



# Management strategy evaluation operating model conditioning: a swordfish case study

Daniela Rosa · Iago Mosqueira · Dan Fu · Rui Coelho

Received: 13 July 2023 / Accepted: 3 June 2024  
© The Author(s) 2024

**Abstract** Evaluation of fish stock status is a key step for fisheries management. Tuna Regional Fisheries Management Organizations (t-RFMOs) are moving towards management strategy evaluation (MSE), a process that combines science and policy and depends on technical aspects, developed by scientists, designed to meet management objectives established by managers and other stakeholders. In the Indian Ocean, the current management advice for swordfish (*Xiphias gladius*) is based on an ensemble of 24 models considering four areas of uncertainty about the stock dynamics. There is an ongoing MSE process for swordfish, and this paper describes the methodology

being applied for the conditioning of the operating model (OM), including model selection and validation. In the MSE, nine sources of uncertainty were considered, each being characterized by 2–3 levels. A partial factorial design was employed to reduce the number of models from a full factorial design to those needed to encompass the overall uncertainty. A selection and validation process was carried out, filtering models that converged, showed good predictive skills, and provided plausible estimates. Overall, the estimated spawning stock biomass (SSB) relative to SSB at maximum sustainable yield (MSY), and fishing mortality (F) relative to  $F_{MSY}$  encompasses the estimates of the stock assessment ensemble at the most optimistic area of the distribution. The MSE for swordfish is an ongoing process that is expected to provide more robust management advice in the future. Further developments to the OM can still occur, but the methods presented herein can be applied to this, or other species, MSE processes.

**Supplementary Information** The online version contains supplementary material available at <https://doi.org/10.1007/s11160-024-09868-w>.

D. Rosa (✉) · R. Coelho  
IPMA – Instituto Português do Mar e da Atmosfera, I.P.,  
Av. 5 de Outubro S/N, 8700-305 Olhão, Portugal  
e-mail: daniela.rosa@ipma.pt

D. Rosa · R. Coelho  
CCMAR – Centro de Ciências do Mar da Universidade  
do Algarve, Universidade do Algarve, Campus de  
Gambelas, 8005-139 Faro, Portugal

I. Mosqueira  
WMR – Wageningen Marine Research, Haringkade 1,  
1976 CP IJmuiden, The Netherlands

D. Fu  
IOTC – Indian Ocean Tuna Commission, Blend Building,  
Providence, Seychelles

**Keywords** Diagnostics · Model validation · Filtering · t-RFMO · *Xiphias gladius* · Indian Ocean

## Introduction

Effective natural resource management is a complex process that involves a comprehensive understanding of both the natural dynamics of resources and the human element of decision-making and management

(Bunnefeld et al. 2011). In the marine realm, fisheries management must deal with multiple (e.g., biological, ecological, economical, and social), and often conflicting, objectives (Kell et al. 2007). Evaluating fish stocks status and providing harvest limits is one of the key elements of the fisheries management cycle. Historically, these evaluations have been based in a scenario where a stock assessment is used to characterize the system using a set of “best” assumptions which are considered to be true (Punt and Donovan 2007; Punt et al. 2016) and uncertainty evaluated through confidence intervals and sensitivity tests (Butterworth 2008). Scientific management advice is then provided based on projections under constant catch or fishing mortality or through a harvest control rule (Butterworth 2008). Under this scenario, uncertainty is either not considered, or not explicitly incorporated in the evaluation process (Kell et al. 2007). Attempts to better capture assessment uncertainty and increase the robustness of advice provision have recently made progress (Jardim et al. 2021; Ducharme-Barth and Vincent 2022), but formal ways of incorporating uncertainty into scientific management measures are still lacking (Punt and Donovan 2007). To overcome these issues, fisheries management advice is moving from the “best assessment” approach to a management strategy evaluation (MSE) approach.

MSE is a simulation framework based on the operating model (OM) which is a set of models that represent the uncertainties about the stock and the fisheries dynamics. For the construction of the OM it is necessary to identify the suitable range of uncertainties and condition, i.e. fit, the models to existing observations. These are used to simulation-test management procedures (MPs), in order to identify those MPs that meet management objectives with a high probability and are robust to uncertainty (Kolody et al. 2019). An MP includes the management action, through a decision rule or harvest control rule (HCR), but also the processes of data collection and analysis to obtain an indicator of stock status (Kell et al. 2007). This process uses a closed-loop simulation with feedback between the two systems, that allows to provide advice that is robust to uncertainty and assess trade-offs of different MPs according to their performance relative to the identified management objectives.

Management strategy evaluation was pioneered by the International Whaling Commission in the 1980s, that first started to calculate potential catch limits

for commercial whaling and determine actual strike limits for aboriginal subsistence whaling (Punt and Donovan 2007). From the early 1990s the use of MSE expanded and has been applied to several fisheries; for example, in South Africa, Australia, USA, New Zealand and the European Union (Punt et al. 2016; ICES 2020; de Moor et al. 2022). MSE efforts have been mainly focused on domestic fish stocks, but regional fisheries management organizations (RFMOs) have also applied MSE for international fisheries management, for example Greenland halibut (*Reinhardtius hippoglossoides*) in the Northwest Atlantic Fisheries Organization (NAFO) (NAFO 2021).

In the case of large migratory species, that are exploited in both international and national waters of many nations, their international management is regulated by the tuna RFMOs (t-RFMOs). There are 5 t-RFMOs worldwide, namely one in the Atlantic (ICCAT—International Commission for the Conservation of Atlantic Tunas), one in the Indian Ocean (IOTC—Indian Ocean Tuna Commission), two in the Pacific (WCPFC—Western and Central Pacific Fisheries Commission, and IATTC—Inter-American Tropical Tuna Commission) and one responsible for the southern bluefin tuna stock (CCSBT—Commission for the Conservation of Southern Bluefin Tuna). All t-RFMOs have begun to implement MSE processes, with multiple MSEs coexisting in the majority of them (Anonymous 2018) and several having been adopted. The CCSBT was the first to pursue an MSE and to adopt a fully specified management procedure tested under MSE, namely in 2011 for southern bluefin tuna (*Thunnus maccoyii*) (Hillary et al. 2016). The IOTC adopted an HCR for skipjack in 2016 (IOTC 2016a) and technical work has continued towards the implementation of an MP<sup>1</sup>. For North Atlantic albacore, managed by ICCAT, an HCR was adopted in 2017, that was converted to a fully specified MP in 2021 (ICCAT 2021). More recently, MPs have been tested under MSEs for the Atlantic bluefin tuna (*Thunnus thynnus*) (ICCAT 2022), Indian Ocean bigeye tuna (*Thunnus obesus*) (IOTC 2022a, 2022b), Western Central Pacific skipjack (WCPFC 2022), and

<sup>1</sup> During the revision process of the manuscript, IOTC has adopted a management procedure for skipjack (Resolution 24/07) and swordfish (Resolution 24/08).

North Pacific albacore, adopted both by IATTC and WCPFC (IATTC 2023; WCPFC 2023).

Conducting an MSE proceeds along several steps (Punt et al. 2016). Amongst the first steps of a MSE, from the technical point of view, is identifying uncertainties and developing operating models (OMs). A recent review on the OM design in t-RFMOs, Sharma et al. (2020) highlighted common issues that could be improved in their technical development, focusing particularly on the operating model development. Although various species used some strategies for model selection and validation, it was noticed that there was a lack of consistency on these practices, and it is important to guarantee that the key uncertainties identified are included in a coherent manner.

In order to address this issue, the purpose of this work is to outline the procedure that has been used to condition, select, and validate an operating model for Indian Ocean swordfish, as well as how uncertainties have been incorporated and examined using a variety of diagnostic techniques. The paper herein encapsulates the research work that is being conducted for swordfish MSE in the Indian Ocean. The expected benefits of having a stock managed through management strategy evaluation are presented and a summarization about the importance of interactions between scientists and managers and how this occurs in the case of IOTC is provided.

## Operating model development

### Stock assessment

The operating models of the Indian Ocean swordfish were constructed based on the stock assessment model used to assess this species. The most recent stock assessment was conducted in 2020 and followed the previous assessments from 2017 using similar specifications. The assessment is based on Stock Synthesis 3 (Methot and Wetzel 2013) which is an age-structured integrated statistical population dynamics model. The population model is sex-specific, with ages ranging from 0 to 30+ years and spatially disaggregated into four areas with no movement dynamics between areas. Data is available for the period from 1950 to 2018. It integrates several sources of fisheries and biological data, including catch information from 15 fisheries, defined by gear and area, and

length composition data for 14 of them. For complete details of the model configurations see Fu (2020) and IOTC (2020). Management advice is currently based on the results of a combination of model configurations, including two options on age and growth (spine or otolith-based age estimates), three values for the Beverton-Holt stock-recruit relationship steepness ( $h$ ), two values for recruitment variability ( $\sigma_R$ ), and alternative effective sample sizes of length composition data (ESS). These resulted in an ensemble of 24 models.

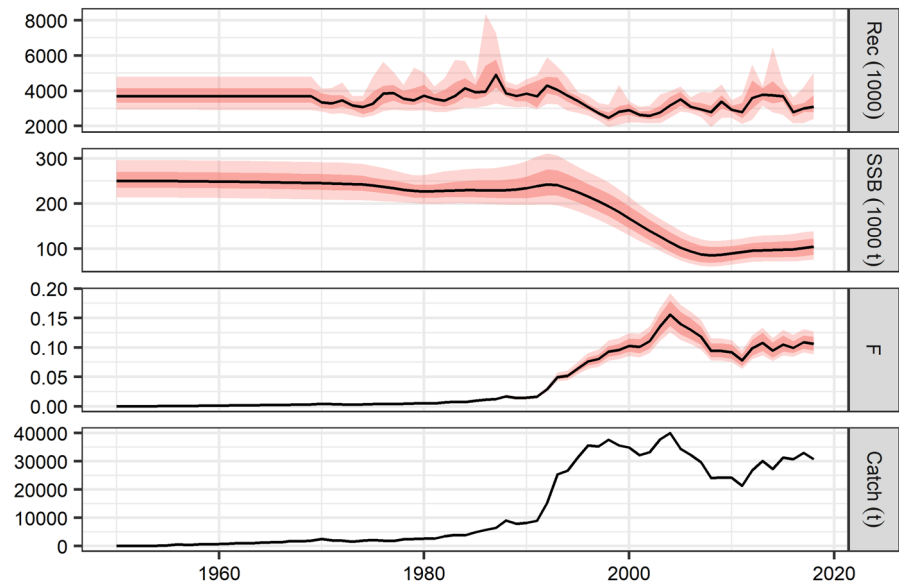
Within this stock assessment model ensemble, fishing mortality rates ( $F$ ) increased sharply from the 1950s, and then decreased after the 2000s, being stable from the 2010s. The spawning stock biomass (SSB) decreased sharply between the 1990s through to the 2000s, having increased slightly since 2010 (Fig. 1). The stock was considered in 2018 (last year of data) not to be overfished and not subject to overfishing, with the final SSB above the value that provides the maximum sustainable yield (MSY) and  $F$  estimated to be lower than the  $F$  at MSY.

### Structural uncertainties

The 2017 session of the IOTC Working Party on Methods (WPM) (IOTC 2017a) discussed and proposed an initial set of elements likely to be responsible for model uncertainty, subsequent to the 2017 stock assessment. Nine factors were identified as key uncertainties for the Indian Ocean swordfish stock. A first iteration of OM conditioning was performed based on the 2017 stock assessment (Rosa et al. 2018; Mosqueira et al. 2017). The stock assessment was updated in 2020 (see section above), but the identified uncertainties remained unchanged. The factors and the levels of each factor are presented in Table 1 and summarized below.

*Selectivity:* In this document, selectivity refers to population selectivity, which is a measure of the length- (or age-) specific mortality due to fishing, resulting from both contact selectivity and availability (Sampson 2014). Selectivity determines how catch is removed from the population and influences estimates of stock productivity. In the swordfish stock assessment and for the MSE work, two possible functional forms were considered for the selectivity-at-length of the longline fleets, namely a double normal in which selectivity decreases in the older ages, and a logistic

**Fig. 1** Time series of recruitment (in thousands), spawning stock biomass (SSB in 1000 tonnes), catch (in tonnes) and fishing mortality for the 24-stock assessment model ensemble used for management advice of Indian Ocean Swordfish. The black line represents the median, the dark shaded red the 75% quantiles and the light shaded red the 95% quantiles



**Table 1** Summary view of the variables and levels of each variable considered as sources of uncertainty for the Indian Ocean swordfish operating model grid

Variable	Values		
Selectivity	Double normal	Logistic	
Steepness ( <i>h</i> )	0.6	0.75	0.9
Growth and maturity	Slow growth, late maturity—spine-based estimates from Wang et al. (2010)	Fast growth, early maturity—otolith-based estimates from Farley et al. (2022)	
Mortality ( <i>M</i> )	0.2	0.3	Sex-specific <sup>a</sup>
Sample size of length frequency data (ESS)	2	20	
Recruitment variability ( $\sigma_R$ )	0.2	0.6	
CPUE <sup>b</sup> scaling	Area effect $\times$ Surface	Catch	Biomass
CPUE series	JPN late <sup>c</sup> + EU.PRT <sup>d</sup>	JPN late	TWN late <sup>e</sup>
Catchability increase (LLQ)	0%	1%/year	

<sup>a</sup>Using the Lorenzen (1996) equation with otolith-based age estimates from Farley et al. (2022); <sup>b</sup>CPUE—catch per unit of effort; <sup>c</sup>JPN late—Japanese CPUE series from 1994 to 2018; <sup>d</sup>EU.PRT—Portuguese CPUE series from 2000 to 2018; <sup>e</sup>Taiwanese CPUE series from 1994 to 2018

function in which selectivity remains flat after reaching its asymptote. It is common in stock assessments to include at least one fleet with an asymptotic selectivity, as having only dome shaped curves can produce a “cryptic” biomass that is not in accordance with the observed proportion of larger fish. The two alternatives considered were to assume: 1) a logistic selectivity for the Japanese longline fleet while the other longline fleets are estimated to have a double

normal selectivity; 2) all longline fleets had a logistic selectivity.

*Steepness (h)*: Steepness from the Beverton and Holt stock-recruit relationship (SRR) is the fraction of virgin recruits produced by 20% of the virgin spawning stock (Mace and Doonan 1988). Steepness is indicative of stock productivity and therefore resilience of the stocks, with higher values indicating higher productivity. It is a very influential parameter which is difficult to estimate in most stock

assessments, as data usually do not have information about it. Therefore, in most tuna and billfish assessments it is a parameter fixed at a certain value. Three values were considered for this parameter, with one being the same as what was used in the base case stock assessment model ( $h=0.75$ ), and plausible lower and upper values for the species, namely  $h=0.6$  and  $h=0.9$ .

*Recruitment variability:* Recruitment deviates, i.e. how much recruitment has deviated from the expected from the SRR, are usually regarded as process error in statistical catch-at-age models as there are various, usually unexplained, reasons why the estimated recruitment differs from the expected. Recruitment deviates are modeled as a lognormal distribution with a mean=0 and an associated standard deviation. Recruitment deviation ( $\sigma_R$ ) is the standard error of the recruitment deviates in log space and is usually fixed in stock-assessments that use penalized likelihood such as Stock Synthesis. For the MSE purpose, two levels were applied, 0.2 and 0.6. Increasing recruitment variability may not influence current stock status, but will inflate future recruitment variability in projections (Kolody et al. 2019).

*Growth and maturity:* Growth and maturity are very important parameters in stock assessment. Growth processes will interact with other processes in the estimation of fishing mortality and is an important factor for key reference quantities (e.g. MSY) (Maunder et al. 2016), while maturity estimates allow for the estimation of the proportion of the population that is able to reproduce. In the case of swordfish, there are concerns in age estimation with differences being found between age estimates obtained from spines versus those coming from otoliths, particularly for older/larger individuals (Farley et al. 2022). This uncertainty also undermines the age at maturity relationship. A slow growth and late maturity option was considered based on the estimates from spines provided by Wang et al. (2010), and alternatively a faster growth with earlier maturation was based on the otolith estimates from Farley et al. (2022). In both cases, sex-specific growth estimates are used as swordfish exhibit a marked difference in growth between males and females.

*Natural mortality:* Natural mortality is difficult to estimate and is commonly unknown for many fish stocks despite its importance in estimating the stock productivity (Punt et al. 2021). Given the difficulties

in estimating this parameter, it is common to fix it at a certain value, or values at age. There are a broad range of  $M$  values assumed in other swordfish assessments worldwide, ranging from at least 0.2–0.5. Three alternatives for  $M$  were considered for the MSE operating model grid, with two constant across ages ( $M=0.2$  and  $M=0.3$ ) and a third with age and sex-specific  $M$  based on the Lorenzen equation (Lorenzen 1996). The age specific mortality was scaled so that  $M$  at age at maturity (age 6) was 0.25.

*Sample size of length frequency data:* Two main data inputs into Stock Synthesis are standardized catch-per-unit-effort (CPUE) and length composition data. These two sources of data inform the model about the stock dynamics and the effects of fishing at length. In the stock assessment, all length composition strata from all fleets were weighted relative to the number of fish measured, with the maximum sample size capped at 20. For the model grid the alternative was to down weight the length composition data further by capping the sample size at 2. As the weights in the total likelihood of the different data sources are related, by down weighting the length frequency data, the weight of the CPUE series is being upweighted, meaning there is the assumption of less confidence in the length data and more on the CPUEs.

*CPUE series:* Available CPUE series used in the stock assessment showed conflicting trends both between fleets and within fleets between areas. The base case considered in the assessment used the Japanese late CPUE series (1994–2018), replaced by the Portuguese index from 2000 to 2018 in the Southwest area (SW), and with an overlap in years 2000–2003. This was maintained in the OM grid, and two additional options were added: 1) using the Japanese CPUE (1994–2018) for all areas, and 2) using the Taiwanese CPUE (1994–2018), also replaced in the SW by the Portuguese index (2000–2018, with an overlap in the years 2000–2003).

*CPUE scaling:* The stock assessment assumed a stock residing in four areas/quadrants of the Indian Ocean (NW, NE, SW, SE), and CPUE scaling was used to convert the area specific CPUE indices to relative abundance indices that are comparable among areas (Hoyle and Langley 2020). Three alternatives for scaling the CPUEs were considered, namely scaling by area, by catch or by biomass of each region.

The biomass-based estimates were derived by fitting region-specific models which included only regional fishery catches and observations, and the resulting stock biomass in each region were used to provide information on regional abundance distribution in the spatially disaggregated Indian Ocean model.

*Catchability increase:* Two scenarios were considered for the effective catchability of the CPUE. One assumes that the fleets have not improved their ability to fish for swordfish over time, or that any increase had been captured by the CPUE standardization process (0% increase). The alternative scenario considers a 1% *per* year increase in catchability which represents potential technology changes not accounted for in the statistical CPUE standardization methods and is therefore added to the CPUE indices to reflect this.

### Conditioning

Following the choice of uncertainties to be considered and their range, models must be conditioned, i.e. operating model parameters are estimated by fitting to available data. In the specific case of Indian Ocean swordfish, the OM was constructed around the Stock Synthesis population model used in the stock assessment. A full factorial design, where all interactions between all factor levels in Table 1 was considered, resulting in 2592 Stock Synthesis models that would compose the OM. The approach explored below results from the application of a partial factorial design to decrease the number of models that compose the OM. Decreasing the number of model runs has two objectives: 1) reducing redundancy, as a full factorial design can result in some models providing essentially the same information about the population, and 2) reducing the computational demand as running the models and their diagnostics is highly computationally demanding, and this demand would also apply to the projections period and the application of the management procedure (MP).

### Software implementation

The Indian Ocean swordfish MSE code is based on the FLR Project R library for quantitative fisheries science (Kell et al. 2007). All analysis were conducted with R (R Core Team 2022). Presented work in this document refers to the development of the

operating model up to the end of 2021 and is available in a public Github repository.<sup>2</sup>

### Partial factorial design

Partial factorial designs require only a fraction of the runs of a full factorial, while preserving many of its attributes in cases where conducting the full factorial design becomes unmanageable (Lawson 2015). In the current case, it can be considered that 3-way (and higher) interactions are rare and that models in the full factorial design might be redundant, not providing any new information on stock trajectories, therefore the partial factorial design was applied. Several design sizes were evaluated, using the R package “AlgDesign” (Wheeler 2022). The minimax normalized variance (Ge) was analyzed as a measure of efficiency with respect to the optimal approximate theory design, with design sizes ranging from 50 to 250 (Fig. 2). Ge varies between 0 and 1, with 1 indicating an optimal design. A cut-off of 0.95 was defined as the minimum acceptable Ge to choose the number of model runs. In this case, 108 model runs were chosen to represent the uncertainties in stock dynamics, as increasing the number of models does not appear to appreciably improve Ge.

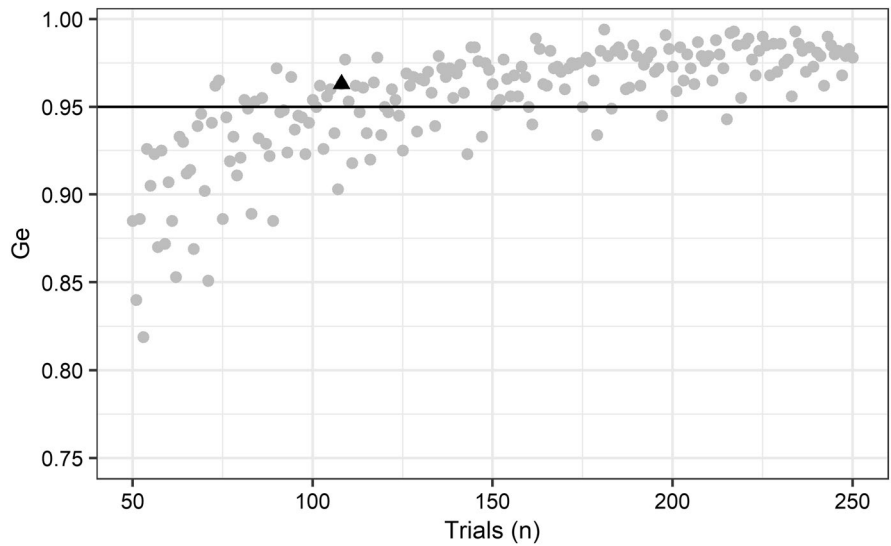
### Model diagnostics and selection

Model validation and filtering after conditioning should occur, through checking for several model diagnostics that evaluate model fits to data. For the 108 models, a drop one-off analysis (i.e. where up to five years of each model are dropped and the model refitted) was performed, and a set of diagnostics calculated (see Supplementary materials for an example). Based on model diagnostics and key population dynamic features, a four-step model selection was conducted to exclude models, or at least to flag models that could be considered less plausible.

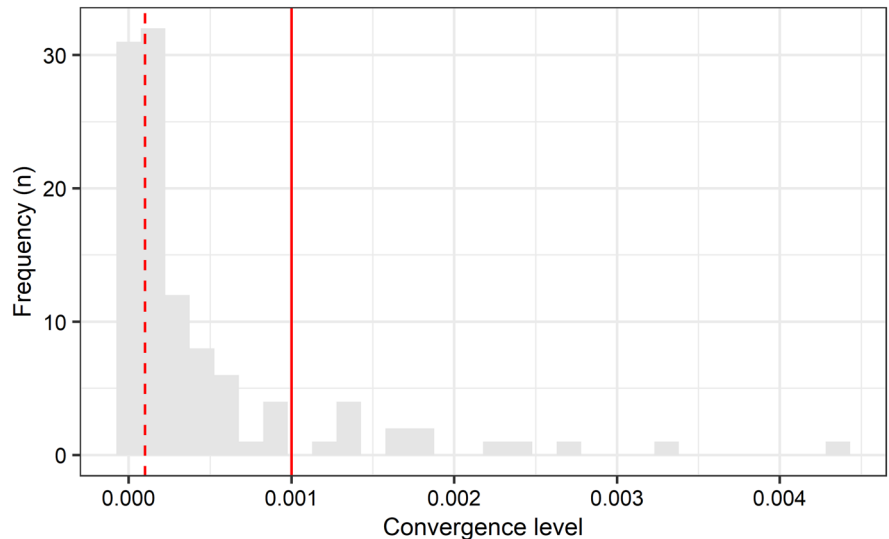
*Model convergence:* The first step was to identify models with non-convergence issues. Convergence is assessed by the inversion of the Hessian matrix, and the value of final gradient of the objective function at the solution, from here forward mentioned as convergence level. Carvalho et al. (2021) suggested

<sup>2</sup> <https://github.com/iotcwpm/SWO>

**Fig. 2** Minimax normalized variance expressed as an efficiency with respect to the optimal approximate theory design ( $G_e$ ) for an increasing number of trials. The triangle represents the sampling design of 108 model runs chosen to be sufficient to represent the uncertainties in stock dynamics for the Indian ocean swordfish



**Fig. 3** Frequency distribution of the convergence level for the 108 model runs. The solid line represents the applied cut-off value ( $1e-3$ ) and the dashed line the  $1e-4$



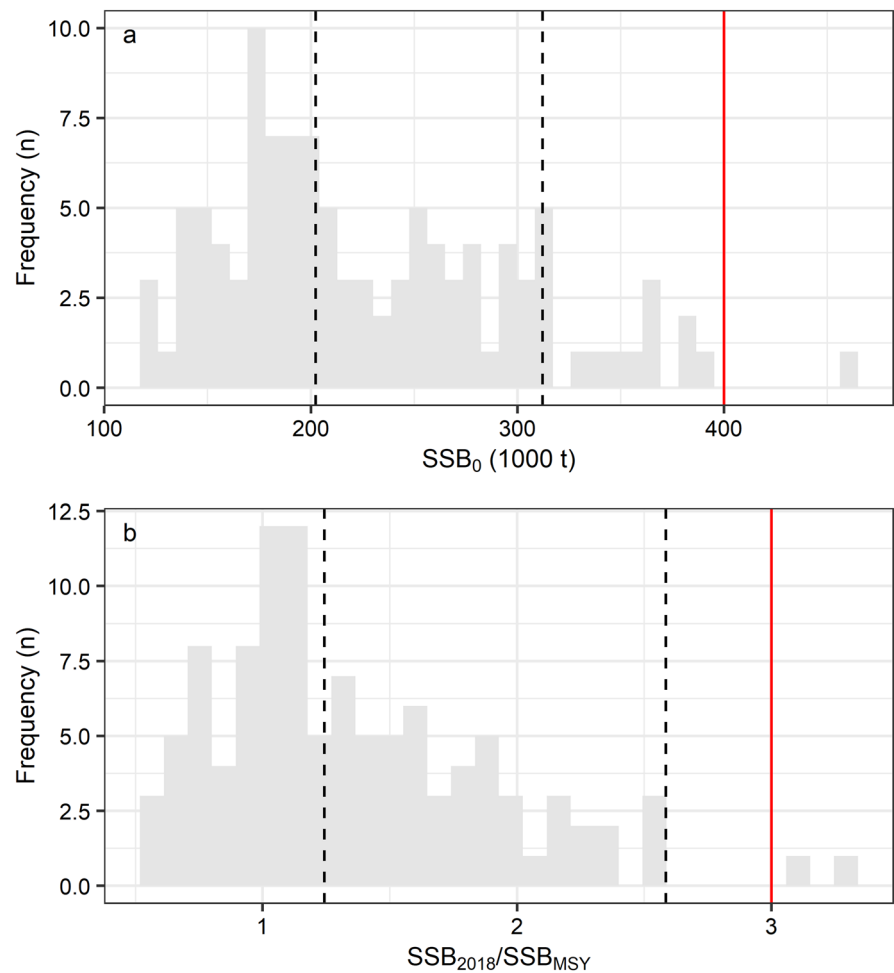
0.0001 as a threshold for the convergence level, noting that a small gradient might not be an absolute requirement. In this case, the frequency distribution of convergence levels was plotted (Fig. 3) to assess a plausible threshold that could be applied to these sets of models, as few models were below the proposed threshold. Based on this distribution the suggested threshold is of 0.001, resulting in 14 excluded models (out of 108).

*Plausibility of population quantities:* The second step is to identify models that yield unrealistic population quantities, for example, virgin biomass or stock status in the last year of data. In the case of swordfish,

it is unlikely that the virgin population biomass is above 400,000 tons (Fig. 4) and that the stock status ( $SSB_{2018}/SSB_{MSY}$ ) in the last assessment year is higher than 3 (Fig. 4) as this would indicate a stock that is in too pristine state considering the known exploitation history. Only three models out of the 108 from the partial factorial design did not meet these assumptions.

*CPUE predictive skill:* The model-free hindcasting technique (HCXval) uses cross-validation to compare observations to their predicted future values (Kell et al. 2016; Carvalho et al. 2021). The prediction skill of a model is then computed from the prediction

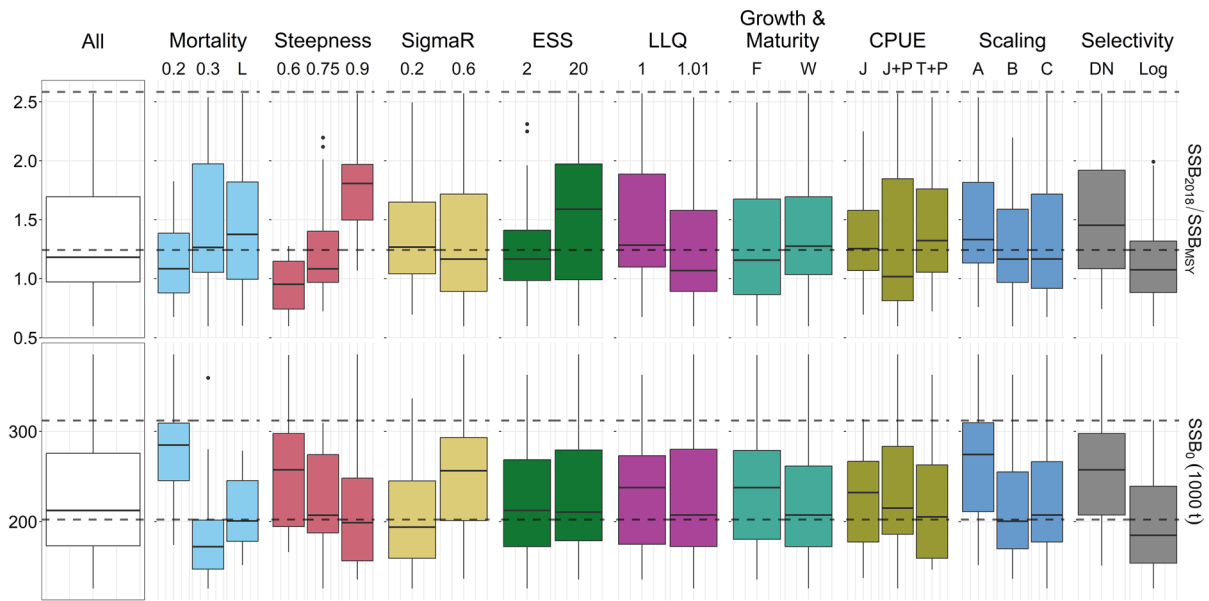
**Fig. 4** Distribution of the 108 models estimated values of **a** virgin spawning stock biomass ( $SSB_0$ ) and **b**  $SSB$  in 2018 relative to  $SSB$  at maximum sustainable yield ( $SSB_{2018}/SSB_{MSY}$ ). Dashed lines represent the minimum and maximum values observed in the 2020 stock assessment and the solid line represents the applied cut-off



residuals. A robust statistic for evaluating the prediction skill can be constructed using the mean absolute scaled error (MASE) of Hyndman et al. (2006). A MASE score larger than 1 indicates the model does only as well as a random walk at predicting the quantity, while a value of 0.5 indicates the model is twice as good as a random walk. In the projection period CPUEs will most likely be used, either as part of model based or empirical MPs, therefore it is important that CPUEs have adequate prediction skills. The MASE score from the CPUEs was used as one of the filters for this model validation and selection procedure. MASE scores were computed using package “ss3diags” in R (Winker et al. 2021). For the OMs, the stock is spatially distributed in four areas, however the projection model does not have spatial structure (only one area is considered). In this case, the CPUE from one of the areas would have to be scaled

to represent the four areas. Given that the NW region presented the CPUEs with lowest MASE score (i.e. better prediction skill), runs with a MASE score  $> 1$  for the CPUE series from this region were also considered to be excluded from the operating model (26 models).

*Catch data update:* Finally, the fourth step for model validation and selection was to investigate if the model can sustain the level of total catch observed after the last year in the model fit. There is a lag between the assessment year and the last data year (i.e. the stock assessment was conducted in 2020 and the last year of data was 2018), so it could happen that some models do not support the most recent catches. In the case of swordfish, the catch of 2019



**Fig. 5** Comparison of virgin spawning stock biomass ( $SSB_0$ ) and SSB in 2018 relative to SSB at maximum sustainable yield ( $SSB_{2018}/SSB_{MSY}$ ) by level of each uncertainty factor. The dashed lines show the minimum and maximum values of  $SSB_0$  and  $SSB_{2018}/SSB_{MSY}$ , respectively, returned by the stock assessment grid. L: age based mortality; sigmaR: recruitment variance; ESS: Sample size of length frequency data; LLQ: Catchability increase; F: otolith-based growth and maturity estimates from Farley et al. (2022); W: spine-based growth

and maturity estimates from Wang et al. (2010); J: Japanese catch-per-unit-of-effort (CPUE) series from 1994 to 2018; J+P: Japanese CPUE series from 1994 to 2018 replaced by the Portuguese CPUE series from 2000 to 2018 in the Southwest (SW) area; T+P: Taiwanese CPUE series from 1994 to 2018 replaced by the Portuguese CPUE series from 2000–2018 in the Southwest (SW) area; A: area; B: biomass; C: catch; DN: double normal; Log: logistic

and 2020<sup>3</sup> was input into the models and all models could support these catches.

### Operating model

#### Key quantities exploration

Given the selection objectives presented above, from the 108 models 41 were excluded, and therefore 67 models are considered to constitute the operating model. Two key quantities were explored that represent scale ( $SSB_0$ ) and final year stock status ( $SSB_{2018}/SSB_{MSY}$ ) (Fig. 5). An ANOVA was used to test for differences in  $SSB_0$  and  $SSB_{2018}/SSB_{MSY}$  between each considered uncertainty axis. It is noted that virgin spawning stock biomass ( $SSB_0$ ) is mostly influenced by mortality, steepness, sigmaR, CPUE scaling

and selectivity (Table 2), while the status of the stock ( $SB_{2018}/SSB_{MSY}$ ) is being mostly affected by mortality, steepness, ESS, LLQ, CPUE scaling and selectivity (Table 2).

#### Time series of the OM

The overall time series plot of the final OM (67 runs) shows values for abundance and fishing mortality to be widely distributed around the stock assessment. The operating model covers and expands the range of these uncertainties in comparison with the stock assessment (Fig. 6), being somewhat more pessimistic. The SSB estimates from the stock assessment tend to be on the higher end of the SSB estimates from the OM grid, while the F estimates from the stock assessment tend to be on the lower end of the OM grid.

Similarly to the stock assessment, most OM models show a stock status in the final year ( $SSB_{2018}/SSB_{MSY}$ ) higher than 1, meaning that stock is not

<sup>3</sup> Catch data available at <https://iotc.org/WPB/20/Data/03-NC>.

**Table 2** Deviance tables of ANOVAs to test the effects of the various uncertainty parameters on virgin spawning stock biomass ( $SSB_0$ ) and  $SSB$  in 2018 relative to  $SSB$  at maximum sustainable yield ( $SSB_{2018}/SSB_{MSY}$ )

For each parameter it is indicated the degrees of freedom (Df), the sum of squares (Sum sq), the mean sum of squares (Mean Sq), the F-test statistic and the significance (*p-value*). *P*-values in bold represent significant differences at  $P < 0.05$

	Df	Sum Sq	Mean Sq	F stat	<i>P</i> -value
<i>Response: SSB<sub>0</sub></i>					
M	2	$1.00 \times 10^{11}$	$5.02 \times 10^{10}$	51.83	<b>&lt; 0.001</b>
Steepness	2	$1.27 \times 10^{10}$	$6.34 \times 10^9$	6.55	<b>&lt; 0.05</b>
sigmaR	1	$3.92 \times 10^{10}$	$3.92 \times 10^{10}$	40.43	<b>&lt; 0.001</b>
ESS	1	$3.49 \times 10^9$	$3.49 \times 10^9$	3.60	> 0.05
LLQ	1	$1.01 \times 10^8$	$1.01 \times 10^8$	0.10	> 0.05
Growth and maturity	1	$1.89 \times 10^9$	$1.89 \times 10^9$	1.96	> 0.05
CPUE	2	$2.08 \times 10^9$	$1.04 \times 10^9$	1.07	> 0.05
CPUE scaling	2	$2.70 \times 10^{10}$	$1.35 \times 10^{10}$	13.93	<b>&lt; 0.001</b>
LLsel	1	$4.95 \times 10^{10}$	$4.95 \times 10^{10}$	51.10	<b>&lt; 0.001</b>
<i>Response: SSB<sub>2018</sub>/SSB<sub>MSY</sub></i>					
M	2	1.784	0.892	14.62	<b>&lt; 0.001</b>
Steepness	2	8.336	4.168	68.31	<b>&lt; 0.001</b>
sigmaR	1	0.015	0.015	0.24	> 0.05
ESS	1	0.847	0.847	13.88	<b>&lt; 0.001</b>
LLQ	1	0.865	0.865	14.18	<b>&lt; 0.001</b>
Growth and maturity	1	0.006	0.006	0.10	> 0.05
CPUE	2	0.004	0.002	0.04	> 0.05
CPUE scaling	2	0.405	0.203	3.32	<b>&lt; 0.05</b>
LLsel	1	2.077	2.077	34.03	<b>&lt; 0.001</b>

overfished. However, the OM grid does provide additional scenarios where the stock is overfished ( $SSB_{2018}/SSB_{MSY} < 1$ ) and where overfishing ( $F_{2018}/F_{MSY} > 1$ ) is occurring. Those additional uncertainties are not accounted for in the current stock assessment grid (Fig. 7).

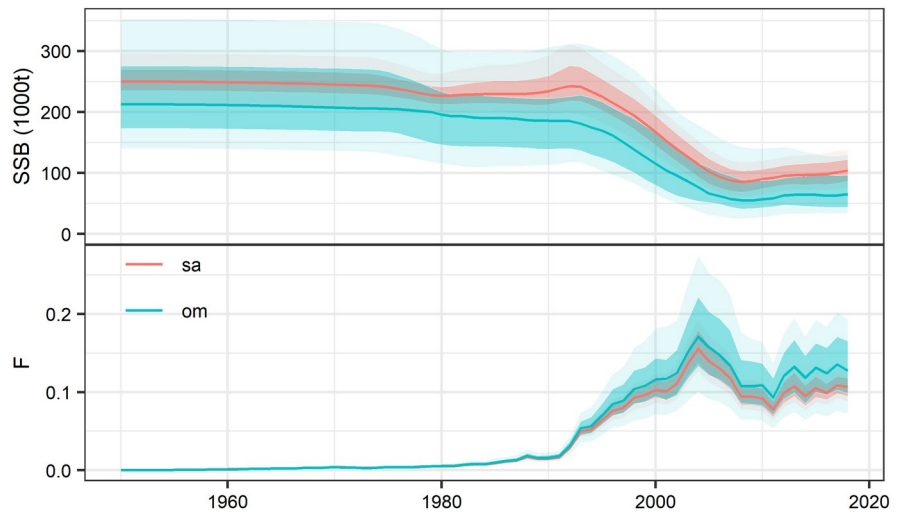
## Discussion and conclusions

Potential advantages on the use of MSE have been previously described (Butterworth 2008; Hillary et al. 2016; Punt et al. 2016; Miller et al. 2019; Holmes and Miller 2022). These advantages include: 1) explicit incorporation of uncertainty and natural variability into management; 2) management objectives that are set at the onset of the process leading to a more transparent decision-making process when setting TACs; 3) management that responds to stocks growth and decline in a suitable manner; 4) improved involvement of stakeholders contributing for a better implementation and the possibility to move from the assessment management cycle allowing both scientists and managers to focus on other scientific or management issues. Additionally, the design of

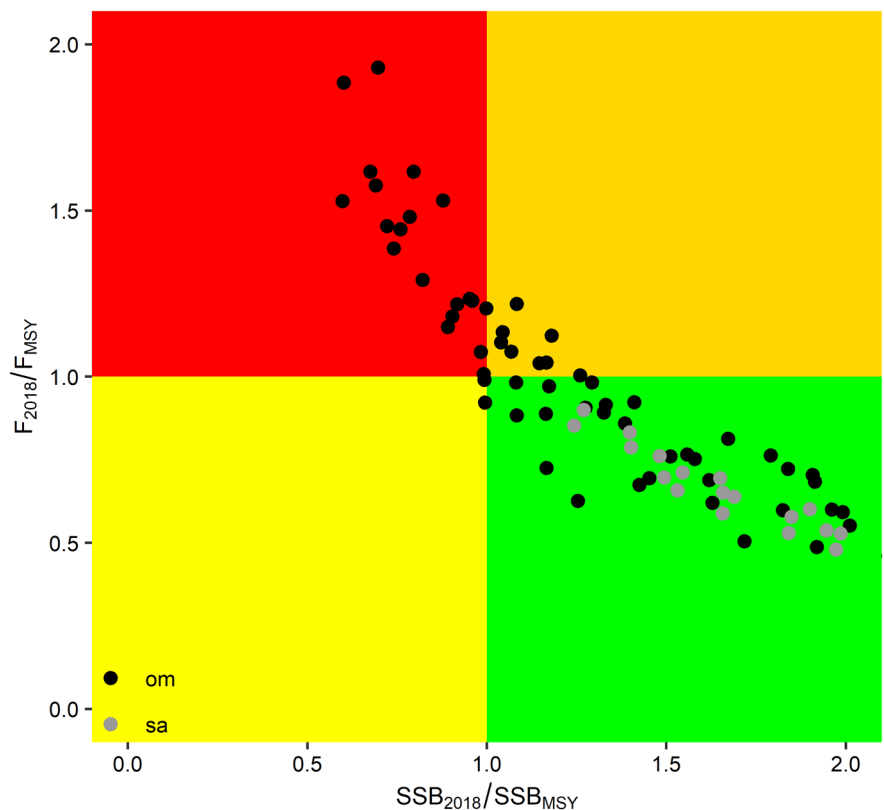
the OMs, either from the reference set or robustness tests with extreme scenarios, can help guide further research regarding its associated uncertainty. From a fisheries perspective, having an adopted HCR or MP allow fisheries to have a higher likelihood of obtaining seafood sustainability certifications, as is the case for some skipjack fisheries in the Indian Ocean which have obtained the Marine Stewardship Council certification (e.g. Maldives pole & line skipjack tuna (MSC 2023); Echebstar Indian Ocean purse seine skipjack tuna fishery (MSC 2024)), which can provide a competitive advantage over non-certified fisheries.

De Moor et al. (2022) argue that even in the case of declaration of exceptional circumstances, i.e. conditions that were not tested at the time of MP development, an adopted MP can still have more benefits than the conventional assessment approach and provide some recommendations from the lessons learned in the 30 years' experience of MSE in South Africa. A successful case of MSE in the t-RFMO world is the southern bluefin tuna, where an MP was adopted in 2011 (and revised in 2019) with the aim of rebuilding the stock; since the implementation the stock status has improved and catch limits have been increasing (CCSBT 2023). Besides southern

**Fig. 6** Time series of spawning stock biomass (SSB in 1000 tonnes) and fishing mortality (F) for the Indian Ocean swordfish, from the stock assessment model ensemble (sa, in red), and the 67 model runs that compose the operating model (om, in blue). The solid lines represent the median, the darker shades the 75% quantiles and the light shades the 95% quantiles



**Fig. 7** Kobe plot showing current spawning stock biomass (SSB) relative to SSB capable of producing maximum sustainable yield (MSY) ( $SSB/SSB_{MSY}$ ) and fishing mortality (F) relative to F capable of producing MSY ( $F/F_{MSY}$ ) for the 24 stock assessment models (sa, grey) and the 67 operating model runs (om, black). The Kobe plot is divided into four quadrants divided at  $SSB = SSB_{MSY}$  and  $F = F_{MSY}$ , graphically representing whether the stock is overfished (values relative to  $SSB/SSB_{MSY}$ ) and/or subject to overfishing (values relative to  $F/F_{MSY}$ )



bluefin tuna, adoption of MP in t-RFMOs is relatively recent. Implementation of MPs in IOTC, IATTC and WCPFC is at more incipient stages and in ICCAT two MPs have been implemented in 2021 and 2023 (Table 3). For the Atlantic, Northern albacore is in good stock status and TACs have been increasing,

while bluefin tuna is at its first management cycle this has led to an increase of the TAC for the eastern stock and a maintenance of TAC in the western. Stability has also been considered in all adopted MPs which include constraints in TAC or effort, depending on the management output provided by the MP, allowing for

gradual changes in case there is the need to increase or reduce the fishing opportunities.

As MSE processes are on the margin between science and policy, its development depends not only on the technical development by scientists but also on the involvement of stakeholders, as many decisions in this process are related to management (Miller et al. 2019). The same authors highlight the importance of the interactions and communication between scientists and managers for effective and successful MSE processes. The yield and status of IOTC mandatory species is monitored against a set of predetermined benchmarks, and one of the primary management objectives is to maintain stocks at or above target levels and avoid breaching the limited reference point with high probability. In a binding resolution, the IOTC recognizes that the management strategy evaluation framework is the best available tool to identify pathways to achieving management goals. To date, management advice for most IOTC stocks still relies on routine stock assessments, and agreement on the best model is difficult to reach, especially for stocks with a high degree of uncertainty, sometimes eroding confidence in the management advice derived from the assessment and in which case members may opt out of management measures. The Commission has supported using MSE as a tool to assess conservation and management strategies since 2010, and it has invested a significant amount of resource in developing species-specific MSE.

In IOTC, the Technical Committee on Management Procedures (TCMP) was established in 2016 to facilitate the dialogue between scientists and decision-makers to enhance the response by the Commission regarding the MSEs in development (IOTC 2016b). The roles of scientists and managers were defined by TCMP in 2017 (IOTC 2017b), with scientists having the tasks of identifying uncertainties, conditioning the OM and test of MP and presenting results, while managers are tasked to define management objectives, performance statistics and selecting MPs. Since 2016, an HCR control has been used to set the yearly catch limit for skipjack tuna, and in 2022, a full management mechanism was put in place for bigeye tuna. In these situations, the role of stock assessment has changed from advising on catches to assessing stock status and determining whether any extraordinary events have occurred that would render the management procedure invalid. Moving from the

stock assessment cycles and the requirement to meet the TAC output by the BET MP has prompted the IOTC Commission to dedicate more time to allocation criteria determination, which has been until now lacking. Absence of allocation criteria has led to compliance issues with skipjack tuna HCR, with catches exceeding the total TAC for several years. The Commission has learned from this and for BET a temporary allocation was set while concurrently developing the allocation criteria. Ultimately, having an allocation scheme will enhance transparency in TAC setting since it has been pre-agreed upon, eliminating the need to revisit discussions for TAC settings at each management cycle. For swordfish significant advances have been made on the technical development of the operating model as shown in this work, but also on the decision-making, with the definition of interim reference points (IOTC 2015), tuning objectives and constrains to be applied in the MP (IOTC 2021).

On the technical development, the operating model of the management strategy evaluation of swordfish in the Indian Ocean is being developed through a grid design that expands the model ensemble from the latest stock assessment to cover a wider range of uncertainties. As in our case, most MSEs being conducted in t-RFMOs are developed using grid designs based on the accepted stock assessment for the species (Sharma et al. 2020). Although this is currently the most common approach, some examples exist where this is not the case. For example, for bluefin tuna in ICCAT a tailor-made model was developed (Carruthers et al. 2016) and for skipjack in IOTC the OM used to test the currently adopted HCR is constructed from a combination of life-history and stock assessment informed priors using the feasible stock trajectories algorithm (Bentley and Adam 2016). Work is also ongoing on the application of Approximate Bayesian Computation methods to condition OMs that are decoupled from the stock assessment model for the Indian Ocean albacore stock (IOTC 2022b).

Hillary et al. (2021) noted that having a high dependence on the stock assessment model can be problematic if there is no accepted stock assessment, or if an adequate suite of OMs based on the stock assessment cannot be established. Additionally, having an OM that is based on the stock assessment may lead to an inclination to update the OM each time the stock assessment is conducted. A clear set of rules should be established to assess the necessity of

**Table 3** Summary table of management objectives, adoption year, implementation state or, when available, effects on status and yield for adopted management procedures (MP) tested under management strategy evaluation (MSE) in each of the tuna Regional Fisheries Management Organizations (t-RFMOs)

CCSBT	Southern Bluefin tuna <i>Thunnus maccoyii</i>	Management objectives	Interim rebuilding target of 20% SSB <sub>0</sub> by 2035, with 70% probability	Achieving 30% SSB <sub>0</sub> by 2035, with 50% probability	1,2	
		Constraint on TAC	Maximum 3000 t increase			
		Adoption	2011/2019			
IOTC	Bigeye tuna <i>Thunnus obesus</i>	Effect on status	Increase from 5% SSB <sub>0</sub> in 2010 to 20% SSB <sub>0</sub> in 2020			
		Effect on yield	TAC remained constant or increased at each management cycle (3 years)			
		Management objectives	Maintain the stock as not overfished and no overfishing occurring while maximizing the average catch from the fishery and reducing the variation in the TAC between management periods	Achieving the target reference point of SSB <sub>MSY</sub> by 2034–2038, with 60% probability;	Avoid breaching the interim limit reference point of 0.5 SSB <sub>MSY</sub> with a high probability;	TAC change between management period should be no more than 15% increase or decrease
		Constraint on TAC	Maximum 15% increase or decrease			
		Adoption	2022			
	Implementation	First implementation of TAC by MP in 2024. TAC for 2024–2025 represents a 15% decrease relative to previous TAC				

**Table 3** (continued)

ICCAT	North Atlantic Albacore <i>Thunnus alalunga</i>	Management objectives	Maintain stock as not overfished and no overfishing occurring, with at least 60% probability while maximizing long-term yield;	Rebuild to or be above SSB <sub>MSY</sub> with at least 60% probability within as short a time as possible, while maximizing average catch and minimizing TAC changes	4
		Constraint on TAC	Maximum 25% increase or 20% decrease. Maximum catch limit set to 50,000t		
		Adoption	2017/2021 (First adoption of an HCR in 2017 and converted to MP in 2021)		
		Effect on status	Stock is estimated not to be overfished and no overfishing occurring with 99% probability		
		Effect on yield	TAC for 2024–2026 increased relative to first management cycle (2022–2023). Increase was limited by the 25% maximum change clause included in the MP		
	Bluefin tuna <i>Thunnus thynnus</i>	Management objectives	Both the western and eastern stocks should not be overfished and no overfishing should be occurring, with 60% probability	There should be a 15% or less probability of either stock falling below the biomass limit reference point	5
		Constraint on TAC	Maximum 20% increase or 35% decrease	Any change in TAC between consecutive management periods in both the western and eastern management areas should be no more than a 20% increase or a 35% decrease	
		Adoption	2022		
		Implementation	First implementation of TAC by MP in 2023. TAC for 2023–2025 represents an increase relative to previous TAC for eastern stock and a maintenance of TAC for the western stock		
WCPFC	Skipjack <i>Katsuwonus pelamis</i>	Management objectives	Maintaining the spawning potential depletion ratio on average at a level consistent with the target reference point	Maintaining the spawning potential depletion ratio above the limit reference point with a risk of the limit reference point being breached no greater than 20%	6
		Constraint on effort	Maximum 10% increase or decrease	Relative stability in fishing levels between management and in the longer term	
		Adoption	2022		
		Implementation	Interim Management Procedure adopted on a trial basis		

**Table 3** (continued)

IATTC/WCPFC	Albacore <i>Thunnus alalunga</i>	Management objectives	Maintain SSB above the Limit Reference Point, with a probability of at least 80% over the next 10 years	Maintain depletion of total biomass around historical (2006–2015) average depletion over the next 10 years	Maintain fishing intensity (F) at or below the target reference point with a probability of at least 50% over the next 10 years	To the extent practicable, management changes (e.g., catch and/or effort) should be relatively gradual between years	7.8
		Constraint on TAC/effort	Maximum 20% increase or decrease in catch or effort				
		Adoption	2023				
		Implementation	Work will be conducted in 2024 to advise how fishing intensity should be interpreted to actual management under this harvest strategy				

1: Commission for the Conservation of Southern Bluefin Tuna (CCSBT) (2019) Resolution on the Adoption of a Management Procedure. Twenty Sixth Annual Meeting of the Commission, Cape Town, South Africa, 14–17 October 2019

2: Commission for the Conservation of Southern Bluefin Tuna (CCSBT) (2023) Report of the Twenty Eighth Meeting of the Scientific Committee. Jeju island, Republic of Korea, 1 September 2023

3: Indian Ocean Tuna Commission (IOTC) (2022a) Resolution 22/03—On a management procedure for bigeye tuna in the IOTC area of competence. 26th Session of the Indian Ocean Tuna Commission, Seychelles, 16–20 May 2022

4: International Commission for the conservation of Atlantic Tunas (ICCAT) (2021) Recommendation 21-04—Recommendation by ICCAT on Conservation and Management Measures, Including a Management Procedure and Exceptional Circumstances Protocol, for North Atlantic Albacore. 27th Regular meeting of the Commission, online, 15–23 November 2021

5: International Commission for the conservation of Atlantic Tunas (ICCAT) (2022) Recommendation 22-09—Recommendation by ICCAT establishing a management procedure for Atlantic bluefin tuna to be used for both the Western Atlantic and Eastern Atlantic and Mediterranean management areas. 23rd Special Meeting of the Commission, Vale do Lobo/hybrid, 13–21 November 2022

6: Western and Central Pacific Fisheries Commission (WCPFC) (2022) CMM 2022-01—Conservation and Management Measure on a Management Procedure for WCPO Skipjack Tuna. Commission 19th regular session, Da Nang City, Vietnam, 28 November–3 December 2022

7: Inter-American Tropical Tuna Commission (IATTC) (2023) Resolution C-23-02—Amendment to harvest strategy for North Pacific albacore in the Eastern Pacific Ocean. 101th Inter-American Tropical Tuna Commission, Victoria, B.C., Canada 7–11 August 2023

8: Western and Central Pacific Fisheries Commission (WCPFC) (2023) Provisional Meeting Outcomes and Attachments. Commission 20th regular session, Rarotonga, Cook Islands, 4–8 December 2023

$SSB_0$  virgin spawning biomass,  $SSB_{MSY}$  spawning stock biomass that produces the maximum sustainable yield (MSY); TAC total allowable catch; HCR harvest control rule

revising the OM when an updated assessment is conducted (Sharma et al. 2020). For example, Preece and Willams (2022) consider that the stock assessment estimates of key reference points should be within the 90% probability interval of the projections of the operating model, as a test for evidence of exceptional circumstances in relation to new information on the status of the stock that may warrant a revision of the MSE/MP. Nevertheless, and as stated by Hillary et al. (2021), having an OM that is based on an accepted stock assessment as is the case presented in this work, does have some advantages: i) highly flexible stock assessment packages being already available, ii) a plausible model has already been fitted to the available data with estimates of key parameters for OM projections, and iii) there is familiarity and understanding within the relevant fora of the stock assessment.

In the case of Indian Ocean swordfish, scientific management advice was based on a 24-model grid which explored several uncertainties (Fu 2020; IOTC 2020) and was used as a base for the operating model as a consequential action. The OM herein described leads, in general, to a more pessimistic view of the stock status compared to the results of the stock assessment ensemble. Although the results of the stock assessment are in the envelope of values of the OM, those tend to be more optimistic, with all models from the stock assessment resulting in a stock that is not overfished ( $SSB_{2018}/SSB_{MSY} > 1$ ) nor suffering from overfishing ( $F_{2018}/F_{MSY} < 1$ ). The fact that the OM developed in this MSE framework provides not only scenarios identical to the stock assessment, but also some others that are more pessimistic, is due to the levels of the uncertainty variables that have been included, and their combinations, leading to scenarios of a less productive stock that is more susceptible to fishing pressure. This will probably lead to a management procedure needing to be more conservative to achieve management objectives, for example, having lower total allowable catches to keep or bring the stock into the not overfished nor overfishing area with a greater probability.

It can be challenging to decide which uncertainties should be included in an MSE because doing so could result in underrepresenting those uncertainties and inadequate testing of the chosen MP (Kolody et al. 2019). However, in practice it can be impractical to manage all uncertainties and, in those cases, the most plausible and most consequential for the dynamics

should be chosen (Punt et al. 2016). Five sources of uncertainty can be considered to be included in an MSE: i) process uncertainty, ii) parameter uncertainty, iii) model uncertainty, iv) implementation uncertainty and, v) observation uncertainty (Punt et al. 2016). In the present case study, the uncertainties considered for the OM conditioning can be classified as model uncertainty (other forms of uncertainty are being considered at other stages of the MSE development, e.g. implementation uncertainty in future catches) and expanded from the assessment in the number of considered factors. Specifically, four factors were considered in the stock assessment ensemble vs nine in the OM ensemble, resulting in a large grid if considering the full factorial design. Most commonly the OMs are based on running the main effects, i.e. choosing a “base” case and changing one factor at a time (Punt et al. 2016), but partial factorial designs have also been applied previously (Schweder et al. 1998; Kolody and Jumppanen 2021), and this is also the case of the Indian Ocean swordfish MSE presented here.

Setting up a grid with all factors, levels and their interactions provides a wider uncertainty range and is simpler to construct but might lead to models with unlikely combinations of parameters. Consideration to the correlations between factors could also be considered to exclude runs with such unlikely combinations. Analyzing the effect of each uncertainty in key parameters that are likely to influence MP could also be applied to reduce the model runs in the reference set, i.e. a set of models that has a higher plausibility and/or influence in MP performance. Less influential uncertainties could be used as “robustness tests”, i.e. a set of models that are less likely or less influential in MP performance but to which the MP should also be robust too. In the present case, univariate analysis of variance was conducted to determine the influence of the uncertainties in scale ( $SSB_0$ ) and final year stock status ( $SSB_{2018}/SSB_{MSY}$ ). It should be noted that to run this analysis a large or larger number of models still must be run to assess the effect of each uncertainty.

In our case study the most influential factors were the values of natural mortality, steepness, shape of the selectivity curve, and the scaling of the CPUEs between areas which affected both the scale and the stock status in the final year. The least influential factors were the choice of CPUE and growth and

maturity, which had no significant impact in either scale or stock status. The effective sample size of the length frequency data and the catchability increase was found to affect the stock status but not the scale of the stock, while recruitment variability had the opposite effect, by affecting the scale of the stock but not its status in the initial year of the projections. Similar results were found for North Atlantic swordfish regarding mortality, steepness, and catchability increase; however for recruitment variability there were contrasting results (Hordyk et al. 2021). In both cases, higher mortality rates and steepness values had lower estimates of  $SSB_0$  and higher estimates of stock status, and increasing catchability did not influence the scale of the stock but resulted in lower estimates of stock status. Regarding recruitment variability, this did not influence the scale of the stock in the case of the North Atlantic swordfish, while increasing the recruitment variability in the Indian Ocean swordfish lead to an increase in  $SSB_0$ . This is important to the evaluation of MP performance due to the significant asymmetries in risk associated with common uncertainties in population models (Hordyk et al. 2019).

In a revision of operating model development in t-RFMOs, Sharma et al. (2020) noted that in several MSEs, filtering was performed to exclude models that produced implausible dynamics, but a standardized procedure is still missing. The joint t-RFMO MSE working group emphasized that the OMs need to be conditioned adequately, and that standard automated model fit diagnostics should be applied to ensure the consistency between the model and the data (Anonymous 2018). With this regard, a 4-step filtering framework is presented in this work using Indian Ocean swordfish as a case study, where automated diagnostic tools for Stock Synthesis were run to exclude operating models that did not meet the validation criteria. Future work could also be developed for the application of OM weighting scheme based on the prediction skill of the CPUEs as applied in Mosqueira and Brunel (2022). It is expected that further changes to the OM can still occur, however the procedures for OM validation and filtering can still be of use for this, and other species, MSE development.

**Acknowledgements** D. Rosa is supported by an FCT Doctoral grant (Ref: SFRH/BD/136074/2018) funded by National Funds and the European Social Fund. R. Coelho and D. Rosa (CCMAR) receives partial support from FCT through projects UIDB/04326/2020, UIDP/04326/2020 and LA/P/0101/2020.

Thanks also to IOTC for providing support to attend IOTC MSE workshops. I. Mosqueira was supported by FAO/IOTC to work on swordfish MSE.

**Funding** Open access funding provided by FCTIFCCN (b-on).

**Data availability** Datasets are available on GitHub at <https://github.com/iotcwpm/SWO>.

**Declarations**

**Conflict of interest** The authors declare that they have no competing interests.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## References

- Anonymous. (2018) Report of the 2018 joint tuna RFMO management strategy evaluation working group meeting. In: The 2018 Joint tuna RFMO management strategy evaluation working group meeting. Seattle, USA–13–15 June 2018
- Bentley N, Adam MS (2016) Management strategy evaluation for the Indian Ocean skipjack tuna fishery. IOTC-2016-WPM07-15 Rev\_1
- Bunnefeld N, Hoshino E, Milner-Gulland EJ (2011) Management strategy evaluation: a powerful tool for conservation? *Trends Ecol Evol* 26(9):441–447
- Butterworth DS (2008) Some lessons from implementing management procedures. In: Tsukamoto K, Kawamura T, Takeuchi T, Beard Jr. TD, Kaiser MJ (eds) Fisheries for global welfare and environment, 5th World Fisheries Congress 2008. TERRAPUB, Tokyo, pp 381–397
- Carruthers T, Kimoto A, Powers JE, Kell L, Butterworth DS, Lauretta MV, Kitakado T (2016) Structure and estimation framework of Atlantic Bluefin Tuna operating models. *Collect Vol Sci Pap ICCAT* 72(7):1782–1795
- Carvalho F, Winker H, Courtney D, Kapur M, Kell L, Cardinale M, Schirripa M, Kitakado T, Yemane D, Piner KR, Maunder MN, Taylor I, Wetzel CR, Doering K, Johnson KF, Methot RD (2021) A cookbook for using model

- diagnostics in integrated stock assessments. *Fish Res* 240:105959
- Commission for the Conservation of Southern Bluefin Tuna (CCSBT) (2019) Resolution on the Adoption of a Management Procedure. Twenty Sixth Annual Meeting of the Commission, Cape Town, South Africa, 14–17 October 2019
- Commission for the Conservation of Southern Bluefin Tuna (CCSBT) (2023) Report of the twenty eighth meeting of the scientific committee. Jeju Island, Republic of Korea, 1 September 2023
- de Moor CL, Butterworth DS, Johnston S (2022) Learning from three decades of management strategy evaluation in South Africa. *ICES J Mar Sci* 79:1843–1852
- Ducharme-Barth ND, Vincent M (2022) Focusing on the front end: a framework for incorporating uncertainty in biological parameters in model ensembles of integrated stock assessments. *Fish Res* 260:106452
- Farley JH, Clear NP, Krusic-Golub K, Evenson P, Young JW (2022) A bone to pick with age estimation using hard parts: a case study of swordfish, *Xiphias gladius*, in the southwest Pacific Ocean. *Fish Res* 254:106413
- Fu D (2020) Preliminary Indian Ocean swordfish stock assessment 1950–2018 (Stock Synthesis). IOTC–2020–WPB18–16
- Hillary RM, Preece AL, Davies CR, Kurota H, Sakai O, Itoh T, Parma AM, Butterworth DS, Ianelli J, Branch TA (2016) A scientific alternative to moratoria for rebuilding depleted international tuna stocks. *Fish Fish* 17(2):469–482
- Hillary RM, Edwards C, Preece AL, Mosqueira I, Preece A, Rosa D, Williams A (2021) Strategies for the conditioning of operating models for IOTC stocks. IOTC–2021–WPM12–16
- Holmes G, Miller S (2022) Harvest strategies and allocation—a chicken and egg scenario? *Mar Policy* 135:104871
- Hordyk A, Huynh Q, Carruthers T (2019) Misspecification in stock assessments: common uncertainties and asymmetric risks. *Fish Fish* 20:888–902
- Hordyk A, Schirripa M, Rosa D (2021) Updates to the operating model uncertainty for the North Atlantic swordfish MSE. *Collect Vol Sci Pap ICCAT* 78(7):240–261
- Hoyle SD, Langley AD (2020) Scaling factors for multi-region stock assessments, with an application to Indian Ocean tropical tunas. *Fish Res* 228:105586
- Hyndman RJ, Koehler AB, Methot RD, Wetzell CR (2006) Another look at measures of forecast accuracy. *Int J Forecast* 22(4):679–688
- Indian Ocean Tuna Commission (IOTC) (2015). Resolution 15/10 on target and limit reference points and a decision framework. 19th Session of the Indian Ocean Tuna Commission, La Reunion, France, 27 April–01 May 2015
- Indian Ocean Tuna Commission (IOTC) (2016a). Resolution 16/02 on Harvest Control Rules for Skipjack Tuna in the IOTC Area of Competence. 20th Session of the Indian Ocean Tuna Commission, La Reunion, France, 23–27 May 2016
- Indian Ocean Tuna Commission (IOTC) (2016b). Resolution 16/09 on establishing a technical committee on management procedures. 20th Session of the Indian Ocean Tuna Commission, La Reunion, France, 23–27 May 2016
- Indian Ocean Tuna Commission (IOTC) (2017a) Report of the 8th session of the IOTC working party on methods. IOTC–2017–WPM08–R[E]
- Indian Ocean Tuna Commission (IOTC) (2017b) Chair report of the 1st IOTC technical committee on management procedures. IOTC–2017–TCMP01–R[E]
- Indian Ocean Tuna Commission (IOTC) (2020) Report of the 18th session of the IOTC working party on Billfish. IOTC–2020–WPM18–R[E]
- Indian Ocean Tuna Commission (IOTC) (2021) Report of the 4th session of the technical committee on management procedures. IOTC–2021–TCMP04–R[E]
- Indian Ocean Tuna Commission (IOTC) (2022a) Resolution 22/03—on a management procedure for bigeye tuna in the IOTC area of competence. 26th Session of the Indian Ocean Tuna Commission, Seychelles, 16–20 May 2022
- Indian Ocean Tuna Commission (IOTC) (2022b) Report of the 13th session of the IOTC working party on methods. IOTC–2022–WPM13–R[E]
- Inter-American Tropical Tuna Commission (IATTC) (2023) Resolution C-23-02—amendment to harvest strategy for North Pacific albacore in the Eastern Pacific Ocean. 101th Inter-American Tropical Tuna Commission, Victoria, B.C., Canada 7–11 August 2023
- International Commission for the conservation of Atlantic Tunas (ICCAT) (2021) Recommendation 21-04—recommendation by ICCAT on conservation and management measures, including a management procedure and exceptional circumstances protocol, for North Atlantic Albacore. 27th Regular meeting of the Commission, online, 15–23 November 2021
- International Commission for the conservation of Atlantic Tunas (ICCAT) (2022) Recommendation 22-09—recommendation by ICCAT establishing a management procedure for Atlantic bluefin tuna to be used for both the Western Atlantic and Eastern Atlantic and Mediterranean management areas. 23rd Special Meeting of the Commission, Vale do Lobo/hybrid, 13–21 November 2022
- International Council for the Exploration of the Sea (ICES) (2020) The third workshop on guidelines for management strategy evaluations (WKG MSE3). *ICES Scientific Reports*. 2:116. p 112
- Jardim E, Azevedo M, Brodziak J, Brooks EN, Johnson KF, Klibansky N, Millar CP, Minto C, Mosqueira I, Nash R, Vasilakopoulos P, Wells BK (2021) Operationalizing ensemble models for scientific advice to fisheries management. *ICES J Mar Sci* 78:1209–1216
- Kell LT, Mosqueira I, Grosjean P, Fromentin J-M, Garcia D, Hillary R, Jardim E, Mardle S, Pastoors MA, Poos JJ, Scott F, Scott RD (2007) FLR: an open-source framework for the evaluation and development of management strategies. *ICES J Mar Sci* 64:640–646
- Kell LT, Kimoto A, Kitakado T (2016) Evaluation of the prediction skill of stock assessment using hind-casting. *Fish Res* 183:119–127
- Kolody DS, Paige Evenson J, Preece AL, Davies CR, Hillary RM (2019) Recruitment in tuna RFMO stock assessment

- and management: a review of current approaches and challenges. *Fish Res* 217:217–234
- Kolody, DS, Jumppanen P (2021) Indian Ocean Bigeye Tuna management procedure evaluation update. March 2021. IOTC-2021-WPM12(MSE)-04
- Lawson J (2015) Design and analysis of experiments with R. Chapman and Hall/CRC, New York
- Lorenzen K (1996) The relationship between body weight and natural mortality in juvenile and adult fish: a comparison of natural eco-systems and aquaculture. *J Fish Biol* 49:627–642
- Mace PM, Doonan IJ (1988) A generalized bioeconomic simulation model for fish dynamics. *N Z Fish Assess Res Doc* 88(4):1988p
- Maunder MN, Crone PR, Punt AE, Valero JL, Semmens BX (2016) Growth: theory, estimation, and application in fishery stock assessment models. *Fish Res* 180:1–3
- Methot RD, Wetzel CR (2013) Stock synthesis: a biological and statistical framework for fish stock assessment and fishery management. *Fish Res* 124:86–99
- Miller SK, Anganuzzi A, Butterworth DS, Davies CR, Donovan GP, Nickson A, Rademeyer RA, Restrepo V (2019) Improving communication: the key to more effective MSE processes. *Can J Fish Aquat Sci* 76(4):643–656
- Mosqueira I, Brunel T (2022) Updates on development of MSE analyses for Indian Ocean albacore tuna and swordfish. IOTC-2022-WPM13-08
- Mosqueira I, Rosa D, Fu D, Coelho R (2017) Update on the conditioning of an operating model for the Indian Ocean swordfish stock. IOTC-2017-SC20-11
- MSC (2023) Maldives pole & line skipjack tuna final draft report. Marine Stewardship Council. p 212
- MSC (2024) Echebstar Indian Ocean purse seine skipjack tuna fishery final draft report. Marine Stewardship Council. p 481
- Northwest Atlantic Fisheries Organization (NAFO) (2021) Report of the NAFO joint commission-scientific council working group on risk-based management strategies (WGRBMS) meeting, online, 24–26 August 2021
- Peerce AL, William A (2022) Consideration of exceptional circumstances for the Bigeye Tuna MP 2022. IOTC-2022-WPTT24-12
- Punt AE, Donovan G (2007) Developing management procedures that are robust to uncertainty: lessons from the International Whaling commission. *ICES J Mar Sci* 64:603–612
- Punt AE, Butterworth DS, de Moor CL, De Olivera JAA, Haddon M (2016) Management strategy evaluation: best practices. *Fish Fish* 17:303–334
- Punt AE, Castillo-Jordán C, Hamel OS, Cope JM, Maunder MN, Ianelli JN (2021) Consequences of error in natural mortality and its estimation in stock assessment models. *Fish Res* 233:105759
- R Core Team (2022) R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL <https://www.R-project.org/>
- Rosa D, Mosqueira I, Fu D, Coelho R (2018) Update on the conditioning of an operating model for the Indian Ocean Swordfish Stock. IOTC-2018-SC21-12
- Sampson DB (2014) Fishery selection and its relevance to stock assessment and fishery management. *Fish Res* 158:5–14
- Schweder T, Hagen GS, Hatlebakk E (1998) On the effect on cod and herring fisheries of retuning the revised management procedure for minke whaling in the greater Barents Sea. *Fish Res* 37:77–95
- Sharma R, Levontin P, Kitakado T, Kell LT, Mosqueira I, Kimoto A, Scott R, Minto-Vera C, De Bruyn P, Ye Y, Kleineberg J, Walton JL, Miller S, Magnusson A (2020) Operating model design in tuna regional fishery management organizations: current practice, issues and implications. *Fish Fish* 21:940–961
- Wang SP, Chi-Hong L, Chiang WC (2010) Age and growth analysis of swordfish (*Xiphias gladius*) in the Indian Ocean based on the specimens collected by Taiwanese observer program. IOTC-2010-WPB-08
- Western and Central Pacific Fisheries Commission (WCPFC) (2022) CMM 2022-01—conservation and management measure on a management procedure for WCPO Skipjack Tuna. Commission 19th regular session, Da Nang City, Vietnam, 28 November - 3 December 2022
- Western and Central Pacific Fisheries Commission (WCPFC) (2023) Provisional meeting outcomes and attachments. Commission 20th regular session, Rarotonga, Cook Islands, 4–8 December 2023
- Wheeler B (2022) AlgDesign: algorithmic experimental design. R package version 1.2.1, <https://CRAN.R-project.org/package=AlgDesign>
- Winker H, Carvalho F, Cardinale M, Kell LT (2021) ss3diags: cookbook for model diagnostics in stock synthesis. R package version 1.0.8

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.