Collect. Vol. Sci. Pap. ICCAT, 81(7), SCRS/2024/137: 1-19 (2024)

RESULTS OF THE NORTH ATLANTIC SWORDFISH MANAGEMENT STRATEGY EVALUATION

A.R. Hordyk[1], C. Brown[2], R. Coelho[3,4], N.M.T. Duprey[5], K. Gillespie[6], A. Hanke[6], S. Miller[7], L. Rueda[8], D. Rosa[3,4], M. Schirripa[2]

SUMMARY

The SCRS has been working on a management strategy evaluation (MSE) for North Atlantic swordfish for the past several years. A set of nine reference operating models were developed based on the key uncertainties in the understanding of the system. A further seven robustness operating models were developed to evaluate the impact of additional uncertainties. A large variety of candidate management procedures (CMPs) were developed and evaluated across these operating models using a set of performance indicators. The final CMPs are described in ***Appendix A****. The results of the MSE are presented in an interactive online application. This paper provides an overview of the methodology used in the MSE, describes examples of the results provided in the online application, and summarizes the key results of this analysis.*

RÉSUMÉ

*Le SCRS travaille depuis plusieurs années sur une évaluation de la stratégie de gestion (MSE) pour l'espadon de l'Atlantique Nord. Un ensemble de neuf modèles opérationnels de référence a été élaboré sur la base des principales incertitudes liées à la compréhension du système. Sept autres modèles opérationnels de robustesse ont été développés pour évaluer l'impact d'incertitudes supplémentaires. Un grand nombre de procédures de gestion potentielles (CMP) ont été mises au point et évaluées par l'ensemble de ces modèles opérationnels à l'aide d'une série d'indicateurs de performance. Les CMP finales sont décrites à l'****appendice A****. Les résultats de la MSE sont présentés dans une application interactive en ligne. Ce document donne un aperçu de la méthodologie utilisée dans la MSE, décrit des exemples de résultats fournis dans l'application en ligne et résume les principaux résultats de cette analyse.*

RESUMEN

El SCRS lleva varios años trabajando en una evaluación de la estrategia de ordenación (MSE) del pez espada del Atlántico norte. Se elaboró un conjunto de nueve modelos operativos de referencia basados en las principales incertidumbres de la comprensión del sistema. Se desarrollaron otros siete modelos operativos de robustez para evaluar el impacto de incertidumbres adicionales. Se desarrollaron una gran variedad de procedimientos de ordenación candidatos (CMP) y se evaluaron a través de estos modelos operativos utilizando un conjunto de indicadores de desempeño. Los CMP definitivos se describen en el ***Apéndice A****. Los resultados de la MSE se presentan en una aplicación interactiva en línea. Este documento ofrece una visión general de la metodología utilizada en la MSE, describe ejemplos de los resultados proporcionados en la aplicación en línea y resume los resultados clave de este análisis.*

KEYWORDS

MSE; simulation; performance indicators

[1] Blue Matter Science, 2150 Bridgman Avenue, North Vancouver, BC, Canada, V7P2T9 adrian@bluematterscience.com
[2] NOAA Fisheries, Southeast Fisheries Center, 75 Virginia Beach Drive, Miami, FL, 33149-1099, USA
[3] Instituto Portugues do Mar e da Atmosfera (IPMA), 8700-305 Olhão, Portugal
[4] Centro de Ciências do Mar do Algarve (CCMAR), Univ. Algarve, 8005-139 Faro, Portugal
[5] Fisheries & Oceans Canada, Fish Population Science, 200-401 Burrard St, Vancouver, BC, Canada
[6] Fisheries & Oceans Canada, Biological Station, 531 Brandy Cove Road, St. Andrews, NB E5B 2L9, Canada
[7] The Ocean Foundation, 1320 19th St., NW, 5th Floor, Washington, DC 20036, United States
[8] Centro Oceanográfico de Málaga (IEO-CSIC). Explanada San Andrés s/n. Puerto de Málaga 29001 Málaga, Spain

1 Introduction

The Commission was scheduled to adopt a management procedure (MP) for the North Atlantic Swordfish (N-SWO) in 2023, but delayed one year to allow the SCRS to complete requested work on the management strategy evaluation (MSE) and present updated results at the 2024 annual commission meeting in Cyprus (Rec. 23-04).

This paper describes the methodology used to conduct the MSE, including details of the candidate management procedures (CMPs) and performance indicators (PIs), presents a summary of the key results of the performance and trade-offs for the CMPs.

2 Methods

The technical specifications of the N-SWO MSE process, including details on the conditioning of the operating models, and the assumptions for the projection period, and the definition of the performance indicators, are described in the Trial Specifications Document (TSD), available online at https://iccat.github.io/nswo-mse/TS/Trial_Specs.html. In this section we provide an overview of the key components of the MSE process, and refer readers to the TSD further details.

2.1 Operating models

Operating models for the N-SWO MSE were based on the 2022 stock assessment (Anon., 2022), conducted with the Stock Synthesis 3 (SS3) assessment software (Methot & Wetzel, 2013). The operating models (OMs) were classified into two categories: the Reference Set, which spanned the key uncertainties in the 2022 stock assessment, and the Robustness OMs, a subset of the Reference Set that were modified to account for additional potential uncertainties.

The OMs were re-conditioned in July 2024 with the most up-to-date information available to the SCRS, which included catch data and CPUE indices through to 2022. Other changes to the methodology since 2023 include generating a new Combined Index using the updated data and a new methodology (Gillespie, this issue), increasing the number of simulations in each OM from 50 to 80, and changing the MP implementation year from 2024 to 2025. Additional Robustness OMs were also developed and are described below.

2.1.1 Reference operating models

The SCRS Swordfish Species Group (hereafter referred to as Group) identified the natural mortality rate (M) and the steepness of the Beverton-Holt stock-recruit relationship (h) as the primary axes of uncertainty that had the greatest impact of the estimated stock dynamics and the performance of candidate management procedures (Hordyk, 2021). Three values were selected for each parameter (M=0.1, 0.2, 0.3 and h=0.69, 0.80, 0.88), and nine operating models were conditioned with these assumed values. These nine OMs are referred to as the Reference OMs. One OM of the Reference Set (M=0.2 & h=0.88) shared had the same values for the biological parameters as the 2022 stock assessment.

The estimated magnitude of the stock varied considerably across the nine OMs, with over a six-fold difference between the smallest magnitude (SB0=66,124 t, M=0.3, h=0.88) and the largest (SB0=430,260 t, M=0.1, h=0.69; **Table 1**). The estimated stock status in the terminal year (2022) in terms of SB/SB_{MSY} ranged from 1.19 (M=0.2, h=0.69) to 2.27 (M=0.3, h=0.88; **Table 1Table 1**). The estimates of F/F_{MSY} in the terminal year ranged from 0.43 to 0.71 for these same models (**Table 1**).

Each individual simulation sharing identical dynamics during the historical period (based on the maximum likelihood estimates of the SS3 model), and stochastic recruitment deviations, conditioned on the recruitment deviations estimated for the historical period, and observation error on the index of abundance in the projection period.

2.1.2 Robustness operating models

A set of Robustness OMs were developed to evaluate the impact of additional uncertainties that were not considered in the Reference Set. The fifth OM from the Reference Set (M=0.2, h=0.8; **Table 1**), referred to as R0, was selected to be used as the base case for the development of robustness OMs. This model is the one that uses the middle assumed values in terms of M and *h*. Seven Robustness OMs were developed by modifying the assumptions of R0 to consider additional uncertainties for the historical and projection periods. **Table 2** provides a summary of the Robustness OMs. More details on the Robustness OMs are available in the TSD.

2.2 Performance indicators

The N-SWO MSE currently includes 11 key performance indicators as a benchmark for evaluation of the Commission's selected management objectives (**Table 3**). These performance indicators were developed based on input received from Panel 4 in March and June 2023. Further details on the Performance Indicators are available in the TSD.

The performance indicators are used to summarise the performance of the candidate management procedures. For the Reference Set, the results were combined across the nine operating models and then the performance indicators were calculated. For example, PGK_{short} was calculated by first combining the results from the 80 simulations from each OM in the Reference Set together, resulting in 720 simulations, and then calculating the proportion of data points from the first 10 years of the projection period where $SB>SB_{MSY}$ and $F<F_{MSY}$.

2.3 Candidate management procedures

The Group worked collaboratively to develop and test a number of CMPs. All CMPs calculate a single total allowable catch (TAC) for the North Atlantic swordfish, and use a 3-year management cycle with the first TAC applying to 2025.

The CMPs use the Combined Index and the reported catches as the primary data sources to determine the TAC for each management cycle. The 2025 TAC is based on catch and index data up to 2022. This 2-year data lag was used in future management cycles in the projection years; e.g., the next TAC will be set for the 2028 – 2030 fishing years using data up to 2025.

A brief description of the CMPs are provided in **Table 4** and a fuller description is provided in **Appendix A**. All the CMPs were tuned across the Reference Set OMs to three levels (0.51, 0.60, and 0.70), referred to as tuning targets a, b, and c respectively, for the PGK_{short}, PGK_{med}, and PGK_{long} performance indicators. The tunings were based on the lowest tuning value that achieved at least 60% PGK for all three 10-year time periods. In most cases this was PGK_{short}. The Commission previously chose not to consider CMPs tuned to 0.51, therefore the results are only shown for the 0.60 and 0.70 tunings. The two tuning variants resulted in 10 different CMP configurations (**Table 4**).

2.4 Presentation of results

An interactive application (App) was developed for examining the MSE results. The App is currently available online (https://shiny.bluematterscience.com/app/swomse). The App can also be run locally by installing the N-SWO MSE R package (https://github.com/ICCAT/nswo-mse) and running the command *Shiny()* after loading the package (*library(SWOMSE)*).

The results of the N-SWO process are summarized as the performance indicators values calculated across the Reference Set and the individual Robustness OMs. A series of plots also shows the performance of the CMPs over time during the projection period. The results for all CMPs developed in the MSE process are available in the NSWO-MSE R package.

Examples of the results presented in the App are provided in the Results section, and we refer readers to the app for more a full examination on the MSE results. The key results of this analysis are summarized in the Results section.

3 Results

3.1 Examples of results presented in the app

3.1.1 Time-series plot

Time-series plots show the trends in F/F_{MSY}, SB/SB_{MSY}, and the TAC over the 30-year projection period for each CMP configuration. These plots are useful for providing a graphical interpretation of the performance indicators that are used to summarize the performance of each CMP configuration.

Figure 1 shows a time-series plot for two tuning configurations of the MCC11 CMP for the Reference Set of operating models (MCC11_b and MCC11_c). The corresponding performance indicator values are shown in a table in the corner of each plot. The probability of being in the green quadrant of the Kobe matrix over the entire projection period can be calculated as the mean of PGK_short, PGK_med, and PGK_long.

This CMP avoided breaching the limit reference point in any of the simulations (nLRP =1), and maintained greater than 60% probability of remaining in the green region of the Kobe space throughout the projection period. On average, the TAC decreased from around 15,000 t in the short-term to an average of 11,000 t in the longer term (**Figure 1**). Time-series plots like this are available in the Shiny app for each CMP configuration, as well as for the results from the Robustness operating models.

3.1.2 Quilt plot

A quilt plot (or quilt table) provides quantitative values for the performance of CMPs using both probability values of achieving performance indicators as well as absolute values for change between management cycles and TAC within various timeframes, assuming the same set of conditions among all CMPs (**Figure 2**). Colour scale is used to provide a visual guide for performance with darker shades of blue indicating better performance. The Shiny app provides sorting and filtering tools where the user can set probability, TAC, and variance thresholds and then sort CMPs by their chosen performance indicator.

The quilt plots can be used to filter CMPs based on minimum performance criteria, or compare the performance of CMPs across the reference and robustness operating models. For example, across the nine OMs in the Reference Set the 10 configurations of the five CMPs did not breach the limit reference point (LRP) in any simulation **(Table 5**). However, the probability of breaching the LRP was more variable across these CMPs in the robustness tests. For example, the probability of breaching the LRP was highest for Robustness test R5, which assumed the first 15 years of the projection period had an environmentally driven period of lower-than-average recruitment (**Table 5**).

3.1.3 Kobe time plot

Kobe time plots show the percentage of simulations for each year of the projection period that are in each quadrant of the Kobe plot for each CMP in the Reference and robustness operating models. For example, **Figure 3** shows a Kobe plot for two configurations of the CE CMP. In this case, for the Reference operating models, there is greater than 50% probability of being in the green region of the Kobe matrix in most of the 30-year projection period, with a higher probability of being in the orange region for the period from 2027 – 2031 (**Figure 3**).

The results from the robustness operating models (R1 – R7) can be compared against the results from Reference models, and the baseline robustness OM (R0). This example shows that both tunings for the CE CMP have a considerably lower probability of remaining in the green region in several of the robustness tests, especially early in the projection period (**Figure 3**).

3.1.4 Trade-off plot

Trade-off plots are used to compare the results of CMPs with respect to two performance indicators in a scatterplot. **Figure 4** provides an example of four trade-off plots showing the trade-offs between the probability of being in the green space of the Kobe matrix (PGK) in the first 10-years of the projection period against the average TAC over this same period (top left), the PGK in years 11 – 20 against the average TAC over this same period (top right), the probability of not breaching the limit reference point against the average TAC in years 11 – 20 (bottom left), and the mean variation in TAC (shown as a negative value so lower values mean more variable) against the median TAC in the medium timeframe (bottom right).

In these plots, higher values (further to the right on x-axis or higher on the y-axis) indicate better performance outcomes. This example shows the results from the 10 configurations of the 5 selected CMPs for the Reference operating models. Results for the robustness operating models are displayed in trade-off plots in the Shiny application.

3.1.5 Violin plot

Violin plots show the density distribution of simulation outcomes for TAC change between management cycles for each CMP configuration under the conditions of the reference and robustness operating models (**Figure 5**). The width of the violin plot is proportional to the frequency of the absolute change in TAC (i.e., wider areas means value is more common). These plots indicate how reactive a CMP may be to new data and thus be driving change in TAC between management cycles relative to other CMPs given the same set of conditions. For example, a CMP may require a relatively large shift in CPUE data before it changes TAC, whereas another CMP may more closely follow the CPUE trend when generating new TAC advice.

3.2 Summary of key results

All of the CMPs did not breach the limit reference point (LRP) in any simulation or year for the Reference Set or for the base Robustness OM (R0), and had <5% probability of breaching the LRP in R2 and R7 (**Table 5**). For R6, all CMPs had <5% probability of breaching the LRP, except CE_b and CE_c which had 6% and 5% respectively (**Table 5**). The most challenging robustness tests with respect to the LRP were R3 and R5, where all CMPs had >15% probability of breaching the LRP (**Table 5**).

For the Reference Set, the four CMPs with the highest short-term yield were MCC11_b, MCC11_c, CE_b, and MCC9_b (**Figure 2**). The MCC methods also had the highest short-term yields for the base Robustness OM (R0) and had > 50% probability of being in the green zone of the Kobe space (PGK) throughout the projection period (**Figure 6**). In this scenario, the CE methods had higher PGK but markedly lower medium-term yields (**Figure 6**).

For R3, the robustness OM that evaluated the ability of the CMPs to rebuild on over-exploited stock, the three CMPs that had the lowest probability of falling below the LRP were MCC9_c, CE_c, and MCC11_c (**Figure 7**). Of these, CE_c had the highest short-term yield and probability of being in the green region of the Kobe space in the long-term, but had the lowest medium and long-term yields (**Figure 7**). The two MCC methods had a lower value for PGK but considerably higher yields compared to CE_c (**Figure 7**).

For R5, the robustness test that considered reduced recruitment for the first 15 years of the projection, the CMPs with the lowest probability of falling below the LRP were SPSSFox2_c, SPSSFox2_b, and SPSSFox_c (**Figure 8**). However, these methods, together with the CE methods, had the lowest medium- and long-term yields (**Figure 8**).

These results suggest that the MCC and CE methods appear to have the best performance with respect to yield and the ability to maintain the stock above the LRP. However, the robustness tests demonstrate that a trade-off exists between the magnitude and stability of the expected TAC and the probability of avoiding the LRP and achieving high probability for PGK. For example, the CE_b method had the highest probability for PGK_long in R3, but had considerably lower TAC in the medium- and long-term compared to the MCC9_b and MCC11_b methods (**Figure 9**).

4 Discussion

The candidate management procedures developed for the north Atlantic swordfish MSE use different sets of rules to convert the fishery data to a total allowable catch recommendation. Consequently, the performance of the candidate management procedures varies considerably across the different performance indicators, and across the different conditions of the reference and robustness operating models. A considerable challenge in the MSE process is the interpretation of the large amount of output from the analysis, and the identification of a candidate management procedure that is robust to uncertainty and most likely to achieve the management objectives under the range of plausible conditions in the future.

Managers can specify minimum performance criteria for some performance indicators, which allows CMPs that fail these requirements to be identified and removed from the list of options. For example, the managers of the swordfish fishery specified that management procedures must have at least an 85% probability of not breaching the limit reference point, and at least a 60% probability of being in the green space of the Kobe matrix throughout the projection period. These criteria were used in the development of the CMPs, and CMPs that fail these minimum requirements are not presented as options to the managers.

It is rare that a MSE process identifies a single CMP that clearly outperforms all other options. The ranking and selection of best performing CMPs can vary across different stakeholders and decision-makers depending on their specific values and objectives for the fishery. More likely, as is the case for swordfish, the CMPs present trade-offs among competing management objectives, such as a desire for high probability of not over-fishing the stock and a desire to maximize the economic output of the fishery. The results presented in the online App, and summarized in this paper, allow different groups of decision-makers to evaluate the performance of the CMPs under the conditions of the reference operating models, compare how well these CMPs perform under the more challenging conditions of the robustness tests, and identify the CMP that they consider to be the best candidate for managing the fishery.

Acknowledgements

We thank the ICCAT, the SCRS, and the Swordfish Species Group for their support on this work. This work was carried out under the provision of the ICCAT Science Envelope and the ICCAT – EU Grant Agreement – Strengthening the scientific basis for decision-making in ICCAT. The contents of this paper do not necessarily reflect the point of view of ICCAT or other funders and in no way anticipates ICCAT future policy in this area.

References

Anon. 2022. Report of the 2022 ICCAT Atlantic Swordfish Stock Assessment Session SCRS/2022/003. Collect. Vol. Sci. Pap. ICCAT, 79

Hordyk A. 2021. Evaluation of the Performance of some Candidate Management Procedures to Prioritize the Key Uncertainties in the North Atlantic Swordfish Operating Models. Collect. Vol. Sci. Pap. ICCAT, SCRS/2021/161

Methot R.D., & Wetzel C.R. 2013. Stock synthesis: A biological and statistical framework for fish stock assessment and fishery management. Fisheries Research, 142, 86–99. https://doi.org/10.1016/j.fishres.2012.10.012

Table 1. Summary of the estimated stock dynamics for the nine operating models (OMs) in the Reference Set. The nine OMs spanned uncertainty in the assumed natural mortality rate (*M*) and the steepness of the Beverton-Holt stock-recruit relationship (*h*). The estimated unfished equilibrium spawning biomass (SB0; ton), and the estimated fishing mortality rate (F) and the spawning biomass (SB) relative to their respective values at maximum sustainable yield (MSY) in the terminal year of the operating models (2022) are reported in the table.

OM #	*M*	*h*	*SB0*	F/F_{MSY}	SB/SB_{MSY}
1	0.1	0.69	430,260	0.71	1.30
2	0.1	0.80	370,240	0.71	1.29
3	0.1	0.88	335,753	0.69	1.32
4	0.2	0.69	154,718	0.74	1.19
5	0.2	0.80	133,280	0.68	1.28
6	0.2	0.88	120,145	0.62	1.45
7	0.3	0.69	82,676	0.59	1.48
8	0.3	0.80	71,069	0.53	1.74
9	0.3	0.88	66,124	0.43	2.27

Table 2. Description of the Robustness operating models (OMs) developed for the North Atlantic swordfish MSE.

Code	*Description*
R0	Reference OM for the Robustness tests. OM 5 from the Reference Set (**Table 1**)
R1	Evaluate impact of an assumed 1 percent annual increase catchability, that is not accounted for in the standardization of the indices of abundance (historical & projection)
R2	Same as R1, but bias in the indices of abundance is only for the historical period
R3	Robustness test to evaluate the ability of the CMPs to recover the stock from a low initial level. The historical indices were modified by adding a persistent slope such that the SB/SBMSY = 0.6 in the terminal year of the OM conditioning
R4	Evaluate impact of cyclical pattern in recruitment deviations in projection period; a proxy for impact of climate change on stock productivity. The recruitment deviations are lower than expected for the first 15 years of the projection period, and then higher than expected in the following 15 years
R5	Evaluate impact of lower than expected recruitment deviations for first 15 years of projection period; a proxy for impact of climate change on stock productivity. Similar to R4, but the recruitment deviations return to average after the first 15 years
R6	Evaluate impact of illegal, unreported, or unregulated catches. The catch is consistently 10% higher than the TAC
R7	Evaluates impact of additional observation error in the index of abundance. The standard deviation of the log-normal observation error in the projection years was doubled from the base robustness OM (R0)

Table 3. Summary of the Management Objectives and corresponding Performance Indicators developed for the North Atlantic swordfish MSE.

Category	*Management Objective*	*PM Name*	*PM Description*
Status	The stock should have a [51, 60, 70]% or greater probability of occurring in the green quadrant of the Kobe matrix.	PGK_{short}	Probability of being in Green Zone of Kobe Space ($SB>SB_{MSY}$ & $F<F_{MSY}$) in years 1-10 (2025-2034)
		PGK_{med}	Probability of being in Green Zone of Kobe Space ($SB>SB_{MSY}$ & $F<F_{MSY}$) in years 11-20 (2035-2044)
		PGK_{long}	Probability of being in Green Zone of Kobe Space ($SB>SB_{MSY}$ & $F<F_{MSY}$) in years 21-30 (2045-2054)
		PGK	Probability of being in Green Zone of Kobe Space ($SB>SB_{MSY}$ & $F<F_{MSY}$) over all years (2025-2054)
		PNOF	Probability of Not Overfishing ($F<F_{MSY}$) over all years (2025-2054)
Safety	There should be a [5, 10, 15]% or less probability of the stock falling below B_{LIM} ($0.4*B_{MSY}$) at any point during the 30-year evaluation period.	LRP	Probability of breaching the limit reference point (SB<0.4SBMSY) in any year (2025-2054)
Yield	Maximize overall catch levels.	TAC1	TAC (t) in the first implementation year (2025)
		$AvTAC_{short}$	Median TAC (t) over years 1-10 (2025-2034)
		$AvTAC_{med}$	Median TAC (t) over years 11-20 (2035-2044)
		$AvTAC_{long}$	Median TAC (t) over years 21-30 (2045-2054)
Stability	Any increase or decrease in TAC between management periods should be less than [25]%. [also test no stability limitation]	VarC	Mean variation in TAC (%) between management cycles over all years and simulations

Table 4. Summary of the candidate management procedures for the North Atlantic swordfish MSE.

Name	*Description*	*Tuning Code*	*Tuning Parameter*
CE	Attempts to maintain a constant exploitation rate in the projection period, based on the mean exploitation rate in the recent historical years. TAC is constrained to change no more than 25% between management cycles.	CE_b	0.8348
		CE_c	0.8157
MCC9	Aims to maintain a mostly constant catch (MCC). The TAC is adjusted between a set of 9 steps based on the ratio of the mean index over the 3 most recent years compared to the mean index from 2017 - 2019.	MCC9_b	0.7483
		MCC9_c	0.7200
MCC11	Similar to MCC9 but the 11 steps are used to adjust the TAC.	MCC11_b	0.7562
		MCC11_c	0.7316
SPSSFox	Surplus production assessment model, using a constant F policy and a linear harvest control rule that reduces fishing mortality when the estimated B/BMSY <1. TAC is constrained to change no more than 25% between management cycles.	SPSSFox_b	0.5939
		SPSSFox_c	0.5682
SPSSFox2	Same as SPSSFox, except there is no constraint on reduction in TAC if estimated B/BMSY < 1.	SPSSFox2_b	0.5939
		SPSSFox2_c	0.5682

Table 5. The probability of breaching the Limit Reference Point (LRP; $0.4SB_{MSY}$) for the tuned versions of the five candidate management procedures for the Reference Set and the Robustness Test OMs.

		Probability of Breaching LRP							
CMP	*Reference Set*	*R0*	*R1*	*R2*	*R3*	*R4*	*R5*	*R6*	*R7*
CE_b	0	0	0.11	0.03	0.58	0.25	0.70	0.06	0.01
CE_c	0	0	0.10	0.01	0.36	0.17	0.61	0.05	0.01
MCC9_b	0	0	0.12	0.03	0.51	0.14	0.61	0.03	0
MCC9_c	0	0	0.06	0.01	0.32	0.06	0.49	0.03	0
MCC11_b	0	0	0.26	0.03	0.59	0.22	0.66	0.03	0
MCC11_c	0	0	0.14	0.01	0.40	0.09	0.54	0.03	0
SPSSFox_b	0	0	0.17	0.03	0.68	0.09	0.60	0	0.01
SPSSFox_c	0	0	0.07	0.01	0.48	0.03	0.42	0	0
SPSSFox2_b	0	0	0.17	0.03	0.66	0.06	0.36	0	0.01
SPSSFox2_c	0	0	0.07	0.01	0.48	0.03	0.22	0	0

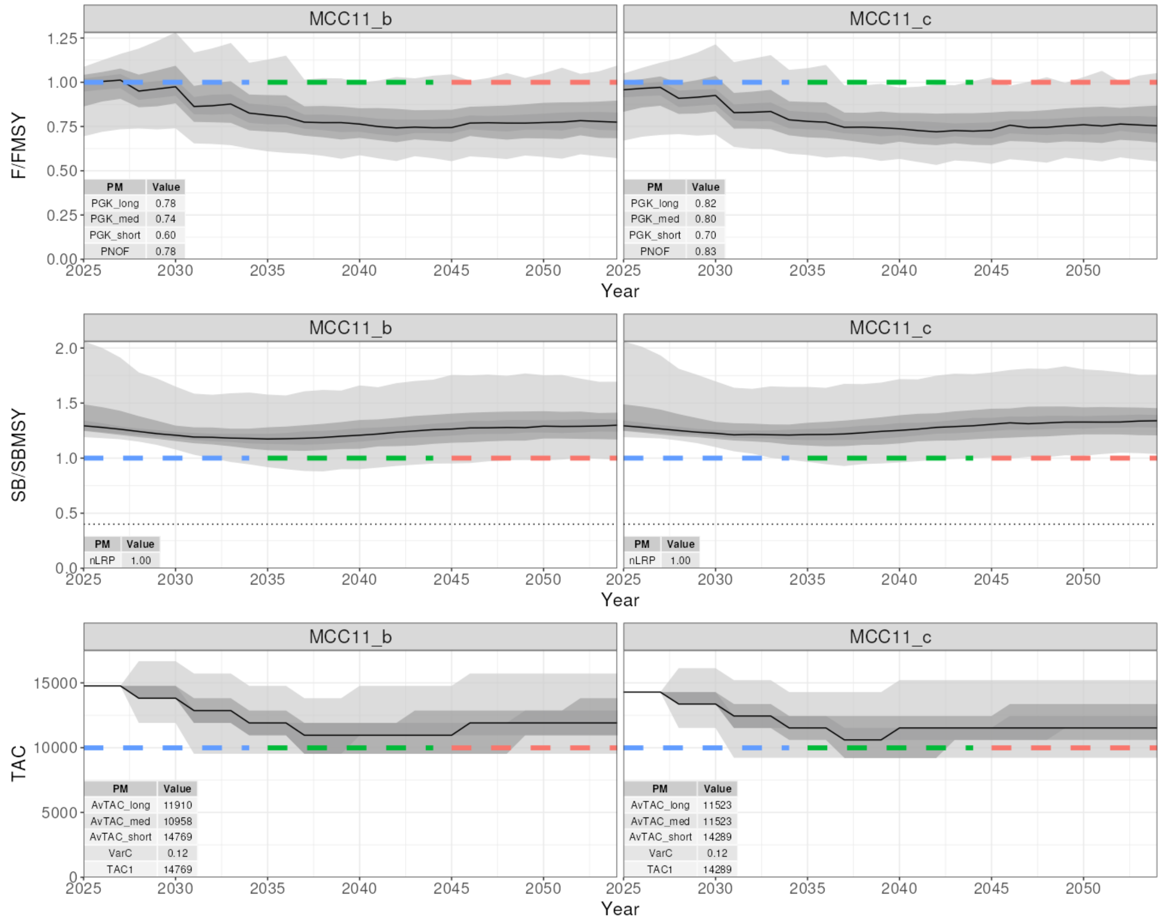

Figure 1. A set of time-series plot for one configuration of the MCC11 CMP, showing the median (black line), 60th, 70th, and 90th percentiles (increasingly lighter shades of grey respectively) for F/F_{MSY} (top), SB/SB_{MSY} (center), and the total allowable catch (TAC; bottom) over the 30-year projection period. This plot shows results for the nine reference operating models. Other plots are available for the robustness models in the Shiny application. The performance indicators associated with this configuration of the MCC11 CMP are shown in tables in the bottom left of each plot.

	MP	AvTAC_long	AvTAC_med	AvTAC_short	nLRP	PGK	PGK_med	PGK_short	PNOF	VarC	TAC1
	All	All	All	All	All	All	All	All	All	All	All
5	MCC11_b	11,911	10,958	14,769	1.00	0.71	0.74	0.60	0.78	0.12	14,769
6	MCC11_c	11,523	11,523	14,289	1.00	0.77	0.80	0.70	0.83	0.12	14,289
1	CE_b	11,820	8,266	14,172	1.00	0.79	0.87	0.60	0.83	0.18	14,172
3	MCC9_b	12,258	11,315	14,144	1.00	0.73	0.78	0.60	0.80	0.12	15,087
7	SPSSFox_b	11,557	11,397	13,869	1.00	0.73	0.75	0.60	0.79	0.14	15,629
9	SPSSFox2_b	11,556	11,397	13,869	1.00	0.73	0.76	0.60	0.80	0.14	15,629
2	CE_c	11,934	8,241	13,846	1.00	0.84	0.91	0.70	0.87	0.18	13,846
4	MCC9_c	11,794	10,887	13,609	1.00	0.80	0.84	0.70	0.85	0.12	14,516
8	SPSSFox_c	11,531	11,336	13,370	1.00	0.81	0.83	0.70	0.85	0.13	14,952
10	SPSSFox2_c	11,522	11,336	13,370	1.00	0.81	0.83	0.70	0.85	0.13	14,952

Figure 2. An example of a quilt plots that are available in the Shiny application that presents the results of the north Atlantic swordfish MSE. This table shows 10 CMP configurations (rows) and 10 performance indicators (columns) for the Reference Set of OMs. The selection of the CMPs and performance indicators can be customized in the Shiny application. The cells are shaded indicating the range of values, with darker colors indicating more desirable outcomes for the various performance indicators.

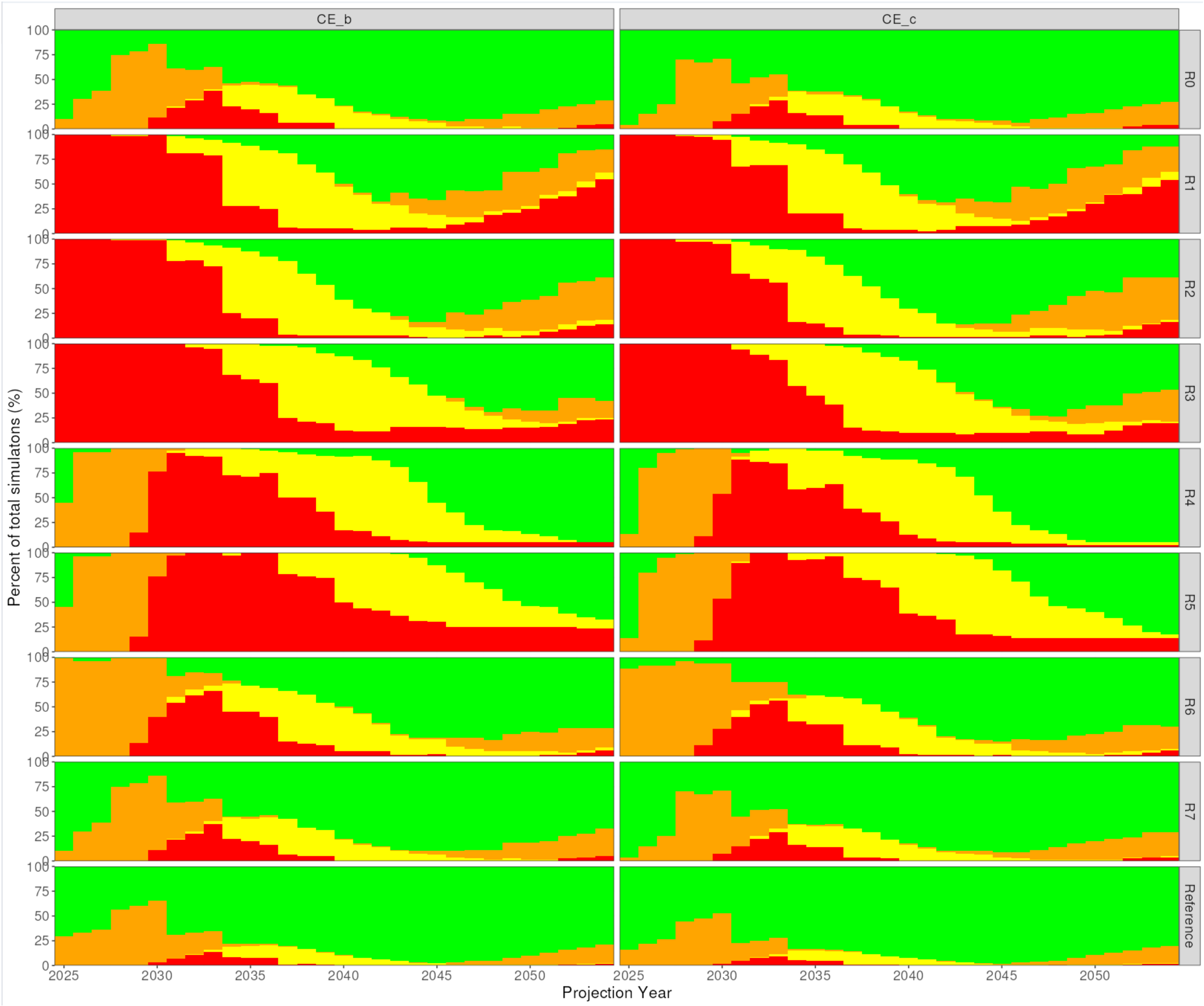

Figure 3. An example of a Kobe time plot for two configuration of the CE CMP, showing the proportion of the simulations in each quadrant of the Kobe matrix in each year of the projection period. The plot on the bottom shows the results for the Reference operating models, and the remaining plots show the results for the baseline (R0) and seven robustness models.

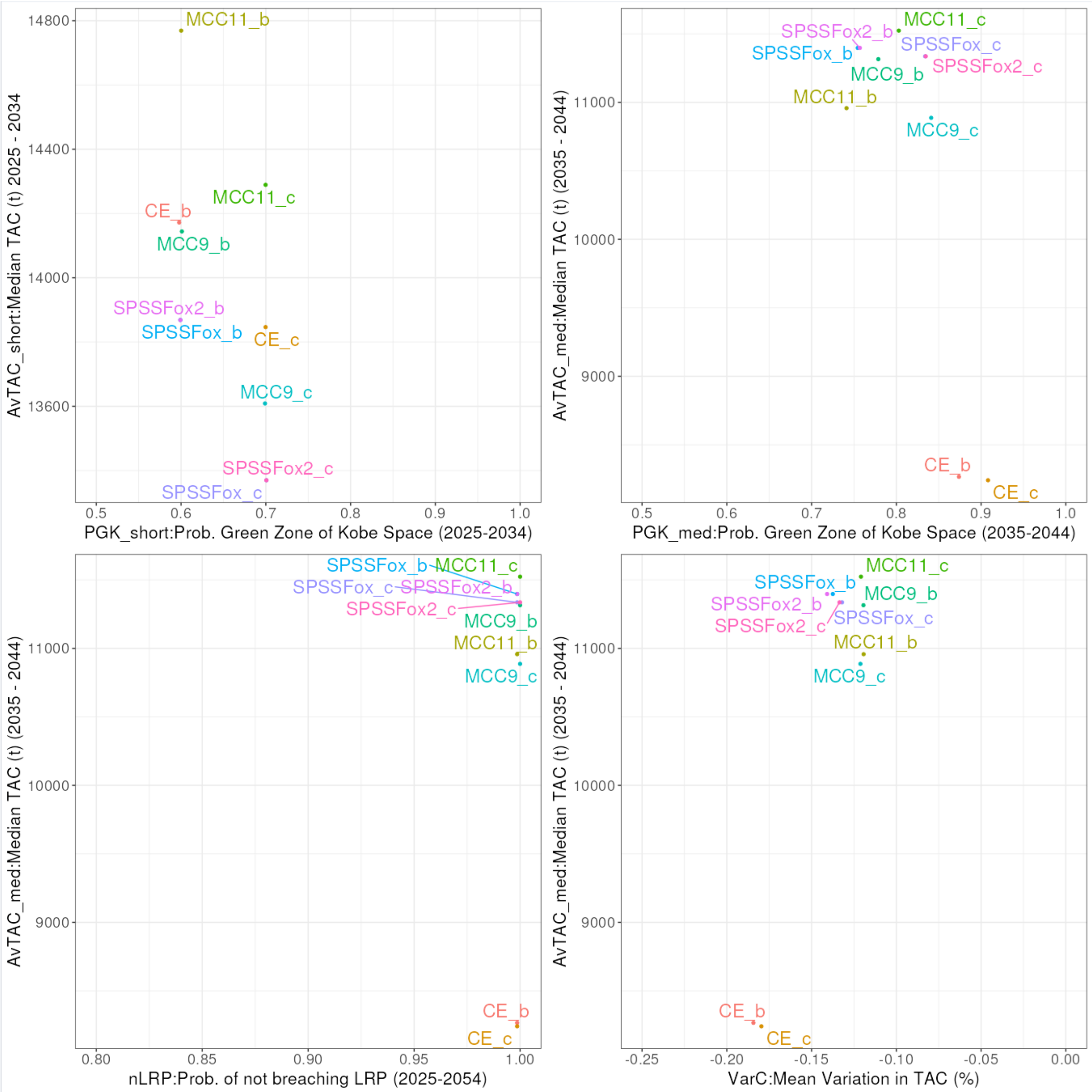

Figure 4. An example of a set of trade-off plots showing the results from 10 configurations of 5 CMPs for the Reference operating models. The plots show the trade-offs between the probability of being in the green space of the Kobe matrix (PGK) in the first 10-years of the projection period against the average TAC over this same period (top left), the PGK in years 11 – 20 against the average TAC over this same period (top right), the probability of not breaching the limit reference point against the average TAC in years 11 – 20 (bottom left), and the mean variation in TAC (shown as a negative value so lower values mean more variable) against the median TAC in the medium timeframe (bottom right).

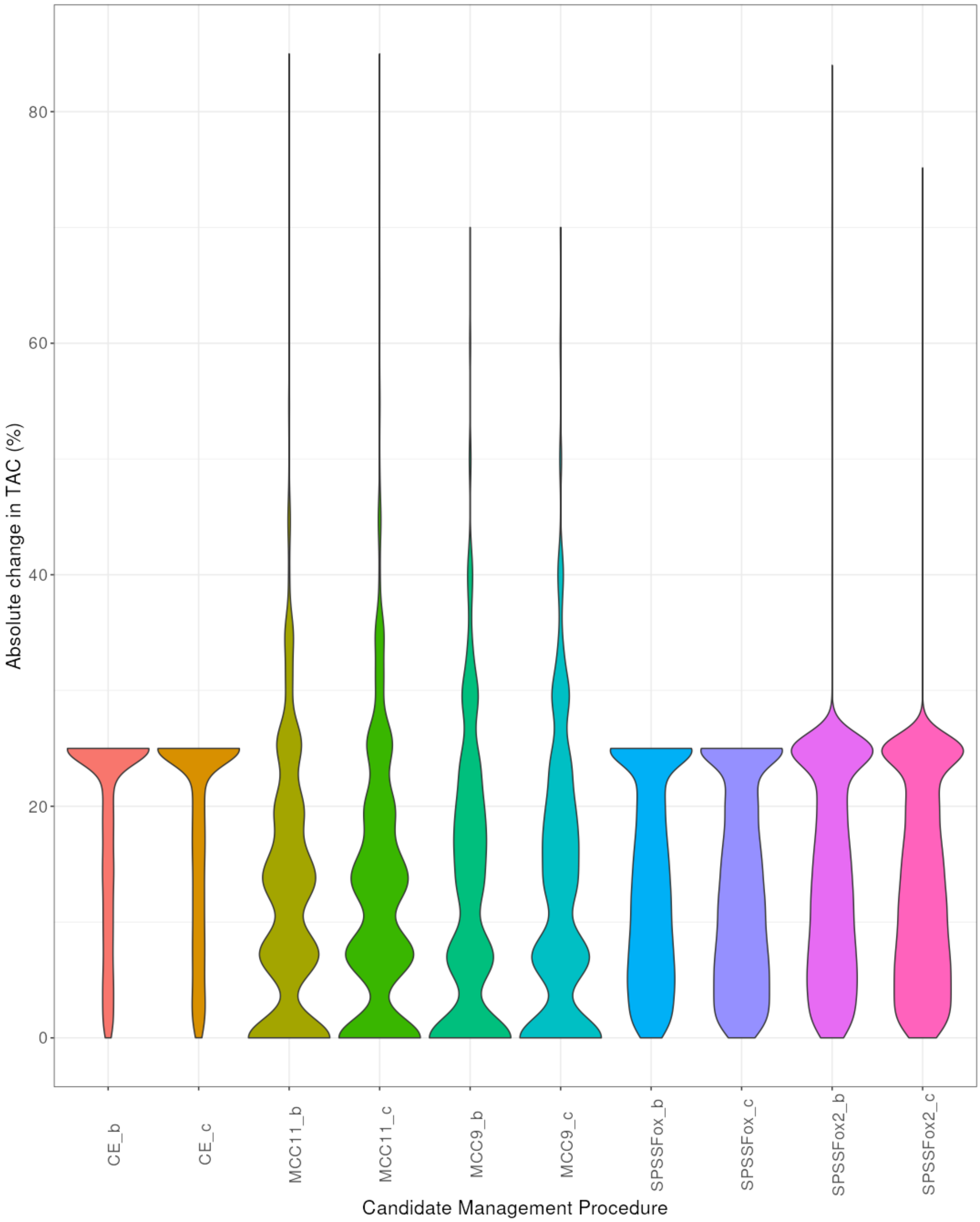

Figure 5. An example of a violin plot showing the distribution of the absolute change in TAC (y-axis) for the CMPs(x-axis). The width of the violin plot is proportional to the frequency of the absolute change in TAC (i.e., wider areas means value is more common).

	MP	AvTAC_long	AvTAC_med	AvTAC_short	nLRP	PGK	PGK_med	PGK_short	PNOF	VarC	TAC1
	All	All	All	All	All	All	All	All	All	All	All
1	CE_b	11,607	8,793	14,172	1.00	0.67	0.72	0.45	0.75	0.19	14,172
2	CE_c	11,899	8,829	13,846	1.00	0.72	0.78	0.56	0.79	0.18	13,846
3	MCC9_b	12,258	11,315	15,087	1.00	0.50	0.52	0.37	0.65	0.12	15,087
4	MCC9_c	11,794	11,794	14,516	1.00	0.63	0.66	0.54	0.74	0.12	14,516
5	MCC11_b	11,911	11,911	14,769	1.00	0.48	0.48	0.39	0.63	0.13	14,769
6	MCC11_c	11,523	11,523	14,289	1.00	0.59	0.58	0.55	0.72	0.13	14,289
7	SPSSFox_b	11,629	11,742	14,286	1.00	0.51	0.52	0.35	0.65	0.14	15,629
8	SPSSFox_c	11,638	11,743	13,782	1.00	0.64	0.64	0.53	0.74	0.13	14,952
9	SPSSFox2_b	11,593	11,742	14,286	1.00	0.52	0.52	0.35	0.65	0.14	15,629
10	SPSSFox2_c	11,638	11,743	13,782	1.00	0.64	0.65	0.53	0.74	0.13	14,952

Figure 6. Quilt plot results for the reference robustness operating model (R0).

	MP	AvTAC_long	AvTAC_med	AvTAC_short	nLRP	PGK	PGK_med	PGK_short	PNOF	VarC	TAC1
	All	All	All	All	All	All	All	All	All	All	All
1	CE_b	9,209	6,527	13,162	0.43	0.25	0.15	0.00	0.50	0.21	14,172
2	CE_c	10,054	6,625	12,639	0.64	0.28	0.22	0.00	0.54	0.21	13,846
3	MCC9_b	9,429	9,429	11,315	0.49	0.01	0.01	0.00	0.17	0.14	15,087
4	MCC9_c	10,887	9,072	10,887	0.68	0.02	0.02	0.00	0.25	0.15	14,516
5	MCC11_b	9,528	9,528	11,911	0.41	0.01	0.01	0.00	0.14	0.14	14,769
6	MCC11_c	11,063	9,219	11,523	0.60	0.01	0.01	0.00	0.20	0.14	14,289
7	SPSSFox_b	9,817	7,726	11,722	0.32	0.13	0.04	0.00	0.42	0.19	15,629
8	SPSSFox_c	10,514	8,246	11,230	0.52	0.14	0.07	0.00	0.45	0.18	14,952
9	SPSSFox2_b	9,152	6,453	11,722	0.34	0.26	0.18	0.00	0.53	0.24	15,629
10	SPSSFox2_c	10,487	7,712	11,230	0.52	0.23	0.18	0.00	0.51	0.21	14,952

Figure 7. Quilt plot results for the robustness OM R3.

	MP	AvTAC_long	AvTAC_med	AvTAC_short	nLRP	PGK	PGK_med	PGK_short	PNOF	VarC	TAC1
	All	All	All	All	All	All	All	All	All	All	All
1	CE_b	3,244	6,457	14,172	0.30	0.17	0.01	0.06	0.39	0.23	14,172
2	CE_c	3,411	6,266	13,846	0.39	0.23	0.01	0.11	0.48	0.22	13,846
3	MCC9_b	9,429	7,072	14,144	0.39	0.13	0.00	0.01	0.47	0.22	15,087
4	MCC9_c	9,072	6,804	13,609	0.51	0.16	0.00	0.04	0.50	0.20	14,516
5	MCC11_b	7,146	7,146	14,769	0.34	0.10	0.00	0.02	0.44	0.19	14,769
6	MCC11_c	9,219	6,914	14,289	0.46	0.15	0.00	0.07	0.48	0.19	14,289
7	SPSSFox_b	4,327	6,426	13,668	0.40	0.21	0.00	0.00	0.49	0.21	15,629
8	SPSSFox_c	5,001	6,482	13,146	0.58	0.24	0.00	0.03	0.53	0.21	14,952
9	SPSSFox2_b	2,977	2,544	13,668	0.64	0.35	0.08	0.00	0.63	0.29	15,629
10	SPSSFox2_c	3,493	3,201	13,146	0.78	0.35	0.07	0.03	0.62	0.27	14,952

Figure 8. Quilt plot results for the robustness OM R5.

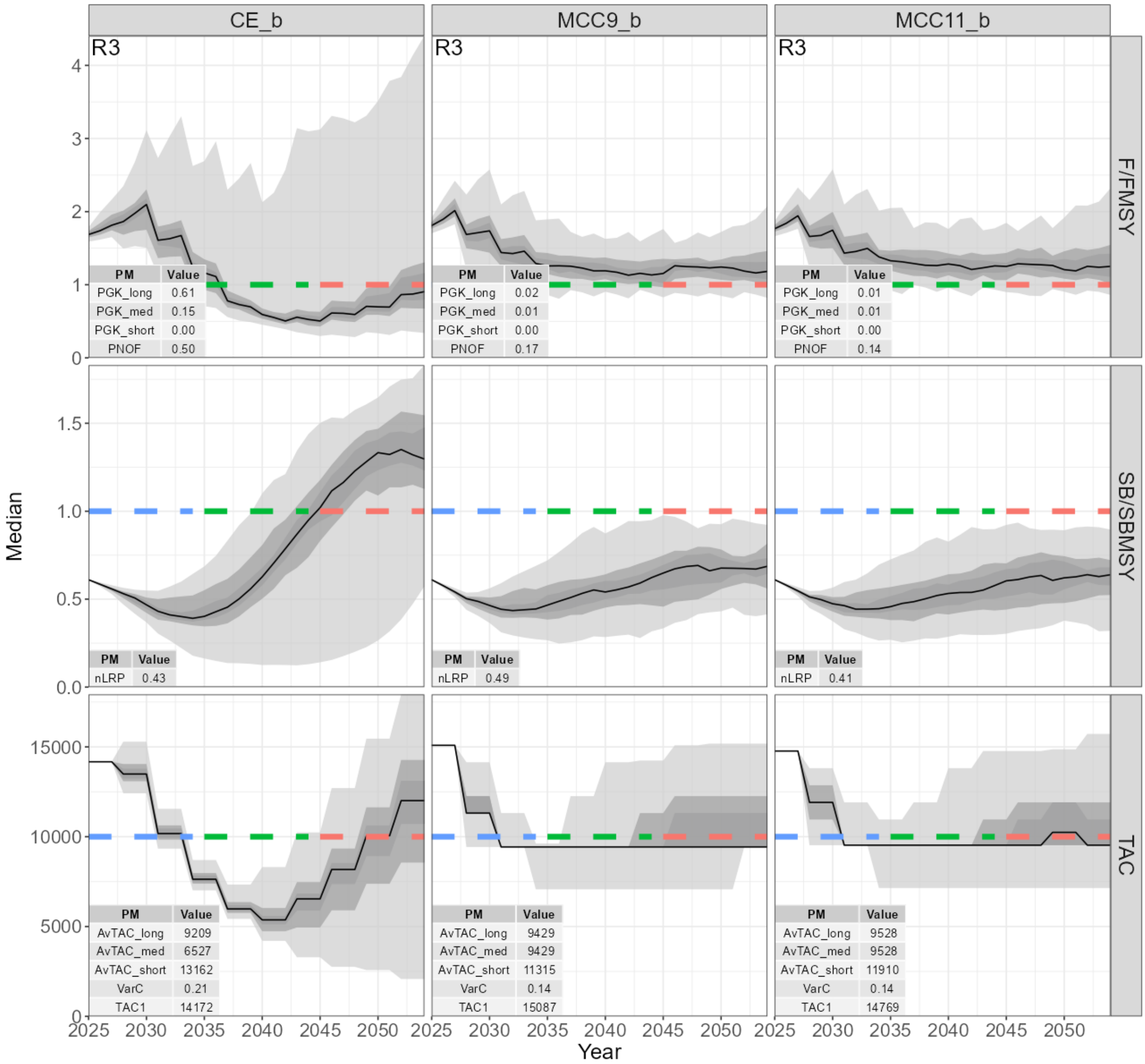

Figure 9. Time series plots with values of the performance indicators for the b tunings of the CE, MCC9 and MCC11 CMPs and the R3 robustness OM.

Appendix A

Details of the Candidate Management Procedures

1. MCC9

The goal of the MCC (Mostly Constant Catch) CMPs is to have the catch remain as constant as possible and only increase if the Combined Index increased substantially and only decrease if the Combined Index declined substantially. The base TAC (constant catch) would be 12,600, this is an approximation of the constant catch that would result in PGK60 and also achieve LRP <15%.

A base TAC ($\mathrm{TAC_{base}}$) is calculated as:

$$\mathrm{TAC_{base}} = \theta 12{,}600$$

where θ is the tuning parameter that results in achieving the desired short-term PGK (currently tested at 51%, 60%, and 70%; **Table 4**).

$\mathrm{TAC_{base}}$ is modified by comparing the ratio of the current 3-year average of the Combined Index (I_{curr}) to a historical 3-year average of the Combined Index (I_{base}):

$$I_{\mathrm{rat}} = \frac{I_{\mathrm{curr}}}{I_{\mathrm{base}}}$$

I_{base} is calculated as the average of the Combined Index from 2017-2019. The value of I_{rat} is used to determine how much $\mathrm{TAC_{base}}$ should be increased or decreased if at all.

If I_{rat} is below 0.5, the total allowable catch (TAC)is set to 4,000 t, otherwise TAC for the following management cycle is calculated as:

$$TAC_{y+1} = \mathrm{TAC_{base}}\Delta_{\mathrm{TAC}}$$

where Δ_{TAC} is calculated as:

$$\Delta_{\mathrm{TAC}} = \begin{cases} 1.7 & \text{if } I_{\mathrm{rat}} \geq 1.7 \\ 1.6 & \text{if } 1.6 \leq I_{\mathrm{rat}} < 1.7 \\ 1.5 & \text{if } 1.5 \leq I_{\mathrm{rat}} < 1.6 \\ 1.4 & \text{if } 1.4 \leq I_{\mathrm{rat}} < 1.5 \\ 1.3 & \text{if } 1.3 \leq I_{\mathrm{rat}} < 1.4 \\ 1.2 & \text{if } 1.2 \leq I_{\mathrm{rat}} < 1.3 \\ 1.0 & \text{if } 0.75 \leq I_{\mathrm{rat}} < 1.2 \\ 0.75 & \text{if } 0.5 \leq I_{\mathrm{rat}} < 0.75 \end{cases}$$

2. MCC11

MCC11 follows the same design as MCC9, but two changes: 1) it does not have a fixed minimum TAC and 2) Δ_{TAC} is calculated as:

$$\Delta_{\mathrm{TAC}} = \begin{cases} 1.85 & \text{if } I_{\mathrm{rat}} \geq 1.85 \\ 1.75 & \text{if } 1.75 \leq I_{\mathrm{rat}} < 1.85 \\ 1.65 & \text{if } 1.65 \leq I_{\mathrm{rat}} < 1.75 \\ 1.55 & \text{if } 1.55 \leq I_{\mathrm{rat}} < 1.65 \\ 1.45 & \text{if } 1.45 \leq I_{\mathrm{rat}} < 1.55 \\ 1.35 & \text{if } 1.35 \leq I_{\mathrm{rat}} < 1.45 \\ 1.25 & \text{if } 1.25 \leq I_{\mathrm{rat}} < 1.35 \\ 1.15 & \text{if } 1.15 \leq I_{\mathrm{rat}} < 1.25 \\ 1.00 & \text{if } 0.75 \leq I_{\mathrm{rat}} < 1.15 \\ 0.75 & \text{if } 0.5 \leq I_{\mathrm{rat}} < 0.75 \\ 0.5 & \text{if } I_{\mathrm{rat}} < 0.5 \end{cases}$$

a. CE

The CE management procedure aims to keep a fixed exploitation rate in the projection years. The Combined Index is used to track to relative changes in the population. A smoothed index is generated by applying Tukey's Running Median Smoother (stats::smooth R function).

The historical relative exploitation rate is calculated as:

$$E_{\text{hist}} = \frac{\bar{C}_{\text{hist}}}{\bar{I}_{\text{hist}}}$$

where $\bar{C}_{\text{hist}}$ and $\bar{I}_{\text{hist}}$ are the mean reported catch and smoothed index respectively over the 5 historical years (2016-2020).

The current relative exploitation rate is calculated as:

$$E_{\text{curr}} = \frac{\bar{C}_{\text{curr}}}{\bar{I}_{\text{curr}}}$$

where $\bar{C}_{\text{curr}}$ and $\bar{I}_{\text{curr}}$ are the mean reported catch and smoothed index respectively over the 5 most recent projection years.

The target relative exploitation rate is set to E_{hist} but subject to a harvest control rule based on the ratio of the current to historical smoothed index (I_{ratio}) (calculated over same years as above):

$$E_{\text{targ}} = \begin{cases} E_{\text{hist}} & \text{if } I_{\text{ratio}} \geq 0.8 \\ E_{\text{hist}}(-1.4 + 3I_{\text{ratio}}) & \text{if } 0.8 > I_{\text{ratio}} > 0.5 \\ 0.1E_{\text{hist}} & \text{otherwise} \end{cases}$$

The ratio of the target to current relative exploitation rate is calculated:

$$E_{\text{ratio}} = \frac{E_{\text{targ}}}{E_{\text{curr}}}$$

The total allowable catch (TAC) for the following year is then calculated as:

$$TAC_{y+1} = \theta E_{\text{ratio}} TAC_y$$

where θ is a tuning parameter (**Table 4**), subject to a constraint where it cannot change by more than 25% from one management cycle to the next.

b. SPSSFox

The SPSSFox management procedure use a state-space surplus production model assuming a Fox production curve, to set the TAC. The Combined Index is used to track to relative changes in the population. A smoothed index is generated by applying Tukey's Running Median Smoother (stats::smooth R function).

The state-space surplus production model from the SAMtool package (SAMtool::SP_SS) is used to fit to the smoothed index and the reported catch. The SP_SS R function is run with the following arguments:

- prior=list(r=c(0.21, 0.1))
- start=list(n=1)
- fix_n=TRUE

The following harvest control rule is used to set the target exploitation rate (E_{targ}):

$$E_{\text{targ}} = \begin{cases} E_{\text{prop}} & \text{if } B_{\text{curr}} \geq B_{\text{thresh}} \\ E_{\text{prop}}\left(-0.367 + 1.167\dfrac{B_{\text{curr}}}{B_{\text{thresh}}}\right) & \text{if } B_{\text{thresh}} > B_{\text{curr}} > B_{\text{lim}} \\ E_{\text{min}} & \text{otherwise} \end{cases}$$

where E_{prop} is the proposed harvest rate, calculated as $\theta 0.15$ where θ is the tuning parameter (**Table 4**), B_{curr} is the estimated biomass from the surplus production model, B_{thresh} is the estimated biomass corresponding with maximum sustainable yield, B_{lim} is $0.4B_{\text{thresh}}$, and E_{min} is $0.1E_{\text{prop}}$.

The total allowable catch (TAC) for the following year is then calculated as:

$$TAC_{y+1} = E_{\text{targ}} B_{\text{curr}}$$

The TAC is subject to a constraint where it cannot change by more than 25% from one management cycle to the next.

c. SPSSFox2

SPSSFox2 is identical to SPSSFox except that the constraint of a maximum 25% change in TAC between management cycles is not used when the assessment model estimates the biomass is less than the biomass corresponding with maximum sustainable yield.