Oceans & Atmosphere
www.csiro.au

Revising the ETBF billfish harvest strategies

Rich Hillary, Ann Preece, Toby Patterson

7 June 2021

CSIRO Oceans & Atmosphere
Battery Point, Hobart 7001, Tasmania, Australia.

Copyright and disclaimer

Important disclaimer

Contents

1 Non-technical summary

Principal Investigator address

Rich Hillary, CSIRO Oceans & Atmosphere, Castray Esplanade, GPO Box 1538, Hobart, TAS 7001, Australia, tel: +61 3 6232 5452

Co-investigator address

Ann Preece, CSIRO Oceans & Atmosphere, Castray Esplanade, GPO Box 1538, Hobart, TAS 7001, Australia. tel: +61 3 6232 5000

Toby Patterson, CSIRO Oceans & Atmosphere, Castray Esplanade, GPO Box 1538, Hobart, TAS 7001, Australia. tel: +61 3 6232 5000

Objectives

1. Update the current suite of Operating Models for both billfish species
2. Reassess existing stock structure and migration hypotheses
3. Restructure and redesign new candidate Harvest Strategies for both species
4. Use Management Strategy Evaluation to fully assess performance of revised Harvest Strategies

Outcomes achieved
The Swordfish Harvest Strategy has been adopted and implemented, along with a suite of custom-designed Operating Models for this stock. Additionally we have provided SPC with updated estimates of movement for the upcoming 2021 full assessment of the WCPO Broadbill Swordfish stock. The current suite of stock structure hypotheses were built into the custom Operating Models and the adopted Harvest Strategy was found to be robust to those hypotheses.

For Striped Marlin, the Operating Models using the previous MSE analyses were updated using the most recent 2019 assessment for the species. A similar structure to the candidate Swordfish Harvest Strategy was explored for this species. While the TTRAG has not fully decided on which parameterisation of the harvest strategy, or whether to explore some constant catch scenarios given its quasi by-catch nature and comparative ease of avoidance from a targeting perspective.

2 Acknowledgements

The members of the SARAG for their advice, patience and suggestions on the work contained herein.

2.1 Author listing

The authors listed below contributed to one or more of the papers in the Appendices:

- **Rich Hillary**: CSIRO Oceans & Atmosphere

- **Ann Preece**: CSIRO Oceans & Atmosphere
- **Toby Patterson**: CSIRO Oceans & Atmosphere

3 Background

The recently revised Commonwealth Harvest Strategy Policy states that if Australia is a major harvester of an international stock and no harvest strategy has been determined internationally, the AFMA must develop and implement a harvest strategy consistent with the objective of the Commonwealth Harvest Strategy Policy. The harvest strategies for Broadbill Swordfish and Striped Marlin have been in use in the ETBF for a number of years. Recent MSE work has shown that, for Broadbill Swordfish, the HS as currently formulated does not perform satisfactorily in either of the key connectivity scenarios. For an essentially isolated ETBF population, the HS cuts the TAC whilst the spawning stock is above the MEY proxy target level; for a connected population across the WCPO it cuts the catches to very low levels with little to no benefit to the wider WCPO spawning population. For Striped Marlin the notion of being able to construct 3 series (small, prime and large) for use in the current HS has been shown to be logically unsound, given there is evidence to suggest a limited number of cohorts present in the ETBF catches.

This project revised the current Harvest Strategies (HS) in light of these two quite different issues for the two billfish species. For Broadbill Swordfish we reconditioned and revised the current suite of Operating Models (OMs), re-evaluated the estimated levels of mixing across the WCPO, and revised the current HS (both in terms of structure and inputs indices). For the Broadbill Swordfish work we created custom-designed OMs that are directly fitted then to the key data sources, not conditioned from the outputs of the WCPFC stock assessments as before. The reasons for this were three-fold: (i) we envisaged a level of potential population structure, given the potential outcomes of the genetics project, required in the OMs that is not currently possible within the MULTIFAN-CL stock assessment model; (ii) we envisaged a need for more flexible movement models to adequately cover the range of plausible migration hypotheses; and (iii) we required the flexibility to include or exclude catch data for the various regions that has been, hitherto, difficult to get done by the assessment team at SPC across the whole suite of uncertainty grid elements. For Striped Marlin we reconditioned and revised the existing OMs (conditioned directly from the revised stock assessment in 2019) and fundamentally restructure the new HS in light of the new reality of what exact year-class coverage we have in the ETBF data, cognisant of the existing level of uncertainty in the growth relationship for this species. For both species a re-evaluation of the current mixing hypotheses was done using the most up-to-date data (tagging) and consideration of outcomes from the current FRDC stock structure project (looking at the five ETBF quota species) as our OMs needed to be significantly revised if any structure below the currently assumed panmictic WCPO stock is uncovered. There was also the issue of what catches within the wider Pacific Ocean are in fact taken from the WCPO population (specifically large recent catches in the far North East of the WCPO), as they have a huge influence on the MSE outcomes (in terms of both ETBF catches and future stock status).

4 Need

Both billfish species are currently managed using fully evaluated harvest strategies and, under the recently revised Commonwealth Harvest Strategy Policy, this remains best practice and something that all key scientists and stakeholders wish to continue. The international element of the problem (the ETBF is part of a wider fishery and population throughout the WCPO and, as such, also falls under the remit of the WCPFC) adds to the complexity of the issue, but previous work has shown that the domestic management of a single fishery on a potential subset/sub-stock of/in the population can dovetail with the international management of the whole population and fisheries thereon.

5 Objectives

1. Update the current suite of Operating Models for both billfish species
2. Reassess existing stock structure and migration hyotheses
3. Restructure and redesign new candidate Harvest Strategies for both species
4. Use Management Strategy Evaluation to fully assess performance of revised Harvest Strategies

6 Benefits/Management Outcomes

The outcomes of the project will form the basis for TACC setting and the management of these two stocks in the ETBF for the coming years, consistent with the requirements of the Commonwealth Harvest Strategy Policy.

The updated estimates of movement for both Broadbill Swordfish and Striped Marlin not only made the stock structure and movement hypotheses used in the MSE work more informed but also made a significant contribution to the updated WCPFC stock assessments for both these species.

7 Conclusions

Meeting the project objectives:

Objective 1

"Update the current suite of Operating Models for both billfish species"

Appendix 4 details the custom designed OMs for Broadbill Swordfish. Appendix 6 details the updated OMs for Striped Marlin.

Objective 2

"Reassess existing stock structure and migration hypotheses"

Appendix 7 details the updated movement estimates for Striped Marlin using the most recent available tagging data (electronic and conventional). Appendix 4 details how the revised stock structure and migration were outlined and implemented in the Broadbill Swordfish OMs.

Objective 3

"Restructure and redesign new candidate Harvest Strategies for both species"

Appendices 1 and 2 detail the processes whereby the general structure of the revised HS for both species was explored and presented to the TTRAG for discussion and eventual adoption.

Objective 4

"Use Management Strategy Evaluation to fully assess performance of revised Harvest Strategies"

Appendices 3 and 5 detail the Broadbill Swordfish MSE work. Appendix 6 details the Striped Marlin MSE work.

Appendix 1

Oceans & Atmosphere
www.csiro.au

Issues and plan for billfish harvest strategy revision project

R. Hillary, A. Preece, T. Patterson

20 March 2019

CSIRO Oceans & Atmosphere
Battery Point, Hobart 7004, Tasmania, Australia.

Copyright and disclaimer

Important disclaimer

Contents

1 Background

This paper is just to outline high-level issues and plans when it comes to updating the billfish harvest strategies. The general items we cover off on are:

- Revising the suite of Operating Models (OMs) we intend to use to test the candidate harvest strategies
- What kind of (potentially species-dependent) harvest strategy structures are we going to explore for both broadbill swordfish and striped marlin
- What are the operational features for the harvest strategies: maximum and minimum TAC changes, possibility for multi-year TACs and - if so - within TAC-block constraints (over/under rules)
- Performance measures to assess HS performance - not just target and limit reference probabilities but inter-annual variation in catch and so on. What features do we want in the harvest strategy basically?

2 Operating Models

Previous MSE work has followed the same general approach for both billfish species: use the WCPFC stock assessments to condition OMs where ETBF catch is managed via a candidate HS and non-ETBF fishing mortality (a proxy for effort) is assumed to be a constant multiple of the recent level. Within these OMs we have explored different hypotheses (e.g. connectivity and migration between ETBF and non-ETBF) areas but always - by necessity - within the confines of what is included within the WCPFC assessments themselves.

2.1 Broadbill Swordfish

For swordfish, we propose a different approach to conditioning the OMs for the MSE analyses. Over the last several years and the last two stock assessments it has become clear that there are several hypotheses that we either are struggling to, or simply cannot, parameterise using the WCPFC stock assessments:

1. Alternate migration models (in terms of magnitude and/or size/age structured) for swordfish across the WCPO, and a reappraisal of the tagging data analyses used to fix migration values in the assessments
2. Stock structure - this is currently impossible to do within MULTIFAN-CL but could well be an outcome of the current FRDC stock structure project
3. Fishery structure: we are essentially beholden to how SPC and WPCFC do this, and their goal is the accurate characterisation of all the relevant (and different) fleets within the WCPO. For us, we simply care about the ETBF and the non-ETBF meta-structure and removing the fish of the right size classes from the population. This also links to the inclusion of catch, particularly in the northern region of Area 2 in the current assessment, that are both highly influential on the assessment outcomes and may not be part of the wider WCPO stock complex (if there is such a thing)

To get around these issues we are going to construcut a custom-designed OM and fit it directly to the same WCPO data as used in the assessment. In terms of fleet structure, obviously we will have the ETBF fleet as a singular fleet. For the others we propose the following: (i) within a given spatial region, if a given fleet has an abundance index that is usable (i.e. assumed constant catchability) this fleet remains a unique fleet in the OM; (ii) in a given spatial region, if a given fleet does not have a usable abundance index group it together with other such fleets that have size or weight frequency composition data and use flexible time-varying selectivity models to model these fleet groups.

2.2 Striped Marlin

The main issue over the last few years with the striped marlin OM has mainly been about it being out of date (the last assessment was done in 2012). Additionally some concerns have arisen that there may be uncertainty in the current growth curve assumed for this stock (J. Farley, pers. comm.). SPC have indicated that they are going to update this assessment this year, and project scientists are going to be in Noumea and able to talk with the analysts in April 2019. We propose to use the revised stock assessment for striped marlin as the basis for the reconditioned OMs for the MSE work.

3 Candidate Harvest Strategies

The MSE work from 2018 [1] demonstrated that the previous HS did not appear to be performing as expected - cutting catches in the ETBF even with the SSB above the target level. Additionally, various members have raised a number of issues with structure of the HS (See Figure 3.1):

1. That, even for level 1, the slope-to-target HCR is both hard to understand and can lead to decreases in TAC even when above the target level
2. That it is hard to understand how a decision 'flows' from level 2 to level 4
3. That there are discrete shifts in the TAC multiplier that happen for arguably small changes in the input indices. As a result, minor changes to data and/or analysis, resulting in tiny changes to the input indices, *could* lead to a 10% difference in the TAC outcome.

3.1 Broadbill Swordfish

For swordfish, the cohort decomposition work [2] suggested that, with modifications to the cut-off points used, we could still construct meaningful CPUE indices for small, prime and large fish. As before, we propose that the prime index is the major driver of the HCR, but we will explore options for including the other two indices (small and large). As for the HCR using the prime index, we propose to do away with the slope-to-target structure of the previous HS and move to a target/limit 'buffered' type HCR as shown in Figure 3.2. The HCR is a multiplier on the current TAC to give the TAC for the next 'block' (if multiple years allowed) and can be simply described:

- If the index is within $\pm x\%$ of the target: **no change** (and $x = 10$ in Figure 3.2)
- If the index is above the upper target buffer: **linearly increase TAC**
- If the index is below the lower buffer, but above the limit level: **linearly decrease TAC**
- If the index is below the limit level: **strongly decrease TAC**

Harvest strategies using these kinds of HCR have been shown to perform consistently better than their slope-based counterparts across a variety of different life-history and fishery status

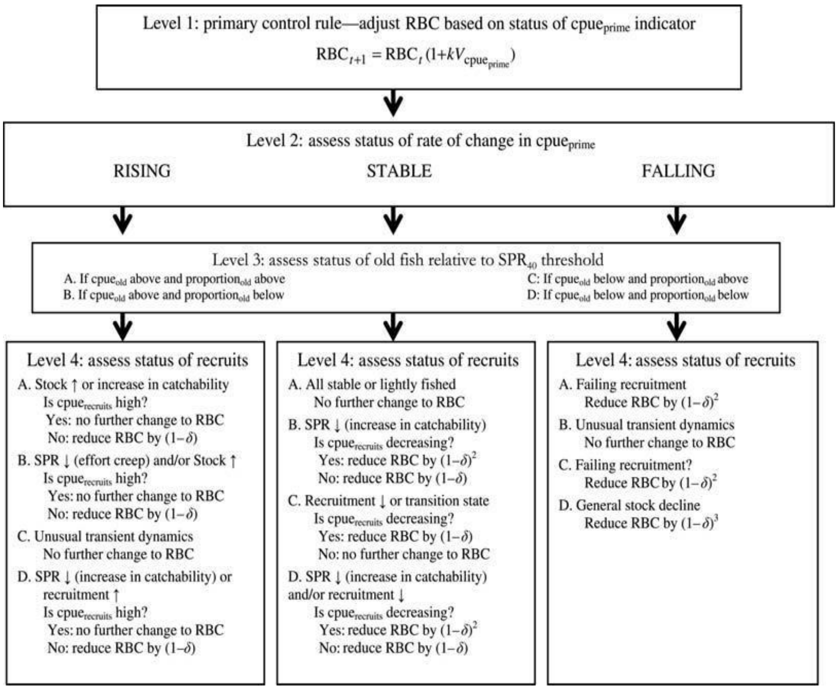


Figure 3.1: *The previous ETBF harvest strategy structure.*

contexts [3] - especially when the parameters of the HCR are tuned (which we intend to do).

When considering using the small and/or large CPUE indices we will explore some simple terms that look to modify the TAC using either the trend or the relative proportion of the small or large CPUE.

3.2 Striped Marlin

The work in [2] cast considerable doubt on whether it as either possible, or even sensible, to decompose the striped marlin CPUE into three distinct indices of non-overlapping cohorts. Given this, and even with some uncertainty as to the growth rates of striped marlin to be explored in the OMs anyway, we propose to explore harvest strategies using a single CPUE index with an HCR such as that proposed in Figure 3.2.

4 Operational features and constraints

The key harvest strategy 'settings' and operational constraints would be something along the lines of the following:

- Frequency of TAC changes: currently annual, but do we explore 2-year blocks? Trade-off reactivity with TAC stability
- Maximum and minimum TAC changes: what percentages do we explore for these two values
- If we do explore multi-year TAC blocks what kind of flexibility to we permit for within-block catch differences. Specifically, we define the TAC block as the TAC multiplied by the number of years it is to be set for and, as long as the total catch taken does not exceed this

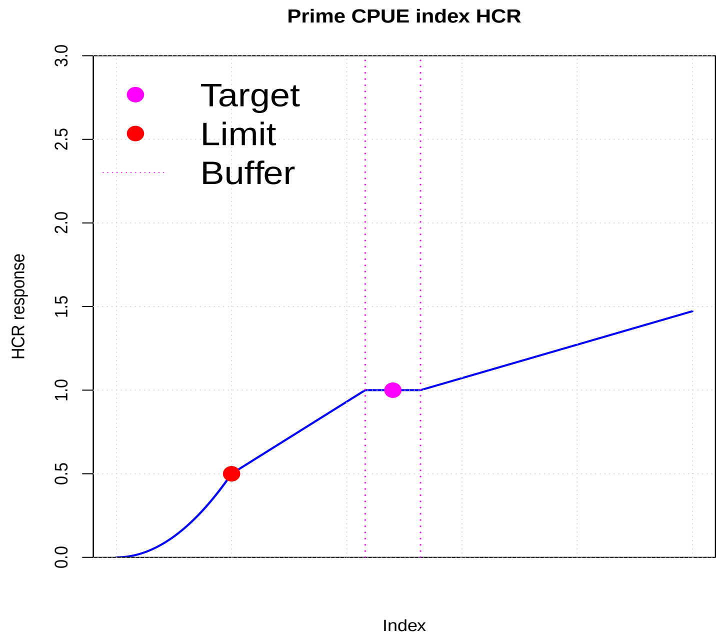


Figure 3.2: *Generic form proposed for the prime CPUE index part of the HCR.*

amount, we permit some freedom to take more or less in some years

5 Performance measures

Previous performance measures have focussed mostly on SSB performance (relative to the target/limit levels of the Commonwealth HSP), average catch levels and so forth. We will obviously include these again but we propose to explore some additional measures such as average interannual/block varition in the TAC.

6 Discussion

This paper, purposefully non-technical in nature, attempts to outline the issues facing, and the plan of action for, the revised ETBF billfish harvest strategy project. We have focused on how we intend to update the current suite of Operating Models, general design features of our envisaged suite of candidate harvest strategies, discussion points around operational setting and constraints, and harvest strategy performance measures. The paper is intended to restart discussions that have been had at previous TTRAG meetings and to debate and refine the ideas that we have outlined at a high level. The technical details will follow as the project gets into full flow after this RAG meeting.

References

[1] R. Hillary (2018) Updating the ETBF Broadbill Swordfish MSE analyses . *TTRAG, Canberra, ACT*.

[2] R. Campbell (2018) Identifying Cohorts in the Eastern Tuna and Billfish Fishery and application to the Harvest Strategy. *TTRAG, Mooloolaba, QLD*.

[3] T. Carruthers *et al.* (2016) Performance review of simple management procedures. *ICES J. Mar. Sci.* **73**(2): 464–482.

CONTACT US

t 1300 363 400
+61 3 9545 2176

e csiroenquiries@csiro.au

w www.csiro.au

WE DO THE EXTRAORDINARY EVERY DAY

We innovate for tomorrow and help improve today for our customers, all Australians and the world.

Our innovations contribute billions of dollars to the Australian economy every year. As the largest patent holder in the nation, our vast wealth of intellectual property has led to more than 150 spin-off companies.

With more than 5,000 experts and a burning desire to get things done, we are Australias catalyst for innovation.

WE IMAGINE. WE COLLABORATE. WE INNOVATE.

Appendix 2

Oceans & Atmosphere
www.csiro.au

Exploring some initial harvest strategy parameters for ETBF billfish MSE work

R. Hillary

16 July 2019

CSIRO Oceans & Atmosphere
Battery Point, Hobart 7004, Tasmania, Australia.

Copyright and disclaimer

Important disclaimer

Contents

1 Background

This document outlines some initial general MSE work to explore and explain the likely ramifications of certain choices for the revised harvest strategy - particularly around the frequency with which TAC changes are made, initial starting status of the stock, and the relative effort share between the ETBF and non-ETBF fleets.

2 ”Toy” Operating Model for broadbill swordfish

Using all the current information from the stock assessment (natural mortality, growth, maturity, weight-at-age, selectivity) we constructed a kind of “toy” OM with all the important characteristics of broadbill swordfish. We make simplifying assumption that there is a single region with the ETBF and non-ETBF fleet having the same selectivity and a pre-defined initial relative effort share (and future trend therein for the non-ETBF fleet) that defines the relative impact of the two fleets in the projection.

Here are the key OM scenarios we explored:

- **Initial stock status**: 30%, 48% and 60% options were explored to mimic a recovery, *status quo*, and under-utilised scenario (assuming 48% as the target SSB depletion level)
- **ETBF effort share**: 100%, 75% and 50% options were explored to cover a (currently) plausible range of values that would be useful in determining general behaviours of the candidate HS

The OMs were run for a 5 year ’burn in’ period to generate CPUE data for use in the main 20 year projection period for when the HS was in action. The single CPUE index was generated for an arbitrary catchability value $q = 10^{-5}$ and using the ETBF selectivity-at-age with a CV of 0.1 in terms of observation error.

3 Candidate Harvest Strategy

We explored one particular formulation of the kind of target-limit HS we outlined at the last TTRAG meeting which uses the prime CPUE index. Figure 1 outlines the general form of the candidate HS explored herein. The HCR has a target CPUE level and a buffer zone around the target (set at 10% either side) to avoid needless tinkering of the TAC around the target level and within the innate stochasticity of the CPUE index used as input to the HCR. Either side of the target, and above the limit level, the HCR linearly decreases or increases the TAC depending on whether the mean recent CPUE is below or above the target/buffer-zone region. In this instance the rate of increase or decrease above or below the target level is the same. Below the limit level the TAC then decreases quadratically towards zero.

It should be noted that, for the upcoming MSE work, a number of these features will be explored. The size of the buffer zone promotes stability around the target, but can decrease reactivity for rapidly evolving scenarios. The gradient of the linear lines above the limit level need not be symmetric - certain scenarios (like recovery phases) can require larger TAC decreases when below the target than increases when above it. The target CPUE here is chosen to be the CPUE level that is expected at 48% SSB_0 and the limit level is chosen to be the 20% analogue of this target CPUE (i.e. $0.2/0.48$ of $CPUE_{\mathrm{targ}}$). One of these could actually be tuned in the MSE work, depending on the real-world OMs and time-frames over which the HS is evaluated.

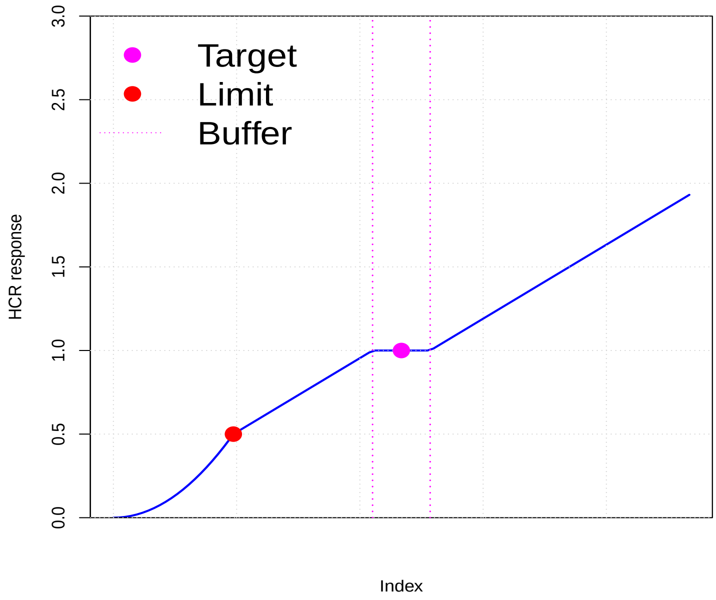


Figure 3.1: *Candidate harvest strategy explored in this exploratory MSE work.*

4 Operational features and constraints

The key harvest strategy 'settings' and operational constraints would be something along the lines of the following:

- Frequency of TAC changes: currently annual, but do we explore 2-year blocks? Trade-off reactivity with TAC stability
- Maximum and minimum TAC changes: what percentages do we explore for these two values

For this work we focussed mostly on the frequency of TAC decisions, and explored a simple minimum change of 2% and maximum change of 20% for the TAC, relative to the previous TAC. For the TAC frequency we explored 1, 2 and 3 years.

5 Performance measures

We explored some obvious but informative statistics:

- The SSB depletion at the end of the 20 year projection period - how well did the HS do at attaining the primary goal?
- Average catch over the projection period
- Interannual average variation in catch (AAV) - a measure of how 'reactive' an HS is
- The relative (to the target) CPUE at the end of the projection period - how well did the HS do in attaining the CPUE target?

6 Results

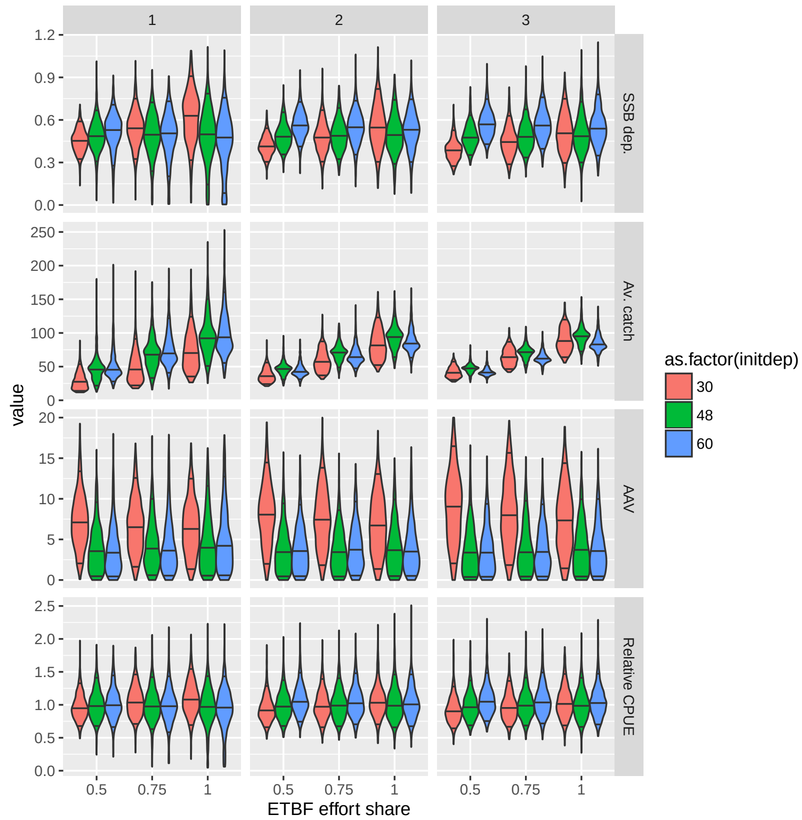


Figure 6.1: *Violin plot summary of the key statistics for the various OM and operational constraint options. The horizontal panels denote the 3 TAC frequency options and the different colours reflect the three different starting stock status scenarios.*

We ran the MSE for the 5 year burn in and 20 year active management phase for the three ETBF effort share options (100%, 75% and 50%), the three initial starting status options (30%, 48% and 60%), and the three TAC frequency options (1, 2 and 4 years). Figure 6.1 summarises the key performance statistics across the 27 different scenarios.

6.1 Role of relative ETBF effort share

In terms of conservation outcomes, it is reasonably clear that for the range of options explored (from 50 to 100% ETBF effort share) that as ETBF effort share decreases, the probability of being above the target level is lower more often when the ETBF effort share is lower, but not in such a manner as to suggest that the effort shares of 50% are pathologically bad for the efficacy of the HS tested. Indeed, some scenarios for 1 and 2 year TAC changes appear to show overshoot of the target (probability of more than 50% of being above the target) resulting in lower average catches than could be taken for these combinations.

6.2 Influence of the starting status

The harvest strategy seems to be clearly able to get the SSB moving towards the target when below it or above it and doesn't seem to result in oscillatory dynamics of going either side of the target for the time frames considered here. For the recovery scenario (initial status of 30%) the target is consistently approached from below so catches are not being reduced beyond levels that would result in a recovery. For the under-utilised scenario (initial status of 60%) the catches are increased but the target is approached from above, thereby not increasing catch so far that future decreases would be needed to maintain the target SSB and CPUE.

6.3 Influence of frequency of TAC change

The frequency of TAC changes appears to have the most obvious impact on performance. Specifically: having a TAC change every year results in fair higher variability in average catch, notably lower tails in the final SSB depletion and CPUE (indicative of instances of excessively high catch trajectories arising over the projection period). When we look at the 2 and 3 year TAC cycles we see this variability in catch and poor performance relative to the SSB and CPUE limit levels disappear. This is because of a mismatch in the time-scale for which *meaningful* signals are detectable in the CPUE series given the life-history, fishery characteristics, and observation error in the index and the freedom of the HS to change things. If the HS can change every year it is more susceptible to simply following the noise not the signal. Even moving to a decision every two years means the HS is using an index that has repeat observations of a mix of age classes and is more likely to receive a meaningful change in the index when not trying to do something every year. There is little to choose between the 2 and 3 year options looking in detail at the statistics as they have similar conservation outcomes, average catches, and AAV (reactivity) statistics. There are obviously more socio-economic effects of 2 relative to 3 year TAC changes that are currently outside the scope of the MSE work, and in the full MSE testing we might see that having 2 year TACs performs better on more pessimistic robustness tests than having a 3 year cycle.

7 Discussion

In this paper we have undertaken some exploratory MSE analyses to look at general outcomes of different high level parts of the MSE and HS structure process: starting depletion level, relative effort share of the ETBF, and a 1, 2 and 3 year TAC decision making cycle. For one instance of the proposed prime CPUE driven HS we see that the HS performed satisfactorily for the range of starting status values and relative ETBF effort shares - conservation outcomes were acceptable, there was no clear overly strong reductions or increases in catch for the different starting status levels and SSB and CPUE targets were approached from where the starting level happened to be.

The biggest effect was the TAC cycle. For a 1 year cycle we tended to see greater variability in average catch, and notable breaches of the SSB limit level and those got worse the bigger the relative ETBF effort share. This is a fairly clear signal that a 1 year TAC cycle, given the fishery and life-history characteristics, is chasing the noise to a level that is effecting performance. When moving to a 2 or 3 year TAC cycle this effect disappeared and the HS performed better. There was little to choose between the 2 and 3 year options in the narrow scope of this work but that is likely to change in the actual MSE work.

References

[1] R. Hillary (2018) Updating the ETBF Broadbill Swordfish MSE analyses . *TTRAG, Canberra, ACT*.

[2] R. Campbell (2018) Identifying Cohorts in the Eastern Tuna and Billfish Fishery and application to the Harvest Strategy. *TTRAG, Mooloolaba, QLD*.

[3] T. Carruthers *et al.* (2016) Performance review of simple management procedures. *ICES J. Mar. Sci.* **73**(2): 464–482.

CONTACT US

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

WE DO THE EXTRAORDINARY EVERY DAY

We innovate for tomorrow and help improve today for our customers, all Australians and the world.

Our innovations contribute billions of dollars to the Australian economy every year. As the largest patent holder in the nation, our vast wealth of intellectual property has led to more than 150 spin-off companies.

With more than 5,000 experts and a burning desire to get things done, we are Australias catalyst for innovation.

WE IMAGINE. WE COLLABORATE. WE INNOVATE.

Appendix 3

Oceans & Atmosphere
www.csiro.au

Update on MSE work and questions to TTRAG on objectives and operating features of revised harvest strategy

R. Hillary

11 Dec 2019

CSIRO Oceans & Atmosphere
Battery Point, Hobart 7004, Tasmania, Australia.

Copyright and disclaimer

Important disclaimer

Contents

1 Background

This paper summarises the progress that has been made on the MSE work - specifically conditioning the various Operating Models (OMs) - as well as putting some questions to the TTRAG members about key features they would like to see in the candidate Harvest Strategies (HSs).

2 MSE work

There are several key decisions effectively already made for the MSE work:

- For the Swordfish we will construct and condition a custom-designed OM to explore alternative migratory hypotheses and fishery inclusions within the model (i.e alternatives to those used in the last stock assessment), as well as some key future robustness tests for the HS
- For the Striped Marlin we will use the revised assessment done in 2019 to condition the previous OMs used in the MSE work
- In terms of indices to be used:
  - For the Swordfish we will use the prime CPUE series as well as exploring the potential utility of a recruitment index from the small fish (to be presented at the March TTRAG meeting)
  - For the Striped Marlin we will use the single revised "prime" index as defined in [2]

At the July TTRAG meeting the generic form of the main part of the HCR proposed to be used for the key indices for both species was presented (see Figure 2.1). It represents a specific variation of a general class of single-index target-threshold HCR explored across a wide variety of scenarios and life-histories in [3]. The key variations are (a) asymmetry above and below the target if required, and (b) a buffer zone around the target level to avoid simply jumping around driven by noise even if the target is effectively reached. Its main advantages, over the previous HS, is that it is continuous in response (there are no jumps in recommended TACs possible for small changes in the index) and is far easier to interpret than the previous decision tree approach.

A very important point we will make many times in the MSE process is this: avoid the temptation to strongly interpret the target level of the index used in the HCR as being somehow synonymous with the suite of possible objectives (in terms of the relative SSB) we explore in terms of assessing the performance of the HSs. It has been demonstrated that there is often a complex relationship between the prime CPUE (for Swordfish in particular) and the SSB as used in the OMs. The HCR structure will be parameterised (tuned) to adjust the ETBF catch (within operational constraints) to move the future SSB to a specific range or value in a given timeframe; the target parameter in the HCR is **not** the CPUE equivalent of the equilibrium SSB at the target objective.

2.1 Swordfish

Movement and reproductive connectivity are two key issues that either cannot yet be addressed within the MFCL stock assessment (reproductive connectivity), or haven't really been addressed satisfactorily thus far (movement scenarios). Additionally, the TTRAG has often questioned the inclusion (or exclusion) of certain fisheries in the non-ETBF spatial region and the resultant im-

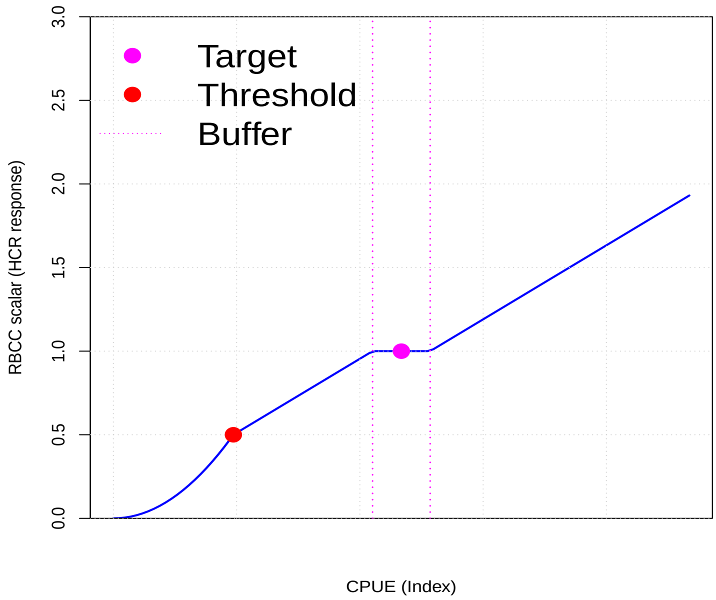


Figure 2.1: *Candidate (primary part of the) HCR to be explored in the MSE work.*

pact on the status of the stock. These three factors were the motivation behind constructing a custom-designed OM to be used in the revised HS MSE work - all these factors have a clear impact on the performance of the OMs but, if we are using what comes from the WCPFC assessments, we have no ability to either include them or define them to our satisfaction.

Unfortunately, we cannot really nail down the reproductive connectivity question given the ongoing work to source enough samples from across the geographic range. This is something that is hopefully going to be achieved but not before March 2020 when we need to consider whether we are going to adopt one of the candidate HSs or not. There are obviously a number of potential hypotheses that could be explored but we suggest to explore these three in the MSE work:

1. Panmictic across the assessment regions with one spawning population in both the regions used in the assessment and MSE work (basically current default)
2. Possibly one genetic population but with spawning populations in each of the regions (not one whole-of-region population) with possible migratory linkages
3. Fully separate “Region 1” population with migration levels and/or differential fishing mortality rates that allow us to consider this essentially an isolated population West of 165E

This will tease out, hopefully, any high-level impacts of some kind of stock structure on the performance of the HS. It is not the job of this project to explore the wider impacts of stock structure on the possible status of whole-of-regiom swordfish abundance and mortality. At least for the part of the population the ETBF harvests, these effects will be apparent in the OMs but we stress that we are **not** performing an alternative assessment of swordfish in the region. We are trying to robustly and informatively condition OMs to use in assessing the relative performance of the candidate HSs we consider.

It’s worth considering the potential difference (in terms of HS performance) of two different hypotheses: (i) panmictic with some movement in terms of migration but one biological and genetic population, (ii) separate genetic populations in each region with little to no migratory movement and no reproductive transfer. In (i) the performance of the HS will be indelibly linked to what happens in both regions: even if we do the right thing in the ETBF when responding to signals in abundance indices, if the fleets in the other region do not and drive down the stock we are going to see worse outcomes from the HS in relation to ETBF catches. In (ii), because we are focussed squarely on our region given it is distinct from its neighbour we have more leverage (what the ETBF catch does dictates the SSB dynamics more than in scenario (i)) and less dependence on fleets we cannot control. So, while these two scenarios can be very close in terms of *historical* dynamics, the projection outcomes in terms of HS performance can be *very* different. This was not really something we could easily look at properly before given MFCL does not permit you to have region-specific spawning populations at present.

In terms of movement scenarios, it might be informative to TTRAG members if what is actually currently included in the WPCFC suite of options is outlined clearly. Three scenarios are labelled as: 0, 11% and 25% **quarterly** diffusivity coefficents are the main scenarios for movement across the 165E boundary. There are some implications of these assumptions that might not be clear given diffusivity rates are used but everyone thinks of movement (including myself) in terms of the fraction of animals that move from one region to another over a given time period:

- **Symmetric diffusion rates does not mean symmetric movement**. Multifan-CL has a unique way of defining movement via random diffusive movement and the size of the box (relative to the others). As region 2 is far bigger than region 1 in the WCPFC swordfish assessment, equal diffusive rates effectively means proportionally **more** movement of animals from region 1 to region 2 than move from region 2 to region 1.
- The **quarterly** diffusivity rates assumed in the assessment imply something quite different in terms of **annual** probabilities of animals moving from one region to another. For the 11% case the annual fraction of animals moving from region 1 (ETBF zone) into region 2 is 24%; from region 2 to region 1 is 8%. For the 25% case region 1 to region 2 is 30 and region 2 to region 1 is 12%. For the case currently considered the diagnostic case in the WCPFC assessment (11%) 24% of the animals in region 1 will move to region 2 at some point within a year.

These scenarios bear little relation to the current (admittedly uncertain) understanding of swordfish migratory dynamics in the region. To the East of 165E there appears to be more latitudinal and longitudinal movement than to the West of 165E. In both regions there appears to be seasonal movement between higher latitude foraging areas and lower latitude spawning areas. There seems to be far more movement *within* regions (defined by 165E boundary) than *between* them. The tag data do not explicitly rule out across-165E movement but suggest it is potentially limited and does not at all suggest that there is 3 times as much going W to E (region 1 to 2) as going E to W (region 2 to 1). The work currently underway in the MSE project is to reassess both if the updated set of tagging data say anything substantively different and get some of upper bound on across-165E migration rates at the annual scale we are working with in the OMs. In the WCPFC assessment the intent is to explore the impact of different levels of movement between the two regions on overall stock status. This project is very much focussed on the degree to which connectivity (be it genetic or migratory) between the two regions affects the performance of the candidate HSs. Those are very different requirements and makes it incumbent on us to best reflect the reality of those dynamics in the MSE work which is why we are effectively discarding the current WCPFC assessment scenarios and developing more quantitatively informed ones.

A full description of the swordfish OM will be given both to and at the March TTRAG but for now we outline where it is similar and where it differs from the most recent WCPFC assessment:

- **Catch**: numbers are used as known inputs not estimated and assigned very small standard errors as done in the MFCL assessment
- **Effort & catchability**: we don't include effort as a covariate for fishing mortality (we use harvest rates given abundance and catch in numbers) and we don't model time-varying catchability for fleets where we don't have a usable CPUE series. This means no effort deviations or catchability parameters dramatically reducing the number of parameters from 000s in the MFCL model to just over one hundred in our OM. This massively reduces run time from hours (for one MFCL grid configuration) to less than 3 seconds. We don't care about the quality of the effort-to-fishing mortality relationship or the catchability for fleets where we have no CPUE index (and are saying that effort is not very meaningful anyway). We can drastically reduce the number of parameters without reducing the complexity of the model by doing this
- **Biology**: use same updated growth, maturity, weight-at-length relationships and steepness values. We propose to reduce the number of natural mortality scenarios from 6 to 2–3 given the early high M scenarios are implausible in a life-history sense
- **Population dynamics**: age-based and size and weight-structured to deal with composition data. We use an annual, not a quarterly, time-step and account for time of year of fishing in catch equations to deal with seasonal dynamics. Recruitment is allowed to be spatiotemporally correlated and treated as a random effect with the correct value of overall variability and the two correlation parameters estimated
- **Migration**: modelled annually, not quarterly, and modelled not using diffusivity but via a transition matrix defined via probabilities of moving from one region given you are currently in another region. Movement is allowed to be asymmetric between regions
- **Length & Weight composition**: we do not use quarterly length and weight data we aggregate (using relative sampling intensity) across quarters to get annual data to reduce seasonal variability. We also remove the somewhat arbitrary minimum and maximum weights applied in the WCPFC assessment and use statistical theory to estimate the correct weighting for each fleet's composition data at run time. This is currently accepted best practice in integrated assessment modelling and used widely in Australia
- **CPUE indices**: as with the composition data we aggregate the CPUE indices from the quarterly to the annual level using relative effort by quarter to reduce the high levels of intra-annual variability in some indices. The same indices are used but, as with length composition, instead of semi-arbitrary subjective weightings (via the effort deviation penalties) in the WCPFC assessment we use statistical theory to estimate the correct weighting for each CPUE index

To demonstrate how the custom swordfish OM performs we show a particular scenario: basically the diagnostic case with zero movement between areas and *all* of the 13 fisheries (including the Northern distant water fleet in area 2) included in [4]. Estimated parameters are: recruitment parameters (total effort deviates, spatiotemporal correlation and overall variance, nad mean spatial recruitment fraction by region), selectivity for each fishery, recruitment for each year and region, over-dispersion values for length and weight frequency by fleet, CPUE process error by fleet. As mentioned, we do not estimate effort deviations or time-varying catchability for each fleet that doesn't have a constant q assumed.

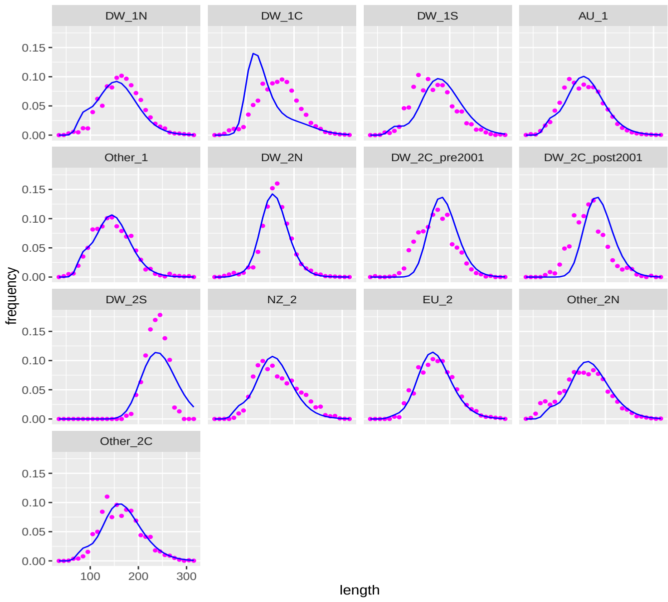


Figure 2.2: *Aggregated fits to the length frequency data: magenta dots are the observed data, blue lines the model predictions.*

Figure 2.2 shows the fits to the length frequency data aggregated across years, and Figure 2.3 shows the fits to the weight frequency data als aggregated across years. Figure 2.4 shows the fits to the CPUE indices. For readers familiar with, or with a copy of, the most recent WCPFC assessment [4] we fit to the length frequency data as well and to the weight frequency data arguably a little better for certain fleets. It is a little harsh to compare our fits to the CPUE data, relative to the assessment, given we have already smoothed the series by averaging across quarters. The series given the highest statistical weighting are the Australian ETBF in region 1 and region 2 the EU fleet with the noisy distant water fleets (in both regions) getting lower statistical weighting.

Figure 2.5 shows a general status summary for the example Swordfish run (zero movement, diagnostic M and steepness values, all fisheries included). In terms of dynamic SSB depletion (given recruitment estimates etc. what the model estimates divided by the SSB in the absence of fishing) the most recent estimates are just above 60% in total, around 73% in region 1 and around 55% in region 2. Total recruitment shows a similar pattern to the WCPFC assessment [4] with a gradual increasing trend from the 50s to the 80s and more variability from the 1970s onwards (more data), a gradual decrease from 1990 to around 2010 with some higher values in the most recent years. The annual trends in region harvest rate multiplier also mirror quite closely those in the WCPFC assessment (not surprisingly given the similarity in both biomass and recruitment trends). In region 1 values gradually increased then rose sharply in the mid 1990s as the ETBF effort and catch increased, decreasing in the late 2000s onwards as effort and catch declined in the ETBF. Region 2 shows a characteristic high spike around 1970 and again around 1980 followed by a sustained increase from just before 2000 onwards as the large distant water catches in the North of the region appear.

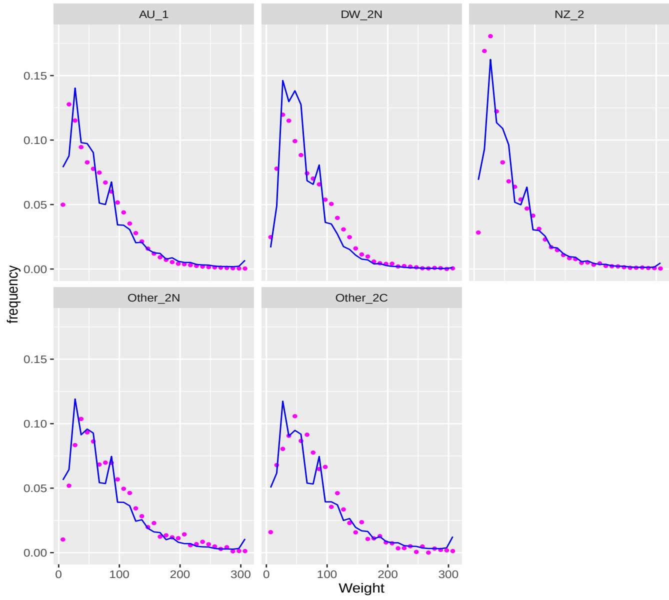


Figure 2.3: *Aggregated fits to the weight frequency data: magenta dots are the observed data, blue lines the model predictions.*

The important trends - SSB depletion, recruitment, harvest rate (fishing pressure) - are very comparable between the custom OM outlined herein and this particular scenario in the WCPFC assessment. What does differ is the estimates of current depletion levels (they are more optimistic in the custom OM herein than in the assessment). This is driven mostly by differing estimates of absolute abundance (the custom OM estimates higher values than the WCPFC assessment). Given all we have for this stock is relative abundance data (CPUE indices, catch composition) this is basically a result of differing interpretations of both what time-scale these data are informative on (annually vs. quarterly) and how to weight these data in the estimation process. In this work we use an annual, not a quarterly, time-step that we think better reflects the reality of the Swordfish life-history and increases the signal-to-noise ratio in the data (by reducing intra-annual variability). In terms of data weighting we take a more rigorous statistical approach and actively estimate the “correct” weighting of each data set, relative to the others. In the WCPFC assessment a pre-set weighting strategy (with some alternatives) is used for both the CPUE indices and the catch composition [4]. Ultimately, with noisy relative abundance data, as we have here, differing interpretations and approaches will lead to different stock status outcomes (to some degree). We have outlined what we think are the reasons behind the choices we have made and consider the estimates demonstrated herein at least as meaningful and - importantly - useful (in the MSE sense) as the WCPFC assessment. Given this, we feel comfortable using this custom OM model to develop our suite of OMs to be used in the MSE work.

2.2 Striped Marlin

The most recent WCPFC Striped Marlin was 2019 [5]. The diagnostic scenario was for a sexually and spatially aggregated model configuration, though some spatially explicit runs were done and were informed by recently updated estimates of movement across the same 165E boundary as assumed for Swordfish [6]. For the Striped Marlin MSE work we propose to use the WCPFC assessments to update the previous MSE code used in the various pieces of work done over the

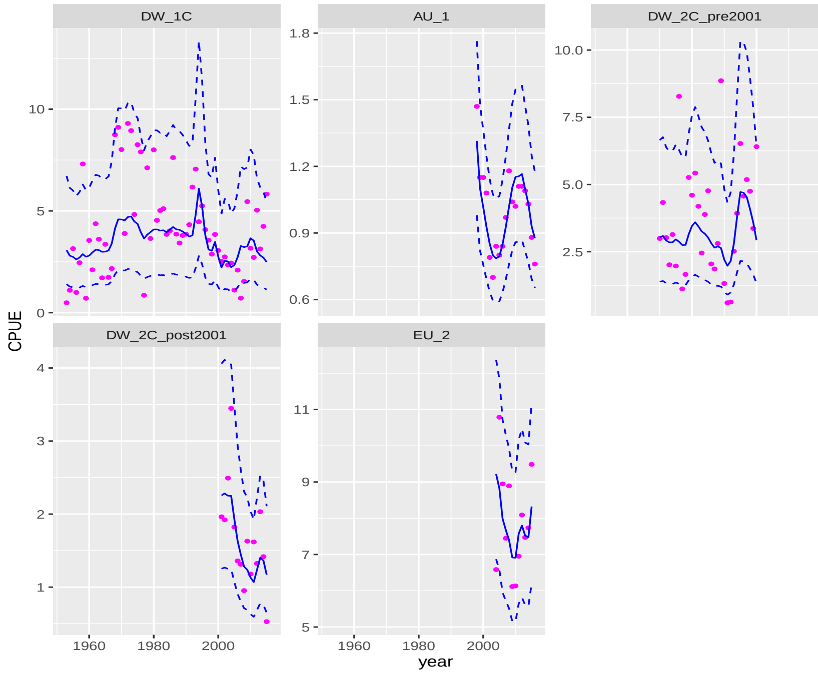


Figure 2.4: *Fits to the CPUE indices. Magenta dots are observed points, the full and dashed blue lines are the predicted value as well as the approximate 95% confidence interval.*

last 10+ years. Beyond efficiency, there is really not many points of contention that the TTRAG has had with the Stiped Marlin stock assessment, as opposed to the Swordfish assessment over at least two previous iterations.

2.3 Multi-species aspects

In the ETBF we do not split effort by species - in terms of generating standardised indices of abundance there is a lot of work done to try and deal with species targeting and seasonal availability. We have a single effort series for the ETBF as a whole and we do not manage effort, we manage total catch. The simple thought experiment for when we are managing two species is this: we get TAC1 for species 1 and TAC2 for species 2 so what are the predicted levels of effort E1 and E2 that relate to these TACs? Consistent mismatches (E1 $>$ E2 or E1 $<$ E2) that cannot be mitigated by technical or operational changes are going to result in the classic multi-species problem of harvest strategies that produce TACs for single species that are tuned to management objectives that are clearly not consistent (we cannot meet both). Economically speaking, the worst case scenario is where the lower effort is predicted for the lower value and/or volume species (e.g. Striped Marlin) than for the higher value/volume species (e.g. Swordfish) and the lower value/volume species becomes the choke species. The best case scenario is that the opposite is true and the TACs for the lower value/volume species result in consistently lower implied effort levels than those for the high value/volume species' TACs. It still results in undercaught TACs (which has been a clear dynamic for certain species in the ETBF for a number of years) but it doesn't economically impede the fishery and only results in certain conservation objectives being overshot.

By developing a probabilistic relationship between historical effort, catch and exploitable abundance coming from the OMs we will be able to construct the implied effort future effort distribu-

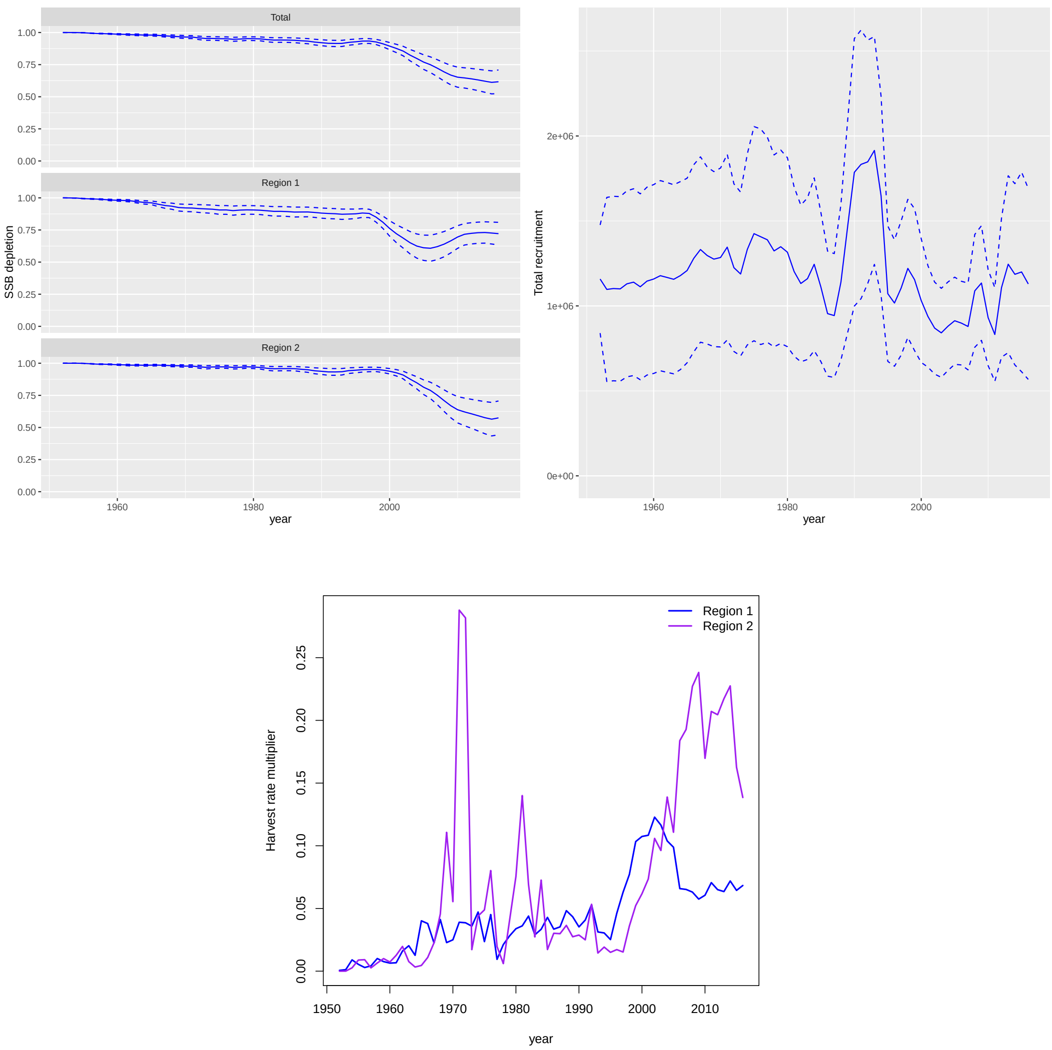


Figure 2.5: *Status summary (depletion, top left; total recruitment, top right; harvest rate multiplier, bottom) for the example run on the Swordfish data: zero movement, diagnostic M and steepness values, all fisheries included.*

tion given abundance and the TAC. By doing this we can, albeit indirectly, explore the operational multi-species aspect of the species specific TACs in terms of consistency of predicted effort and possible economic and conservation implications that might flow on from those.

3 Questions to TTRAG

This section tries to lay out the key points of feedback that we are seeking (or actively require) from the rest of the TTRAG in order to do the MSE work.

3.1 Objectives & time frame for the candidate harvest strategies

The notion of objectives for the candidate harvest strategies is clearly of high importance: what do we want the HS to achieve and how would we like it to achieve it?

What do we want to focus on: do we focus on a single objective (say relative SSB depletion at a given point in time) or multiple objectives more rooted in real world things we actualy monitor (like average CPUE levels) and control (like average catches)? We cannot, in this MSE framework, include very low level objectives such as say average catch per boat or other more operational factors so it is worth bearing in mind to consider only higher level processes as objectives. For example, one set of objectives might be maintaining some level of mean (or even minimum) future CPUE without average catches varying too much outside some pre-specified range. Setting objectives based on relative SSB depletion can have clear benefits, chief among them ensuring an HS is consistent with the intent of the Commonwealth Harvest Strategy Policy (HSP). The main downside is that we do not have a particularly accurate and robust estimate of what the current SSB depletion is for Swordfish, thereby weakening the relationship between what we can measure (CPUE) and what we wish to influence (SSB). As a consequence of this, the levels (and trajectories) of future CPUE resulting from attaining say a future SSB depletion of 48% can differ quite markedly - possibly from levels so high that require untenable cuts in catch to achieve, or levels so low they are fundamentally uneconomic. We would suggest that, for the ETBF case, focussing on a combination of some measure of future CPUE and ETBF catch might be preferable. We will look at relative SSB depletion as a key performance statistic in the MSE work. However we will **not** use a stock assessment (and associated *estimates* of SSB depletion) in the HS so focussing on using level of relative SSB depletion as something we actively want to achieve will have some problems.

A secondary part of the setting objectives process after “what do we want to achieve” is “how do we want to achieve it”. Given a set of high level objectives, there are a number of possible ways a harvest strategy could achieve those objectives. Giving HS developers some preferences can assist in developing a HS that not only meets the objectives, but does it in such a way that is preferred by the stakeholders. In the development of the previous SBT HS and the currently adopted one developers in the CCSBT Scientific Committee did a lot of stakeholder consultations (inside and outside formal meetings) to get information on preferable characteristics of any candidate HS so that the end result has acceptable features. For example: if we need a HS to rebuild the SSB from the current level then, almost certainly, this involves cutting catches. Depending on exogenous factors (such as discount rates or current market conditions and availability of other species) it might be preferable to take the cuts early and flatten out the TAC later on (so called early pain) or it might be better to offset the cuts until later on (late pain). Flipping this scenario around if we have to increase catches to meet the objectives we might prefer to take increases earlier than later. Knowing of these preferences in advance can be very useful to HS developers.

The third part is timeframe: over how long in the future do we want to test the HS? There are some logical boundaries we can place around the timeframe given what we know about the species life-history and the lags in the system itself (delay between getting information and acting on it). In terms of lower bounds on timeframe we would suggest at least 15 years. It takes 5–10 years for future random recruitment dynamics to really begin to have an impact on the exploitable abundance and SSB. Given we use a moving average CPUE in the HS it takes another at least 5 years between getting a signal in the CPUE, acting on it through the HS and it having an effect on the abundance and SSB. So we would suggest a minimum of 15 years for future projections. Maximum timeframe should be guided by two factors: (i) give the HS enough time to do what it needs to do to meet the objectives, and (ii) we have to be realistic about how far into the future our projections are going to be meaningful. Going beyond 25 years will begin to stretch the reliability of the projections but 20–25 years *should* give the HS enough time to act to meet the objectives. So we suggest between 15–25 years for the timeframe.

3.2 Operational constraints

There are still a number of key operational parameters of the harvest strategy that need to be decided or at least bounded in terms of possibilities in order to complete the MSE work:

- *Frequency of change*: previously an annual TAC cycle was used in the operation of the now-defunct harvest strategy, but there are advantages as well as possible downsides to exploring multi-annual TACs
- *Constraints on TAC variability*: specifying a maximum and a minimum percentage or actual tonnage amount of change for the TAC is advisable to avoid the HCR just chasing too much noise in the signal from the indices used in the HS

Certainly for longer lived species (like swordfish) there are often no clear positives in having an annual TAC cycle. Their life-history can be such that there is no real reason annual changes of TAC will actually pick up true changes in the underlying population variables of interest. Additionally, economically speaking for ITQ fisheries where multi-annual capital investment is important stability in TAC can often be preferable to possible short term increases or decreases in TAC. Reducing volatility is a key concern in capital-heavy resource economics and that is the principle at play here. We suggest exploring 1, 2 and 3 year options. For the annual TAC (1 year) scenario the 10% maximum change rule probably makes sense, but for 2 and 3 year options it should be higher given the potential need to act stronger when permitted given we only change the TAC every 2 or 3 years. This could take the form of a percentage or an actual tonnage cap on the amount of change allowed.

3.3 HCR control parameters

The "target" parameter in the HCR is our primary control parameter and will be actively changed, or tuned, to meet the various objectives so is not something we advocate simply fixing based on heuristic principles or even simple regression analyses of CPUE against predicted SSB. Similarly, we will explore different levels of the threshold parameter also to try and ensure the HS has good performance on avoiding the limit SSB depletion levels as defined in the Commonwealth Harvest Strategy Policy (HSP). Two feature we seek advice on are: (i) do we want a symmetric or asymmetric response of the HCR either side of the target buffer zone; and (ii) how wide do we want the buffer zone to be (in terms of a percentage) either side of the target?

Symmetric responses are often good at maintaining a kind of *status quo* scenario where do not require the HS to move the SSB (via the catch) too much from where it is at the start of the projection period. Also, they are easier to visualise what they are going to do: if index X changes by $Y\%$ up or down the catch will change by $Z\%$ up or down. Asymmetric responses can often be advantageous when we want the HS to move the SSB in a specific direction (be it up or down). There are no hard-written rules here, but these would be the high level general properties of symmetric/asymmetric responses.

In terms of the buffer this is about balancing a perennial trade-off: we want stability (in terms of inter-annual catch variability) balanced with reactivity and ability to react fast enough to signals (be they positive or negative). The wider the buffer, the more stable the TAC when the index is around the target (a good thing); once signals begin to move away from the target (especially negative signals) the wider the buffer the slower the HS is at changing the TAC in response to those signals. There's no optimal here: it is about expressing how much you wish to trade off stability with reactivity. Given we generally work with an informed guesstimate of a CV of 0.15

and the CPUE observation error and we suggest a 4 year moving average for the CPUE. That would suggest a rough estimate of the standard error of the mean CPUE we will use in the HS is $0.15/\sqrt{4} = 0.075$ suggesting that a change in the moving average CPUE *outside* of the range of $\pm 10 - 15\%$ will me far more likely to be a *true* change in CPUE as opposed to be just noise. That would suggest a buffer zone of $\pm 10 - 15\%$ at least would avoid us being close to the "target" yet jumping around following noise.

3.4 Robustness trials & future non-ETBF fishery scenarios

Best practice in MSE is not simply about tuning one base case scenario and then assuming that the performance on this reference case scenario is enough to tell us the likely efficacy of the HS in the real world. We must also explore a number of plausible future hypotheses about what might happen to the population and the fisheries (not just the ETBF) in the future.

Arguably the consistently most influential future scenario is where recruitment does something different in the future than what it has done in the past. These scenarios can be either positive or negative but you do want a HS that can deal with these scenarios adequately. For the SBT MSE work, for example, we have a scenario where the first n years in the projections mean recruitment is altered by a fixed percentage (up and down) and the most important scenario in the MSE work is often were recruitment in the first 5 years is 50% of the expected value. We should consider similar scenarios for the ETBF MSE work.

Another key factor has always been: what do the non-ETBF fleets do in the future? In the previous MSE work this has perhaps been over-simplified to a single multiplier to adjust the future non-ETBF fleets (considered all together) fishing mortality (a proxy for effort). It would be worth considering a more nuanced approach for the various fisheries in the different regions given their differing recent historical catch and effort dynamics (some are increasing, others are decreasing and have differing selectivity patterns). The TTRAG should think on specifying either fixed F or catch possibilities for each of the various fleets to explore a wider range of possible hypotheses.

3.5 Exceptional circumstances and meta-rules

With the adoption of a harvest strategy, the TTRAG and TTMAC may wish to consider adoption of meta-rules that outline the agreed schedule of activities for HS implementation including a safety-net check on the TAC recommendation. The meta-rules schedule of activities can include the frequency of: evaluation of exceptional circumstances, TAC setting, assessment of stock status and HS review. A key component is identification of exceptional circumstances, which are events, or observations, that are outside the range of values used in the simulation testing of the HS which may make application of the HS recommended TAC risky, or highly inappropriate. The exceptional circumstances process under the meta-rules involves the following three steps: 1) Determining whether exceptional circumstances exist; 2) A process for action that examines the severity (and implications) of the exceptional circumstances for the operation of the HS, and the types of actions that may be considered; 3) Principles for action that determine how recommendations from the HS might be altered, if at all. This process is not overly specified, as it is intended to deal with circumstances not foreseen or tested against. The process should examine whether: 1) the inputs to the HS are affected, 2) the population dynamics are potentially significantly different from those for which the HS was tested (as defined by the Reference and Robustness sets of OMs), 3) the fishery or fishing operations have changed substantially, 4) total removals are greater than the HS recommended TACs, and 5) if there are likely to be impacts on the performance of the HS relative to the broad HS objectives.

The meta-rules process is used in the SBT fishery, where TACs are set every 3 years using the adopted management procedure, and meta-rules checks for exceptional circumstances are conducted each year to ensure that the TAC decisions are still sound. An SBT example of evaluation of exceptional circumstances is from 2012, when close-kin adult abundance data were included in the SBT models and our understanding of the historical status of the stock and population dynamics changed: the depletion level was not as low as previously thought and the HS had not been tested under these conditions. The impact of this exceptional circumstance on performance of the adopted MP was examined through testing the MP in updated operating models. The results showed that the MP was still going to drive recovery of the stock, and in a faster time frame than before. Additionally average catches would be better than previously predicted. As this was a positive impact of the exceptional circumstance, the Commission agreed to keep the MP as it was adopted and there was no change the recommended TAC.

The meta-rules process and evaluation of exceptional circumstances provides a framework for orderly implementation and review of the performance of the HS, and provides a transparent and clearly reasoned TAC recommendation from TTRAG and TTMAC to stakeholders and AFMA.

4 Discussion

In this paper we have tried to crystallise the key questions the MSE team have for the TTRAG about various aspects of the candidate harvest strategies (objectives, features, control parameters etc.). We have also outlined the progress made in conditioning the custom suite of OMs for the Swordfish MSE work as we have moved away from using the WCPFC assessment to condition the previous OMs.

References

[1] R. Hillary (2018) Updating the ETBF Broadbill Swordfish MSE analyses . *TTRAG, Canberra, ACT*.

[2] R. Campbell (2018) Identifying Cohorts in the Eastern Tuna and Billfish Fishery and application to the Harvest Strategy. *TTRAG, Mooloolaba, QLD*.

[3] T. Carruthers *et al.* (2016) Performance review of simple management procedures. *ICES J. Mar. Sci.* **73**(2): 464–482.

[4] Y. Takeuchi, G. Pilling, and J. Hampton (2017) Stock assessment of swordfish (*Xiphias gladius*) in the southwest pacific ocean. *WCPFC-SC13-2017/SA-WP-13*.

[5] N. Ducharme-Barth, G. Pilling, and J. Hampton (2019) Stock assessment of SW Pacific striped marlin in the WCPFC. *WCPFC-SC15-2019/SA-WP-07*.

[6] R. Hillary and T. Patterson (2019) Initial estimates of striped marlin WCPFC movement dynamics. *TTRAG, Mooloolaba, July 2019*

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

WE DO THE EXTRAORDINARY
EVERY DAY

We innovate for tomorrow and help improve today for our customers, all Australians and the world.

Our innovations contribute billions of dollars to the Australian economy every year. As the largest patent holder in the nation, our vast wealth of intellectual property has led to more than 150 spin-off companies.

With more than 5,000 experts and a burning desire to get things done, we are Australias catalyst for innovation.

WE IMAGINE. WE COLLABORATE.
WE INNOVATE.

Appendix 4

Oceans & Atmosphere
www.csiro.au

CSIRO

Conditioning of the Broadbill Swordfish Operating Models

R. Hillary

25 Feb 2020

CSIRO Oceans & Atmosphere
Battery Point, Hobart 7004, Tasmania, Australia.

Copyright and disclaimer

Important disclaimer

Contents

1 Background

This paper details the structure and conditioning of the Operating Models (OM) for both the broadbill swordfish (*Xiphias gladius*) the key billfish species targeted in the ETBF. For swordfish we constructed custom-designed OMs so as to be able to accommodate the scenarios we have discussed previously that were not either included or even possible in the WCPFC stock assessment.

2 Candidate Operating Models

At a (moderately) non-technical level we make the following changes to the assumptions used in the most recent WCPFC stock assessment:

- **Catch**: numbers are used as known inputs not estimated and assigned very small standard errors as done in the MFCL assessment
- **Effort & catchability**: we don't include effort as a covariate for fishing mortality (we use harvest rates given abundance and catch in numbers) and we don't model time-varying catchability for fleets where we don't have a usable CPUE series. This means no effort deviations or catchability parameters dramatically reducing the number of parameters from 000s in the MFCL model to just over one hundred in our OM. This massively reduces run time from hours (for one MFCL grid configuration) to less than 3 seconds. We don't care about the quality of the effort-to-fishing mortality relationship or the catchability for fleets where we have no CPUE index (and are saying that effort is not very meaningful anyway). We can drastically reduce the number of parameters without reducing the complexity of the model by doing this
- **Biology**: use same updated growth, maturity, weight-at-length relationships and steepness values. We propose to reduce the number of natural mortality scenarios from 6 to 2–3 given the early high M scenarios are implausible in a life-history sense
- **Population dynamics**: age-based and size and weight-structured to deal with composition data. We use an annual, not a quarterly, time-step and account for time of year of fishing in catch equations to deal with seasonal dynamics. Recruitment is allowed to be spatiotemporally correlated and treated as a random effect with the correct value of overall variability and the two correlation parameters estimated
- **Migration**: modelled annually, not quarterly, and modelled not using diffusivity but via a transition matrix defined via probabilities of moving from one region given you are currently in another region. Movement is allowed to be asymmetric between regions
- **Length & Weight composition**: we do not use quarterly length and weight data we aggregate (using relative sampling intensity) across quarters to get annual data to reduce seasonal variability. We also remove the somewhat arbitrary minimum and maximum weights applied in the WCPFC assessment and use statistical theory to estimate the correct weighting for each fleet's composition data at run time. This is currently accepted best practice in integrated assessment modelling and used widely in Australia
- **CPUE indices**: as with the composition data we aggregate the CPUE indices from the quarterly to the annual level using relative effort by quarter to reduce the high levels of intra-annual variability in some indices. The same indices are used but, as with length

composition, instead of semi-arbitrary subjective weightings (via the effort deviation penalties) in the WCPFC assessment we use statistical theory to estimate the correct weighting for each CPUE index

The three major changes, in terms of expected flow-on effect on estimates of current relative SSB depletion:

1. Treatment of the size and weight frequency data: by statistically estimating the appropriate weighting of these data we change their likely contribution, relative to the current WCPFC assessment, to how big the stock actually is. This will, conditional on recruitment deviations and the CPUE indices likelihood, likely change estimates of SSB depletion
2. The same goes for how we treat the CPUE indices: conditional on the recruitment deviations and the catch frequency data likelihoods, our different weighting scheme (statistical weighting) will mean the same (generally speaking) data will likely suggest different overall abundance levels and, as such, different SSB depletion levels
3. The particulars of the recruitment penalty we assume explicitly embeds both spatial and temporal variation around the stock-recruitment curve (be it a singlular population or multiple spawning populations). As we estimate both the degree of variation and correlation (in both time and space), as opposed to the WCPFC assessment which fixes both, this will also have a knock-on effect to how the recruitment assumptions alter the estimates of abundance and SSB depletion

All of the data available for assessing this stock of swordfish, whatever the particulars of the model, are relative in nature when it comes to abundance information. With multiple sources of data this effectively means that the relative weighting of those data, for a given model configuration, will very often (strongly) affect the outcome in terms of abundance, mortality and status estimates. Anyone familiar with assessments will have seen this many times: changing assumptions results in changed outcomes. The point of constructing this custom model was to get around what the TTRAG thought were a number of problems with the WCPFC assessment in terms of representing plausible hypotheses about the stock and the particulars of the ETBF fishery. We did not set out to replicate the most recent assessment so we hope people understand that, as a result of the changes we made and feel are justified, we do not always get the same results. This is not an alternative assessment - it is a process for generating plausible OMs for the MSE work that are compatible with the available data and understanding. The technical specifications of the OM can be found in the Appendix.

2.1 OM conditioning scenarios

Previously for swordfish there were a number of elements within the uncertainty grid employed by WCPFC in their assessments (whilst still having an overall diagnostic case as termed). Over the years both the growth and maturity questions (which curves were correct) have effectively been answered, so we only use the most recent growth and maturity curves. As we are statistically weighting the length/weight frequency and the CPUE indices at estimation time we do not have any alternative data weighting scenarios. That leaves the following five scenario "groups":

1. Steepness of the stock-recruit function (region independent): we embed the three candidate values, $h \in \{0.65, 0.8, 0.95\}$, within the uncertainty grid
2. Natural mortality-at-age: we have discarded the very high early-age M scenarios from

the previous assessment given what we would suggest is a lack of comparative plausibility with other similar sized animals in a similar ecosystem. That leaves two alternative M scenarios: the diagnostic case and the alternative lower mortality case. For reasons that will hopefully be made clear later on, we actually separate these two options into the reference case (the WCPFC diagnostic case) and a sensitivity scenario (lower M option); we do not embed them within the uncertainty grid. Figure 2.1 shows the summary of the two natural mortality scenarios

3. Migration: we explored 0, 5 and 10% movement probabilities (symmetric in terms of direction) for **annual** migration hypotheses. Rather than embed them within the uncertainty grid (like steepness) we use the 0 movement as the reference case and look at the others as sensitivities

4. Stock structure: we have two options and they are panmictic (i.e. a single reproductive stock across the South West Pacific) and separate stocks (i.e. a distinct spawning population in each region)

5. Fishery options: the reference case **excludes** the distant water fleet in the north of region 2 (DW2N) - the one with the recent high catches right on the "border" of the north of region 2 - and the sensitivity we explore is where we include it (so all fisheries included as the diagnostic case in the WCPFC assessments)

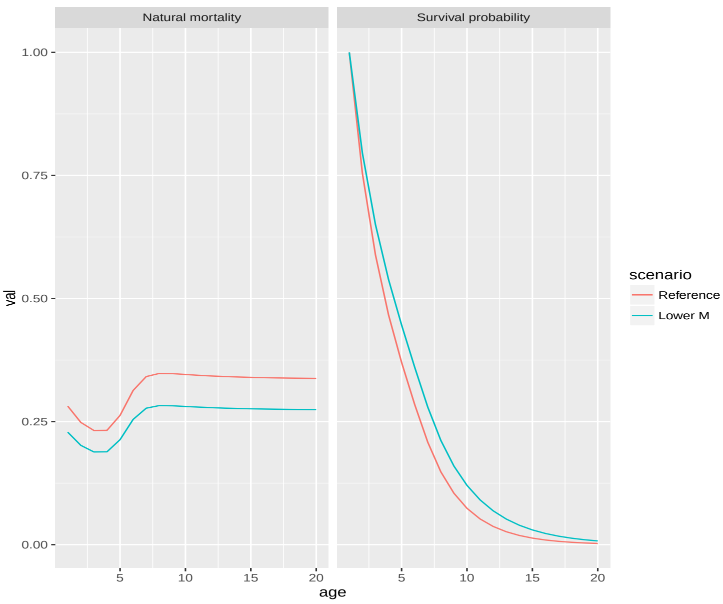


Figure 2.1: *Natural mortality (left) and associated survival probability (right) summary for the Reference and Lower M scenarios explored herein.*

So, in summary, the steepness is the only core element within the estimation uncertainty grid - that is every scenario used in the MSE runs has all three steepness values included no matter what the other scenarios are. We have run "crosses" for all the migration scenarios and the two fishery scenarios given their noted interaction in the previous work. For the natural mortality and stock structure scenarios we do not cross them with any other scenarios for the moment. This gives a total of 9 scenarios each with 3 steepness values so, in total, 27 possible OM configurations used in the initial MSE work herein. There is not technical impediment to running a far more detailed full cross of all the scenarios, but given this is the first look at the revised

MSE work, for now this gives us more than enough possibilities with which to make some useful insights and possible recommendations.

2.2 Fits to data for reference OM

In this section we outline how the "median" reference case OM fits to the available data (no DW2N, steepness of 0.8, no migration, single stock). Estimated parameters are: recruitment parameters (total effort deviates, spatiotemporal correlation and region-level variance, and mean spatial recruitment fraction by region), selectivity for each fishery, recruitment for each year and region, over-dispersion values for length and weight frequency by fleet, CPUE process error by fleet. As mentioned, we do not estimate effort deviations or time-varying catchability for each fleet that doesn't have a constant q assumed.

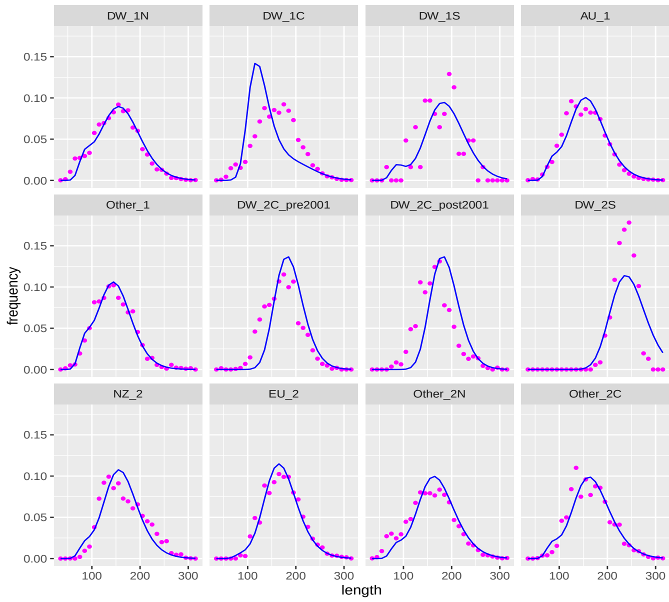


Figure 2.2: *Aggregated fits to the length frequency data: magenta dots are the observed data, blue lines the model predictions.*

Figure 2.2 shows the fits to the length frequency data aggregated across years, and Figure 2.3 shows the fits to the weight frequency data als aggregated across years. Figure 2.4 shows the fits to the CPUE indices. For readers familiar with, or with a copy of, the most recent WCPFC assessment [4] we fit to the length frequency data as well and to the weight frequency data arguably a little better for certain fleets. It is a little harsh to compare our fits to the CPUE data, relative to the assessment, given we have already smoothed the series by averaging across quarters. The series given the highest statistical weighting are the Australian ETBF in region 1 and region 2 the EU fleet with the noisy distant water fleets (in both regions) getting lower statistical weighting.

Figure 2.5 shows a general status summary for median reference OM. In terms of dynamic SSB depletion (given recruitment estimates etc. what the model estimates divided by the SSB in the absence of fishing) the most recent estimates are just above 68% in total, around 75% in region 1 and around 62% in region 2. Total recruitment shows a similar pattern to the WCPFC assessment [4] with a gradual increasing trend from the 50s to the 80s and more variability from the 1970s onwards (more data), a gradual decrease from 1990 to the mid 2000s with some

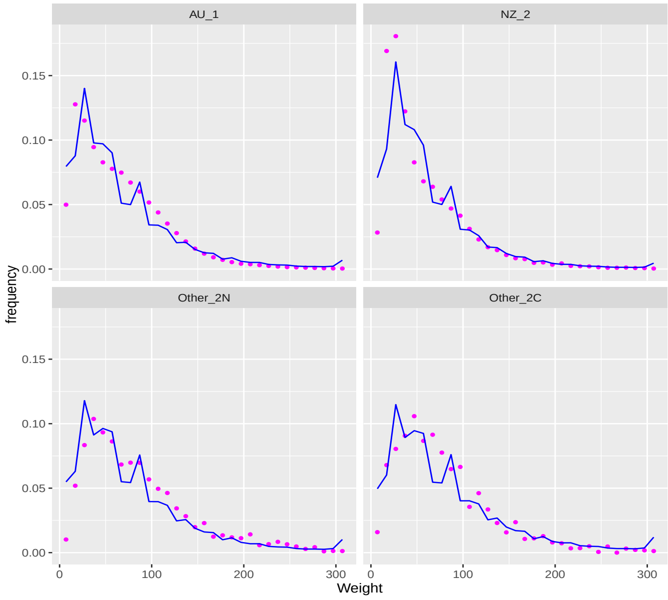


Figure 2.3: *Aggregated fits to the weight frequency data: magenta dots are the observed data, blue lines the model predictions.*

higher values in the most recent years. Contrary to the diagnostic case scenario which includes both migration and the DW2N catches, the overall recruitment fraction in each region is about 50/50. The annual trends in regional weighted mean harvest rate also mirror quite closely those in the WCPFC assessment (not surprisingly given the similarity in both biomass and recruitment trends). In region 1 values gradually increased then rose sharply in the mid 1990s as the ETBF effort and catch increased, decreasing in the late 2000s onwards as effort and catch declined in the ETBF. Region 2 shows a characteristic high spike around 1970 and again around 1980 followed by a sustained increase from just before 2000 onwards as the catches in region 2 (including the EU) begin to increase (even without the DW2N fishery included).

Table 2.1 details the depletion summary (total as well as per region) for the "median" (i.e. steepness of 0.8) grid run across a number of scenarios alternative to the reference case. In terms of total depletion (i.e. across all regions) the range in median levels is 0.44–0.74 with most numbers around the 0.65–0.7 - it is the lower M scenario that results in a clearly lower depletion estimates across the board. For region 1 median estimates range between 0.56–0.8 with the majority of estimates between 0.75–0.8; for region 2 median estimates range between 0.35–0.7 with the majority of estimates between 0.6–0.7. With regards to the recruitment parameters: estimates of σ_R were around 0.25–0.3, estimates of ρ_1 (temporal correlation) were around 0.45–0.5, and estimates of ρ_2 were effectively zero (very small) so not apparent evidence of spatial correlation of recruitment variables but strong and consistent autocorrelation. Tentatatively one might suggest that the drivers of recruitment variation within the regions show strong temporal correlation, but those drivers appear to differ across regions given they show little to no spatial correlation. Overall estimates of total recruitment variability (which is comparable to the value used in the WCPFC assessments – see Appendix as to why) were around 0.35–0.4 so slightly lower than the value of 0.5 assumed in the WCPFC assessments.

In terms of high-level conclusions we suggest the following:

- The separate stock hypothesis does nothing to change overall or region depletion levels

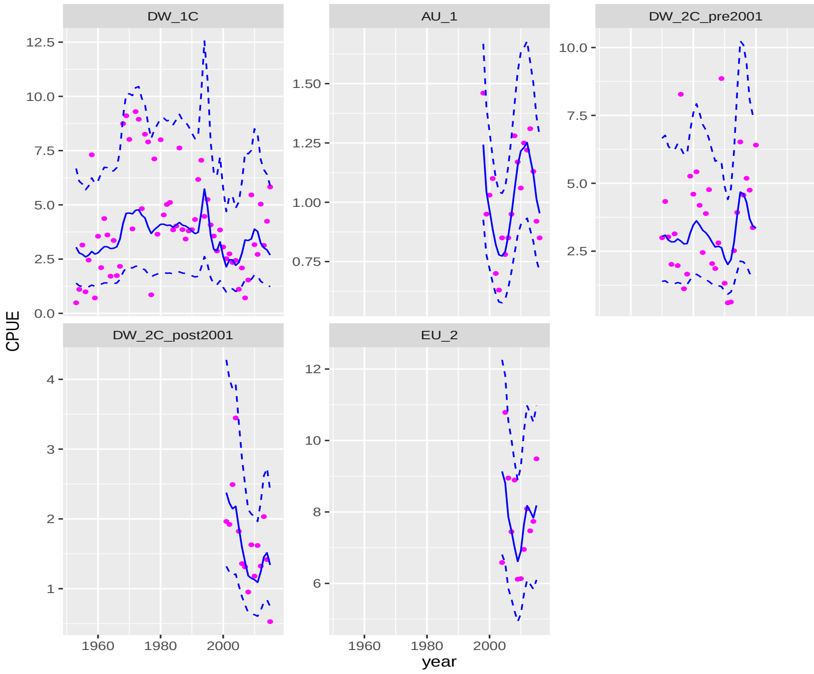


Figure 2.4: *Fits to the CPUE indices. Magenta dots are observed points, the full and dashed blue lines are the predicted value as well as the approximate 95% confidence interval.*

(somewhat predictably). What it does is start a debate on what depletion level we focus on for the ETBF performance summary in this case: do we stay with the total or do we move to region 1?

- Migration, at least for the scenarios considered herein, actually tends to result in higher levels of status across the regions and in total. What is interesting is that the 5% migration scenario results in the highest improvement for region 1 but reduces a little as we move to 10%; for region 2 you get a clear continual improvement in status as the migration increases. What we can conclude is that, eventually, region 1 begins to pay the price for "subsidising" region 2's increase in status.
- Including the DW2N fishery is generally detrimental for overall and region 2 status but, by very definition, with zero migration it makes no difference to the status in region 1
- The lower M scenario has a clear effect which is reducing status across the board. The reason for this is because the lower M vector (across all ages) results in lower estimates

Run	Total	Region 1	Region 2
Reference	0.68 (0.61–0.75)	0.77 (0.68–0.85)	0.6 (0.5–0.71)
Inc. DW2N	0.64 (0.55–0.75)	0.77 (0.68–0.85)	0.58 (0.44–0.71)
Sep. stock	0.68 (0.61–0.75)	0.77 (0.68–0.85)	0.6 (0.5–0.71)
5% migration	0.73 (0.67–0.8)	0.8 (0.74–0.86)	0.66 (0.58–0.73)
10% migration	0.74 (0.69–0.8)	0.78 (0.73–0.83)	0.7 (0.64–0.76)
Lower M	0.44 (0.38–0.51)	0.56 (0.49–0.63)	0.35 (0.25–0.45)

Table 2.1: *Depletion summary (total, regional) for the "median" grid option across the key OM scenarios.*

of unfished recruitment, R_0, as with lower M values it takes less initial fish to populate the exploitable abundance that we see in the various CPUE indices. The flow-on effect is that the catch that has been taken out has a stronger effect on the depletion over time and this is why we see the lower levels of depletion for this scenario

3 Discussion

This paper details the conditioning of the sworfish OMs to be used for the revision of the billfish harvest strategies. We have constructed custom-designed OMs that better fit the requirements the TTRAG has discussed and agreed on over the last few years. For the swordfish we successfully fitted all the variant OM configurations to the available data (CPUE indices, length and weight frequency data). We explored alternative fishery inclusions, migration hypotheses, the plausible alternative lower M vector, and separate spawning stocks in the 2 regions. The range of recent median depletion levels is around 0.45–0.75 but most around the 0.65–0.75 level. Depletion levels in region 1 were consistently above those in region 2 and, as such, consistenly higher than the overall stock depletion levels which does have potential implications for the separate stock scenario.

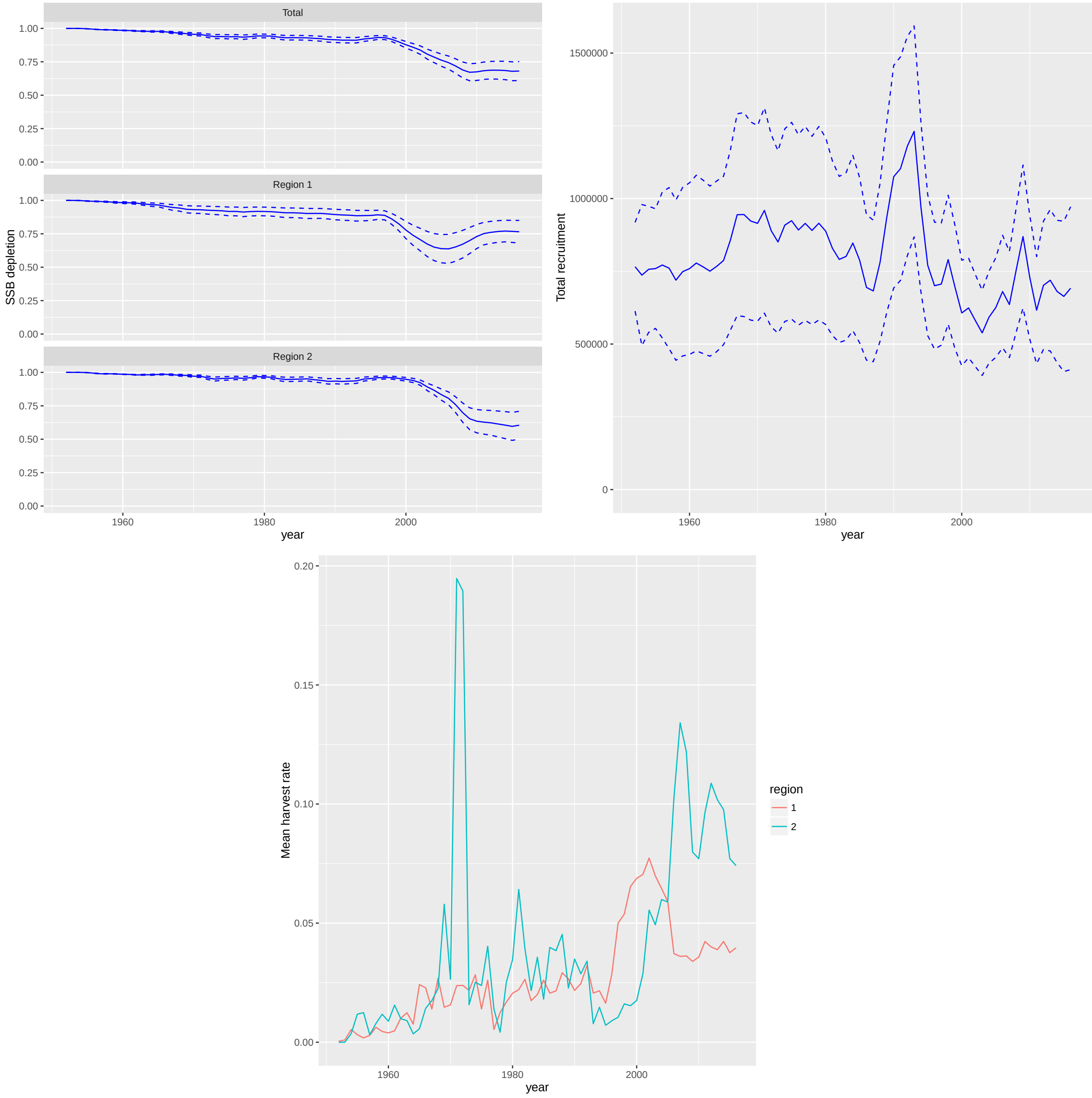


Figure 2.5: *Status summary (depletion, top left; total recruitment, top right; weighted mean harvest rate, bottom) for the example run on the Swordfish data: zero movement, diagnostic* M *and steepness values, all fisheries included.*

References

[1] R. Hillary (2018) Updating the ETBF Broadbill Swordfish MSE analyses . *TTRAG, Canberra, ACT*.

[2] R. Campbell (2018) Identifying Cohorts in the Eastern Tuna and Billfish Fishery and application to the Harvest Strategy. *TTRAG, Mooloolaba, QLD*.

[3] T. Carruthers *et al.* (2016) Performance review of simple management procedures. *ICES J. Mar. Sci.* **73**(2): 464–482.

[4] Y. Takeuchi, G. Pilling, and J. Hampton (2017) Stock assessment of swordfish (*Xiphias gladius*) in the southwest pacific ocean. *WCPFC-SC13-2017/SA-WP-13*.

[5] N. Ducharme-Barth, G. Pilling, and J. Hampton (2019) Stock assessment of SW Pacific striped marlin in the WCPFC. *WCPFC-SC15-2019/SA-WP-07*.

[6] R. Hillary and T. Patterson (2019) Initial estimates of striped marlin WCPFC movement dynamics. *TTRAG, Mooloolaba, July 2019*

Appendix

The model is spatially structured, sex-structured, age-based and size and weight structured with an assumed annual time-step. The initial population is assumed to be in an unfished equilibrium state. Movement is annual and considered at this stage to be independent of age or size, and is characterised via a time-independent matrix, Φ_{ij}: the probability that an animal in region i will migrate to region j from one year to the next (and we assume a closed population where $\sum_j \Phi_{ij} = 1$). Recruitment is assumed at age 1 and, for the single stock assumption, the mean value of total recruitment (across all areas) is governed by the following:

$$\tilde{R}_y = \frac{\alpha S_{y-1}}{1 + \beta S_{y-1}}$$

where S_{y-1} is the biomass of mature animals (across all regions) in year $y-1$ (calculated assuming the revised maturity-at-age used in the diagnostic case assessment). The parameters α and β are defined in terms of the steepness, h, unfished total female mature biomass, B_0, and the unfished total recruitment, R_0:

$$\alpha = \frac{4hR_0}{B_0(1-h)},$$
$$\beta = \frac{5h-1}{B_0(1-h)}.$$

The spatially and sexually disaggregated recruitment, $R_{y,s,r}$, is defined using the time-invariant spatial recruitment faction η_r ($\sum_r \eta_r = 1$):

$$R_{y,r} = \tilde{R}_y \eta_r,$$

and if annual recruitment deviations, $\epsilon^R_{y,r}$, are being estimated then there is an additional multiplier to $R_{y,r}$:

$$R_{y,r} = \tilde{R}_y \eta_r \exp\left(\epsilon^R_{y,r} - \sigma^2_R/2\right),$$
$$\epsilon^R_{y,r} \sim N\left(\mathbf{0}, \Sigma_R\right).$$

For the separate stock scenarios we set up the stock-recruit relationship differently. In that case the mean recruitment in each spatial region is defined by

$$\tilde{R}_{y,r} = \frac{\alpha_r S_{y-1,r}}{1 + \beta_r S_{y-1,r}}$$

where $S_{y-1,r}$ is the spatially-specific mature biomass and the (spatial) stock-recruit parameters are given by

$$\alpha_r = \frac{4hR_{0,r}}{B_{0,r}(1-h)},$$
$$\beta_r = \frac{5h-1}{B_{0,r}(1-h)}.$$

so we assume that steepness is not region-specific, but unfished recruiment (and mature biomass) is. The spatial recruitment values for the separate stock scenario are now defined as follows:

$$R_{y,r} = \tilde{R}_{y,r} \exp\left(\epsilon^R_{y,r} - \sigma^2_R/2\right).$$

The Gaussian process prior for $\epsilon^R_{y,r}$ permits both spatial and temporal correlation to be included in one multivariate prior distribution for the recruitment deviations. The prior covariance matrix, Σ_R, is defined via covariance matrices for both time and space: Σ_1 and Σ_2, respectively. The temporal covariance matrix is defined as follows:

$$\Sigma_1^{ij} = \sigma^2_R \rho_1^{|i-j|},$$

where ρ_1 is the temporal autocorrelation and i and j are the years in the model. This would be (marginally) equivalent to an AR(1) process often used in non-spatial auto-correlated recruitment process. The spatial covariance matrix is defined as follows:

$$\Sigma_2^{ij} = \rho_2^{|i-j|},$$

and here i and j are the spatial regions and ρ_2 is the spatial correlation coefficeint, so it is the spatial “analogue” of the AR(1) process. The full spatiotemporal covariance is defined as follows:

$$\Sigma_R = \Sigma_2 \otimes \Sigma_1$$

where $\otimes$ is the Kronecker matrix product, so we assume $\epsilon^R_{y,r}$ has a matrix normal distribution. In this particular formulation, σ_R in this case refers to overall variation in *region-specific* recruitment variations not *total* recruitment variation. To get from the area-specific formulation to an overall level of recruitment variation, σ^2_{tot}, we use the following: $\sigma^2_{\text{tot}} = 2\sigma^2_R(1+\rho_2)$. The main difference with the WCPFC (and most other) stock assessment is that we estimate σ_R, ρ_1 and ρ_2, we do not simply fix σ_R and ignore potential spatiotemporal correlation.

For ages 2 to $A^+ - 1$ (where A^+ is the plus group) the abundance dynamics are

$$N_{y,a,r} = \sum_{r'} \Phi_{r',r} N_{y-1,a-1,r'} e^{-M_{a-1}} \left(1 - h_{y,a,r'}\right)$$

where M_a is the age-dependent natural mortality rate, and $h_{y,a,r}$ is the fishery-aggregated harvest rate by age, year and region. For the plus group we have that

$$\begin{aligned} N_{y,A^+,r} &= \sum_{r'} \Phi_{r',r} N_{y-1,A^+-1,r'} e^{-M_{A^+-1}} \left(1 - h_{y,A^+-1,r'}\right) \\ &+ \sum_{r'} \Phi_{r',r} N_{y-1,A^+,r'} e^{-M_{A^+}} \left(1 - h_{y,A^+,r'}\right) \end{aligned}$$

The harvest rates are calculated on a fishery-specific basis as follows:

$$\begin{aligned} X_{y,f} &= \sum_a N_{y,a,r_f} e^{-\tau_{y,f} M_a} s_{a,f}, \\ h_{y,a,f} &= \frac{C_{y,f}}{X_{y,f}} \times s_{a,f}. \end{aligned}$$

where:

- $C_{y,f}$ is the catch **in numbers** by year and fishery
- $s_{a,f}$ is the selectivity-at-age by fishery
- $\tau_{y,f}$ is the fraction of the year at which the median fishing operations occurred
- r_f is the region in which fishery f occurs, and the regionally aggregated harvest rates, $h_{y,a,r}$, are just the sum (within year and age) of all the individual fishery-specific harvest rates

The distribution of length-at-age is simply defined from the growth relationship. The mean length-at-age is defined via the Schnute parameterisation of the von Bertalanffy growth curve:

$$\mathbb{E}\left(l(a)\right) = l_1(a_1) + \left(l_2(a_2) - l_1(a_1)\right)\frac{1-\exp\left(-k\left(a-a_1\right)\right)}{1-\exp\left(-k\left(a_2-a_1\right)\right)},$$

where $l_1(a_1)$ and $l_2(a_2)$ are the lengths at reference ages a_1 and a_2 ($a_2 > a_l$), and k is the growth rate. To generate the distribution of length-at-age we assume a lognormal distribution (with a given standard deviation σ_l) around this mean length-at-age. distribution of length-at-age, $\pi_{l\,|\,a}$. The model-predicted fishery-specific catch-at-age is given by $C_{y,a,f} = h_{y,a,f}N_{y,a,r_f}$ and we get the model-predicted catch-at-length (and associated frequency) as follows:

$$C_{y,l,f} = \sum_a \pi_{l\,|\,a} C_{y,a,f},$$
$$p_{y,l,f} = \frac{C_{y,l,f}}{\sum_l C_{y,l,f}}.$$

To get the weight frequency data we first need to define the distribution of weight-at-length, $\pi_{w\,|\,l}$ defined via a lognormal distribution with log-scale mean $\ln a + b\ln l$ and standard deviation σ_w. We get the distribution of weight-at-age as follows:

$$\pi_{w\,|\,a} = \sum_l \pi_{w\,|\,l}\pi_{l\,|\,a}$$

and, similarly to the length frequency, we get the catch-at-weight and associated frequency:

$$C_{y,w,f} = \sum_a \pi_{w\,|\,a} C_{y,a,f},$$
$$p_{y,w,f} = \frac{C_{y,w,f}}{\sum_w C_{y,w,f}}.$$

For both the size and weight frequency we assume a Dirichlet-multinomial distribution for the likelihood function. Without going into mathematical specifics this distribution uses the initial sample sizes of the observed data, the model-predicted frequencies and something called the over-dispersion parameter $\varphi_{f,\bullet}$ (• represents length or weight). The overdispersion parameter, which is estimated, represents the degree to which the variance in the observed data relative to the model-prediction. This is the parameter that ensures that the catch frequency data are correctly weighted statisticaly speaking.

The predicted CPUE, $\widehat{I}_{y,f}$, is given by calculating the exploitable abundance of the given fishery and its associated catchability, q_f:

$$\widehat{I}_{y,f} = q_f \sum_a N_{y,a,r_f} e^{-\tau_{y,f}M_a} s_{a,f},$$

and, for the ETBF case and depending on the particular index being used, we additionally account for the size range used in the construction of the index. A lognormal distribution is then assumed for the observed CPUE indices, $I_{y,f}$, with log-scale mean $\ln \widehat{I}_{y,f}$ and associated standard deviation σ_f^2. The total variance σ_f^2 is partitioned into a fixed observation level, $\sigma^2_{\text{obs},f}$, and an estimated process error variance, $\sigma^2_{\text{proc},f}$ and $\sigma_f^2 = \sigma^2_{\text{obs},f} + \sigma^2_{\text{proc},f}$.

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

WE DO THE EXTRAORDINARY EVERY DAY

We innovate for tomorrow and help improve today for our customers, all Australians and the world.

Our innovations contribute billions of dollars to the Australian economy every year. As the largest patent holder in the nation, our vast wealth of intellectual property has led to more than 150 spin-off companies.

With more than 5,000 experts and a burning desire to get things done, we are Australias catalyst for innovation.

WE IMAGINE. WE COLLABORATE. WE INNOVATE.

Appendix 5

Oceans & Atmosphere
www.csiro.au

Management Strategy Evaluation of the Broadbill Swordfish ETBF harvest strategies

R. Hillary

10 Mar 2020

CSIRO Oceans & Atmosphere
Battery Point, Hobart 7004, Tasmania, Australia.

Contents

1 Introduction

This paper details the initial results of the MSE work done to explore the first options for the candidate revised harvest strategies for the ETBF billfish species. The Operating Model (OM) specifications and scenarios used to condition them were detailed in [1].

2 Candidate Harvest Strategies

Figure 2.1 details the "base" harvest control rule (HCR) proposed for the revised HS. It has three broad levels to how it behaves:

- **Within the buffer zone**: centered around a notional "target" level there is a zone within which the HCR does *not* change the TAC and it stays at the previous level
- **Outside the buffer, above the threshold**: the TAC linearly increases/decreases above/below the buffer zone, and is permitted to do so asymmetrically
- **Below the threshold**: the HCR acts to strongly (quadratically) decrease the TAC down to zero as the key input index decreases to zero

There is the potential, at least for the swordfish, to include an additional recruitment term in the HCR but that is not explored in this initial work. In terms of the operational constraints around the HCR we explore 1 and 3 year TAC cycles and we permit a maximum of 10% and 27% maximum TAC changes, respectively with no minimum change.

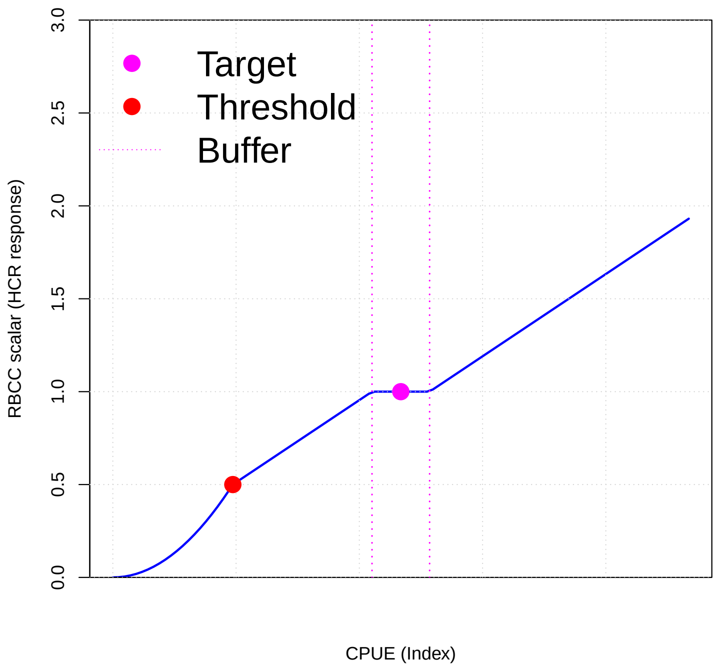


Figure 2.1: *Candidate (primary part of the) HCR to be explored in the MSE work.*

3 Implementation model

An inspection of the TAC and actual catch time series for either of the billfish clearly shows that there is a consistent, but variable, undercatch of the TAC from year to year. The HCR will recommend a TAC but we really want to account for the catch that is taken out of the population, which will almost always be less than the TAC. The variation suggests there is some randomness to the process and it's something we can embed within the projections. To do this we look at the bootstrapped mean and standard deviation in the ratio of catch taken to TAC for the years 2013–2018 (we omit 2019 as something of an outlier and exclude it in this initial analysis). We use the mean estimates of these bootstrapped mean and SD to parameterise a beta distrubtion for the multiplier applied to the TAC to get the actual catch taken. The median (and approximate 90% CI for this distribution) is 0.86 (0.78–0.92), so the actual catch taken by the ETBF is almost always between 80–90% of the TAC in the projections.

4 Tuning criteria

In the out-of-session phone discussion the group generally settled on two quite different general tuning groupings:

1. **CPUE tuning**: by 2035 (15 years from the first TAC decision) attain the mean CPUE, measured between 2011–2016, and also $\pm$ 20% of this value
2. **SSB depletion tuning**: by 2035 (15 years from the first TAC decision) attain SSB depletion levels of 38%, 48% and 58%

5 Key MSE runs

The reference set of OMs, the one in which we tune the candidate HS to, si defined as follows: three steepness options (0.65, 0.8, 0.95), the diagnostic case M-at-age vector as per the WCPFC assessment, no DW2N fishery inclueded, zero migration, and one spawning stock across both regions. To define a particular scenario we have defined the following hierarchy:

- Two fishery options: remove the DW fleet in the North of region 2 (**noDW2N**), and include all 13 fisheries (**allf**)
- Two tuning criteria: tuned to ETBF CPUE (**CPUEtuned**) and to SSB relative depletion (**deptuned**)
- The particular tuning value: recent mean (2011–2016) CPUE (**recent**), 20% higher (**1.2up**), and 20% lower (**1.2down**) by 2035; for a given percentage value $x\%$ of SSB depletion (**xpc**)
- Two TAC schedules: every year (**1yrtac**) and every 3 years (**3yrtac**)
- Several robustness options: 5% and 10% symmetric migration (**mig5** and **mig10**) scenarios, respectively; low (50%) mean future recruitment for 5 years (**reclow**); lower M-at-age (**msens**); separate spawning stocks in each region (**sepstk**); change in migration to $m\%$ in year n (**xmigmyrn**); and a fractional z change in EU fishing pressure in *each* region (**zEUF**)

Obviously, all these different unique classifiers could all be pairwise run and would result in an enormous amount of results we would have to summarise. Indeed, arguably so much it would be hard to really see the standout significant factors. Table 5.1 details all the key combinations run in this initial round of MSE work which we suggest will be representative of all the key features we want to explore, without unnecessarily repeating very similar combinations.

Key MSE runs
noDW2N_CPUEtuned_recent_3yrtac
noDW2N_CPUEtuned_recent_1yrtac
noDW2N_CPUEtuned_1.2up_3yrtac
noDW2N_CPUEtuned_1.2up_1yrtac
noDW2N_CPUEtuned_1.2down_3yrtac
noDW2N_CPUEtuned_1.2down_1yrtac
allf_CPUEtuned_recent_3yrtac
noDW2N_deptuned_60pc_3yrtac
noDW2N_CPUEtuned_recent_3yrtac_msens
noDW2N_CPUEtuned_recent_3yrtac_mig5
noDW2N_CPUEtuned_recent_3yrtac_mig10
noDW2N_CPUEtuned_recent_3yrtac_reclow
noDW2N_CPUEtuned_recent_3yrtac_sepstk
noDW2N_CPUEtuned_recent_3yrtac_xmig10yr5
noDW2N_CPUEtuned_recent_3yrtac_2EUF
noDW2N_CPUEtuned_recent_3yrtac_0.5EUF

Table 5.1: *Particular naming convention covering all the MSE runs included in this paper.*

6 Interpretation of tuning principle

One of the key issues that has to be addressed in this MSE, and indeed in all the previous MSE work, is how much being able to change the ETBF TAC allows you to move (in whatever direction) the SSB depletion in the population. In cases where one controls effectively *all* the fishing pressure (be it in terms of a quota or effort), often the strict tuning principle is used (e.g. for SBT [2]) and a given target in a given year is met *exactly* via changing the parameters of the HCR.

For the ETBF case, clearly we do not have control of all the effort or catch taken from the stock. To explore this quantitatively for the reference case we look at: (i) ratio of (weighted) mean harvest rate for the ETBF and the whole of region 1; (ii) ratio of mean harvest rate for the ETBF and the whole of both region 1 and 2; and (iii) the mean ratio of loss rate to natural mortality and harvest rate averaged across both regions. For ratio (i) in the most recent assessment year is 0.7 (i.e. we control 70% of the fishing pressure via the ETBF); for ratio (ii) the value is 0.24 so we ony control 24% of the fishing pressure across the whole of regions 1 and 2; and for ratio (iii) the value is 0.25 so this shows that, averaged across the exploitable population, natural mortality (at least for the diagnostic case M vector) is more dominant than the fishing pressure. What this demonstrates is that, clearly, in region 1 we control most of the fishing pressure via the ETBF catch but that, across both regions, this reduces fairly strongly to only 24% so basically a quarter of the fishing pressure. Obviously, for the separate stock scenario where region 1 is a separate stock we would be in control of 70% of the relevant fishing pressure, but this reduces to 24% with a single stock. What is quite revealing is the comparison with natural mortality: for the reference case natural mortality seems to be quite a bit larger than fishing pressure but, even for the lower M-at-age vector, it is still larger than estimated fishing pressure (with a ratio of 60% instead of 24% for the reference case). What this suggests is that, no matter how much of the fishing pressure we control, our ability to really move the mature biomass in any given direction is somewhat limited given that estimated total fishing pressure is noticeable lower than the natural rate of attrition.

When it comes to interpretation of the tuning principle, we take a more nuanced approach given the - what we hope are now clear - limitations we have in terms of affecting fishery and stock outcomes via changing the ETBF catch in some fashion via the candidate HS structure. When it comes to tuning via historical mean CPUE levels we do the following: tuning is really trying to get the 2035 mean CPUE within the 90% CI of the CPUE tuning level (some fraction of the 2011–2016 mean CPUE). If we were interpreting the tuning principle *exactly* we would have to get the 2035 median CPUE to be exactly the mean value of the tuning level defined via the mean historical CPUE. We chose to relax this principle a little so we could explore variation in HSs while trying to attain the general tuning levels requested. For tuning to the SSB depletion level this is far stricter, as it is very linked to the distance between the current depletion and the target depletion level in the given year. For example, for the reference OM getting from the current estimated depletion level of around 0.65–0.7 to the 48%, a "strict" interpretation of the Commonwealth Harvest Strategy Policy (HSP), would require *very* strong changes in the ETBF TAC over the next 15 years.

7 Summary statistics

In this section we outline the initial summary statistics we use to look at the performance of the various HSs. In terms of (dynamic) SSB depletion we look at the depletion in the initial year (2016), the tuning year (2035), and the final year (2040). For the CPUE performance statistics we look at the *ratio* of the future simulated CPUE and the target CPUE (the mean from 2011–2016) for the same three years as in the depletion statistics. For the TACs we look at three statistics: distribution of TAC in the same three years (2016, 2035, 2040); the average inter-annual variation (AAV) over the time period for which the TAC is changing (i.e. 2020 onwards); and the average TAC change on the same time period as the AAV is calculated.

The secondary graphical summary we use are so-called worm plot [2]. Alongside the median and approximate confidence interval (90% in this case) we plot 10 randomly chosen trajectories (or worms) and plot them on top of the time summary of the following variables: SSB depletion, simulated CPUE, TAC and actual catch taken. The reason for doing this is both simple and at the same time complex: the median trajectory will **never** happen. It is a useful summary for the *range* of possible options but, given this is an intrinsically stochastic system, what will **actually** happen is one of the individual trajectories. For this reason we show the worm plots for several key scenarios to make it clear what "real" SSB, CPUE, TAC and actual catch trajectories will probably look like.

8 Results

The result section is split so as to organise the performance statistic summaries in useful groups to compare across the options without overpopulating the graphics.

8.1 Comparing tuning options

Figures 8.1 and 8.2 summarise the SSB/CPUE and TAC performance for the three CPUE tuning options, respectively. In the CPUE plots it is the ratio (relative to the target mean CPUE) that is plotted but, for reference back to “real world” values the current, 1.2up and 1.2 down expected CPUE values are 1.1, 1.32, and 0.88, respectively. For the recent CPUE tuning option the SSB depletion largely stays where it began (decreasing slightly over time); for the 1.2up option we see a slight improvement in depletion levels; for the 1.2 down option we see a far more significant movement downwards to around 0.58 by 2040. The most recent sub-adult CPUE for swordfish has been estimated to be below the average for the most recent years and, while this correlative effect is taken into account in the simulations going forward, eventually the model predicts that the CPUE *should* rise to around the 1-1.2 level in a few years. Subsequently, it takes substantial increases in catch to decrease the CPUE in the tuning year to the lower tuning level (0.88 with s.e 0.05). Figure 8.2 makes this clear in terms of the TAC summary statistics: future TACs for the recent case are generally around 1,500t by the tuning year, around 1,200 tonnes for the 1.2up scenario, but around 2,500 tonnes for the 1.2 down scenario. The OMs predict that counteracting the eventual increase in CPUE due to an increase in the future mean recruitment (relative to the 2011-2016 period) by increasing the TAC takes a lot more increase than the reverse case of trying to increase the CPUE by the same amount. The AAV levels for the 1.2down case are much higher (*ca.* 22%) than for the other cases where less TAC movement is required (*ca.* 5–10%); the same is true for the average TAC change of around 550t in the 1.2 down case, versus 50-100t for the other tuning cases.

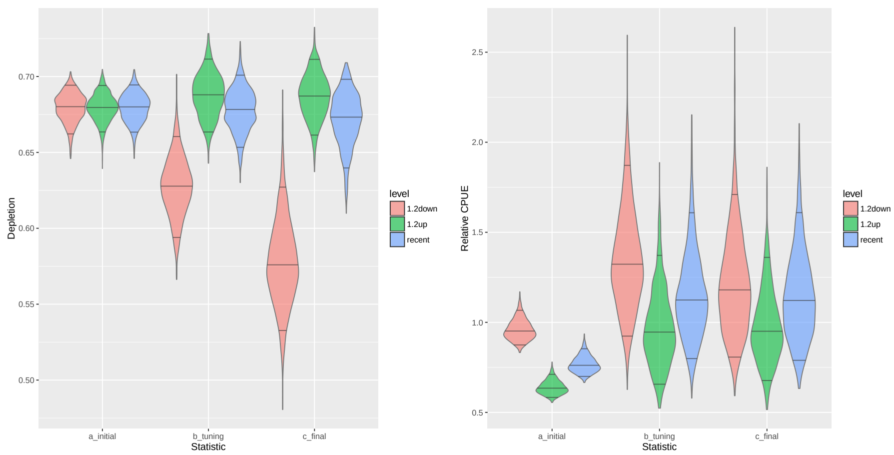


Figure 8.1: *SSB depletion (left) and relative CPUE (right) summaries for CPUE tuning options.*

8.2 Alternative TAC schedules

A major choice in a HS is how often we want to make a change to the current advice (be it a TAC or whatever). For this work the TTRAG requested we look at 1 year and 3 year options for the frequency of TAC change. Looking at Figure 3 for basically the same CPUE tuning

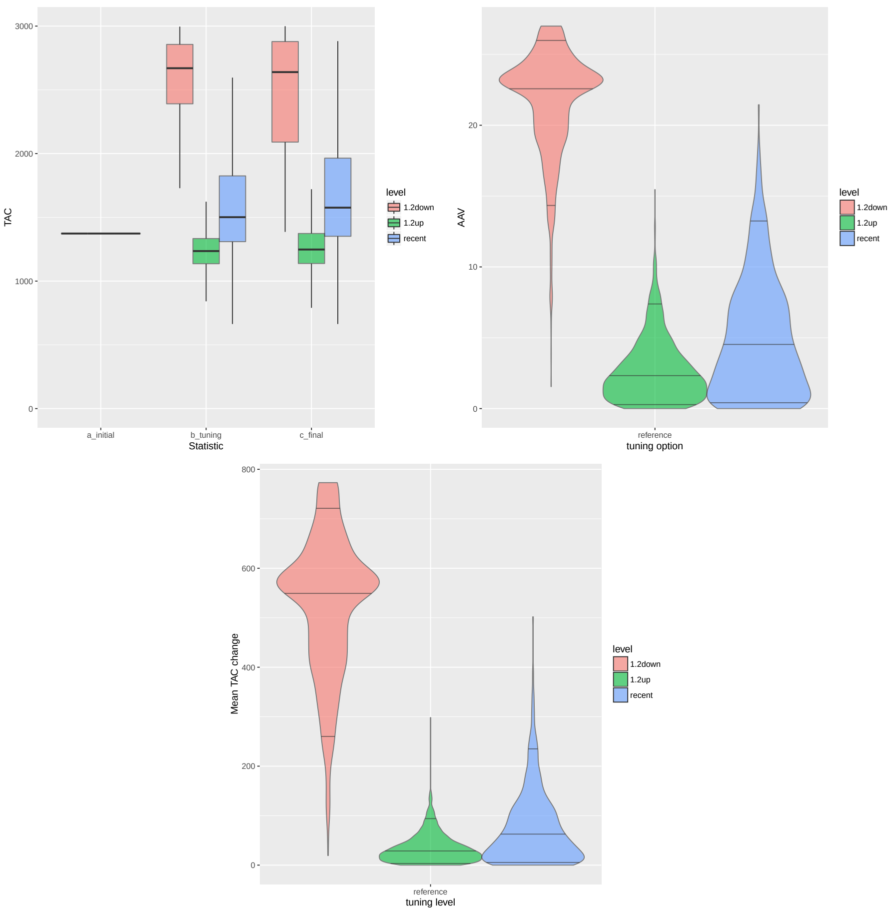


Figure 8.2: *Average TAC (top right), AAV (top left) and mean TAC change (bottom) summaries for CPUE tuning options.*

behaviour (the parameters of the HS are basically identical between TAC options) we see lower SSB depletion levels by the tuning period and end of the simulations when changing TAC every year - in particular lower tails in this distribution. Figure 8.4 shows how this is happening: average TACs are around 200t higher for the annual cycle and with longer tails in the upper distribution of the TAC (this causes the lower tails in the SSB depletion we see). In terms of AAV, in median terms, they are actually very similar around the 5% level but with a higher upper tail for the 3 year cycle. Not surprising given the 1 year cycle is more constrained (10% maximum change vs. 27% for the 3 year cycle). Average maximum changes are also very similar (around 50t in median terms) but with a higher tail for the 3 year cycle as one would expect. A very important point is this: the AAV and average change in TAC are very similar but the annual cycle does this every year so the one year cycle may be simply jumping around more than the 3 year cycle does over time while ending up with comparable average TACs over the projection period.

8.3 Alternative fisheries included in the OM

When focussing on the reference case fishery configuration (now DW2N) and the alternative scenario (include all fisheries) Figures 8.5 and 8.6 show the SSB/CPUE and TAC summaries, respectively. There is very little difference in the CPUE performance for these cases, but clear consistent differences in the SSB depletion performance. When including all the fisheries we see

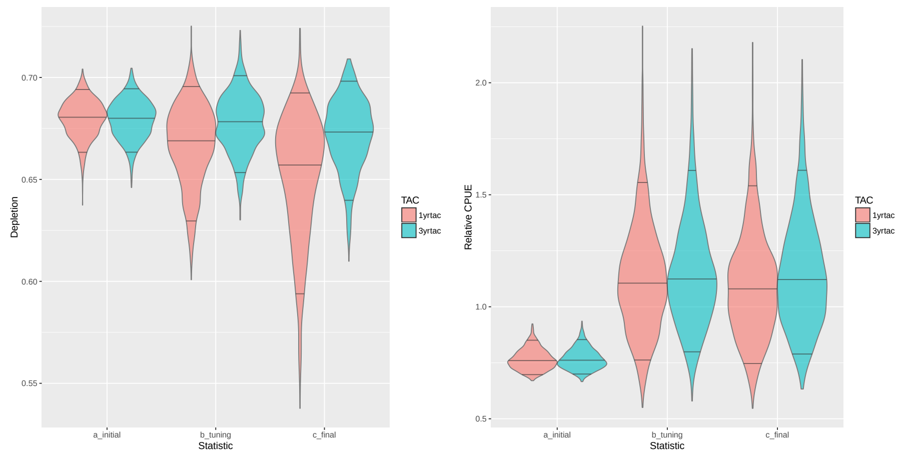


Figure 8.3: *SSB depletion (left) and relative CPUE (right) summaries for TAC frequency options.*

a lower starting level of depletion (*ca.* 62%) followed by some increase by the tuning year and a slight decrease by the final year. In terms of TAC comparisons we get slightly higher TACs for the reference case (given better starting status and overall population size) but generally slightly higher variance in mean TACs, annual variability and mean change.

8.4 Key robustness tests

Figures 8.7 and 8.8 summarise the SSB/CPUE and TAC performance, respectively, on the key robustness tests outlined previously. The robustness tests clearly have a wide spread in terms of initial SSB depletion estimates (median estimates from 0.45–0.75). Those initial status levels *generally* follow through to the tuning and final projection years for those scenarios where nothing fundamentally changes in the future relative to the past. For these cases we generally see a slight decline in the SSB depletion levels but they never get lower than around 40% and most are above 60%. For the scenarios where something really changes in the future (EU effort halves/doubles, migration changes into the future, mean recruitment is 50% lower for the first 5 years) the direction of the change is quite different depending on the test, but all the SSB depletion levels are almost always above 55%. In general, relative CPUE levels are fairly similar across all robustness scenarios and future points in time.

In terms of future TAC performance - apart from the **reclow**, **msens**, and **sepstk** - average TAC levels by the tuning and final years are at or above the current TAC and generally between 1,250–2,000 tonnes. One would hope/expect that the HS would reduce the overall TAC for the **reclow** robustness trial given what this tests but it does not do so pathologically (the lower 25%ile is above 1,000t) with regards to the ETBF and its viability. In terms of catch variability year to year the median AAV levels are between 3–8% and rarely exceed 15%, with average TAC changes between 40–110t and rarely exceeding 250t (every 3 years).

8.5 Key worm plot summaries

The previous summaries have given a high-level feel for how the candidate HS will perform across the various scenarios. However, for a more detailed look at how the HS might perform over time we now move to looking at specific scenario worm plots: individual possible time-trajectories of depletion, ETBF CPUE, TAC and actual catch taken.

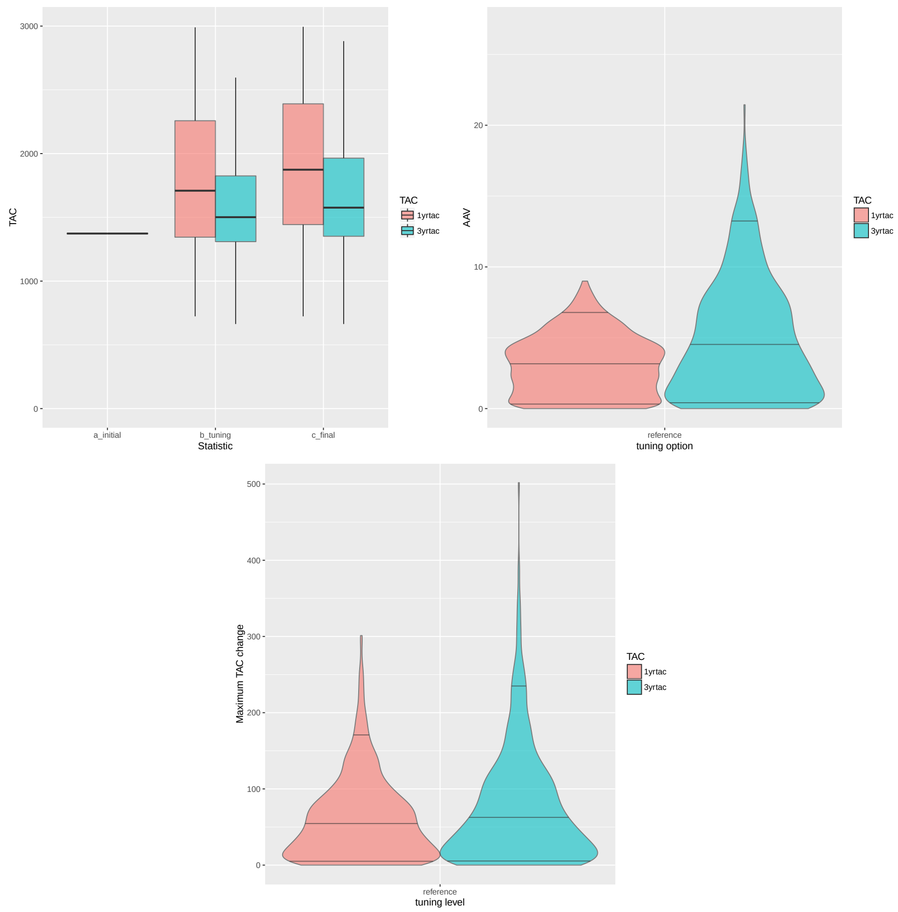


Figure 8.4: *Average TAC (top right), AAV (top left) and mean TAC change (bottom) summaries for TAC frequency options.*

Figure 8.9 shows the worm plot summary for the reference set of OMs and the 3 year TAC frequency. With regards to depletion, we see the fairly constant overall time trend but with the recruitment stochasticity-driven variation over time. In terms of CPUE we see the fairly quick increase in CPUE from the current level (estimated to be below average given stock abundance) to a mean level of around 1.05–1.35. The TAC on average stays fairly constant until around 2025, and then begins to increase as the CPUE increase flows into the HS index. This pattern is reflected in the actual catch taken, but with the variability due to the random under-catch we see historically and which is modelled into the future. Figure 8.10 shows the reference OM worms but for an annual TAC frequency. The depletion summaries are basically identical up to 2030, but the annual TAC scenario shows a lower tail in the SSB depletion by the tuning (2035) and final (2040) years and with individual worms showing clear downward trends from 2030 onwards. The main difference is that, although the TAC begins increasing by around 2025 as with the 3 year cycle, it is permitted to change the TAC every year and, given the on average increasing and above-target CPUE by 2025–2030, the TAC begins increasing on average at a faster rate than for the 3 year cycle. This more rapid increase (and fatter upper tail to the TAC distribution) is what drives the lower tail in the SSB depletion.

Figure 8.11 shows the worm plot summary for the lower natural mortality robustness test. As we know, the starting SSB depletion is lower than for the reference set of OMs (around 0.44 *vs.* 0.68 at the start). That being said, the time-trends in depletion are basically identical. Given the,

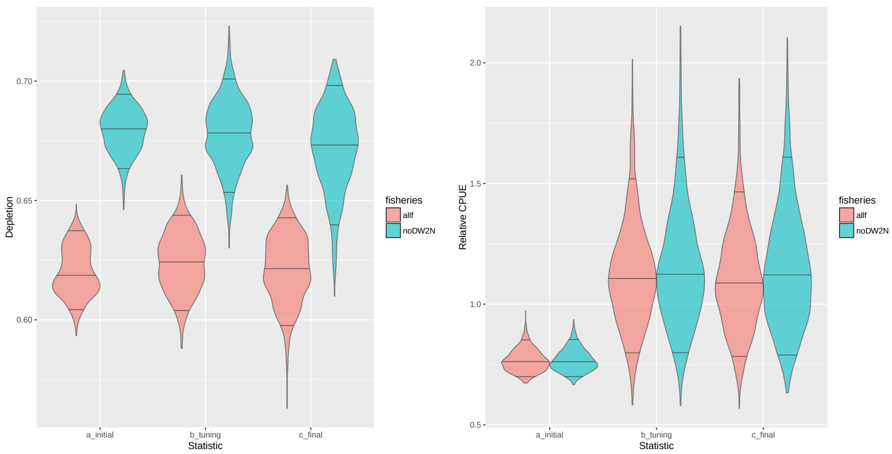


Figure 8.5: *SSB depletion (left) and relative CPUE (right) summaries for fishery inclusion options.*

on average, lower mean CPUE for this particular scenario, the average TAC stays constant (with low variability over time) and gradually increasing from the tuning year. To show the difficulty with tuning the swordfish HS to an actual depletion target, look at the worm plot summary in Figure 8.12. The tuning involves changing key parameters in the HCR so that the *median* SSB depletion by 2035 is 0.6. To be able to do this by changing *only* the ETBF TAC we have to strongly increase the TAC and keep increasing it to reach the tuning target. Two obvious issues with this behaviour are: (i) it is unclear that TACs of 4,000t and above is even catchable given the current fleet size and multi-species facets; and (ii) this increase in TAC eventually causes the average CPUE to begin to decrease (backing up the theory of economic efficiency of catches well above historical levels) and the SSB depletion is going down rapidly and would likely continue to decrease until the ETBF TAC is eventually decreased (as would happen eventually given the CPUE decreasing up to the tuning and final year). These are the problems encountered when trying to tune to an SSB depletion of 60% (not a million miles away from the starting value of 0.68). One can imagine that it would be basically impossible to tune the depletion to the 40% and 48% MSY and MEY proxies - even when tuning the CPUE to current and $\pm$20% levels we necessarily invoked a weaker notion of tuning (making sure they CI of the means overlap). This is fundamentally driven by the limitations of the HS to move the stock in any given direction because of the small fraction of both overall fishing pressure we control in the ETBF TAC and the generally low ratio of fishing pressure to attrition in biomass due to natural mortality. Figure 8.13 shows the worm plot summary for the **reclow** robustness trial. We see depletion, with around a 5 year delay due to maturation, decrease and then begin to increase as normal mean recruitment returns and the ETBF TAC is slowly decreasing making sure the return in recruitment numbers is allowed to rebuild the SSB (one can see how the TAC is only strongly decreased for the very low recruitment scenarios - the HS seems suitably responsive without being punitive).

9 Overall summary & Future Work

In this document we have outlined the key robustness scenarios and the performance of the various candidate harvest strategies on those particular tests. While often discussed in general terms by the RAG many times we have also outlined in quantitative terms exactly how much "leverage" we have in obtaining certain SSB depletion outcomes for the whole stock, when con-

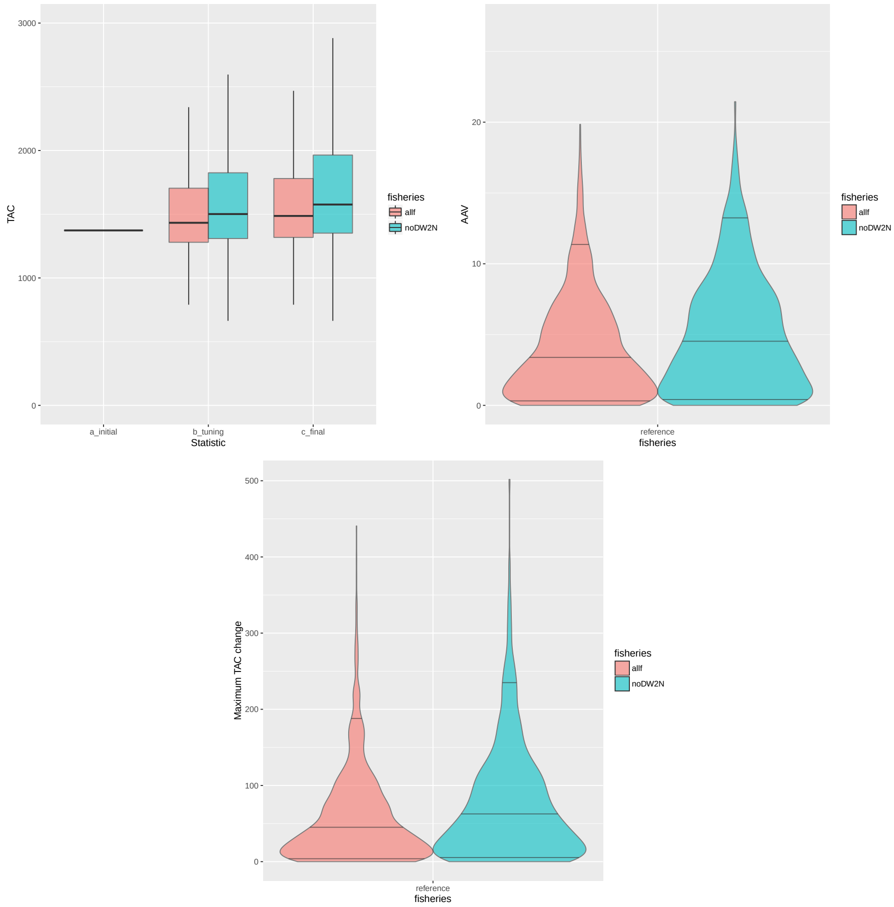


Figure 8.6: *Average TAC (top right), AAV (top left) and mean TAC change (bottom) summaries for fishery inclusion options.*

trolling only the ETBF catch. In the current reference scenario (i.e. one single stock across both spatial regions) that amounts to around 25% of the total fishing pressure, but for our spatial region that number rapidly increases to 70%. Levels of fishing pressure are also fairly low compared to any of the natural mortality scenarios, re-enforcing the notion that natural attrition is higher than fishing mortality for the most part and across a wide range of age-classes. This further limits the ability of any HS that controlled even all the fishing pressure in the region to move the SSB in a given direction over medium time frames such as those explored herein. The main point is this: we can have a sensible and responsive HS which is apparently well-informed by the sub-adult CPUE, but we have to be cognisant of the fact that there are fairly strong limits on what we can accomplish at the full regional scale in the given time-frame.

We also showed that this limitation really makes it difficult to tune the HS in the sense of achieving the 38%, 48% and 58% SSB depletion targets. We did manage to tune (in the weaker sense of this concept) to the future mean CPUE targets and we did see clear differences in the candidate harvest strategies for these three possible tuning options. Overall, future levels of depletion in the tuning year were positive ranging from a low of around 44% to a high of almost 75% but mostly within the 60–70% range. In terms of TAC performance levels of AAV (reactivity from change to change) were fairly low (in the 5–10% range or mean TAC changes between 50–200t). Only for more negative scenarios did the mean TAC decrease by the tuning year and were mostly around the 1,250–1,800t level, with the attendant variation seen in the worm plots.

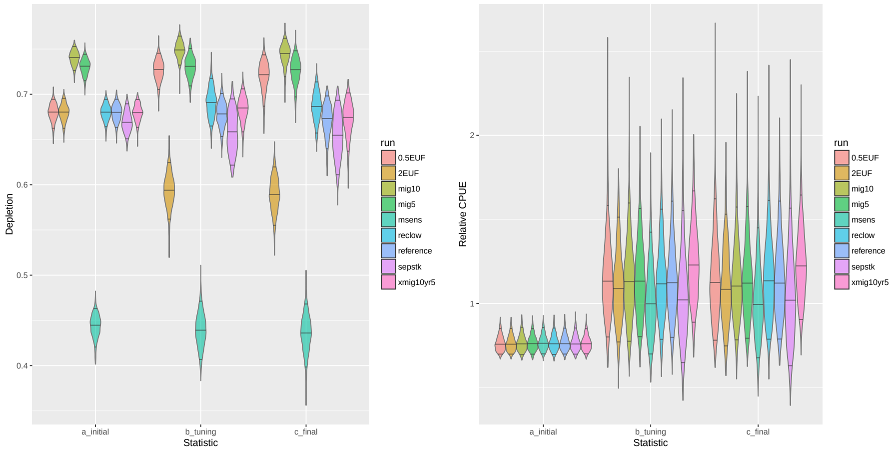


Figure 8.7: *SSB depletion (left) and relative CPUE (right) summaries for the key robustness tests.*

In terms of TAC frequency, given we limit the 1 year TAC change to 10% relative to 27% for the 3 year cycle, there is not huge variation in performance for the 1 year cycle relative to the 3 year cycle. In general, there is more overall variation in TAC when one accounts for the fact that it changes annually in this case. Additionally, the more sustained increase in TAC in the later years for the annual cycle results in lower tails in the SSB depletion by the tuning and final years. One would probably see more to like in the 3 year cycle performance wise, but because we have limited the amount by which the annual TAC can change relative to a 3 year cycle, this difference is not as large as it probably would be if they both had the same maximum percentage change values.

The alternative EU effort scenarios we explored (double or half the current level) show that SSB depletion outcomes and, to a lesser degree, average TAC values are still equally dependent on what happens with the other major players (in terms of fishing pressure) in the region. In terms of low recruitment in the future (5 years at 50% of the expected level) the HS seems to deal well with this scenario, reducing TACs when needed (and strongly for very low recruitment trajectories) but not unnecessarily so and, as a result, future depletion levels actually slightly increased. Even for the most pessimistic robustness scenario, the lower natural mortality scenario, starting depletion levels are at around 45% and decrease very little over the full projection period and with an effectively zero probability of every being lower than the 20% limit level. For the migration scenarios and stock structure scenarios we explored none of these appeared to result in significant negative outcomes relative to the reference case - for migration levels of 5% and 10% in both directions it is arguably slightly more positive than the reference scenario.

In terms of future work, do we need to explore additional parts to the harvest strategy or alternative indices? To justify exploring say a recruitment type index, such as the CPUE or mean length of the size ranges lower than the sub-adult cut-off we would have to say: what current problem would it solve? For the fairly extreme low recruitment scenario we explored, where mean recruitment halves for 5 years in a row, the HS seemed to perform fairly well and did what we would probably want it to do. Given this, it seems unlikely at this stage that the addition of a more recruitment focussed index embedded within the HS will result in the kind of performance improvement that would justify the additional complexity we have introduced. Additionally, should we explore using the adult/sub-adult compound index or even an all size class CPUE index?

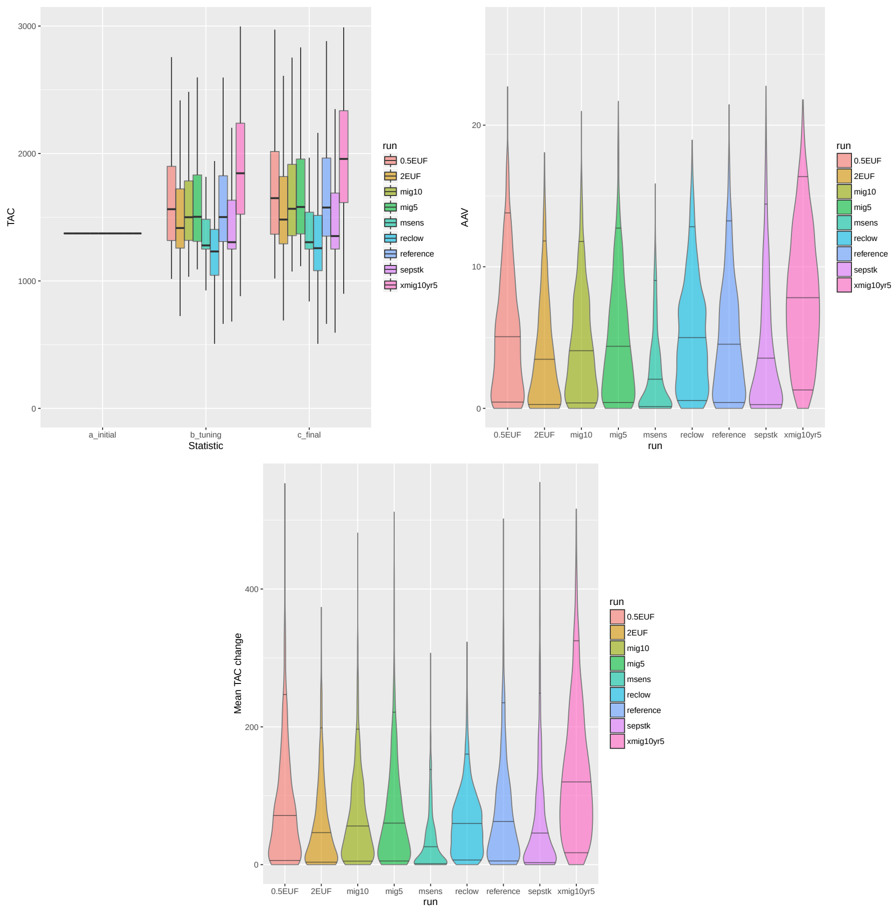


Figure 8.8: *Average TAC (top right), AAV (top left) and mean TAC change (bottom) summaries for the key robustness tests.*

Again, the HS using the sub-adult index seems to be able to respond to both (significant) recruitment variation and changes in the adult biomass of the stock. It is not clear that exploring additional types of CPUE indices in the harvest strategy will solve any obvious problems that we can currently identify.

Summing up, it seems we can construct a harvest strategy using *only* the sub-adult CPUE index that can perform adequately in relation to the constraints outlined by the TTRAG and the CPUE targets it recommended we explore. When it comes to tuning the HS to explicit SSB depletion, we run into the limits on how much changing ETBF TAC can actually achieve in terms of whole-stock outcomes. We can apparently control the ETBF TAC with a full-feedback tested HS that appears to do sensible things when confronted with positive and negative alternative scenarios, but we cannot really drive specific SSB depletion outcomes that are far away from where we currently think the stock might be in terms of status. The general levels of status estimated in the MSE work are usually within the 60-70% range, and are never estimated to be below around the 40% level over the whole projection period. This suggests there are not any obvious issues in terms of not being able to justify any of the candidate harvest strategies when considering the overarching objectives of the Commonwealth Harvest Strategy policy.

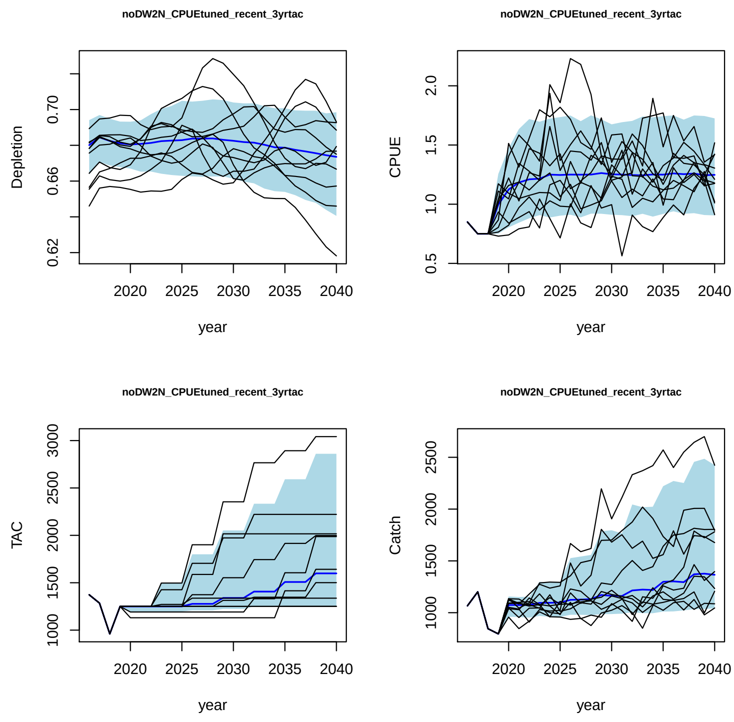


Figure 8.9: *Median (and 90% CI) and 10 random worms for the reference set of OMs.*

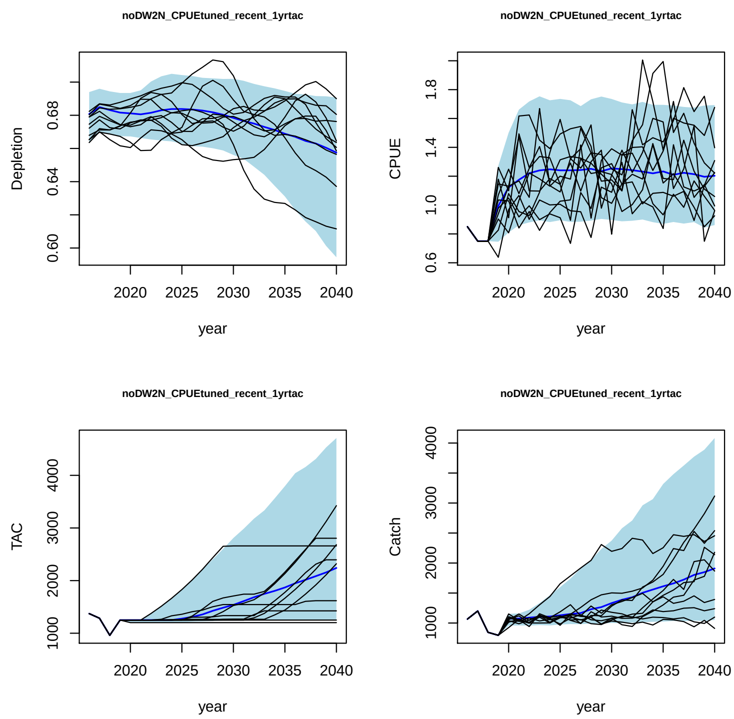


Figure 8.10: *Median (and 90% CI) and 10 random worms for the reference set of OMs and for the 1 year TAC frequency.*

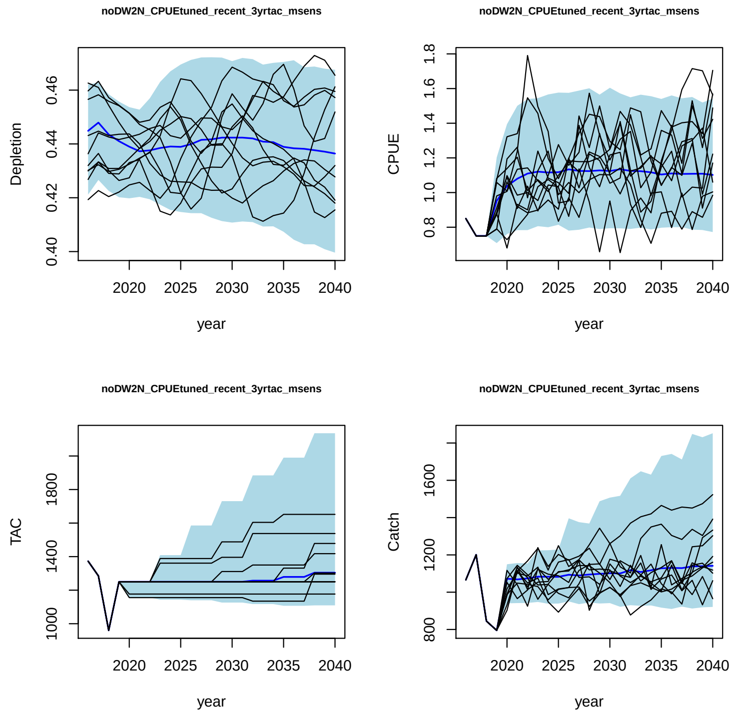


Figure 8.11: *Median (and 90% CI) and 10 random worms for the lower M robustness trial.*

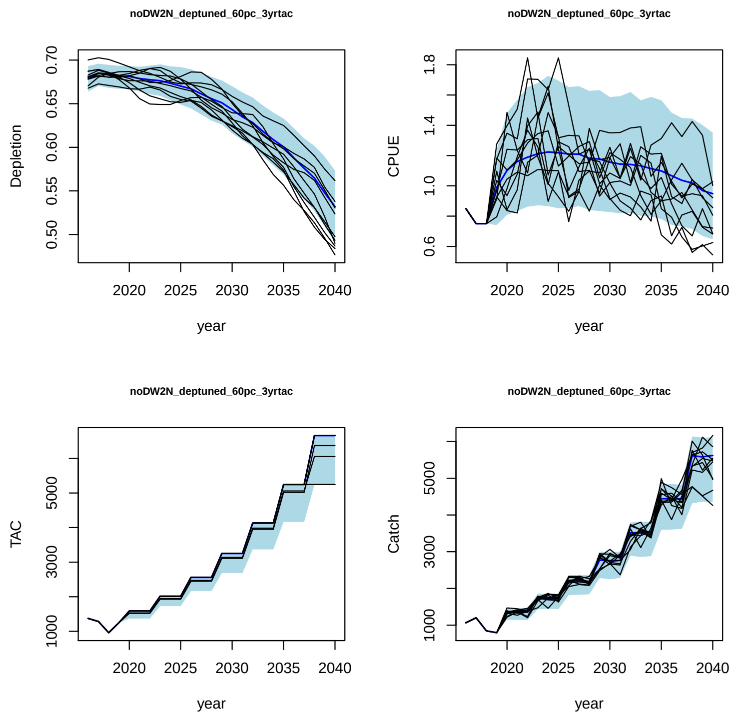


Figure 8.12: *Median (and 90% CI) and 10 random worms for the reference set of OMs with an HS tuned to reach 60% depletion by the tuning year of 2035.*

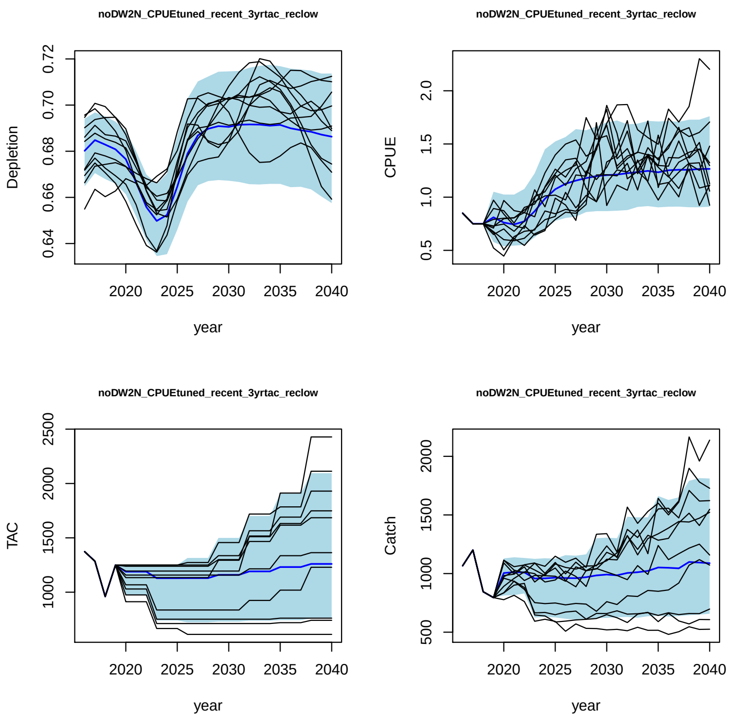


Figure 8.13: *Median (and 90% CI) and 10 random worms for the low recruitment robustness trial.*

References

[1] R. Hillary and A. L. Preece (2020) Conditioning of the revised Operating Models for the ETBF harvest strategy revision. *TTRAG, Canberra, February 2020*

[2] R. Hillary *et al.* (2016) A scientific alternative to moratoria for rebuilding depleted international tuna stocks. *Fish and Fisheries* **17**: 469–482.

CONTACT US

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

WE DO THE EXTRAORDINARY EVERY DAY

We innovate for tomorrow and help improve today for our customers, all Australians and the world.

Our innovations contribute billions of dollars to the Australian economy every year. As the largest patent holder in the nation, our vast wealth of intellectual property has led to more than 150 spin-off companies.

With more than 5,000 experts and a burning desire to get things done, we are Australias catalyst for innovation.

WE IMAGINE. WE COLLABORATE.
WE INNOVATE.

Appendix 6

Australia's National Science Agency

Management Strategy Evaluation for Striped Marlin

March 2021

Ann Preece

Contents

1 Introduction

The Operating Model software has been updated for Management Strategy Evaluation (MSE) of a new Harvest Strategy (HS) for striped marlin. The operating models use the most recent assessment of the striped marlin stock in the SW Pacific Ocean (SWPO) as the initial conditions for projections into the future for testing HSs under different conditions. This work builds on previous projects examining performance of harvest strategies for striped marlin and the other key species in Australia's Eastern Tuna and Billfish Fishery (ETBF). The operating model structure is described in Preece and Hillary (2020).

Candidate harvest strategies use the form of the Harvest Control Rule (HCR) agreed to by the TTRAG members in December 2019 for swordfish. Multiple candidate harvest strategies can be developed by varying the parameters of the HCR. The TTRAG 29 (Sept 2020) provided advice on the types of candidate HCRs to examine, and the scenarios that they should be tested against. Summary figures and statistics on performance of the harvest strategies compare performance primarily in future CPUE, catch and biomass. A single harvest strategy will be chosen for future implementation.

2 A new striped marlin harvest control rule

The TTRAG has agreed to use a single CPUE index (all sizes combined) in candidate harvest strategies for striped marlin. Details on the CPUE standardisation are in Campbell (2019), and the 2020 update is in Dell et al (2020), and no further examination of the CPUE standardisation is considered here.

The form of the HCR is specified in Hillary (2020) and has been built into the striped marlin operating model code (Figure 1). This HCR uses the one CPUE series (all sizes) for striped marlin, as the only input data, and specifies a target reference level of CPUE which is used to define actions to increase or decrease total allowable commercial catch (TACC). The HCR incorporates a buffer zone around the reference level of CPUE, so that the HS is not overly reactive to variation in CPUE. The HCR also incorporates a threshold at half the reference CPUE level, which triggers stronger decreases in TACC. The TTRAG defined the reference level as the average CPUE in 2012-2015. The previous striped marlin HS (Kolody et al. 2010, Hillary et al. 2013) specified average CPUE in 1998-2002 as the target.

The HCR reacts to recent CPUE to adjust the recommended Catch in the ETBF:

- If recent CPUE is in a buffer zone around the defined target reference level, then recommended total allowable commercial catch (TACC) is unchanged.
- If recent CPUE is higher than the CPUE in the buffer zone, then increase recommended TACC.
- If recent CPUE is lower than the CPUE in the buffer zone, then decrease recommended TACC.
- If recent CPUE is very low, less than or equal to half the CPUE levels in the buffer zone (threshold), then strongly decrease recommended TACC.

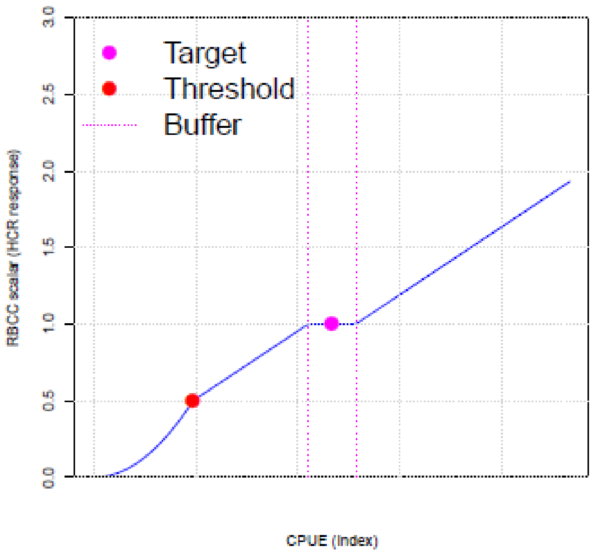


Figure 1 The form of the harvest control rule. There is a buffer zone around the average CPUE reference level, where TACC is unchanged, and a threshold point at half the CPUE reference level, below which TACC multiplier is strongly decreased. Source: Hillary (2020).

3 Candidate Harvest Strategies

Alternative candidate harvest strategies are based on the agreed form of HCR and differ via alternative values for several parameters:

- HS1 has the same settings as the swordfish HS adopted in 2020 and is the base HCR from which others vary. The CPUE input to the HCR is the average of the 3 most recent years of data, the maximum TAC change is 10% and TAC is set annually.
- HS2 uses the average CPUE over most recent 2 years as input to the HCR
- HS4 uses average CPUE over most recent 4 years.
- HS3 sets the TAC every 3 years, with maximum TAC change of 27%.
- HS5 uses a narrower and symmetrical buffer zone, set at 10% above and below the target level.
- HS6 uses a narrower and symmetrical buffer zone, set at 20% above and below the target level.

4 Reference set and projection options

Each HS is tested using the reference set of operating models. The reference set of operating models aims to encompass the range of plausible scenarios of future dynamics of the stock and fisheries. The MSE simulation process tests the performance of the HS by using it to set TACCs into the future under the different operating model conditions. The performance statistics are generated by combining results from each operating model replicate simulation (180 models). The simulations are run for 20 years (to 2040) and performance statistics are measured at 2035.

The results of the 2019 SWPO stock assessment indicate that the spawning stock is currently at 0.196 (0.1-0.34, 80%PI) of unfished biomass levels, and concludes that the stock is likely overfished and is close to overfishing (Ducharme-Barth et al., 2019).

From the set of stock assessment models used to provide stock status advice at WCPFC SC 2019, a subset of 18 are used as a reference set for MSE. These are the cross-combination of the key factors that affect stock status estimates: 3 values for Steepness (productivity of the stock, 0.65, 0.8, 0.95), 3 values for natural mortality (0.3, 0.4, 0.5), and 2 CPUE options (Japanese LL 2 CPUE series, and Australian LL CPUE series). The other factors (growth, size frequency weighting and recruitment penalty cv) are set at the same level used in the stock assessment diagnostic case. Additional terms define a projection set and 10 replicates are run for each model, giving a reference set of 180 models. Proj1 defines the reference set, and the other projection sets are used to test robustness of the HS to alternative projection scenarios.

Proj1 uses the reference set of models x 10 stochastic replicates and has the following settings:

- 2 areas, ETBF area and non-ETBF area
- Migration 20%/quarter (both directions)
- Non-ETBF effort is fixed at current levels
- Additional error in future estimates: Implementation error (catches taken can vary from the TAC), CPUE error, recruitment error, autocorrelation and additional error on first 2 age classes at the start of projections.
- The selectivity in the ETBF area is Australian LL and in the non-ETBF area is Japanese LL2.

Proj2 has migration of 1% / q.

Proj3 has higher autocorrelation in recruitment. The assessment model recruitment is highly variable.

Proj4 has higher effort in the non-ETBF area (effort outside the ETBF gradually doubles over the first 5 years of projections and stays at that level).

Proj1 has a narrower range of estimates of current stock status than the full set of models in the 2019 stock assessment. Additional variability is incorporated into the projections to ensure that the HS is robust to a wide range of conditions.

5 Harvest Strategy Performance

The data used in the HSs has been updated with the new CPUE series and recent ETBF STM catches through to 2019 (Campbell, 2000). The projections start from 2017 (the end year of the assessment models), TACs are fixed at 351t through to 2019 and ETBF catches are fixed at actual catches taken through to 2019. The HS decision rule is used to set the TAC from 2020 onwards.

The performance of candidate harvest strategies is examined in the figures below.

5.1 Constant Catch 0 projection

To provide an estimate of the impact of the ETBF catches on stock indicators and status measures, a zero constant future catch in the ETBF HS is examined (CC0_stm_2020 Figure 2). The results indicate that median SSB depletion at 2035 is at the 1980s-2000s levels, and the lower bound (10th percentile, blue area) is above median estimates of depletion in 2017. Historic and future recruitment is highly variable and 2035 median recruitment is at 1980-2000 levels. CPUE in 2035 would be higher than the target and recent levels.

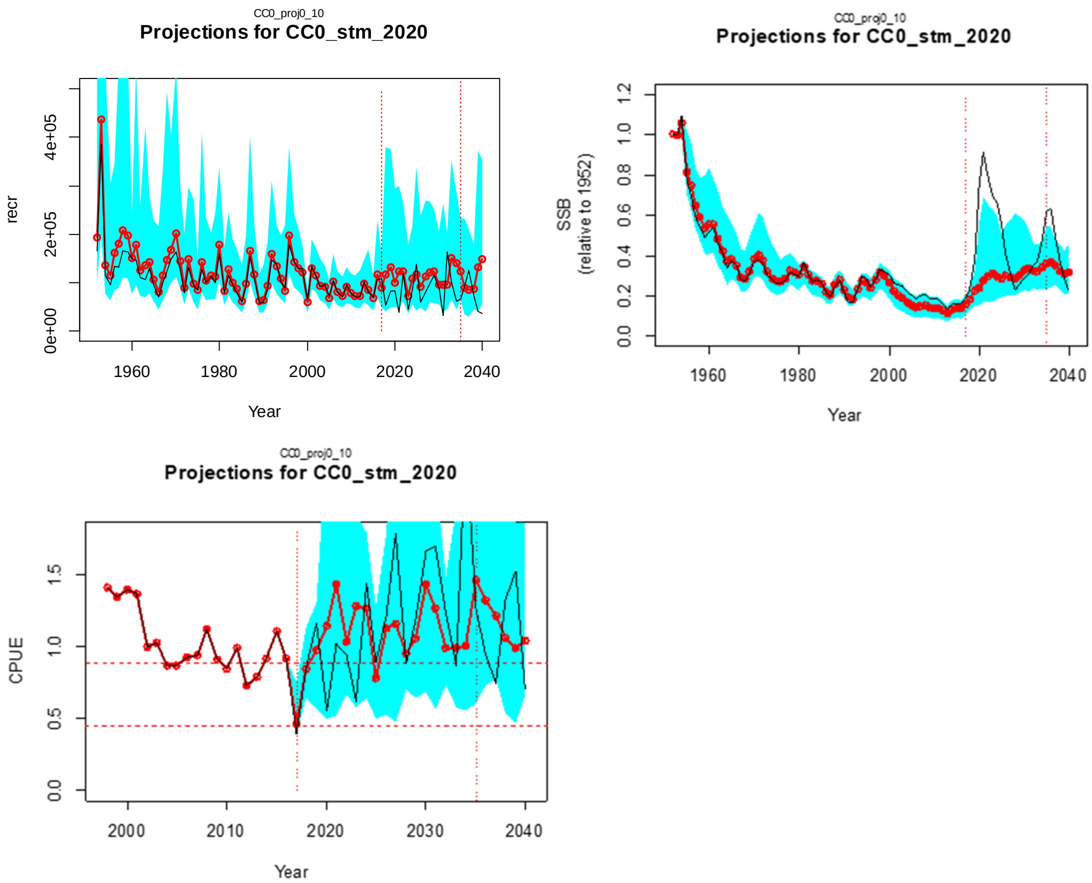


Figure 2 Examination of recovery of the stock under a constant zero catch in the ETBF, with continued fishing in the non-ETBF area at current effort levels. SSB, CPUE and recruitment. Red circles indicate median projections, shaded blue region represents the 10th-90th percentiles of the distribution, the thin black line represents a random trajectory. The vertical dashed line at 2017 indicates the start of projections and the vertical line at 2035 indicates the tuning year.

5.2 Exploration of HS1 - same settings as the swordfish HS.

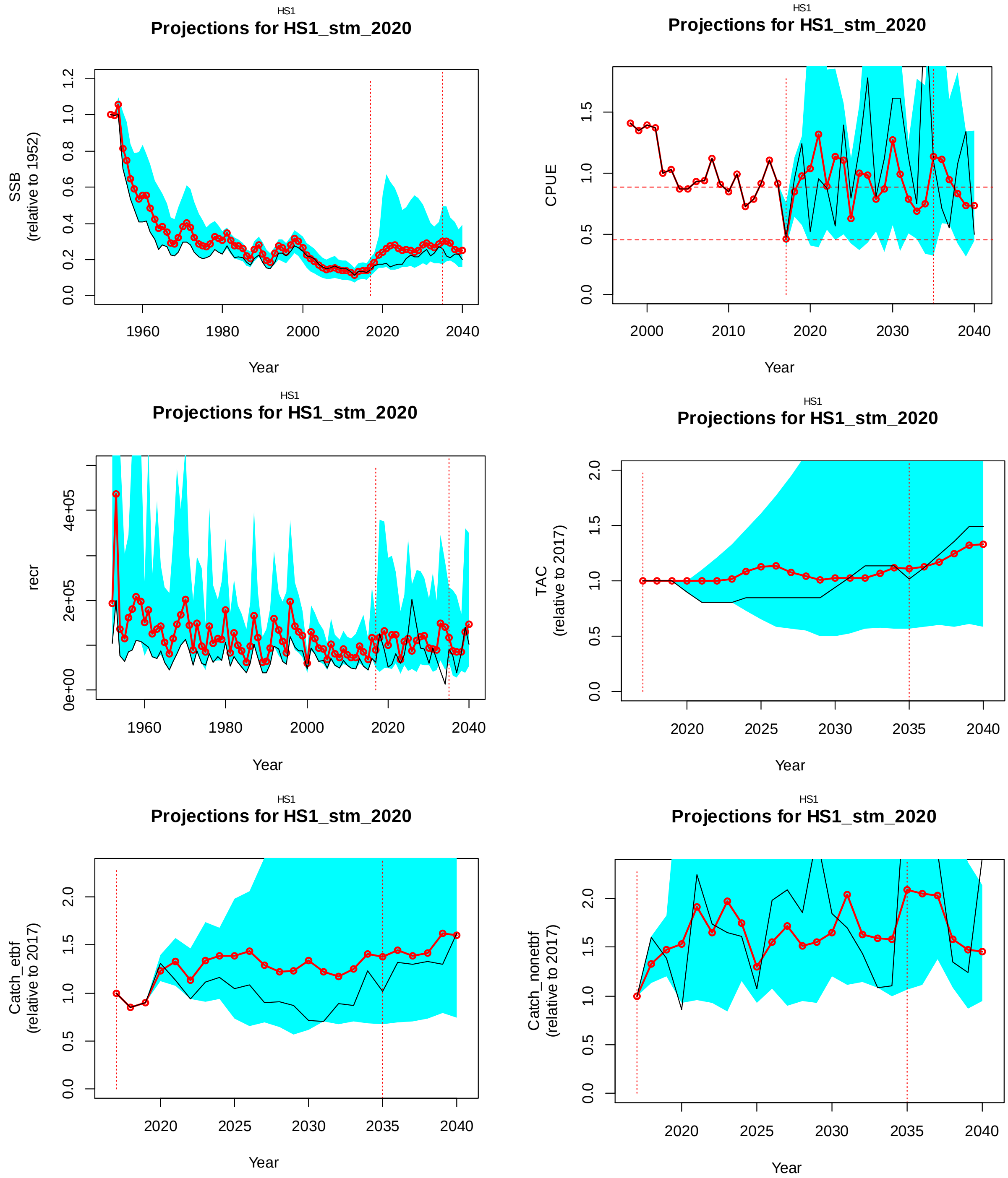


Figure 3 Examination of HS1. Red circles indicate median projections, shaded blue region represents the 10th-90th percentiles of the distribution, the thin black line represents a random trajectory. The vertical dashed line at 2017 indicates the start of projections and the vertical line at 2035 indicates the tuning year.

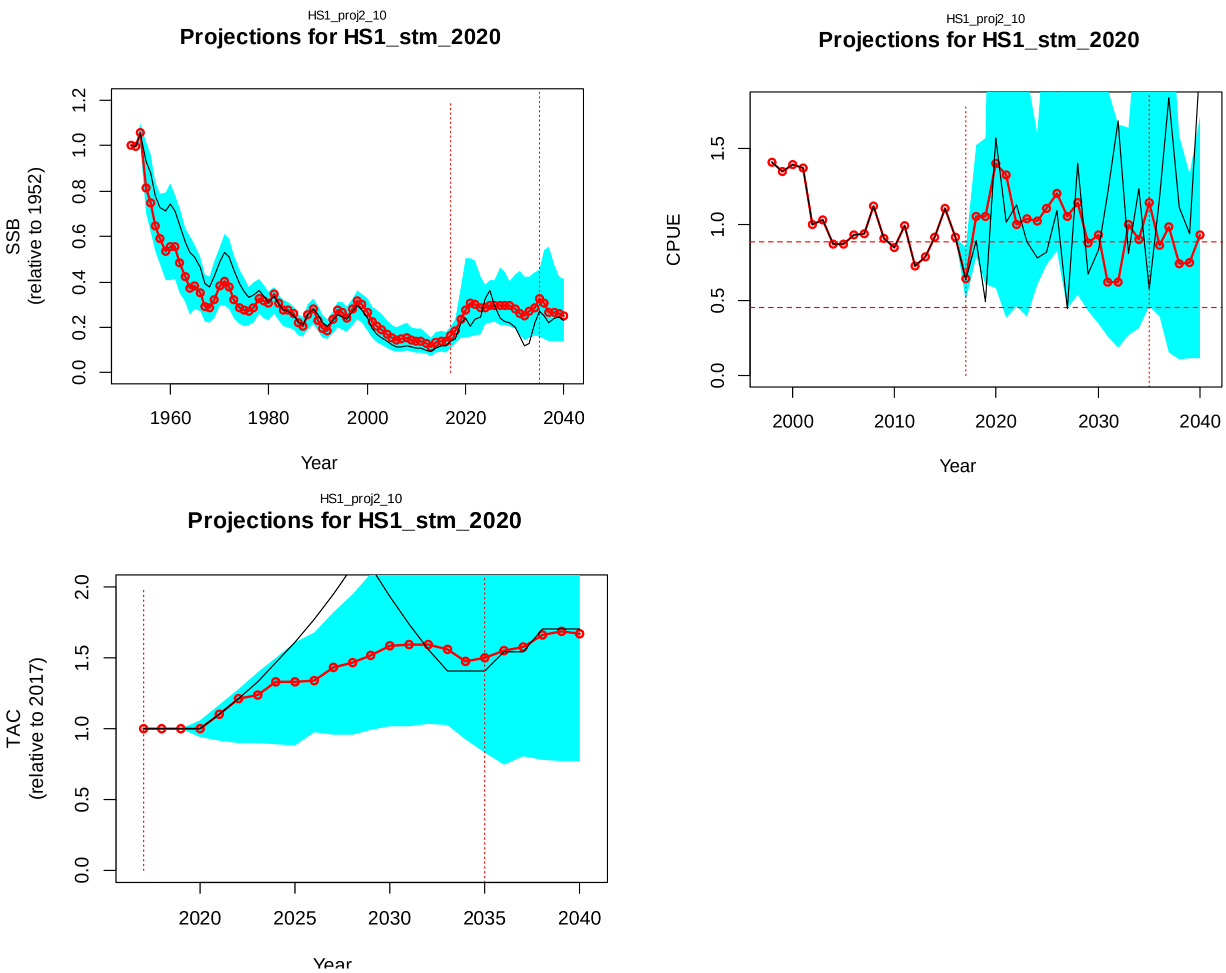
HS1 has the same settings as the swordfish HS adopted in 2020 and is used to set future TACs (from 2020 onwards) under the conditions defined in 'Proj1' which is the reference set of operating models and additional uncertainty settings. Results (Figure 3) of projections indicate that SSB depletion in 2035 and the lower bound (10th percentiles, area in blue) are above median estimates for 2017. Median CPUE oscillates around the target CPUE level and is above the target in 2035. The lower 10th percentile of CPUE estimates occasionally drops below the threshold level in the HCR. There are no unusual patterns in future recruitment. Median TAC in 2035 is above current levels, has an increasing trend over the projection period, but the range of values is wide. The catch figures show the comparative catch estimates in the ETBF and non-ETBF, and both are scaled relative to actual catches in 2017.

HS1 provides a base against which performance of other HSs, or HS1 under different conditions, can be explored.

5.3 Reduced migration in projections - HS1-proj 2: migration 1%

In Proj2 migration between the ETBF and non-ETBF area is 1%, which represents a scenario where the ETBF has lower connectivity with the non-ETBF area. In contrast, in Proj1 there is higher connectivity between areas with migration 20%.

For this scenario (Figure 4), there appears to be a decreasing trend in CPUE over the projection period, but median estimates of CPUE and SSB depletion in 2035 are above the Proj1 scenario estimates. The TAC estimates under this scenario are substantially larger. Total biomass in ETBF area is higher than the historic 1980-2000 period, whereas non-ETBF total biomass returns to 1980-2000 historic levels.



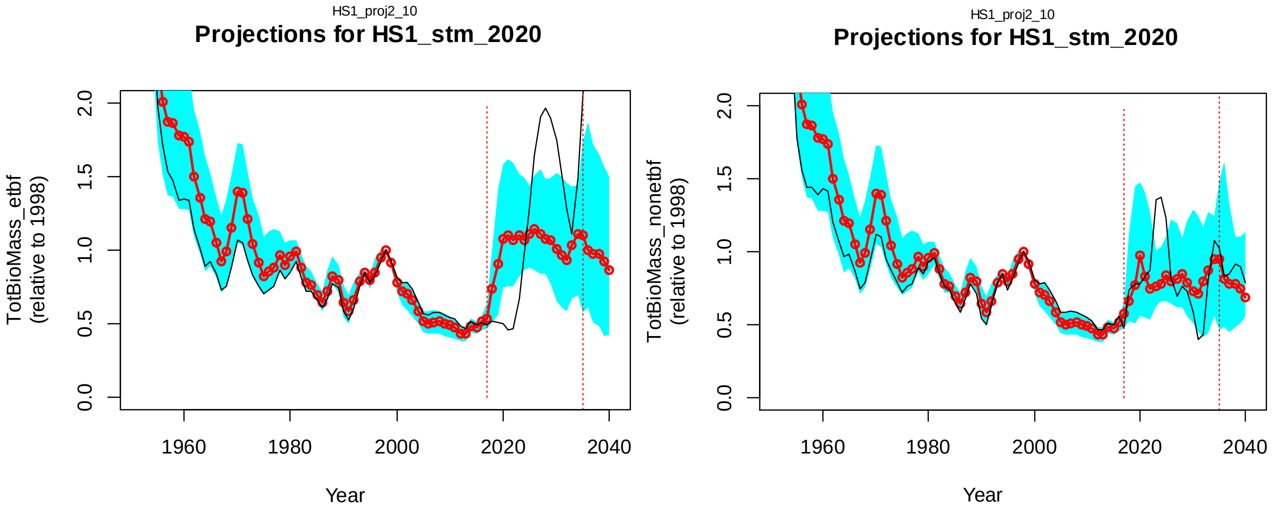


Figure 4 Examination of reduced migration effects on HS1. Red circles indicate median projections, shaded blue region represents the 10th-90th percentiles of the distribution, the thin black line represents a random trajectory. The vertical dashed line at 2017 indicates the start of projections and the vertical line at 2035 indicates the tuning year.

5.4 Frequency of TAC changes – HS3 Proj1: 3yr TAC changes

HS3 examines the effect of a 3 year TAC change (set 3 yr TAC block, max TAC change = 27%) and results are compared to HS1 which has a 1 year TAC cange with max of 10%.

The SSB and CPUE performance by 2035 were very similar (not shown) to HS1, but median TAC is lower (Figure 5).

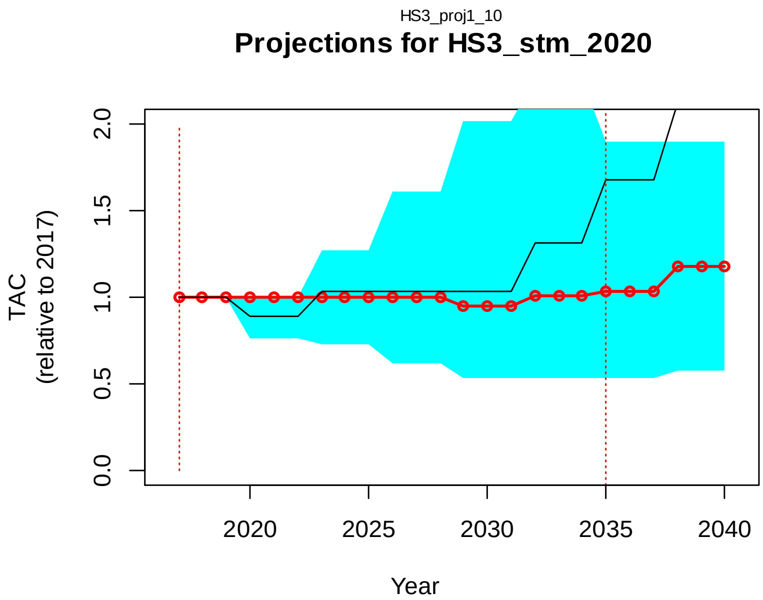


Figure 5 Examination of HS3 – 3-year TAC change with max TAC change 27%. Red circles indicate median projections, shaded blue region represents the 10th-90th percentiles of the distribution, the thin black line represents a random trajectory. 2017 is the start of projections and the vertical line at 2035 indicates the tuning year.

5.5 Examining recruitment autocorrelation uncertainty -Proj3

Changing the recruitment autocorrelation has an effect on the performance of HS1 (Figure 6). If there is higher autocorrelation, the median CPUE estimate in 2035 reaches the target CPUE but does not exceed it as it does in the Proj1 scenario results.

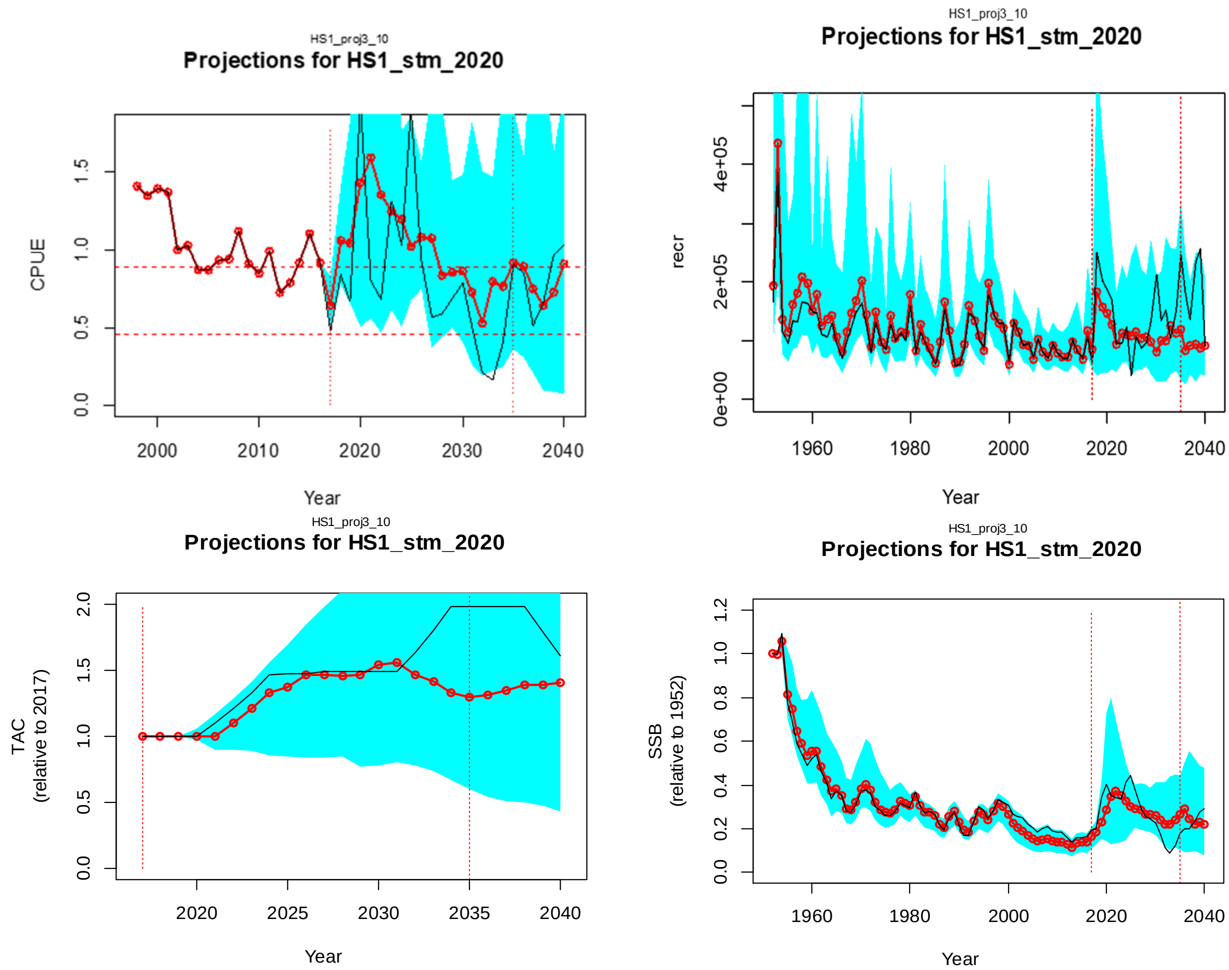


Figure 6 Examination of effect of changing the recruitment autocorrelation variable. Red circles indicate median projections, shaded blue region represents the 10th-90th percentiles of the distribution, the thin black line represents a random trajectory. The vertical dashed line at 2017 indicates the start of projections and the vertical line at 2035 indicates the tuning year.

5.6 Exploring number of years for recent average CPUE- H2 & H4

The HCR uses the recent average CPUE as the input data and in HS1 this is averaged over 3 years. Two alternatives are examined. In HS2 the average is over the most recent 2 years, and in H4 the average is over 4 years. There are some small differences in these results (Figure 7); median 2035 TAC is higher for HS4 and median CPUE 2035 is lower and closer to the target. Median SSB depletion is the same in both HS2 and HS4.

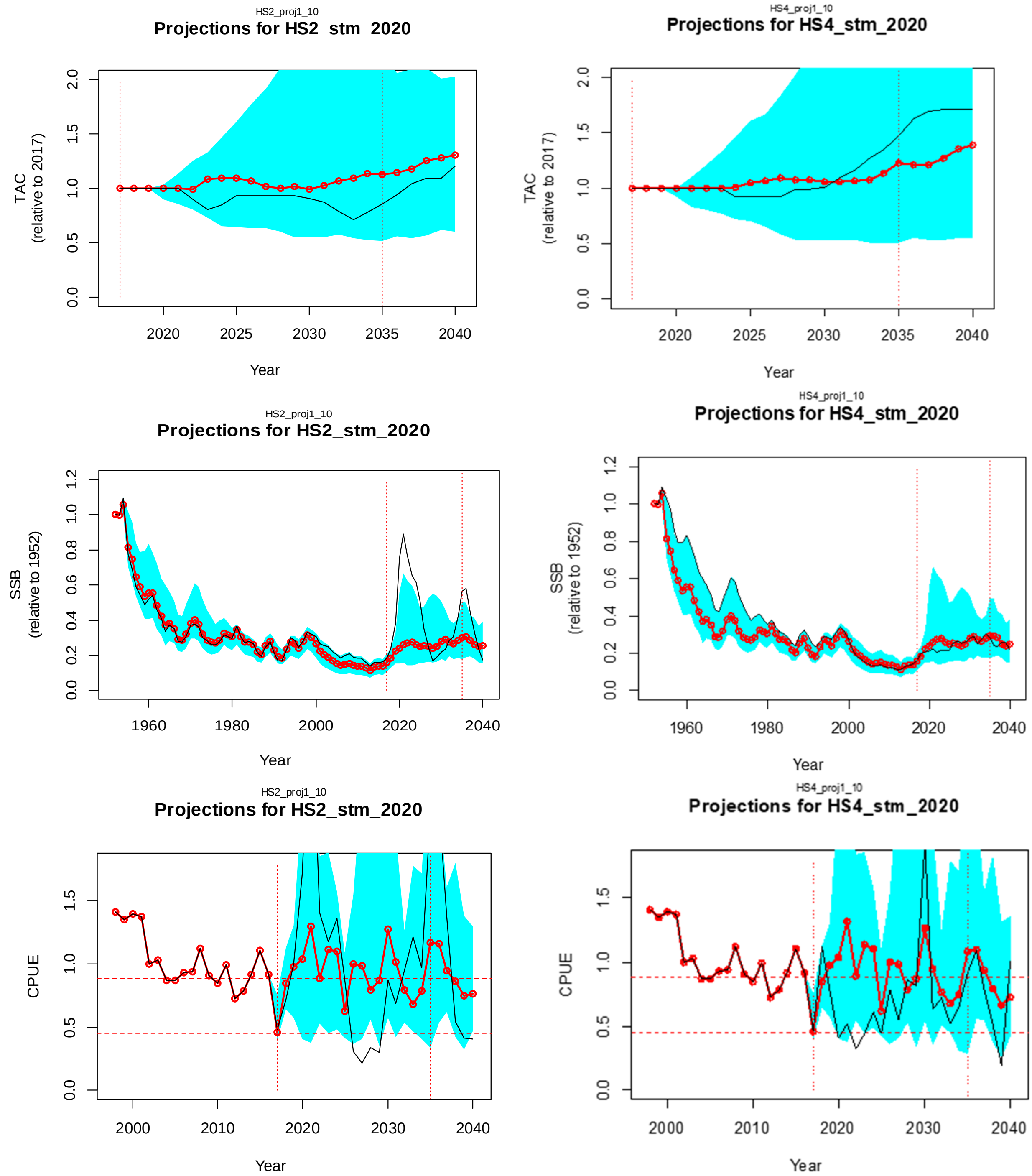


Figure 7 Examination of number of years for the average recent CPUE. Red circles indicate median projections, shaded blue region represents the 10th-90th percentiles of the distribution, the thin black line represents a random trajectory. Projections start in 2017 and the vertical line at 2035 indicates the tuning year.

5.7 Exploration of alternative HCR settings – HS5 and HS6

The HS1 rule has asymmetrical bounds (lower bound of 0.8 and upper bound of 1.25), around the target CPUE of 0.89. Two alternative settings were requested by TTRAG29 (Sept 2020) for 10% and 20% symmetrical bounds around the CPUE target. HS5 10% bounds are 0.8 and 0.97, and HS6 20% bounds are 0.71 and 1.07. In both these cases the upper bound is lower than the setting in HS1, which means that TAC increases will start to occur at lower levels of CPUE.

For HS5 (10% bounds) median 2035 TAC estimates are higher than HS1 and very slightly lower than HS6 (Figure 8). Median 2035 SSB depletion estimates are very similar for HS1, HS5 and HS6 (slightly lower for HS6). Median CPUE at 2035 is lower for HS5 and HS6 than HS1, and there is a decreasing trend in CPUE over the projections period for HS5 and HS6.

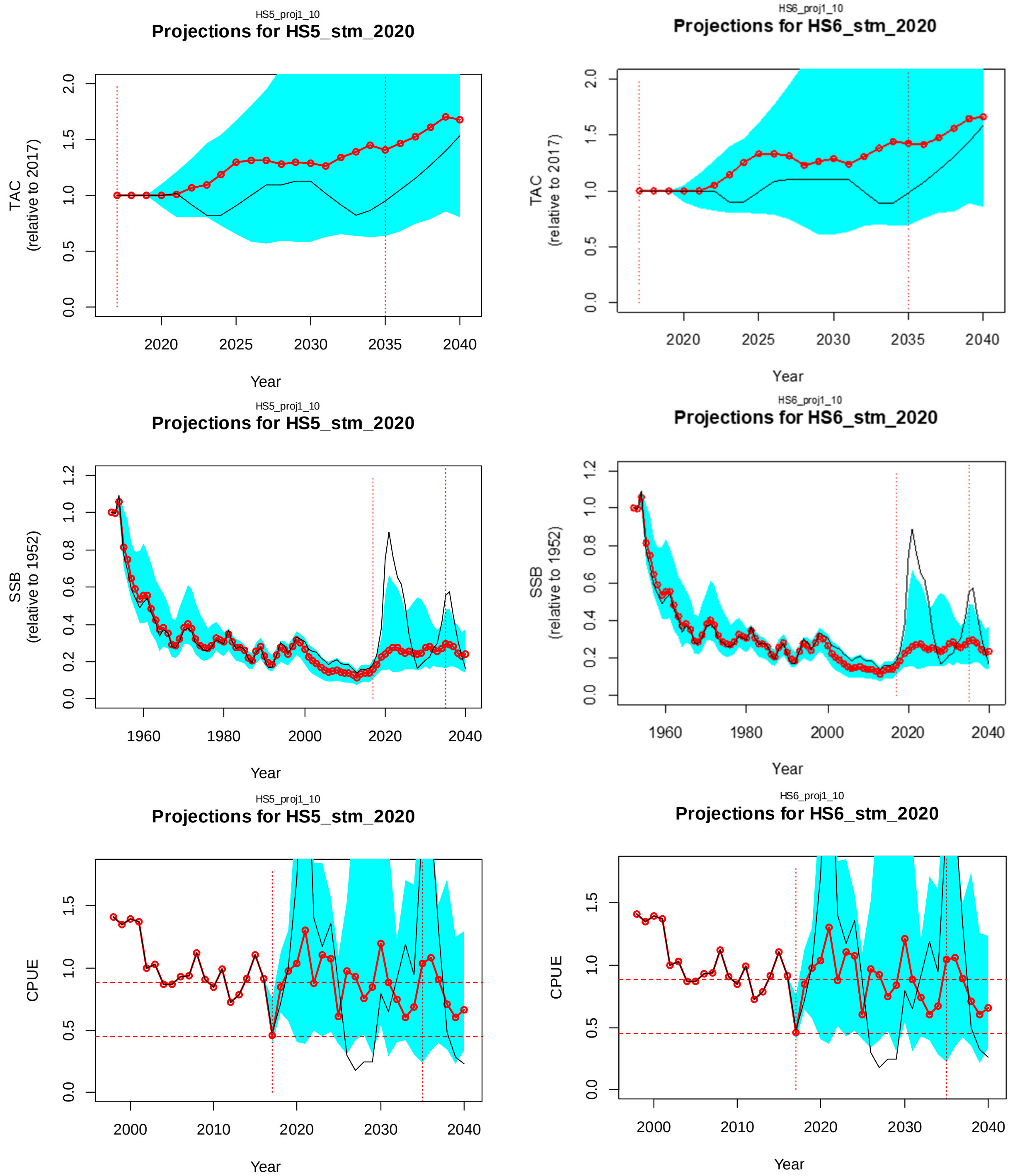


Figure 8 Alternative HCRs HS5 and HS6, with differing bounds (10% and 20% respectively). Red circles indicate median projections, shaded blue region represents the 10th-90th percentiles of the distribution, the thin black line represents a random trajectory. The projections start in 2017 and the vertical line at 2035 indicates the tuning year.

5.8 Examining effect of non-ETBF effort – HS1 Proj 4

In the Proj4 scenario effort in the non-ETBF area gradually increases to twice the current levels over a period of 5 years. In the previous tests (Proj1 settings), the non-ETBF effort is assumed to remain at current (2017) levels. This scenario investigates the impact on the ETBF fishery and TAC setting if there is an increase in effort in non-ETBF fleets.

The increased fishing effort in the non-ETBF area has an impact on the median of SSB depletion estimates in 2035, which are lower than the median level in all other scenarios (which have fixed effort in the non-ETBF area), but are higher than recent low levels (Figure 9). Median TAC in 2035 is the lowest of all scenarios. The CPUE estimates are lower, but remain above the target level in 2035, fluctuating around the target level. The lower 10th percent interval indicates a trend in lower estimates of CPUE in the later years of the projections, and a decreasing trend.

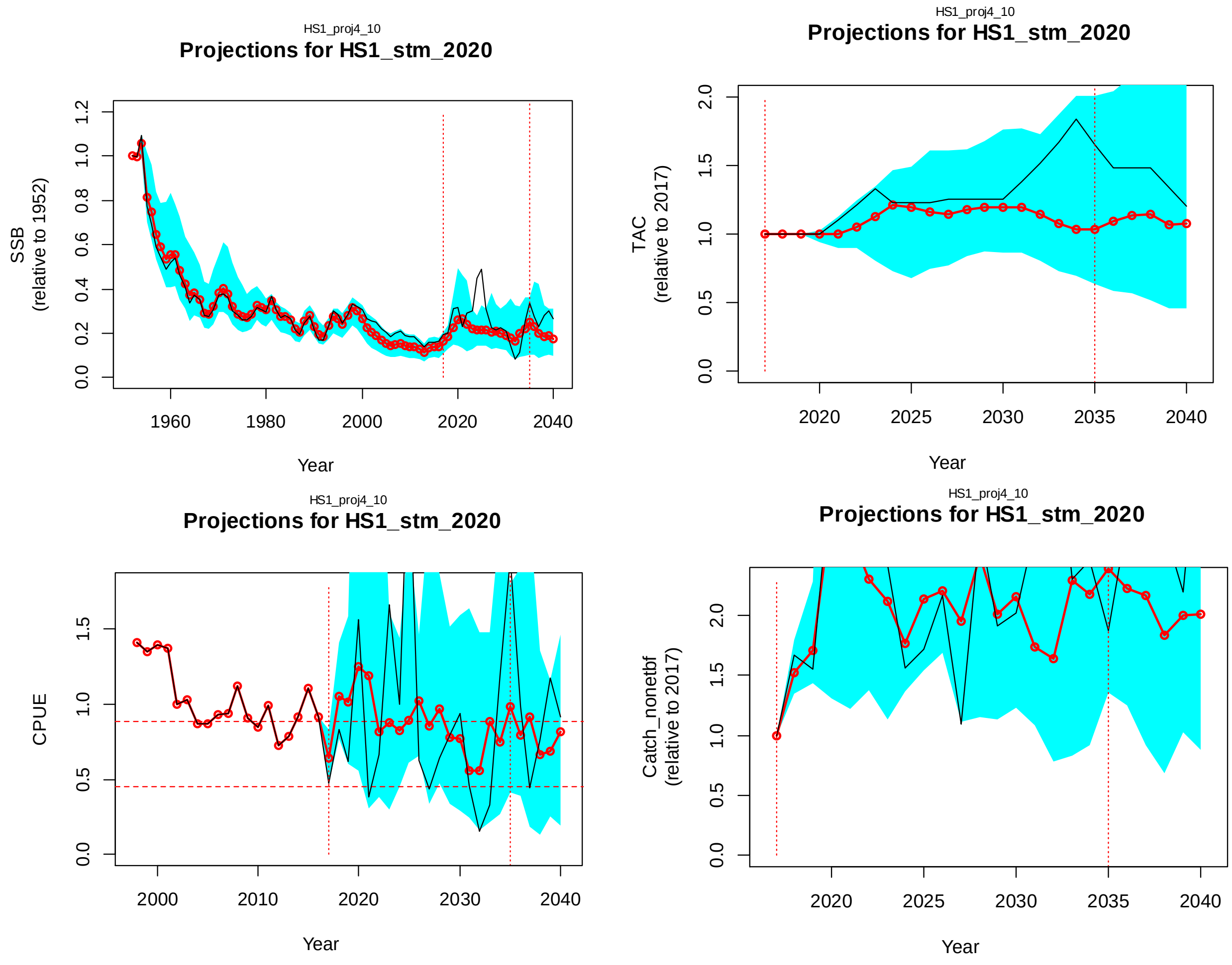


Figure 9 Examination of effects of increased effort in the non-ETBF area. Red circles indicate median projections, shaded blue region represents the 10th-90th percentiles of the distribution, the thin black line represents a random trajectory. The vertical dashed line at 2017 indicates the start of projections and the vertical line at 2035 indicates the tuning year.

5.9 Summary Statistics

Summary statistics for each of the HS and projection scenarios described above are summarised in Table 1 (median and 80% percentile range).

Table 1 Summary statistics, median and 80th percentile range, for the HS and projection scenario combinations described above.

HS	Proj	Change from HS1, Proj1	SSB2035/SSB0	CPUE 2035	TAC 2035
HS1	proj1		0.30 (0.17, 0.50)	1.27 (0.36,3.16)	392 (200, 832)
HS1	proj2	Migration 1%	0.33 (0.16, 0.45)	1.29 (0.52, 2.50)	527 (293, 979)
HS3	proj1	3yr TAC, 27% max TAC change	0.31 (0.18, 0.51)	1.32 (0.44, 3.23)	364 (188, 667)
HS1	proj3	Higher recruitment autocorrelation	0.27 (0.11, 0.45)	1.03 (0.41, 2.15)	456 (212, 976)
HS2	proj1	Recent CPUE ave 2 yrs	0.30 (0.18, 0.50)	1.32 (0.38, 3.24)	397 (183, 777)
HS4	proj1	Recent CPUE ave 4 yrs	0.30 (0.17, 0.50)	1.22 (0.32, 3.15)	431 (181, 992)
HS5	proj1	10% bufferzone 0.8, 0.97,	0.30 (0.16, 0.48)	1.17 (0.27, 3.08)	495 (223, 982)
HS6	proj1	20% bufferzone 0.71, 1.07	0.29 (0.16, 0.49)	1.17 (0.26, 3.07)	501 (243, 982)
HS1	proj4	non-etbf effort X2 over 5 yrs	0.25 (0.11, 0.36)	1.11 (0.47,2.02)	363 (224, 704)

6 Conclusion

Candidate HSs have been evaluated using a reference set of operating models with additional uncertainty in future projections, and alternative future projection scenarios. The alternative projection scenarios check for robustness to a wider set of uncertainties than covered in the reference set.

The HSs considered demonstrate feedback in that they respond to adjust TAC when CPUE, and related biomass, increase or decline. This is true even in the scenario where non-ETBF effort doubles, and in the scenario when migration is reduced to 1% which mimics a localised Australian stock.

As expected, there are trade-offs in performance between higher potential TACs and higher potential CPUE. SSB results are relatively stable. To assist with selecting a HS for setting TAC in future, the trends in the figures above and the median and lower percentile of estimates, summarised in Table 1, should be considered.

Acknowledgments

Nicholas Ducharme-Barth, Graham Pilling and John Hampton, at SPC, for access to the striped marlin stock assessment outputs.

Robert Campbell, CSIRO, for provision of CPUE data from the ETBF fishery.

This work is funded by AFMA and CSIRO.

References

Ducharme-Barth N., G. Pilling, J. Hampton. 2019. Stock assessment of SW Pacific striped marlin in the WCPO. Scientific Committee fifteenth regular session, Pohnpei, Federated States of Micronesia, 12-20 August 2019. WCPFC-SC15-2019/SA-WP-07.

Ducharme-Barth, N. and Pilling, G. (2019). Background analyses for the 2019 stock assessment of SW Pacific striped marlin. Technical Report, Pohnpei, Federated States of Micronesia. WCPFC-SC15-2019/SA-IP-07.

Hillary R. 2019. Update on MSE work and questions to TTRAG on objectives and operating features of revised harvest strategy. Working paper to TTRAG members.

Preece and Hillary (2020). The Striped Marlin Operating Model. TTRAG29, Sept 2020.

As Australia's national science agency and innovation catalyst, CSIRO is solving the greatest challenges through innovative science and technology.

CSIRO. Unlocking a better future for everyone.

Contact us

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

For further information

Oceans and Atmosphere
Ann Preece
+61 3 6232 5336
ann.preece@csiro.au
csiro.au/O&A

Appendix 7

Oceans & Atmosphere
www.csiro.au

CSIRO

Initial estimates of striped marlin WCPO movement dynamics

R. Hillary & T. Patterson

12 July 2019

CSIRO Oceans & Atmosphere
Battery Point, Hobart 7004, Tasmania, Australia.

Copyright and disclaimer

Important disclaimer

Contents

1 Background

This document details an initial analysis of both the conventional tagging data and the most recent set of satellite tracks for striped marlin (*Kajikia audax*) to estimate large scale movement dynamics for use in the upcoming revision to both the ETBF billfish harvest strategies [1] and the last stock assessment [2] for this species in 2012. Although the previous stock assessment [2] was not explicitly spatial in nature, and the reference case of the revised assessment is proposed to be non-spatial, SPC scientists intend to explore a spatial sensitivity model within the suite of options. Given that we intend to explore spatial operating models for the recently initiated MSE work, and intend to use the SPC assessment models as the basis for those OMs, we have undertaken these analyses to both assist in the upcoming assessment and assist ourselves in having a stock assessment model that can be morphed into our existing OM structure.

The method of analysis is novel in that for the conventional tagging data it uses *only* recaptures, not both releases and recaptures as traditionally done, and integrates these data with spatiotemporally aggregated satellite data to estimate movement rates. The reason for using recaptures only is that we can then remove complex processes such as natural and fishing mortality, tag shedding, tag mortality and absolute reporting rates and focus on migration only. The only additional information required is an estimate of the *relative* probability of recapture and reporting across the spatial regions at the time of recapture.

2 Data

We have access to two sources of data that we can use to try and estimate migration dynamics:

1. Conventional tagging data, from both Australian and New Zealand releases so spanning the range assumed in the assessment and MSE work
2. Satellite tagging data, again from both Australian and New Zealand releases

2.1 Conventional tagging data

For these analyses as already outlined we use *only* recaptures, not releases and recaptures. We also only use recaptures where the time-at-liberty was at least 3 months, given the quarterly time-step of the movement model we employ later on. The model itself has only two spatial regions: region 1 (West of 165E) and region 2 (East of 165E). The recapture data spans a wide range of years (1989 to 2018) and quarters, though within-year patterns in fishing pressure mean that the recaptures-by-quarter are not equitably spread across the years. Table 2.1 details the high level summary of these data. What is abundantly clear is that **none** of the tags released in area 1 were recaptured in area 2, yet just short of 40% of the tags released in area 2 were recaptured in area 1. We know of no existing information that would suggest that tags released in area 1 are less reported than releases from area 2. Also, both data sets have similar ranges in terms of time-at-liberty (with a maximum of around 4 years for both). The strong qualitative implications of these data are that there is *very* little apparent movement from region 1 to region 2, but clear evidence of movement from region 2 to region 1.

Release Area	Region 1	Region 2
Recapture Area		
Region 1	61	10
Region 2	0	16

Table 2.1: *Summary of release and recapture location for both Australian (Region 1) and New Zealand (Region 2) releases for tags with a time-at-libert of at least 3 months.*

2.2 Satellite tagging data

There are currently 73 satellite tagged fish released across the South West Pacific Ocean. Of these, 60 were tagged around New Zealand and 13 were tagged off Eastern Australia. The times-at-liberty/transmission for these fish are, by the nature of the tags, more limited than the conventional tagging data: they range from 13 to 226 days, so very few with more than half a year of data.

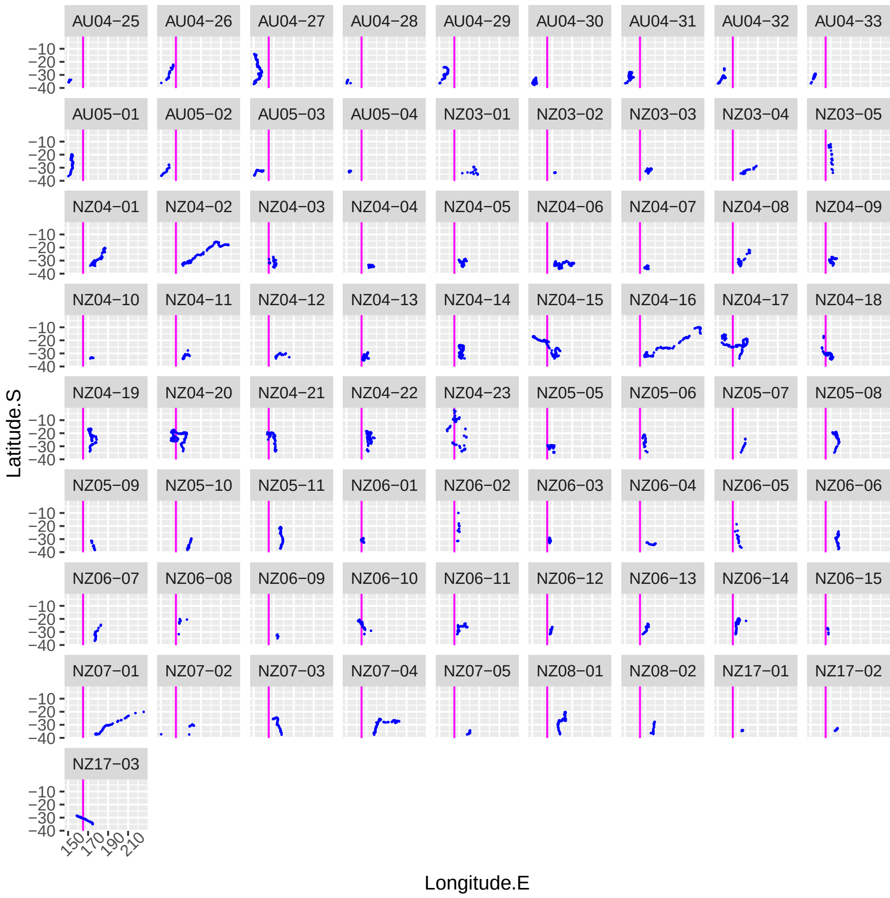


Figure 2.1: *Individual tracks for each of the 73 fish tagged with PSATs. The full vertical magenta line denotes the 165E longitude i.e. the line that demarks region 1 (to the left) from region 2 (to the right).*

Figure 2.1 shows the individual tracks of each of the 73 PSAT-tagged animals. Looking at Figure 2.1 it is reasonably clear that none of the Australian tagged fish (including those two fish with more than 90 days of tracking) appear move across the 165E boundary. Of the 20 NZ releases with more than 90 days of tracking 7 of those fish appear to cross the 165E boundary - 35% which is actually close to the 38% of the conventional tags released in area 2 and recaptured in area 1. In summary, the qualitative features of the PSA data appear to be consistent with those of the conventional tagging data. In the integrated movement models we use these 22 PSAT tracks with at least 90 days of data and where their quartlery location is defined by the region in which they spent the majority of their time. In the near future, we intend to use a state-space geolocation model to get appropriate corrected location estimates (and error estimates thereof) to use in the analyses but for this initial attempt we use the raw tracks to define their quarterly location.

3 Methods

The estimation method integrates the conventional tagging data and the spatiotemporally aggregated PSAT data to estimate movement probabilities between the two spatial regions. The fundamental movement model is assumed to have a quarterly time-step, with a time-invariant spatial transition matrix, $\mathbf{\Phi}$ and the rows of $\mathbf{\Phi}$ sum to 1.

From the year of release up until the year of recapture, only movement is a factor, so for a recaptured fish released in time period t_0, released in region r_0, then the spatial distribution can be represented in vector form: $\mathbf{u}_0 = \delta_{r,r_0}$ (where $\delta()$ is the Kronecker delta function). For $t \in [t_0 + 1, t_1]$ we have that

$$\mathbf{u}_t = \mathbf{u}_{t-1}\Phi$$

Using this we have the spatial probability distribution at the time of recapture, $\mathbf{u}_{t_1}$.

3.1 Conventional tagging recapture model

When using the conventional tagging data we do, however, need some additional information to realistically reflect the relative spatial recapture probability. Specifically we need to account for the following two things:

1. In the year of recapture, a region with a relatively higher/lower exploitation rate will have a commensurately higher/lower chance of a tag being caught there, conditional on their probability of being in that region at that time
2. In the year of recapture a fleet with higher/lower chances of reporting a recaptured tag will increase/decrease the probability of being reported in that area, conditional on their probability of being in that region *and* relative chance of recapture at that time

All we have to do to work out the probability of recapture in region r_1 for the recaptured fish as follows:

$$\mathbb{P}\left(r = r_1 \,|\, t_0, t_1, r_0\right) = \frac{u_{t_1,r_1}\eta_{t_1,r_1}}{\sum_j u_{t_1,j}\eta_{t_1,j}},$$
$$\eta_{t,r} = \frac{\xi_{t,r}\rho_{t,r}}{\sum_j \xi_{t,j}\rho_{t,r}},$$

which, in words, is the probability of being in the recapture region, r_1, in the recapture period, t_1, weighted by the relative probability of recapture and reporting across all regions in the recapture period, t_1. For a given set of spatial recaptures in time period t_1, $\mathbf{x}_{t_1}$, the reference likelihood would be multinomial:

$$\Lambda^c\left(\mathbf{x}_{t_1} \mid \cdots\right) \propto \prod_j \mathbb{P}\left(r = j \mid t_0, t_1, r_0\right)^{x_{t_1,j}}.$$

3.2 Satellite tagging observation model

So $\mathbf{u}_t$ has the same meaning as for the conventional tagging model. For each of the n_e electronically tagged animals there is an aggregated discrete location history $r_{t,i}$, where i denotes an individual tagged animal. The likelihood at any given time period for which we have an observation of location is categorical, with location $r_{t,i}$ the observation and probability vector $\mathbf{u}_t$. If each animal as T_i observations then the total likelihood across all electronically tagged animals would be

$$\Lambda^e = \prod_{i=1}^{n_e} \prod_{j=1}^{T_i} u_{t_j, r_{t_j,i}}$$

The electronic and conventional tagging data are independent, so we just add up their respective log-likelihoods and use this for inference about $\boldsymbol{\Phi}$:

$$\ln \Lambda^{\text{tot}} = \ln \Lambda^c + \ln \Lambda^e.$$

The unfortunate reality of the PSAT data is that their position at a given time and a given time period will be probabilistic: we will not have a clean observation $r_{t,i}$ but more a discrete probability observation, $\pi_{t,i,r}$, for whic region r the tag is most likely in for that time perid t. The good news is that these basically 'fall out' of the geolocation HMMs we intend to use. This would require the following modification to the eletronic tagging data likelihood:

$$\Lambda^e = \prod_{i=1}^{n_e} \prod_{j=1}^{T_i} \left(\sum_{k=1}^{R} u_{t_j,k} \pi_{t,i,k} \right)$$

4 Results

We have already summarised the two data sets (and how they are subsetted and aggregated for use in the models) but the final step for using the conventional tagging is to define the spatiotemporal relative recapture and reporting probability: $\eta_{t,r}$. Figure 4.1 shows the proposed spatial regions for the upcoming spatial sensitivity model to be used in the upcoming striped marling assessment. Both the conventional recapture data and the PSAT tracks suggest there are very few animals either going to sub-region 2 - all but one of the NZ tags used in the analyses were recaptured in sub-regions 3, 4, 5, 6 and 7.

Ideally, we would somehow have both reporting rate estimates and exploitation rates for each of the fisheries in those areas. We do not have that information, but we can explore some plausible scenarios using relative catches by quarter and region (not sub region but spatial region) and which fleets we think are reporting recaptured tags. For spatial area 1 this corresponds to sub-regions 3 and 5 in Figure 4.1; for spatial area 2 this corresponds to sub-regions 4, 6, and 7 in Figure 1. We explored the following three options for relative reporting rates:

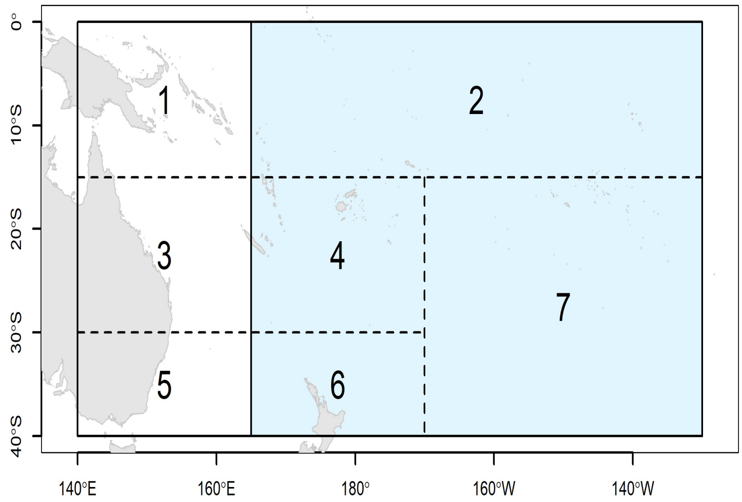


Figure 4.1: *Proposed spatial regions (and sub-regions therein) for the upcoming striped marlin assessment. The 165E line separates the two areas between which there is migration permitted, whereas the sub-regions are used to place particular fisheries in though there is no movement between sub-regions in a given region.*

1. **Option 1**: *only* the Australian and New Zealand recreational and long-line fleets report their recaptured tags
2. **Option 2**: the recreational/long-line Australian and New Zealand fleets and other long-line fleets report recaptured tags but not the Japanese long-line fleet
3. **Option 3**: all the fleets in those areas report their recaptured tags at the same *relative* level

Right off the bat Option 1 is both extreme and, given the locations where some of the tags have been reported, the most *a priori* unlikely of the three. Options 2 and 3 have more relative plausibility but might over-state the reporting rates of the long-line fleets, relative to the recreational fleets. For each of these three options the value of $\eta_{t,r}$ for each time-period (year-quarter combination) is then simply the relative distribution in total catch for the fleets assumed to be reporting recaptured tags.

For the models we use Template Model Builder (TMB) [3] and the associated MCMC add-on package `tmbstan` [4] to obtain a sample from the posterior probability distribution of the spatial transition matrix, $\mathbf{\Phi}$. We did this for each of the three options for the relative spatial recapture and reporting probabilities. Table 4.1 details the posterior summaries for the movement probabilities between regions 1 and 2, and Figure 4.2 detials the posterior predictive distribution of the number of recaptures in each region, given their region of release and for each of the three relative recapture and reporting scenarios outlined above.

From Table 4.1 it is clear that, whatever the relative recapture and reporting scenario we choose, there is very little evidence for any kind of significant movement from region 1 to region 2. Of 61 conventional tags, none were recaptured in region 2 even when at liberty for a number of years. None of the PSAT tracks from animals tagged in region 1 appeared to move East of 165E either. For movement from region 2 to region 1 there is clear evidence this occurs, with median estimates ranging from 0.07 to 0.14 depending on the scenario chosen. Figure 4.2 summarises

	$\Phi_{1,2}$	$\Phi_{2,1}$
Option 1	0.001 (0–0.005)	0.07 (0.04–0.13)
Option 2	0.001 (0–0.005)	0.07 (0.04–0.11)
Option 3	0.001 (0–0.004)	0.14 (0.09–0.21)

Table 4.1: *Posterior summary in terms of median (and 95% credible interval) for the quarterly movement probabilities between regions 1 and 2.*

the posterior predictive distribution of the number of recaptures by recapture location, given the region of release and scenario option chosen. What is clear is that relative recapture and reporting rate options 1 and 2 (arguably the less *a priori* plausible options) do not seem consistent with the observed data given the assumed movement model - at least for the relative number of recaptures in regions 1 and 2 given their release in region 2. Options 1 and 2 appear to result in consistent over/under estimation of recaptures in regions 1 and 2, respectively.

The reason for this is because, when we assume that neither of the non-Australian and New Zealand long-line fleets (Option 1) or the Japanese long-line fleet (Option 2) don't report tags, it becomes increasingly more probable for quarters 2, 3 and 4 that a tag would be recaptured and reported in region 1, thereby suggesting the movement probability from region 2 to region 1 would not need to be very large for a tag to be released there and recaptured in region 1. This, however, does not tally with the PSAT data that suggests movements from region 2 to region 1 are more common than the conventional tagging would suggest, *given* we assume in Options 1 and 2 that these tags are overwhelmingly more likely to be caught and recaptured in region 1 for quarters 2, 3 and 4. As a result, the model 'splits the difference' between what the conventional tag data are saying for Options 2 and 3 and what the PSAT data are saying (independently of the reporting options). For Option 3, where this preference for tags to be recaptured and reported in region 1 relative to region 2 is far far weaker, the model is able to fit the conventional data well (see Figure 4.2) *and* fit to the PSAT data also given their consistent information on movement from region 2 to region 1 for this scenario. Given the PSAT data's independence from the conventional tagging data, this appears like *a posteriori* evidence that Option 3 is more likely than Options 1 and 2.

5 Discussion

This paper details an initial attempt at an integrated movement model using both the conventional and satellite tagging data for Striped Marlin in the WCPO. The model for the conventional tagging data uses recaptures *only* - not releases and recaptures - focussing singularly on movement dynamics, needing only estimates of relative recapture and reporting probability by spatial region to work. The model for the satellite data initially uses spatiotemporally aggregated tracks as observations of where that animal as been at the model's spatiotemporal scales.

What is abundantly clear, in a qualitative sense, from both data sets is that animals do not appear to move very much from region 1 (the ETBF region) to region 2 (NZ and the South Pacific). For releases in region 2 and for times-at-liberty of at least 90 days around 35% of the tags either moved to, or were recaptured, in region 1, suggesting there is clear movement from region 2 to region 1. We explored three scenarios for the relative recapture and reporting scenarios: (1) where only the recreational fleet returned tags, (2) where the recreational and other long-line fleets returned tags, and (3) where the recreational, other and Japanese long-line fleets returned

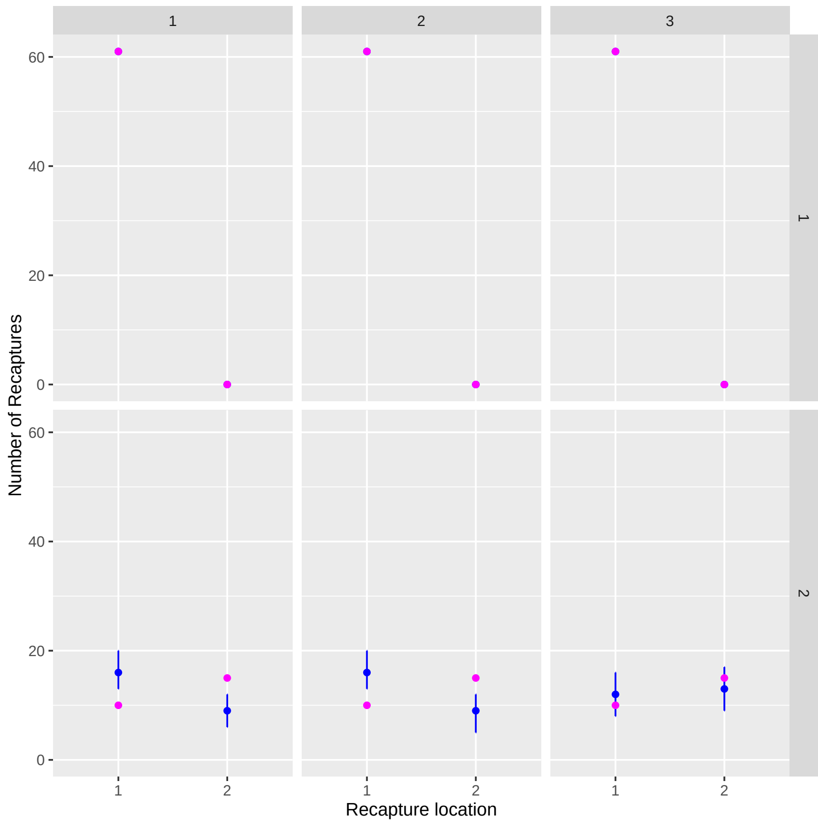


Figure 4.2: *Posterior predictive summary (median and 95% credible interval) for the number of recaptures in a given location, conditional on the region of release (vertical panels) and the three relative recapture and reporting scenarios (the horizontal panels).*

tags. Given the recapture location and quarter it was assumed that the *a prior* ranking of these scenarios from least to most probable was 3, 2, 1.

Estimates of the movement probabilities from region 1 to region 2 were, irrespective of the reporting scenario, less than 1% per quarter. This should not be a surprise given **not one** conventional tag was recaptured in region 2 given release in region 1 and none of the PSAT tracks appeared to move into region 2 either. For the movement probabilities from region 2 to region 1, these depended on the relative reporting scenario. For options 1 and 2, which heavily favoured the recapture and reporting of tags in region 1 in quarters 2, 3 and 4, median quarterly movement probabilites of 0.07 were obtained. For Option 3, which assumes reporting and recapture in both regions is far more equiprobable, median quartlerly movement probabilities of 0.14 were obtained. In terms of uncertainty the estimates were fairly accurate with a 95% credible range of 0.04–0.13 for Options 1 and 2, and 0.09–0.21 for Option 3.

In terms of *a posteriori* ranking of the relative recapture and reporting scenarios, Options 1 and 2 were not compatible with the PSAT data resulting in an over/under estimation pattern in the recaptures in regions 1 and 2 given release in region 2. Option 3 fitted both the conventional and PSAT data well and resulted in movement estimates more compatible with both sets of data. At least at this stage, some kind of scenario more like Option 3 (where tags are being reported in vaguely similar patterns across fleets) seems more probable than one were fleets that comprise a smaller subset of the catch, especially outside of quarter 1, were assumed to be the only ones recapturing and reporting tags in meaningful numbers.

There is still a reasonable amount of work that needs to be done to finalise this analyis: finessing the relative recapture and reporting scenarios and including the spatial uncertainty in the PSAT tracks. That being said, this does perhaps give us enough information to explore more nuanced migrations scenarios - particularly around the apparently direction nature of it - in both the stock assessment revision and the MSE work.

6 Acknowledgements

We thank Dr. Julian Pepperell, NSW DPI and Phil Bolton and Clay Hibert in particular for organising access to the Australian conventional tagging data. We thank John Holdsworth for providing us with both the New Zealand tagging data and the PSAT data. We also thank the Marine Conservation Science Institute and New Zealand Marine Research Foundation for the use of the PSAT data.

References

[1] R. Hillary (2018) Updating the ETBF Broadbill Swordfish MSE analyses . *TTRAG, Canberra, ACT*.

[2] N. Davies, S. Hoyle and J. Hampton (2012) Stock assessment of striped marlin (*Kajikia audax*) in the South West Pacific Ocean. *WCPFC-SC8-2012/SA-WP-05*.

[3] Kristensen, K. *et al.* (2016) TMB: Automatic Differentiation and Laplace Approximation. *J. Stat. Soft.* **70**(5): 1–21.

[4] Monnahan, C. C. and Kristensen, K. (2018) No-U-turn sampling for fast Bayesian inference in ADMB and TMB: Introducing the adnuts and tmbstan R packages. *PLoS ONE*. **13**(5): e0197964.

CONTACT US

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

WE DO THE EXTRAORDINARY EVERY DAY

We innovate for tomorrow and help improve today for our customers, all Australians and the world.

Our innovations contribute billions of dollars to the Australian economy every year. As the largest patent holder in the nation, our vast wealth of intellectual property has led to more than 150 spin-off companies.

With more than 5,000 experts and a burning desire to get things done, we are Australias catalyst for innovation.

WE IMAGINE. WE COLLABORATE.
WE INNOVATE.

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

WE DO THE EXTRAORDINARY EVERY DAY

We innovate for tomorrow and help improve today for our customers, all Australians and the world.

Our innovations contribute billions of dollars to the Australian economy every year. As the largest patent holder in the nation, our vast wealth of intellectual property has led to more than 150 spin-off companies.

With more than 5,000 experts and a burning desire to get things done, we are Australias catalyst for innovation.

WE IMAGINE. WE COLLABORATE. WE INNOVATE.