STOCK ASSESSMENT OF INDIAN OCEAN SWORDFISH (*XIPHIAS GLADIUS*) USING BAYESIAN SURPLUS PRODUCTION MODELS (JABBA): MODEL DEVELOPMENT, VALIDATION, SENSITIVITY AND STOCK STATUS

RUI COELHO[1,2,*]

Discussions: The main purpose of this paper is to inform the WPB that a Bayesian state-space surplus production model (JABBA) has been developed for the Indian Ocean swordfish, as a structurally simpler complement to the integrated age-structured assessment (SS3) and in support of the monitoring of the adopted Management Procedure (MP) for this stock. The analysis comprises a single ocean-wide base case fitted to the total catch and all standardized CPUE series available, a continuity run updating the configuration of previous JABBA assessments, and 4 area-specific models by Indian Ocean region.

Feedback: The author of the paper requests the WPB to provide review and comments on the configuration choices of the base case and the alternative runs made. Comments are also requested on whether the results seem consistent with the perception of stock status using other assessment frameworks, and on any further sensitivity scenarios that should be explored in the future.

Conclusion: The author of the paper seeks confirmation from the WPB that the JABBA models presented are appropriately characterised and presented, and can be used as a complement to the other modelling frameworks used and as support to the monitoring of the adopted Management Procedure (MP).

[1] Portuguese Institute for the Ocean and Atmosphere (IPMA), Av. 5 de Outubro s/n, 8700-305 Olhão, Portugal.
[2] Centro de Ciências do Mar do Algarve (CCMAR/CIMAR LA), Campus de Gambelas, Univ. Algarve, 8005-139 Faro, Portugal.
* Corresponding author: Rui Coelho (rpcoelho@ipma.pt)

Contents

Abstract

This paper provides a stock assessment for Indian Ocean (IOTC) swordfish using Bayesian state-space surplus production models (JABBA). The base case is developed as a single ocean-wide model fitted to the total swordfish catch and using all standardized CPUE series available (13 series), each with its own catchability. An extensive sensitivity analysis was carried out, covering the production function, the productivity priors, the CPUE error assumptions, the process error formulation and the influence of each individual CPUE index, including runs using one CPUE at a time, removing one CPUE at a time, and removing CPUEs for particular areas at a time. A continuity run reproducing the configuration of the last JABBA assessment (5 CPUE series), using updated catches and CPUEs and maintaining consistent with the previous remaining configurations, was also carried out and compared to the new base case. Finally, four area-specific models were fitted and used as a further diagnostic to the base model, assuming that each index used is proportional to and can track total stock biomass. The final base case model estimated that in the final year (2024) the stock is not overfished (B/Bmsy = 1.528) and not subject to overfishing (F/Fmsy = 0.511). Current MSY is estimated at 33,816 t. The final stock status perception tended to be robust across multiple sensitivity analyses carried out and when using the continuity run.

Key-words

Bayesian statistics, Indian Ocean, model diagnostics, stock assessment, surplus production models, swordfish, *Xiphias gladius*.

1. Introduction

The swordfish (*Xiphias gladius*) is a large, highly migratory billfish circumglobally distributed in temperate and tropical waters. In the Indian Ocean it is caught mainly by pelagic longline fleets, mainly on targeted species and to a lower extent as bycatch of other pelagic species fisheries.

Previous stock assessments for Indian Ocean (IOTC) swordfish have been carried out using production models (ASPIC and JABBA) and age-structured integrated statistical models (Stock Synthesis SS3). The last assessment was carried out in 2023 (Xu et al., 2023; Fu, 2023).

Since 2025, the management of swordfish in IOTC has been carried out using a Management Procedure (MP) that has been put in place after a Management Strategy Evaluation (MSE) process. Still, the Working Party on Billfishes (WPB) has scheduled a stock assessment for swordfish during 2026, with the main intent to verify the status of the stock.

The purpose of this paper is to present the configuration, inputs, diagnostics and results of Bayesian surplus production models (JABBA) applied to Indian Ocean swordfish, in support of the provision of management advice for this species to IOTC.

2. Material and methods

2.1. Catch data

The catch data used in this assessment were the raised catch estimates that are prepared by the IOTC Secretariat for swordfish. The dataset is provided at the level of year, month, 5° x 5° spatial area, fleet, gear and fishery, and covers the period 1950 to 2024. Catches are reported in tonnes of live weight.

The base case model is fitted to the total swordfish catches from the entire IOTC area. For the region-specific models, the catches were mapped onto the four IOTC swordfish assessment areas used for the CPUE standardizations, on the basis of the latitude and longitude of the corresponding fishing grounds.

It is noted that in IOTC the swordfish catches were very low until the 1960s, increased slowly until the late 1980s, and then rose very rapidly during the 1990s, which is linked to the rapid expansion of swordfish-directed longline fisheries in that period. The total catch peaked at more than 40,000 t in 2004 and has since declined, averaging around 26,000 t over the last three years of the series (2022–2024), with specifically 27,735 t in the terminal year of 2024 (**Figure 1**).

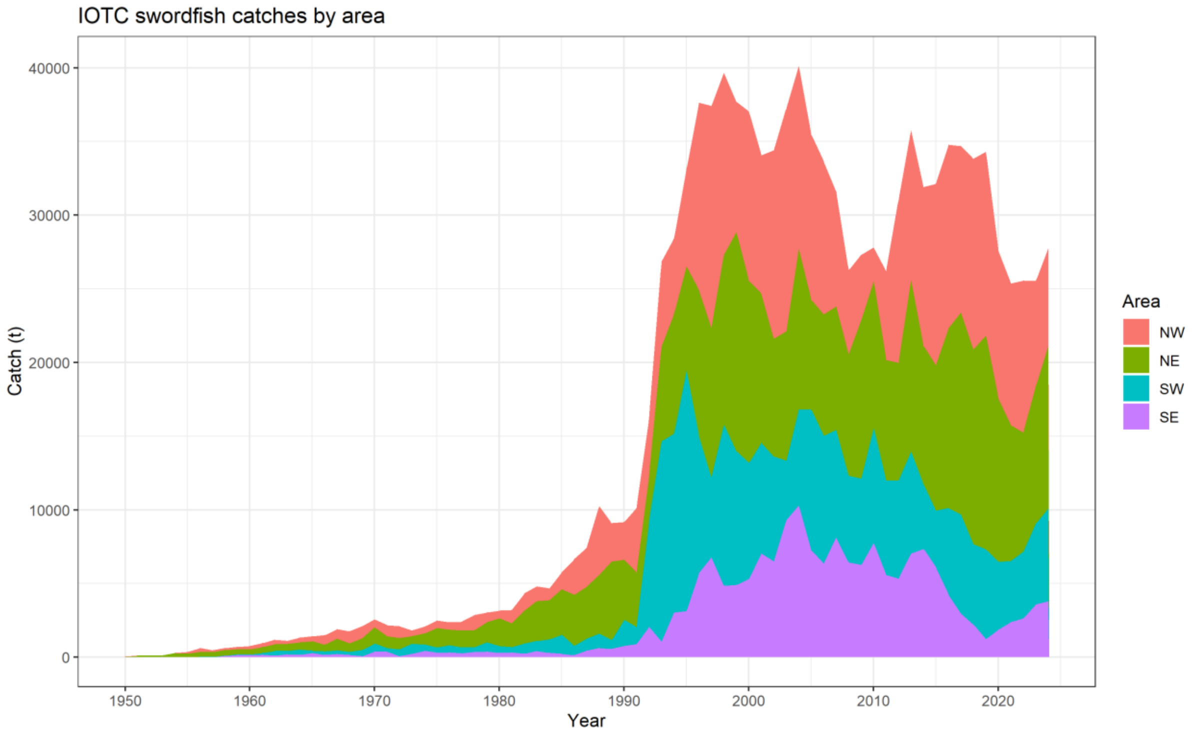

Figure 1: Total raised catch of Indian Ocean swordfish used as the input of the base case model. The figure also shows the area-specific catches, that are used in the area-specific models.

2.2. Productivity

In surplus production models the intrinsic rate of population increase, r, is the parameter that carries the biological information, and it is normally set as an informative prior derived from life history or demographic analysis models.

For this assessment, the base case used the prior adopted in the previous JABBA assessment of Indian Ocean swordfish (Parker, 2020), namely a lognormal prior with log(r) ~ N(log(0.42), 0.4).

Given that the productivity of this stock remains a major axis of uncertainty, two alternative productivity scenarios were also run as sensitivities, centred at r = 0.30 and r = 0.55 with a lower CV of 0.3, spanning the plausible range implied by the base case prior. In addition, a run using a vague uniform range prior on r, between 0.1 and 0.7, was carried out to assess how much of the estimated stock status is driven by the biological prior rather than by the data.

2.3. Standardized CPUE series

The standardized CPUE series considered were those made available for the present 2026 assessment. A total of 13 series are available, from six pelagic longline fleets, and each series is specific to one of the four assessment areas (**Table 1**).

The Japanese and Taiwanese series are available for all four areas, the Indonesian series is specific to the northeast area, and the EU-Spain, EU-Portugal, EU-France (La Réunion) and South Africa series are specific to the southwest area.

All 13 series are used in the base case model, each with its own catchability coefficient. Before being input, each series was scaled by its own mean so that the indices enter the model on a comparable scale.

Overall, there are some consistent patterns on all series, with a decreasing trend until around 2010-2015, following by what looks like an increasing trend in the more recent years (**Figure 2**). When analysed by IOTC area, this overall pattern seems more evident in the SW area and less in the other regions (**Figure 3**).

The coverage of the indices is uneven between areas, given that 6 of the 13 series come from the SW area, 3 from the NE, and only 2 each from the NW and SE areas. Because each index contributes its own term to the likelihood, the SW could therefore influence the fit disproportionately. Two sensitivity analyses were prepared to address this, namely one in which the observation variance is estimated per area rather than per index, and one in which all the indices of a given area are removed together.

A constant CV of 0.20 was assumed for all indices, which gives each index the same nominal weight in the likelihood, irrespective of the precision achieved in the standardization process, and removes any within-series variation in precision over time. Additional observation variance was estimated internally by the model for each index.

Table 1. Summary of the CPUE series available for the 2026 IOTC swordfish stock assessment. The series label is the name used in the figures and results tables along this paper.

Area	CPUE index	Series label	Period
NW	Japan LL	JPN_NW	1994–2024
	Taiwan,China LL	TWN_NW	2005–2023
NE	Indonesia LL	IDN_NE	2006–2024
	Japan LL	JPN_NE	1994–2024
	Taiwan,China LL	TWN_NE	2005–2023
SW	EU–Spain LL	EU_ESP_SW	2001–2024
	EU–Portugal LL	EU_PRT_SW	2000–2024
	EU–France (Réunion) LL	EU_REU_SW	2005–2024
	Japan LL	JPN_SW	1994–2024
	Taiwan,China LL	TWN_SW	2005–2023
	South Africa LL	ZAF_SW	2004–2024
SE	Japan LL	JPN_SE	1994–2024
	Taiwan,China LL	TWN_SE	2005–2023

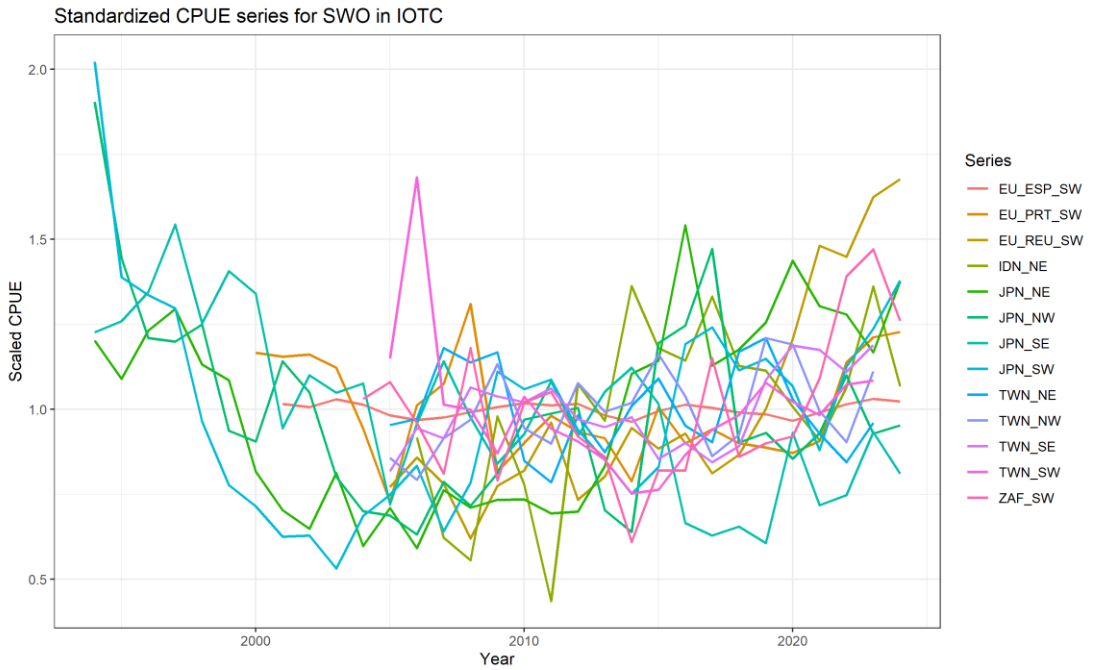

Figure 2: Standardized CPUE series available for the IOTC swordfish stock assessment, shown together for all series and areas. For better visualization and comparison, each series is scaled by its respective mean.

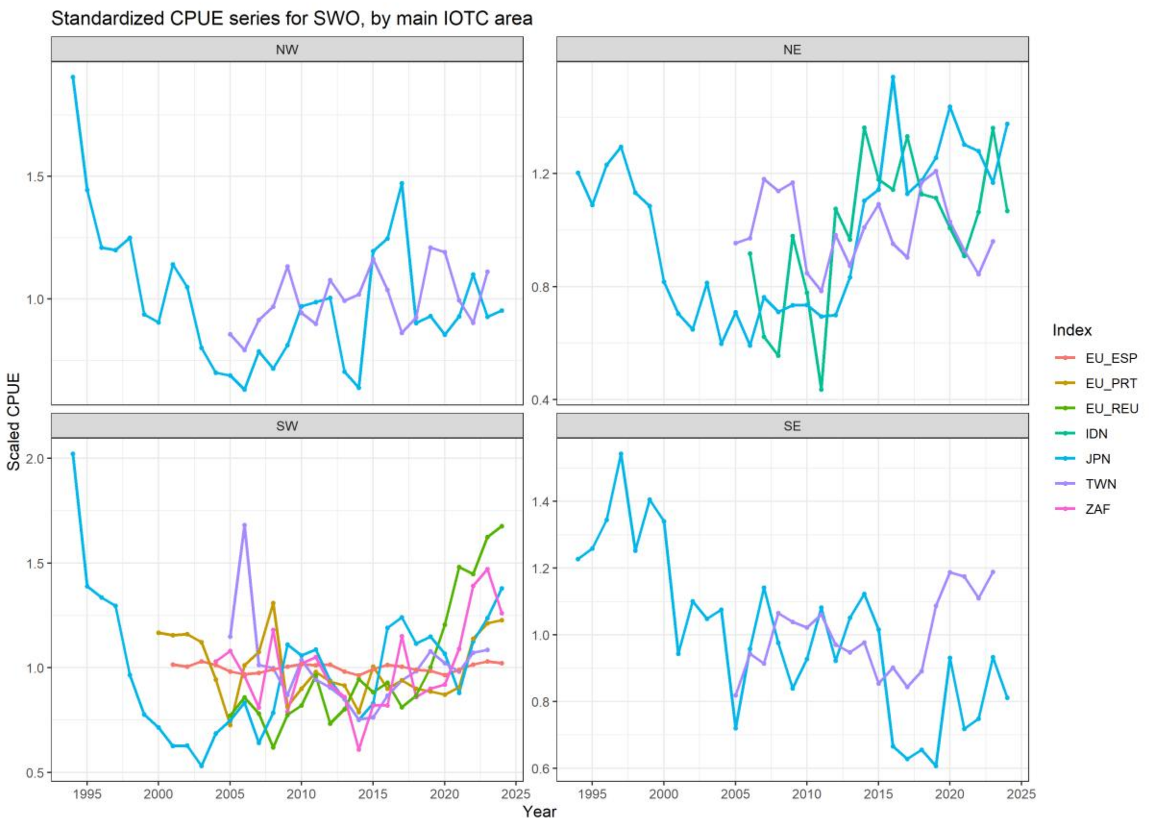

Figure 3: Standardized CPUE series available for the IOTC swordfish stock assessment, shown by IOTC area. For better visualization and comparison, each series is scaled by its respective mean.

2.4. Model development and structure

Three sets of models are presented in this paper, that are summarized in a grouped way in **Table 2**.

The base case is a single ocean-wide model, fitted to the total catch of swordfish in the entire IOTC area, and fitted to all 13 available indices. This configuration matches the scale at which IOTC manages the stock, and it uses all the abundance information available. Its central assumption is that each index is proportional to the total stock biomass, i.e. that a decline observed in one area reflects a decline of the stock as a whole rather than a purely local change. This is the standard assumption of surplus production models fitted to multiple indices.

The continuity run reproduces the configuration of the previous JABBA assessment of this stock, using only the five indices that were used at the time, namely the 4 Japanese series (one for each area) and the EU-Portugal series specific to the SW). The main purpose of this run is to provide continuity with the previous JABBA configuration, using a comparable model structure with updated data.

Finally, 4 area-specific models were developed, each fitted to the catch and the CPUE indices of each area separately. The main purpose was to compare the relative depletion and fishing mortality trajectories across areas and against the base case. This check is complemented by the residual patterns and runs tests of the base case itself, which would be expected to show systematic, area-correlated departures if the assumption were badly violated.

Table 2. The three main sets of JABBA model configurations developed and presented in this paper, for the IOTC swordfish stock assessment.

Model	Catch input	CPUE indices
Base case	Total IOTC (Indian Ocean) catches	All available 13 CPUE series
Continuity run	Total IOTC (Indian Ocean) catches	Previously used CPUE series (4 Japanese + 1 EU-Portugal)
Area-specific models	Specific catches of each area separately	Specific indices of each area used separately

2.5. Stock assessment

2.5.1. Assessment platform

The assessment models were implemented in JABBA, a Bayesian state-space surplus production model framework (Winker et al., 2018). JABBA is implemented in R and available from github.com/jabbamodel/JABBA.

JABBA is a flexible Bayesian stock assessment modelling framework with various options, that include: 1) automatic fitting of multiple CPUE time series and associated standard errors, 2) estimating or fixing the process variance, 3) optional estimation of additional observation variance for individual or grouped CPUE time series, 4) specifying the production function, i.e. Fox, Schaefer or Pella-Tomlinson, this last one by setting the inflection point from BMSY/K and converting it into the shape parameter m, 5) setting priors for various parameters, including r and K, that can range from more to less informative depending on the confidence in the previously available information, 6) model diagnostics and goodness-of-fit features with associated tests and plots (e.g. residual runs tests, hindcast and retrospective analysis) and, 7) projections for constant catches in the future to achieve management objectives over certain timeframes.

2.5.2. Base case model specifications

The base case specifications are summarized in **Table 3**. The production function used was a Pella-Tomlinson model with the inflection point of the production curve fixed at BMSY/K = 0.4, following a previous JABBA assessment of this stock configured by Parker (2020). The shape parameter m of this Pella-Tomlinson model was set as a prior and estimated around that inflection point with a CV of 0.3.

The K prior was kept vaguely informative, as regularly done in most production models, given the lack of prior knowledge of these values and to allow more emphasis to be placed on the r parameter, which is derived from biological information. Specifically, the K prior used the default settings of JABBA, namely a lognormal prior with a large CV (100%) and a central value corresponding to 8 times the maximum catch. This is consistent with other types of models, such as the approach used in catch-MSY (Martell and Froese, 2013).

The model was started in 1950, the first year of the catch series, when catches were negligible. The initial depletion (B1950/K) was therefore given a beta prior with a mean of 0.95 and a CV of 4%, corresponding to a stock close to but not exactly at the unfished level. The CV of this prior is constrained by the requirement that the resulting beta distribution remains bounded at its upper limit, as the B1/K cannot be higher than 1.

Catches were input with an associated CV of 0.20, therefore allowing some deviation from the reported swordfish catches, to reflect the likelihood that the catches may not be accurately estimated, recorded and reported to IOTC. Catchability parameters were formulated as uninformative priors, one per index, and the CPUE series were scaled externally by their respective means before being input to the model.

The process error was defined by a vaguely informative inverse-gamma distribution with the shape and scaling parameters set at 4 and 0.01, which represents the JABBA default. Sensitivities were carried out by setting both parameters at 0.001 (Gelman, 2006; Mourato et al., 2023), and by fixing the sigma of the process error to CVs of 5% and 10%.

In addition to the CPUE variance associated with the data, the base case configuration allowed the internal estimation of additional observation variance for each CPUE, therefore allowing for a larger divergence between the observed and model-predicted CPUEs. Sensitivity analyses were carried out by disabling this process, and by estimating the additional variance per area rather than per index.

For the parameter estimation, 3 MCMC chains were used with 40,000 iterations each, a burn-in period of 10,000 and a thinning rate of 5, resulting in a total of 18,000 posterior samples. This specification was used consistently in the base case, the continuity run, the area models and all the sensitivity runs.

Table 3. Summary of the base case model specifications and prior assumptions, for the JABBA 2026 IOTC swordfish stock assessment.

Model component	Base-case specification
Spatial structure	Single IOTC ocean-wide model

Model component	Base-case specification
Model period	1950–2024
Catch series	Total IOTC swordfish catches
CPUE indices	13 series, each with its own catchability and observation variance
Production function	Pella-Tomlinson, BMSY/K at 0.4, shape CV 0.3
r prior	Lognormal, log(r) ~ N(log(0.42), 0.4)
K prior	Lognormal, mean = 8 × maximum catch, CV = 100%
Initial depletion (B1950/K)	Beta, mean 0.95, CV 0.04
Catch CV	0.20
CPUE observation error	Fixed CV of 0.20, plus additional variance estimated internally
Minimum observation error	0.05
Process error	Estimated, inverse-gamma (4, 0.01)
MCMC	3 chains * 40,000 iterations, 10,000 burn-in, thinning 5 (18,000 samples)

2.5.3. Continuity run specifications

The continuity run follows the configuration of the previous assessment. It uses the total catch with only five indices (Portugal from the SW area plus the 4 Japanese indices, one from each area). The model uses a two-stage r/K prior approach, in which a first model is fit with vague range priors on r (uniform between 0.1 and 0.7) and K (uniform between 100,000 and 500,000 t), and supplies the lognormal priors for the final fit, with the r CV fixed at 0.4 and the K CV at 1. The initial depletion prior is lognormal with a mean of 1 and a CV of 0.1.

2.5.4. Area-specific model specifications

Four additional models were fitted, one for each of the IOTC swordfish assessment areas (NW, NE, SW and SE). Each model used the catch series of that specific area together with the standardized CPUE indices of that same area, i.e. 2 indices in the NW, 3 in the NE, 6 in the SW and 2 in the SE.

All settings and priors of these models were identical to those of the base case (**Table 3**), so that any differences between areas reflect differences in the data rather than in the model configuration. The only quantity that differs by construction is the K prior, which is set at 8 times the maximum catch and therefore scales automatically with the size of each area fishery.

The main purpose of these area-specific models is an additional diagnostic to the base case and continuity models. The base case assumes that each of the 13 indices is proportional to the total stock biomass, and that assumption is most likely to fail if the four areas are at substantially different levels of depletion. Comparing the relative stock trajectories (e.g., B/B0, B/BMSY and F/FMSY) estimated by the four area models therefore provides a direct check. Similar trajectories support the assumption made in the base case, whereas clearly divergent trajectories would indicate that the base case results should be interpreted with more caution.

It should be noted that each of these models treats its area as a closed population, with its own carrying capacity, productivity and reference points, which is a strong assumption that cannot be maintained for a highly migratory species such as swordfish. Any movement of fish into or out of an area cannot be represented in a closed-area production model, and will instead be absorbed into the estimates of other paraments, such as r and K. For this reason, the absolute quantities estimated by these specific area models, and particularly the area-specific MSY values, should not be interpreted as biological quantities of those areas.

2.6. Model diagnostics

Basic diagnostics of model convergence included MCMC trace plots and other statistics (Geweke, 1992; Gelman and Rubin, 1992), together with the effective sample sizes of the key parameters.

To evaluate the CPUE fits, the model-predicted CPUE indices were compared to the observed CPUE. Additionally, residual plots were used to examine the residuals of observed versus predicted CPUE indices for all fleets, with boxplots showing the median and quantiles of all residuals for each year and a loess smoother through all residuals to aid detection of systematic residual patterns.

Additionally, the root-mean-squared-error (RMSE) was used as a goodness-of-fit statistic, and runs tests were conducted to quantitatively evaluate the randomness of the residuals (Carvalho et al., 2017). The runs test diagnostic was applied to the residuals of the CPUE fit on the log scale, considering the 2-sided p-value of the Wald-Wolfowitz runs test, and is visualized in JABBA to illustrate which time series passed or failed the test, as well as highlighting individual data points that fall outside the three-sigma limits (Anhøj and Olesen, 2014).

To check for systematic bias in the stock status estimates, a retrospective analysis was carried out by sequentially removing one year of data at a time, over a total period of 5 years, and refitting the model without those years. The parameters of interest were then compared to the original model fitted using the full time series. The presence of possible retrospective bias was analysed visually and statistically with the Mohn's rho (ρ) statistic (Mohn, 1999), using the formulation defined by Hurtado-Ferro et al. (2014). The more the values diverge from zero, the stronger the presence of a retrospective bias, noting that in general the values that fall between -0.15 and 0.20 are widely deemed as having an acceptable retrospective bias.

Finally, hindcast cross-validation was conducted for the recent period of the indices, with one-year-ahead forecasts of the CPUE values produced when the last years are removed one at a time. Prediction skill was quantified with the mean absolute scaled error (MASE), where values below 1 indicate that the model forecasts are better than those of a naive random walk (Carvalho et al., 2021).

2.7. Sensitivity analysis

All sensitivity analysis was carried out on the base case model, with the various scenarios tested listed in **Table 4**. In addition to the usual axes of uncertainty in the production function, the priors and the error assumptions, three groups of runs address the influence of the individual indices: one model per index in turn, one model with each index removed in turn, and one model with all the indices of a given area removed together. The last of these is the direct test of whether the six southwest indices drive the ocean-wide result.

Table 4. Sensitivity analyses conducted on the base case model of the 2026 IOTC swordfish JABBA stock assessment.

Type	Scenarios
Production function	Schaefer, Fox, and Pella-Tomlinson with BMSY/K of 0.3 and 0.5
Productivity	r ~ lognormal(0.30, 0.3); r ~ lognormal(0.55, 0.3); uniform range 0.1–0.7
Initial depletion	B1950/K beta prior with mean 0.90, CV 0.05
CPUE CV	CPUE CV of 0.3
Additional CPUE variance	Internal estimation of additional CPUE variance turned off
CPUE variance grouping	Observation variance grouped by area rather than by index
Process error prior	Uninformative inverse-gamma (0.001, 0.001)
Process error fixed	Process error sigma fixed at CVs of 5% and 10%
Use one CPUE index at a time	13 runs, one per index
Leave out one CPUE index at a time	13 runs, dropping one index at a time
Leave out CPUEs from one area at a time	4 runs, dropping all the indices of each one area at a time

2.8. Software

We used JABBA v2.3.2 (Winker et al., 2026) run from R v4.5.1 (R Core Team, 2025). JABBA uses the JAGS software (Plummer, 2003) to estimate model parameters in a Bayesian framework, by means of Markov Chain Monte Carlo (MCMC) simulation. JAGS is executed from R using the library "r2jags" (Su and Yajima, 2024). Additional libraries used in this paper for data preparation and plotting, included "tidyr" (Wickham et al., 2025), "tidyverse" (Wickham et al., 2019), "ggplot2" (Wickham, 2016), "dplyr" (Wickham et al., 2026), "gridExtra" (Auguie, 2017), "cowplot" (Wilke, 2025) and CODA (Plummer et al., 2006).

3. Results

3.1. Base case model

3.1.1. Model convergence

The MCMC convergence tests all passed with regards to the estimation of the parameters of the base case model. Adequate convergence of the MCMC chains was also corroborated visually by checking the trace plots, which showed good mixing and random deviations around the parameter space, without any detectable bias or patterns that could result from autocorrelations in the estimations (**Figure 4**).

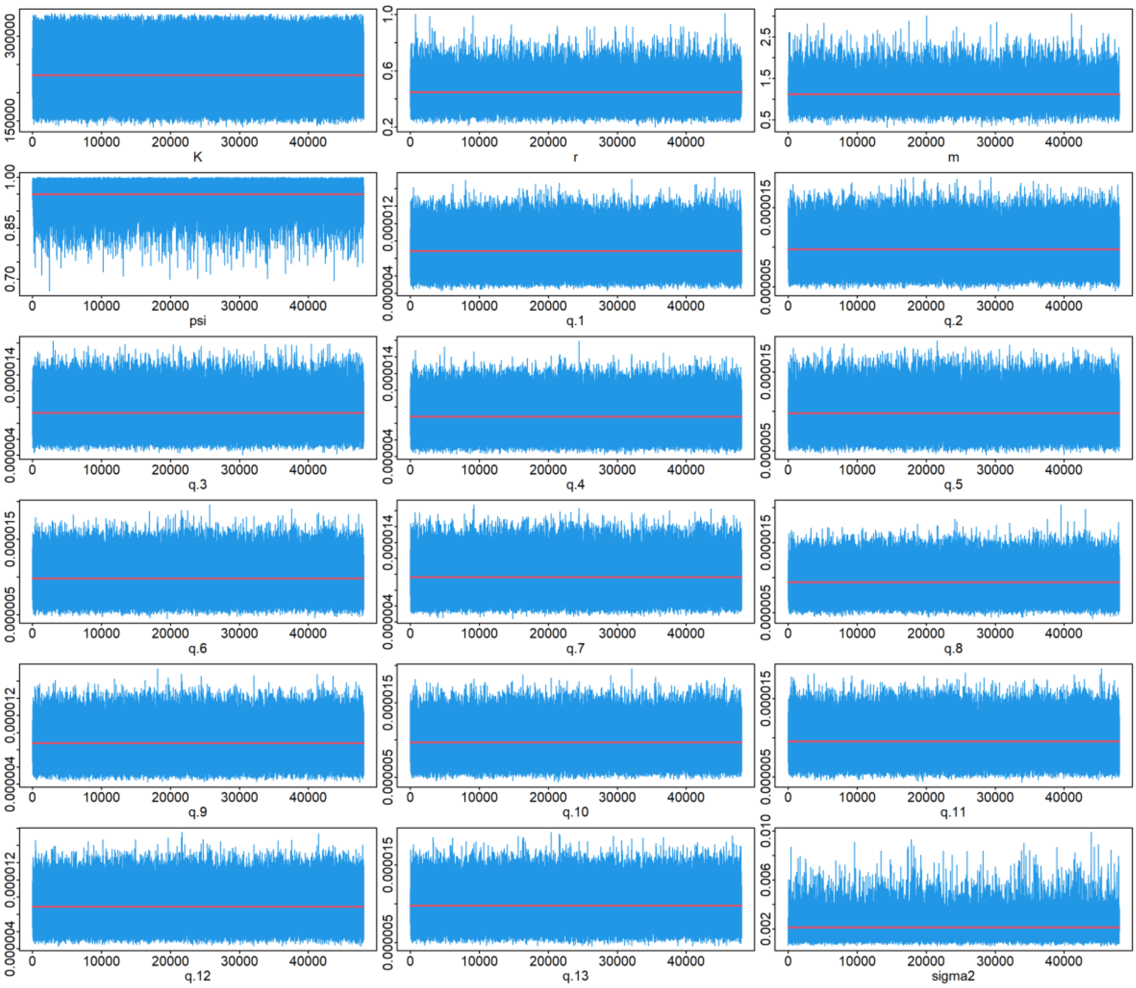

Figure 4. MCMC trace plots of the base case JABBA model used for the 2026 IOTC swordfish stock assessment.

3.1.2. Fits to CPUE indices

The fits of the base case model to each of the 13 standardized CPUE indices are shown in **Figure 5**, and the residual diagnostics in **Figure 6**. The overall RMSE was 19.5%.

The runs tests are shown in **Figure 7**, with 6 of the 13 indices passing the test. The runs test failures are concentrated in the series that carry the strongest downward-then-upward trend signal (e.g., Japanese indices and several of the SW series), while the flatter series tend to pass this test. This pattern can be expected and explained given the model is a single-area model fitted to multiple indices that are not fully consistent with each other. The model produces a compromise trajectory that cannot simultaneously reproduce the trend in each individual series. It is therefore an indication of conflict between indices rather than necessarily of misspecification of the population dynamics.

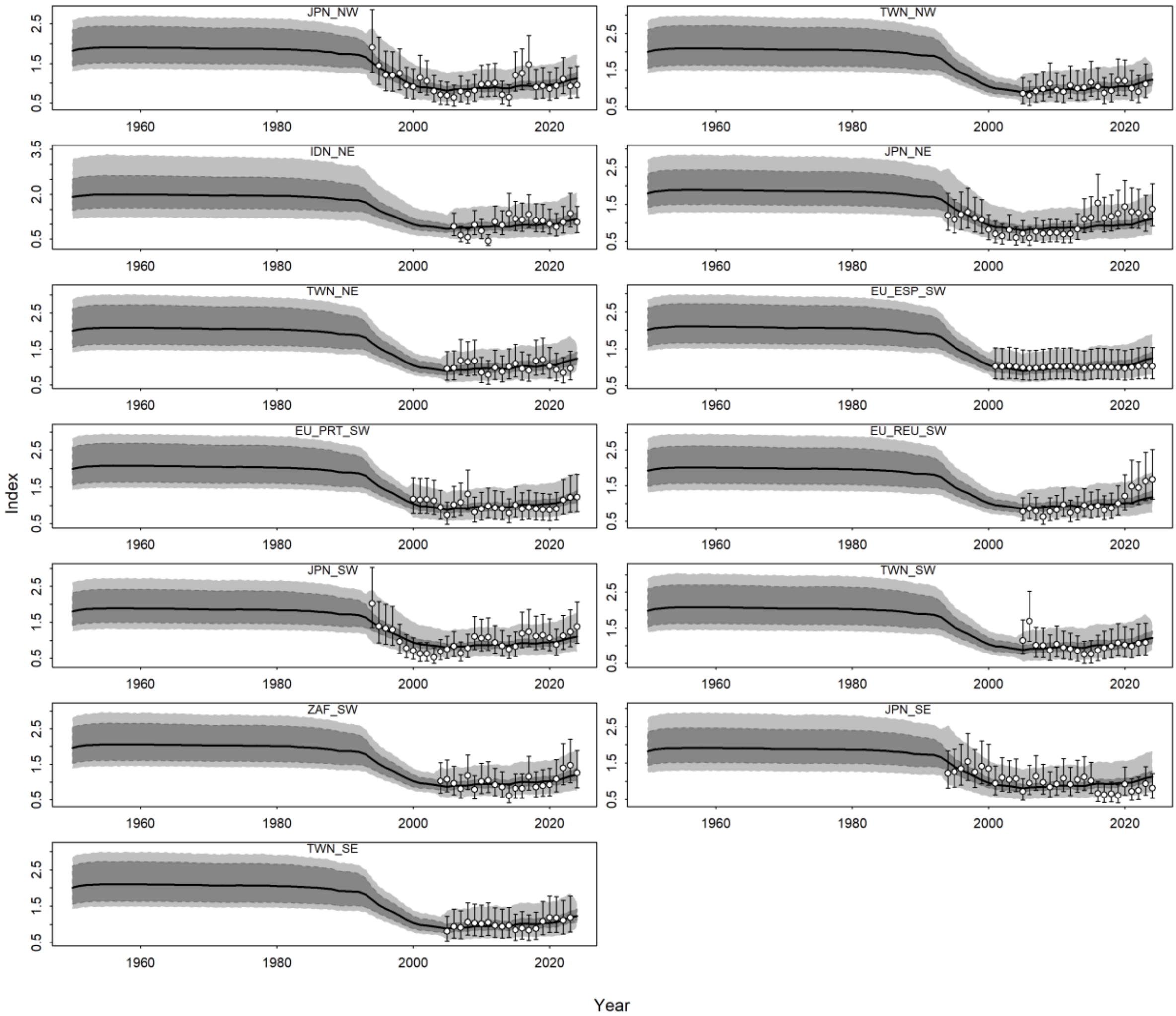

Figure 5. Time series of observed (circles) and predicted (solid line) CPUEs for the base case model. The dark shaded areas represent the 95% credibility intervals of the expected mean CPUE, and the light shaded areas represent the 95% posterior predictive distribution intervals. The error bars are the 95% confidence intervals from the CPUE observations.

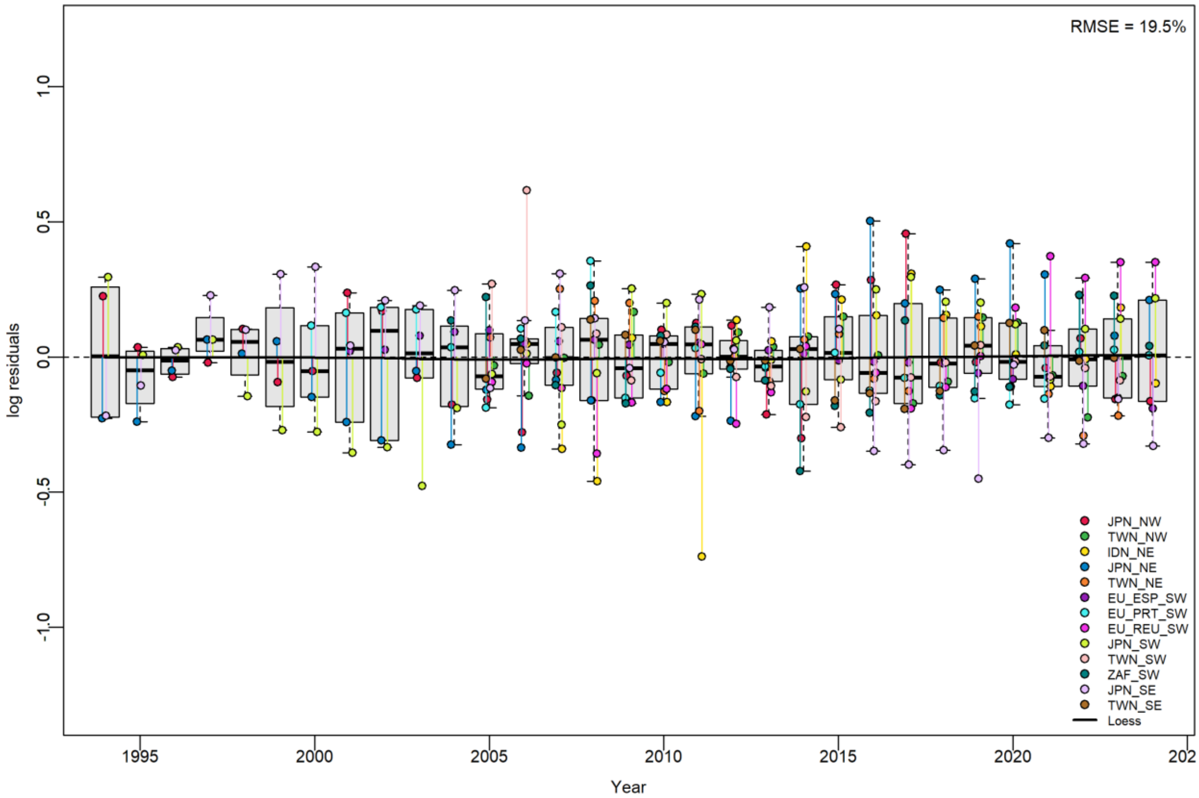

Figure 6. Residual diagnostic plot for the base case model. Each individual CPUE index and its respective residuals are represented by a different colour. The solid black line represents a loess smoother through all residuals combined.

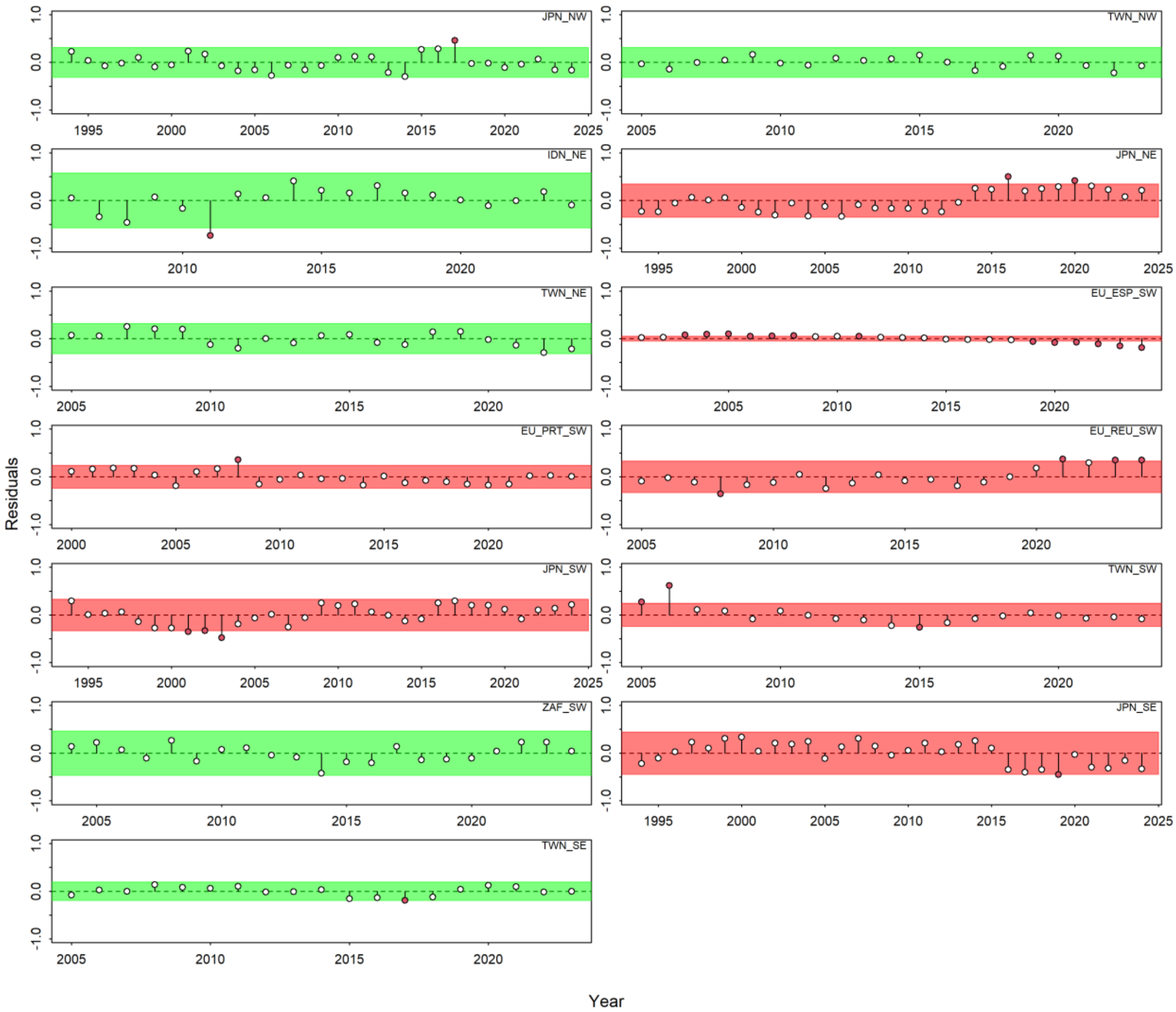

Figure 7. Runs tests for the 13 CPUE indices of the base case model. CPUE plots that are highlighted in green passed the formal statistical test and the ones highlighted in red failed the test.

3.1.3. Process error

The process error deviations of the base case are shown in **Figure 8**. The line is close and oscillates around the value of zero, with the 95% credibility intervals always including the zero value throughout the series. This suggests no major evidence of structural model misspecification.

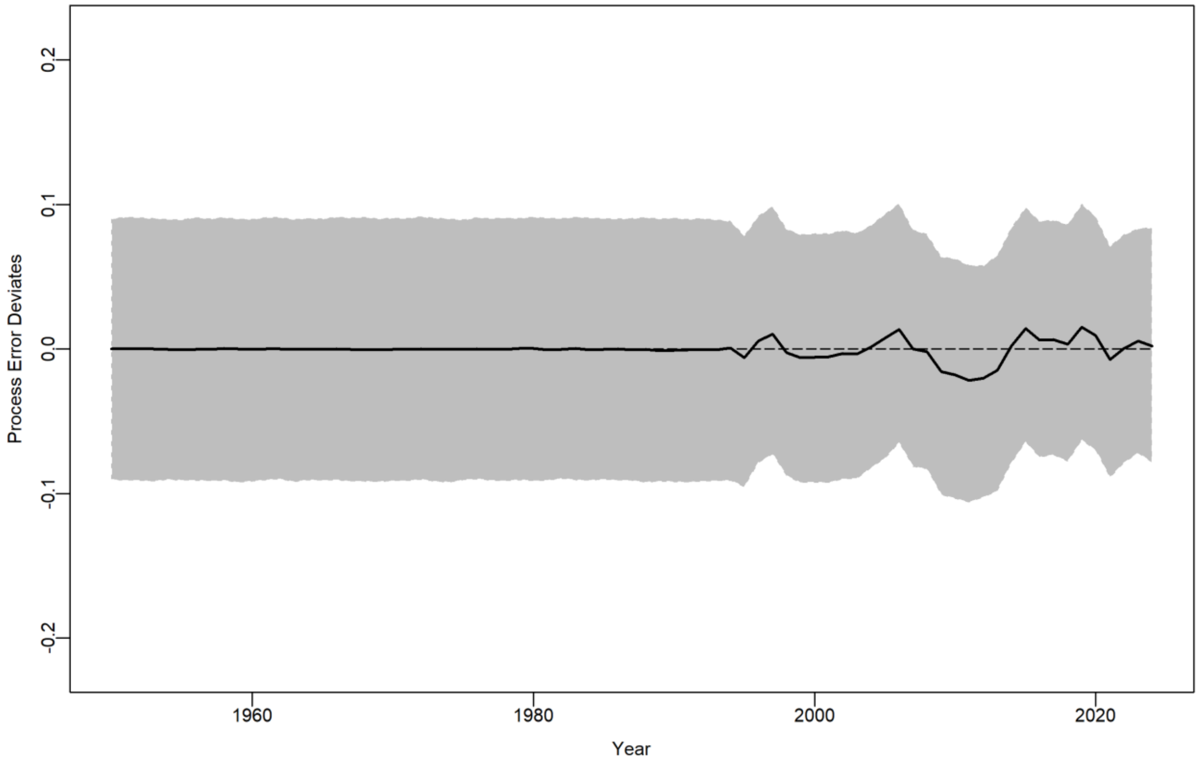

Figure 8. Process error deviates from the base case model. The solid black line represents the median, and the shaded grey area the 95% credibility intervals.

3.1.4. Retrospective analysis

The results of the retrospective analysis are shown in **Figure 9**, and the Mohn's rho statistics are presented in **Table 5**. Almost all cases passed the statistical test, highlighting that in general there are no major retrospective patterns in the base case model.

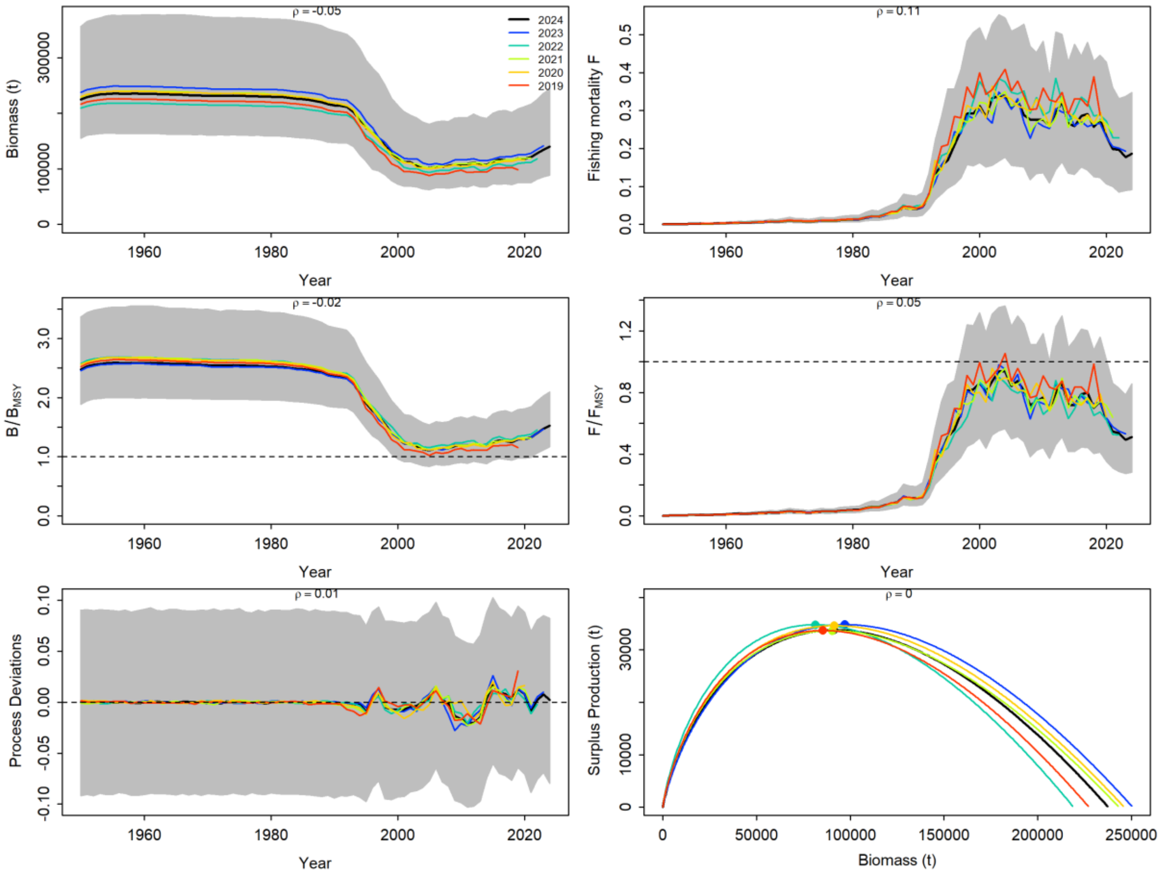

Figure 9. Retrospective analysis conducted for the base case model, by removing 1 year at a time sequentially (n = 5) and predicting the trends in overall biomass and fishing mortality, biomass and fishing mortality relative to MSY, process error deviations and surplus production.

Table 5. Mohn's rho statistics for the retrospective analysis conducted for the base case model. Values that fall within the accepted values are highlighted in green and the cases that failed the test are highlighted in red.

Year	B	F	Bmsy	Fmsy	procB	MSY
2024	0.051	0.086	-0.002	0.083	0.003	0.013
2023	-0.075	0.161	0.041	-0.022	-0.004	0.029
2022	-0.017	0.210	-0.008	0.163	0.007	-0.017
2021	-0.021	0.041	-0.008	0.024	0.009	0.005
2020	-0.169	0.047	-0.109	0.005	0.018	-0.007
rho.mu	-0.046	0.109	-0.017	0.050	0.006	0.005

3.1.5. Hindcasting

The hindcast cross-validation results are shown in **Figure 10**. Two indices (IDN_NE and TWN_NW) had MASE lower than 1, meaning that their prediction skill better than a naive random walk, while the remaining indices had MASE values above 1, ranging from 1.12 to 7.25. It should be noted, however, that MASE is scaled by the one-step-ahead error of a random walk, so that very smooth and highly autocorrelated series (e.g., EU_ESP_SW) will produce a very small denominator and consequently highly inflated MASE values, even when the absolute prediction errors are small.

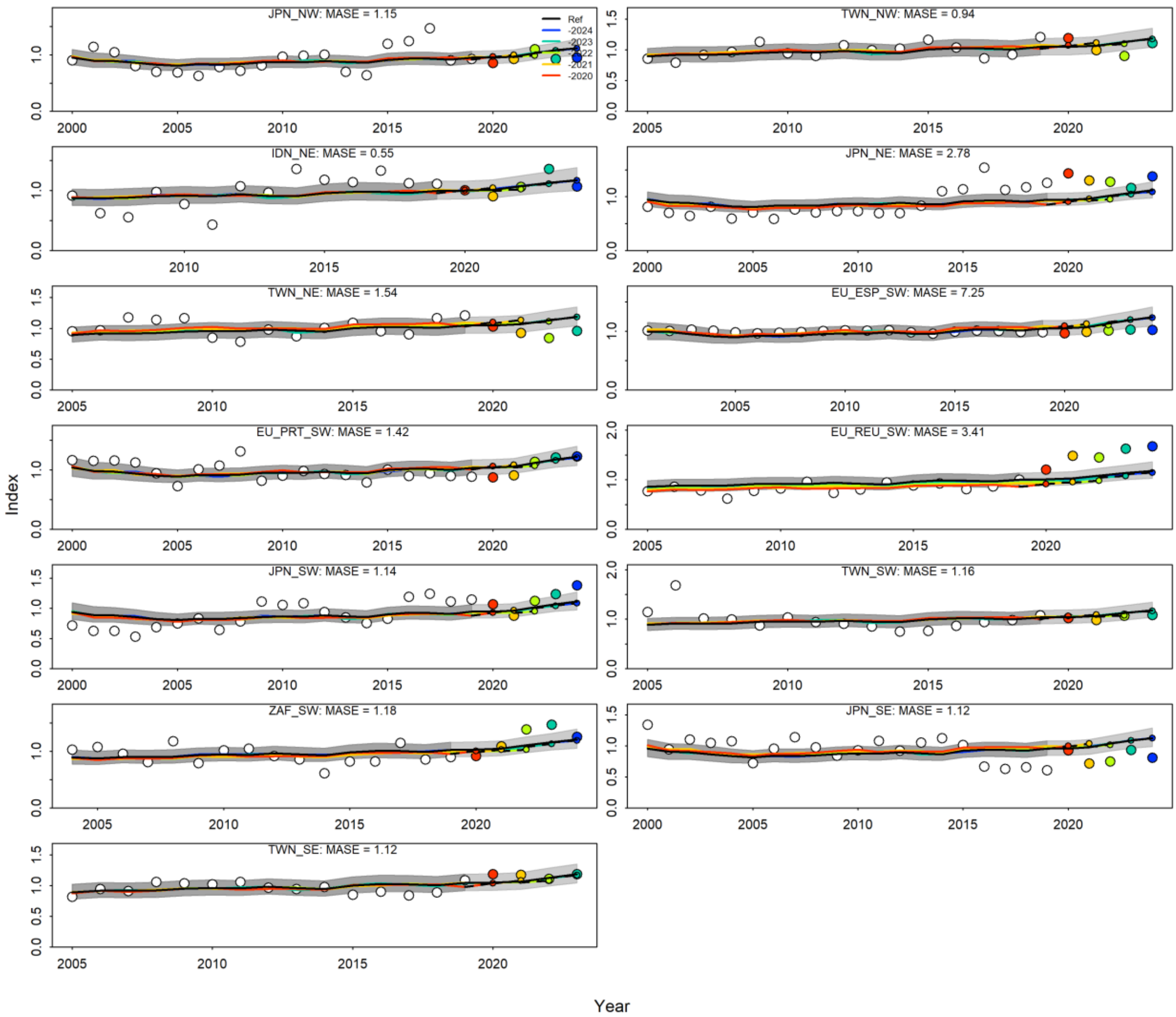

Figure 10. Hindcast cross-validation results for the base case model. The plots show 1-year-ahead forecasts of CPUE values, when the last years are removed one at a time, relative to the observed CPUE using all data.

3.2. Continuity run

The continuity run, using the five indices and the configuration of the previous JABBA assessment, is compared with the base case in **Figure 11**. The 2 models are very similar, along most of the time series of the trajectories.

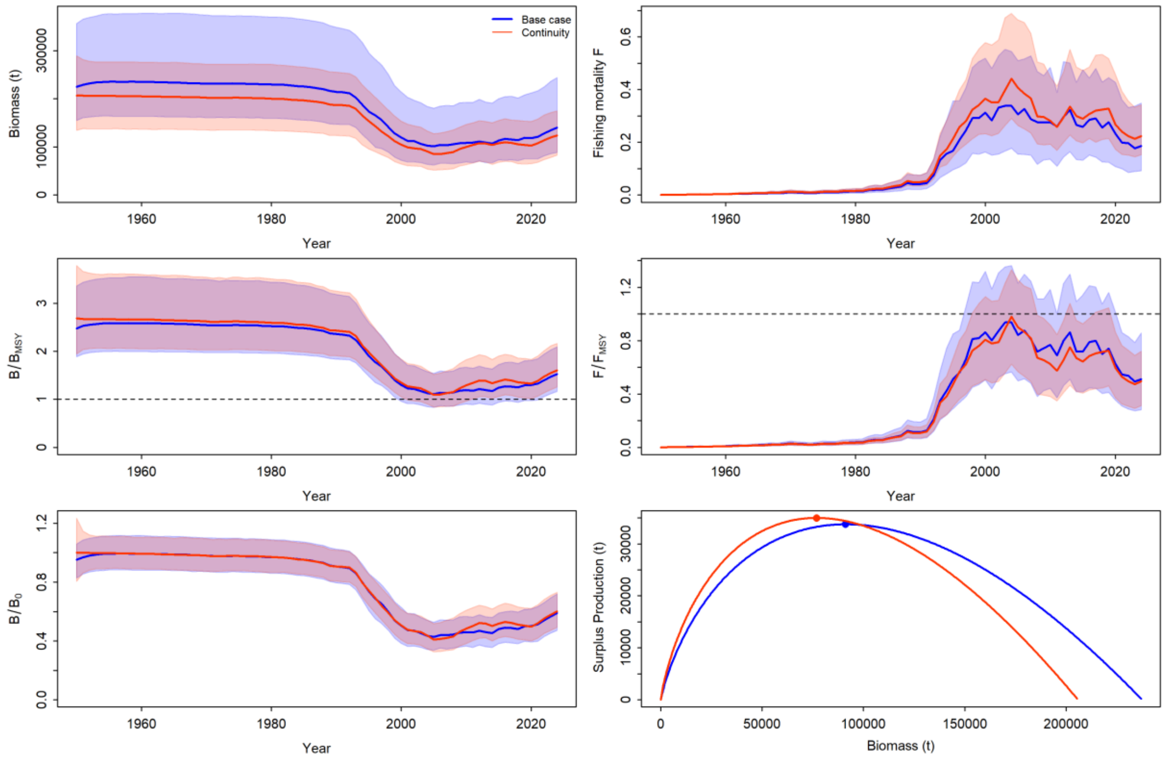

Figure 11. Comparison of the estimated trajectories of the base case model and the continuity run for the IOTC swordfish stock assessment.

3.3. Area-specific models

The 4 area-specific models were fitted and their relative trajectories compared with each other and with the base case, as well as with the continuity run (**Figure 12**).

In general, the relative trajectories (B/B_0, B/B_{MSY} and F/F_{MSY}) estimated by the 4 area-specific models are broadly consistent with each other and with the base case, with no single area showing a markedly divergent depletion trajectory. This provides some support for the base case assumption that the individual indices can be treated as proportional to total stock biomass.

It is emphasised, however, that the absolute quantities estimated by these area-specific models (K, BMSY, FMSY and particularly MSY) are conditional on the assumption that each area would be a closed unit population, which is not realistic for a migratory species such as swordfish. Therefore, the absolute quantities are not additive across the areas and should not be interpreted as area-specific reference points or catch limits.

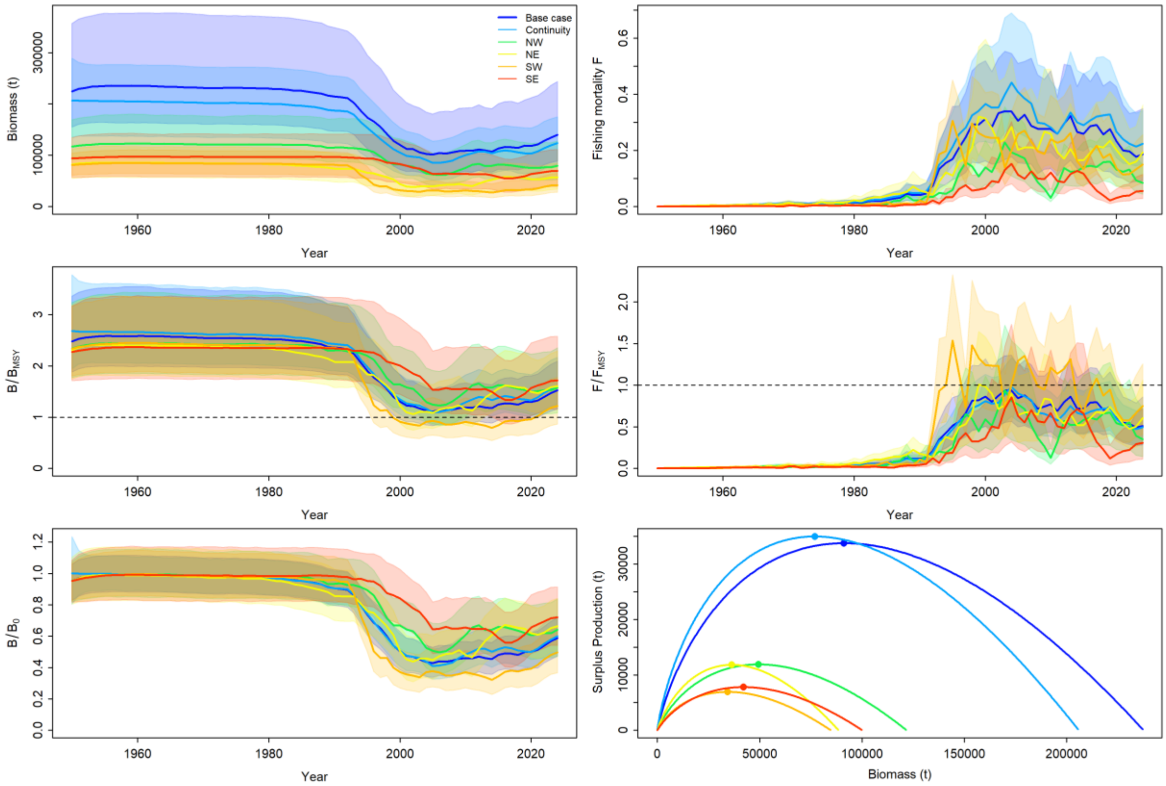

Figure 12. Comparison of the estimated trajectories of the base case model and the continuity run, with the 4 area-specific models, for the IOTC swordfish stock assessment.

3.4. Sensitivity analysis

3.4.1. Production function

The comparison of the various assumptions and sensitivity runs on the production function, carried out in the base case model, are shown in **Figure 13**.

It is interesting to note that the model maintains largely its consistency and overall trends, even when this function is modified. The final stock status also remains unchanged when compared to the base case model

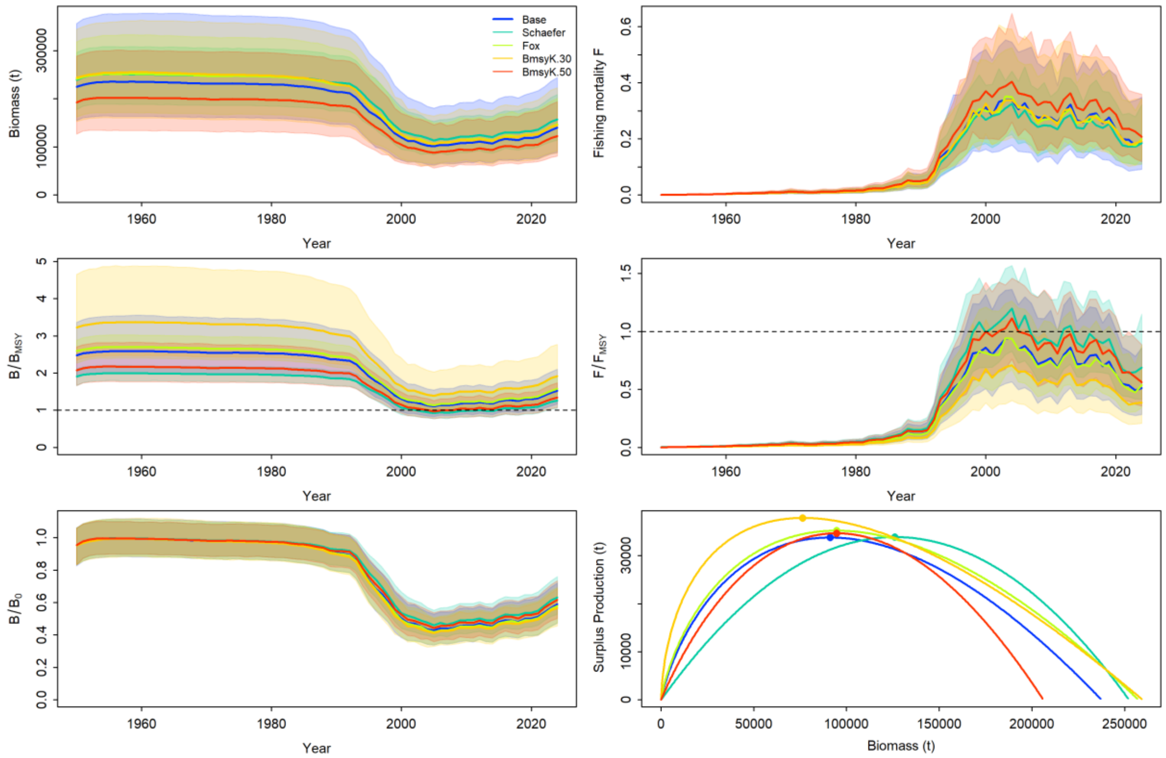

Figure 13. Sensitivity analysis comparing the various production functions to the base case model, used for the IOTC swordfish stock assessment using JABBA.

3.4.2. Productivity

The comparison of the species assumed productivity, in the form of the r prior, is shown in **Figure 14**.

It is interesting to note that the model maintains also its consistency and overall trends very close to the base case when the r prior is given different formulations, from more productive to less productive scenarios. Even when a non-informative r prior was used, the model still resulted in trends that were very similar to the base case.

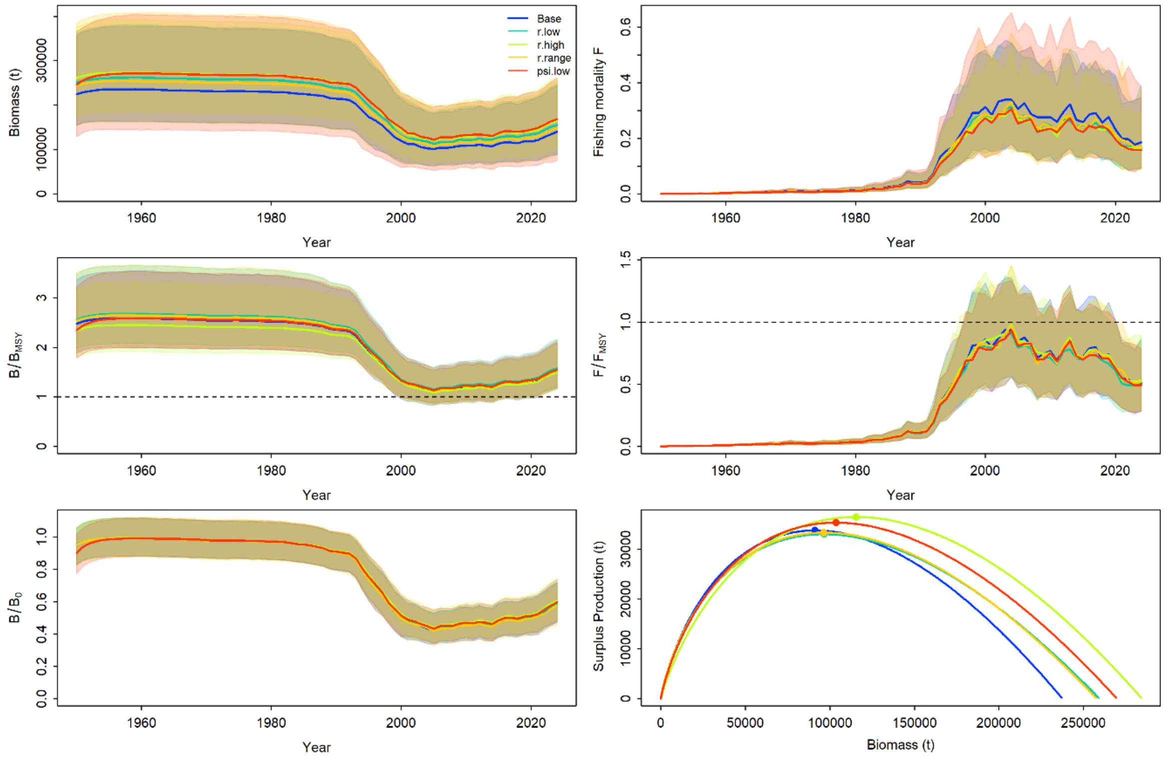

Figure 14. Sensitivity analysis comparing various options in terms of the r prior that is used for the IOTC swordfish stock assessment using JABBA.

3.4.3. Process error

The comparison of various process error options is shown in **Figure 15**.

In this case, again there were no major differences when comparing the final trajectories to the base case model specification.

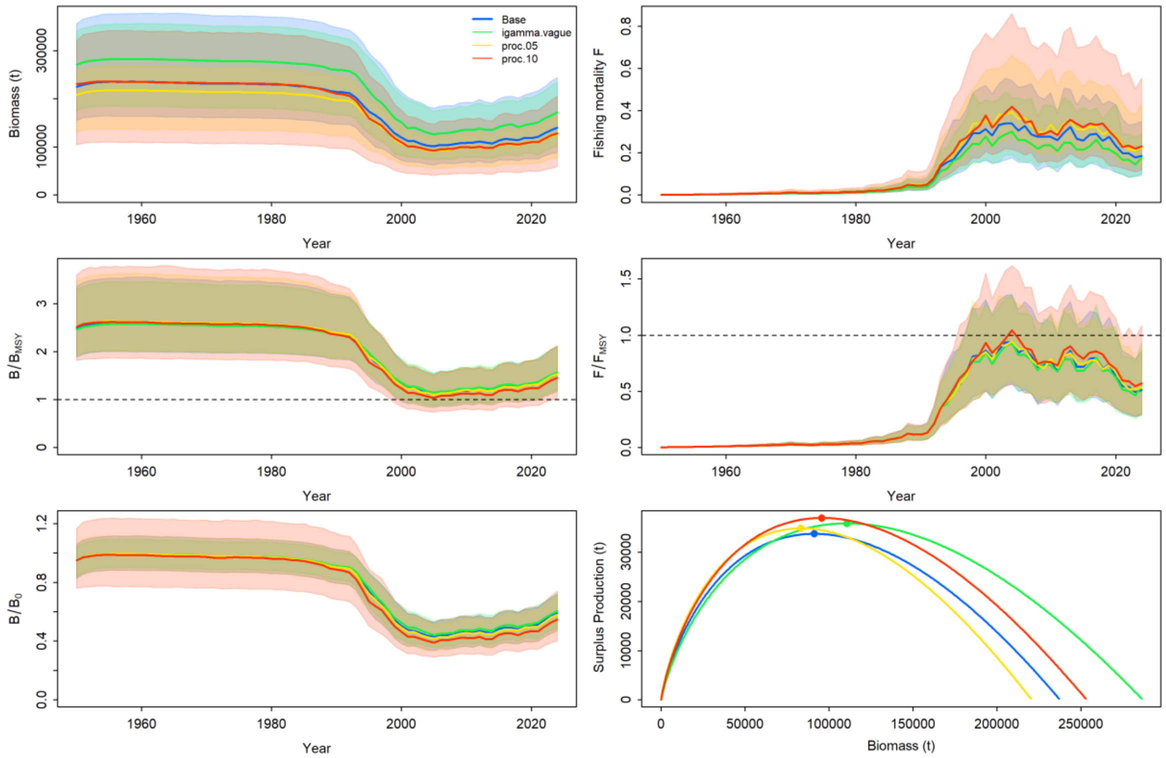

Figure 15. Sensitivity analysis comparing various options in terms of process error specifications used for the IOTC swordfish stock assessment using JABBA.

3.4.4. Leave out one CPUE at a time

The results of the runs leaving one CPUE index out at a time are shown in **Figure 16**.

The results are largely consistent, meaning that even if one of the indices used is missing and is not used, the final results would be largely similar, especially in the trends of B and F relative to MSY.

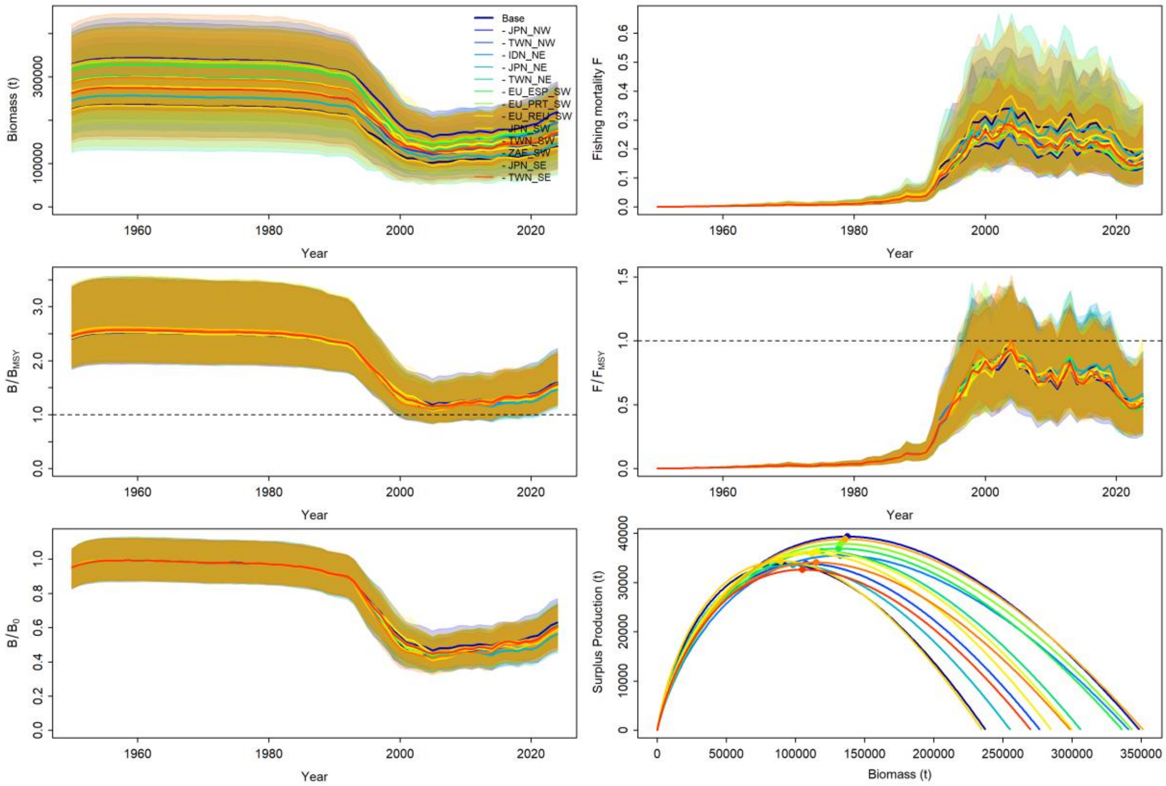

Figure 16. Sensitivity analysis leaving out one CPUE index at a time from the base case model, used for the IOTC swordfish stock assessment using JABBA.

3.4.5. Use one CPUE at a time

The results of the runs using one CPUE index at a time are shown in **Figure 17**. In this case there are more differences observed, and different indices tend to have more different signals in the model.

There seems to be 2 clusters of indices, one producing the results more aligned to the base case and one that produces flatter trends. The indices that produce the results more aligned with the base case are JPN_SW, IDN_NE, JPN_NE, JPN_NW, EU_REU_SW and JPN_SE. The indices that produce the flatter time series trajectories are TWN_NE, TWN_SE, EU_ESP_SW, TWN_SW, TWN_NW, EU_PRT_SW and ZAF_SW.

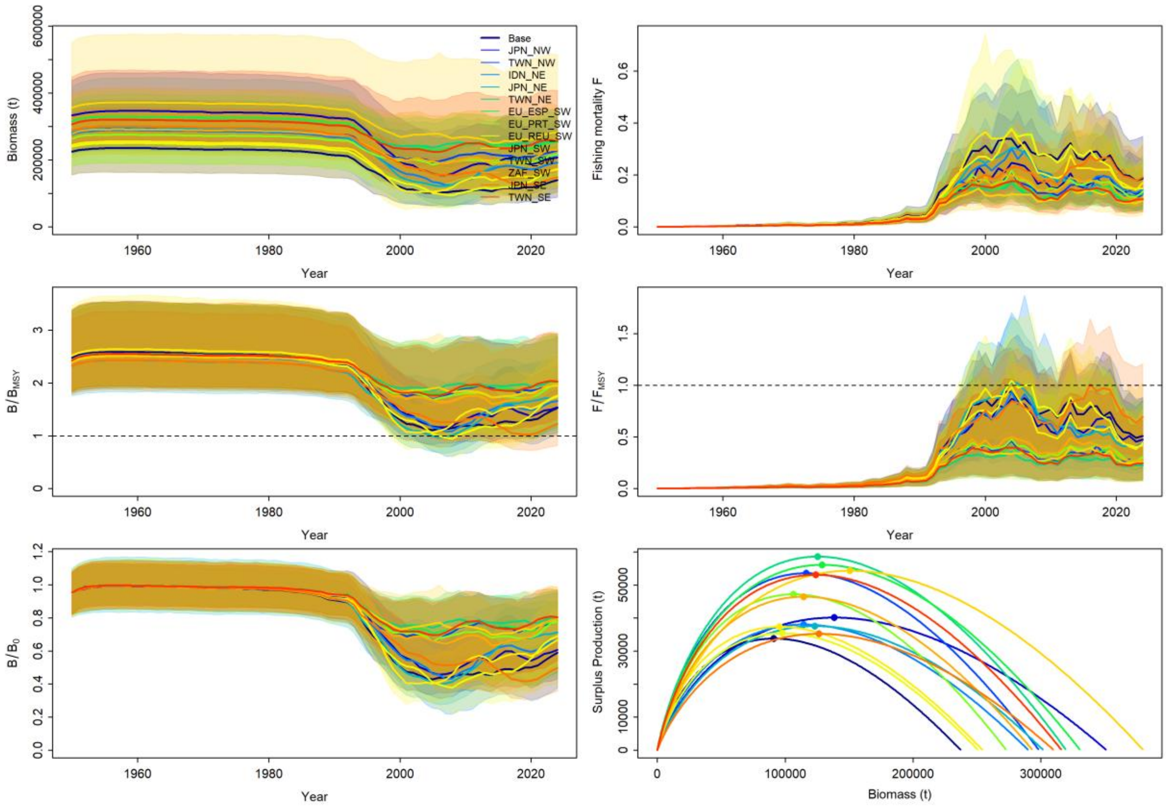

Figure 17. Sensitivity analysis using one CPUE index at a time, used for the IOTC swordfish stock assessment using JABBA.

3.4.6. Removing CPUEs from specific areas

The runs in which all the indices of each specific area were removed together are shown in **Figure 18**.

In this case, the overall trends are largely aligned and similar to the base case, that uses the CPUEs from all areas combined. This means that each single area is not too influential in the final result.

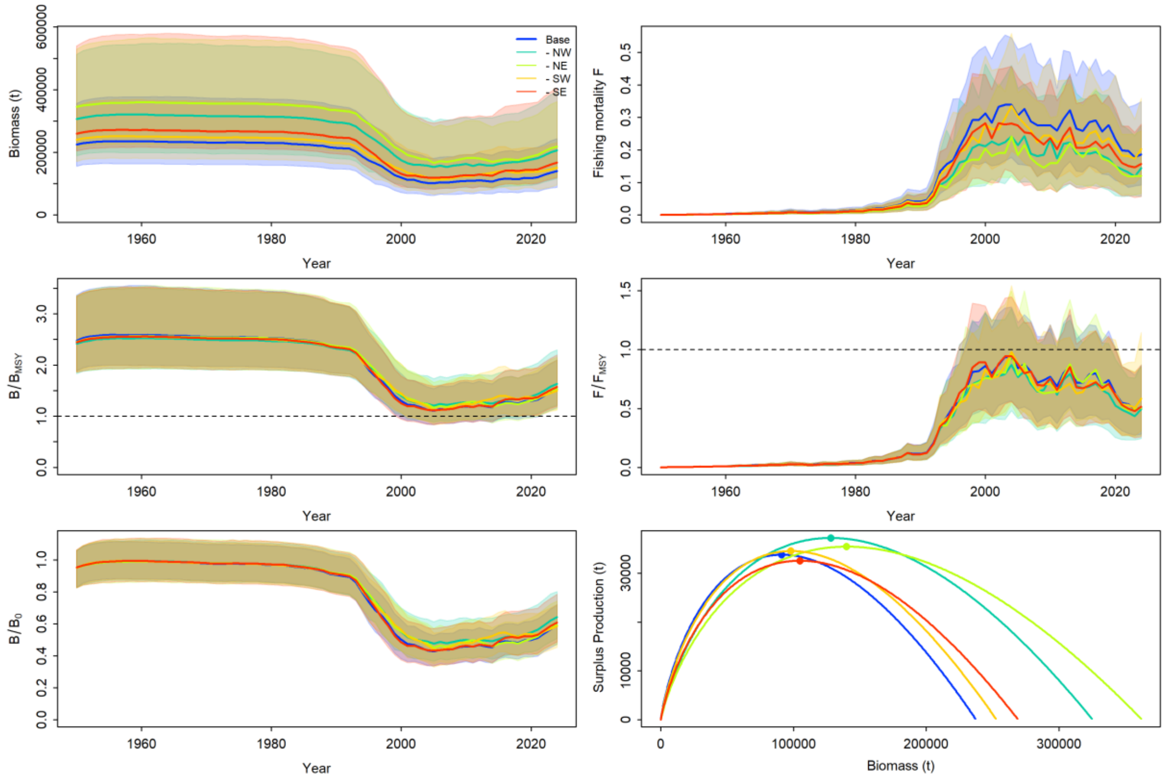

Figure 18. Sensitivity analysis removing all the CPUE indices of one assessment area at a time from the base case model.

3.5. Stock status

The final model stock trajectories and stock status is represented in the Kobe phase plot in **Figure 19**. The main parameters estimated from this final base model are reported in **Table 6**.

There was an increase in F/Fmsy and a decrease in B/Bmsy from the start of the model until around 2000, then a relatively stable period until 2010, and since then F/Fmsy has been decreasing and B/Bmsy has been increasing.

The stock is currently estimated to not be overfished (B/Bmsy = 1.528; 95%CIs: 1.168 - 2.096) and not subject to overfishing (F/Fmsy = 0.511; 95%CIs: 0.282 - 0.859). MSY is estimated at 33,816 t (95% CIs: 30,038 – 39,884).

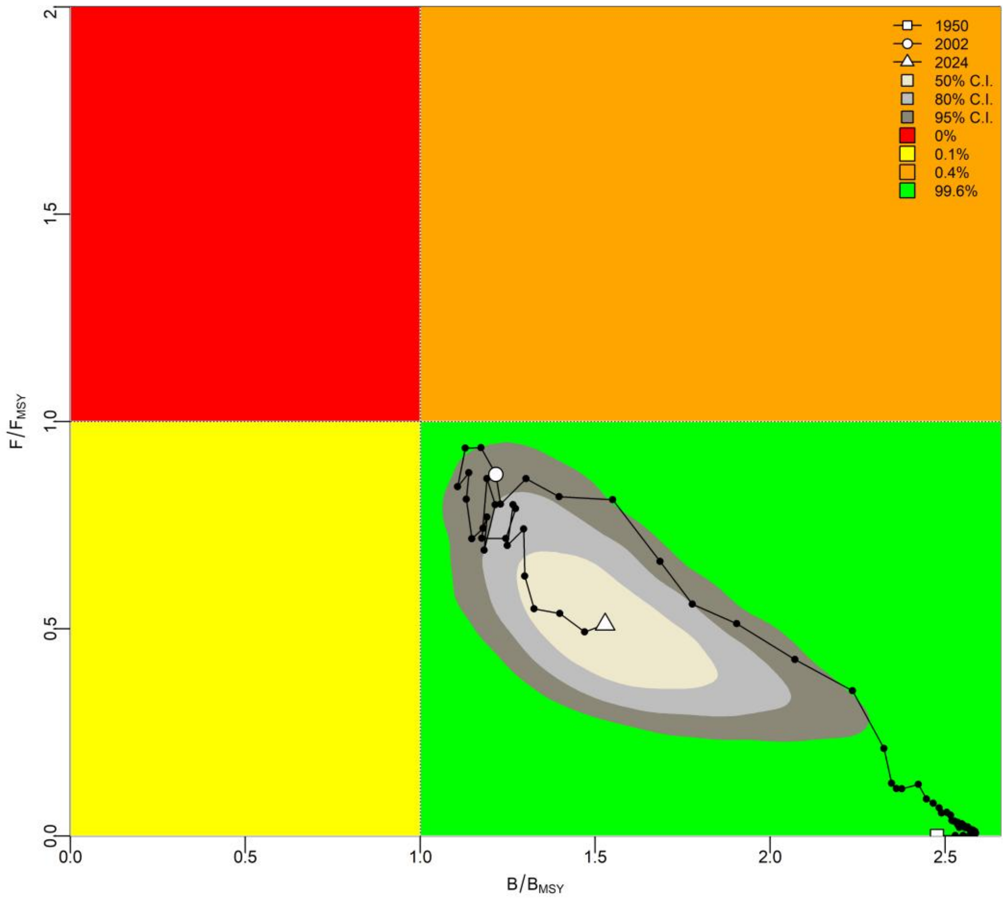

Figure 19. Kobe phase plot with the estimated trajectory of B/BMSY and F/FMSY for the base case model. The different grey shaded areas denote the 50%, 80% and 95% credibility intervals for the terminal year of the assessment data (2024). The probability of the terminal year stock status falling within each quadrant is indicated in the figure legend.

Table 6. Main quantities estimated by the final base case model, for the IOTC swordfish JABBA stock assessment.

Parameter	mu	lci	uci
K	237566	167495	355473
r	0.400	0.240	0.619
Psi	0.958	0.850	0.996
sigma.proc	0.044	0.031	0.065
m	1.091	0.633	1.859
Fmsy	0.370	0.202	0.610
Bmsy	91090	57184	161700
MSY	33816	30038	39884
Bmsy/K	0.384	0.288	0.486
P1950	0.952	0.827	1.059
P2024	0.589	0.472	0.719
B/Bmsy.cur	1.528	1.168	2.096
F/Fmsy.cur	0.511	0.282	0.859

4. Discussion

Bayesian surplus production models (JABBA) were applied to determine the stock status of Indian Ocean swordfish, in support of the provision of management advice for this species to IOTC. The final base case was a single ocean-wide model fitted to the total raised catch and to all 13 standardized CPUE series available.

An extensive sensitivity analysis was carried out, and the estimated final stock status tended to be robust to the sensitivity scenarios tested. A continuity run, reproducing the configuration of the previous JABBA assessment with 5 indices, and using updated catches and CPUEs until the new terminal year, also resulted in a similar final stock status and overall perception.

Several main limitations of this analysis should be taken into account. In terms of the CPUEs, it is important to note that the CPUE coverage by area is uneven, with more information in the SW area (6 of the 13 series) and less from the remaining regions. This unbalance was addressed through the sensitivity analysis by the leave-one-area-out and variance grouping, but it remains a data characteristic that should be highlighted. Another characteristic of the way we used the CPUE data is that the CVs of the standardized CPUE indices were set to a constant CV of 0.20, which gives every index the same nominal weight irrespective of its actual precision. Finally, and as a more general limitation, we need to highlight that surplus production models such as JABBA track exploitable biomass and do not represent the age or size structure of the stock, so they cannot detect changes driven by shifts in selectivity or in the age composition of the catch.

The final stock status for the IOTC swordfish using JABBA, considering a terminal year of 2024, is estimated to not be overfished (B/Bmsy = 1.528) and not subject to overfishing (F/Fmsy = 0.511). This final perception of the stock tented to be robust across most sensitivities and scenarios that were run using JABBA productions models.

5. Acknowledgments

We thank the IOTC secretariat for preparing and providing the data used in this stock assessment, and CPC scientists that have prepared and provided standardized CPUE series from their national fleets. R. Coelho (CCMAR) receives Portuguese national funds from FCT - Foundation for Science and Technology through contracts UID/04326/2025 (DOI https://doi.org/10.54499/UID/04326/2025), UID/PRR/04326/2025 (DOI https://doi.org/10.54499/UID/PRR/04326/2025) and LA/P/0101/2020 (DOI:10.54499/LA/P/0101/2020).

6. References

Anhøj, J., Olesen, A.V., 2014. Run charts revisited: A simulation study of run chart rules for detection of non-random variation in health care processes. PLoS One, 9: 1–13. https://doi.org/10.1371/journal.pone.0113825.

Auguie, B., 2017. gridExtra: Miscellaneous Functions for "Grid" Graphics. R package version 2.3, https://CRAN.R-project.org/package=gridExtra.

Carvalho, F., Punt, A.E., Chang, Y.J., Maunder, M.N., Piner, K.R., 2017. Can diagnostic tests help identify model misspecification in integrated stock assessments? Fisheries Research, 192: 28–40. https://doi.org/10.1016/j.fishres.2016.09.018

Carvalho, F., Winker, H., Courtney, D., Kapur, M., Kell, L., Cardinale, M., Schirripa, M., Kitakado, T., Yemane, D., Piner, K.R., Maunder, M.N., Taylor, I., Wetzel, C.R., Doering, K., Johnson, K.F., Methot, R.D., 2021. A cookbook for using model diagnostics in integrated stock assessments. Fisheries Research, 240: 105959. https://doi.org/10.1016/j.fishres.2021.105959.

Fu, D., 2023. Indian Ocean swordfish stock assessment 1950-2021 (Stock Synthesis). IOTC–2023–WPB21–19.

Gelman, A., 2006. Prior distributions for variance parameters in hierarchical models. Bayesian Analysis, 1(3): 515–533.

Gelman, A., Rubin, D.B., 1992. Inference from iterative simulation using multiple sequences. Statistical Science, 7: 457–472. https://doi.org/10.1214/ss/1177011136.

Geweke, J., 1992. Evaluating the accuracy of sampling-based approaches to the calculation of posterior moments. In: Berger, J.O., Bernardo, J.M., Dawid, A.P., Smith, A.F.M. (Eds.), Bayesian Statistics 4. Clarendon Press, Oxford, pp. 169–193.

Hurtado-Ferro, F., Szuwalski, C.S., Valero, J.L., Anderson, S.C., Cunningham, C.J., Johnson, K.F., Licandeo, R., McGilliard, C.R., Monnahan, C.C., Muradian, M.L., Ono, K., Vert-Pre, K.A., Whitten, A.R., Punt, A.E., 2014. Looking in the rear-view mirror: Bias and retrospective patterns in integrated, age-structured stock assessment models. ICES Journal of Marine Science, 72 (1): 99–110. https://doi.org/10.1093/icesjms/fsu198.

Martell, S., Froese, R., 2013. A simple method for estimating MSY from catch and resilience. Fish and Fisheries, 14(4): 504–514. https://doi.org/10.1111/j.1467-2979.2012.00485.x

Mohn, R., 1999. The retrospective problem in sequential population analysis: an investigation using cod fishery and simulated data. ICES Journal of Marine Science, 56 (4): 473–488. https://doi.org/10.1006/jmsc.1999.0481.

Mourato, B., Sant'Ana, R., Kikuchi, E., Cardoso, L.G., Ngom, F., Arocha, F., Kimoto, A., Ortiz, M., 2023. Assessment of the eastern Atlantic sailfish stock using JABBA model. Collective Volumes of Scientific Papers ICCAT, 80(8): 315–336.

Parker, D., 2020. Preliminary stock assessment of Indian Ocean swordfish (*Xiphias gladius*) using the Bayesian state-space surplus production model JABBA. IOTC-2020- WPB18-18.

Plummer, M., 2003. JAGS: A Program for Analysis of Bayesian Graphical Models using Gibbs Sampling. 3rd International Workshop on Distributed Statistical Computing (DSC 2003), Vienna, Austria.

Plummer, M., Best, N., Cowles, K., Vines, K., 2006. CODA: Convergence Diagnosis and Output Analysis for MCMC, R News, vol 6, 7-11.

R Core Team, 2025. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/.

Su, Y., Yajima, M., 2024. R2jags: Using R to Run 'JAGS'. doi:10.32614/CRAN.package.R2jags, R package version 0.8-9. https://CRAN.R-project.org/package=R2jags.

Xu, W-Q., Wang, S-P., Lin, C-H., 2023. Stock assessment of swordfish (*Xiphias gladius*) in the Indian Ocean using A Stock-Production Model Incorporating Covariates (ASPIC). IOTC–2023–WPB21–18.

Wickham, H., 2016. ggplot2: Elegant Graphics for Data Analysis. Springer-Verlag, New York.

Wickham, H., Averick, M., Bryan, J., Chang, W., McGowan, L.D., François, R., Grolemund, G., Hayes, A., Henry, L., Hester, J., Kuhn, M., Pedersen, T.L., Miller, E., Bache, S.M., Müller, K., Ooms, J., Robinson, D., Seidel, D.P., Spinu, V., Takahashi, K., Vaughan, D., Wilke, C., Woo, K., Yutani, H., 2019. Welcome to the tidyverse. Journal of Open Source Software, 4(43): 1686. https://doi.org/10.21105/joss.01686

Wickham, H., François, R., Henry, L., Müller, K., Vaughan, D., 2026. dplyr: A Grammar of Data Manipulation. https://doi.org/10.32614/CRAN.package.dplyr, R package version 1.2.0. https://CRAN.R-project.org/package=dplyr.

Wickham, H., Vaughan, D., Girlich, M., 2025. tidyr: Tidy Messy Data. doi:10.32614/CRAN.package.tidyr. R package version 1.3.2. https://CRAN.R-project.org/package=tidyr.

Wilke, C., 2025. cowplot: Streamlined Plot Theme and Plot Annotations for 'ggplot2'. doi:10.32614/CRAN.package.cowplot. https://doi.org/10.32614/CRAN.package.cowplot, R package version 1.2.0.

Winker, H., Carvalho, F., Kapur, M., 2018. JABBA: Just Another Bayesian Biomass Assessment. Fisheries Research, 204: 275–288. https://doi.org/10.1016/j.fishres.2018.03.010

Winker, H., Carvalho, F., Kapur, M., 2026. JABBA: Just Another Bayesian Biomass Assessment. R package version 2.3.2. https://github.com/jabbamodel/JABBA.