

Specifications for MSE Trials for North Atlantic Swordfish

Adrian Hordyk adrian@bluematterscience.com

11 September, 2023

Contents

1	Introduction	2
1.1	About this document	2
2	MSE Framework	3
3	Stock Assessment	3
3.1	Data	3
3.2	Model Structure	5
3.3	Fixed Biological Parameters	5
3.4	Selectivity	5
3.5	Retention and Discard Mortality	5
3.6	Stock-Recruitment	5
4	Overview of Operating Model Conditioning	6
4.1	OM Uncertainty Grid	6
4.2	Evaluation of the OM Grid	6
5	Operating Models	7
5.1	Base Case	7
5.2	Reference Operating Models	7
5.3	R0. Reference Model	8
5.4	R1. Hyperstable Indices in Historical and Projection Periods	9
5.5	R2. Hyperstable Indices in Historical Period	9
5.6	R3a and R3b. Effect of Climate Change on Recruitment Deviations in Projection Period	9
5.7	R4. TAC Implementation Error	10
6	OM Validation	10
6.1	OM Summary Report	10
6.2	OM Diagnostic Reports	10
7	Historical Spool-Up Period	12
7.1	Historical Data	12
8	Closed-Loop Simulation Testing	13
8.1	Simulation Specifications	13
8.2	Assumptions	13
8.3	Generation of Future Data	16
8.4	Management Interval	16
9	Performance Metrics	17
9.1	Summary of Performance Metrics	17

9.2	Peformance Metrics Examples	17
9.3	Status	17
9.4	Safety	22
9.5	Yield	29
9.6	Stability	29
10	Candidate Management Procedures	32
10.1	Tuning of Candidate Management Procedures	32
11	Reporting of MSE Results	33
11.1	Interactive Application	33
References		33

1 Introduction

The North Atlantic swordfish fishery, under the management of the [International Commission for the Conservation of Atlantic Tuna](#) (ICCAT), is undergoing a management strategy evaluation (MSE) process.

ICCAT describes [MSE](#) as:

a collaborative process between Scientists and decision-makers that involves using computer simulation to compare the relative ability to achieve a set of management objectives using alternative Management Strategies, defined as different combinations of data collection schemes, methods of analysis, harvest control rules and subsequent processes leading to management actions.

There are three main components in an MSE process:

1. **Operating models (OMs):** a collection of mathematical/statistical models that describe alternative hypotheses of the historical fishery dynamics and specifications for simulating the collection of data and implementation of management measures in the future;
2. **Candidate management procedures (CMPs):** a set of proposed algorithms that generate management recommendations from fishery data, and will be evaluated in the MSE;
3. **Performance metrics (PMs):** statistics use the quantitatively evaluate the CMPs against specified management objectives.

The operating models, candidate management procedures, and performance metrics are developed as a collaborative effort between scientists, decision-makers, and other stakeholders in the fishery.

1.1 About this document

This document describes the specifications for the OMs, CMPs, and PMs that have been proposed and developed for the North Atlantic swordfish (hereafter swordfish) fishery.

It is a living document and will be continued to be updated so that it reflects the current state of the swordfish MSE process.

Members of the Swordfish Species Group (hereafter the Group) are encouraged to provide feedback, comments, or edits to any part of this document.

The document is written using the [Markdown](#) format and can edited in any text editor. The source document is available on the [ICCAT/nswo-mse](#) GitHub repository.

Group members can make edits to the document either [directly in the online repository](#) or by cloning the repository and submitting [pull requests](#) with their edits. Alternatively, they can email questions or comments to [Adrian](#). The former approach has the advantage that all comments, questions, and edits are immediately visible to all members of the Group.

Group members can also use the [Discussions](#) feature on the Github repository to post questions, comments, or points for discussion related to any aspect of this document or the MSE process in general.

This document is available at the North Atlantic Swordfish MSE [homepage](#).

2 MSE Framework

The `SWOMSE` R package has been developed to conduct the MSE for the Atlantic swordfish fishery. All code used for the MSE is open-source and reproducible. Group members can install the package and reproduce the analysis on their own machines.

The `SWOMSE` package used for the swordfish MSE uses the `openMSE` framework. The `openMSE` package is automatically installed when `SWOMSE` is installed, and all `openMSE` functions are available to the user when the `SWOMSE` package is loaded.

`openMSE` is an R package that has been developed for conducting fast, flexible, and transparent, MSE for a wide range of fisheries. `openMSE` is an umbrella package that includes the `MSEtool`, `SAMtool` and `DLMtool` packages. A non-technical description of `openMSE` and its key features is available on the [openMSE website](#).

The operating model in `openMSE`, including assumptions and equations, is described in detail in Carruthers & Hordyk (2018).

The [SWOMSE User Manual](#) includes instructions for installing the `SWOMSE` package and running the MSE analyses.

3 Stock Assessment

Previously, the swordfish operating models were based on the 2017 stock assessment (Anon., 2017) using Stock Synthesis 3 (SS3, Methot & Wetzel, 2013).

A new stock assessment was conducted in 2022. The swordfish operating models have been updated based on this new assessment.

The report for the 2022 stock assessment is published in [SCRS/2022/012](#). Details of the 2022 assessment are published in [SCRS/2022/124](#). The data used in the 2022 assessment, and the structure and assumptions of the assessment model are summarized in the sub-sections below.

3.1 Data

The assessment used landings data from 8 longline fleets (7 national fleets and ‘Other’ representing all catches not accounted for in the national fleets) (Table 1 and Figure 1).

There were 6 indices of abundance derived from catch-per-unit-effort (CPUE) data from the national fleets (Table 1 and Figure 1). Data from the Japan longline fleet was split into two phases: Early (1950 - 1993), and Late (1994 - 2020) due to changes in fishing practices in these periods. Similarly, the Chinese-Taipei longline fleet was split into two phases: Early (1968 - 1989) and Late (1997 - 2020). Five additional survey indices were developed from the age-specific (ages estimated from length data) CPUE from the Spain longline fleet (Table 1 and Figure 1).

Length composition data from 7 national longline fleets (Spain, U.S., Canada, Japan (Early, Late), Portugal, Chinese-Taipei (Early, Late), and Morocco) and the Canada/USA harpoon fleet (Figure 1). Mean weight data from two fleets (U.S. and Canada) were also used (Figure 1).

The catchability coefficient (q) for the CPUE indices for Canada, Morocco, and the Age-1, Age-2, and Age-4 Spain age-specific survey indices, was made a function of the Atlantic Multidecadal Oscillation (AMO) (see [SCRS/2022/124](#) for details).

The effective sample size (ESS) for the length composition data was established by adjusting ESS until unity was reached between modeled ESS and the Francis suggested sample size (See Anon., 2017 for details).

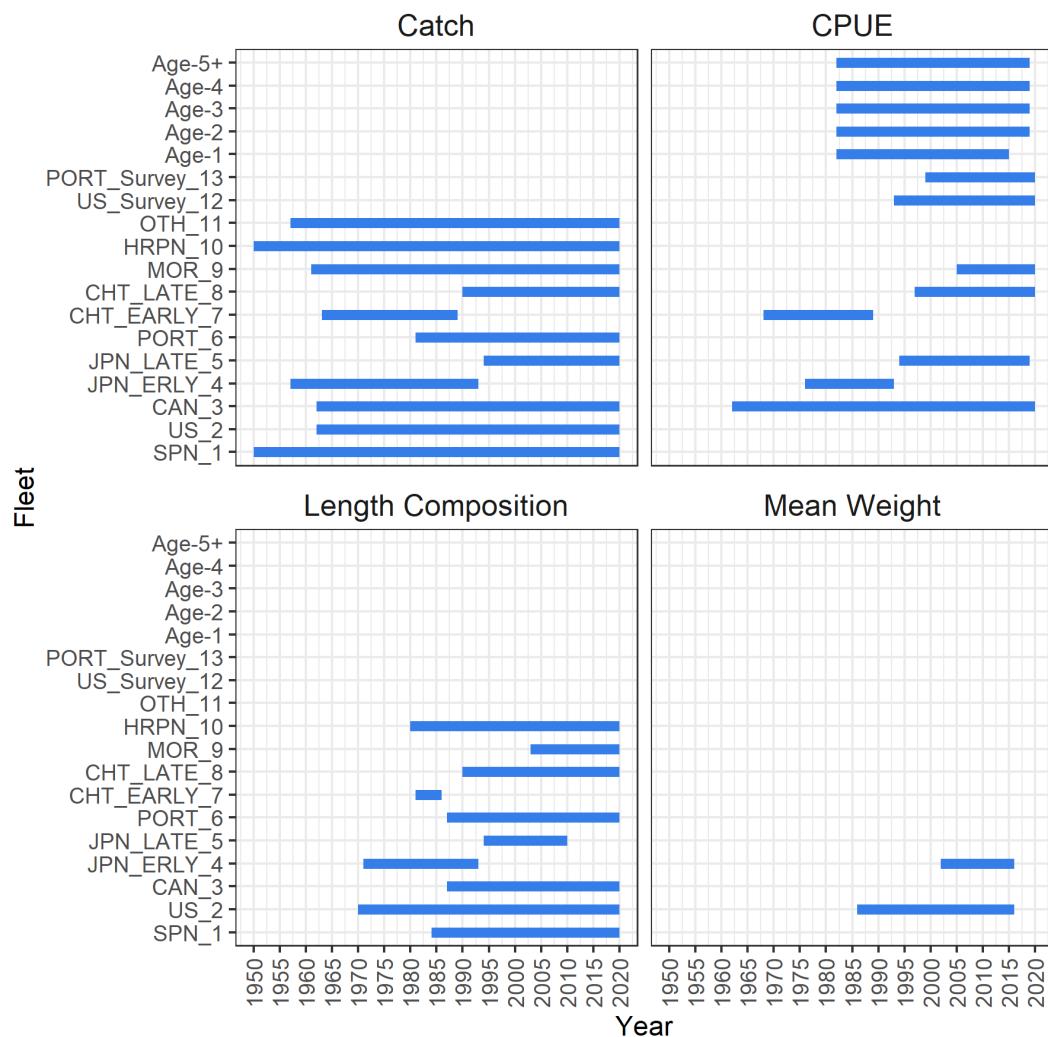


Figure 1: Summary of the data used in the 2022 stock assessment.

Table 1: Summary table of the fishing fleets and indices of abundances included the 2022 stock assessment of North Atlantic swordfish.

Name	Code
EU Spain longline (LL)	SPN 1
USA LL	US 2
Canada LL	CAN 3
Japan LL Early	JPN ERLY 4
Japan LL Late	JPN LATE 5
EU Portugal LL	PORT 6
Chinese Taipai LL Early	CHT EARLY 7
Chinese Taipai LL Late	CHT LATE 8
Morocco LL	MOR 9
Canada USA Harpoon	HRPN 10
Other LL by the other CPCs, and all other gears except HP	OTH 11
USA Survey	US Survey 12
EU Portugal Survey	PORT Survey 13
Age 1 Survey	Age 1
Age 2 Survey	Age 2
Age 3 Survey	Age 3
Age 4 Survey	Age 4
Age 5 plus Survey	Age 5 plus

3.2 Model Structure

The SS3 model used one season, one area, and two sexes.

3.3 Fixed Biological Parameters

Natural mortality for both male and female was fixed at 0.2 for all age classes. Maturity-at-age was knife-edge, with 50% at age-5 and 100% thereafter. Fecundity was proportional to body weight. The growth parameters for the 2022 assessment were fixed at same used in the 2017 assessment, which were developed during the 2017 ICCAT Swordfish Data Preparatory Meeting (See Anon., 2017 for details).

3.4 Selectivity

Selectivity was modeled as a function of length. Dome-shaped selectivity was allowed for five fleets: EU-Spain, USA, Japan, EU-Portugal, and Morocco. Asymptotic selectivity was assumed for Canada, Chinese-Taipei and Other. The age-specific survey CPUEs were modeled with a fixed age-based selectivity.

3.5 Retention and Discard Mortality

The 2022 assessment assumed a minimum legal length of 119 cm lower jaw fork length (LFFL) for all fleets from 1993 - 2020, and estimated the selectivity and retention curves from the available data. Discard mortality was either estimated from the observer data (USA and Canada) or fixed at values taken from the literature (see SCRS/2022/124).

3.6 Stock-Recruitment

Expected recruitment to age-0 was calculated from the total spawning stock biomass using the Beverton-Holt stock-recruit function. The standard error for the log-normally distributed recruitment deviations (σ_{R}) was fixed to 0.2. Steepness (h) was fixed at the value estimated in the 2017 assessment (0.88).

4 Overview of Operating Model Conditioning

The 2022 stock assessment was used as the Base Case model for developing the operating models (OMs) evaluated in the MSE.

4.1 OM Uncertainty Grid

In 2019, the Swordfish Species Group developed an OM uncertainty grid using a full factorial design with 6 axes of uncertainty:

1. Natural mortality (M) - three levels: 0.1, 0.2, 0.3
2. Recruitment variability (σ_R ; σ_R) - two levels: 0.2, 0.6
3. Steepness (h) - three levels: 0.6, 0.75, 0.90
4. CPUE Lambda - three levels: 0.05, 1, 20
5. Directional trend in catchability (llq) - two levels: TRUE, FALSE
6. Environmental covariate (env) - two levels: TRUE, FALSE

The directional increase in catchability was modeled by assuming a 1% average annual increase in catchability throughout the history of the fishery, and adjusting the CPUE indices accordingly.

The environmental covariate was included by modeling catchability as a function of the AMO, as was done in the 2017 and 2022 stock assessments.

This factorial design of the 6 axes of uncertainty, each with 2 or 3 levels, resulted in an uncertainty grid of 216 OMs.

4.2 Evaluation of the OM Grid

Previous analyses of the OM uncertainty grid based on the 2017 assessment revealed that the three levels of natural mortality and steepness had the largest impact on the predicted stock dynamics [Hordyk et al., 2021](#) and are therefore the most important axes of uncertainty in the OM grid.

The second level of recruitment variability ($\sigma_R = 0.6$) had a minor impact on the predicted stock dynamics [Hordyk et al., 2021](#), but did influence the relative performance of candidate management procedures [Hordyk, 2021](#). This second level is now treated as a robustness test (see below for more details).

The fourth axis of uncertainty was intended to evaluate the effect of alternative relative weightings of the length composition data and the indices of abundance. The three levels reflect a complete down-weighting of the indices of abundance (0.05; effectively only fitting the model to the length composition data), leaving the relative weighting of the two data sources unchanged from that used in the assessment (1), and up-weighting the indices of abundance so the model ignores the length composition data (20). This was done because of apparent conflicting signals between the length composition data and some of the indices of abundance, and the high computation demand of conducting the recommended iterative re-weighting procedure across all OMs in the grid (Francis, 2011). However, this iterative re-weighting procedure has now been conducted for the new operating models based on the 2022 assessment, and therefore this axis of uncertainty has been modified to two levels: 1) fit the assessment to both length and CPUE data and conduct the iterative re-weighting procedure, and 2) only fit the model to the CPUE data. This second level is now treated as a robustness test (see below for more details).

The fifth axis of uncertainty with the assumed 1% increase in catchability for the indices of abundance had a relatively minor influence on both the predicted stock dynamics [Hordyk et al., 2021](#) and the performance of candidate management procedures [Hordyk, 2021](#). Therefore it was decided that the default assumption for the operating models was to use the same indices as were used in the 2022 stock assessment; i.e., the indices were not adjusted for an assumed 1% annual increase in changeability. The second level of this axis,

where the indices are adjusted for an assumed increase in catchability, is now treated as a robustness test (see below for more details).

In both the 2017 and 2022 stock assessments, the catchability coefficient (q) for the CPUE indices for Canada, Japan, Portugal, Morocco, and the Spain age-specific survey indices, was made a function of the Atlantic Multidecadal Oscillation (AMO). Including this environmental covariate resulted in a better statistical fit to the data for these indices. The sixth axis of uncertainty examined the impact of not including this environmental covariate in the stock assessment. The analyses revealed that removing the environmental covariate from the assessment model had no detectable influence on either the predicted stock dynamics [Hordyk et al., 2021](#) or the performance of candidate management procedures [Hordyk, 2021](#). Therefore, the environmental covariate was included in all models in the OM grid. Further examination of the impact of changing environmental conditions on the performance of the candidate management procedures may be examined in additional robustness tests (see below for more details).

This same pattern in results was found when the analyses were re-conducted with the new OM grid based on the 2022 assessment.

5 Operating Models

The OMs, conditioned using SS3, were imported into MSE framework using the `SS2MOM` function in `MSEtool` (see [here](#) for details on the function).

Previously, the 2-sex SS3 models were imported into a combined single-sex operating model in the MSE framework. The MSE framework was updated in 2022, and the imported operating models now maintain the separate sexes in the same manner as the SS3 assessment. Since the fishery is managed with a single overall total allowable catch (TAC), the individual fishing fleets were aggregated together in the operating models.

The 2-sex operating model are available in the `SWOMSE` package as objects of class `MOM` (multi-stock operating model). For example, the 2022 stock assessment is available as the base case operating model named `MOM_000`:

```
library(SWOMSE)
MOM_000@Name
```



```
## [1] "M:0.2 sigmaR:0.2 steepness:0.88 Include CAL:TRUE llq:1 env:1 Class:Base Case"
```

The operating models have been classified into two categories: Reference OMs and Robustness OMs. Table 2 provides a general overview of the operating models includes in the MSE analysis. The following sub-sections describe the Reference and Robustness OMs in more detail.

5.1 Base Case

The Base Case operating model is the 2022 Stock Assessment. See the [Stock Assessment section](#) for more details.

5.2 Reference Operating Models

The Reference OMs are based on the Base Case Model and span the plausible range for the key uncertainties. The Candidate Management Procedures (CMPs) are tuned across the Reference OMs (see the [CMP Tuning section](#) for more details).

Based on the analyses described [above](#), a set of 9 operating models were identified as the Reference OMs. These operating models spanned the three levels of M (0.1, 0.2, and 0.3) and three levels h (0.69, 0.80, and 0.88).

The values for steepness in the Reference Set were determined at the 2023 Intersessional Meeting of the Swordfish Working Group. The lower and upper values were set at the 95% confidence intervals of a prior distribution for the steepness parameter used in the 2022 assessment.

Table 2: Summary of the Base Case, Reference, and Robustness Operating Models.

Class	Description
Base Case	The 2022 Stock Assessment, considered the Base Case OM
Reference	Nine models spanning the key plausible uncertainties in natural mortality (M) and steepness of the Beverton-Holt stock-recruit relationship (h). Same as Base Case with 3 levels of M (0.1, 0.2, 0.3) and 3 levels of h (0.69, 0.80, 0.88)
R0	Reference OM for the Robustness Tests. OM 5 of the Reference Set
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 and projection)
R2	Same as R2, but only for the historical period
R3a	Evaluate impact of cyclical pattern in recruitment deviations in projection period; a proxy for impact of climate change on stock productivity
R3b	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
R4	Evaluate impact of illegal, unreported, or unregulated catches

The center value was calculated by first converting steepness to the Goodyear compensation ratio (CR) with $CR = \frac{4h}{1-h}$. Using this equation, the CR for $h = 0.69$ and $h = 0.88$ are 8.90 and 29.33 respectively. A mid-point between these two CR values is calculated as:

$$CR_{mid} = e^{\frac{\sum_i \log CR_i}{2}} = 16.16, \text{ where the lower and upper values are } 16.16/1.815 \text{ and } 16.16 \times 1.815 \text{ respectively.}$$

Back-calculating to steepness with $h = \frac{CR}{CR+4}$ gives $h = 0.80$.

Initially the OMs with the lower steepness value ($h = 0.6$) was also considered to be in the Reference OMs. However, recent analysis and discussion with the MSE Technical Team has revealed that such a low steepness is unlikely for North Atlantic swordfish. Consequently, the OMs with $h = 0.6$ have been re-classified into a Robustness Set.

The Reference OMs had the following assumptions:

1. Following the assessment, $\sigma_R = 0.2$
2. The Francis iterative re-weighting procedure was conducted on each operating model to find the appropriate weighting between the length composition data and the indices of abundance
3. The standardized indices of abundance were used (i.e., no assumed 1% increase in catchability)

5.3 R0. Reference Model

The robustness tests are conducted for a single operating model from center of the Reference Set ($M=0.2$, $h=0.8$). R0 describes the performance of the CMPs for this single operating model, and is used for evaluating the results of the other robustness tests.

5.4 R1. Hyperstable Indices in Historical and Projection Periods

The R1 Robustness test evaluates the impact of a consistent increase in catchability that is not accounted for in the standardization of the indices, resulting in a hyper-stable index.

This test is implemented by assuming that the indices used in the OM conditioning were ‘corrected’ for an average 1% annual increase in catchability. This was done by decreasing the observed values of the indices by 1% per year. Figure 2 shows the estimated F/FMSY and SB/SBMSY for the Reference model (R0) and the this robustness test (R1).

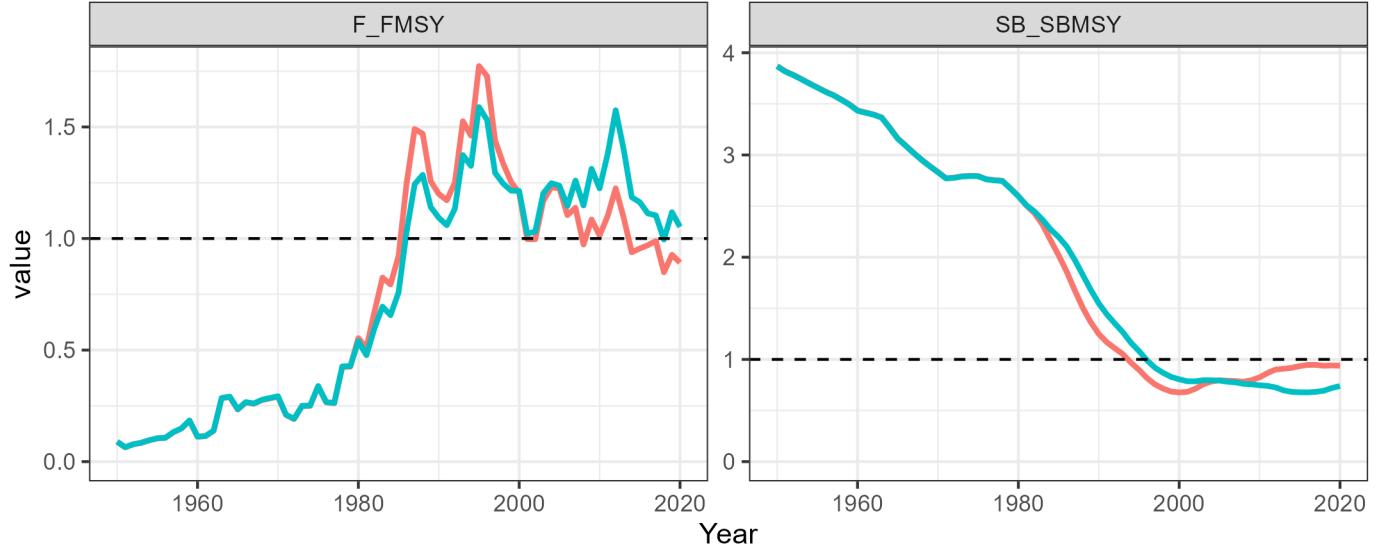


Figure 2: The fishing mortality and spawning biomass relative to MSY reference points for R0 (reference OM) and R1, the robustness test that assumed the indices were corrected for an assumed 1 percent annual increase in catchability.

The MPs were provided with the observed historical indices (rather than the ‘corrected’ indices). The unaccounted 1% annual increase in catchability was assumed to continue in the projection period. This was done by assuming 1% annual positive observation error for the indices in the projection period (Figure 3).

5.5 R2. Hyperstable Indices in Historical Period

The R2 Robustness test is identical to R1 (Figure 2), except that the indices in the projection period are assumed to be proportional to the population; i.e., the 1% increase in catchability that was not accounted for in the standardization of the historical indices no longer occurs in the projection period.

5.6 R3a and R3b. Effect of Climate Change on Recruitment Deviations in Projection Period

Climate change is generally predicted to increase temperatures throughout the Atlantic Ocean, which could have significant effects on swordfish recruitment. The Atlantic Multidecadal Oscillation (AMO) is an indicator of long-duration changes in the sea surface temperature of the North Atlantic Ocean. Thus, it could serve as the basis for possible future patterns in recruitment. This is not postulating that the AMO is in fact driving swordfish recruitment. Rather it is using the 32-year time series as a representative of actual oceanographic conditions and one possible pattern of a cyclical productivity that we wish to test the robustness of a CMP. In essence, these robustness tests are asking whether the CMP is robust to an extended period (~15 years) of negative recruitment deviations without a change in the underlying productivity of the stock.

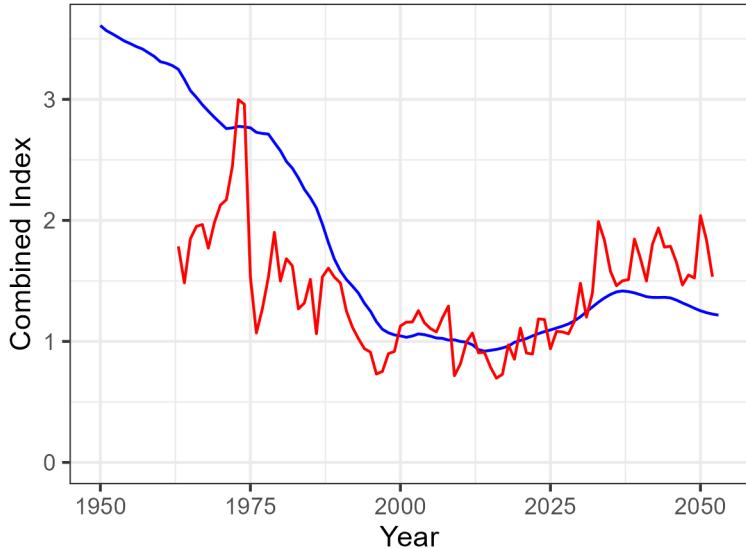


Figure 3: An example of the standardized biomass (blue) and the observed combined index (red) for R1, where an assumed 1 percent annual increase in catchability has not been accounted for in the standardization of the index.

The entire AMO long time series (1856-present) was considered and a 32-year period (1980-2011) was selected to represent a time series that contained an extended period of consecutive negative and consecutive positive deviations. The AMO values multiplied by an implied annual inflation factor of 2.0 to simulate the hypothesis that climate change could result in larger recruitment deviations in the future than in the past.

Two versions of this approach were developed. R3a uses the AMO values directly, with a period of lower than average recruitment for the first fifteen years, followed by a period of higher than average deviations (Figure 4 center panel). R3b considers the impact of just a period of negative recruitment deviations for the first fifteen years (Figure 4 right panel).

In both cases (R3a and R3b), the climate-induced pattern in recruitment deviations is multiplied by the recruitment deviations from the reference model (R0; Figure 5).

5.7 R4. TAC Implementation Error

The R4 Robustness test evaluates the impact of a 10% implementation error in the TAC. The catches are assumed to be 10% higher than the TAC, but these catches are assumed to be unreported (i.e., the observed catches provided to the CMPs is equal to the TAC and ~90% of the actual landings).

6 OM Validation

6.1 OM Summary Report

A [Summary Report](#) summarizes the diagnostic checks, the calculated biological reference points, and the estimated stock status relative to those reference points, across the Reference and Robustness operating models.

6.2 OM Diagnostic Reports

Individual diagnostic reports with objective function values and plots of model fits and patterns in residuals are available for each of the Reference and Robustness OMIs (and the 2022 assessment) on the North Atlantic

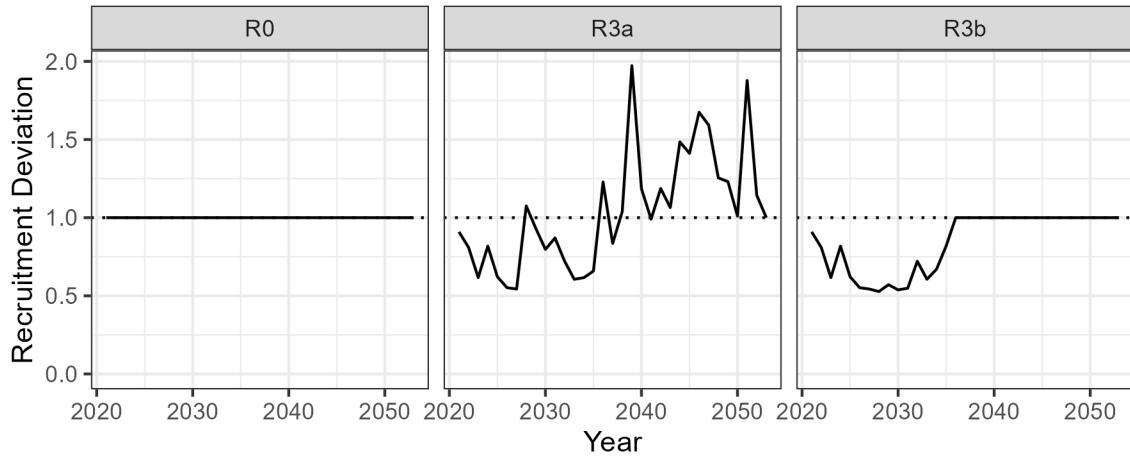


Figure 4: The cyclical (R3a) and initial negative pattern (R3b) in recruitment deviations for the robustness tests that investigate the impact of climate change on the performance of the CMPs. R0 is the reference model where no additional recruitment variability is added.

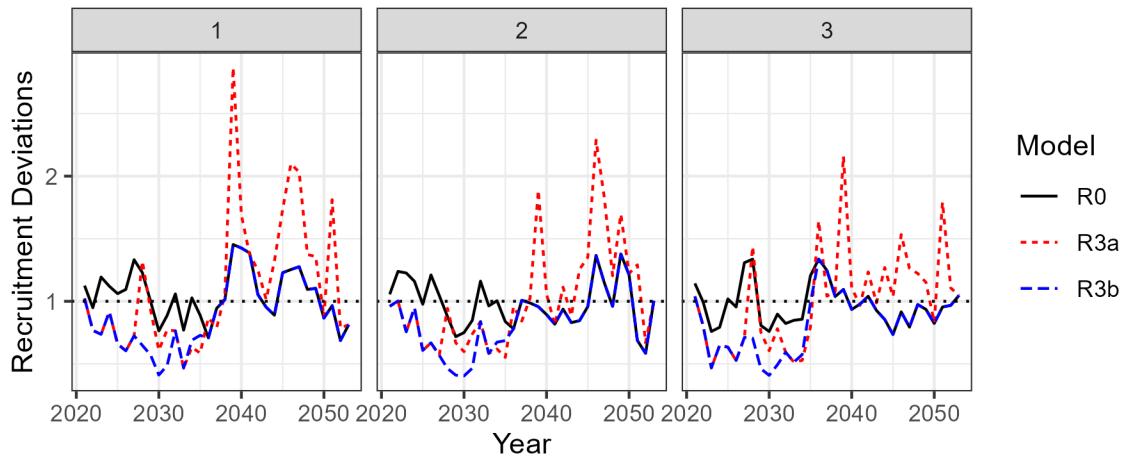


Figure 5: An example of the recruitment deviations from three simulations. R0 is the reference model (OM 5 from the Reference Set). R3a superimposes a cyclical pattern in recruitment deviations over those from R0. R3b superimposes negative recruitment deviations over those from R0 for the first fifteen years, before reverting back to the recruitment pattern used in R0.

Swordfish MSE [homepage](#).

7 Historical Spool-Up Period

The historical spool-up period was generated by importing the output from the SS3 assessment models and running the MSE simulation framework to recreate the historical fishery dynamics. No additional uncertainty was added to the historical simulations; that is, all 50 simulations were identical (e.g., same recruitment deviations) for the historical spool-up period.

The operating models have been conditioned on data up to and including 2020 (see the [Data section](#)), and therefore the projection period for the MSE framework starts in the following year, i.e., 2021. Currently, model uses assumed catches for the first three years (2021 - 2023) and then implements the management procedures in the following year (i.e., 2024). See the [Future Catches](#) section for more information.

7.1 Historical Data

The observed fishery data was imported into the MSE framework and made available to the CMPs. Three sources of data available to the CMPs: 1) catch data, 2) a primary index of abundance, and 3) additional individual indices.

7.1.1 Catch Data

The catch data used in the OM conditioning are made available in the simulated data that is provided to the CMPs (Figure 6).

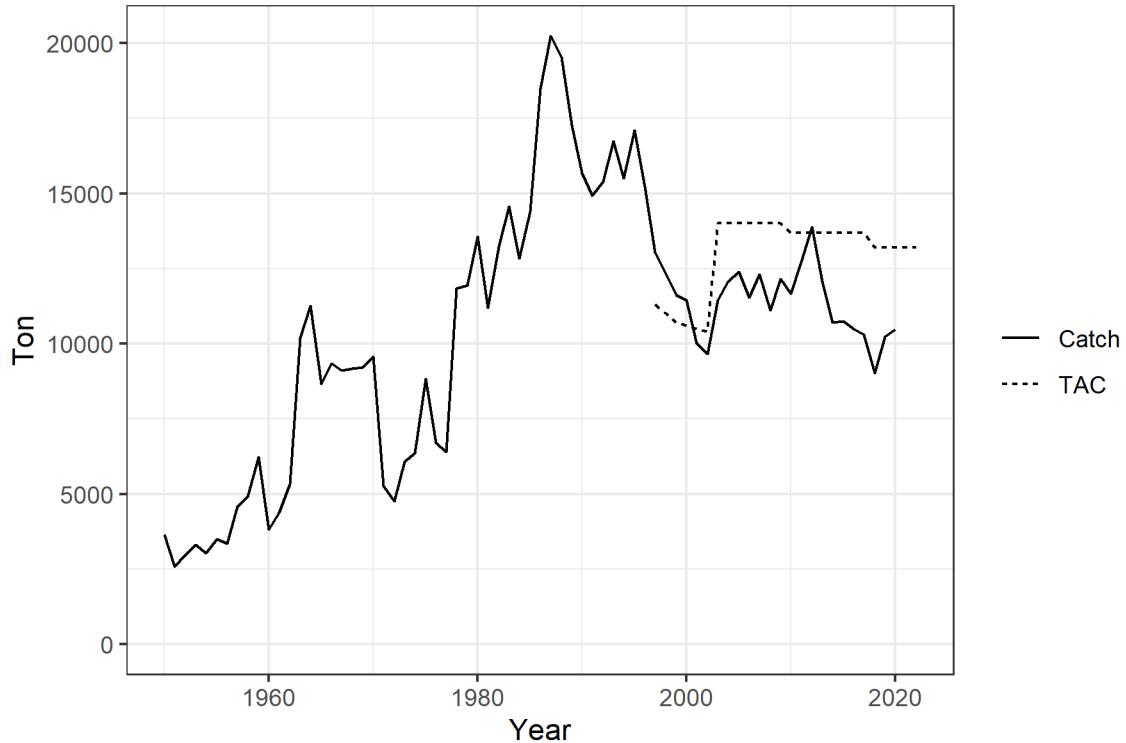


Figure 6: The historical total catch and total allowable catch (TAC) for the North Atlantic swordfish.

7.1.2 Indices of Abundance

7.1.2.1 Primary Index At the **2020 Swordfish MSE technical meeting (4 – 5 June 2020)**, the Group chose to use the Combined Index (Ortiz et al., 2017) as the primary index for the development of CMPs. This index was updated for the 2022 assessment. The Combined Index was imported into the SWOMSE framework and made available to the CMPs (Figure 7).



Figure 7: The Combined Index for the North Atlantic swordfish fishery.

7.1.2.2 Other Indices The additional individual fishery-dependent and survey indices are also imported into the MSE framework and made available to the CMPs (Figure 8).

8 Closed-Loop Simulation Testing

8.1 Simulation Specifications

The current simulation specifications for the MSE are:

1. Management interval: 3 (i.e., TAC updated every 3rd year)
2. Number of projection years: 33 (2021 – 2053)
3. Number of simulations per OM: 50

8.2 Assumptions

8.2.1 Recruitment

The stock-recruitment relationship in the projection years was modeled using the Beverton-Holt function, with steepness fixed at the value assumed in the OM conditioning.

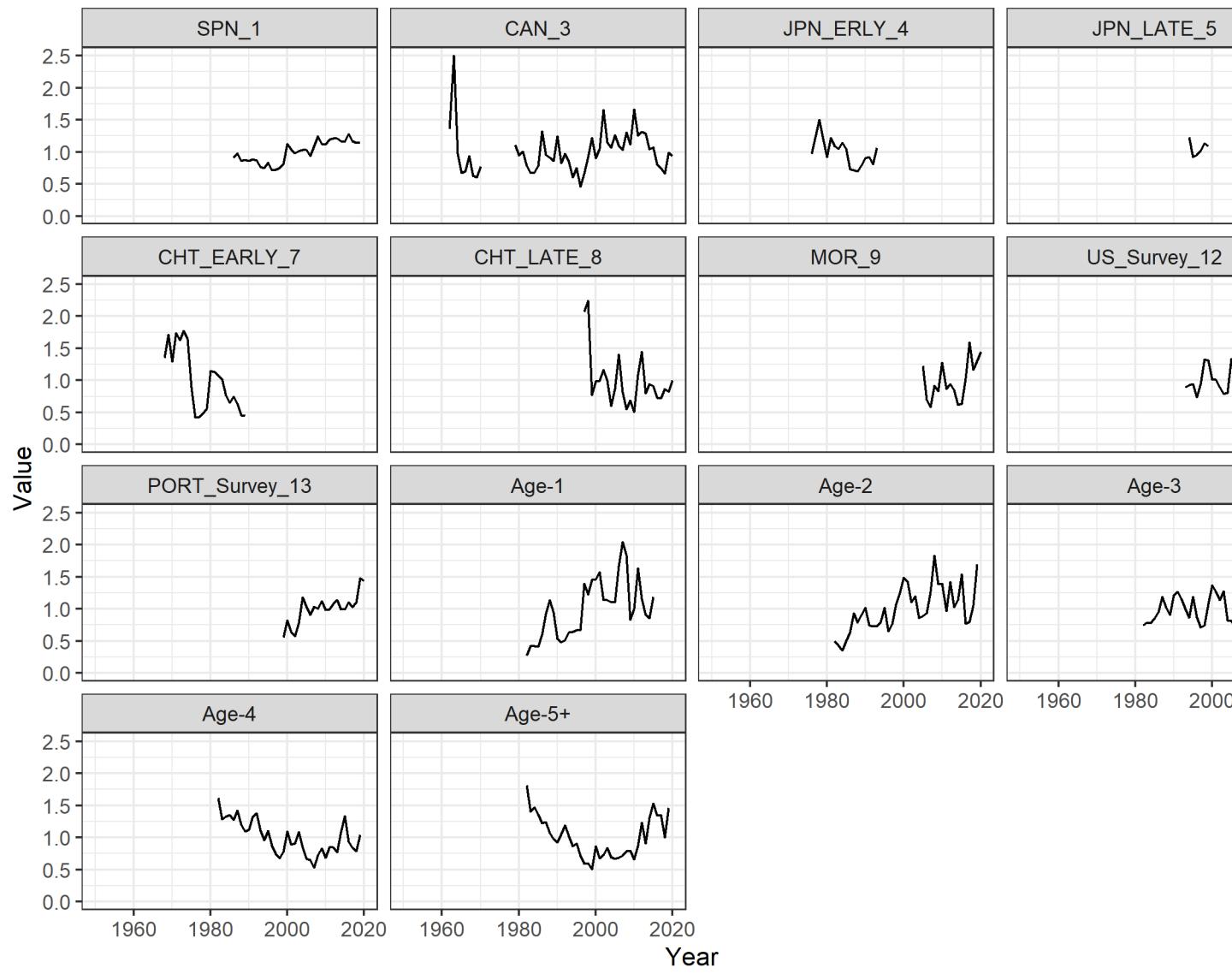


Figure 8: The additional individual fishery-dependent and survey indices for the North Atlantic swordfish fishery.

Table 3: The assumed landed catch (ton) for the first 3 projection years.

Year	Catch	Details
2021	9729	Reported Catch
2022	10770	Assumed Catch
2023	10770	Assumed Catch

Recruitment deviations for the projection period were generated assuming a log-normal distribution with mean μ_R and variance σ_R^2 , calculated as:

$$\mu_R = -0.5\sigma_R^2 \left(1 - \frac{AC}{\sqrt{(1 - AC^2)}} \right)$$

where AC is the lag-1 autocorrelation factor calculated from the historical recruitment deviations and σ_R^2 is the variance of the recruitment deviations specified in the OM (i.e., sigmaR^2 defined [above](#)).

The recruitment deviations for the projection period were generated independently for each simulation.

8.2.2 Life-History

Biological parameters such as mean weight-at-age and maturity-at-age were constant for all years (historical and projection).

8.2.3 Selectivity

Currently, the selectivity-at-length for the projection years is fixed to the estimated selectivity pattern from the last historical year for each OM. The reviewer of the SWOMSE R code highlighted this as a possible uncertainty and recommended that the Group consider testing alternative assumptions for the selectivity pattern in the projection years (Anon., 2021).

The Working Group will discuss this assumption in more detail and decide if it wishes to run scenarios with alternative assumptions for the selectivity pattern in the projection years.

8.2.4 Catchability

No time-varying catchability is assumed for the projection years.

8.2.5 Future Catches

The total allowable catch (TAC) recommended by the CMPs is assumed to be implemented without error. That is, annual catches are equal to the TAC recommendations (provided sufficient vulnerable biomass was available). The Group determined that this assumption was appropriate based on the fact that historically the total landed catch has been below the total allowable catch (TAC) (see Figure 6).

A Robustness Test (R4) has been designed to evaluate the consequences of violating this assumption.

The operating models also assume that the TACs recommended by the CMPs apply to the landed catch and do not include discards; that is, the actual removals will be higher than the TACs if there is discard mortality on the sub-legal fish.

The CMPs are first implemented in 2024, and every 3rd year after that, with the TAC held constant in the interim years. Since the MSE projection period begins in 2021 (the year after the conditioning), the landed catches for the first few projection years until the CMPS are implemented (2021, 2022, and 2023) are fixed and not set by the CMPs. The catch for 2021 is set at the reported level, and the catches for 2022 and 2023 are set to the mean from previous 10 years (2012 – 2021) (Table 3).

8.3 Generation of Future Data

8.3.1 Indices

The primary index of abundance (Combined Index) was generated in the projection years by adding observation error to the projected stock biomass, and standardizing to be on the same scale as the historical observed Combined Index.

The additional individual indices were generated in the projection years in the same manner, using the simulated stock biomass (or abundance in some cases).

The observation error was generated by calculating the statistical properties of the residuals from the fit of each respective index to the simulated stock (biomass or abundance) during the historical period. The procedure for calculating the residuals for the projection years is described in detail in the [openMSE documentation](#).

It was decided at the 2023 Intersessional Meeting of the Swordfish Working Group to calculate the statistical properties of the observation error for the Combined Index for the years 1999 – 2020. This period was chosen as it includes only the years where the data used to generate the index included the majority of the fleets.

Following the assumptions of the stock assessment, it is assumed that all indices are proportional to the respective stock biomass or abundance (i.e., it is assumed that the indices are not hyper-stable or hyper-deplete). The only exception to this is **Robustness Set 5** where the indices are assumed to have a 1% annual average increase in catchability.

8.3.2 Catches

Following the assumption of the stock assessment, the observed catches in the projection years were generated with no observation error. That is, the observed catches in the MSE framework are equal to the actual simulated landed catches (provided there is sufficient biomass in the simulated population to support the prescribed TAC). The model does not include a bio-economic component and inter-annual change in fishing effort is unconstrained.

8.3.3 Data Lags

In the MSE framework, the CMPs are always provided with the simulated fishery data up to the previous year. That is, if a CMP is implemented in year t , the catch and indices of abundance data provided to the CMP are up to and including year $t - 1$.

Data lags are then handled in the CMP code, where the default assumption is that the catch and index data are lagged by 2 years from the year where the TAC will be implemented; e.g., for setting the TAC for 2024, the data is up to and including 2022. See the CMP (see the [CMP Development Guide](#) for more details).

8.4 Management Interval

At the 2021 Intersessional meeting, the Group proposed a 3-year MP advice interval (see Table 5 in the [Report of the 2021 ICCAT Intersessional Meeting of the Swordfish Species Group](#)).

The operating models have been conditioned on data up to 2020. As described [above](#), the catches in 2021 – 2023 are fixed to assumed values and the CMPs are first applied in 2024.

The 3-yr management interval is applied within the CMP code (see the [CMP Development Guide](#) for more details).

Robustness Set 9 (R9) evaluates the impact of alternative management intervals.

9 Performance Metrics

23 performance metrics (PMs) have been developed for the North Atlantic Swordfish MSE (Table 4). These PMs are grouped into four families:

1. **Status:** the probability the stock is in the green quadrant of the Kobe matrix
2. **Safety:** the probability of the stock not falling below the biological limit reference point
3. **Yield:** the TAC in the projection years (ton)
4. **Stability:** the variation in the TAC between management cycles

9.1 Summary of Performance Metrics

The performance metrics are calculated over all the simulations (50) in each operating model. For example, *PGK_short* is calculated as the proportion of the 50 simulations where the spawning biomass (*SB*) is above *SB_{MSY}* and fishing mortality (*F*) is below *F_{MSY}*.

The performance metrics can also be calculated over several operating models, for example the entire **Reference Set**. This is done by first combining the MSE results for each OM (e.g., the 9 Reference OMs) into a single set of results (in this example, resulting in 450 simulations (9 OMs multiplied by 50 simulations per OM)), and then calculating the performance metrics over all results.

The Minimum Acceptable Performance values listed in Table 4 will be used as alternative values to determine if candidate management procedures meet the minimum performance criteria.

More details on the individual performance metrics are in the sections below.

9.2 Peformance Metrics Examples

The following sections show figures and statistics for examples from each family of Performance Metrics. These results include two example candidate management procedures (MP1, MP2) (see **Candidate Management Procedures**), with a single operating model that includes 10 stochastic simulations.

9.2.1 Summary Table

Table 5 shows the summary performance statistics for the example CMPs. More details for each individual performance metric are in the sub-sections below.

9.3 Status

9.3.1 PGK_short

PGK_short calculates the probability of being in Green Zone of Kobe Space (*SB*>*SB_{MSY}* & *F*<*F_{MSY}*) in years 1-10 (2024-2033). This is calculated as the proportion of the total number of simulations that are in the Green Zone of Kobe Space in those years.

Figure 9 shows an example of two CMPs (MP1, MP2) for an operating model with 10 simulations. The values for *PGK_short* in this example are: 0.06, 0.31.

9.3.2 PGK_med

PGK_med calculates the probability of being in Green Zone of Kobe Space (*SB*>*SB_{MSY}* & *F*<*F_{MSY}*) over years 11–20 (2034–2043). This is calculated as the proportion of the total points (number of simulations x number of years) that are in the Green Zone of Kobe Space.

Figure 10 shows an example of two CMPs (MP1, MP2) for an operating model with 10 simulations. The values for *PGK_med* in this example are: 0.15, 0.58.

Table 4: Summary of the North Atlantic Swordfish Performance Metrics.

Family	Name	Description	Minimum Acceptable Values
Status	PGK_short	Probability of being in Green Zone of Kobe Space ($SB > SBMSY \& F < FMSY$) in years 1-10 (2024-2033)	51, 60, 70
	PGK_med	Probability of being in Green Zone of Kobe Space ($SB > SBMSY \& F < FMSY$) in years 11-20 (2034-2043)	51, 60, 70
	PGK_long	Probability of being in Green Zone of Kobe Space ($SB > SBMSY \& F < FMSY$) in years 21-30 (2044-2053)	51, 60, 70
	PGK	Probability of being in Green Zone of Kobe Space ($SB > SBMSY \& F < FMSY$) over all years (2024-2053)	51, 60, 70
	PGK_30	Probability of being in Green Zone of Kobe Space ($SB > SBMSY \& F < FMSY$) in year 30 (2053)	51, 60, 70
	POF	Probability of Overfishing ($F > FMSY$) over all years (2024-2053)	
	PNOF	Probability of Not Overfishing ($F < FMSY$) over all years (2024-2053)	
Safety	LRP_short	Probability of breaching the limit reference point ($SB < 0.4SBMSY$) in any of the first 10 years (2024-2033)	5, 10, 15
	LRP_med	Probability of breaching the limit reference point ($SB < 0.4SBMSY$) in any of years 11-20 (2034-2043)	5, 10, 15
	LRP_long	Probability of breaching the limit reference point ($SB < 0.4SBMSY$) in any of years 21-30 (2044-2053)	5, 10, 15
	LRP	Probability of breaching the limit reference point ($SB < 0.4SBMSY$) in any year (2024-2053)	5, 10, 15
Yield	TAC1	TAC (t) in the first implementation year (2024)	
	AvTAC_short	Median TAC (t) over years 1-10 (2024-2033)	
	AvTAC_med	Median TAC (t) over years 11-20 (2034-2043)	
	AvTAC_long	Median TAC (t) over years 21-30 (2044-2053)	
Stability	VarC	Mean variation in TAC (%) between management cycles over all years and simulations	No minimum value and 25 ¹⁸

Table 5: Summary of Performance Metric values for some example CMPs.

Family	Name	Description	MP1	MP2
Status	PGK_short	Probability of being in Green Zone of Kobe Space ($SB > SBMSY \& F < FMSY$) in years 1-10 (2024-2033)	0.06	0.31
	PGK_med	Probability of being in Green Zone of Kobe Space ($SB > SBMSY \& F < FMSY$) in years 11-20 (2034-2043)	0.15	0.58
	PGK_long	Probability of being in Green Zone of Kobe Space ($SB > SBMSY \& F < FMSY$) in years 21-30 (2044-2053)	0.16	0.66
	PGK	Probability of being in Green Zone of Kobe Space ($SB > SBMSY \& F < FMSY$) over all years (2024-2053)	0.12	0.52
	PGK_30	Probability of being in Green Zone of Kobe Space ($SB > SBMSY \& F < FMSY$) in year 30 (2053)	0.10	0.80
	POF	Probability of Overfishing ($F > FMSY$) over all years (2024-2053)	0.65	0.27
Safety	LRP_short	Probability of breaching the limit reference point ($SB < 0.4SBMSY$) in any of the first 10 years (2024-2033)	0.00	0.00
	LRP_med	Probability of breaching the limit reference point ($SB < 0.4SBMSY$) in any of years 11-20 (2034-2043)	0.30	0.10
	LRP_long	Probability of breaching the limit reference point ($SB < 0.4SBMSY$) in any of years 21-30 (2044-2053)	0.30	0.10
	LRP	Probability of breaching the limit reference point ($SB < 0.4SBMSY$) in any year (2024-2053)	0.30	0.10

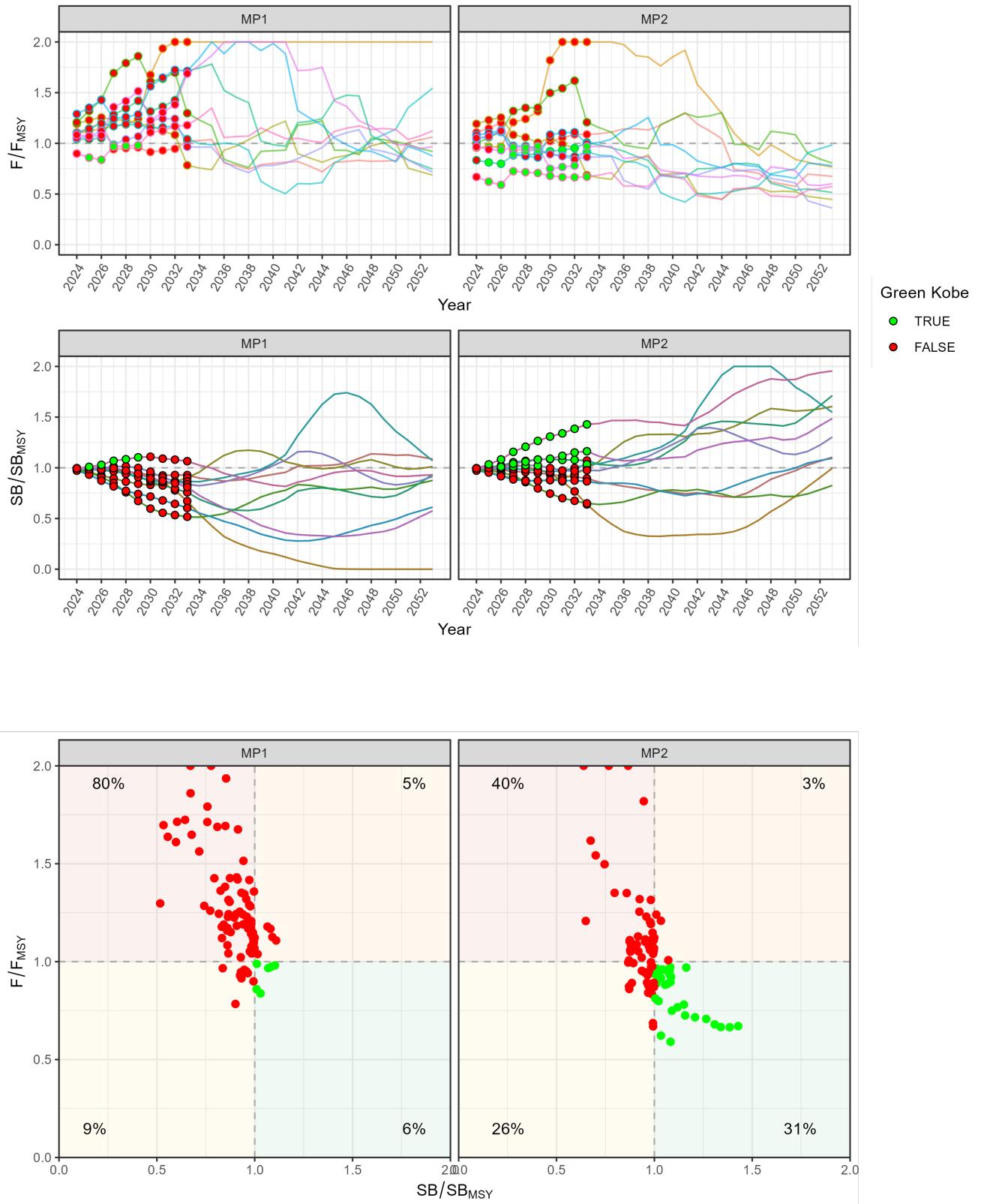


Figure 9: An example of time-series plots of F/F_{MSY} and SB/SB_{MSY} for two CMPs. This OM has 10 simulations, shown as the colored lines. Simulations where $SB > SB_{MSY}$ and $F < F_{MSY}$ in 2033 are indicated with a green point, and those outside of the green space in the Kobe plot are indicated in red. The Kobe plot on the bottom shows these points in the Kobe space. The numbers in the corners indicate the percent of points in each area of the Kobe space.

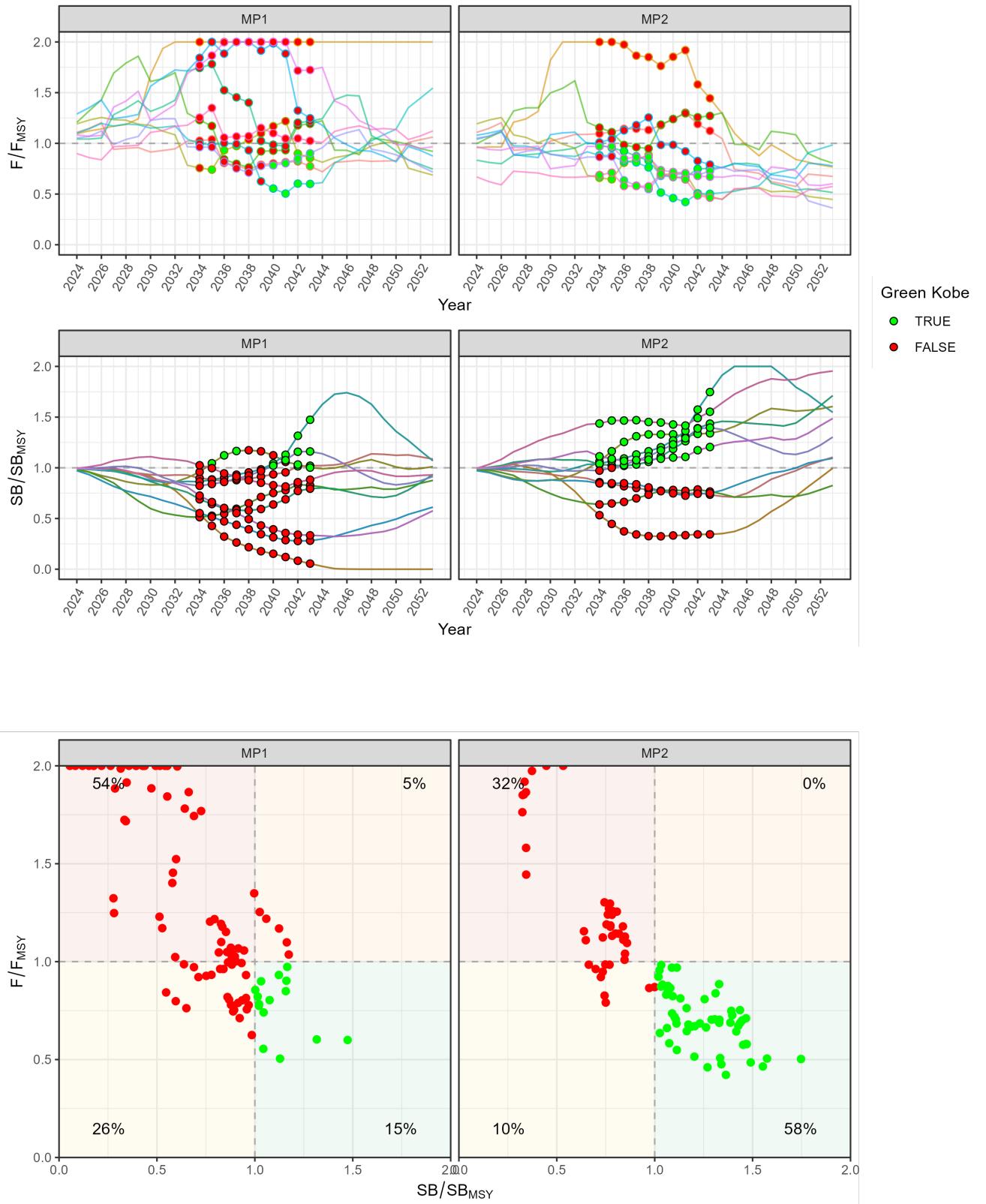


Figure 10: An example of time-series plots of F/F_{MSY} and SB/SB_{MSY} for two CMPs. This OM has 10 simulations, shown as the colored lines. Simulations where $SB > SB_{MSY}$ and $F < F_{MSY}$ in years 2029–2033 are indicated with a green point, and those outside of the green space in the Kobe plot are indicated in red. The Kobe plot on the bottom shows these points in the Kobe space. The numbers in the corners indicate the percent of points in each area of the Kobe space.

9.3.3 PGK_long

PGK_long calculates the probability of being in Green Zone of Kobe Space ($SB > SBMSY \& F < FMSY$) over years 21–30 (2044–2053). This is calculated as the proportion of the total points (number of simulations x number of years) that are in the Green Zone of Kobe Space.

Figure 11 shows an example of two CMPs (MP1, MP2) for an operating model with 10 simulations. The values for PGK_long in this example are: 0.16, 0.66.

9.3.4 PGK

PGK calculates the probability of being in Green Zone of Kobe Space ($SB > SBMSY \& F < FMSY$) over all years (2024–2053). This is calculated as the proportion of the total points (number of simulations x number of years) that are in the Green Zone of Kobe Space.

Figure 12 shows an example of two CMPs (MP1, MP2) for an operating model with 10 simulations. The values for PGK in this example are: 0.12, 0.52.

9.3.5 PGK_30

PGK_30 calculates the probability of being in Green Zone of Kobe Space ($SB > SBMSY \& F < FMSY$) in year (2053). This is calculated as the proportion of the total number of simulations that are in the Green Zone of Kobe Space.

Figure 13 shows an example of two CMPs (MP1, MP2) for an operating model with 10 simulations. The values for PGK_30 in this example are: 0.1, 0.8.

9.3.6 POF and PNOF

POF calculates the probability of over-fishing ($F > FMSY$) over all years (2024–2053). This is calculated as the proportion of the total number of points (number of simulations x number of years) where F is greater than FMSY.

Figure 14 shows an example of two CMPs (MP1, MP2) for an operating model with 10 simulations. The values for POF in this example are: 0.65, 0.27.

PNOF reports the probability of not overfishing, calculated as $1 - POF$, i.e. 0.35, 0.73.

9.4 Safety

9.4.1 LRP_short

LRP_short calculates the probability of breaching the limit reference point (LRP; $SB < 0.4SBMSY$) in any of the first 10 years (2024–2033).

This is calculated as the proportion of simulations where $SB < 0.4SBMSY$ in any of the first 10 years.

Figure 15 shows an example of two CMPs (MP1, MP2) for an operating model with 10 simulations. The values for LRP_short in this example are: 0, 0.

9.4.2 LRP_med

LRP_med calculates the probability of breaching the limit reference point (LRP; $SB < 0.4SBMSY$) in any of years 11–20 (2034–2043).

This is calculated as the proportion of simulations where $SB < 0.4SBMSY$ in any of years 11–20.

Figure 16 shows an example of two CMPs (MP1, MP2) for an operating model with 10 simulations. The values for LRP_med in this example are: 0.3, 0.1.

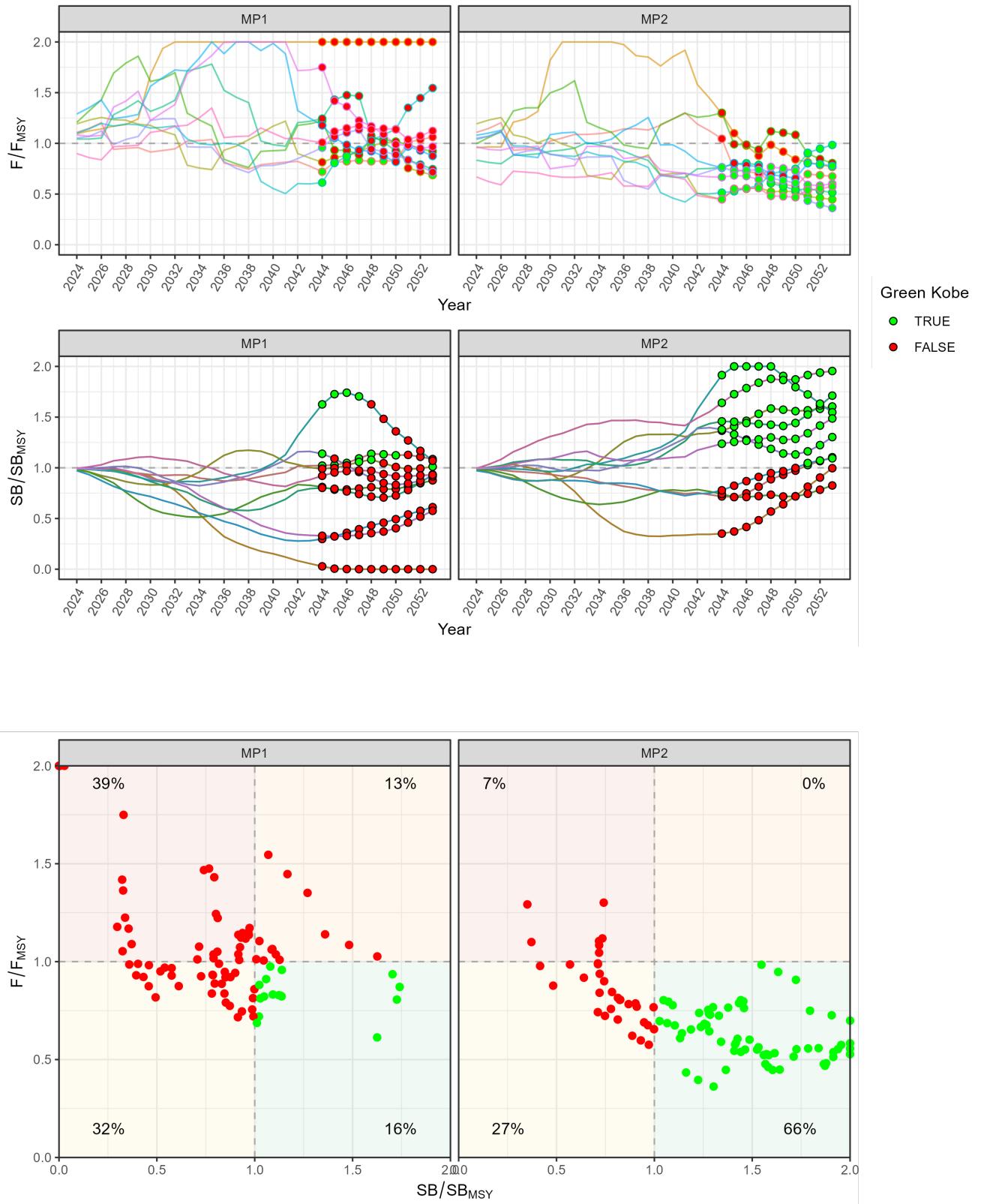


Figure 11: An example of time-series plots of F/F_{MSY} and SB/SB_{MSY} for two CMPs. This OM has 10 simulations, shown as the colored lines. Simulations where $SB > SB_{MSY}$ and $F < F_{MSY}$ in years 2034–2053 are indicated with a green point, and those outside of the green space in the Kobe plot are indicated in red. The Kobe plot on the bottom shows these points in the Kobe space. The numbers in the corners indicate the percent of points in each area of the Kobe space.

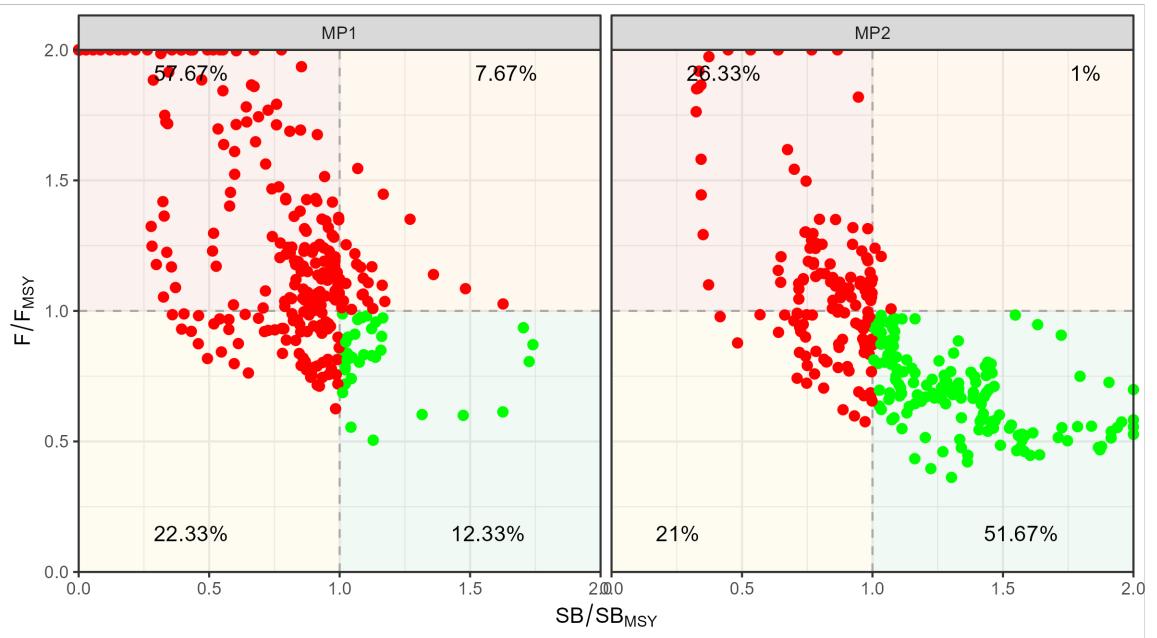
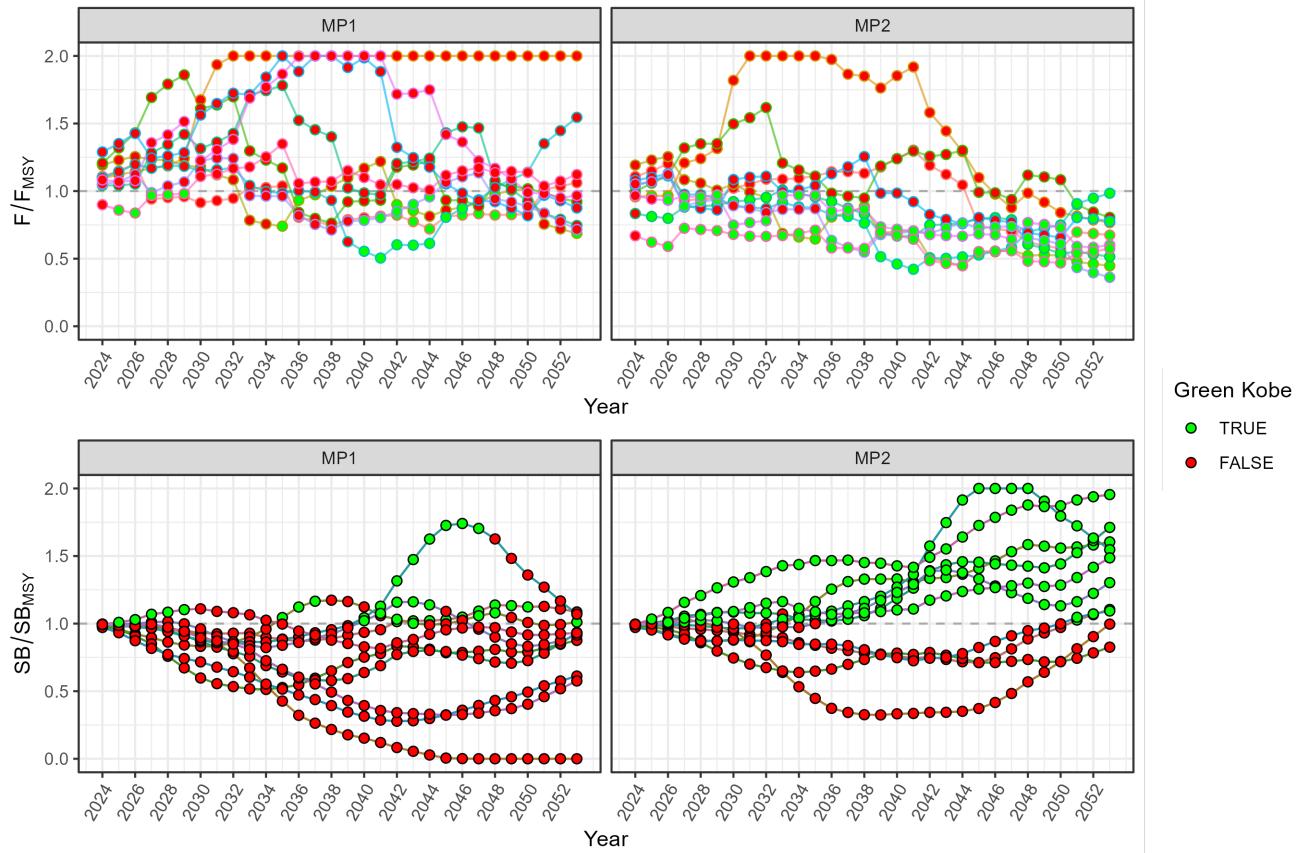


Figure 12: An example of time-series plots of F/F_{MSY} and SB/SB_{MSY} for two CMPs. This OM has 10 simulations, shown as the colored lines. Simulations where $SB > SB_{\text{MSY}}$ and $F < F_{\text{MSY}}$ are indicated with a green point, and those outside of the green space in the Kobe plot are indicated in red. The Kobe plot on the bottom shows these points in the Kobe space. The numbers in the corners indicate the percent of points in each area of the Kobe space.

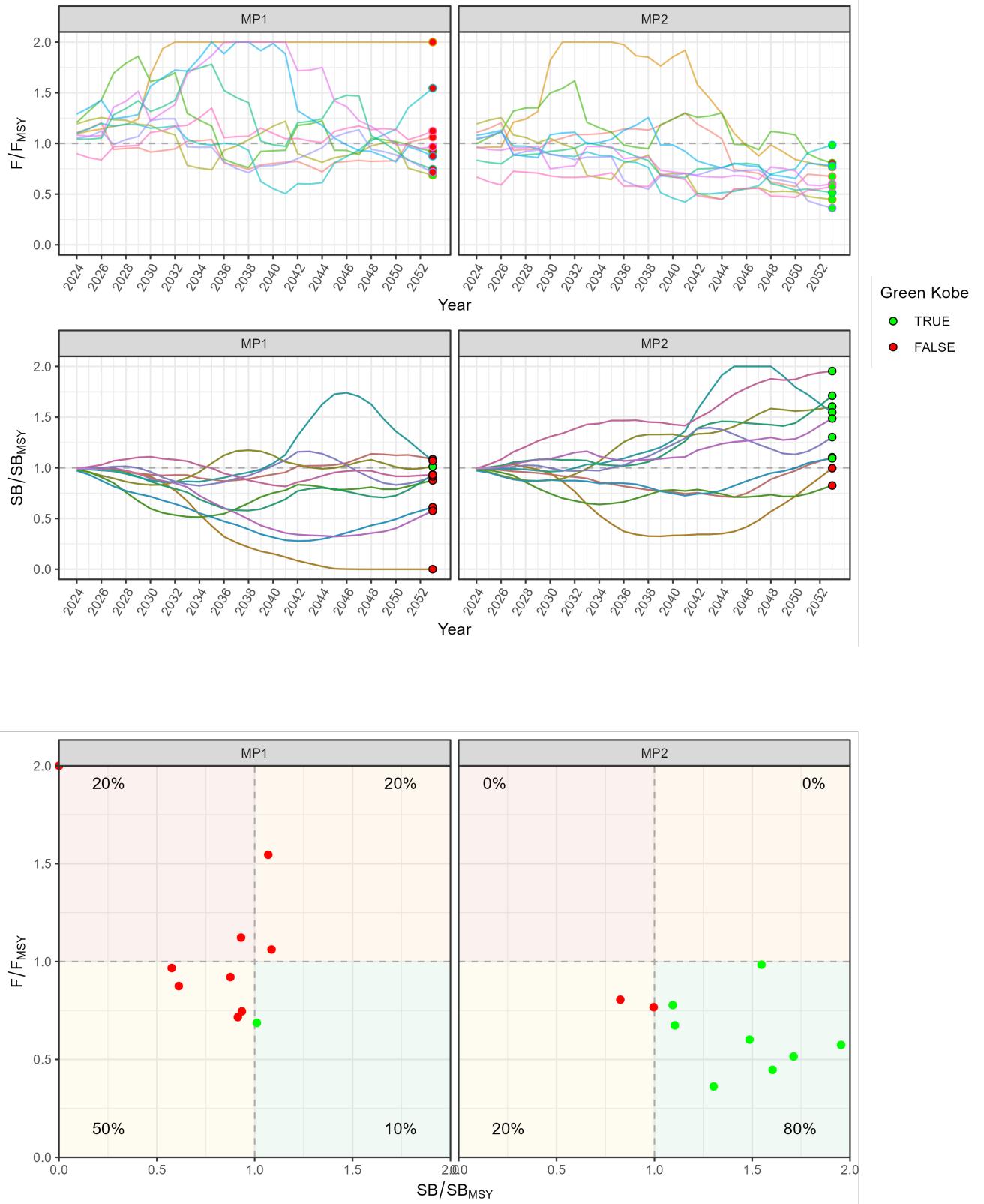


Figure 13: An example of time-series plots of F/F_{MSY} and SB/SB_{MSY} for two CMPs. This OM has 10 simulations, shown as the colored lines. Simulations where $SB > SB_{MSY}$ and $F < F_{MSY}$ in 2053 are indicated with a green point, and those outside of the green space in the Kobe plot are indicated in red. The Kobe plot on the bottom shows these points in the Kobe space. The numbers in the corners indicate the percent of points in each area of the Kobe space.

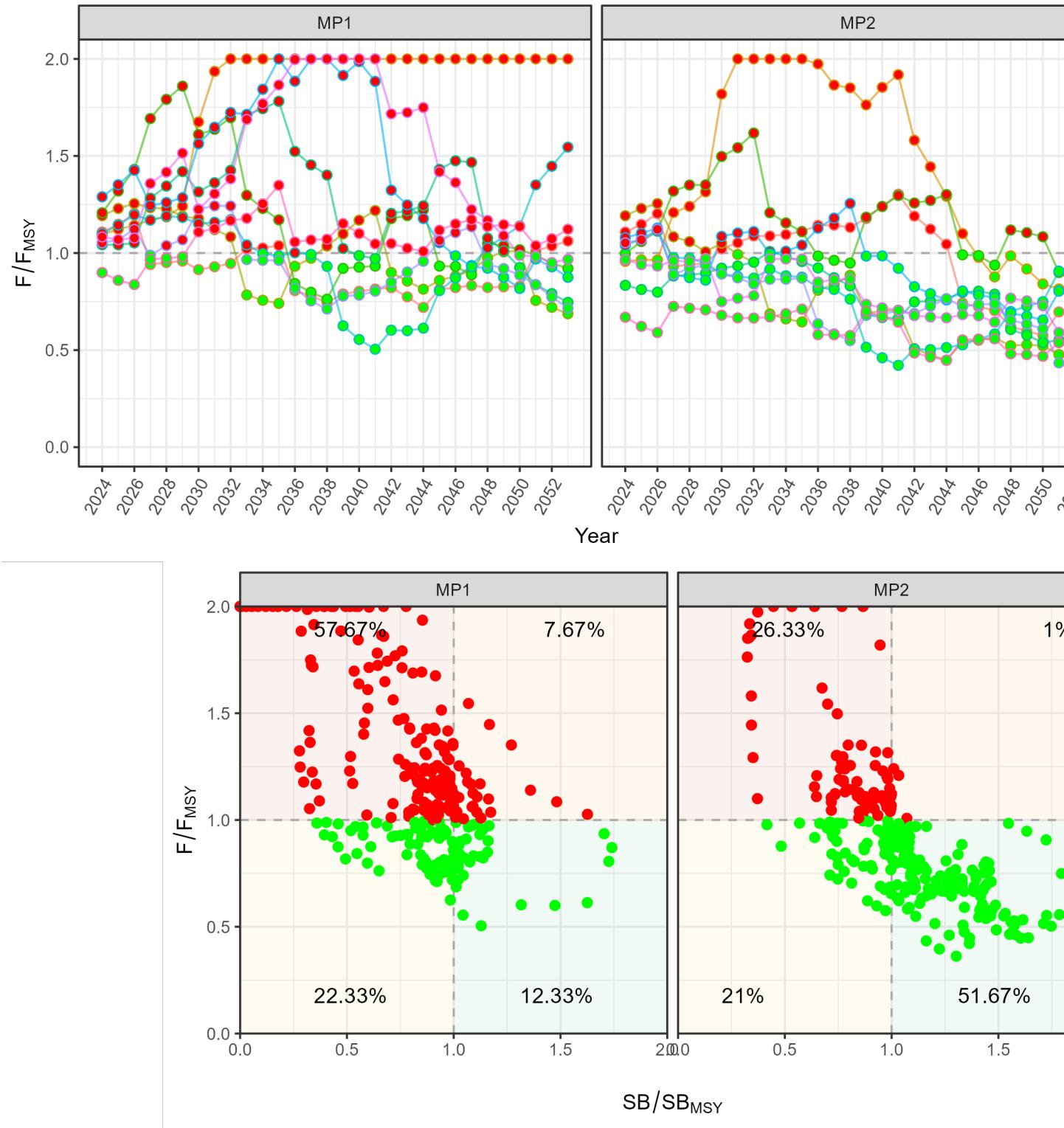


Figure 14: An example of a time-series plot of F/F_{MSY} for a two CMPs. This OM has 10 simulations, shown as the colored lines. Simulations where $F < F_{\text{MSY}}$ are indicated with a green point, and red points indicate $F > F_{\text{MSY}}$. The Kobe plot on the bottom shows these points in the Kobe space. The numbers in the corners indicate the percent of points in each area of the Kobe space.

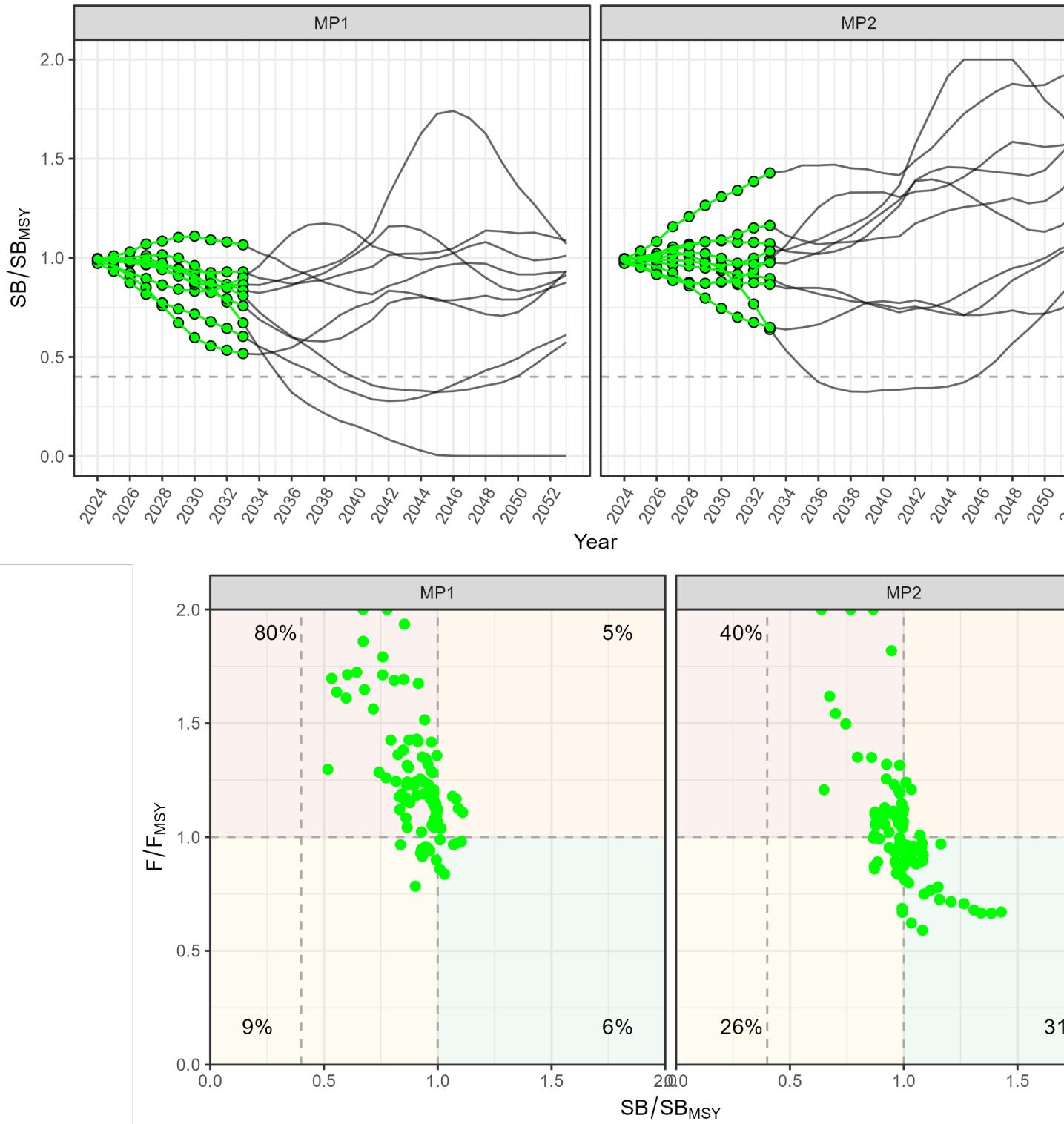


Figure 15: An example of a time-series plot of SB/SB_{MSY} for two CMPs. This OM has 10 simulations, shown as the colored lines. Simulations where $SB > 0.4SB_{MSY}$ are indicated with a green point, and red points indicate $SB < 0.4SB_{MSY}$. The Kobe plot on the bottom shows these points in the Kobe space. The numbers in the corners indicate the percent of points in each area of the Kobe space.

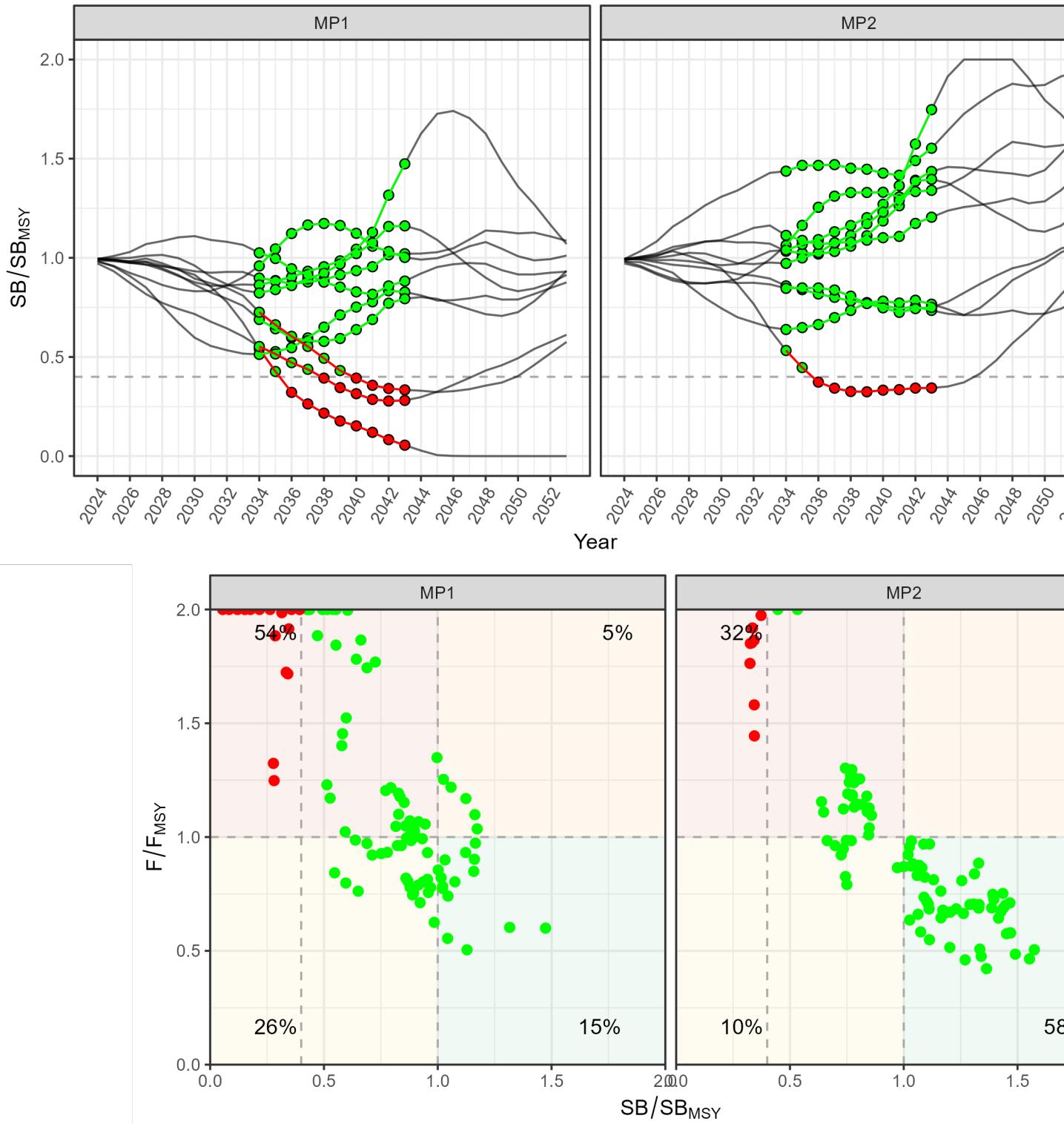


Figure 16: An example of a time-series plot of SB/SB_{MSY} for two CMPs. This OM has 10 simulations, shown as the colored lines. Simulations where $SB > 0.4SB_{MSY}$ are indicated with a green point, and red points indicate $SB < 0.4SB_{MSY}$. The Kobe plot on the bottom shows these points in the Kobe space. The numbers in the corners indicate the percent of points in each area of the Kobe space.

9.4.3 LRP_long

LRP_long calculates the probability of breaching the limit reference point (LRP; SB<0.4SBMSY) in any of years 21-30 (2044-2053).

This is calculated as the proportion of simulations where SB<0.4SBMSY in any of the last 10 years.

Figure 17 shows an example of two CMPs (MP1, MP2) for an operating model with 10 simulations. The value for LRP_long in this example are: 0.3, 0.1.

9.4.4 LRP

LRP calculates the probability of breaching the limit reference point (LRP; SB<0.4SBMSY) in any of the projection years (2024-2053).

This is calculated as the proportion of simulations where SB<0.4SBMSY in any year.

Figure 18 shows an example of two CMPs (MP1, MP2) for an operating model with 10 simulations. The value for LRP in this example are: 0.3, 0.1.

9.5 Yield

9.5.1 TAC1

TAC1 calculates the median TAC (t) in the first year (2024). This is calculated as the median across the simulations of the 2024 TAC.

Figure 19 shows an example of two CMPs (MP1, MP2) for an operating model with 10 simulations. The values for TAC1 in this example are: 13462, 11987.

9.5.2 AvTAC_short

AvTAC_short calculates the median TAC (t) in the first 10 years (2024–2033). This is calculated as the median TAC across the simulations and years.

Figure 20 shows an example of two CMPs (MP1, MP2) for an operating model with 10 simulations. The values for AvTAC_short in this example are: 13097, 11201.

9.5.3 AvTAC_med

AvTAC_med calculates the median TAC (t) in years 11 – 20 (2034–2043). This is calculated as the median TAC across the simulations and years.

Figure 21 shows an example of two CMPs (MP1, MP2) for an operating model with 10 simulations. The values for AvTAC_med in this example are: 10332, 10200.

9.5.4 AvTAC_long

AvTAC_long calculates the median TAC (t) in the last 10 years (2044–2053). This is calculated as the median TAC across the simulations and years.

Figure 22 shows an example of two CMPs (MP1, MP2) for an operating model with 10 simulations. The values for AvTAC_long in this example are: 9883, 8793.

9.6 Stability

9.6.1 VarC

VarC calculates the mean absolute change in the TAC between management cycles, calculated across all years and simulations.

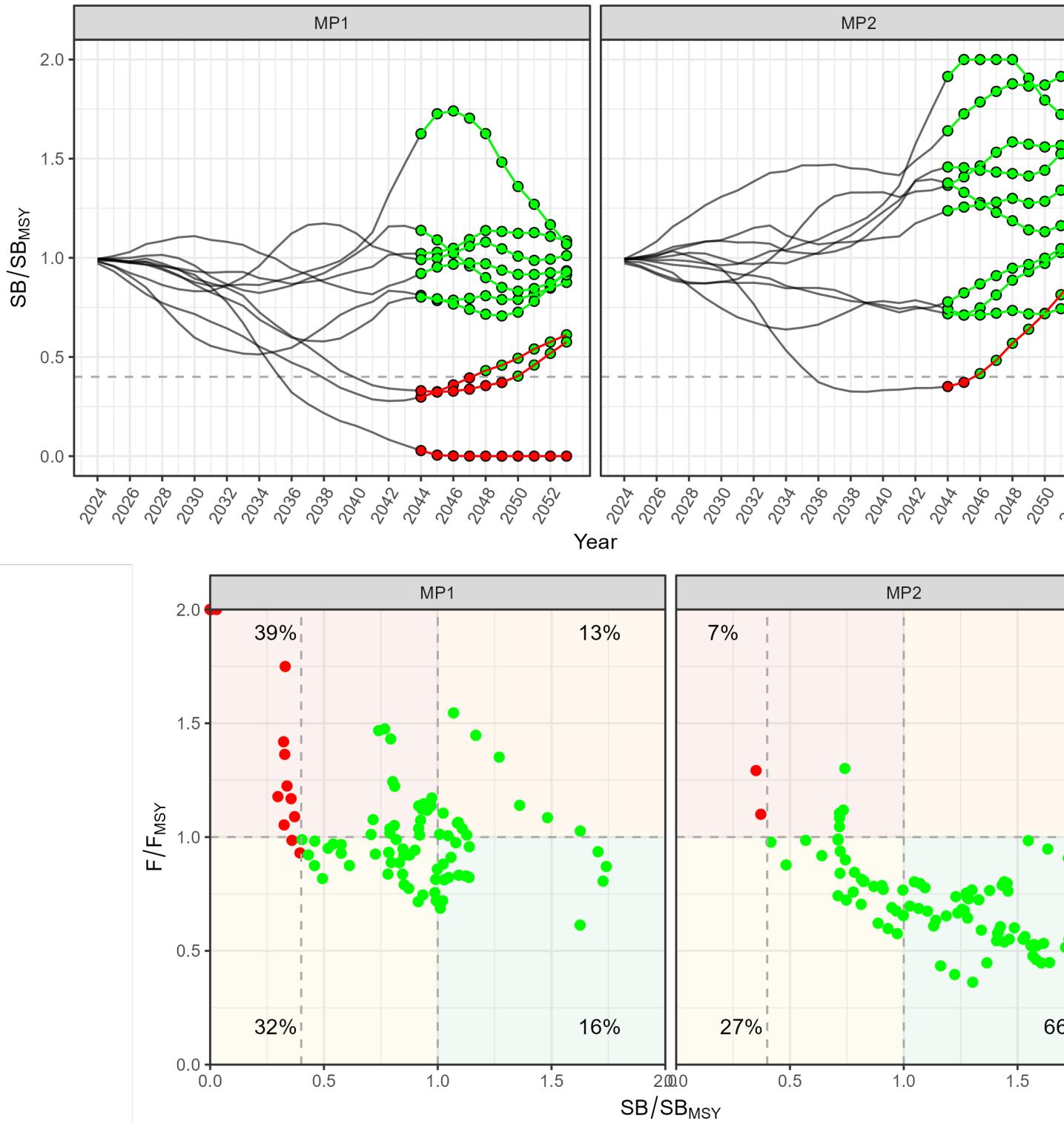


Figure 17: An example of a time-series plot of SB/SB_{MSY} for two CMPs. This OM has 10 simulations, shown as the colored lines. Simulations where $SB > 0.4SB_{MSY}$ are indicated with a green point, and red points indicate $SB < 0.4SB_{MSY}$. The Kobe plot on the bottom shows these points in the Kobe space. The numbers in the corners indicate the percent of points in each area of the Kobe space.

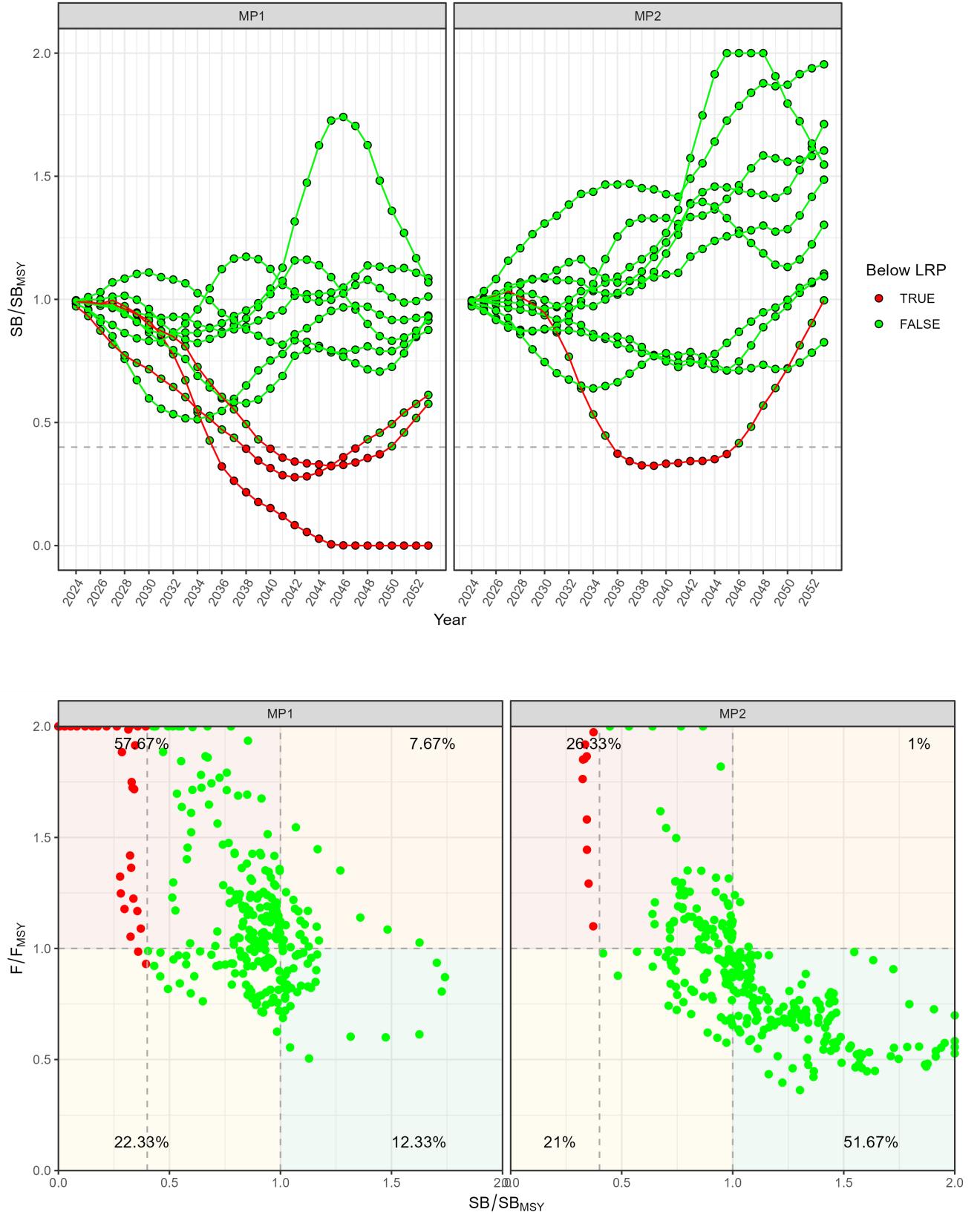


Figure 18: An example of a time-series plot of SB/SB_{MSY} for a two CMPs. This OM has 10 simulations, shown as the colored lines. Simulations where $SB > 0.4SB_{MSY}$ are indicated with a green point, and red points indicate $SB < 0.4SB_{MSY}$. The Kobe plot on the bottom shows these points in the Kobe space. The numbers in the corners indicate the percent of points in each area of the Kobe space.

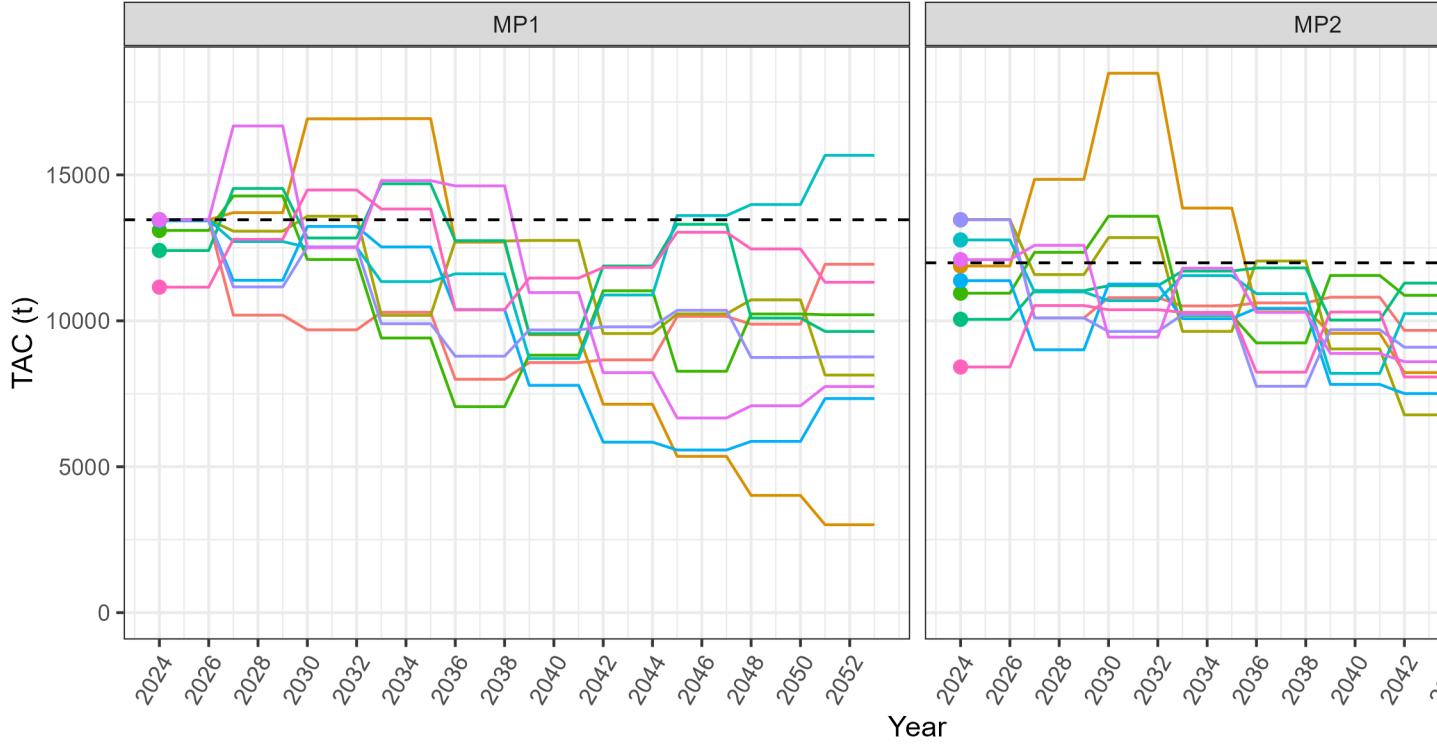


Figure 19: An example of time-series plots of the TAC for two CMPs. This OM has 10 simulations, shown as the colored lines. The points indicate the values used to calculate the median value.

Table 6: The codes, metrics, and targets used for the tuning of the CMPs.

Code	Metric	Target
a	PGK short	0.51
b	PGK short	0.60
c	PGK short	0.70

Figure 23 shows an example of two CMPs (MP1, MP2) for an operating model with 10 simulations. The values for VarC in this example are: 0.15, 0.14.

10 Candidate Management Procedures

10.1 Tuning of Candidate Management Procedures

During development, the Candidate Management Procedures are tuned to achieve specific performance outcomes. These performance outcomes are calculated across all the Reference OMs; i.e., the MSE results from each Reference OM are aggregated together. 3 different tuning targets have been specified (Table 6).

There will be 3 versions of each CMP corresponding to metrics and targets described in Table 6. The names of the CMPs include the tuning code (i.e., ‘a’, ‘b’).

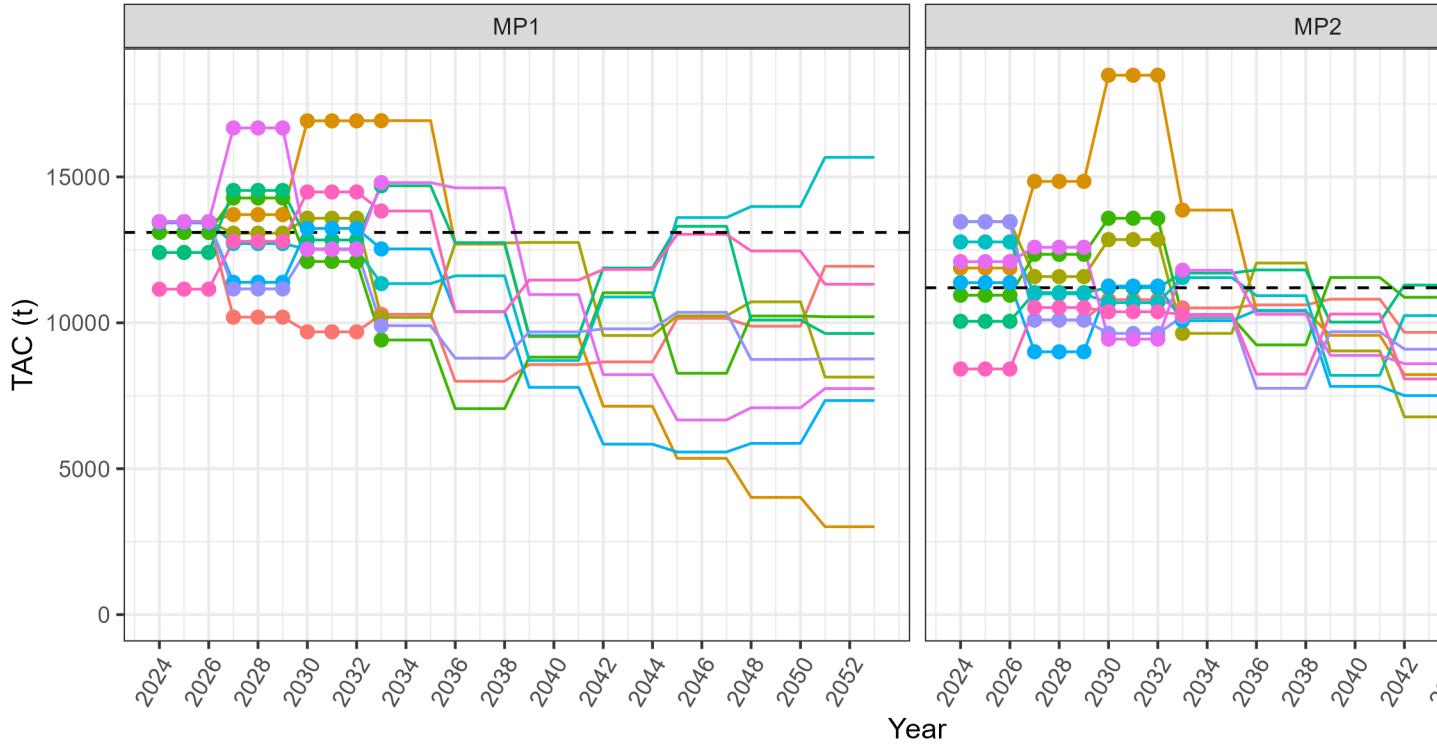


Figure 20: An example of time-series plots of the TAC for a two CMPs. This OM has 10 simulations, shown as the colored lines. The points indicate the values used to calculate the median value.

11 Reporting of MSE Results

11.1 Interactive Application

The current MSE results are available in an [Interactive Application](#)

References

- Anon. (2017). Report of the 2017 ICCAT Atlantic Swordfish Stock Assessment Session. SCRS/2017/008. *Collect. Vol. Sci. Pap. ICCAT*, 74(3), 841–967. https://www.iccat.int/Documents/CVSP/CV074_2017/n_3/CV074030841.pdf
- Anon. (2021). Peer review of the North Atlantic Swordfish Management Strategy Evaluation (MSE) code and algorithms SCRS/2021/097. *Collect. Vol. Sci. Pap. ICCAT*, 78(7).
- Carruthers, T., & Hordyk, A. R. (2018). The Data-Limited Methods Toolkit (DLMtool): An R package for informing management of data-limited populations. *Methods in Ecology and Evolution*, 9(12), 2388–2395. <https://doi.org/10.1111/2041-210X.13081>
- Francis, R. I. C. C. (2011). Data weighting in statistical fisheries stock assessment models. *Canadian Journal of Fisheries and Aquatic Sciences*, 68(6), 1124–1138. <https://doi.org/10.1139/f2011-025>
- 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>
- Ortiz, M., Mejuto, J., Hanke, A., Ijima, H., Walter, J., Coelho, R., & Ikkiss, A. (2017). Updated combined biomass index of abundance of North Atlantic swordfish stock 1963 - 2015. SCRS/2017/137. *Collect. Vol. Sci. Pap. ICCAT*, 74(3), 1275–1294. https://www.iccat.int/Documents/CVSP/CV074_2017/n_3/CV074031275.pdf

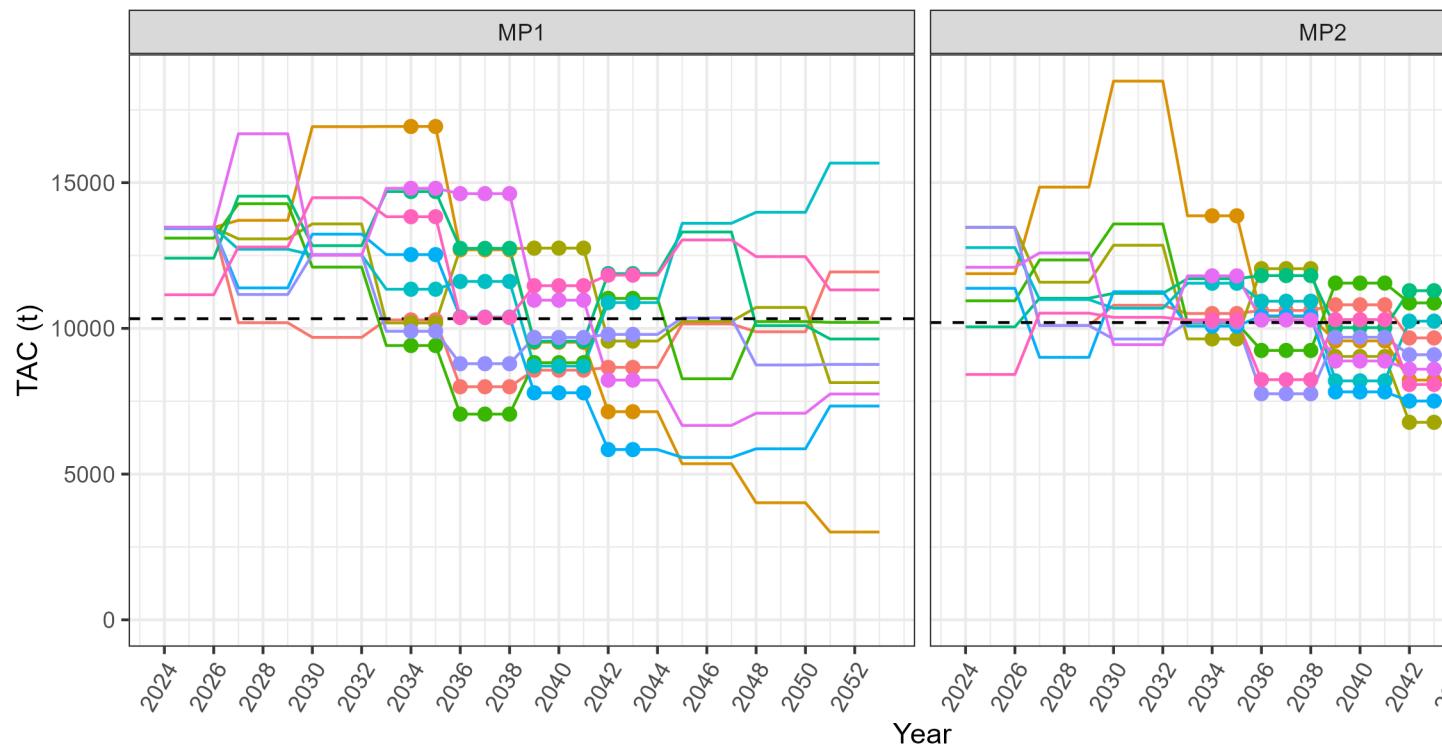


Figure 21: An example of time-series plots of the TAC for two CMPs. This OM has 10 simulations, shown as the colored lines. The points indicate the values used to calculate the median value.

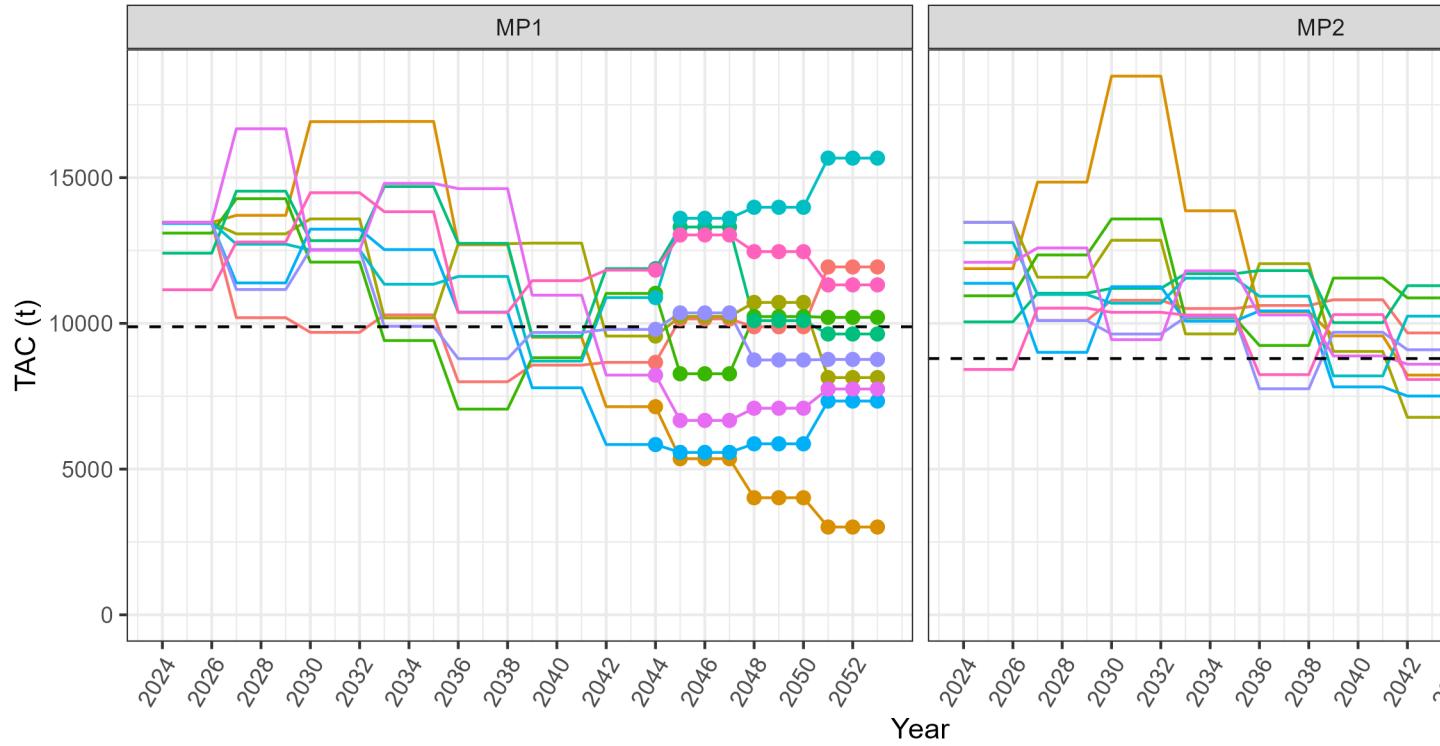


Figure 22: An example of time-series plots of the TAC for two CMPs. This OM has 10 simulations, shown as the colored lines. The points indicate the values used to calculate the median value.

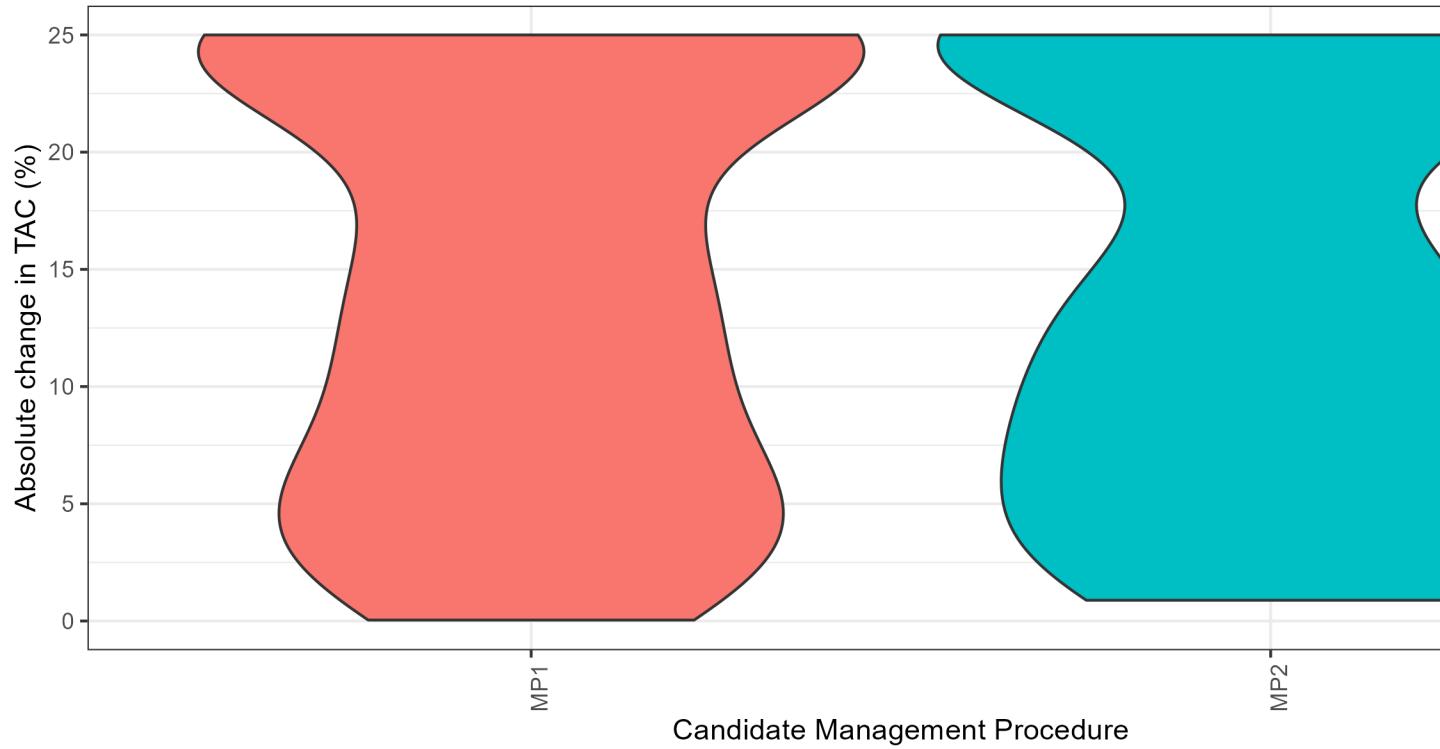


Figure 23: Violin plots of the relative absolute change in the TAC between management cycles for two CMPs.