plots for the five CMPs investigated tuned to achieve a median SSB which is 35% of its pristine value by 2040 for the base18 reference set when these are applied to robustness test reclow5. Medians and 90% probability envelopes are shown.

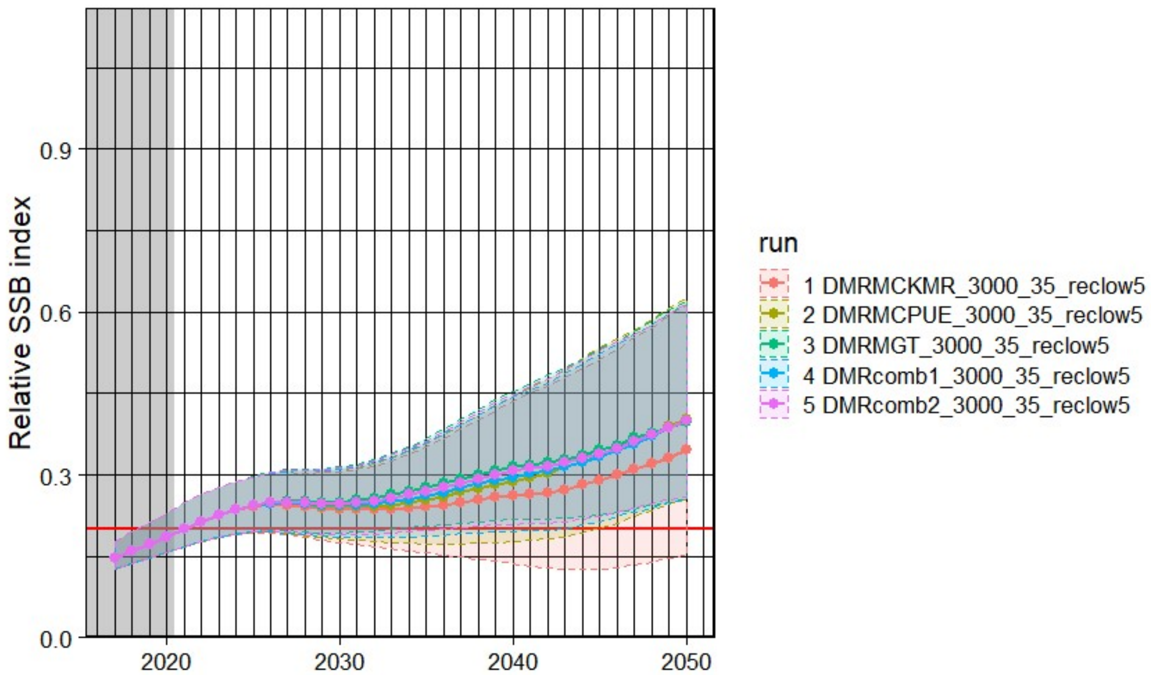

Figure 4.15: SSB trajectory plots for the five CMPs investigated tuned to achieve a median SSB which is 35% of its pristine value by 2040 for the base18 reference set when these are applied to robustness test reclow5. Medians and 90% probability envelopes are shown. The red horizontal line here indicates 20% SSB_0. This reflects the interim rebuilding target for which the CCSBT specified that any CMP should achieve at least 70% probability of reaching this interim target by 2035.

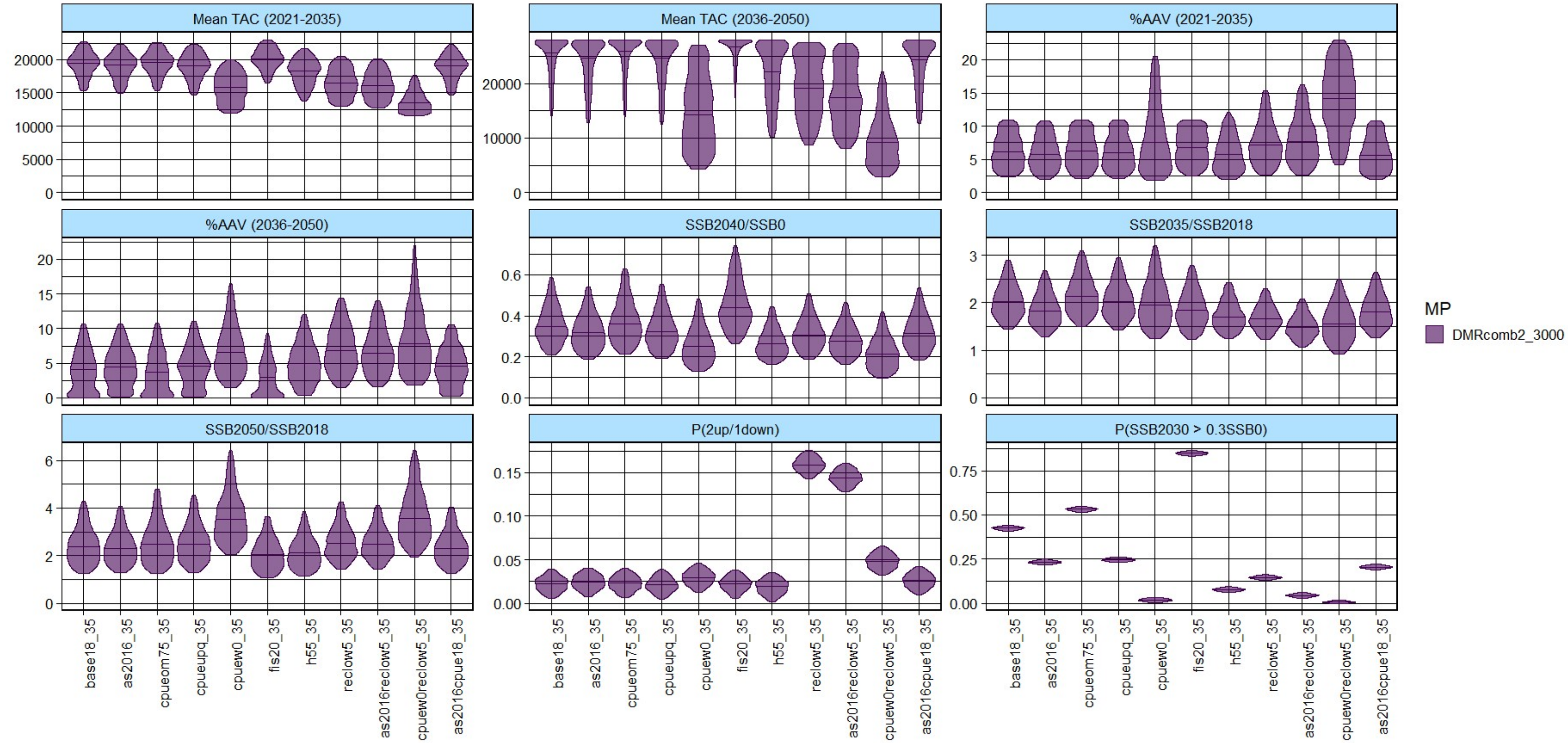

Figure 4.16: Guitar plots of performance statistics for the CMP DMRcomb2 which achieves a median SSB which is 35% of its pristine value by 2040 for the base18 reference set when applied to base18 and to ten robustness tests. The robustness tests are described in Appendix B1. Here "cpue18" means remove the 2018 cpue point; thus the robustness test as2016cpue18 means remove both (CCSBT, 2019a).

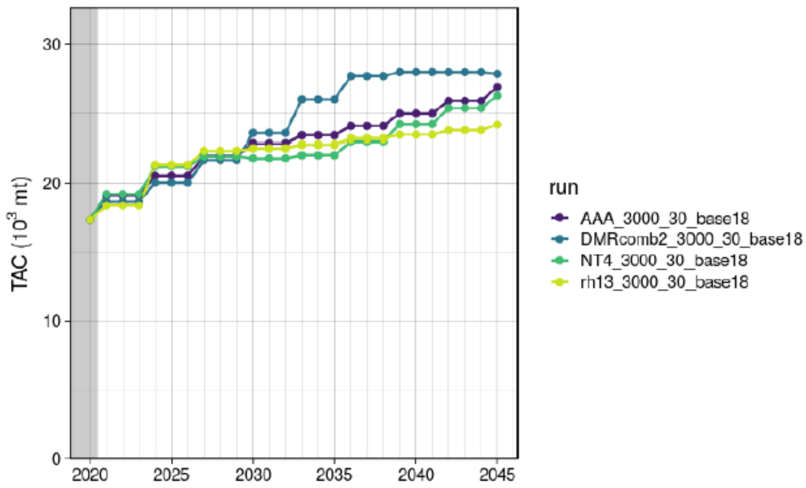

Figure 4.17: The median TAC plots for the four final selected CMPs tuned to achieve a median SSB which is 30% of its pristine value by 2035 under the base18 reference set.

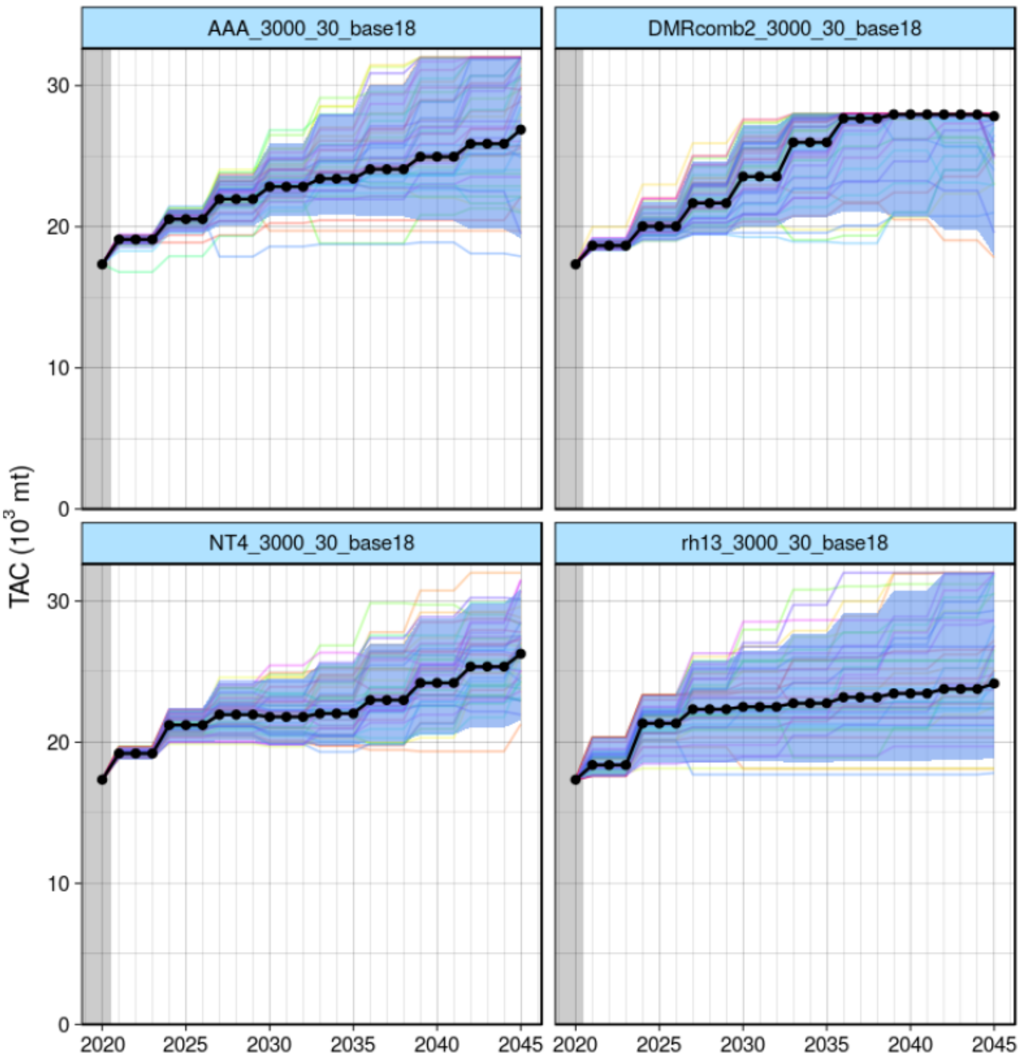

Figure 4.18: The TAC plots for the four final selected CMPs tuned to achieve a median SSB which is 30% of its pristine value by 2035 under the base18 reference set. Medians and 90% probability envelopes are shown. Here and sometimes below individual trajectory realisations are shown, with each having a different colour.

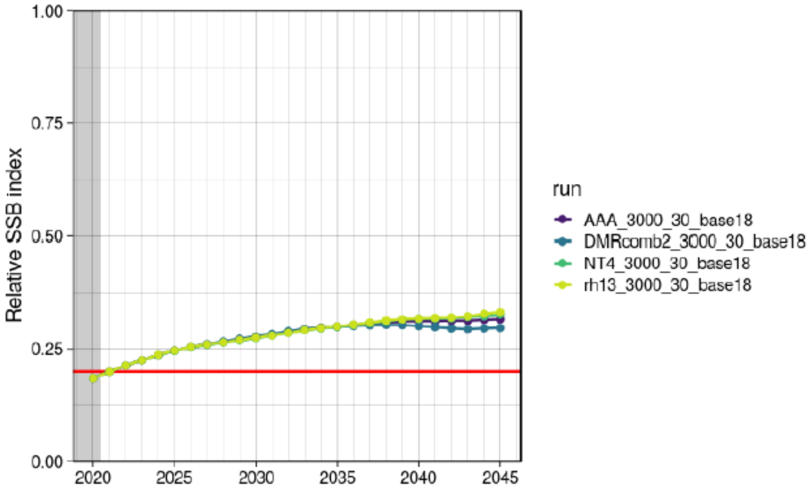

Figure 4.19: The median SSB trajectory plots for the four final selected CMPs tuned to achieve a median SSB which is 30% of its pristine value by 2035 under the base18 reference set.

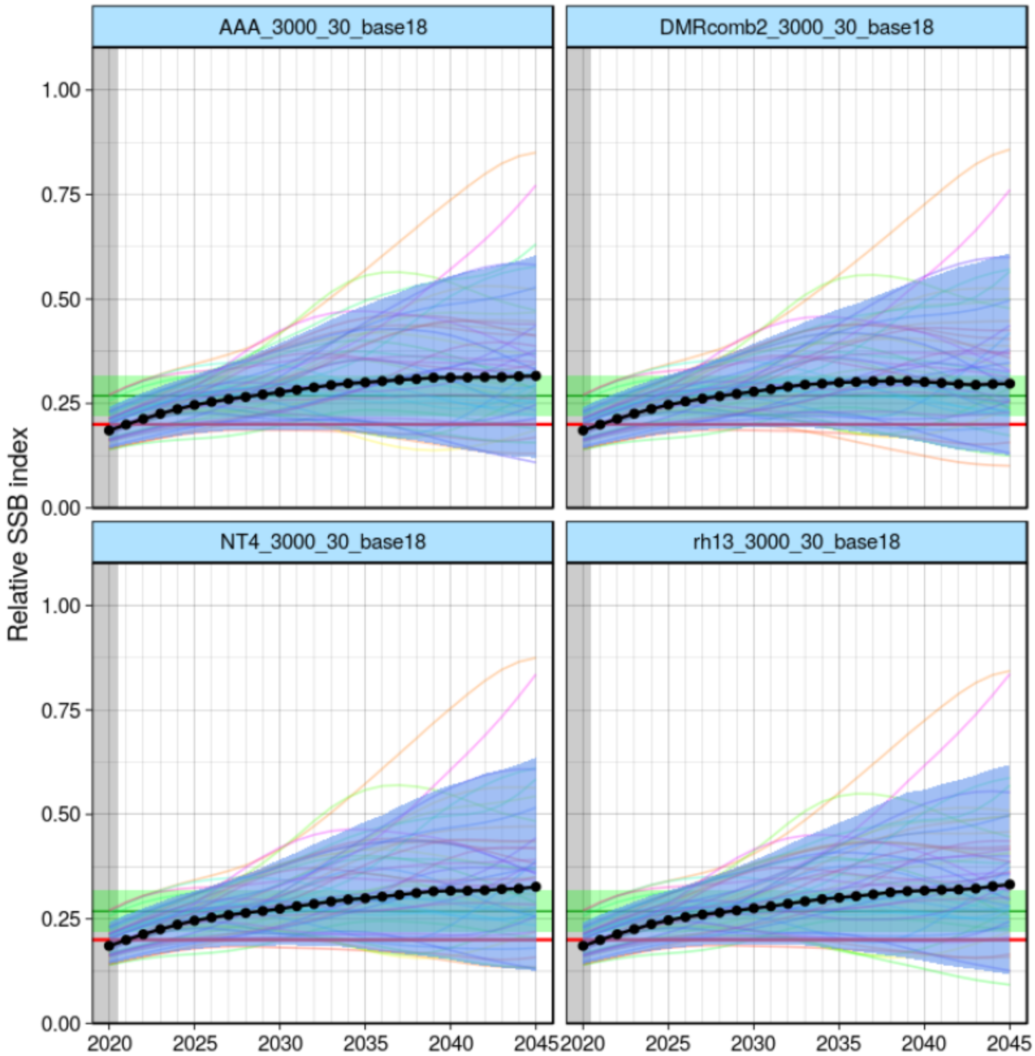

Figure 4.20: SSB trajectory plots the four final selected CMPs tuned to achieve a median SSB which is 30% of its pristine value by 2035 for base18 reference set. Medians and 90% probability envelopes are shown. The red horizontal lines here indicate 20% SSB_0. This reflects the interim rebuilding target for which the CCSBT specified that any CMP should achieve at least 70% probability of reaching this interim target by 2035.

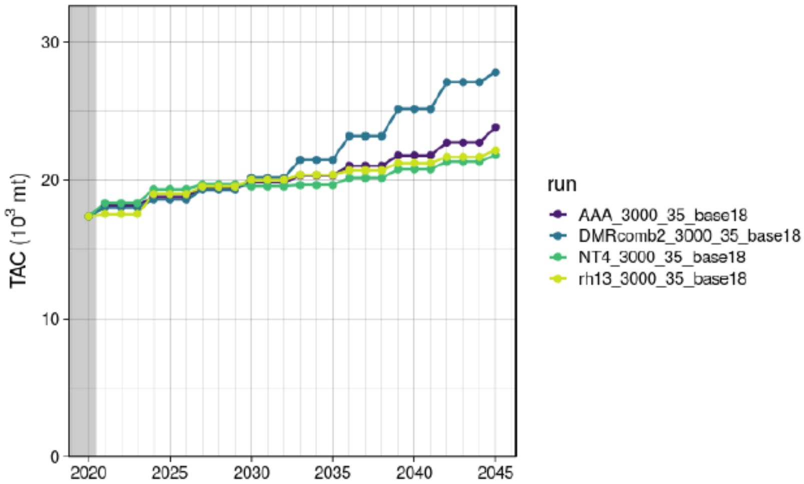

Figure 4.21: The median TAC plots for the four final selected CMPs tuned to achieve a median SSB which is 35% of its pristine value by 2040 under the base18 reference set.

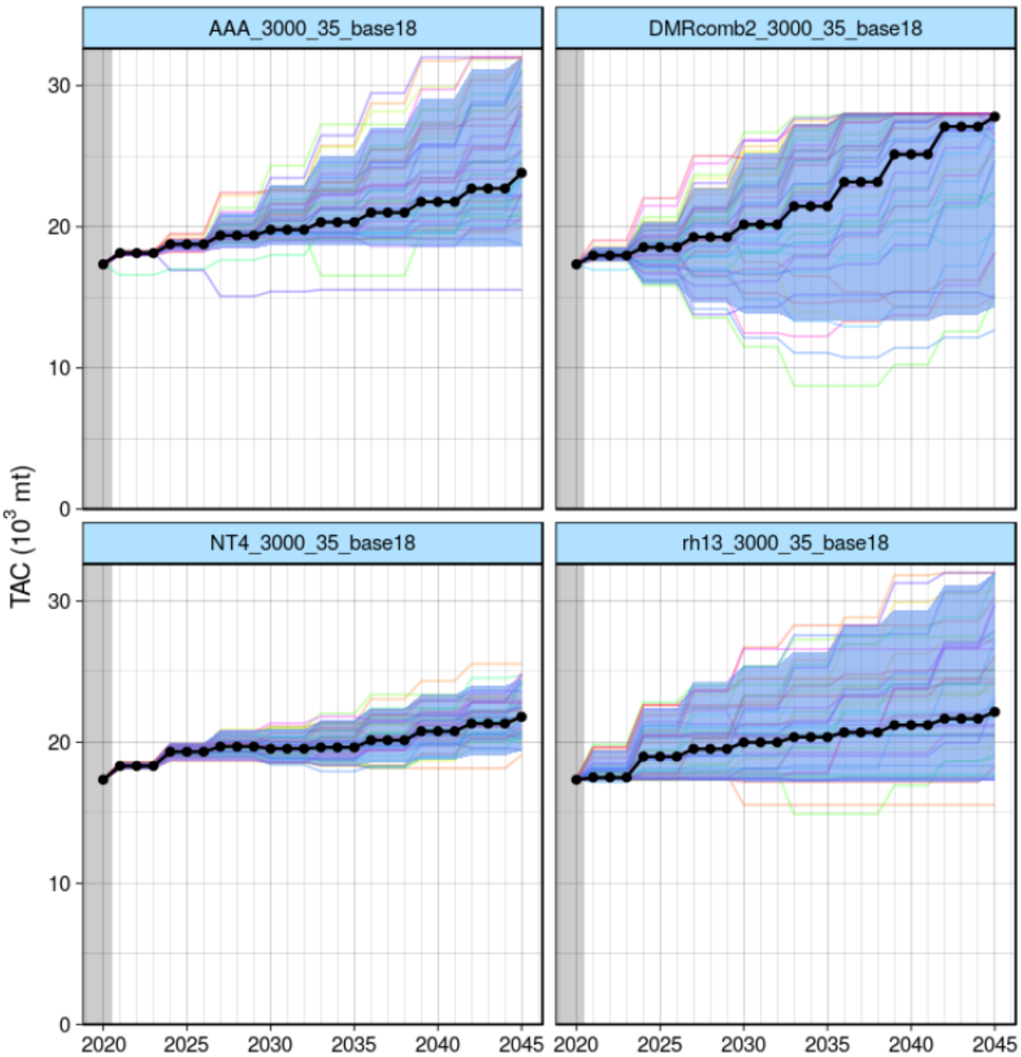

Figure 4.22: The TAC plots for the four final selected CMPs tuned to achieve a median SSB which is 35% of its pristine value by 2040 under the base18 reference set. Medians and 90% probability envelopes are shown.

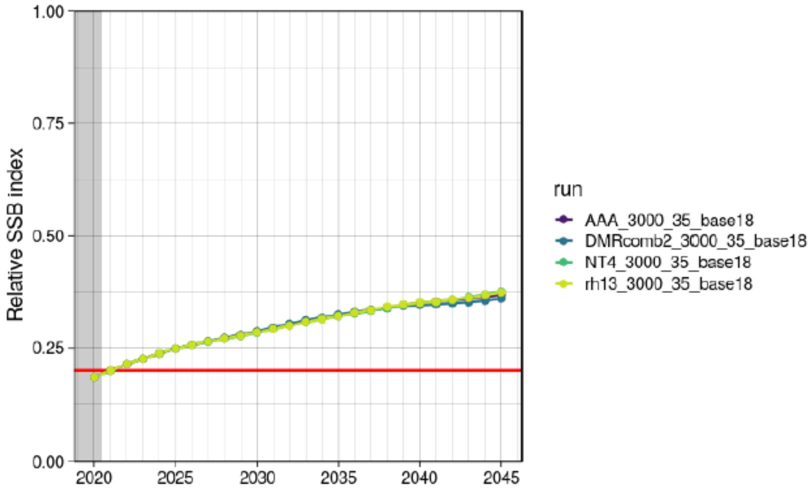

Figure 4.23: The median SSB trajectory plots for the four final selected CMPs tuned to achieve a median SSB which is 35% of its pristine value by 2040 under the base18 reference set.

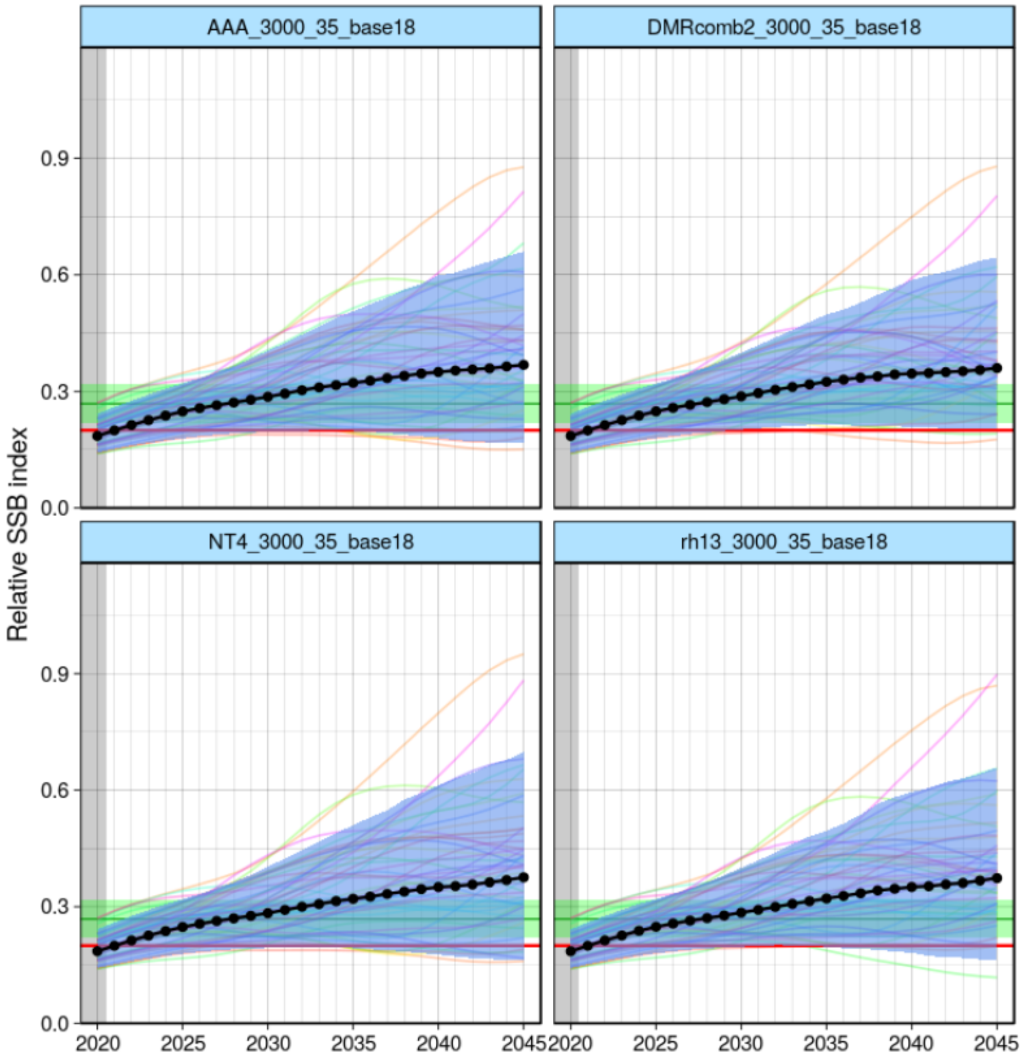

Figure 4.24: SSB trajectory plots the four final selected CMPs tuned to achieve a median SSB which is 35% of its pristine value by 2040 for base18 reference set. Medians and 90% probability envelopes are shown. The red horizontal lines here indicate 20% SSB_0. This reflects the interim rebuilding target for which the CCSBT specified that any CMP should achieve at least 70% probability of reaching this interim target by 2035.

4.2 Atlantic Bluefin Tuna

4.2.1 Results

Results are shown for CMPs FXP_1 and FXP_2 applied to the deterministic and stochastic versions of OM1 to OM15[2], where OM1 to OM12 are from the reference set of Operating Models (OMs) as shown in Table 3.5, and ROM13 to ROM15 are primary robustness tests which are listed in Table 3.6.

The "C=0" scenario represents an extreme CMP in which no catch is taken, denoted "0 catch" in Figures 4.25 and 4.26, by setting β and α to zero in the CMP equations 3.24 and 3.25. The reason that this scenario is simulated is to show the final stock levels when no catch is taken, which provides the "best possible outcome" in terms of stock recovery. However, the first three years' catches in the 30-year projection period are non-zero due to TACs already having been set, and hence being in place, for the initial years of this period. Only after those three years, when management under the CMP starts, are future catches set to zero for this scenario. Br0 is the depletion value relative to the dynamic biomass[3] corresponding to maximum sustainable yield (B_{MSY}) at the start of the 30-year projection period, which is shown to allow conclusions to be drawn as to whether the resource abundance increased or decreased following the CMPs being applied to the OMs. Br30 is the projected depletion value for stock concerned after 30 years, and AvC30 is the projected average annual catch for the West/East area over those 30 years.

Deterministic simulation runs refer to CMPs applied to deterministic OMs for which there is no observation error associated with the projected abundance indices, and where there is no variation of recruitment about the stock-recruitment relationship in the future. In contrast, stochastic simulation runs include variation in both these respects, as described in Section 3.3.3. "Perfect data" are provided to CMPs for deterministic runs, whereas more realistic "noisy data" are provided to these for stochastic runs. Deterministic results are considered before stochastic results, because if a CMP does not perform adequately under deterministic runs, it would hardly do so under stochastic runs, so that this allows for the speedy elimination of poor CMPs.

Table 4.4 shows the values of the parameters used in the final round of CMP testing applied to OM package v5.2.3: control parameters β and α which were selected based on tuning, parameters $TAC_{W,max}(mt)$ and $TAC_{E,max}(mt)$ which are used in the CMPs to make allowance for the possibility that a future regime shift occurs (see equation 3.30), and the parameter T that is used if the western stock abundance drops too low (see equation 3.31).

Table 4.5 shows the deterministic results for depletion and average catch over the 30-year projection period for FXP_1 and FXP_2 for OM1 to OM15. The results in this Table are shown graphically in Figure 4.25.

[2]ROM13, ROM14 and ROM15 are interchangeably referred to as OM13, OM14 and OM15.

[3]Dynamic biomass (B_0) refers to the biomass that would be present if no exploitation had taken place; note that this changes, including with transient effects, when a regime shift occurs.

Tables 4.6 and 4.7. show the stochastic results for the same situations as reported in Table 4.5 for deterministic runs. The results in Tables 4.6 and 4.7 are shown graphically in Figure 4.26, where the probability intervals are shown for FXP_2 only to avoid cluttering.

Figures 4.27 to 4.30 show the stochastic projections for West catch and SSB on the left, and for the East on the right, for OM5 and for each of the robustness tests ROM13 to ROM15. Projections are displayed for the "C=0" scenario, denoted "ZeroC" in the top panels, for FXP_1 in the middle panels denoted FXP_W1 and FXP_E1, and for FXP_2 in the bottom panels denoted FXP_W2 and FXP_E2.

4.2.2 Discussion

Deterministic results in Table 4.5, as well as the stochastic results in Tables 4.6 and 4.7, reflect higher Western and Eastern stock SSBs for the "C=0" scenarios than for FXP_1 and FXP_2, which is as to be expected given the impact made by the catches. These "C=0" scenarios reflect the upper bounds on the extent of stock recovery that is possible. Unless otherwise specified, comments made below apply to both to deterministic and stochastic results, and also to both Western and Eastern stocks.

The ROM14 primary robustness test dominated the selection of the control parameter values. The control parameters for FXP_1 were set to ensure no resource decline for that scenario, but that meant that most of the other reference set scenarios all reflected lower catches and underexploited resources. For FXP_2 the control parameter values were adjusted to get higher catches overall, while still ensuring that the Western stock was not rendered extinct for ROM14. Nevertheless, both stocks also generally remained underutilised (i.e. biomass remains well above B_{MSY}) for the reference set scenarios under FXP_2.

There is no consistent pattern when comparing the stock levels of Br0 and Br30 for the "C=0" scenario, even though one might (at first) expect that there must be some increase if no catch is taken. For many of the reference set OMs (OM1 to OM12) there is some decrease after 30 years, but this occurs mainly for the Recruitment 1 and Recruitment 3 scenarios where there has been a regime shift and transient effects are still playing a role.

For the A-group OMs that are representative of fish that mature at a young age and have high natural mortality, all the SSB levels for the Western stock are predicted to fall over the projection period, i.e., $Br30 < Br0$. However, for the B-group OMs for which fish mature later and have low natural mortality, half fall over the projection period ($Br30 < Br0$) while half increase ($Br30 > Br0$). Most of these reductions are not a problem, since the stock remains above its maximum sustainable yield level (B_{MSY}), which is denoted by the horizontal black lines in Figures 4.25 and 4.26 (which reflect $Br30 = 1$) at the end of the projection period. The cases where problems do arise due to the stock abundance dropping below B_{MSY} are discussed below.

4.2.2.1 Analysis of OM5 and the primary robustness tests ROM13 to ROM15 with regard to the CMP restrictions in Section 3.3.3

Stock depletion levels:

For ROM13 after 30 years, the Eastern stock is above B_{MSY} for FXP_1 and FXP_2, but the Western stock is essentially at B_{MSY} for FXP_1 and below for FXP_2. More positively, the Eastern and Western stocks for ROM15 are above B_{MSY} after 30 years for FXP_1 and FXP_2. This behaviour can be seen clearly in Figures 4.28 and 4.30, and more broadly in Figures 4.25 and 4.26.

ROM14 (2wBI) is a variation of OM5 (2BI), which incorporates the Recruitment 2 scenario, fish that mature at an older age with a lower natural mortality and a lower level of mixing, as shown in Table 3.4 but importantly with a lower initial Western stock abundance for ROM14 compared to OM5. Figure 4.27 shows that for OM5 under FXP_1 (denoted FXP_W1) the Western stock abundance varies only slightly while remaining a little above B_{MSY}, whereas the Eastern stock starts much further above B_{MSY} and decreases gradually over time. Under FXP_2 (denoted FXP_W2), the Western stock drops to below B_{MSY} by the end of the 30-year projection period, whereas the Eastern stock decreases gradually but remains above B_{MSY} at the end of this period. Given that this is the case for OM5, worse projection behaviour is to be expected for the ROM14 variant, and this is indeed the case especially for the Western stock as shown in Figure 4.29. The Western stock Br30 results for ROM14 are appreciably less than those for the other robustness tests as well as for OM1 to OM12. This can be seen in the deterministic and stochastic results in Tables 4.5 and 4.6, and graphically in Figures 4.25 and 4.26, where ROM14 in particular displays outlier behaviour compared to the other OMs for the Western stock Br30 values.

Unfortunately, the Western stock under FXP_1 and FXP_2 (denoted FXP_W1 and FXP_W2) is well below B_{MSY} at the end of the projection period, and for FXP_2 is concerningly near extinction. The Eastern stock for FXP_1 and FXP_2 (denoted FXP_E1 and FXP_E2) shows a steady decrease in abundance, but remains above B_{MSY} by the end of the 30-year projection period. Furthermore, the abundance for FXP_E2 is below B_{MSY} at the end of the projection as shown in Figure 4.29, which extends for 50 rather than for only 30 years. As for OM5, the catches for ROM14 generally decrease throughout the projection period, but not sufficiently to counter the risk of serious stock depletion.

Catch:

The TAC restrictions of equation 3.30 were introduced to avoid the West and East area TACs getting too large to be able to reduce them fast enough later to avoid undesirable depletion in the event of a future shift to a lower productivity regime. When these restrictions come into play, they result in an overall decrease in the TACs which would have applied in the absence of these restrictions, and can also lead to periods where TACs are fixed at their respective maximum levels for the West and the East. This behaviour is evident for some projections in Figures 4.27 to 4.30. In median terms for

ROM13 and ROM15, there is an initial increase in catch, followed by a constant catch at 4000 mt for West areas and 30 000 mt for East areas, except for some fluctuations where catch initially decreases and then increases for FXP_1.

The development stage of these CMPs emphasized addressing concerns for the conservation of the Western stock; consequently, starting from the third round of CMP testing in Section 3.3.3, the need for further TAC restrictions in the West became evident as soon as poor Western stock depletion levels became apparent for certain OMs. Low Western stock abundance persisted in the final round of CMP testing, so that yet further restrictions were placed on the TACs resulting from equation 3.22 and equation 3.24, as described by equations 3.31 to 3.33. The parameter T in equation 3.31 was introduced to try to be more effective in avoiding depletion of the Western stock, especially under the primary robustness tests and particularly under ROM14. This can be seen in Figures 4.27 and 4.29 where catches in the West and East initially increase but then decrease for most simulated trajectories for OM5 and ROM14.

These results are as expected given modifications 1 and 2 that are explained in Section 3.3.3 (equations 3.30 to 3.33), leading to the TAC decreasing if there is a regime shift and/or Western or Eastern stock abundance drops too low.

Table 4.4: The parameters used in the final round of CMP testing applied to OM package v5.2.3.

CMPs	β	α	$TAC_{W,max}(mt)$	$TAC_{E,max}(mt)$	T
FXP_1	0.5	0.5	4000	30 000	1.0
FXP_2	1.0	1.0	4000	30 000	1.0

Table 4.5: Deterministic results for depletion (Br30) and average catch (AvC30) over a 30-year projection period for two CMPs FXP_1(β=0.5, α=0.5) and FXP_2(β=1.0, α=1.0) for OM1 to OM15. Br0 shows the depletion at the start of the projection period.

				West							East						
				Br0	**Br30**			**AvC30**			**Br0**	**Br30**			**AvC30**		
					C=0[4]	FXP_1	FXP_2	C=0	FXP_1	FXP_2		C=0	FXP_1	FXP_2	C=0	FXP_1	FXP_2
A-group	OM	1	1AI	2.343	2.788	2.160	1.901	0.223	2.867	3.624	1.947	2.963	2.501	2.092	2.652	19.548	29.546
		7	1AII	3.099	2.942	2.576	2.402	0.223	2.796	3.621	1.926	2.959	2.490	2.083	2.652	19.675	29.546
		2	2AI	2.344	2.311	1.963	1.733	0.223	2.610	3.616	2.100	2.349	2.017	1.628	2.652	17.662	27.829
		8	2AII	2.811	2.452	2.268	2.135	0.223	2.551	3.614	2.226	2.388	2.088	1.725	2.652	17.271	27.366
		3	3AI	2.343	2.820	2.306	2.091	0.223	3.078	3.625	1.947	2.541	1.811	1.210	2.652	19.603	28.674
		9	3AII	3.099	2.954	2.593	2.451	0.223	3.078	3.622	1.926	2.557	1.857	1.243	2.652	19.086	28.381
B-group		4	1BI	1.342	2.794	1.722	1.439	0.223	3.078	3.631	1.587	2.752	2.259	1.948	2.652	21.462	29.546
		10	1BII	2.188	3.142	2.396	2.176	0.223	3.075	3.629	1.644	2.703	2.175	1.859	2.652	21.745	29.546
		5	2BI	1.356	1.850	1.227	0.852	0.223	2.615	3.579	2.768	2.245	1.937	1.560	2.652	17.612	27.760
		11	2BII	2.921	2.637	2.451	2.318	0.223	2.548	3.600	3.534	2.351	2.093	1.761	2.652	17.271	28.636
		6	3BI	1.343	2.966	2.216	1.936	0.223	3.078	3.632	1.587	2.341	1.693	1.306	2.652	21.868	29.546
		12	3BII	2.188	3.217	2.639	2.424	0.223	3.078	3.629	1.644	2.262	1.597	1.166	2.652	21.441	29.494
ROM		13	1wBI		2.358	1.106	0.820	0.223	3.111	3.638		2.768	2.263	1.956	2.652	21.718	29.546
		14	2wBI		1.077	0.410	0.149	0.223	2.534	3.204		2.230	1.910	1.527	2.652	17.730	27.640
		15	3wBI		2.592	1.720	1.427	0.223	3.111	3.638		2.360	1.675	1.326	2.652	22.661	29.546

[4] The reason the average catch value is not zero for the zero catch scenario ("C=0") is because for the first three years of the projection period, catches are always set to the TACs already in place, resulting in non-zero averages over the whole period even though projected catches are zero.

The OMs and ROMs are described in section 3.3.2 but this is as a summary of the factors and levels that formulate each OM in the Tables and Figures:
Levels 1/2/3: Reflect different Beverton-Hold stock recruitment relationships and regime shift scenarios.
Level A/B: Reflect the assumption that fish that mature young and have a higher natural mortality, and fish that mature older have a lower natural mortality.
Levels I/II: Reflect the extent of mixing, where level II represent a higher level of mixing than level I.
Western Contrast. Increased precision (CV of 15%) of the GOM larval survey index to create scenarios with lower current western stock status as indicated in Table 3.6.

Table 4.6: Stochastic results for depletion (Br30) and average catch (AvC30) for Western stock and West area, respectively, over a 30-year projection period for two CMPs FXP_1($\beta = 0.5, \alpha = 0.5$) and FXP_2($\beta = 1, \alpha = 1$) for OM1 to OM15. The results shown are medians with lower 5%-ile and upper 95%-ile values in parenthesis.

				West					
				Br0	**Br30**			**AvC30**[5]	
					C=0	FXP_1	FXP_2	FXP_1	FXP_2
A-group	OM	1	1AI	2.343	2.706(1.821, 3.690)	2.074(1.490, 3.154)	1.828(1.199, 2.755)	2.610(1.405, 3109)	3.476(2.134, 3.638)
		7	1AII	3.099	2.846(1.782, 4.004)	2.497(1.574, 3.718)	2.281(1.404, 3.423)	2.389(1.294, 2.914)	3.442(1.978, 3.638)
		2	2AI	2.344	2.375(1.360, 3.517)	2.051(1.146, 3.143)	1.824(1.036, 2.908)	2.346(1.469, 2.910)	3.443(2.411, 3.638)
		8	2AII	2.811	2.547(1.457, 3.698)	2.345(1.335, 3.487)	2.146(1.259, 3.295)	2.286(1.390, 2.789)	3.405(2.253, 3.638)
		3	3AI	2.343	2.765(2.209, 3.641)	2.232(1.702, 3.114)	1.996(1.461, 2.841)	3.078(2.731, 3.295)	3.556(3.323, 3.638)
		9	3AII	3.099	2.868(2.218, 3.894)	2.470(1.858, 3.380)	2.334(1.693, 3.235)	3.068(2.435, 3.242)	3.518(3.214, 3.638)
B-group		4	1BI	1.342	2.729(2.173, 3.535)	1.737(1.336, 2.516)	1.375(1.000, 2.005)	2.906(1.723, 3.236)	3.527(2.607, 3.638)
		10	1BII	2.188	3.077(2.270, 4.173)	2.415(1.745, 3.336)	2.16(1.484, 2.929)	2.750(1.512, 3.179)	3.509(2.321, 3.638)
		5	2BI	1.356	1.865(1.210, 2.826)	1.368(0.773, 2.186)	0.962(0.485, 1.613)	2.077(1.477, 2.546)	3.240(2.287, 3.622)
		11	2BII	2.921	2.649(1.849, 3.863)	2.483(1.699, 3.665)	2.328(1.560, 3.444)	2.097(1.280, 2.519)	3.327(2.115, 3.638)
		6	3BI	1.342	2.938(2.369, 3.947)	2.168(1.649, 3.085)	1.923(1.373, 2.802)	3.078(2.752, 3.297)	3.557(3.274, 3.638)
		12	3BII	2.188	3.175(2.421, 4.282)	2.571(1.888, 3.583)	2.417(1.680, 3.354)	3.078(2.635, 3.303)	3.565(3.169, 3.638)
ROM		13	1wBI		2.319(1.834,2.996)	1.053(0.707, 1.653)	0.749(0.452, 1.249)	3.077(2.141, 3.329)	3.570(2.771, 3.638)
		14	2wBI		1.068(0.577, 1.851)	0.502(0.221, 1.082)	0.187(0.053, 0.595)	2.072(1.519, 2.603)	3.023(2.068, 3.562)
		15	3wBI		2.530(2.020, 3.336)	1.625(1.167, 2.451)	1.366(0.918, 2.070)	3.078(2.752, 3.345)	3.597(3.285, 3.638)

[5] There are no average catch (AvC30) values for the stochastic simulations when catch is equal to zero "C=0" because there are only data available for the first three years; therefore there is no noise and the AvC30 "C=0" here is the same as for the deterministic West AvC30 "C=0" values in Table 4.5.

Table 4.7: Stochastic results for depletion (Br30) and average catch (AvC30) for Eastern stock and East area, respectively, over a 30-year projection period for two CMPs FXP_1($\beta = 0.5, \alpha = 0.5$) and FXP_2($\beta = 1, \alpha = 1$) for OM1 to OM15. The results shown are medians with lower 5%-ile and upper 95%-ile values in parenthesis.

				East					
				Br0	**Br30**			**AvC30**[6]	
					C=0	FXP_1	FXP_2	FXP_1	FXP_2
A-group	OM	1	1AI	1.947	2.916(2.260, 3.994)	2.385(1.738, 3.306)	2.012(1.403, 2.946)	19.694 (15.82, 24.378)	29.054(24.692,29.546)
		7	1AII	1.926	2.904(2.085, 3.959)	2.421(1.666, 3.558)	2.071(1.319, 3.148)	20.4 (16.013, 24.403)	29.328(25.596,29.546)
		2	2AI	2.100	2.100(1.316, 3.291)	1.797(1.003, 2.809)	1.524(0.660, 2.606)	19.629 (15.853, 24.949)	28.625(24.615,29.546)
		8	2AII	2.226	2.098(1.235, 3.377)	1.810(0.860, 2.891)	1.39(0.264, 2.595)	18.754 (15.039,24.317)	27.884(22.899,29.546)
		3	3AI	1.947	2.385(1.688, 3.317)	1.577(1.021, 2.326)	1.072(0.473, 1.731)	19.586 (16.172, 24.593)	28.438(25.292, 29.546)
		9	3AII	1.926	2.477(1.761, 3.515)	1.706(1.147, 2.575)	1.141(0.618, 2.037)	20.037 (15.962, 24.055)	28.336(25.269, 29.546)
B-group		4	1BI	1.587	2.724(1.986, 3.594)	2.257(1.632, 3.105)	1.958(1.322, 2.808)	21.819 (17.011, 25.857)	29.546(27.200, 29.546)
		10	1BII	1.644	2.652(1.922, 3.599)	2.145(1.474, 3.016)	1.836(1.146, 2.752)	21.715 (16.668, 26.002)	29.544(26.408, 29.546)
		5	2BI	2.768	2.145(1.255, 3.179)	1.824(0.796, 2.792)	1.524(0.477, 2.535)	19.828 (15.723, 26.496)	29.087(24.491, 29.546)
		11	2BII	3.534	2.196(0.957, 3.244)	1.905(0.758, 2.91)	1.306(0.581, 2.616)	19.126 (15.175, 25.684)	28.481(23.512, 29.546)
		6	3BI	1.587	2.314(1.664, 3.214)	1.598(1.081, 2.428)	1.323(0.711, 2.146)	22.622 (18.021, 26.364)	29.512(28.247, 29.546)
		12	3BII	1.644	2.169(1.558, 3.116)	1.518(0.925, 2.369)	1.146(0.582, 1.992)	21.654 (17.167, 26.299)	29.388(27.201, 29.546)
ROM		13	1wBI		2.784(2.117, 3.518)	2.256(1.679, 3.038)	1.963(1.379, 2.679)	21.817 (17.022, 26.088)	29.546(26.845, 29.546)
		14	2wBI		2.130(1.267, 3.157)	1.787(0.697, 2.730)	1.438(0.478, 2.440)	20.157 (15.857, 26.797)	28.945(24.589, 29.546)
		15	3wBI		2.270(1.720, 3.100)	1.578(1.064, 2.283)	1.278(0.785, 2.03)	23.796 (19.519, 26.839)	29.546(28.412, 29.546)

[6]There are no average catch (AvC30) values for the stochastic simulations when catch is equal to zero "C=0" because there are only data available for the first three years; therefore there is no noise and the AvC30 "C=0" here is the same as for the deterministic East AvC30 "C=0" values in Table 4.5.

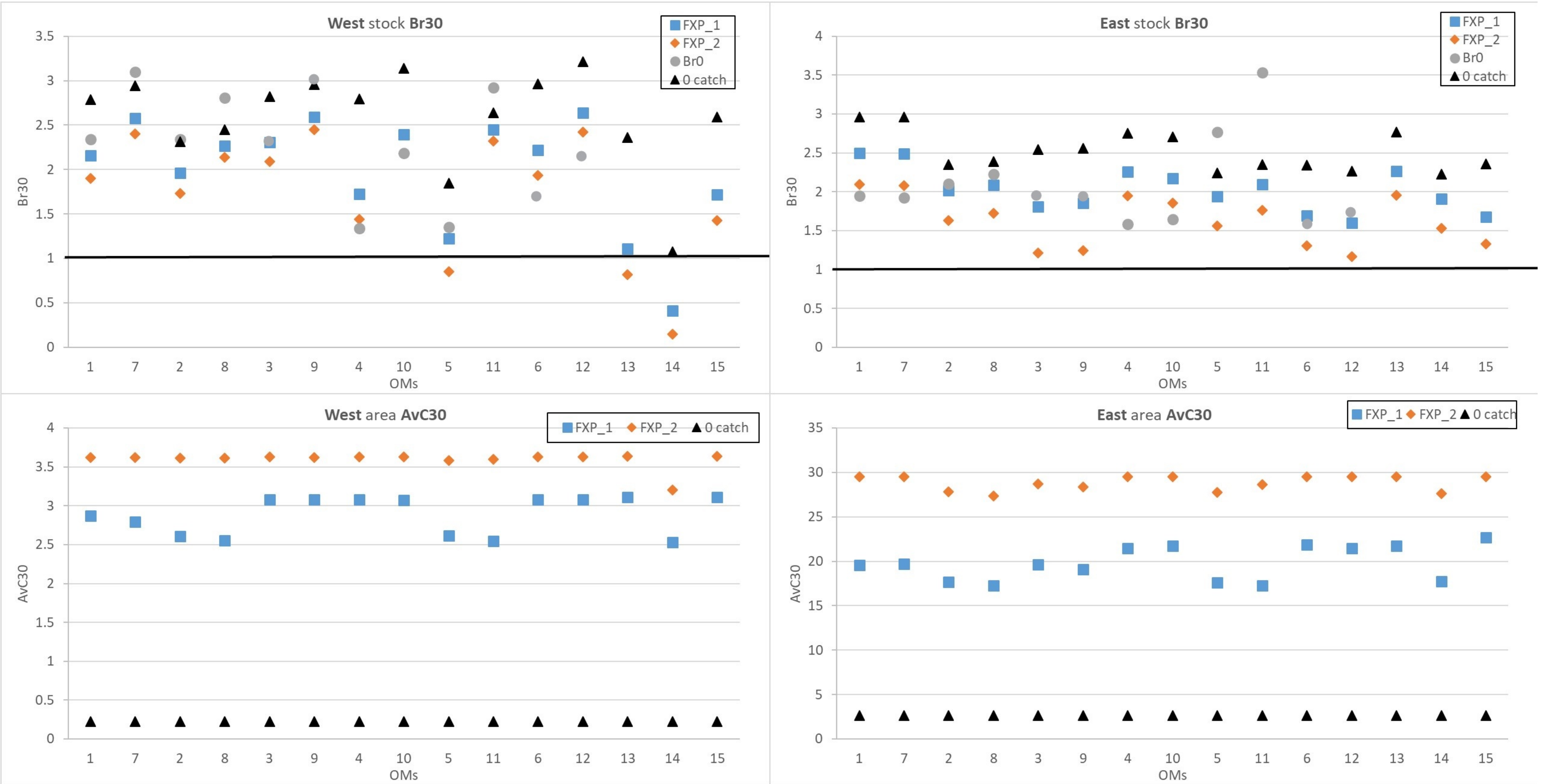

Figure 4.25: Deterministic results for depletion (Br30) and average catch (AvC30) for the "0 catch" scenario and the two CMPs FXP_1($\beta = 0.5, = 0.5$) and FXP_2($\beta = 1, \alpha = 1$) for OM1 to OM15.

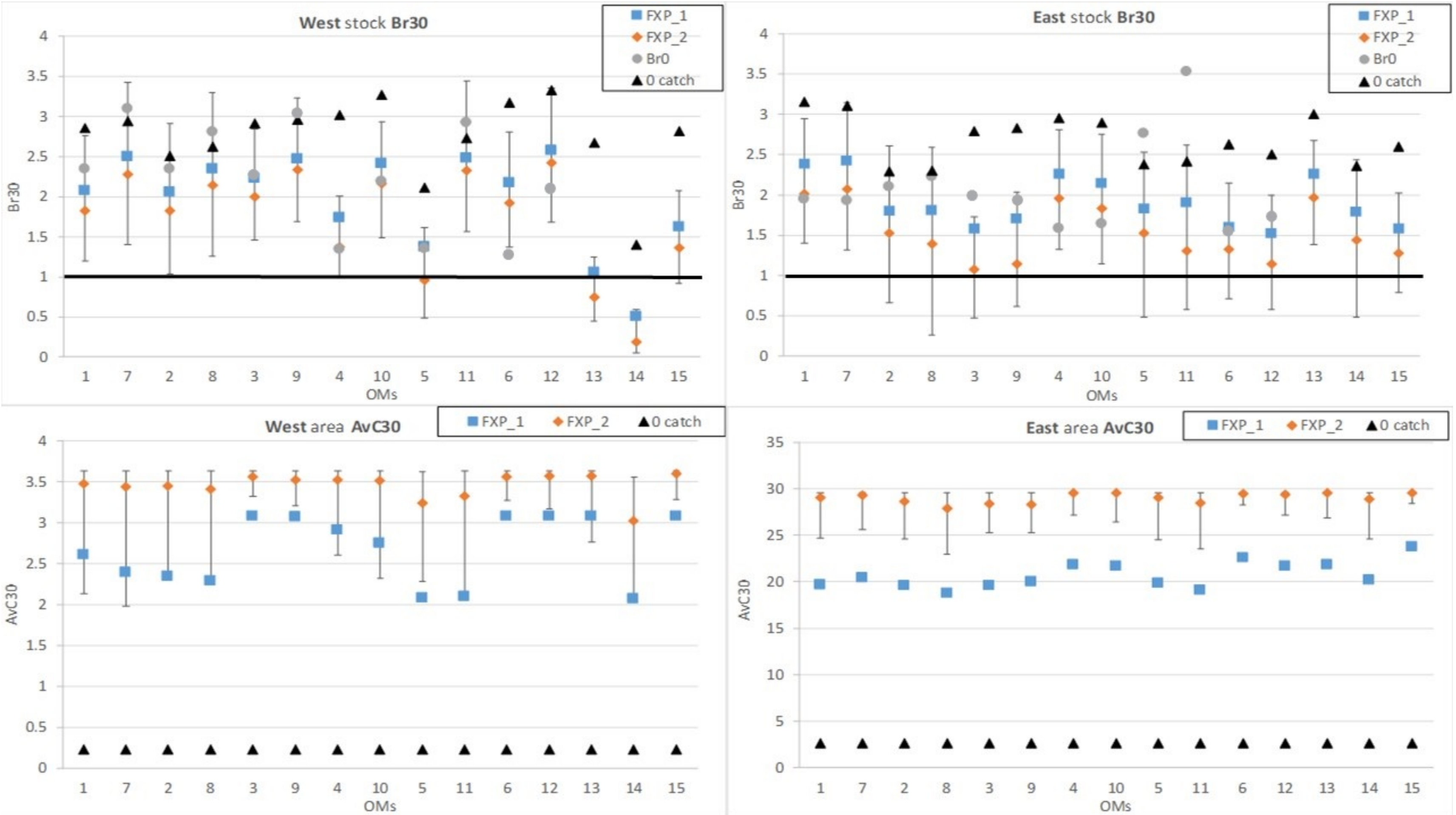

Figure 4.26: Stochastic results for depletion (Br30) and average catch (AvC30) for the "0 catch" scenario and the two CMPs FXP_1($\beta = 0.5, = 0.5$) and FXP_2($\beta = 1, \alpha = 1$) for OM1 to OM15. Medians are shown except that for FXP_2 the 95% PIs are also shown.

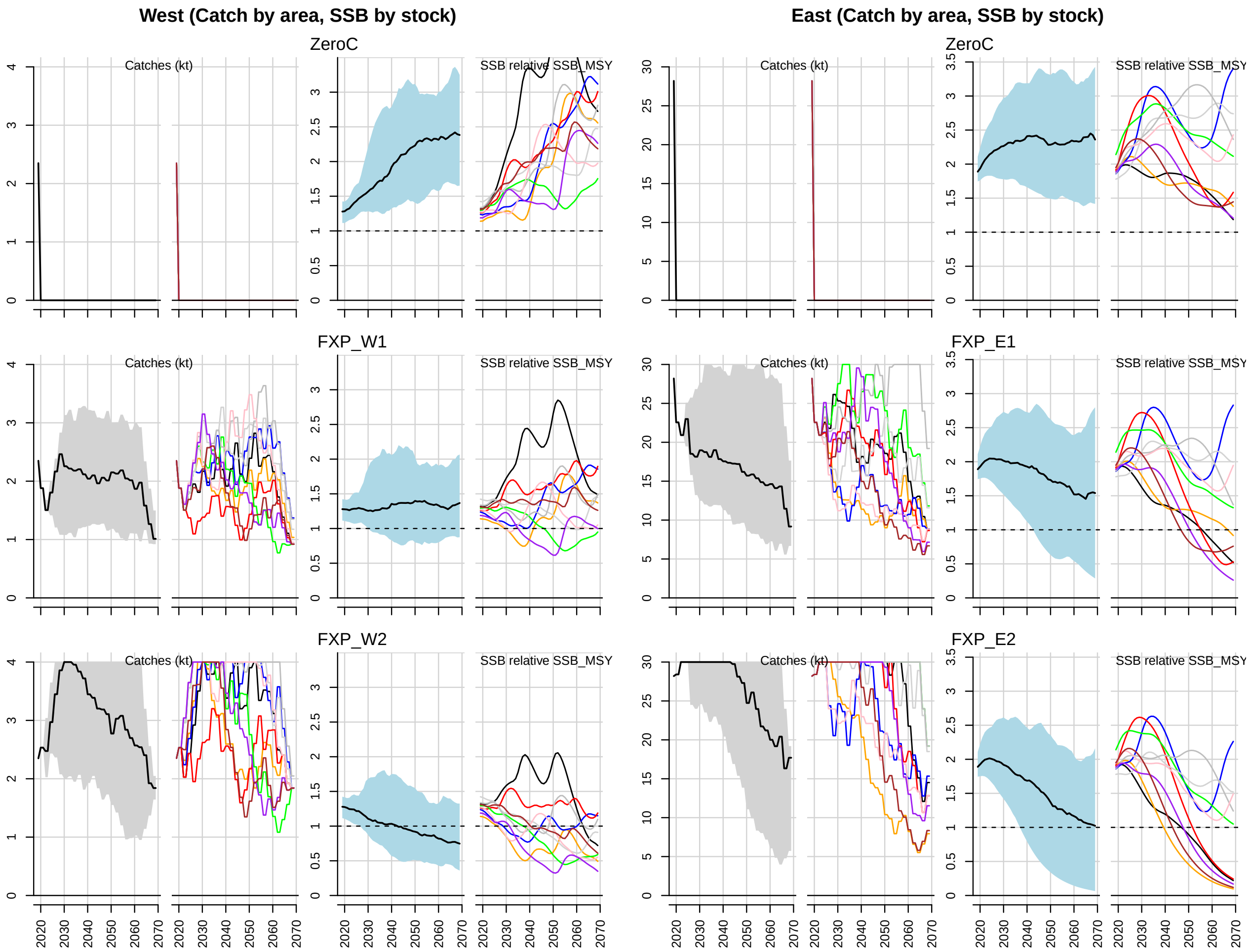

Figure 4.27: "Worm" plot projections for stochastic catch and SSB that show individual trajectories as well as the 95% probability envelopes (grey and blue shading) for OM5 for the "C=0" scenario, and for FXP_1 and FXP_2.

Features of T.Carruthers' design of the ABTMSE package to be noted for Figures 4.27 to 4.30:

1. The side by side projections is a feature to allow for comparison of two stochastic trajectory realisations.
2. The projections display until the year 2070, but the 30-year projection period for which results are given in the Tables ends in the year 2050.

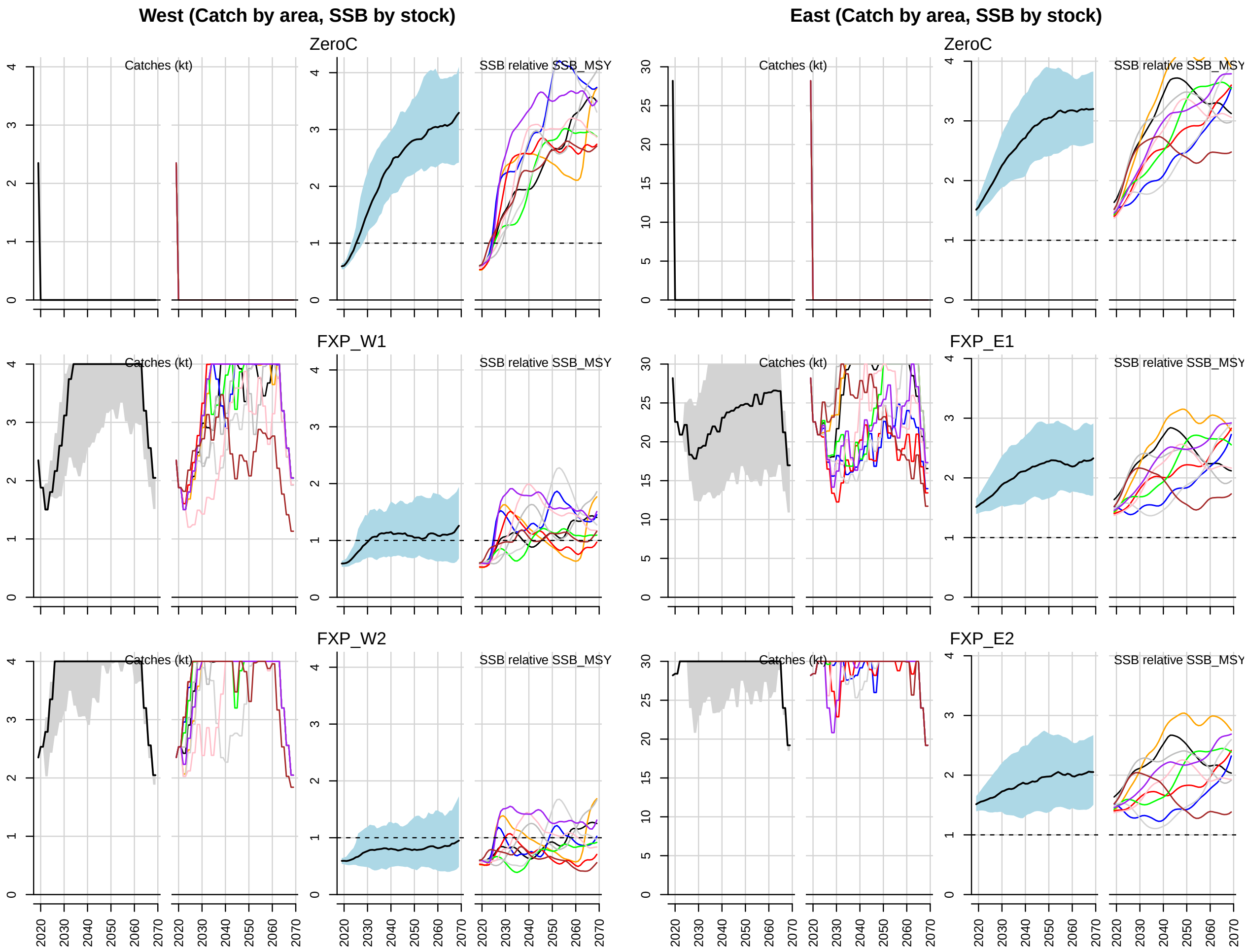

Figure 4.28: "Worm" plot projections for stochastic catch and SSB that show individual trajectories as well as the 95% probability envelopes (grey and blue shading) for ROM13 for the "C=0" scenario, and for FXP_1 and FXP_2.

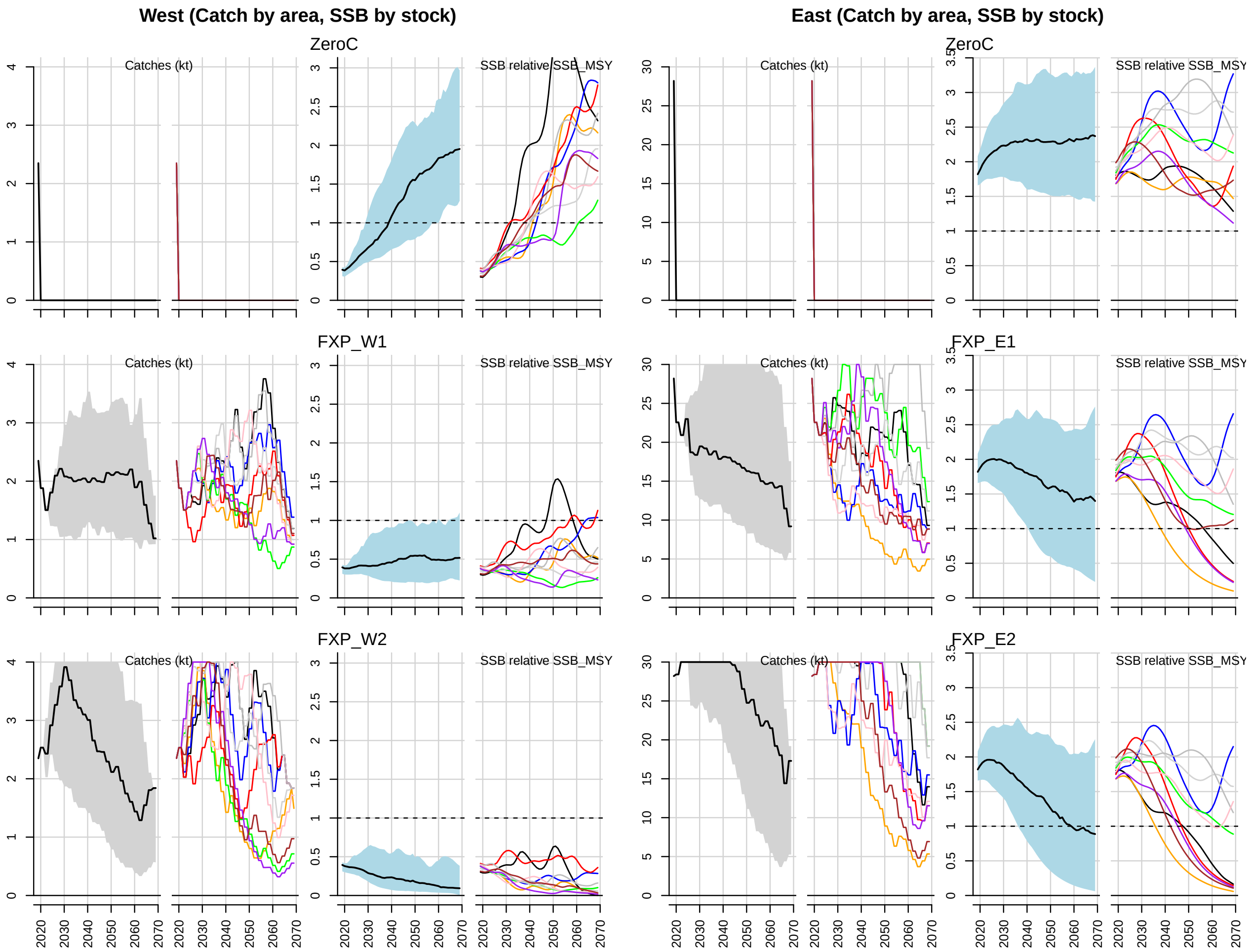

Figure 4.29: "Worm" plot projections for stochastic catch and SSB that show individual trajectories as well as the 95% probability envelopes (grey and blue shading) for ROM14 for the "C=0" scenario, and for FXP_1 and FXP_2.

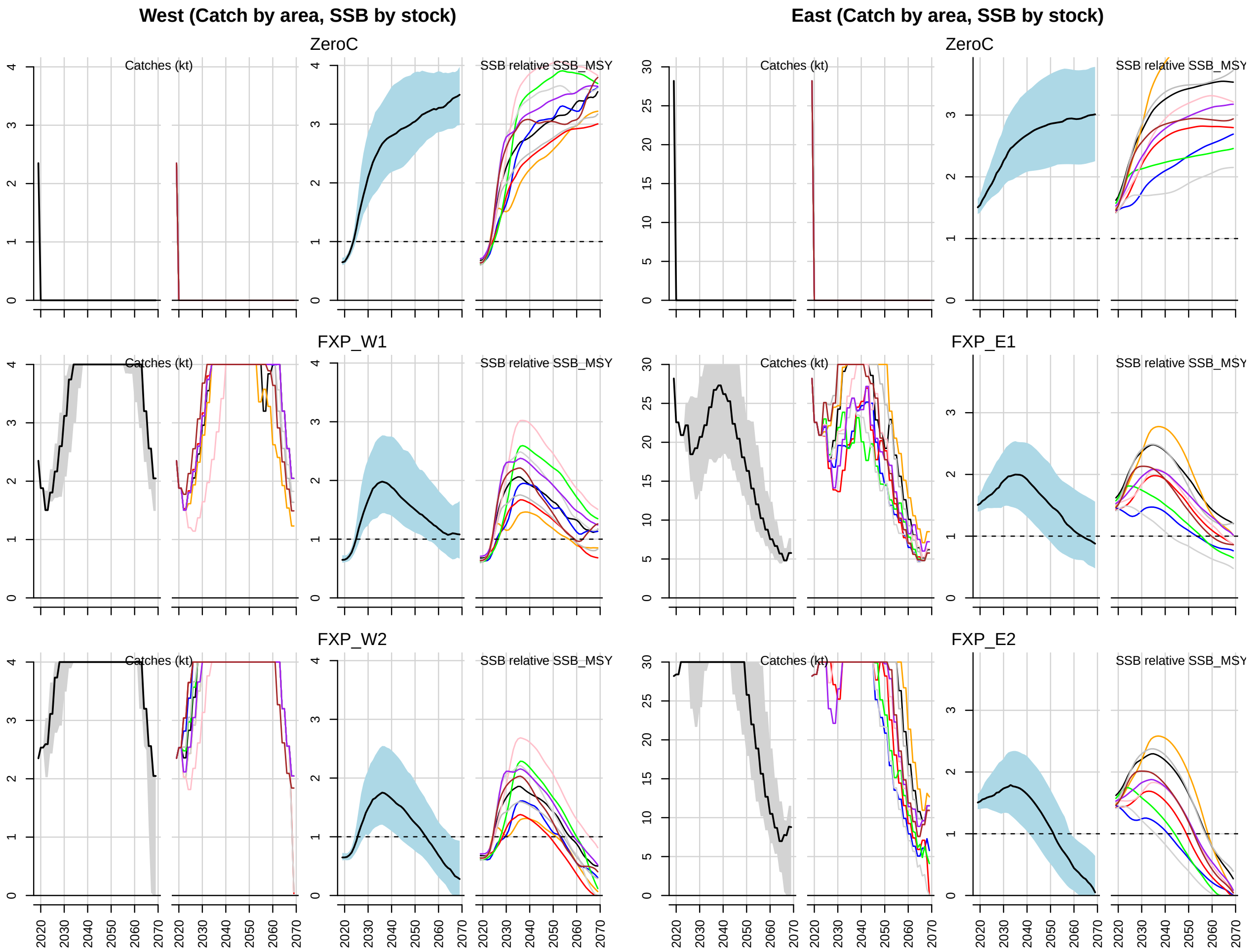

Figure 4.30: "Worm" plot projections for stochastic catch and SSB that show individual trajectories as well as the 95% probability envelopes (grey and blue shading) for ROM15 for the "C=0" scenario, and for FXP_1 and FXP_2.

Chapter 5

Conclusions and Future Work

5.1 Southern Bluefin Tuna

The development, simulation testing and evaluation of five CMPs was conducted with the intention to select a final CMP to be considered internationally, together with CMPs developed by other groups, for the implementation of a revised management procedure approach to set catch limits for Southern Bluefin Tuna. These five CMPs are DMRMCPUE, DMRMGT, DMRMCKMR, DMRcomb1 and DMRcomb2, where the first three each use only the CPUE, GT and CKMR indices of abundances respectively and the other two are weighted combinations of the first three. These CMPs were tuned to achieve two targets recovery levels specified by the Commission: 30% of pristine SSB by 2035 and 35% by 2040 under a base18 reference set of OMs. Special attention was paid to performance for two of the robustness tests which had a large impact on results: CPUE variable squares (cpuew0) and a future period of low recruitment (reclow5).

Considering first the CMPs for each input data type separately proved advantageous, as this showed the different trade-offs in performance statistics between these three. This assisted in the decision on how to weight the three in the combination CMP. DMRMGT showed the best performance in averting unintended low depletion (especially for the reclow5-related robustness tests), it was accorded the highest weight (of 60%) in the CMP (DMRcomb2) eventually put forward for consideration by the CCSBT ESC.

Compared to the other three CMPs submitted to the CCSBT ESC meeting, the main strength of the DMRcomb2 CMP was less depletion in circumstances of a period of low recruitment in the future, However, its main comparative weakness was increasing TACs too fast during the period to the first tuning target year of 2035. This led (in median terms) to little further TAC increase, and possible slight reduction in SSB, after 2035. Furthermore, for the 2040 tuning year, there was a comparatively high probability of the TAC dropping to levels which would have reflected economic risk for the fishery. These weaknesses contributed to the CCSBT SC eventually preferring the RH13 CMP to recommend for implementation.

The underlying reason for these weaknesses was that the DMR CMPs were deliberately kept as simple as possible, with TACs changing linearly with the abundance index in question (see equation 3.6 and following). This made it unavoidable that the TAC had to be fairly large at the time (2035) when the abundance recovery target had to be met. The next CCSBT MP review will take place in 6-9 years' time. If DMR-like CMPs are considered then, possibly the most important improvement needed would be some relaxation of this simple linearity to be able to address these weaknesses.

5.2 Atlantic Bluefin Tuna

The results reported in Chapter 4 for ABFT show that the conservative approach taken to avoid Western stock depletion to obtain acceptable stock risk performance for every OM considered leads to the under-exploitation of resources for most of the OMs because the TACs for those are much lower than could be taken sustainably. This is similar to, but more severe than, the situation for the final DMRcomb2 selected for SBT, for which there is a trade-off between low risk of stock depletion but poor fishery performance.

In particular, the results show that the tuning and selection of the control parameters for the CMPs reflect the domination of a primary robustness test, ROM14, that reflects a lower Western stock abundance. Since the work presented in this thesis corresponds to an intermediate stage of CMP development, it is clear that future work will depend on the plausibility to be accorded in discussions in ICCAT scientific meetings to this robustness test, in other words a decision on whether this is a sufficiently realistic possibility to retain. If it is considered to be realistic in those discussions, then this low abundance must lead to consequent low catches, i.e. very conservative management is necessary. On the other hand, if it is not considered to be realistic, then the CMPs developed here would need to be re-tuned to allow higher catches, even if the risk of high stock depletion would persist under the then discarded ROM14 test.

Deterministic results for robustness tests ROM_1 to ROM_30 are shown in Appendix C1 and serve as guidance for further work. Table C3 shows such results for depletion (in terms of Br30) and average catch over the 30-year projection period for FXP_1 and FXP_2 for these tests, which are also shown graphically in Figure C1. These results are suggestive of the necessity for further evaluation of the plausibility of the robustness tests, which has since been in progress. The results reported in this thesis are as pertained in September 2019 using the most recent OM package (version 5.2.3) available at that time. However, further discussions in ICCAT scientific meetings since that time have resulted in important changes made in developing an updated interim reference set (grid) of OMs. These are are detailed in the Report of the February 2020 Intersessional meeting of the ICCAT Bluefin tuna MSE Technical Group (ICCAT, 2020b).

Mechanisms are also required to allow for better utilisation of the Eastern stock, considering that that this resource is under-exploited for most of the OMs under the FXP_1 and FXP_2 CMPs; however, it is also necessary to be able to reduce the TAC fast enough if this stock drops too low in abundance. Three suggestions are made below for the development of improved CMPs to be pursued once the OM updates are in place:

1. Testing a wider range of control parameters to attain better overall CMP performance, with the intention of attaining higher catches particularly for the East area, but still with a low risk of undue stock depletion.

2. It is evident from the results in Chapter 4 that projections showing low Western stock depletion are a persisting problem. Even though CMP adjustments were put in place in an attempt to deal with this, they did not provide sufficient reduction in risk. Furthermore, depletion may also become a problem for the Eastern stock under heavier exploitation levels, because under a future regime shift (corresponding to Recruitment level 3) abundance may drop too low if TACs are not reduced sufficiently rapidly. A solution to this may be to incorporate a parameter that adjusts the TAC for the East area by modifying the formula for the TAC to reflect parabolic behaviour at low abundance, similar to the approach for the West area detailed in equation 3.31 and the corresponding Figure 3.4. The aim of this adjustment is to prevent the Eastern stock dropping too low by setting the TAC lower than previously for low abundance levels.

3. Allow more flexibility for the limit on the maximum extent of TAC decrease allowed, compared to the current specification that imposes a 20% minimum/maximum change. In addition to this, the maximum TACs of 4000 mt and 30 000 mt set for the West and East areas respectively might be reduced; by not allowing the TAC to increase too high, it becomes easier to subsequently reduce it sufficiently to offset the risk to the stock if abundance becomes low (especially following a shift to a lower productivity regime).

Appendices

Appendix A

Table A.1: Glossary of terms for harvest strategies, management procedures and management strategy evaluation (Joint Tuna RFMO Management Glossary, 2018). These originate, in the main, from a glossary developed by the 2018 Joint Tuna RFMO Management Strategy Evaluation Working Group Meeting, with some embellishments and a few additions.

Term	Definition	Abbreviation/Symbol
Average Annual Variation (in catch/TAC)	The absolute value of the proportional TAC change each year, averaged over the projection period.	AAV
Biomass	Stock biomass, which may refer to various components of the stock. Often spawning stock biomass (SSB) of females is used, as the greatest conservation concern is to maintain the reproductive component of the resource.	B
Candidate Management Procedure	An MP (defined below) that has been proposed and is under evaluation, but not yet adopted.	CMP
Conditioning	The process of fitting an Operating Model (OM) of the resource dynamics to the available data on the basis of some statistical criterion, such as a Maximum Likelihood. The aim of conditioning is to select those OMs consistent with the data and reject OMs that do not fit these data, and to reject satisfactorily and, as such, are considered implausible. As time progresses and further data become available, the conditioning process may be repeated to take these into account – this is referred to as reconditioning	
Error	Differences, primarily reflecting uncertainties in the relationship between the actual dynamics of the resource (described by the OMs) and observations. Four types of error may be distinguished, and simulation trials may take account of one or more of these: Estimation error: differences between the actual by the estimator when fitting a model to the available data; Implementation error: differences between intended management actions (as output by an MP) and those actually achieved (e.g. which may for example reflect over-catch); Observation error (or measurement error): differences between the measured value of some resource index and the corresponding value calculated by the OM; Process error: natural variations in resource dynamics (e.g., fluctuations about a stock- recruitment curve, or variation variation in fishery or survey selectivity /catchability).	

Term	Definition	Abbreviation/Symbol
Estimator	The statistical estimation process within a population model (assessment or OM); in a Management Strategy Evaluation (MSE) context, the component that provides information on resource status and productivity from past and generated future resource-monitoring data for input to the Harvest Control Rule (HCR) component of an MP in projections.	
Exceptional circumstances	Specifications of circumstances (primarily related to future monitoring data falling outside the range covered by simulation testing) where overriding of the output from a Management Procedure should be considered, together with broad principles to govern the action to take in such an event.	EC
Feedback Control	Rules or algorithms based, directly or indirectly, on trends in observations of resource indices, which adjust the management actions (such as a TAC change) in directions that will change resource abundance back towards a level consistent with decision makers' objectives when it moves away from these.	
Harvest Control Rule (also Decision Rule)	A pre-agreed and well-defined rule or action(s) that describes how management should adjust management measures in response to the state of specified indicator(s) of stock status. This is described by a mathematical formula.	HCR
Harvest Strategy	Some combination of monitoring, assessment, harvest control rule and management action designed to meet the stated objectives of a fishery. Sometimes referred to as a Management Strategy (see below). A fully specified harvest strategy that has been simulation tested to show satisfactory performance for performance and adequate robustness to uncertainties is often referred to as a Management Procedure.	HS
Implementation	The practical application of a Harvest Strategy to provide a resource management recommendation.	

Term	Definition	Abbreviation/Symbol
Kobe plot	A plot that shows the current stock status, or a trajectory over time for a fished population, with abundance on the horizontal axis and fishing mortality on the vertical axis. These are often shown relative to B_{MSY} and to F_{MSY}, respectively. A Kobe plot is often divided into four quadrants by a vertical line at $B = B_{MSY}$ and a horizontal line at $F = F_{MSY}$.	
Management Objectives	The social, economic, biological, ecosystem and political (or other) goals for a given management unit (i.e. stock). These typically conflict, and include concepts such as maximising catches over time, minimising the chance of unintended stock depletion, and enhancing industry stability through low inter-annual variability in catches. For the purposes of Management Strategy Evaluation (MSE), these objective need to be quantified in the form of Performance statistics (see below).	Objectives, MOs
Management Procedure	A management procedure has the same components as a harvest strategy. The distinction is that each component of a Management Procedure is formally specified, and the combination of monitoring data, analysis method, harvest control rule and management measure has been simulation tested to demonstrate adequately robust performance in the face of plausible uncertainties about stock and fishery dynamics.	MP
Management strategy	Synonymous with harvest strategy. (But note that this is also used with a broader meaning in a range of other contexts.)	
Management Strategy Evaluation	A process whereby the performances of alternative harvest strategies are tested and compared using stochastic simulations of stock and fishery dynamics against a set of performance statistics developed to quantify the attainment of management objectives.	MSE
Maximum Economic Yield	The (typically annual) yield that can be taken continuously from a stock sustainably (i.e. without reducing its size) that maximizes the economic yield of a fishery in equilibrium. This yield occurs at the effort level that creates the largest positive difference between total revenues and total costs of fishing (including the cost of labor, capital, management and research etc.), thus maximizing profits.	MEY

Term	Definition	Abbreviation/Symbol
Maximum Sustainable Yield	The largest (typically annual) yield that can be taken continuously from a stock sustainably (i.e. without reducing its size). In real, and consequently stochastic situations, this is usually estimated as the largest average long-term yield that can be obtained by applying a constant fishing mortality F, where that F is as denoted as F_{MSY}.	MSY
Observation Model	The component of the OM that generates fishery-dependent and/or fishery-independent resource monitoring data from the underling true status of the resource provided by the OM, for input to an MP.	
Operating Model(s)	A mathematical–statistical model (usually models) used to describe the fishery dynamics in simulation trials, including the specifications for generating simulated resource monitoring data when projecting forward in time. Multiple models will usually be considered so as to reflect the uncertainties about the dynamics of the resource and fishery.	OM(s)
Performance statistics/measures	A set of statistics used to evaluate the performance of Candidate MPs (CMPs) against specified management objectives, and the robustness of these MPs to important uncertainties in resource and fishery dynamics.	
Plausibility (weights)	The likelihood of a scenario considered in simulation trials representing reality, relative to the other scenarios also under consideration. Plausibility may be estimated formally based on some statistical approach, or specified based on expert judgement, and can be used to weight performance statistics when integrating over results for different scenarios (OMs).	
Precautionary Approach	An approach to resource management in which, where there are threats of serious irreversible environmental damage, lack of full scientific certainty is not used as a reason for postponing cost-effective measures to prevent environmental degradation.	PA
Reference case (also termed reference scenario or base case)	A single, typically central, conditioned OM for evaluating Candidate MPs (CMPs) that provides a pragmatic basis for comparison of performance statistics across different CMPs.	RC (or BC)
Reference set (also termed base-case or evaluation scenarios)	A limited set of scenarios, with their associated conditioned OMs, which include the most important uncertainties in the model structure, parameters and data (i.e. alternative scenarios which have both high plausibility and major impacts on the performance statistics of Candidate MPs).	RS

Term	Definition	Abbreviation/Symbol
Research-conditional option	Temporary application of an MP that does not satisfy conservation performance criteria, accompanied by both a research programme to check the plausibility of the scenarios that gave rise to this poor performance and an agreed subsequent reduction in catches should the research prove unable to demonstrate implausibility.	
Robustness tests	Tests to examine the performance of an MP across a full range (i.e. beyond the range of the Reference Set of models alone) of plausible scenarios. While plausible, robustness test OMs are typically considered to be less likely than the reference set OMs, and often focus on particularly challenging circumstances with potentially negative consequences which are to be avoided.	
Scenario	A hypothesis concerning resource status and dynamics or fishery operations, represented mathematically as an OM.	
Simulation trial/test	A computer simulation to project stock and fishery dynamics for a particular scenario forward for a specified period, under controls specified by a HS or MP, to ascertain the performance of that HS or MP. Such projections will typically be repeated a large number of times to capture the implications of stochasticity.	
Spawning Biomass, initial	Initial spawning biomass prior to fishing, as estimated from a stock assessment.	SSB_0
Spawning Biomass, current	Spawning biomass (SSB) in the last year(s) of the stock assessment.	$SSB_{current}$
Spawning Biomass at MSY	The equilibrium spawning biomass that results from fishing at F_{MSY}. In the presence of recruitment variability, fishing a stock at F_{MSY} will result in a biomass that fluctuates above and below SSB_{MSY}.	SSB_{MSY}

Term	Definition	Abbreviation/Symbol
Stationarity	The assumption that population parameter values are fixed (at least in expectation), and not varying systematically, over time. This is a standard assumption for many aspects of stock assessments, OMs and management plans.	
Stock assessment	The process of estimating stock abundance and the impact of fishing on the stock, similar in many respects to the process of conditioning OMs.	
Target Reference Point	The point which corresponds to a state of a fishery and/or resource which is considered desirable and which management aims to achieve.	TRP
Trade-offs	A balance, or compromise, achieved between desirable but conflicting objectives when evaluating alternative MPs. Trade-offs arise because of the multiple objectives in fisheries management, and the fact that some objectives conflict (e.g. maximizing catch vs minimizing risk of unintended depletion).	
Tuning	The process of adjusting values of control parameters of the Harvest Control Rule in a Management Procedure to achieve a single, precisely-defined performance statistic value in a specified simulation test. This reduces confounding effects to allow the performance of different candidate MPs to be compared more readily with respect to other management objectives. For example, in the case of evaluating the same rebuilding plans, all candidate MPs might be tuned to meet the rebuilding objective for a specified simulation trial; then the focus of comparisons among MPs is performance and behaviour with respect to the catch and CPUE dimensions.	
Weight(s)	Either qualitative (e.g. high, medium, low) or quantitative measures of relative plausibility accorded across a set of scenarios.	
Worm plot	Time series plots showing a number of possible realizations of simulated projections of, for example, catch or spawning biomass under the application of an MP for a specific OM or weighted set of OMs.	

Appendix B1

Table B.1: The Table below lists the robustness tests for the SBT developed by the CCSBT ESC, some of which have been considered in this thesis (CCSBT, 2019b).

Robustness test	Person[1]	Software syntax[2]	Description	Priority[3]	Hard/Easy
SFOC40	Ana	sfo40	40% overcatch by Australian surface fishery: ramps up from 1% in 1992 to 40% by 1999 and onwards to 2016. Adjust the age composition as was done for the 20% method. Continued 40% overcatch in projections	M	
SFO00	Ana	sfo00	No historical additional catch in surface fishery. No future additional catch in surface fishery	L	
Corr Sel		selrev	Reversing order of estimates at decadal scale "Corrugated selectivity"	L	Hard
		selalt	Five year blocks of Alternate bimodal and recent selectivity, most extreme case of bimodality should be used (for projections).	M	Hard
lowR10		reclow10	Reduce future recruitment by half during the first n years. For 2018, n was set to 10.	L	
lowR5		reclow5	Reduce future recruitment by half during the first n years. For 2018, n was set to 5.	H	
highR		rechigh	Increase future recruitment by 50% during the first n years. For 2018, n was set to 5.	M	Easy
q_{hsp1}	Ana	hspq1	Set HSP proportionality coefficient to 1, to be moved to reference set, next year	M	

Robustness test	Person	Software syntax	Description	Priority	Hard/Easy
q_hsp1	Ana	hspq1	Set HSP proportionality coefficient to 1, to be moved to reference set, next year	M	
h=0.55	Ana	h55	Just check any estimation tweaks that might be required	M	
GT qtrend		gtqtr	1% increase per year, note that an increasing q leads to over-estimated abundance	M	Easy
GT q low		gtql	q=0.85, Specifics and rationale to be determined	M	
GT q high		gtqh	q=1.15 Specifics and rationale to be determined	M	
GT overdisp.		gtod	Use over-dispersion as applied to conventional tagging	M	
GTI	Norio	troll	Includes the grid type trolling index as additional recruitment index. Increase CV of aerial survey to preclude aerial survey dominating the fit, given apparent conflicts in the data.	L	
IS20	Ann	fis20	Indonesian selectivity flat from age 20+	M	
Const sq. CPUE	Norio	cpuew1	Constant squares	L	
Var sq. CPUE	Norio	cpuew0	Variable squares	L	
Upq2008		cpueupq	CPUE q increased by 25% (permanent in 2008)	H	
S50CPUE	Ann	cpues50	50% of LL1 overcatch associated with reported effort	M	
S00CPUE	Ann	cpues00	Overcatch had no impact on CPUE	L	
Omega75		cpueom75	Power function for biomass-CPUE relationship with power = 0.75	H	
Drop q increase		cpuenocrp	of 0.5% yr-1 in future years – no continuous effort creep	L	Easy
High fut. CPUE CV	Rich	cpuehcv	Increase the future CPUE CV to 30% (currently 20%)	M	
		cpue59	Age range from 5-9, check connection between OM and projections...seem to be passed through so okay	M	
LL1 Case 2 of MR	Rich	case2	LL1 overcatch based on Case 2 of the 2006 Market Report	L	
Aerial2016		as2016	Remove the 2016 aerial survey data point.	H	

[1]Person: Scientist/MP developer of robustness test.

[2]Software syntax for robustness test used when setting up the code.

[3]Priority given to apply CMP to specific robustness test, where H is high, M is medium and L is low.

Appendix B2

This Appendix lists important information from the CCSBT Report of the Ninth Operating Model and Management Procedure (OMMP9) Technical Meeting which incorporates the following points made by the fifth Strategy and Fisheries Management Working Group (SFMWG5) meeting (CCSBT, 2018c).

Details pertaining to CMP testing:

Following extensive discussion, the SFMWG5 meeting agreed to the following objectives for use in the initial round of CMP testing:

- Tuning biomass levels of 0.25, 0.30, 0.35 and 0.40 of unfished spawning biomass SSB_0 (here interpreted as initial Total Reproductive Output; TRO_0);
- CMPs be tuned to a 50% probability of achieving the tuning biomass levels;
- The tuning year set to 2035, provided the projection period was not too short and did not lead to numerical issues;
- Projections should be extended to 2045 to evaluate post-2035 performance;
- All CMPs should achieve the current objective of providing at least a 70% probability of reaching 20% of SSB0 by 2035. Once the current interim rebuilding target of 20% of unfished spawning biomass has been reached, there should be a high probability that the stock would not fall below this level after 2035.

The following performance statistics were recommended by the SFMWG:

- Spawning biomass in medium term relative to SSB_0;
- Spawning biomass in short and medium terms relative to current;
- Minimum spawning biomass relative to current;
- Proportion of runs above the current biomass at the tuning year;
- SSB lower (10th) percentile continuing to increase (no decline over 2013-2035);
- Lower 10th SSB percentile in year t, e.g. in 10 years;
- Probability of meeting the interim rebuilding target by 2035 (aim to have at least 70% of the simulated trajectories rebuild to higher than 0.2 SSB_0 by 2035);
- Probability of dropping below 0.2 SSB_0 in any future year beyond 2035;
- Year at which 70% of simulations reach 0.2 SSB0;
- Median year that SSB_{MSY} is reached; and
- Probability of being above SSB_{MSY} in last 10 years (i.e., after 2035)

In terms of features of the CMP, the meeting agreed to conduct the test with the following specifications:

- Set TACs in 3-year blocks;
- Set the TAC for 2021-2023 in 2020 as the first TAC decision, noting that the usual lag between TAC setting and implementation will be reduced by 1 year to allow more time for MP development. The usual schedule would be used after that (i.e., in 2022 set TAC for 2024-2026);
- Set maximum TAC changes of 2000 mt, 3000 mt and 4000 mt, and add 5,000 t if the previous three did not provide sufficient contrast. Each level of maximum TAC change would not necessarily be applied in combination with all tuning levels. The OMMP group would decide on the appropriate scenarios to test each level of Maximum TAC change in this initial round.

Details pertaining to OM and data input that is updated in the code[4]

The Chair noted that after SFMWG5, the projection code and control files were updated to:

- run projections up to 2045;
- use UAM1 estimates as default for base projections;
- conduct the first TAC calculation in 2020 for 2021-2022 with no extra lag, and use the standard 2-year lag after that;
- simulate gene-tagging data; and
- simulate close-kin data

[4]Code updates have been provided via Github (CCSBT, 2018a)

Details pertaining to MP testing workplan

The following table lays out the elements of a workplan for MP developmet and implementation from 2018 to 2020. EC is the Extended Commission and ESC is the Extended Scientific Committee (note that "Extended" means including Taiwan.)

Table 5. Slightly modified table of work plan from SFMWG report.

2018		
March	SFMWG5	Initial discussions of rebuilding goals and MP features
June	OMMP9	First presentation of candidate MPs (CMPs) evaluated using 2017 OMs.
September	ESC + 1 day informal OMMP	Evaluation of refined CMPs.
October	EC	Results on CMP performance and trade-offs presented to EC. Consultation with stakeholders. Commission decides or amends broad recovery objectives and longer term performance based on advice from the ESC (and SFMWG).
2019		
June/July	OMMP10	**Recondition the OM** and review initial updated versions of CMPs to develop a limited set to put forward to the ESC. The week of June 17-21st.
September	ESC + 1 day informal OMMP	Review and advice on set of CMPs and a session for interaction with stakeholders.
October	EC	Aim to select and **adopt MP**.
2020		
June	Special ESC/EC meeting	Contingency placeholder in case more time is needed to complete evaluation
September	ESC	Implementation of adopted MP to provide **TAC advice for 2021** (i.e., no standard 1-year lag) (note, this MP implementation will include the 2020 data exchange). **Updated assessments including projections using adopted MP**
October	EC	Agrees TAC for 2021-2023.

Appendix B3

This Appendix lists relevant information from the Report of the Tenth Operating Model and Management Procedure (OMMP10) Technical Meeting (CCSBT, 2019a).

The updated reconditioned reference set of OM for the base case on which 2019 CMPs were applied to. Table 3.1 in section 3.2.2.4 was based on a slightly simpler version of Table B.2. Here, there is an updated additional row for unaccounted mortality (UAM1) and is described below. The update has affected the final results in Chapter 4.

Table B.2: This Table[5] is the grid configuration for the agreed reference set of OMs for the 2019 stock assessment with the addition of UAM1.

Paramter	**Value**	**CumuIN**	**Prior**	**Sampling**
h	$0.6, 0.7, 0.8$	3	Uniform	Prior
M_0	$0.35, 0.4, 0.45, 0.5$	12	Uniform	ObjFn
M_{10}	$0.05, 0.085, 0.12$	36	Uniform	ObjFn
Omega (ω)	1	36	Uniform	Prior
CPUE series	w0.5, w0.8	72	Uniform	Prior
CPUE age range	$4-18, 8-12$	144	$0.67, 0.33$	Prior
Psi (ψ)	$1.5, 1.75, 2$	432	$0.25, 0.5, 0.25$	Prior
UAM1	Described below	432		

[5]This is Table 1 from the Report of the Tenth Operating Model and Management Procedure Technical Meeting (CCSBT, 2019a).

The ESC had also agreed to include the UAM1 scenario in the reference set for the purposes of MP-testing (this is labelled base18 in this report), to account for uncertainty in total catches. This is as agreed in 2016 with respect to discussion of the “MP approach”. The “added-catch” (UAM1) scenario is currently implemented as unaccounted catch increasing linearly from 0 to 1000 mt over the period 1990 to 2013 and constant at 1000 mt for 2014-2018 for smaller fish and larger fish. These unaccounted mortalities were added to those already included in the reference set (e.g., 20% for the surface fishery). For future projections, the added catch was to remain at the same proportion of the TAC as in 2016. The unaccounted mortalities are assigned to the fisheries to whose size distributions there is the closest match (fishery 1 and 4 in the projection model). It should be understood that these fisheries may not be the source of the unaccounted mortality; rather this is an expedient way to implement the scenario.

Appendix B4

This Appendix provides detail for the model structure, and population model and stock-recruitment equations that are used to condition the OM developed for SBT MP testing. However, some details related to tagging and projections have been omitted because they are long, complex and not of key importance. It is taken from Attachment 6 of the Report of the Operating Model and Management Procedure Technical Meeting (CCSBT, 2009).

Attachment 6

Operating Model developed for SBT MP testing

Conditioning model

(sbtmod22.tpl, July 2009)

Model Structure

The SBT population is modeled as a single, age-structured stock. Historical trends in growth are allowed and fixed from parameters (mean and variances at age) estimated externally. The stock-recruitment relationship is given by a Beverton-Holt function with log-normal auto-correlated errors. Six fisheries are distinguished in the conditioning analysis, occurring in two pulses, according to:

Fishery	Catch data included	Pulse (season)	Actual period used for compiling statistics
LL1	Primarily Japanese LL areas 4-9 plus all LL catches not covered in LL2-LL5	(2) 1 July	Jan 1 through Dec 31
LL2	SBT caught in Taiwanese albacore LL fishery and Taiwanese gillnet catches	(2) 1 July	Jan 1 through Dec 31
LL3	Japanese LL in Area 2	(1) 1 Jan	Jan 1 through Dec 31
LL4-size	Japanese spawning fishery (Area 1)	(1) 1 Jan	July 1 through June 30
Indonesian	Indonesian spawning	(1) 1 Jan	July 1 through June 30
Australian Surface		(1) 1 Jan	July 1 through June 30

Population Model

The model is age-structured. Fishing for each fishery is treated as a pulse that takes place in one of two fishing seasons (see Table above). The population dynamics are:

$$N_{y+1,a+1} = N_{y,a}\left(1-\sum_{f\in f^1} H_{f,y,a}\right)\left(1-\sum_{f\in f^2} H_{f,y,a}\right)e^{-M_a} \quad \text{for } 0 \le a \le A-2, \quad y_{n1} \le y \le y_{n2}$$

$$N_{y+1,A} = N_{y,A-1}\left(1-\sum_{f\in f^1} H_{f,y,A-1}\right)\left(1-\sum_{f\in f^2} H_{f,y,A-1}\right)e^{-M_{A-1}} + N_{y,A}\left(1-\sum_{f\in f^1} H_{f,y,A}\right)\left(1-\sum_{f\in f^2} H_{f,y,A}\right)e^{-M_A} \quad \text{for } y_{n1} \le y \le y_{n2}$$

$$N_{y+1,0} = R_{y+1}$$

$$N^*_{y,a} = N_{y,a}\left(1-\sum_{f\in f^1} H_{f,y,a}\right)e^{-M_a/2}$$

$$H_{f,y,a} = s_{f,y,a}F_{f,y}$$

$$F_{f,y} = \frac{C_{f,y}}{\sum_a w_{f,y,a} s_{f,y,a} N_{y,a}} \quad \text{for} \quad f \in f^1$$
$$F_{f,y} = \frac{C_{f,y}}{\sum_a w_{f,y,a} s_{f,y,a} N^*_{y,a}} \quad \text{for} \quad f \in f^2$$

where:

- $N_{y,a}$ is the number of fish of age a at the start of year y,
- $N^*_{y,a}$ is the number of fish of age a at mid-year y,
- M_a denotes the natural mortality rate on fish of age a,
- $C_{f,y}$ is the catch of fish (biomass) in fishery f in year y,
- $F_{f,y}$ is the age-averaged fishing proportion of fishery f in year y,
- $H_{f,y,a}$ is the fishing proportion of fishery f in year y for fish of age a,
- $s_{f,y,a}$ is the standardized selectivity of fish of age a in fishery f in year y,
- $w_{f,y,a}$ is the average weight of fish of age a in year y in fishery f,
- R_y is the age-0 recruitment in year y,
- f^1 is the set of fisheries that occur in the first season (I33),
- f^2 is the set of fisheries that occur in the second season (I33), and
- A is the maximum age considered (I6, taken to be a plus-group).
- y_{n1}, y_{n2} are the first (I1) and the last (I2) years for the stock reconstruction.

Note that solutions are constrained so that the maximum harvest rate on an age-class during a fishing season is 0.9. For the MP reference case the maximum age considered, A, is 30.

Stock-Recruitment

The number of recruits at the start of year y $\left(R_y\right)$ is related to the spawning stock size by a stochastic Beverton-Holt stock-recruitment relationship. The relationship includes a parameter that allows for depensatory effects and has the option for serial correlation (AC) in a terminal sub-set of the residuals:

$$R_y = \frac{\alpha^r S_y}{\beta^r + S_y} \exp(\tau_y - 0.5\sigma_R^2)\left(1 - \exp\left(\frac{\ln(0.5) S_y}{\nu B_0^r}\right)\right)$$
$$\tau_y = \begin{cases} \delta_y & \text{if no AC} \\ \delta_y & \text{if AC and } y < y_{AC} \\ \varpi \tau_{y-1} + \delta_y & \text{if AC and } y \geq y_{AC} \end{cases}$$

where S_y is the spawning stock biomass in year y,

- α^r, β^r are Beverton-Holt stock-recruitment parameters for regime r,
- τ_y is the stock-recruitment residual for year y, $\tau_y \sim N(0, \sigma_R^2)$,
- ν is a depensation parameter (I23), (Note that setting ν at a very small number corresponds in the limit to no depensation),
- B_o^r is the equilibrium spawning stock biomass expected during regime r in the absence of fishing,

δ_y are stock-recruitment residual parameters estimated in the fitting procedure for years $y_{n1} > y \leq y_{n2} + 1$

ϖ is the empirical autocorrelation in the recruitment residuals, $\omega = Cor(\tau_y, \tau_{y-1})$, for $1966 \leq y \leq (y_{AC} - 4)$

y_{AC} is the year initiating the serial correlation in the stock-recruitment residuals (must be 1996 or later to activate this option, I13).

Spawning stock biomass is estimated as:

$$S_y \Sigma_{a=1}^{A} m_a \left(w_{y,a}^s\right)^{\kappa} N_{y,a}$$

where m_a is the proportion of fish of age a that are mature, w_{ya}^s, the spawning weight at age a in year y is assumed to be the same as the mean weight-at-age for the Indonesian spawning fishery, and κ is the exponent for a non-linear relationship between body size and reproductive potential (I24). Note that all these parameters, m_a, w_{ya}^s, and κ are specified as model inputs.

In order to work with parameters that are more meaningful biologically, the stock-recruitment relationship is reparameterized in terms of the equilibrium spawning biomass expected in the absence of fishing, B_0^r, and the "steepness", *h*, of the stock-recruitment relationship (steepness is defined as the fraction of the average spawning biomass expected in the unfished stock, which is obtained when recruitment is 20% of the recruitment expected in the unfished stock):

$$\alpha^r = \frac{4hR_0^r}{5h-1} \quad \text{and} \quad \beta^r = \frac{B_0^r(1-h)}{5h-1}$$

where

$$R_0^r = \frac{B_0^r}{\Sigma_{a=1}^{A-1} m_a \left(w_{y_{n1},a}^s\right)^{\kappa} \exp\left(-\Sigma_{a'=0}^{a'=a-1} M_{a'}\right) + m_A \left(w_{y_{n1},A}^s\right)^{\kappa} \dfrac{\exp\left(-\Sigma_{a'=0}^{a'=A-1} M_{a'}\right)}{1-\exp(-M_A)}}$$

Only a very limited regime shift option is currently coded in the SBT conditioning model. When the regime shift option (called carrying capacity) is invoked (I8b), an alternate stock-recruitment relationship, based on a different B_0^r, is used from 1978 onward. The two regimes share a common steepness parameter.

Selectivities

The parameterization of selectivity is age-specific and the model structure allows the selectivity to change slowly over time. For the first year in which there is catch data (I3) ($y = y_{c1}$), selectivities are functions of the estimated parameters:

$$s'_{f,y_{c1},a} = \begin{cases} 0 & \text{for } a < a_f^{mins} \\ \exp(\lambda_{f,a}) & \text{for } a_f^{mins} \leq a \leq a_f^{maxs} \\ \exp(\lambda_{f,a_f^{maxs}}) & \text{for } a > a_f^{maxs} \text{ and } f \in z \\ 0 & \text{for } a > a_f^{maxs} \text{ and } f \notin z \end{cases}$$

where a_f^{mins} and a_f^{maxs} are the minimum (I34) and maximum (I35) age-classes for which selectivity parameters are estimated for fishery f, z is the set of fisheries for which age-classes greater than a_f^{maxs} have the same selectivity as age-class a_f^{maxs} (I36). The code has two options for normalizing selectivities, controlled by a parameter hardwired in the code. The default option normalizes selectivities with respect to a reference age

$$a_f^{med} = \text{int}\left(\frac{a_f^{mins} + a_f^{maxs}}{2}\right) + 1$$

by

$s_{f,y_{c1},a} = \dfrac{s'_{f,y_{c1},a}}{s'_{f,y_{c1},a_f^{med}}}$ and forcing $\lambda_{f,a_f^{med}}$ to 0 using a quadratic penalty.

In the alternative parameterization (set jim_select = 1 in tpl code), selectivities are normalized with respect to the mean over the age range $a_f^{mins} \le a \le a_f^{maxs}$, and a quadratic penalty is added to the log of the mean selectivity for the first year to force the mean selectivity to 1. Also, in this version, when age-specific harvest rates exceed 0.90 during the minimization (i.e. "kludge" message) only the harvest rate where the bound is exceeded is reduced. In previous versions, all harvest rates were adjusted. This is a preferable way to constrain harvest rates but it is ~30% slower to run.

For other years ($y > y_{c1}$),

$$s'_{f,y,a} = \begin{cases} s_{f,y-1,a}\exp(\gamma_{f,y,a}) & \text{for } y \in c^f, \quad \gamma_{f,y,a} \sim N(0, \sigma^2_{s_y^f}) \\ s_{f,y-1,a} & \text{for } y \notin c^f \end{cases}$$

where c^f is the set of years in which the fishing selectivity ogive can change for fishery f (non-zero I40), and γ_{fya} reflects the amount of change in the age effect of fishery f for age a.

After each update, selectivities are again normalized according to the parameterization chosen:

$$s_{f,y,a} = \frac{s'_{f,y,a}}{s'_{f,y,a_f^{med}}} \qquad \text{or} \qquad s_{f,y,a} = \frac{s'_{f,y,a}}{\text{mean}(s'_{f,y,a_f^{mins}}, \cdots, s'_{f,y,a_f^{maxs}})}.$$

The stochastic error terms, $\gamma_{f,y,a}$ are treated as free parameters subject to the constraints of their input variances, $\sigma^2_{S_y^f}$ (I40).

If the age effects of fishing ($s_{f,y,a}$) are constant over time, this results in a decomposition of the fleet-specific fishing mortality rate into an age component and a year component. This assumption creates what is known as a separable model. If the age effect of fishing in fact changes over time, then the separable model can mask important changes in fish abundance. The constraints imposed through the variance terms can restrict the selectivity to change only slowly over time, thus improving the ability to estimate the $\gamma_{f,y,a}$'s. Also, to provide smoothness in the age component there is a curvature penalty on the age-specific coefficients. This can be based on either the logarithm of the selectivity parameters (I38=0), or a non-negative power of the selectivity parameters (I38>0):

$$x_{f,y,a} = \begin{cases} \ln(s_{f,y,a}) & \text{for I38} = 0 \\ (s_{f,y,a})^{\text{I38}} & \text{for I38} > 0 \end{cases}$$

Then a penalty term, based on either squared second-differences or squared third-differences, is added to the negative log-likelihood function for each fishery:

$$g^f\left(x_{fya};\sigma_{b^f}^2\right)=\begin{cases}\displaystyle\sum_{y\in\left(y_{c1},c^f\right)}\sum_{a=a_f^{mins}}^{a=a_f^{maxs}-2}\frac{\left(x_{f,y,a+2}-2x_{f,y,a+1}+x_{fya}\right)^2}{2\sigma_{b^f}^2} & \text{for } \mathrm{I39}=2\\ \displaystyle\sum_{y\in\left(y_{c1},c^f\right)}\sum_{a=a_f^{mins}}^{a=a_f^{maxs}-3}\frac{\left(x_{f,y,a+3}-3x_{f,y,a+2}+3x_{f,y,a+1}-x_{fya}\right)^2}{2\sigma_{b^f}^2} & \text{for } \mathrm{I39}=3\end{cases}$$

This prevents irregular shifts between adjacent age classes. A selection of the third differences penalty function encourages selectivity to be dome-shaped with age while the second difference penalty function favours linear behaviour with age.

Growth

Growth is not estimated in the model, but is fixed with assumed known length-age relationships. The mean length at age is input for each year y and season t, so growth can change over time. Also, fixed length-weight relationships are assumed for each fishery. The length frequency distributions for each age are calculated assuming normal distributions. The standard deviation ($\sigma_{t,y,a}$) of length-at-age is linearly related to the mean length-at-age ($\mu_{t,y,a}$) based on the relationship of Kolody and Polacheck (2001): $\sigma_{t,y,a}=2+\mu_{t,y,a}/30$.

Natural Mortality

Natural mortality is assumed to vary over age as a function of four parameters: m^1, m^4, m^{10} and m^{30}, which correspond respectively to the instantaneous rate of natural mortality at ages 1, 4, 10, and 30+.

$$M_a=\begin{cases} m^1 & \text{for } a=0\\ m^1+\dfrac{m^4-m^1}{3}(a-1) & \text{for } 1\le a<4\\ m^4+\dfrac{m^{10}-m^4}{6}(a-4) & \text{for } 4\le a\le 10\\ m^{10} & \text{for } 10\le a\le 25\\ M_{25}+\dfrac{m^{30}-M_{25}}{5}(a-25) & \text{for } 25<a<30\\ m^{30} & \text{for } a\ge 30\end{cases}$$

Two of the parameters (m^1 and m^{10}) are fixed while the other two (m^4 and m^{30}) are estimated. The parameter m^4 is bounded between m^1 and m^{10}. Conditioning trials showed that the estimates of m^4 and m^{30} had low coefficients of variation, while there was considerable uncertainty around m^1 and m^{10}. Thus, a range of values are selected for m^1 and m^{10} to reflect that uncertainty in model projections.

Further details on Tagging Model and its corresponding parameter descriptions have been omitted here and the next section of this Appendix relates to details related to future data predicted by the OM and followed by tables listing model input parameter descriptions, quantities used in the CCSBT OM code and quantities estimated through objective function.

PREDICTED QUANTITIES

Catch-at-age and Catch-at-length

Observations of either catch-at-age or catch-at-length are available for each of the fisheries, and are fitted in the model. The predicted catch-at-age a in fishery f and year y is:

$$\hat{C}_{f,y,a} = s_{f,y,a} F_{f,y} N_{y,a} \quad \text{for} \quad f \in f^1$$
$$\hat{C}_{f,y,a} = s_{f,y,a} F_{f,y} N^*_{y,a} \quad \text{for} \quad f \in f^2$$

For fisheries with length-based data, the predicted catch-at-length l in fishery f and year y is given by:

$$\hat{L}_{f,y,l} = \sum_a p^t_{y,a,l} \hat{C}_{f,y,a} \text{ for } f \in f^1, t = 1 \text{ and for } f \in f^2, t = 2$$

where $p^t_{y,a,l}$ is the proportion of fish of age a that are length l in season t, calculated assuming normal distributions for length-at-age with known means and variances.

CPUE

Catch per unit effort (CPUE) is fitted as an aggregate index (i.e. not age-based) for the LL1 fishery only. The relationships between CPUE and abundance and between CPUE and effort allow for a number of non-linear effects. These effects are not estimated in the model fitting procedure, but rather are determined by control parameters input by the user. The predicted CPUE in year y is given by:

Further details to tag returns omitted here.

$$CPUE_y = q_y \widetilde{N}_y^{\varpi}\left(1+\beta\left(\frac{E_y - E_{2000}}{E_{2000}}\right)+\gamma\left(\frac{E_y - E_{2000}}{E_{2000}}\right)^2\right)$$

where $\widetilde{N}_y = \Sigma_a\left(\frac{s_{LL1,y,a}}{\mathrm{mean}(S_{LL1,y,a_1},\cdots,S_{LL1,y,a_2})}\right)^{\psi} N_{y,a}$

and $E_y = \frac{C_{LL1,y}}{CPUE_y}$

In this model, parameters β, γ, ϖ, ψ, q_y and a_1 and a_2 are specified by the user. Current default values are: $\beta = 0$, $\gamma = 0$, $\omega = 1$, $\psi = 1$, $(a_1, a_2) = (4,18)$ or $(8,12)$.
Parameters β and γ : changing the values of β and γ had little or no effect in the conditioning (CCSBT-MP/0304/07).
Parameter ω : Is one of the axes in the grid, with values 1 and 0.75.
Parameters a_1 and a_2 (age range to standardize selectivity for CPUE predictions) are included as one grid axes with two alternative ranges: (1) a_1=4 and a_2=18 (2) a_1=8 and a_2=12. The rational for changing a_2 from 30 to 18 was that selectivities estimated for ages 19-30 are very low.

The only parameter in the above equations that is estimated through the minimization is $\ln(q_y)$ for the first year of the CPUE series. In the current version a fixed 0.5% annual increase in q is assumed.

The analyses looking at historical CPUE trend based on a linear increase (CCSBT-MP/0304/07) showed that no improvement was obtained by imposing this relationship. A test assuming a linear increase in catchability of 1% per year throughout the whole time series was examined. This test was later dropped but an increase in q of 0.5 % a year (half way between Q0 and Q1) was kept in both the conditioning and in the projections in the core set.

Tag Returns

Let $N_{c,a,g}$ denote the number of fish from cohort c tagged at age a by taggers in group g. We refer to this set of tag releases as set (c, a, g). Let $R_{c,a,g,i}$ be the observed number of fish from release set (c, a, g) that were recaptured at age i and had at least one of their tags returned (for simplicity, we will refer to this as the number of tag returns). Then, the predicted number of tag returns is given by

$$\hat{R}_{c,a,g,i} = N_{c,a,g}\, p_{c,a,g,i}$$

Aerial survey

The aerial survey data are treated as a relative index of biomass of age classes 2 to 4, predicted as

$$\hat{I}_i = q_{\mathrm{aerial}} \sum_{a=2}^{a=4} s_a w_{y_i,a} N_{y_i,a},$$

where s_a indicate selectivities-at-age and w_{y_i} are weights at age for season 1 in year y. An initial attempt to estimate the selectivity parameters produced unrealistic results. Three alternative fixed selectivity scenarios are available and can be chosen in the control file:

Option	s_2	s_3	s_4
1	1	1	1
2	0.5	1	1

3	0.33	1	0.33

Trolling survey

Treated as a relative index of abundance at age 1.

OBJECTIVE FUNCTION

Likelihood Components for Data Fits

The model is fitted to a CPUE index series, fishery catch-at-age and catch-at-length data, and tag return data. The estimates of total catch for each fishery are assumed to be without error. The negative of the log-likelihood (-lnL) for each of the data components are described below. Note that constant terms of the negative log-likelihood are ignored.

CPUE data

The likelihood is calculated assuming that the observed abundance index (I14) is log-normally distributed about its expected value with variance σ_I^2:

$$-\ln L = n_I \ln(\sigma_I) + \frac{\sum_{y=y_{I1}}^{y=y_{I2}} \left(\ln(I_y) - \ln(\hat{I}_y)\right)^2}{2\sigma_I^2}$$

where y_{I1} and y_{I2} are the first (I4) and the last (I5) years with CPUE data and n_I $(n_I = y_{I2} - y_{I1} + 1)$ is the number of CPUE observations. The variance parameter, σ_I^2, is estimated through the fitting procedure, assuming a normal distribution with a minimum value of $(0.1)^2$.

Catch-at-age and catch-at-length

For fitting to catch-at-age and catch-at-length data a multinomial sampling distribution is assumed. Under this assumption, the log-likelihood function for the catch-at-age or catch-at-length data (in numbers) from each fishery can be written:

$$-\ln L = n_y^f \sum_y \sum_k p_{fyk} \ln(\hat{p}_{fyk})$$

where $k = a$ for catch-at-age data, $k = l$ for catch-at-length data, n_y^f is the effective sample size for fishery f in year y, and

$$p_{fya} = \frac{O_{fya}}{\sum_a O_{fya}}, \quad \hat{p}_{fya} = \frac{\hat{C}_{fya}}{\sum_a \hat{C}_{fya}} \quad \text{for age-based data}$$

$$p_{fyl} = \frac{O_{fyl}}{\sum_l O_{fyl}}, \quad \hat{p}_{fyl} = \frac{\hat{L}_{fyl}}{\sum_l \hat{L}_{fyl}} \quad \text{for length-based data}$$

The O_{fya}, O_{fyl}, $\hat{C}_{fya}$, $\hat{L}_{fyl}$ are the observed and predicted catch-at-age or catch-at-length for fishery *f*. The effective sample sizes, n_y^f, are quantities input for each fishery and year (I41).

Different methods are used for inputting age-frequency and length-frequency data in the SBT conditioning model code. Input of length-frequency data is hard-wired such that the code expects input of data for 110 length-frequency bins each of 2cm width and beginning at 32cm. The user controls the fitting of these data by specifying the minimum length category fitted in the model (I25, fish in bins of smaller length than the minimum are aggregated in the first bin), the width of the length bins used in the fitting procedure (I26, best to specify this in 2 cm increments, consistent with how the data is input), and the number of bins used in the fitting (I27, note that any fish of length greater than the length of the terminal bin are aggregated in the terminal bin). The specified binning values apply to the length-frequency data from all fisheries. An additional option allows for fishery-specific aggregation of a specified number of the smallest length bins seen by the model (I28).

For age-frequency data, the data input controls the age range in the model fit. For each fishery with age-frequency data (I29 and I30) the user specifies the minimum (I31) and the maximum (I32) ages in the data set. The only aggregation of age-classes that is allowed is when the maximum age specified for fitting the age-frequency is the same as the maximum age in the model, which is a plus group.

<u>Tag Returns</u>

If all assumptions of a Brownie tagging model are met (e.g., complete mixing; independence between tagged fish), then the numbers of tags returned at ages *a* to *I*, plus the number not returned by age *I*, corresponding to the $N_{c,a,g}$ releases from release set (*c*, *a*, *g*) have a multinomial distribution; i.e.,

$$\left\{R_{c,a,g,a}, \ldots, R_{c,a,g,I}, N_{c,a,g} - R_{c,a,g,\bullet}\right\} \sim \text{Multinom}\left(N_{c,a,g}, \left\{p_{c,a,g,a}, \ldots, p_{c,a,g,I}, 1 - p_{c,a,g,\bullet}\right\}\right)$$

where a dot in the subscript denotes summation over the index it replaces (e.g., $R_{c,a,g,\bullet} = \sum_{i=a}^{I} R_{c,a,g,i}$).

However, in practice, the tag return data will almost certainly be over-dispersed relative to a multinomial distribution (i.e., more variable). To account for this, we model the tag returns for release set (*c*, *a*, *g*) using a Dirichlet-multinomial distribution, parameterized such that the amount of variance in the data is φ times that of multinomial data (for details refer to Polacheck et al. 2006. Can. J. Fish. Aquat. Sci. **63**: 534–548). Then, the likelihood function for the observed numbers of returns from all release sets is the product of Dirichlet-multinomials.

Aerial survey

A log-normal likelihood with autocorrelated error and added process error was used:

$$-\ln L_{\text{aerial}} = 0.5 \ \ln|\Sigma| \ + \ 0.5 \ \ res^T \Sigma^{-1} res$$

where

$res = \ln(I) - \ln(\hat{I})$ is a vector of residuals computed using a maximum likelihood estimate of the log of the proportionality coefficient

$$\ln \hat{q}_{\text{aerial}} = \frac{1^T \Sigma^{-1} res}{1^T \Sigma^{-1} 1}$$

and $\Sigma = S \ + \ 1 \ \tau^2_{\text{aerial}}$ is a variance-covariance matrix with S the empirical variance-covariance matrix for the logged survey indices and τ_{aerial} an estimated parameter representing added process error (which would impact projections for MP considerations).

Trolling survey

This index (for 1996 onwards) is used in sensitivity trials. A normal likelihood with constant estimated variance σ^2_{piston} is assumed.

Likelihood Components for Priors

Stock-recruitment relationship

The stock-recruitment relationship used in the SBT model requires prior assumptions about the stock-recruitment steepness parameter, the magnitude of the change in carrying capacity, and the magnitude of the recruitment residuals. The steepness parameter can either be fixed (I9<1) or estimated in the analysis (I9 $\geq$ 1). When estimated, the steepness is assumed to be normally distributed $h \sim N\left[\tilde{h}, 0.1^2\right]$, but the user can specify a tighter area of support (i.e bounds, I10 and I 11) than would be expected for a normal distribution. The negative log-likelihood for the steepness prior is:

$$\frac{\left(h-\tilde{h}\right)^2}{2\left(0.1\right)^2} \quad \text{where } \tilde{h} = 0.5^*(\text{I10+I11})$$

A normal distribution (in log space) is also assumed for the stock-recruitment residuals, $\tau_y \sim N\left[0, \sigma_R^2\right]$. The variance of the residuals can either be fixed (I12 <1, σ_R=I12) or estimated (I12 ≥ 1). In either event, the negative log-likelihood for the normal distribution prior is:

$$(y_{n2} - y_{n1} + 1)\ln(\sigma_R) + \frac{\sum_{y=y_1+1}^{y=y_n+1} \tau_y^2}{2\sigma_R^2} \quad .$$

Note that when estimated there is a lower bound of 0.4 on the σ_R parameter.
The likelihood assumes no autocorrelation except for the last three years (e.g. 2007-2009 when the last year of data is 2008). The empirical autocorrelation of recruitment residuals estimated over the period 1965-2003 is applied from 2007 onward. Let τ_y represent the lognormal recruitment deviate in year y and $\hat{\tau}_y$ its MPD estimate. The initial abundances passed to the projection code (when troll data are not included) correspond to

$\hat{\tau}_{2006}$ estimated from model fit

$\hat{\tau}_{2007} = \rho\hat{\tau}_{2006}$

$\hat{\tau}_{2008} = \rho^2\hat{\tau}_{2006}$

$\hat{\tau}_{2009} = \rho^3\hat{\tau}_{2006}$

where $\hat{\rho}$ is the empirical estimate of autocorrelation based on recruitments for years 1965-2003.

An uninformative prior is assumed for the change in the carrying capacity (*i.e.* uniform), so the contribution to the objective function is a constant.

Selectivity

The age-specific selectivity parameterization incorporates two type of assumption that reflect prior belief about the form of the selectivity function. For all fisheries either a dome-shaped or linear relationship between selectivity and age can be specified. The negative log-likelihood for the prior is:

$$\sum_f g^f\left(x_{fya}; \sigma_{b^f}^2\right)$$

where the variance term, $\sigma_{b^f}^2$ (I37), reflects belief about the degree to which the selectivities for fishery *f* follow the expected shape (domed or linear).

For some or all fisheries, the age-specific selectivity functions can change over time. The amount of change is controlled by input parameters (fishery and year specific, I40) related to the variance of the changes $\left(\sigma_{S_y^f}^2\right)$.

$$\sum_f \sum_{y \in c^f} \frac{\left(\gamma_{fya}\right)^2}{2\sigma_{S_y^f}^2}$$

where c^f is the set of years in which selectivity changes for fishery *f*. Note that for fisheries with time-invariant selectivity this set will be empty.

Natural Mortality

Additional components are added to the likelihood function if natural mortality at ages 1 and/or 10 are also estimated. In that case normal prior distributions are assumed for both parameters. The negative log-likelihoods for these priors are:

$$0.5\frac{(m^1-0.4)^2}{0.04^2} \text{ and } 0.5\frac{(m^{10}-0.10)^2}{0.06^2}$$

Table 1. Fixed quantities determined through model inputs.

Quantity	Description	Control file code
y_{n1}, y_{n2}	first and last years for reconstruction	I1, I2
y_{c1}	first year for catch data	I3
y_{I1}, y_{I2}	first and last years for CPUE index data	I4, I5
y_{AC}	the year that initiates serial correlation in the stock-recruitment residuals	I13
A	last age class in model	I6
	number of fisheries	I7
a_f^{mins}, a_f^{maxs}	minimum and maximum age-class for which selectivity parameters are estimated for fishery f	I34, I35
f^1, f^2	the set of fisheries in season 1 and in season 2	I33
c^f	the set of years in which selectivity changes for fishery f	I40
z	the set of fisheries where selectivity for fish older than a_f^{maxs} is equal to that of a_f^{maxs}	I36
$\beta, \gamma, \varpi, \psi,$ q_y, a_1, a_2	parameters determining the relationship between CPUE and stock abundance	I16, I17, 118, I19, I20, I21, I22
ν	stock-recruitment depensation parameter	I23
κ	parameter for non-linear body weight-reproductive potential relationship	I24a
h	stock-recruitment steepness parameter (Note: also can be estimated)	I9 $\geq$ 1, then $h = 0.5(\text{I10+I11})$
M_a	natural mortality (Note: also can be estimated)	I8a
m_a	fraction mature at age	
σ_R^2	variance of stock-recruitment residuals (Note: can be estimated)	I12<1
$\sigma_{b^f}^2$	variance for the shape of the selectivity function for fishery f	I37
$\sigma_{S_y^f}^2$	variance of the selectivity change in year y for fishery f	I40
n_y^f	multinomial sample size for length or age sample from fishery f in year y	I41
φ	over-dispersion factor for tagging Dirichlet-multinomial	

Table 2. Quantities "hardwired" in code (i.e. you will need to change the code if you want to change these)

Quantity	Description
b_a	proportion of fish mature at age a
w_{fya}	mean weight at age a in fishery f in year y – dependent on input mean lengths-at-age, but weight-length relationship for each fishery is hardwired
w^s_{ya}	mean weight of spawning fish at age a in year y is set equal to mean weight-at-age for fishery 1 (LL1 fishery)
ϖ	empirical correlation of S-R residuals used in "hard-wired" AC is based on residuals from 1966 to last year of data minus 5. It is applied from control parameter rec_AC_sw onwards (usually set to last year of data minus 1)

Table 3. Quantities estimated through the function minimization. Note that with the exception of the stock-recruitment steepness parameter and the variances of the stock-recruitment residuals and the selectivity changes, the prior distributions are "hardwired" in the code. (ie. you will need to change the code if you wish to change the prior).

Quantity	Description	Prior
B_0^r	Equilibrium spawning stock biomass in the absence of fishing for regime r	$B_0^r \sim U[0,\infty]$
h	stock-recruitment steepness (Note: can also be a fixed quantitiy)	$h \sim N\left[\tilde{h}, 0.1^2\right]$ $\tilde{h} = 0.5*(\text{I10+I11})$
$\overline{q}$	Log of catchability	$U[-\infty,\infty]$
$\ln \hat{q}_{\text{aerial}}$	logarithm of "catchability" of aerial survey	$U[-\infty,\infty]$
τ_{aerial}	standard deviation of added process error for aerial survey	$U[0,0.8]$
m^1	natural mortality at age 1	$N(0.4,0.4^2)$
m^4	natural mortality at age 4	$U[m^1,m^{10}]$ or fixed
m^{10}	natural mortality at age 10	$N(0.1,0.6^2)$
m^{30}	natural mortality at age 30	$U(0.20,0.50)$ or fixed
δ_y	parameters related to the stock-recruitment residuals – note that the prior distribution is for the s-r residuals, τ_y, not the estimated parameters	$\tau_y \sim N\left[0,\sigma_R^2\right]$
λ_{fa}	selectivity parameter for age a in fishery f	$\lambda_{fa} \sim U[0,\infty]$
γ_{fya}	logarithm of the parameter governing the change in selectivity at age a in year y and fishery f	$\gamma_{fya} \sim N\left[0,\sigma_{S_y^f}^2\right]$
σ_I^2	variance of the CPUE index data	$\sigma_I^2 \sim U[0.2,\infty]$
σ_R^2	variance of stock-recruitment residuals (fixed in reference set)	$\sigma_R^2 \sim U[0.4,\infty]$
$h_{1.c.i}^*$	proportion of age *i* fish from cohort *c* that were tagged at age *i* and recaptured in the season directly following release	
$\upsilon_{c.i}$	Tag reporting rates for fish of age *i* cohort *c*	

Appendix C1

This Appendix lists all the robustness tests for ABFT that are available for the final round CMP testing, updated and corresponding to v5.2 updates. There are 35 robustness tests in total, where Table C.1 lists 11 priority robustness tests and Table C.2 lists 24 the other robustness tests. The tests designated as ROM 13, ROM 14 and ROM 15 in the main text (section 3.2.2) correspond to ROM_1, ROM_2 and ROM_3 from the priority robustness tests listed in Table C.1 below. Please note the notation of the robustness tests in Table C.1 is to link to the factors and levels specicified ed in Table 3.4 in Ca.

Table C.1: Priority Robustness Tests.[6]

	One factor deviation from OM		
	OM_4: 1wBI	**OM_5: 2wBI**	**OM_6: 3wBI**
Western Contrast. Increased precision (CV of 15%) of the GOM_LAR_SUV index to create greater contrast in current Western stock status	ROM_1	ROM_2	ROM_3
	OM_1: 1AI	**OM_2: 2AI**	
Gulf of Mexico SSB. Prior on higher GOM SSB in quarter 2 and lower GOM SSB in quarter 3	ROM_4	ROM_5	
'Brazilian catches'. Catches in the South Atlantic during the 1950s are reallocated from the West to the East.	ROM_6	ROM_7	
Time varying mixing. Future movement switches from half stock mixing (robustness scenario 1) to 150% stock mixing every three years.	ROM_8	ROM_9	
Persistent change in mixing. Future movement permanently switches from half mixing to 150% mixing after 10 years.	ROM_10	ROM_11	

[6] Table C.1 is Table A.1 from ICCAT paper: *Application of "fixed proportion" candidate management procedures for North Atlantic Bluefin Tuna using operating model package version 5.2.3* (Butterworth *et al.*, 2020).

Table C.2: Other suggested robustness tests[7] Upweighting refers to a five-fold increase in the likelihood weighting component for a particular data type.

	One factor deviation from OM	
	OM_1: 1AI	OM_2: 2AI
Senescence. An increase in natural mortality rate for older individuals as applied in CCSBT	ROM_12	ROM_13
Upweighting of CPUE indices.	ROM_14	ROM_15
Upweighting of 'fishery independent' indices.	ROM_16	ROM_17
Upweighting of genetic stock of origin data. 5x likelihood factor on genetics, ignore microchemistry SOO data by increasing imprecision to a logit CV of 500%	ROM_18	ROM_19
Greater influence of microchemistry stock of origin data. 5x likelihood factor on genetics, ignore microchemistry SOO data by increasing imprecision to a logit CV of 500%	ROM_20	ROM_21
Greater influence of the Length composition data.	ROM_22	ROM_23
Greater influence of the historical landings data.	ROM_24	ROM_25
Unreported overages. Future catches in both the West and East are 20% larger than the TAC as a result of IUU fishing (not accounted for by the MP).	ROM_26	ROM_27
Catchability Increases. CPUE-based indices are subject to a 2% annual increase in catchability.	ROM_28	ROM_29
Non-linear indices. Hyperstability / hyper depletion in OM fits to data is simulated in projection years for all indices.	ROM_30	ROM_31
Probabilistic recruitment shifts. The same recruitment shift as Factor 1 level 3, but with prob of 0.05 for each of the first 20 years of projection.	ROM_32	ROM_33
Decreasing catchability. 2% annual decline in the catchability of CPUE-based indices.	ROM_34	ROM_35

[7]Table C.2 is from Table A.2 from ICCAT paper: *Application of "fixed proportion" candidate management procedures for North Atlantic Bluefin Tuna using operating model package version 5.2.3* (Butterworth *et al*, 2020).

Table C.3: Deterministic results for depletion (Br30) after and average catch (AvC30) over a 30-year projection period for for the "C=0" scenario and two CMPs FXP_1($\beta = 0.5, \alpha = 0.5$) and FXP_2($\beta = 1, \alpha = 1$) for robustness tests ROM_1 to ROM_30 as described in Appendix F. Note that ROM_1, ROM_2 and ROM_3 as indicated below correspond to what are termed ROM13, ROM14 and ROM15 in Chapter 3.

	West						East					
	Br30			**AvC30**			**Br30**			**AvC30**		
	C=0	FXP1	FXP2	C=0	FXP1	FXP2	C=0	FXP1	FXP2	C=0	FXP1	FXP2
ROM_1	2.690	1.106	0.820	0.223	3.111	3.638	2.991	2.263	1.956	2.652	21.718	29.546
ROM_2	1.398	0.410	0.149	0.223	2.534	3.204	2.445	1.910	1.527	2.652	17.730	27.640
ROM_3	2.867	1.716	1.423	0.223	3.113	3.638	2.677	1.655	1.305	2.652	22.712	29.546
ROM_4	2.931	2.222	1.975	0.223	2.862	3.624	3.157	2.499	2.091	2.652	19.560	29.546
ROM_5	2.450	2.106	1.919	0.223	2.587	3.615	2.544	2.029	1.643	2.652	17.644	27.839
ROM_6	2.909	2.128	1.870	0.223	2.891	3.625	3.163	2.510	2.105	2.652	19.539	29.546
ROM_7	2.404	1.959	1.727	0.223	2.612	3.616	2.535	2.008	1.616	2.652	17.685	27.842
ROM_8	3.019	2.528	2.342	0.223	2.709	3.622	3.157	2.411	2.411	2.652	21.856	29.546
ROM_9	2.502	2.233	2.065	0.223	2.456	3.615	2.568	2.043	1.675	2.652	18.821	28.967
ROM_10	3.019	2.500	2.296	0.223	2.644	3.621	3.157	2.387	2.112	2.652	22.321	29.546
ROM_11	2.502	2.210	2.020	0.223	2.430	3.614	2.568	2.034	1.668	2.652	19.045	29.354
ROM_12	2.830	2.346	2.124	0.223	2.711	3.624	3.101	2.580	2.175	2.652	18.648	29.218
ROM_13	2.380	2.147	2.013	0.223	2.540	3.619	2.541	2.178	1.878	2.652	17.301	27.549
ROM_14	2.998	2.466	2.140	0.223	2.417	3.569	3.033	2.085	1.631	2.652	21.377	29.546
ROM_15	2.420	1.974	1.628	0.223	2.299	3.562	2.295	1.502	0.971	2.652	17.546	25.858
ROM_16	2.705	1.402	1.090	0.223	3.078	3.633	3.196	2.541	2.185	2.652	20.238	29.546
ROM_18	2.784	1.902	1.683	0.223	3.056	3.618	3.233	2.676	2.345	2.652	19.653	29.546
ROM_19	2.148	1.480	1.191	0.223	2.694	3.608	2.642	2.216	1.892	2.652	17.411	27.773
ROM_20	2.894	1.953	1.631	0.223	2.809	3.621	3.128	2.432	2.023	2.652	19.984	29.546
ROM_21	2.000	1.089	0.774	0.223	2.948	3.621	2.491	1.890	1.465	2.652	18.358	28.315
ROM_22	2.965	2.366	2.162	0.223	2.871	3.630	3.209	2.705	2.371	2.652	19.121	29.546
ROM_23	2.533	2.399	2.323	0.223	2.505	3.616	2.732	2.499	2.310	2.652	16.721	27.173
ROM_24	2.902	2.157	1.913	0.223	2.891	3.625	3.152	2.476	2.079	2.652	19.867	29.546
ROM_25	2.385	1.926	1.691	0.223	2.636	3.616	2.542	2.014	1.622	2.652	17.870	28.118
ROM_26	2.922	2.109	1.901	0.223	2.975	3.627	3.158	2.424	2.092	2.652	21.152	29.546
ROM_27	2.405	1.931	1.713	0.223	2.724	3.618	2.539	1.961	1.551	2.652	18.966	29.546
ROM_28	2.922	2.238	1.916	0.223	2.656	3.621	3.158	2.578	2.140	2.652	17.991	28.576
ROM_29	2.405	2.006	1.750	0.223	2.420	3.613	2.539	2.062	1.691	2.652	16.654	16.342
ROM_30	2.922	2.657	2.455	0.223	1.253	2.032	3.158	2.927	2.822	2.652	10.696	13.927

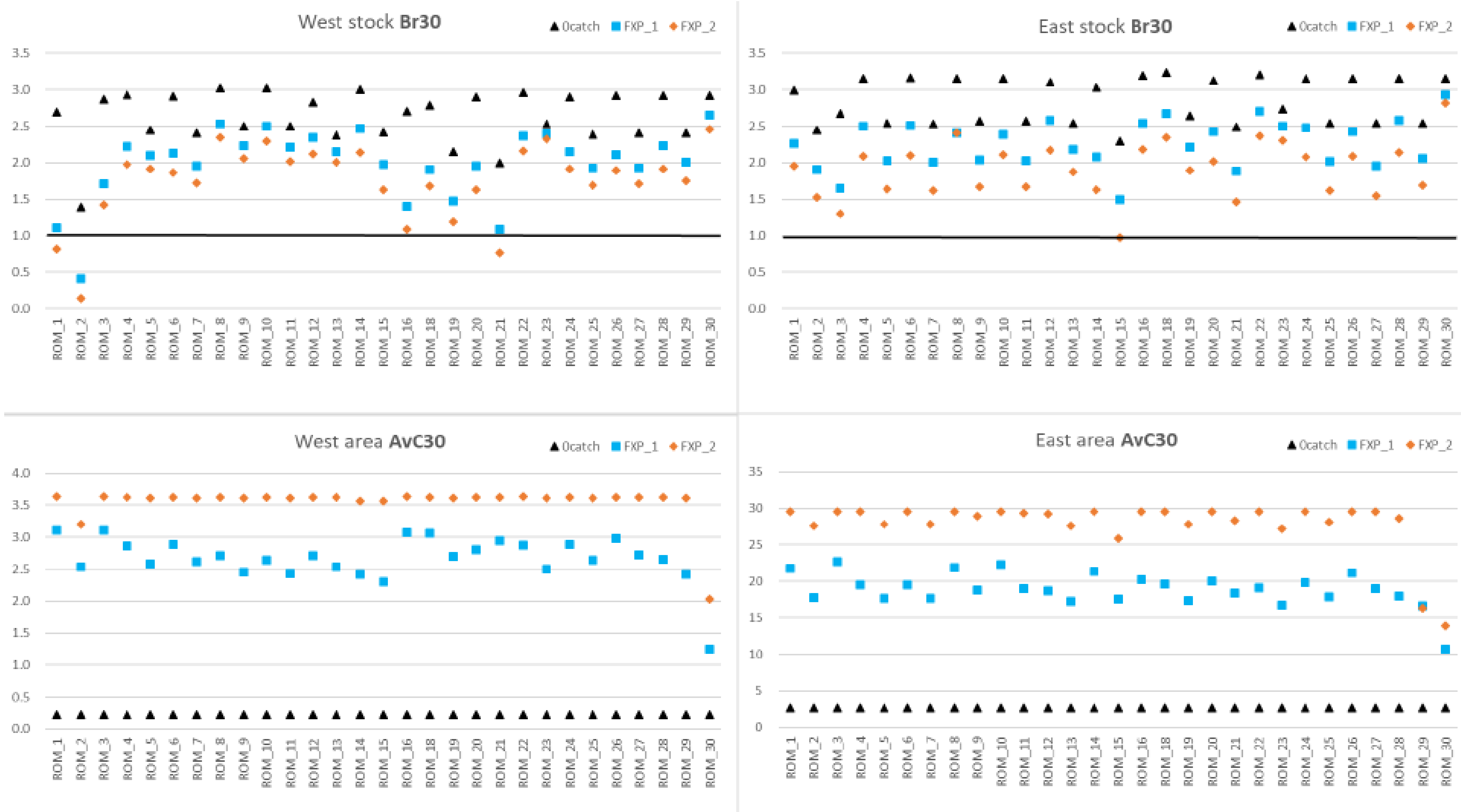

Figure C.1: Deterministic results for depletion (Br30) and average catch (AvC30) over a 30-year projection period for the "C=0" scenario and the two CMPs FXP1($\beta = 0.5, \alpha = 0.5$) and FXP2($\beta = 1, \alpha = 1$) for robustness tests ROM_1 to ROM_30 as described in Appendix C1. Note that ROM_1, ROM_2 and ROM_3 as indicated below correspond to what are termed ROM13, ROM14 and ROM15 in Chapter 3.

Appendix C2

Table C.4: Fishing fleets[8] included in the operating model, based on the selectivities of fleets active historically in the Atlantic.

No.	Name	Gear	Flag	Strata	Quarter	Start-End	Selectivity type/Bound[9] on fleet selectivity
1	LLOTH	LL	All except Japan	All	All	1964-2016	DN; 12.5 - 412.5
2	LLJPNold	LL	Japan	All	All	1964-2009	DN; 12.5 - 362.5
3	BBold	BB	EU.Spain, EU.France	Bay of Biscay (EATL)	2,3,4	1960-2006	DN; 12.5 - 237.5
4	BBnew	BB	EU.Spain, EU.France	Bay of Biscay (EATL)	2,3,4	2007-2016	DN; 12.5 - 287.5
5	PSMEDold	PS	All except EU.Croatia	MED	1,3,4	1960-2008	DN; 12.5 - 362.5
6	PSMEDoldQ2	PS	All except EU.Croatia	MED	2	1960-2008	DN; 12.5 - 312.5
7	PSMEDnew	PS	All except EU.Croatia	MED	All	2009-2016	DN; 12.5 - 362.5
8	PSNOR	PS	Norway	NATL, EATL	3,4	1964-2016	DN; 112.5 - 337.5
9	PSHRV	PS	EU.Croatia	MED	All	1991-2016	DN; 12.5 - 337.5
10	PSWold	PS	USA, Canada	ATW	2,3,4	1964-1984	DN; 12.5 - 337.5
11	PSWnew	PS	USA, Canada	ATW	All	1985-2015	DN; 62.5 - 312.5
12	TPold	TP	EU.Spain, Morocco, EU.Portugal	St. Gibrartar (SATL, MED)	All	1964-2011	DN; 37.5 - 337.5
13	TPnew	TP	USA, EU.Spain, Morocco, EU.Portugal	St. Gibrartar (SATL, MED)	2,3,4	2012-2016	DN; 37.5 - 362.5
14	CAN RR	RR	USA, Canada	ATW, GSL	All	1964-2016	DN; Logistic; 12,5 - 362.5
15	RRUSAFS	RR	USA	ATW	2,3,4	1964-2016	DN; 12.5 - 162.5
16	RRUSAFB	RR	USA	ATW	2,3,4	1964-2016	DN; 62.5 - 362.5
17[10]	OTH	other	other	All	All	1964-2016	DN; 12.5 - 362.5
18	LLJPNew	LL	Japan	WATL, SATL, NATL, EATL	All	2010-2016	DN; 62.5 - 312.5

[8]Table C.3 is from Table 3.1 in TSD (Version 20-02) with Catch and length composition by fleet are prepared by year, quarter, and strata from the revised CATDIS (Kimoto et al., 2020) and screened Task 2 Size (Carruthers, 2020).
[9]Selectivity type DN means double normal. Bounds are the middle point in a length bin (width of length bin is 25cm) that are the lowest and highest for which lengths have be observed for each fleet.
[10]This table was updated in 2020. However, the fishing fleets included in the 2019 OM from the TSD (version 19-4) is based a 17-fleet model instead of a 18-fleet model, therefore the information pertaining to row 18 here is not included in the 2019 OM. The reason the table 2019 information is not shown is because it was incomplete.

Appendix C3

Table C.5: Parameter values of Base Case and alternative options. This table specifies the numerical values for the base case and alternative options discussed in section 3.3.2.1 in the main text(Carruthers, 2019c).

Parameter	**Western stock**	**Eastern stock**
	0.6 changing to 0.9 in 1975	0.98
	0.6	0.7

Type	**Richards growth** (Ailloud *et al.*,2017)	**von Bert. Growth** (Cort, 1991)	
$A2$	34		
$L1$ **(cm)**	33.0		
$L2$ **(cm)**	270.6	L_∞ **(cm)**	318.8
K	0.22	K	0.093
p_0	-0.12	t_0	-0.97

Natural morality rate at age (Western and Eastern)

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15+
High	0.38	0.30	0.24	0.20	0.18	0.16	0.14	0.13	0.12	0.12	0.11	0.11	0.11	0.10	0.10
Low	0.36	0.27	0.21	0.17	0.14	0.12	0.11	0.10	0.09	0.09	0.08	0.08	0.08	0.08	0.07

Selectively of at least one fleet

Fleets #13 'TPnew' and #14 'CAN RR' is assumed to be logistic.

Spawning fraction

Age class	1	2	3	4	5	6	7	8	9	10	11	12	13	14+
Younger	0	0	0	0.25	0.5	1	1	1	1	1	1	1	1	1
Older (East)	0	0	0.15	0.3	0.45	0.6	0.75	0.9	1	1	1	1	1	1
Older (West)	0	0	0	0	0	0	0	0.01	0.04	0.19	0.56	0.88	0.98	1

References

Ailloud, L. E., Lauretta, M. V., Hanke, A. R., Golet, W. J., Allman, R. J., Siskey, M. R., Secor, D.H. and Hoenig, J. M. 2017. Improving growth estimates for Western Atlantic bluefin tuna using an integrated modeling approach. Fisheries Research, 191, 17–24.

Bravington, M.V., Eveson, J.P., Grewe, P.M. and Davies, C.R. 2017. SBT Close -Kin Mark-Recapture with Parent-Offspring and Half-Sibling Pairs: update on genotyping, kin-finding and model development. Unpublished CCSBT document CCSBT-ESC/1709/12.

Butterworth, D. S. 2007. Why a management procedure approach? Some positives and negatives. ICES Journal of Marine Science, 64: 613 - 617.

Butterworth, D.S., Miyagawa, M. and Jacobs, M.R.A. 2018a. Initial exploratory investigations of some simple candidate management procedures for Southern Bluefin Tuna. Paper presented to the CCSBT OMMP meeting, Seattle, USA, 18-22 June 2018. CCSBT-OMMP/1806/12.

Butterworth, D.S., Miyagawa, M. and Jacobs, M.R.A. 2018b. Further exploratory investigations of some simple candidate management procedures for Southern Bluefin Tuna. Paper presented to the CCSBT ESC meeting, San Sebastian, Spain, 3-8 September 2018. CCSBT-OMMP/1809/43.

Butterworth, D.S., Miyagawa, M. and Jacobs, M.R.A. 2019a. Results for initial explorations of simple candidate "fixed proportion" MPs for Atlantic Bluefin Tuna based on the operating models package circulated. Collective Volume of Scientific Papers. ICCAT, 75(6): 1182–1200.

Butterworth, D.S., Miyagawa, M. and Jacobs, M.R.A. 2019b. Further investigations of simple "fixed proportion" candidate management procedures for North Atlantic Bluefin using operating model package version 3.3.0. Collective Volume of Scientific Papers. ICCAT, 75(6): 1552-1575.

Butterworth, D.S., Miyagawa, M. and Jacobs, M.R.A. 2019c. Application of "fixed proportion" candidate management procedures for North Atlantic Bluefin Tuna using operating model package version 4.2.15. Unpublished ICCAT document. SCRS/2019/018.

Butterworth, D.S, Jacobs, M.R.A, Rademeyer, R.A. and Miyagawa, M. 2019d. Further exploratory investigations of some simple candidate management procedures for Southern Bluefin Tuna. Paper presented to the CCSBT OMMP meeting, Seattle, USA, 24-28 June 2019. CCSBT-OMMP/1906/11.

Butterworth, D.S, Jacobs, M.R.A, Rademeyer, R.A. and Miyagawa, M. 2019e. Simple target-based CMPs for southern bluefin tuna. Paper presented to the CCSBT ESC meeting, Cape Town, South Africa, 2-7 September 2019. CCSBT-

ESC/1909/30.

Butterworth, D.S, Jacobs, M.R.A, Rademeyer, R.A. and Miyagawa, M. 2020. Application of "fixed proportion" candidate management procedures for North Atlantic Bluefin Tuna using operating model package version 5.2.3. Collective Volume of Scientific Papers. ICCAT, 76(1): 139-143.

Campbell, R.A. 2004. CPUE standardisation and the construction of indices of stock abundance in a spatially varying fishery using general linear models. Fisheries Research 70: 209–227.

Carruthers, T.R. 2019a. ABT-MSE: Atlantic Bluefin Tuna Management Strategy Evaluation (v5.2.2) [online]. ICCAT Atlantic Wide Research Programme for Bluefin Tuna (GBYP). Available at: `file:///C:/ABT-MSE/ABT-MSE\%20User\%20Guide\%205\_2.html`

Carruthers, T.R. 2019b. ABT-MSE: Candidate Management Procedure Developers Guide [online]. ICCAT Atlantic Wide Research Programme for Bluefin Tuna (GBYP). Available at: `file:///C:/ABT-MSE/CMP\_Developers\_Guide\%205\_2.html#`

Carruthers, T.R. 2019c. Trial Specifications Document (Draft annex Version 19-4: July 7 2019). Specifications for MSE trials for Bluefin Tuna in the North Atlantic.

Carruthers, T.R. 2020. Trial Specifications Document (Draft annex Version 20-02: May 10, 2020). Specifications for MSE trials for Bluefin Tuna in the North Atlantic.

CCSBT. 2009. Report of the Operating Model and Management Procedure Technical Meeting, July 2009, Seattle, USA.

CCSBT. 2010. Report of the Third Operating Model and Management Procedure Technical Meeting, June 2010, Seattle, USA.

CCSBT. 2017a. Report of the Eighth Operating Model and Management Procedure Technical Meeting, June 2017, Seattle, USA.

CCSBT. 2017b. Report of the Twenty Second Meeting of the Scientific Committee, September 2017, Yogyakarta, Indonesia.

CCSBT. 2018a. Report of the Ninth Operating Model and Management Procedure Technical Meeting, June 2018, Seattle, USA.

CCSBT. 2018b. Report of the Twenty Third Meeting of the Scientific Committee, September 2018, San Sebastian, Spain.

CCSBT. 2018c. Report of the Fifth Meeting of the Strategy and Fisheries Management Working Group, March 2018, Canberra, Australia.

CCSBT. 2019a. Report of the Tenth Operating Model and Management Procedure Technical Meeting, June 2019, Seattle, USA.

CCSBT. 2019b. Report of the Twenty Fourth Meeting of the Scientific Committee, September 2019, Cape Town, South Africa.

The Commission for the Conservation of Southern Bluefin Tuna (CCSBT) [online]. 2020. Available at: https://www.ccsbt.org/

Cort, J.L. 1991. Age and growth of the bluefin tuna, *Thunnus thynnus* (L.) of the North East Atlantic. SCRS/1990/066. Collective Volume of Scientific Papers. ICCAT. 35 (2): 213-230.

Cramer Fish Sciences [online]. 2015. "Genetic tagging." YouTube, 21 May 2015. Retrieved from: https://www.youtube.com/watch?v=I6pkqv6B9Ug

Davies, C., Preece, A., Hillary, R. and Parma, A. 2018. Desirable Behavior and Specifications for the Development of a New Management Procedure for SBT. Prepared for the CCSBT Strategy and Fisheries Management Working Group, Canberra, Australia, 6-9 March 2018. CCSBT-OMMP/1809/BGD 01.

Francis, R.I.C.C. 2016. Growth in age-structured stock assessment models. Fisheries Research. 180: 77-86.

Fromentin, J-M. and Powers, J.E. 2005. Atlantic bluefin tuna: population dynamics, ecology, fisheries and management. Fish and Fisheries 6: 281–306.

Fromentin, J-M. 2010. ICCAT Manual. Updated 2016. International Commission for the Conservation of Atlantic Tuna. In: ICCAT Publications (online). Chapter 2.15: Atlantic Bluefin Tuna, January 2010, 1st Edition. Available at: https://www.iccat.int/en/iccatmanual.html

Glossary of Fishery Terms. 2000. ICCAT Manual. Updated 2016. International Commission for the Conservation of Atlantic Tuna. In: ICCAT Publications (online). Appendices: A7. Glossary of ICCAT Terms. Available at: https://www.iccat.int/en/iccatmanual.html

Hilborn, R. and Ovando, D. 2014. Food for Thought – Reflections on the success of traditional fisheries management. ICES Journal of Marine Science, 17(5): 1040–1046.

Hilborn, R. and Walters, C.J. 1992. Quantitative Fisheries Stock Assessment – Choice, Dynamics Uncertainty.

Hillary, R., Preece, A. and Davies, C. 2018. Initial MP structure and performance. Paper presented to the CCSBT OMMP meeting, Seattle, 18-22 June 2018. CCSBT-OMMP/1806/05: 14pp.

IUCN. 2020. The IUCN's Red list of threatened species [online]. Version 2020-1. Available at: https://www.iucnredlist.org/

The International Commission for the Conservation of Atlantic tunas (ICCAT) [online]. 2020a. Available at: https://www.iccat.int/

ICCAT. 2020b. Report of the 2020 Intersessional meeting of the ICCAT Bluefin tuna MSE Technical Group. February 2020. Madrid, Spain. In press.

Joint Tuna RFMO Management Glossary. 2018. Joint Tuna RFMO Management Glossary Strategy Evaluation Working Group Meeting. Glossary of terms for harvest strategies, management procedures and management strategy evaluation, June 2018, Seattle, USA.

Kimoto, A., Carruthers, T., Walter, J. F., Mayor, C., Hanke, A., Abid, N., Arrizabalaga, H., Rodríguez-Marín, E., Palma, C., and Ortiz, M. 2020. Summary of input data (catch, size and indices) used in the Atlantic bluefin tuna Operating Models (Version 5.2.3). Collective Volume of Scientific Papers. ICCAT, 76(1): 144-164.

Kituyi, M. and Thomson, P. 2018. 90% of fish stocks are used up – fisheries subsidies must stop [online]. World Economic Forum. Available at: `https://www.weforum.org/agenda/2018/07/fish-stocks-are-used-up-fisheries-subsidies-must-stop/`

Kolody, D and Polacheck, T. 2001. A statistical catch-at-age/length integrated analysis model for southern bluefin tuna stock assessment. Paper presented to the CCSBT SC meeting, Tokyo, 28-31 August 2001. CCSBT/SC/0108/27.

Polacheck, T., Hobday, A., West, G., Bestley, S and Gunn, J. 2006. Comparison of East-West Movements of Archival Tagged Southern Bluefin Tuna in the 1990s and early 2000s. Paper presented to the CCSBT ESC meeting, Tokyo, 12-15 September 2006. CCSBT-ESC/0609/28.

Preece, A., Eveson, P., Davies, C., Hillary, R. and Bravington, M. 2015. Report on gene-tagging design study. Paper presented to the CCSBT ESC meeting, Incheon, 1-5 September, 2015. CCSBT-ESC/1509/18.

Punt, A.E., Butterworth, D.S., de Moor, C.L., De Oliveira, Jose, A.A. and Haddon, M. 2016. Management strategy evaluation: best practices. Fish and Fisheries 17(2):303-334

Punt, A. E. and Donovan, G. P. 2007. Developing management procedures that are robust to uncertainty: lessons from the International Whaling Commission. ICES Journal of Marine Science, 64: 603–612.

Takahashi, N., Kurota, H., Sakai, O., Itoh, T. and Butterworth, D. 2015a. A check of operating model predictions with discussion of aerial survey index issues related to continuing use of the Bali Management Procedure.

Takahashi, N., Kurota, H., Sakai, O., Itoh, T. and Butterworth, D. 2015b. Some initial considerations for the review of the Bali Management Procedure in 2017.

Takahashi, N., Kurota, H., Sakai, O., Itoh, T. and Butterworth, D. 2016. A check of operating model predictions from the viewpoint of the management procedure implementation in 2016.