PA4_OCT_02/i2023
04/10/2023 11:29

Original: English

North Atlantic Swordfish MSE: Background, Structure, Results and Key Decisions

Executive Summary

This document describes core concepts of the North Atlantic swordfish management strategy evaluation (MSE). The intention is to provide sufficient knowledge to facilitate decision-making and discussion among scientists, fishery managers and other stakeholders at the Panel 4 meeting on 10-11 October 2023 and continuing in the lead up to scheduled adoption of a management procedure (MP) in November 2023. This document summarizes the MSE structure, process, results, and key decisions for the October PA4 meeting.

Background

The SCRS Swordfish Species Group has been developing a management strategy evaluation (MSE) framework for North Atlantic swordfish (SWO-N) for a decade (see **Appendix D** for key terminology). In 2009, ICCAT called for development of a limit reference point (LRP) for swordfish (Rec. 09-02), and the Commission adopted $0.4*B_{MSY}$[1] as the interim limit reference point in 2013 (Rec. 13-02). Recommendation 13-02 also tasked the SCRS with development of a harvest control rule for SWO-N. In 2015, the Commission called for adoption of a management procedure based on an MSE for 8 priority stocks, including SWO-N (Rec. 15-07). In 2017, the SCRS developed an integrated, size-structured stock assessment model for SWO-N on which a future MSE would be based. Funds were provided by the Commission in 2018 to develop the simulation framework, and following initial work by the SCRS, an MSE expert was contracted in 2019 to develop the SWO-N MSE. MSE development by the SCRS then began in earnest. The Commission adopted conceptual management objectives for SWO-N in 2019 (Res. 19-14) to help guide MSE development. In 2022, the SCRS carried out a new stock assessment in which the base case model was modified to incorporate discard mortality of undersized fish, and the MSE was updated with this new model. MSE development has continued in 2023, incorporating feedback provided by Panel 4 at its March and June meetings. The MSE work is on track for ICCAT to adopt an MP in 2023, in accordance with the Commission's MSE workplan. Development of an exceptional circumstances protocol and additional robustness testing, as requested by Panel 4, will be completed in 2024.

MSE overview

The SWO-N MSE is built using an open-source MSE software package called openMSE. The package can input information from Stock Synthesis stock assessments (the 2022 SWO-N stock assessment, in this case) to efficiently create – and then customize – an MSE framework for testing candidate management procedures (CMPs).

Indices of abundance

Data from 6 different longline indices were used in the stock assessment and are used to condition the MSE. A combined index that incorporates raw data from seven CPCs is being used as the primary index for CMP development. The MSE's historical period is from 1950 through to 2020, and projections cover the subsequent 33 years.

Operating Models

Each operating model (OM) in the MSE represents a plausible scenario/a potential truth for the dynamics of the stock and fishery. The SWO-N MSE includes nine main operating models (i.e., the "reference set or grid of OMs") based on two major sources of uncertainty:

[1] Spawning stock biomass (SSB; biomass of mature females), is used in this MSE.

1. Stock productivity: Steepness of the relationship between stock size and recruitment potential is one of the most important and uncertain inputs into stock assessments. Practically, this is often thought of as a measure of the stock's ability to rebuild biomass when depleted to a low level (3 options);

2. Natural mortality: the rate at which individuals die of natural causes (3 options).

The nine OMs allow for all combinations of these options (3x3=9). All OMs are considered to be equally plausible, so they are weighted equally. The three steepness values were updated in May 2023 in response to a Panel 4 request, and the new reference set was reconditioned.

There are also five sets of robustness tests to evaluate the performance of the CMPs under alternative scenarios, similar to "sensitivity runs" in a stock assessment. These include 1) an assumed 1% annual increase in catchability in both the historical and projections; 2) an assumed 1% annual increase in catchability in only the historical period; 3a) climate change effects in the projection period through a cyclical pattern in the recruitment deviations; 3b) climate change effects where recruitment deviations are lower than expected for the first 15 years of the projection period; and 4) a scenario with 10% underreporting of catch.

Management objectives

The SWO-N MSE currently includes 10 key performance metrics as a benchmark for evaluation of the Commission's selected management objectives. **Appendix A** shows the current management objectives and performance metrics based on input received from Panel 4 in March and June 2023.

Importantly, all yield performance metrics calculate the total allowable catch (TAC) as landings plus dead discards (as estimated in the 2022 stock assessment).

Candidate Management Procedures

The SCRS Swordfish Species Group has worked collaboratively to develop and test a number of CMPs. Five (5) CMP types remain, as described in **Appendix B**. These 'short-listed' CMPs currently assume a 3-year management cycle and calculate a single TAC for the North Atlantic. This short-list includes both model-based and empirical CMPs (empirical CMPs use an index of abundance to directly set the TAC rather than running a statistical model). The North Atlantic albacore MP (Rec. 21-04) is model-based, whereas the Atlantic bluefin tuna MP (Rec. 22-09) is empirical.

In addition to representing both model-based and empirical CMPs, the 5 remaining CMP types are SCRS-recommended because they cover a wide range of the performance tradeoff space, use a variety of TAC-setting rules, and because they use the combined index, which includes data from the broadest geographic and fleet coverage. CMPs are tuned to 51%, 60%, and 70% probability of being in the Kobe green quadrant in years 1-10 (i.e., the PGK_{SHORT} performance metric). Tuning means that all CMPs must achieve this performance standard; further, tuning allows for comparison among the CMPs across the full suite of performance metrics. The Safety minimum threshold requires that CMPs have greater than 85% probability of not breaching the limit reference point at any point in the projection period. All CMPs achieve the Safety minimum threshold, with all achieving 95% or greater probability of not breaching the LRP. Performance against other objectives is then compared. In addition, three of the CMPs (CE, FX4 MCC7) were tested with 1) a 4-year management cycle and 2) a minimum 200 t TAC change threshold between management cycles (i.e., where the TAC is not changed between management cycles when the CMP generates a TAC that is within 200 t of the TAC of the previous cycle).

The proposed schedule for CMP implementation is included in **Appendix C** and includes data requirements for each step, as well as a schedule for review of the MSE model assumptions.

Final results

Included here are the key performance results for the 5 remaining CMP types. The full suite of results is available online in the interactive application (see "Other resources" below). Each CMP type has three tuning variants for Stock Status: 'a' is PGK_{SHORT}=51%; 'b' is PGK_{SHORT}=60%; and 'c' is PGK_{SHORT}=70%. For three of the CMPs, the 'a' variants did not meet performance metric minimum standards and were omitted from the plots, resulting in a final list of 12 CMPs from the 5 CMP families.

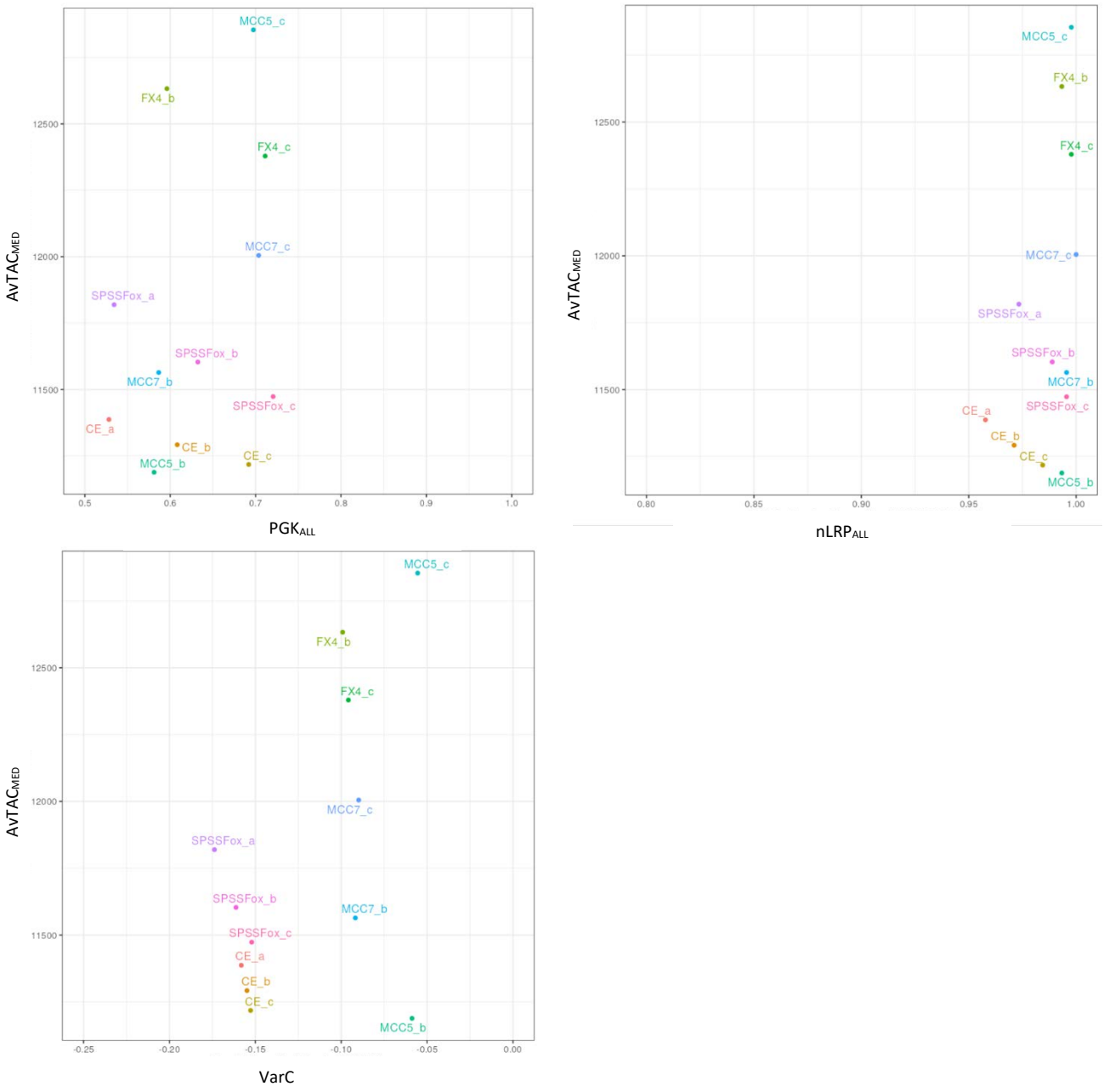

Figure 1. Plots showing the key tradeoffs between Yield ($AvTAC_{MED}$) on the vertical axis and a) Status (PGK_{ALL}), b) Safety ($nLRP_{ALL}$), and c) Variability (VarC) on the horizontal axis for the 5 short-listed CMPs. VarC is shown as a negative value so lower values mean more variable. For all plots, each colour indicates a different CMP type. Performance metrics are described in **Appendix A**, and CMP types are described in **Appendix B**.

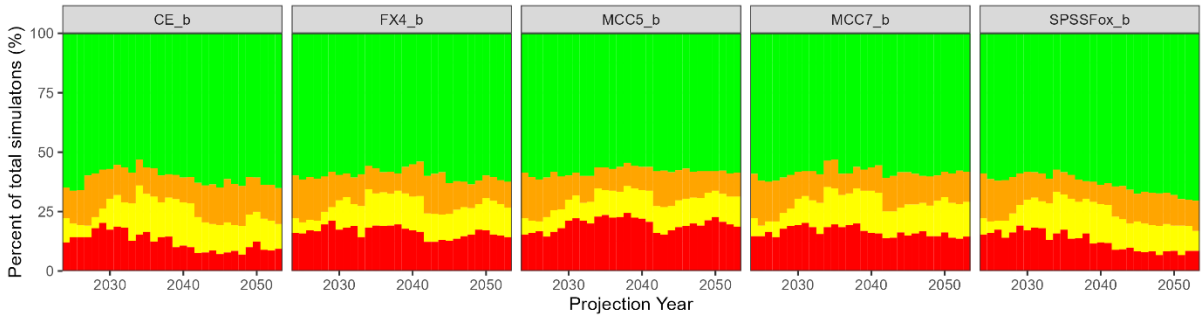

Figure 2. Kobe time plot showing the median percentage (vertical axis) of simulations across all reference operating models that fall in each of the Kobe quadrants in each projection year (horizontal axis). Green indicates that the stock is neither overfished nor subject to overfishing. Orange means that the stock is subject to overfishing but not overfished. Yellow indicates that the stock is overfished but not subject to overfishing. Red means that the stock is both overfished and subject to continued overfishing. CMP types are described in **Appendix B**. Results for only the 'b' tuning (PGK_{SHORT}=60%) are shown here.

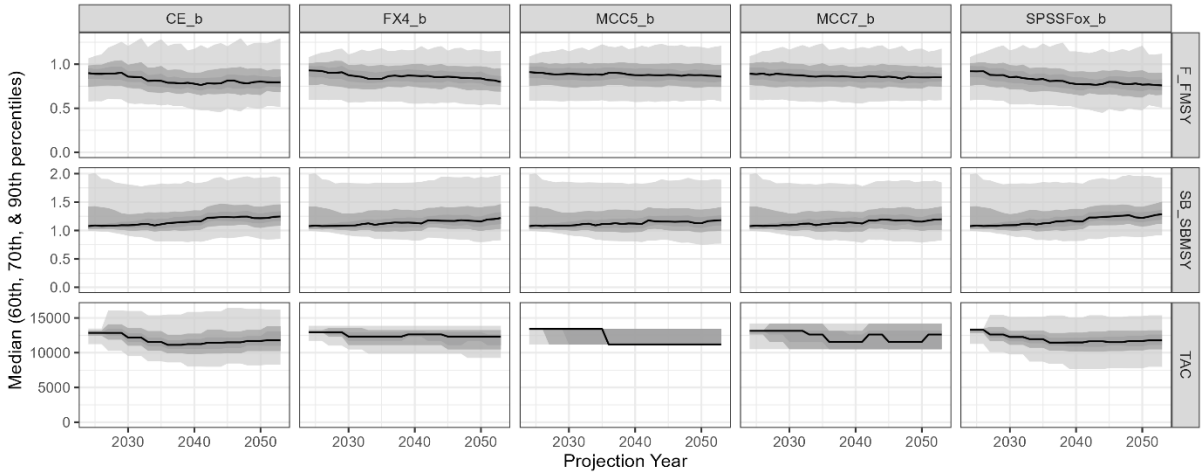

Figure 3. Trajectory of a) fishing mortality (F) relative to F_{MSY} (top row), b) spawning stock biomass (SSB) relative to SSB_{MSY}, and c) TAC (in tons, bottom row) for the 5 short-list CMP types. Results are summarized across all reference operating models. CMP types are described in **Appendix B**. Results for only the 'b' tuning (PGK_{SHORT}=60%) are shown here.

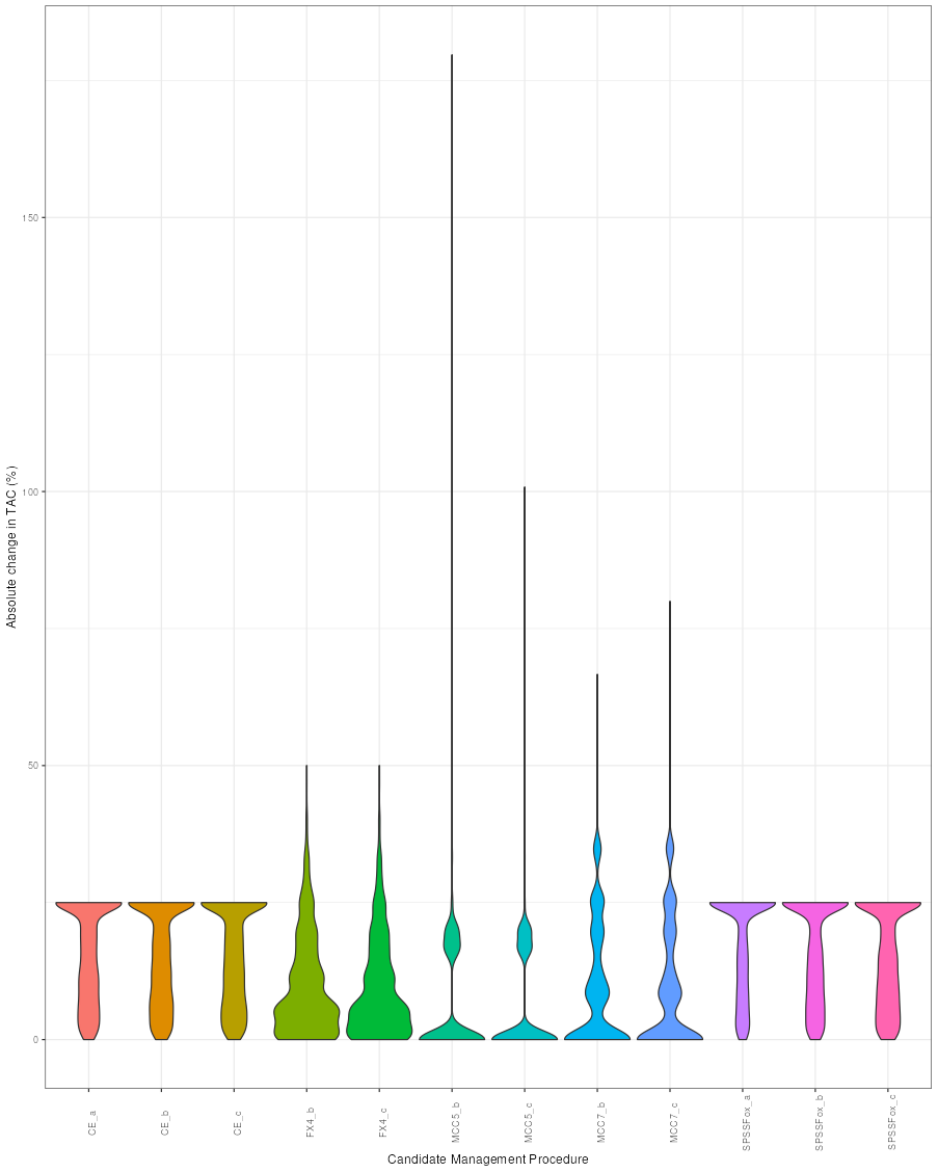

Figure 4. Violin plot for the change in TAC between management cycles. Note that some of the CMPs (CE, SPSSFox) include a 25% cap on TAC change. The width of the violin plot indicates the proportion of data points that are in each region of the plot (i.e., wide areas of the plot indicate a relatively large number of data points in that region, while narrow areas of the plot indicate few data points). CMP types are described in **Appendix B**.

	MP	AvTAC_long	AvTAC_med	AvTAC_short	nLRP	PGK	PGK_med	PGK_short	PNOF	TAC1	VarC
1	CE_a	11660	11390	13450	0.96	0.53	0.51	0.51	0.68	13460	0.16
2	CE_b	11650	11290	12770	0.97	0.61	0.59	0.6	0.74	12860	0.15
3	CE_c	11560	11220	12160	0.98	0.69	0.68	0.7	0.79	12250	0.15
4	FX4_a	12230	12870	13520	0.99	0.49	0.47	0.51	0.61	13520	0.1
5	FX4_b	12320	12630	12940	0.99	0.6	0.57	0.6	0.71	12940	0.1
6	FX4_c	12080	12380	12380	1	0.71	0.7	0.7	0.82	12380	0.1
7	MCC5_a	11710	11710	14050	0.97	0.48	0.47	0.51	0.57	14050	0.06
8	MCC5_b	11190	11190	13430	0.99	0.58	0.56	0.6	0.68	13430	0.06
9	MCC5_c	12850	12850	12850	1	0.7	0.68	0.7	0.8	12850	0.06
10	MCC7_a	11030	11030	13780	0.99	0.49	0.48	0.51	0.61	13780	0.09
11	MCC7_b	11560	11560	13140	1	0.59	0.57	0.6	0.71	13140	0.09
12	MCC7_c	12510	12010	12510	1	0.7	0.69	0.7	0.81	12510	0.09
13	SPSSFox_a	11790	11820	13460	0.97	0.53	0.51	0.51	0.67	13460	0.17
14	SPSSFox_b	11680	11600	12750	0.99	0.63	0.62	0.6	0.75	13290	0.16
15	SPSSFox_c	11570	11470	12190	1	0.72	0.7	0.7	0.82	12520	0.15

Figure 5. Quilt plot showing results for the 5 remaining CMPs types (each with up to three Status tuning options: PGK_{SHORT}=51% - 'a', 60% - 'b', or 70% - 'c') against the 10 key performance metrics. CMPs are listed in alphabetical order. See **Appendix A** for performance metric descriptions and **Appendix B** for CMP descriptions. For three of the CMPs, the 'a' variants do not meet performance metric minimum standards (noted in red), resulting in a final list of 12 CMPs from the 5 CMP families. The nLRP performance metric is the probability of not breaching the limit reference point; this modification of the LRP performance metric means that higher values are better for all metrics except VarC. Darker shading indicates better performance, but some of the values are very similar, despite different shading.

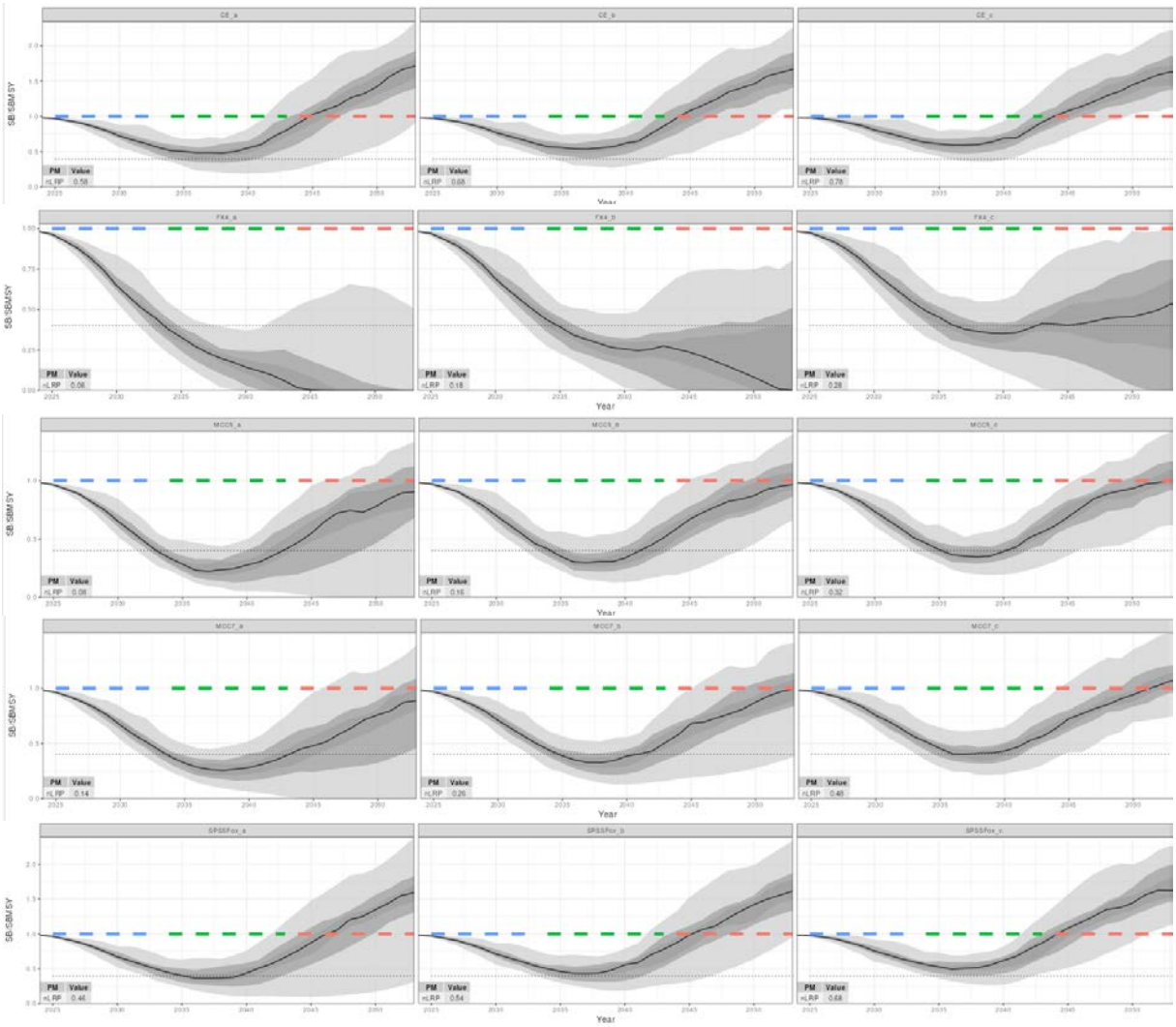

Figure 6. Trajectory of spawning stock biomass (SSB) relative to SSB_{MSY} for all CMPs under the climate change robustness test 3b (features a decline in recruitment in the first fifteen years, followed by a return to average recruitment for the remainder of the projection period). This robustness operating model, R3b, presents the biggest challenge to CMPs compared to all other OMs. The first column is the 'a' CMPs (PGK_{SHORT}=51%), the second column is the 'b' CMPs (PGK_{SHORT}=60%), and the last column is the 'c' CMPs (PGK_{SHORT}=70%). The small-dashed black horizontal line indicates the LRP of $0.4*SSB_{MSY}$. The coloured horizontal line shows the SSB_{MSY} target over the short (blue), medium (green) and long (red) terms. The dark black trend line shows the median value of SSB, while the increasingly lighter shades of grey show the 50th, 60th, and 90th percentiles, respectively.

Next steps and key decisions

The 10-11 October meeting is the third and final Panel 4 meeting for the exchange of information between the SCRS and Panel 4 in advance of the 2023 Commission meeting. The Swordfish Species Group is also using ambassador sessions to help improve understanding of the MSE and answer questions (12 June and 5 October 2023).

At the 10-11 October 2023 meeting, Panel 4 should prepare to make the following decisions to select a final MP to recommend to the Commission for adoption in November:

a) Choice of final operational management objectives (See Appendix A), including:

- Minimum acceptable threshold for the Status objective.[2] Options are 51%, 60% or 70% probability of occurring in the green quadrant of the Kobe matrix.
    - See **Figures 1, 2, 3, 4** and **5** to compare performance across the three options.
- Minimum acceptable threshold for the Safety objective. Options are 85%, 90% or 95% probability of the stock not falling below B_{LIM} ($0.4*B_{MSY}$) at any point during the 30-year evaluation period.
    - See **Figures 1, 2, 3** and **5** to compare performance across the three options.
    - Note that all CMPs in the short-list meet the most stringent safety objective threshold (95%).
- Maximum percent allowable change in TAC between management periods. Options are 25% and no limit.
    - See **Figures 3, 4** and **5** to explore the impacts of the various stability options.
    - CMPs were tested with both no caps and 25% caps on the percent change in TAC allowable between adjacent management cycles. These variations sometimes produced CMPs that performed poorly, and these were eliminated from consideration.

b) Final CMP type

- There are 5 remaining CMP types – CE, FX4, MCC5, MCC7 and SPSSFox tuned to 51%, 60%, and 70% PGK, for a total of 15 variants.
- Each CMP uses the combined index.
- Three of the CMP variants (FX4_a, MCC5_a, MCC7_a) do not meet the minimum operational management objectives for Stock Status and Safety. However, the other 12 viable CMPs do meet the minimum operational objectives but with varying performance on the Yield and Stability tradeoffs.
- The relative performance results are provided above in **Figures 1-6**. Although the performance is relatively similar, the FX4 CMPs have notably lower performance under the climate change robustness test.

[2] Tuning objective: CMPs are currently tuned to 51%, 60%, and 70% probability of being in the Kobe green quadrant in years 1-10 (i.e., PGK_{SHORT} performance metric), as agreed by Panel 4 at the March meeting. The final minimum probability agreed for the operational objective for Status will be assumed as the final tuning objective.

c) Final MP specifications

- Management cycle length: All CMPs were tested using a 3-year management cycle (as outlined in **Appendix C**). In addition, 3 CMPs (CE, FX4, MCC7) were evaluated using a four-year management cycle. The results indicated very little difference in performance between the 3-year and 4-year options. However, the testing for the 4-year management cycle was done only for the reference set of OMs, which is not as challenging as the robustness set. The lower level of responsiveness from a 4-year management cycle might have lower performance under the robustness OMs.

- Minimum TAC change: At each application of the MP, it may be desirable to set a minimum bound for TAC change for administrative purposes. A minimum TAC change is part of the bluefin MP but not the albacore MP. The SCRS tested a 200 t minimum TAC change for 3 CMPs (CE, FX4, MCC7) and found identical performance since all TAC changes are projected to be either 0 t or greater than 200 t.

d) MP implementation schedule

- A key element of the process of management procedure implementation is the process of its review. Such a review can occur at regular, prescheduled intervals or following the declaration of exceptional circumstances. In most cases, such a review would not constitute a wholesale revision to the operating model structure, full reconditioning of the OMs or substantial changes to the CMPs, though it offers that opportunity should the need arise. In most cases, such reviews could implement index revisions or relatively minor improvements to the operating models or MPs; indeed, the outcome may leave the MP unchanged. The proposed MP implementation schedule is included in **Appendix C** for Panel 4's review and approval.

Other resources

North Atlantic Swordfish MSE splash page
North Atlantic Swordfish MSE interactive Shiny App (includes final results)
Harveststrategies.org MSE outreach materials (multiple languages)

Appendix A

Current management objectives and corresponding performance metrics based on input received at the March and June 2023 Panel 4 meetings. Importantly, all yield performance metrics calculate the TAC as landings plus dead discards.

Management objectives	*Corresponding key performance metrics*
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 the Kobe green quadrant (i.e., $SSB \geq SSB_{MSY}$ and $F < F_{MSY}$) in years 1-10 **PGK_{MED}**: Probability of being in the Kobe green quadrant (i.e., $SSB \geq SSB_{MSY}$ and $F < F_{MSY}$) in years 11-20 **PGK_{ALL}**: Probability of being in the Kobe green quadrant (i.e., $SSB \geq SSB_{MSY}$ and $F < F_{MSY}$) over years 1-30 **PNOF:** Probability of not overfishing ($F < F_{MSY}$) over years 1-30
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_{ALL}**[3]: Probability of breaching the limit reference point (i.e., $SSB < 0.4*SSB_{MSY}$) in any of years 1-30
Yield Maximize overall catch levels.	**TAC1** – TAC in the first management cycle (years 1-3) **$AvTAC_{SHORT}$** – Median TAC (t) over years 1-10 **$AvTAC_{MED}$** – Median TAC (t) over years 11-20 **$AvTAC_{LONG}$** – Median TAC (t) over years 21-30
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 years 1-30

[3] nLRP (not breaching the LRP) is used when it is more appropriate for higher values of performance metrics to indicate a 'safer' outcome, such as in trade-off plots. For example, a 15% LRP threshold is equivalent to a nLRP threshold of 85%.

Appendix B

Final Candidate Management Procedures (CMPs). Stability – Yes indicates a 25% cap on TAC changes between management cycles, except when the stock is outside the Kobe green quadrant, at which time there is no cap on TAC decreases. CMPs are named according to their tuning target: 'a' for PGK_{SHORT}=51%, 'b' for PGK_{SHORT}=60%, and 'c' for PGK_{SHORT}=71%.

MP	*Type*	*PGK tuning target*	*Indices*	*Stability*	*Description*
CE	Empirical	51%, 60%, 70%	Combined	Yes	Attempts to maintain a constant exploitation rate in the projection period, based on the mean exploitation rate in the recent historical years.
FX4	Empirical	51%, 60%, 70%	Combined	Yes	Index ratio method using the Combined Index for the most recent 3 years, smoothed and scaled by the inverse variance before averaging.
MCC5	Empirical	60%, 70%	Combined	No	Mostly Constant Catch 5 (MCC5) focuses on trying to provide a stable TAC. To do this it uses a base TAC which has the possibility of increasing by one step and decreasing by 2 steps. These steps are selected depending on the value of the current 3-year average of the Combined Index compared to a 3-year historical average (2017-2019). The TAC is set at a minimum (4kt) when the current 3-year average of the Combined Index is less than half of the 3-year historical average.
MCC7	Empirical	51%, 60%, 70%	Combined	No	Mostly Constant Catch 7 (MCC7) focuses on trying to provide a stable TAC. To do this it uses a base TAC which has the possibility of increasing by four small steps and decreasing by 2 steps. These steps are selected depending on the current 3-year average of the Combined Index compared to a 3-year historical average (2017-2019). The TAC is set at a minimum (50% of the base TAC) when the current 3-year average of the Combined Index is less than half of the 3-year historical average. When the 3-year average of the Combined Index is calculated, a smoother is used to reduce its variability year-to-year.
SPSSFox	Model	51%, 60%, 70%	Combined	Yes	Fox Surplus Production model with a hockey-stick HCR where fishing mortality decreases linearly from $X*B_{MSY}$ to $Y*B_{MSY}$.

Appendix C

Proposed schedule for data provision, updating MPs and stock assessments

		Activity					***Data inputs***	
Year	Management cycle	MP run	MP advice implemented	Stock assessment	MSE Review	Exceptional circumstances evaluated	Combined index	Exceptional circumstance indicators
2023		x					x	x
2024	1		x			x		x
2025	1					x		x
2026	1	x				x	x	x
2027	2		x			x		x
2028	2			x (alternative)		x		x
2029	2	x		x		x	x	x
2030	3		x	x (alternative)		x		x
2031	3					x		x
2032	3	x			x	x	x	x

Appendix D

Key terminology used in this document

Limit reference point (LRP): A benchmark for an indicator that defines an undesirable biological state of the stock such as the B_{LIM} or the biomass limit which is undesirable to be below. To keep the stock safe, the probability of violating an LRP should be very low. In many cases, nLRP (not breaching the LRP) is used when it is more appropriate for higher values of performance metrics to indicate a 'safer' outcome, such as in trade-off plots. For example, a 15% LRP threshold is equivalent to an nLRP threshold of 85%.

Management objectives: Formally adopted social, economic, biological, ecosystem, and political (or other) goals for a stock and fishery. They include high-level or conceptual objectives often expressed in legislation, conventions or similar documents. They must also include operational objectives that are specific and measurable, with associated timelines. When management objectives are referenced in the context of management procedures, the latter, more specific definition applies, but sometimes conceptual objectives are adopted first (e.g., Rec. 19-14 for SWO-N).

Management procedure (MP): Some combination of monitoring, assessment, harvest control rule and management action designed to meet the stated objectives of a fishery, and which has been simulation tested for performance and adequate robustness to uncertainties. Also known as a harvest strategy.

Management strategy evaluation (MSE): A simulation-based, analytical framework used to evaluate the performance of multiple management procedures relative to the pre-specified management objectives.

Operating model (OM): A model representing a plausible scenario for stock and fishery dynamics that is used to simulation test the management performance of CMPs. Multiple models will usually be considered to reflect the uncertainties about the dynamics of the resource and fishery, thereby testing the robustness of management procedures.

Performance statistic: A quantitative expression of a management objective used to evaluate how well an objective is being achieved by determining the proximity of the current value of the statistic to the objective. Also known as a performance metric or performance indicator.

Reference Grid: The operating models that represent the most important uncertainties in stock and fishing dynamics, which are used as the principal basis for evaluating CMP performance. The reference operating models are specified according to factors (e.g., natural mortality rate) that have multiple levels (possible scenarios for each factor, e.g., high / low natural mortality rate). Reference OMs are usually organized in a fully crossed orthogonal 'grid' of all factors and levels.

Robustness Set: Other potentially important uncertainties in stock and fishing dynamics may be included in a Robustness Set of tests that provide additional tests of CMP performance robustness. They can be used to further discriminate between CMPs. Compared to the Reference Grid operating models, the Robustness Set will be typically less plausible and/or influential on performance.