

Stock assessment of blue marlin (*Makaira nigricans*) in the Indian Ocean using JABBA

Chih-Yu Lin, Wen-Qi Xu, Sheng-Ping Wang*

Department of Environmental Biology and Fisheries Science, National Taiwan Ocean University, Keelung, Taiwan.

* Corresponding author: wsp@mail.ntou.edu.tw

ABSTRACT

This study assessed the stock status of blue marlin (*Makaira nigricans*) in the Indian Ocean using the Bayesian state–space biomass dynamic model JABBA. Updated catch data (1950–2023) and standardized CPUE indices from Taiwanese, Japanese, and Indonesian longline fleets were analyzed under ten scenarios combining alternative CPUE series, production model types (Schaefer or Fox), and treatments of process error variance. Posterior diagnostics and residual analyses indicated that the Fox model with the Taiwanese sdmTMB index and process error standard deviation (σ_{proc}) fixed at 0.15 (S10) provided the most robust and internally consistent fit. Biomass declined across scenarios since the mid-1980s, with depletion falling below B_{MSY} and fishing mortality exceeding F_{MSY} in recent decades. Kobe plots indicated probabilities exceeding 95% that the stock is both overfished and subject to overfishing, although the Fox model with fixed process error variance produced slightly less pessimistic outcomes compared to the Schaefer alternatives. Projections under constant catch levels further confirmed that rebuilding prospects depend on reducing catches. Overall, the assessment concludes that the Indian Ocean blue marlin stock is overfished and undergoing overfishing, with findings intended to inform future IOTC WPB discussions and management advice.

1. INTRODUCTION

The stock status of blue marlin (*Makaira nigricans*) in the Indian Ocean was evaluated based on the results of the 2022 benchmark assessment, which applied both the Bayesian state-space biomass production model JABBA and the integrated age-structured model Stock Synthesis (SS). Both models produced consistent results, indicating that the stock was overfished ($B_{2020}/B_{MSY} = 0.73$) and subject to overfishing

($F_{2020}/F_{MSY} = 1.13$). According to the weight of evidence, the stock was placed in the red quadrant of the Kobe plot, with an estimated 72% probability of being subject to overfishing and overfished (IOTC, 2022; 2024).

The biomass trajectory (B/B_{MSY}) has shown a long-term declining trend since the mid-1980s. A temporary increase occurred between 2007 and 2012, which coincided with the period of reduced fishing activity in the northwestern Indian Ocean due to piracy, after which the biomass declined again in recent years. Fishing mortality (F/F_{MSY}) has remained above 1 since the mid-1980s, despite some short-term fluctuations.

Based on this evidence, the stock is currently considered to be both overfished and subject to overfishing. Average catches from 2019 to 2023 (7,049 t) were lower than the estimated MSY (8,740 t), although the probability of rebuilding depends on further catch reductions.

This study conducted the stock assessment for blue marlin in the Indian Ocean using the base-case specifications adopted by the IOTC Working Party on Billfish (WPB), consistent with the 2022 assessment. In addition to the reference case, supplementary scenarios were developed using the most recent catch data and updated standardized CPUE indices in order to evaluate the robustness of the assessment outcomes.

2. MATERIALS AND METHODS

2.1. Data used

Catch data from 1950 to 2023 were obtained from the IOTC Secretariat, and the aggregated total catch from all fleets was used in the assessment (Fig. 1).

The standardized CPUE series from Taiwanese, Japanese, and Indonesian longline fleets served as the primary abundance indices for the assessments presented to the IOTC WPB in 2022 and 2025 (Fig. 2). In this study, standardized CPUE indices were available from three sources:

- Taiwanese longline fishery (TWN), separated into two areas and covering the period 2005–2023 (Lin et al., 2025),
- Japanese longline fishery (JPN), separated into two time periods, 1979–2010 and 2011–2023 (Kai, 2025),
- Indonesian longline fishery (IDN), covering 2006–2023 (Setyadji et al., 2025).

For the Taiwanese fleet, CPUE standardization was conducted using both delta-GLM and sdmTMB approaches. Both indices were included in sensitivity analyses,

while the delta-GLM index was adopted as the reference case to ensure continuity with Parker (2022).

2.2. Assessment model

The stock assessment analysis was conducted using JABBA (version 2.3.0), which is implemented as an R package available at github.com/jabbamodel/JABBA. JABBA is a Bayesian state-space biomass production model that integrates catch data and abundance indices within a flexible prior specification framework. A comprehensive description of the model formulation, state-space implementation, prior specification, and diagnostic tools is provided in Winker et al. (2018).

2.3. Model specifications

Following the recommendations of the IOTC WPB, the assessment time series began in 1950, when the stock was assumed to be close to its unfished biomass level (IOTC, 2022). All catchability parameters were specified as uninformative uniform priors. For the continuity scenario adopted in the 2022 stock assessment (IOTC, 2022; Parker et al., 2022), the same range of plausible prior distributions for key parameters was assumed, including carrying capacity (K), intrinsic growth rate (r), Initial depletion (ψ , i.e. B_1/K) and observations error variance, consistent with scenario S3 of Parker (2022) (Table 2):

In the 2022 assessment, only the Schaefer production model ($B_{MSY}/K = 0.5$) was applied, and the process error standard deviation (σ_{proc}) was estimated using an inverse-gamma prior distribution with both scaling parameters set to 0.001. In this study, we extended the model framework to also include the Fox production model ($B_{MSY}/K = 0.38$), as recommended by the IOTC WPB (IOTC, 2022). In addition, we conducted sensitivity analyses in which the σ_{proc} was fixed at plausible values rather than estimated, recognizing that estimating process error variance can sometimes lead to unrealistically large fluctuations in annual biomass trajectories (Ono et al., 2012).

Accordingly, a total of ten scenarios were developed by combining alternative model specifications. These scenarios incorporated different Taiwanese CPUE indices (delta-GLM or sdmTMB), model types (Schaefer or Fox), and treatments of process error variance (either estimated or fixed). The full set of scenarios is summarized in Table 3.

3. RESULTS AND DISCUSSION

3.1. Posterior distribution diagnostics

The marginal posterior and prior distributions for scenarios S1 to S4 were

examined to evaluate the information content of the data under alternative treatments of process error variance ($\sigma.\text{proc}$) (Fig. 3). For the carrying capacity (K), the posterior distributions were consistently narrower than their priors, and the prior-to-posterior variance ratios (PPVRs) were relatively small. This indicates that the data were moderately informative for K . However, the occurrence of multimodal posterior shapes across scenarios suggests that substantial uncertainty remains in estimating this parameter.

For the intrinsic growth rate (r), the posterior distributions shifted away from the priors, with posterior-to-prior mean ratios (PPMRs) ranging from approximately 0.9 to 1.2. These shifts demonstrate that the data provided meaningful updates to the prior assumptions for r . In contrast, the initial depletion parameter (ψ) showed strong prior influence, as the posterior distributions closely resembled the priors. Both PPMR and PPVR values were close to unity, indicating that ψ was largely constrained by the prior rather than informed by the data.

The treatment of process error variance (σ^2) had a notable influence on model behavior. When σ^2 was estimated (S1), the posterior distribution of σ^2 was centered around 0.02 with substantial uncertainty, and the posterior of K exhibited more pronounced multimodality. By contrast, when σ^2 was fixed at plausible values of 0.15, 0.10, or 0.20 (S2–S4), the posterior distributions of other parameters became narrower and more stable. Fixing σ^2 reduced uncertainty in K and r , as indicated by smaller PPVR values, although multimodality in K persisted across all scenarios.

Other scenarios that estimated or fixed the process error variance produced similar qualitative results, and therefore only representative scenarios (S1–S4) are presented here.

3.2. Goodness-of-fit

The model fits to the CPUE data varied across scenarios, as reflected by the root mean squared error (RMSE) values (Fig. 4). Among the Schaefer production model scenarios, S5 (TWN-sdmTMB CPUE with the $\sigma.\text{proc}$ fixed at $\sigma.\text{proc} = 0.15$) provided the best fit, with an RMSE of 19.9%. In contrast, S1–S4 (based on TWN-GLM CPUE) showed relatively poorer fits, with RMSE values generally exceeding 20%, particularly when the process error variance was estimated rather than fixed.

The Fox production model scenarios generally exhibited improved fits compared with the Schaefer model. In particular, S10 (TWN-sdmTMB CPUE with $\sigma.\text{proc}$ fixed at 0.15) yielded the lowest RMSE at 18.5%, representing the best fit across all scenarios. Other Fox scenarios, such as S7 and S9, also demonstrated relatively low RMSE values, further indicating that the Fox model framework tends to better capture the CPUE dynamics.

Overall, the results suggest that the Fox production model provides more robust fits to the CPUE data than the Schaefer model, and that fixing the process error variance yields consistently stable performance. Within the Schaefer framework, S5 appears to be the most reliable configuration. When considering all scenarios, however, S10 represents the optimal fitting configuration, achieving the lowest RMSE and offering the most consistent agreement with the observed CPUE trends.

3.3. Runs tests of CPUE residuals

Runs tests were conducted to evaluate whether the time series of CPUE residuals were randomly distributed across fleets under three representative scenarios (Fig. 5): S1 (the continuity setting from the 2022 assessment, Schaefer model with process error variance estimated), S6 (Schaefer model with σ_{proc} fixed at 0.15, identified as the best-performing Schaefer case), and S10 (Fox model with σ_{proc} fixed at 0.15, identified as the overall optimal scenario).

The runs tests of CPUE residuals revealed clear differences among these scenarios. In S1, significant departures from randomness were evident for the Taiwanese western CPUE index derived from the delta-GLM and for the Japanese early-period index, suggesting temporal structure in these residuals. By contrast, the Taiwanese eastern CPUE index and the Indonesian index were consistent with randomness. When σ_{proc} was fixed at 0.15 in S6, residual diagnostics improved, as the Taiwanese western index no longer displayed significant departures from randomness, although the Japanese early-period index continued to exhibit temporal correlation. Other indices (Taiwanese eastern, Japanese later-period, and Indonesian) remained consistent with random variation. The most favorable outcomes were observed in S10, which applied the Fox production model with σ_{proc} fixed at 0.15. In this scenario, all CPUE indices except the Japanese early-period index passed the randomness test, indicating that the Fox model combined with a fixed process error variance provided the most robust fit with minimal signs of temporal structure compared to the Schaefer alternatives.

Overall, these results suggest that the Japanese early-period CPUE series is the primary source of non-random residual structure across scenarios, whereas both Taiwanese and Indonesian indices demonstrated greater stability. Moreover, the Fox model framework combined with a fixed $\sigma_{\text{proc}} = 0.15$ (S10) provided the most consistent residual diagnostics across fleets.

3.4. Estimates of key quantities

The temporal trajectories of biomass, fishing mortality, process deviations, and management reference ratios (B/B_0 , B/B_{MSY} , and F/F_{MSY}) displayed consistent overall patterns across scenarios S1–S10, although notable differences were evident depending

on the model type and treatment of process error variance (Fig. 6).

Biomass estimates showed a gradual decline from the 1950s, with more pronounced reductions after the mid-1980s, driven by the sustained increase in fishing mortality. Depletion (B/B_0) and B/B_{MSY} followed similar trajectories, indicating a decline below the MSY-based reference level in the late 1990s to early 2000s, after which both indices remained at low levels across all scenarios. Conversely, fishing mortality and F/F_{MSY} increased steadily after the 1980s, with F/F_{MSY} consistently above unity in recent decades, suggesting persistent overfishing.

Comparisons across model configurations revealed that the Fox production model generally produced higher and more stable biomass trajectories and smoother estimates of process deviations compared to the Schaefer model. Within the Schaefer framework, scenarios with fixed process error variance, particularly S6 ($\sigma.\text{proc} = 0.15$), provided more stable trajectories relative to those in which $\sigma.\text{proc}$ was estimated. The Fox scenarios (S7–S10) exhibited greater consistency, with S10 (Fox model with $\sigma.\text{proc}$ fixed at 0.15) emerging as the most stable and reliable configuration, characterized by reduced variability in biomass, depletion, and process deviations.

Table 4 provides quantitative estimates of key parameters, including MSY , F_{2023} , F_{MSY} , B_0 , B_{2023} , B_{MSY} , and ratios of B_{2023}/B_0 , B/B_{MSY} , and F_{2023}/F_{MSY} across scenarios. The results indicate that while absolute values varied with model assumptions, the overall patterns were consistent with biomass depletion and elevated fishing mortality. The Schaefer scenarios with estimated process error variance (e.g., S1) tended to produce more variable estimates, whereas fixing $\sigma.\text{proc}$ (e.g., S6) improved stability. In contrast, the Fox scenarios (S7–S10) generally yielded lower estimates of MSY and B_0 but provided more stable and internally consistent outputs. Among them, S10 (Fox model with $\sigma.\text{proc}$ fixed at 0.15) offered the most balanced and robust estimates across key quantities.

The Kobe plot (Fig. 7) further illustrates the stock status across scenarios. All scenarios consistently placed the stock in the overfished and overfishing quadrant ($B/B_{MSY} < 1$ and $F_{2023}/F_{MSY} > 1$). However, outcomes differed between model frameworks: the Schaefer scenarios typically indicated more pessimistic conditions, while the Fox scenarios provided slightly more optimistic outcomes, particularly S10, which placed the stock closest to the boundary line. Taken together, these findings suggest that while uncertainty remains regarding the absolute scale of biomass and MSY , the weight of evidence strongly indicates that the Indian Ocean blue marlin stock is both overfished and subject to ongoing overfishing.

Kobe plots with confidence surfaces around the 2023 estimates (Fig. 8) reinforced these findings, showing probabilities greater than 95% that the stock is simultaneously overfished ($B_{2023} < B_{MSY}$) and subject to overfishing ($F_{2023} > F_{MSY}$). Among scenarios,

the Fox model with the sdmTMB index and fixed process error variance (S10) yielded slightly less pessimistic outcomes than the Schaefer alternatives (S1 and S6), placing the stock marginally closer to the threshold.

3.5 Hindcasting cross-validation

Hindcasting cross-validation (HCxval) was conducted for three representative scenarios (Fig. 9): S1 (the continuity baseline from the 2022 assessment, Schaefer model with the Taiwanese delta-GLM index and estimated process error variance), S6 (Schaefer model with the Taiwanese sdmTMB index and σ_{proc} fixed at 0.15, identified as the optimal-performing Schaefer case), and S10 (Fox model with the Taiwanese sdmTMB index and σ_{proc} fixed at 0.15, identified as the overall optimal scenario). One-year-ahead forecasts of CPUE values (2010–2023) were compared with observed data, and prediction skill was evaluated using the mean absolute scaled error (MASE). Results showed that S1 produced mixed performance, with good predictive skill for the Taiwanese western index but poor accuracy for the eastern index. In S6 the use of sdmTMB indices improved stability for the western index, but prediction errors for the eastern index remained high, suggesting structural limitations of the Schaefer framework. By contrast, S10 produced the most consistent predictive performance across fleets, with all indices except the Taiwanese eastern series achieving MASE values below or close to 1.0.

The comparisons reveal three insights. First, sdmTMB-based CPUE indices generally yield more stable predictions than delta-GLM, although their performance differs by fleet. Second, the Fox production model provides structurally more robust fits than the Schaefer model, particularly in reducing systematic prediction bias. Third, fixing σ_{proc} at 0.15 improves predictive reliability compared with estimating the variance. Compared to S1, which represents the 2022 continuity baseline, S10 offers a more balanced and consistent predictive framework. These findings support the use of sdmTMB-standardized CPUE indices in combination with the Fox model and fixed $\sigma_{\text{proc}} = 0.15$ as the optimal configuration for providing robust scientific advice in future assessments. Importantly, these hindcasting results are consistent with the goodness-of-fit and residual diagnostics, further reinforcing S10 as the most reliable reference case scenario.

3.6 Projections of future stock status

Projections of blue marlin stock status under constant future catch levels ranging from 40% to 160% of the recent catch (2020–2023 average) revealed consistent outcomes across scenarios S1, S6 and S10, although the magnitude of recovery or decline varied depending on the model specification (Fig. 10).

Under the current catch level (100%), the stock is projected to remain in an overfished and overfishing state, with biomass (B/B_{MSY}) staying below unity and fishing mortality (F/F_{MSY}) above unity throughout the projection period. Catch levels at or above 120% are expected to further exacerbate stock depletion, driving biomass to lower levels and pushing fishing mortality well above the sustainable threshold. By contrast, catch reductions to 80% or lower lead to gradual stock rebuilding, with the potential for B/B_{MSY} to exceed unity after 2030, particularly under the Fox model (S10).

Comparisons among scenarios highlighted the influence of model type and process error variance specification. S1 (Schaefer with estimated process error variance) projected relatively slower rebuilding and greater variability in outcomes, reflecting higher uncertainty. S6 (Schaefer with σ_{proc} fixed at 0.15) provided more stable trajectories but still indicated limited potential for rebuilding under current catch levels. In contrast, S10 (Fox model with σ_{proc} fixed at 0.15) consistently produced the most optimistic projections, showing faster recovery of biomass and earlier declines in fishing mortality under reduced catch scenarios.

These projections suggest that meaningful recovery of blue marlin in the Indian Ocean is conditional on substantial reductions in catch, and that differences in model structure and process error treatment strongly influence the projected rebuilding potential.

3.7 Reference scenarios selected by the WPB

Based on the performance of CPUE fits, prediction skill and Mohn's ρ (retrospective analysis of the influence of adding new data) shown in Table 5, S10 is selected as the base scenario and recommended as the reference case. Prediction skill (MASE) comparisons indicate that S1 performs well for the Taiwanese western CPUE but poorly for the eastern index, while S6 improves the western index but eastern errors remain high. S10 shows the most consistent predictive skill, with all indices except the Taiwanese eastern series achieving $MASE \leq 1.0$. Key insights are that the sdmTMB CPUE indices provide more stable predictions than the delta-GLM series, the Fox production model is structurally more robust and less prone to prediction bias than the Schaefer model, and fixing σ_{proc} at 0.15 improves predictive reliability. Overall, S10 (Fox model with the Taiwanese sdmTMB index and σ_{proc} fixed at 0.15), as also suggested by the 2022 WPB, offers the most balanced and consistent predictive framework and is therefore recommended as the optimal reference scenario for future assessments. In contrast, the Schaefer model is considered unsuitable for routine stock assessments because it ignores key age- and size-structure dynamics, which can bias results (Wang et al., 2014; Winker et al.,

2018).

3.8 Additional scenarios requested by the WPB

Building on the S10 reference case, the WPB requested a set of additional sensitivity analyses to evaluate how alternative data sets and model assumptions might influence the assessment results (Tables 5 and 6). These scenarios varied the CPUE series included and certain population–model parameters (Table 7). Specifically, 10a excluded the Taiwanese (TWN) CPUE series, 10b excluded the Japanese late-period and Indonesian CPUE series, 10c used an estimated CV for observation error instead of a fixed value, 10d applied the Fox model with ϕ fixed at 0.99, and 10e applied the Pella-m model with the shape parameter m estimated. Except where otherwise noted, all scenarios retained the Fox or Pella-m production model and a fixed process error standard deviation of 0.15. These additional runs were designed to test the robustness of the S10 conclusions under different CPUE inputs and structural assumptions, providing the WPB with a clearer understanding of the sensitivity of the stock assessment outcomes.

Among the additional WPB scenarios, the CPUE fit indicators show that S10b achieved the lowest standardized RMSE (0.125), followed by S10c (0.167) and S10 (0.171), all of which indicate very good fits. Bayesian p -values were high across the three scenarios (>0.83), with S10b extremely high (0.99), suggesting a possible tendency toward over-smoothed fits. However, in terms of predictive skill, S10b had a MASE of 1.41 (>1), a clear sign of weaker predictive ability, whereas S10 and S10c had MASE values around 0.85–0.88, reflecting good prediction performance. Mohn’s ρ values for B , F , B_{MSY} , F_{MSY} , $procB$, and MSY were all within ± 0.15 for these scenarios, which is generally considered acceptable, and although S10b produced the lowest ρ for F_{MSY} (0.036), the overall differences among S10, S10b, and S10c were minor. Taken together, S10b shows the best standardized RMSE (Std_RMSE) and Bayesian p but suffers from weak predictive skill; S10c is well balanced with low Std_RMSE, good MASE, and stable Mohn’s ρ ; and S10 is close to S10c, with slightly higher Std_RMSE but stable MASE and the advantage of being the established reference scenario.

Using estimated CV, excluding the JPN-late and IDN indices, fixing ϕ with an extremely small CV, and estimating the shape parameter m all helped reduce uncertainty and sharpen the posterior estimates of K and r . Considering these improvements together with the overall model-performance statistics (Tables 5–6; Fig. 11), scenario S10c was selected as the most plausible sensitivity case, while S10 remains the primary reference scenario for the IOTC WPB assessment.

These tables present the final results for the WPB-selected scenario S10c. Table 8

summarizes the estimated 2023 stock status of Indian Ocean blue marlin from the JABBA assessment, including MSY (8.35 thousand t), F_{MSY} (0.30 yr^{-1}), unfished biomass B_0 (≈ 75.9 thousand t), 2023 biomass B_{2023} (≈ 17.5 thousand t), and key ratios ($B_{2023}/B_0 \approx 0.23$, $B_{2023}/B_{MSY} \approx 0.62$, $F_{2023}/F_{MSY} \approx 1.54$). Tables 9 and 10 show the projected probabilities of meeting or violating MSY-based management targets under a range of constant catch scenarios (60–140 % of the recent 2020–2023 average catch of 7,262 t) over 3- and 10-year horizons. These results indicate that only substantial catch reductions (≈ 80 % or lower of recent catch) would give a high probability of rebuilding the stock to B_{MSY} and achieving the Kobe “green” quadrant within the next decade.

REFERENCE

- IOTC Secretariat, 2025. Annual amount of fish caught and retained. IOTC-2025-WPB23-DATA02.
- IOTC, 2022. Report of the 19th Session of the IOTC Working Party on Billfish. 13–16 September 2021, Microsoft Teams Online. IOTC–2021–WPB19–R[E].
- IOTC, 2024. Report of the 27th Session of the IOTC Scientific Committee. 2–6 December 2024, South Africa. IOTC–2024–SC27–R[E].
- Kai, M., 2025. Spatio-temporal model for CPUE standardization: application to blue marlin caught by Japanese tuna longline fishery in the Indian ocean from 1979 to 2023. IOTC-2025-WPB23-16.
- Ono, K., Punt, A.E., Rivot, E. 2012. Model performance analysis for Bayesian biomass dynamics models using bias, precision and reliability metrics. Fish. Res. 125: 173–183.
- Parker, D. 2022. Updated stock assessment of blue marlin (*Makaira nigricans*) in the Indian Ocean using JABBA. IOTC–2022–WPB20–12.
- Setyadji, B., Spencer, M., Kell, L., Wright, S., Ferson, S., 2025. Update on CPUE Standardization of Blue Marlin (*Makaira nigricans*) from Indonesian Tuna Longline Fleets 2006-2023. IOTC-2025-WPB23-15
- Winker, H., Carvalho, F., Kapur, M., 2018. JABBA: Just Another Bayesian Biomass Assessment. Fish. Res. 204, 275–288.
- Xu, W.Q., Lin, C.Y., Wang, S.P., 2025. CPUE standardization of blue marlin (*Makaira nigricans*) caught by the Taiwanese large-scale longline fishery in the Indian Ocean using GLM and sdmTMB. IOTC-2025-WPB23-17.

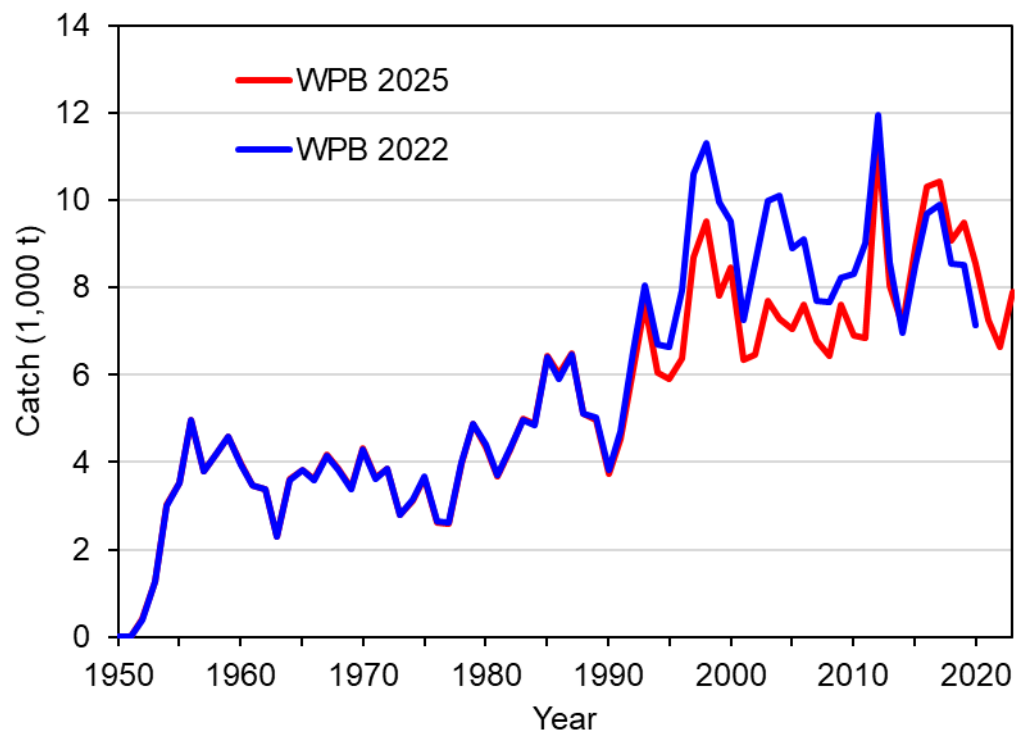


Fig. 1. Time-series of estimated catch in metric tons (t) for Indian Ocean blue marlin.

Taiwanese fleet

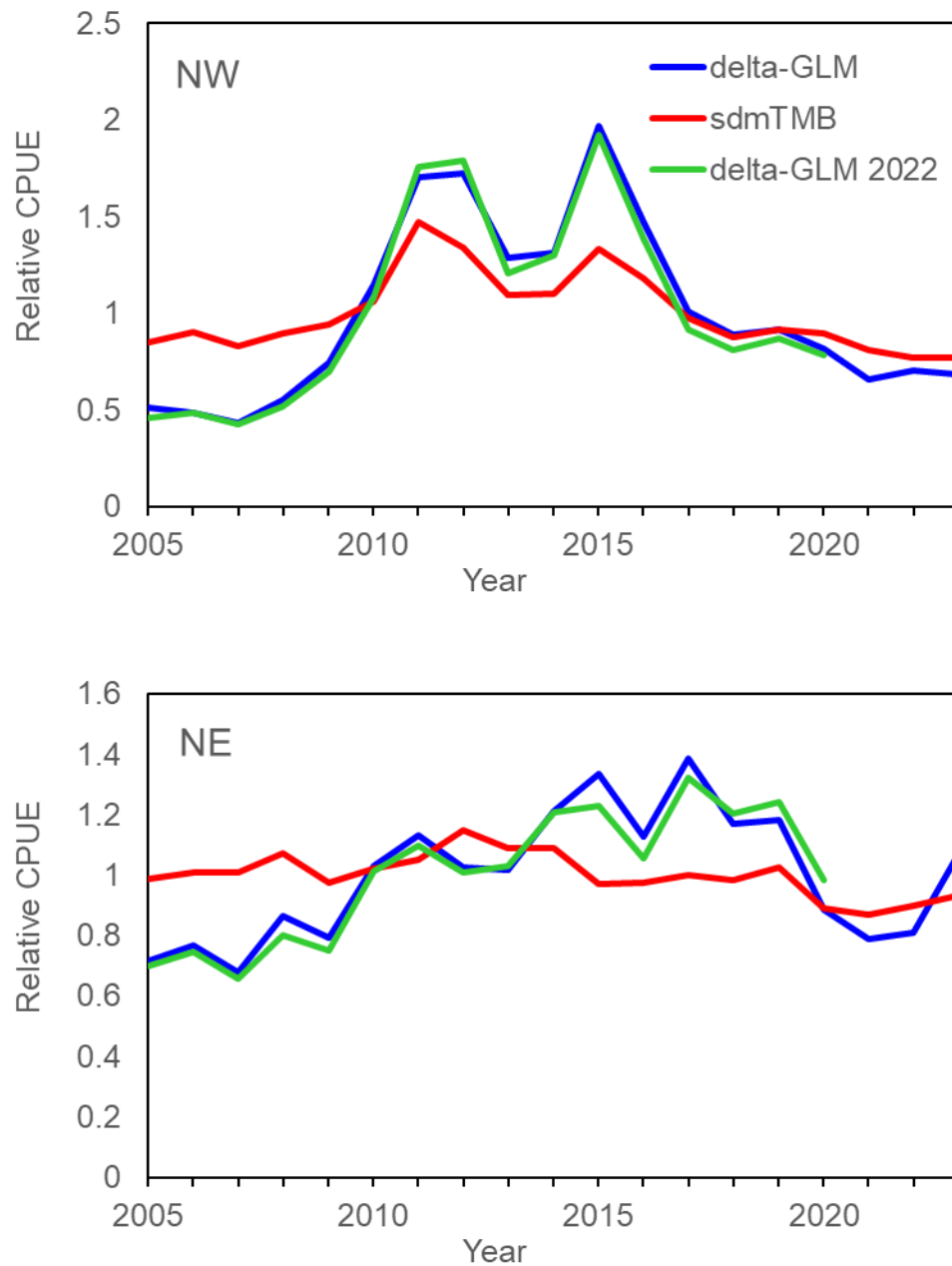
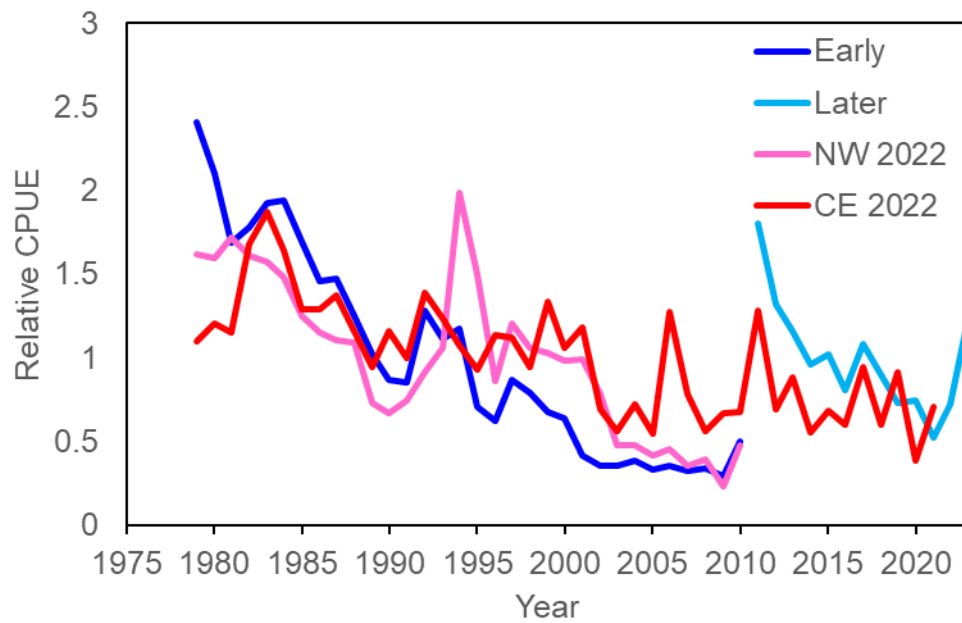


Fig. 2. Standardized CPUE series from Taiwanese, Japanese, and Indonesian longline fleets used in the stock assessments of blue marlin in the Indian Ocean for the IOTC WPB in 2022 and 2025.

Japanese fleet



Indonesian fleet

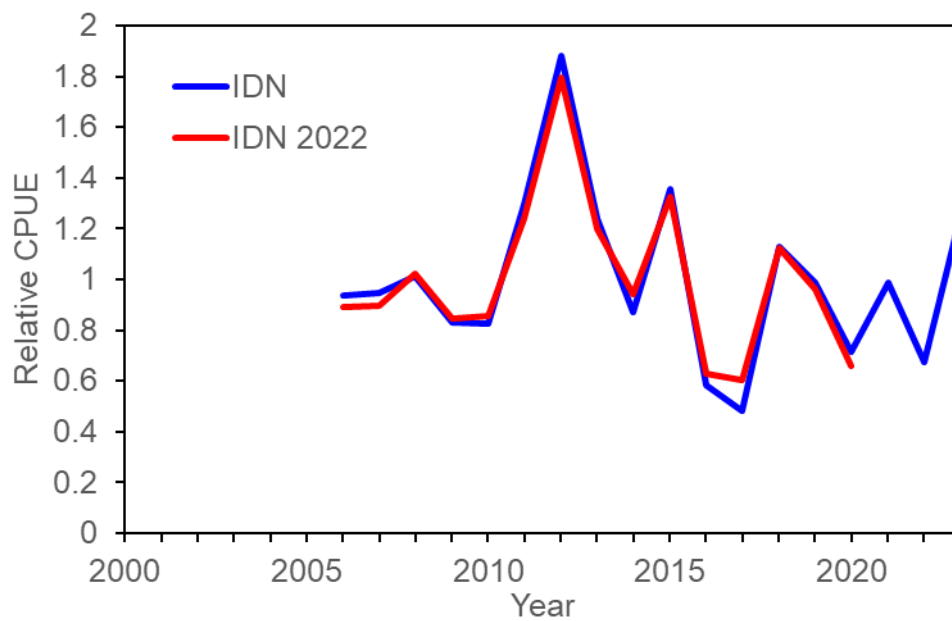
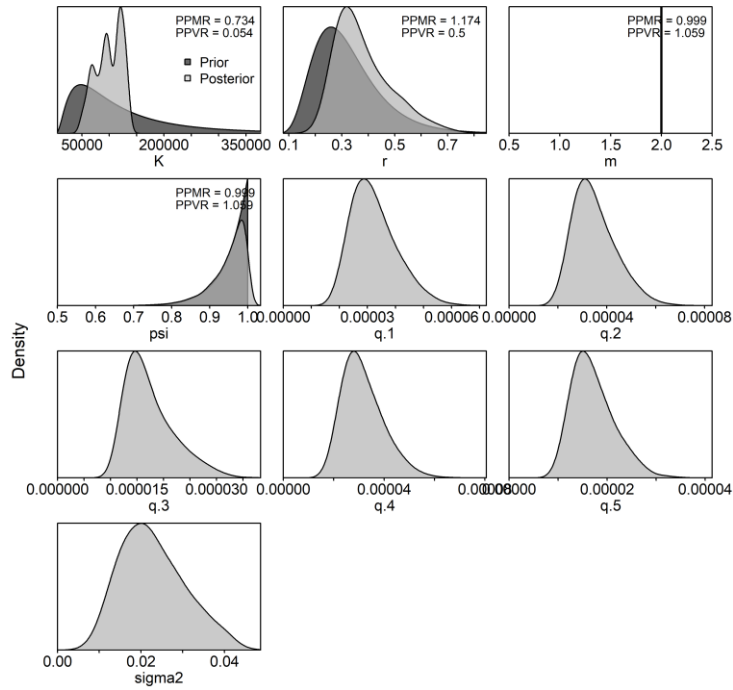


Fig. 2. (Continued).

S1



S2

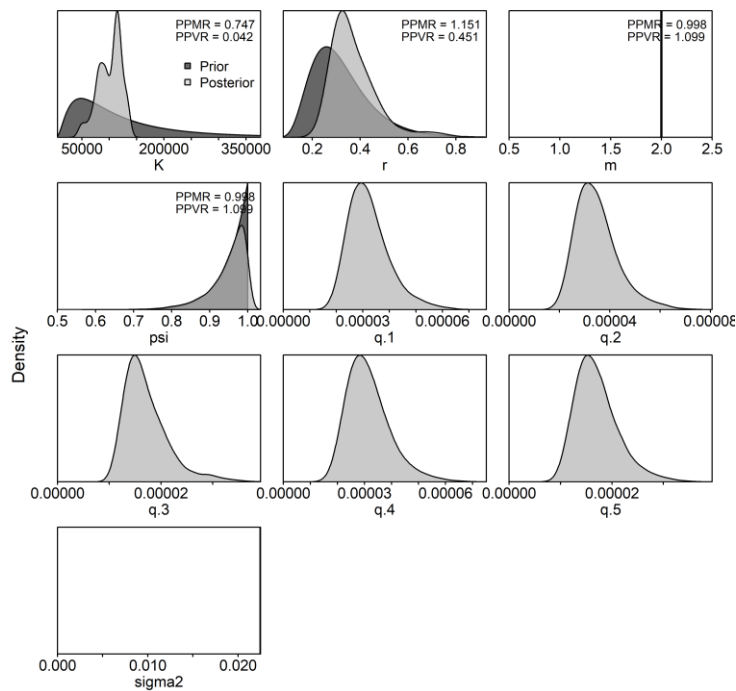
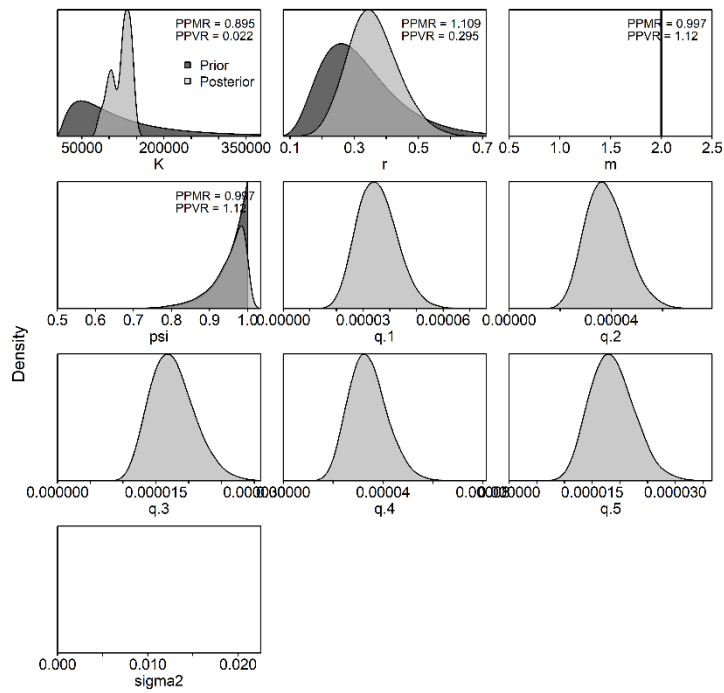


Fig. 3. Prior and posterior distributions for scenario S1~S4 for blue marlin in the Indian Ocean. PPMR: Posterior to Prior Ratio of Means; PPVR: Posterior to Prior Ratio of Variances.

S3



S4

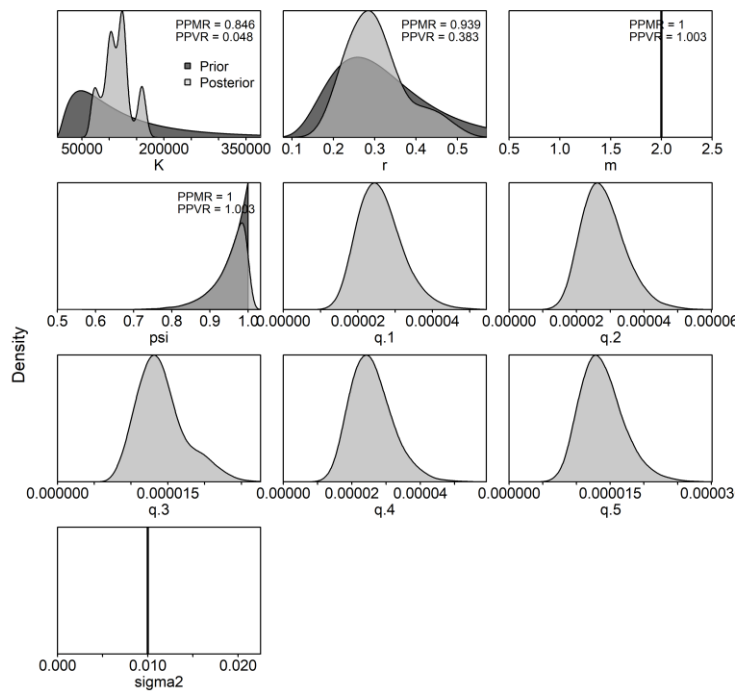


Fig. 3. (Continued).

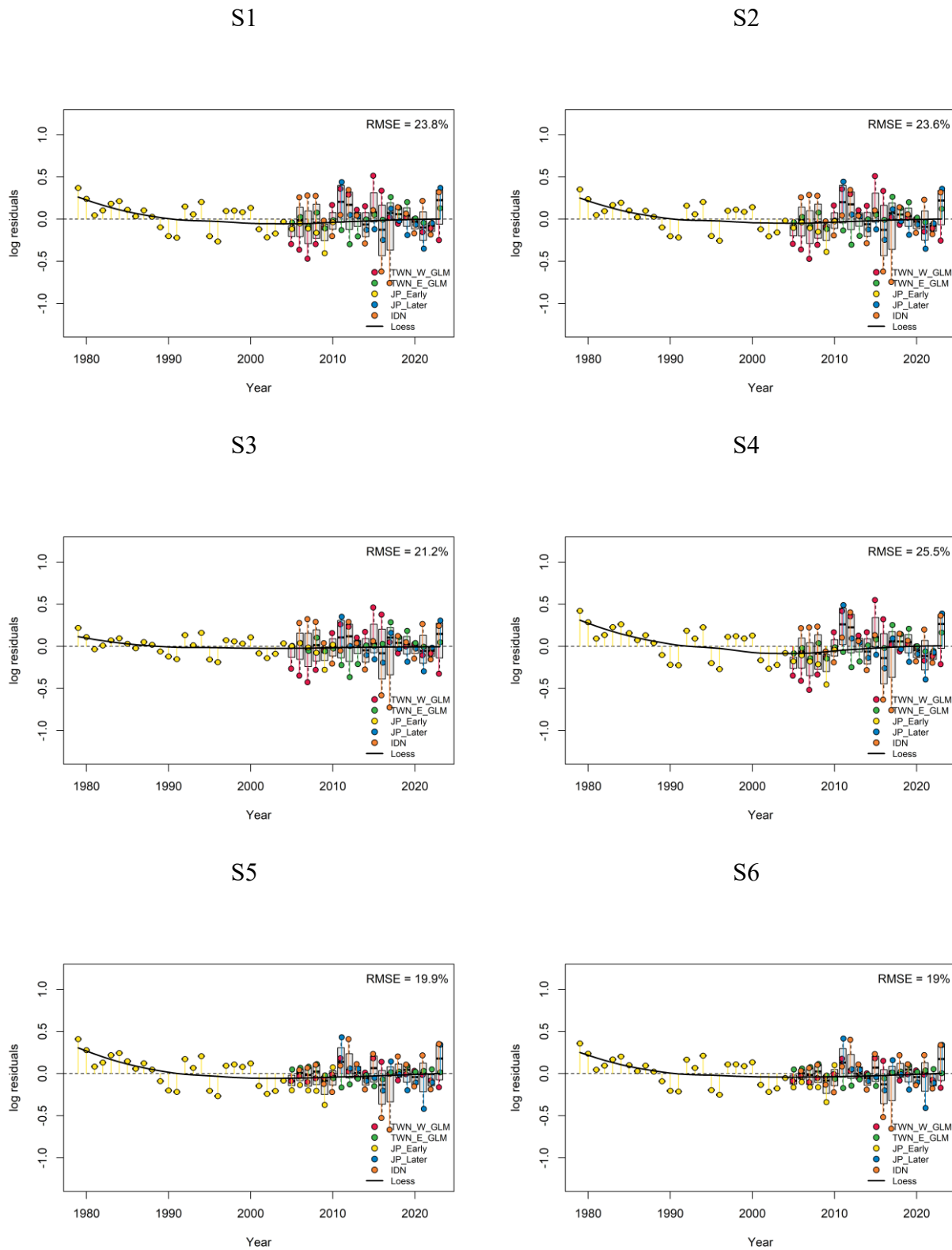
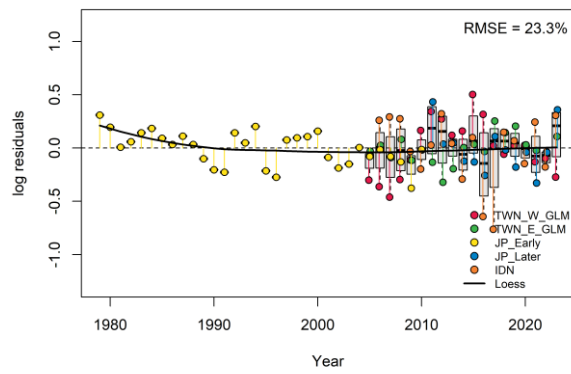
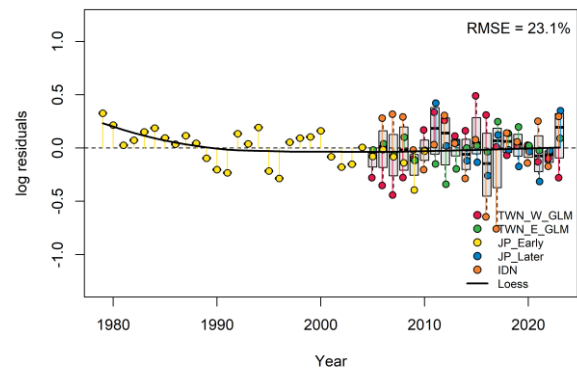


Fig. 4. Residual diagnostic plots of JABBA for all scenarios for CPUE indices for blue marlin in the Indian Ocean. Boxplots indicating the median and quantiles of all residuals available for any given year, and solid black lines indicate a loess smoother through all residuals.

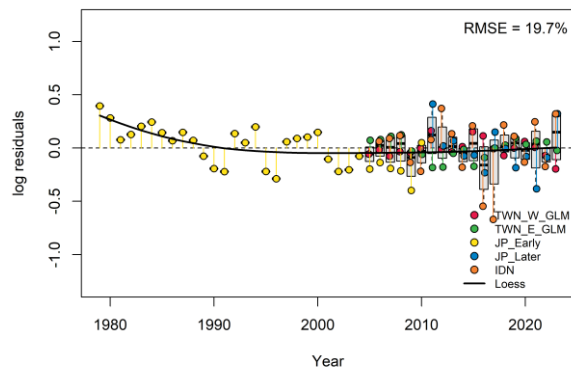
S7



S8



S9



S10

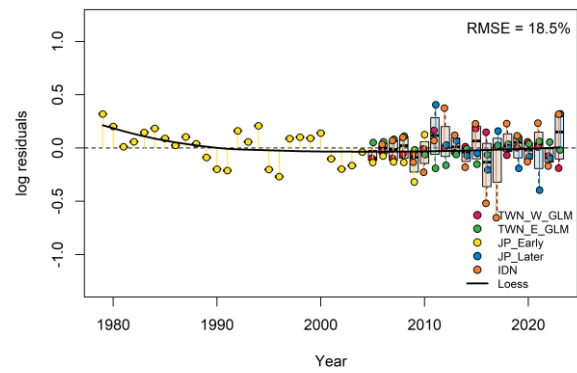


Fig. 4. (continued).

S1

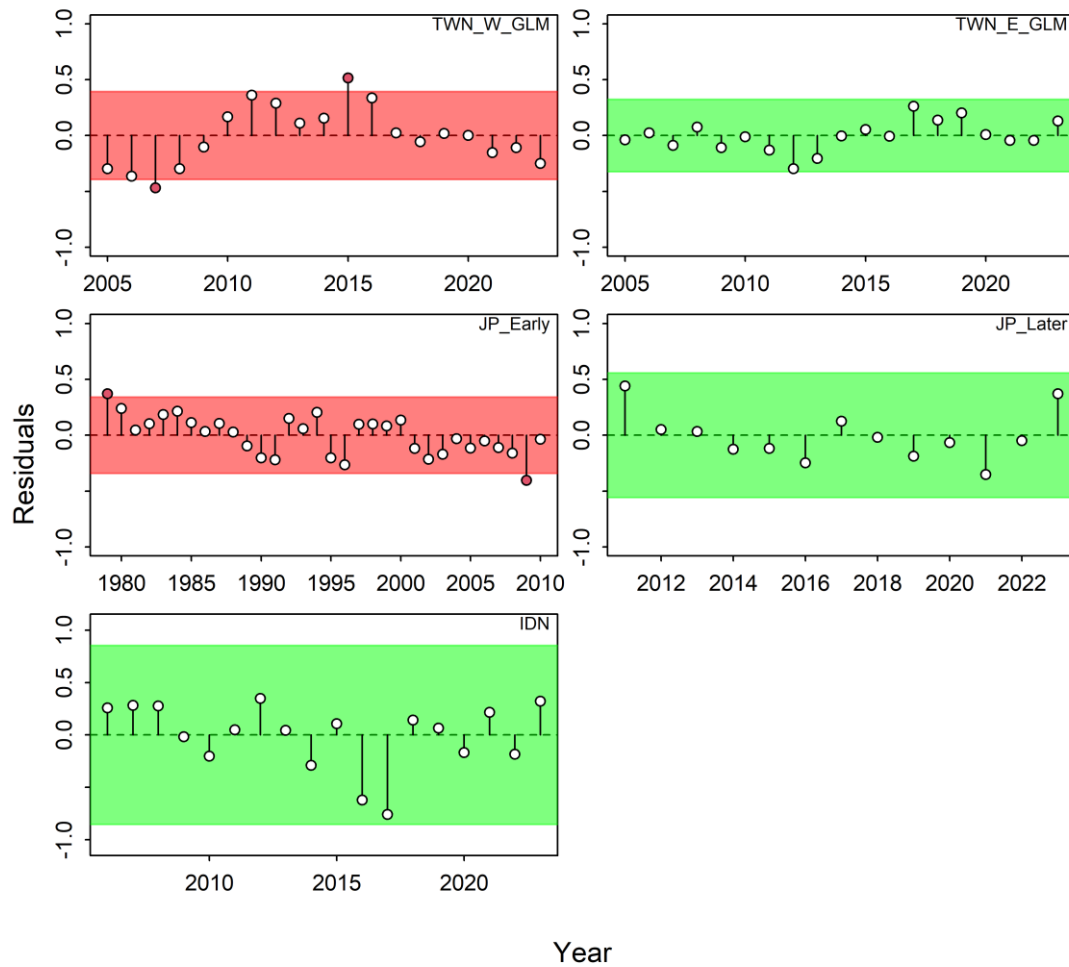


Fig. 5. Runs tests of JABBA for the randomness of the time series of CPUE residuals by fleet for blue marlin in the Indian Ocean under scenarios S1, S6 and S10. Green panels indicate no evidence of lack of randomness of time series residuals ($p > 0.05$) while red panels indicate the opposite. The inner shaded area shows three standard errors from the overall mean and red circles identify a specific year with residuals greater than this threshold value (3x sigma rule).

S6

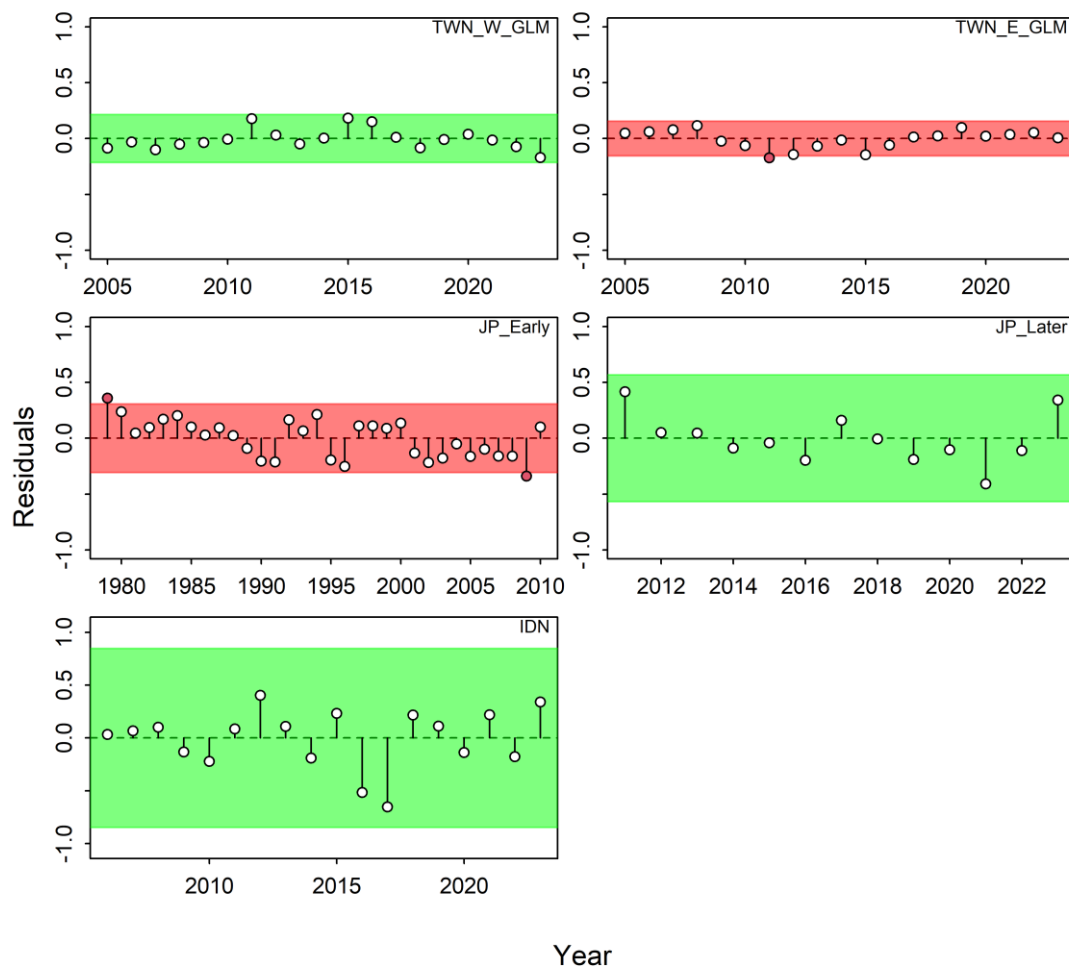


Fig. 5. (Continued).

S10

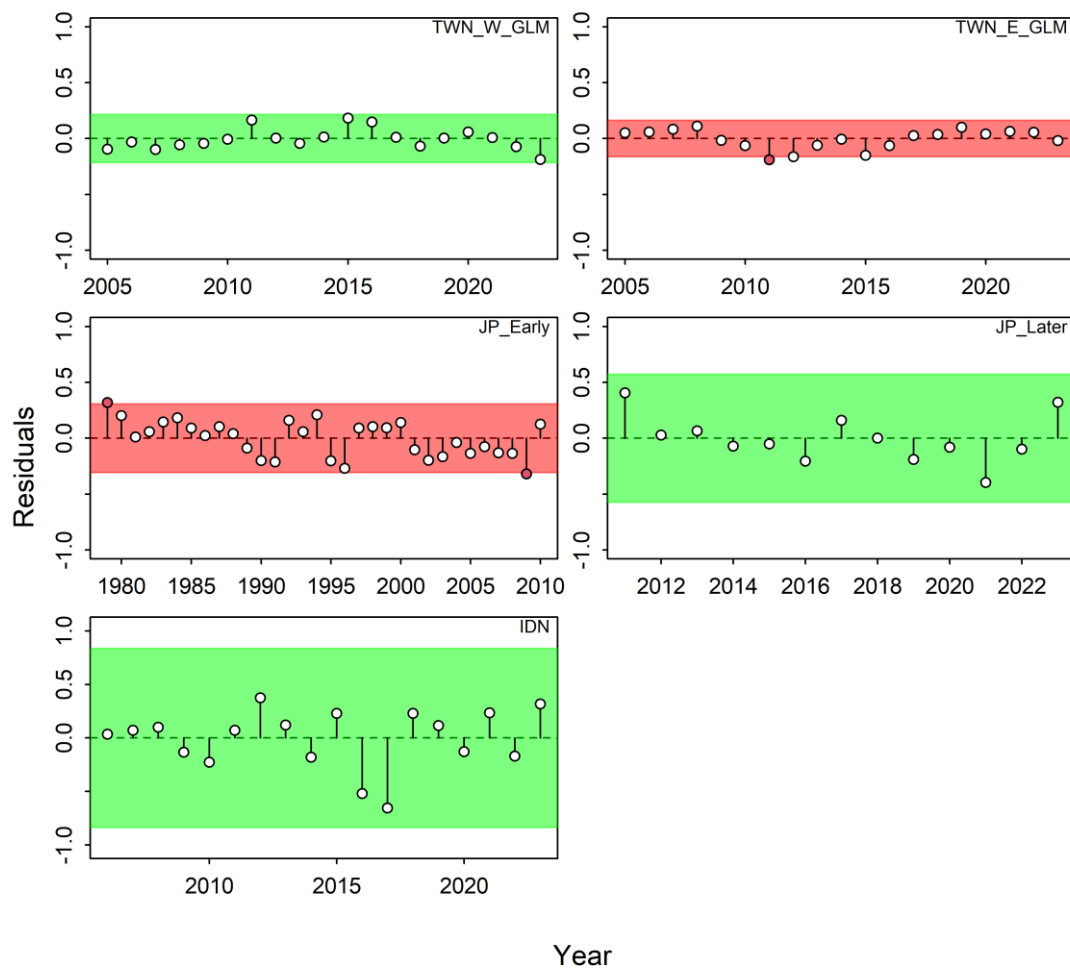


Fig. 5. (Continued).

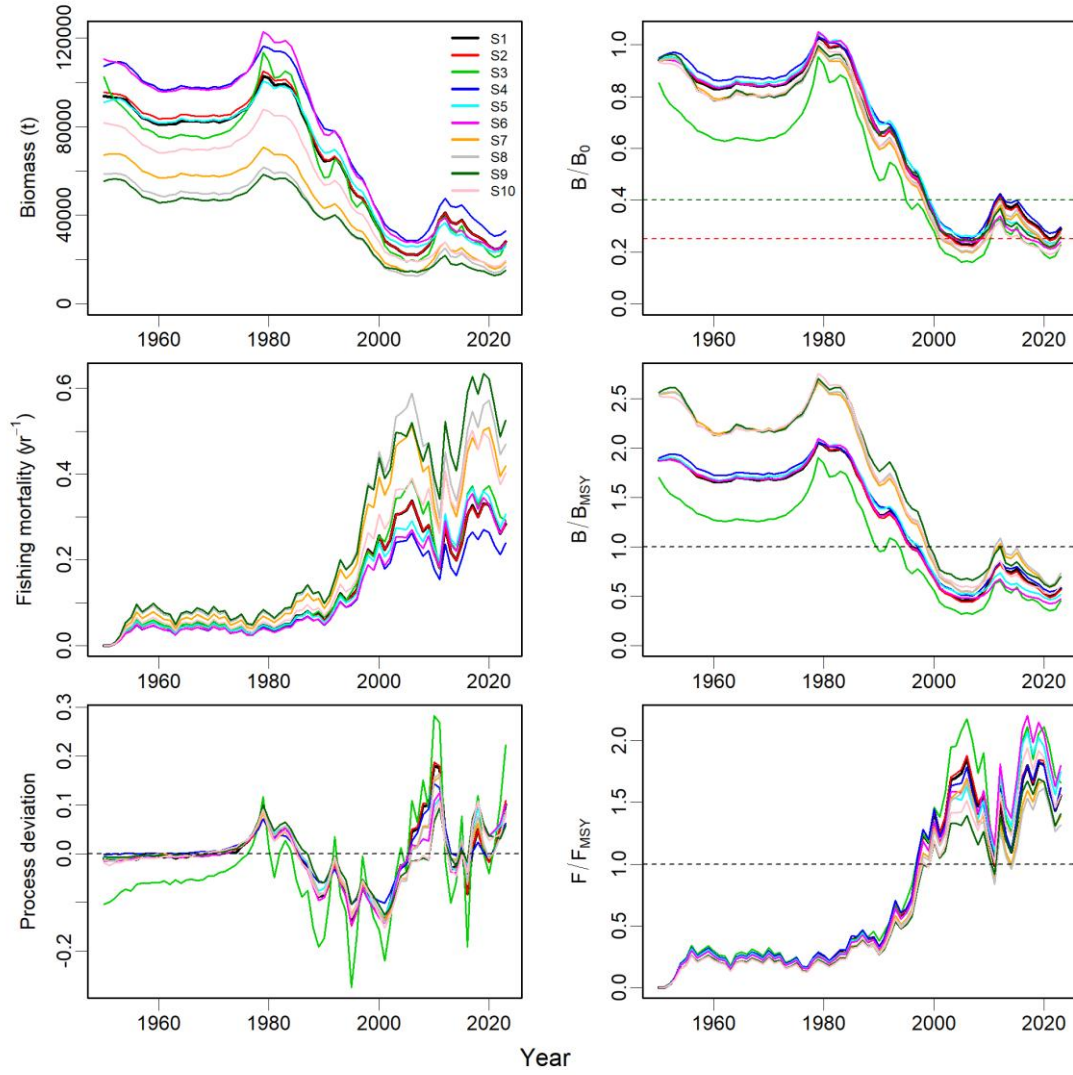


Fig. 6. Temporal trajectories of biomass, fishing mortality, process deviations, and management reference ratios (B/B_0 , B/B_{MSY} , and F/F_{MSY}) for blue marlin in the Indian Ocean, estimated by JABBA under all scenarios.

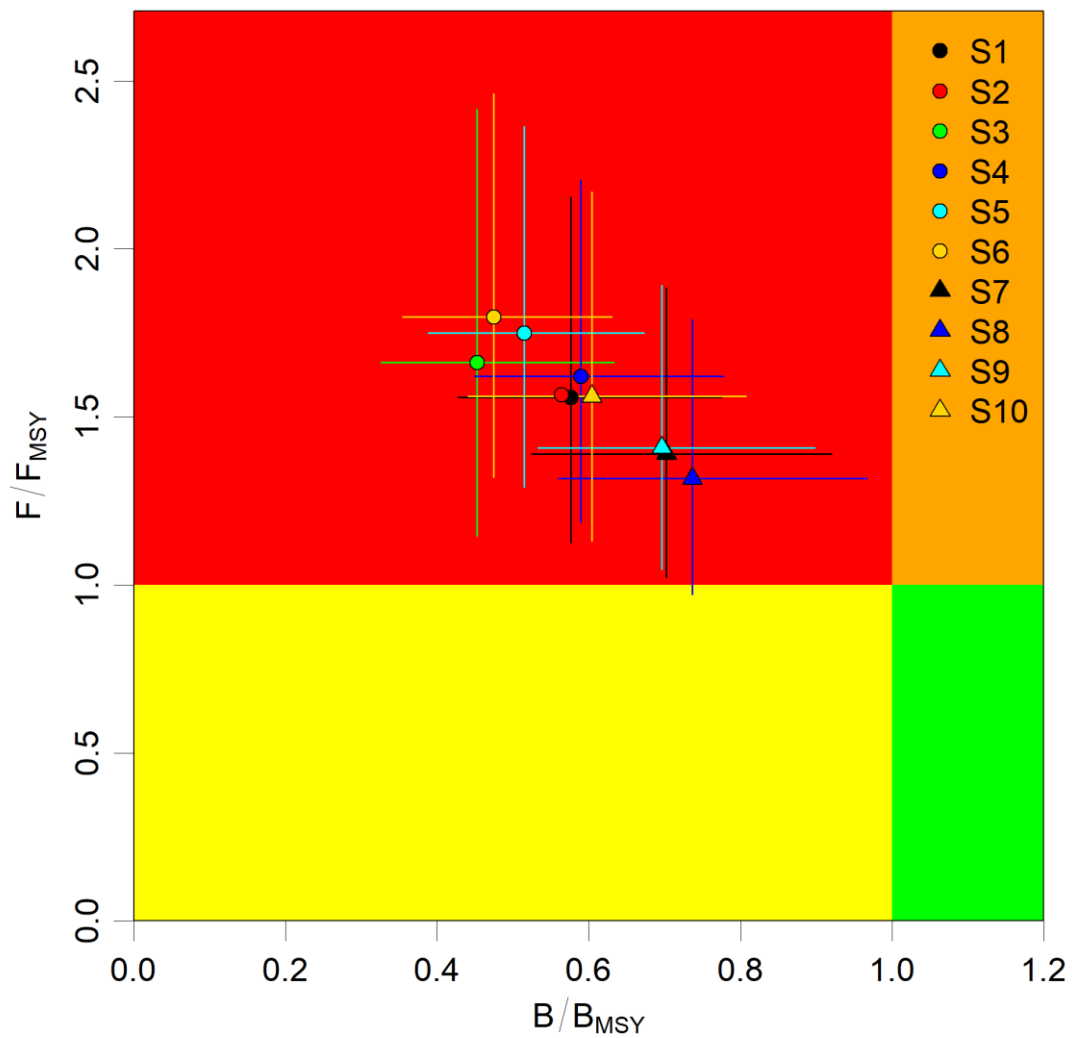
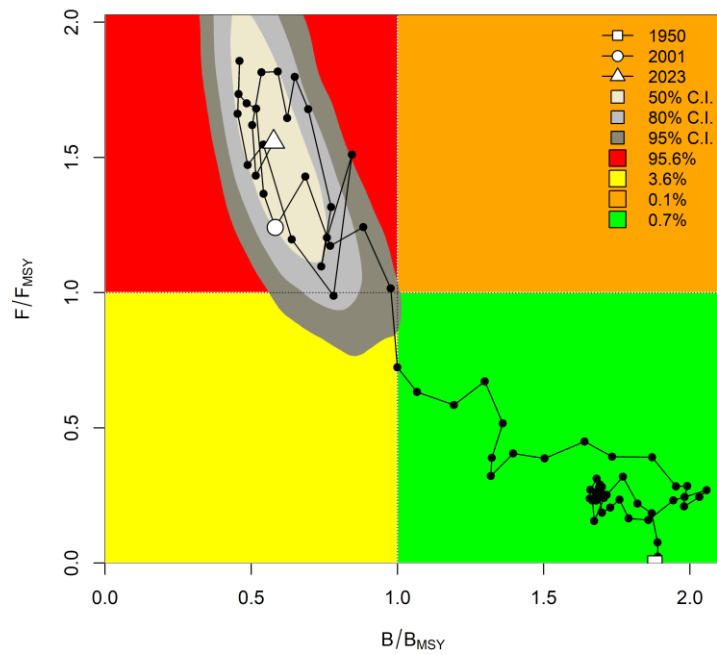


Fig. 7. Kobe plot of 2023 estimates of spawning biomass and fishing mortality relative to their MSY reference points from all scenarios for blue marlin in the Indian Ocean. The error bars represent the 80% confidence interval of the estimates.

S1



S6

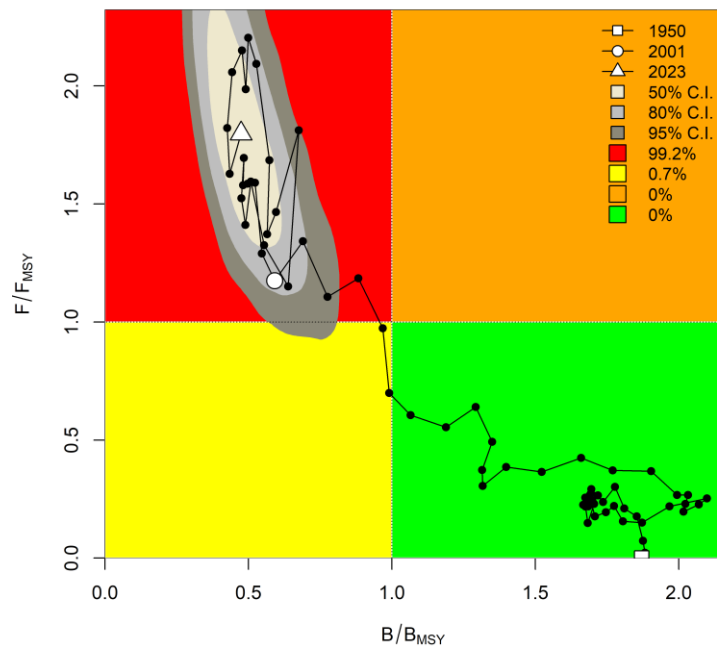


Fig. 8. Kobe plot with confidence surfaces around the 2023 estimates of stock status for blue marlin in the Indian Ocean, obtained from JABBA under scenarios S1, S6 and S10.

S10

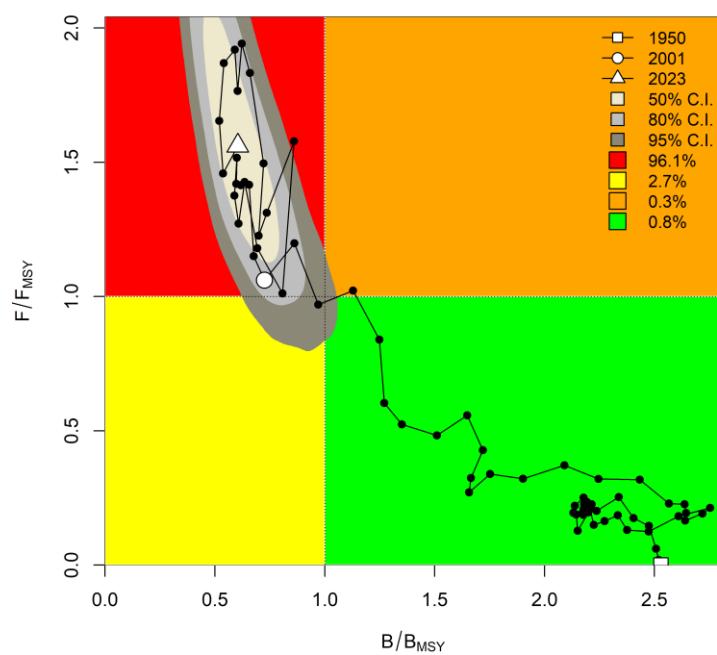


Fig. 8. (Continued).

S1

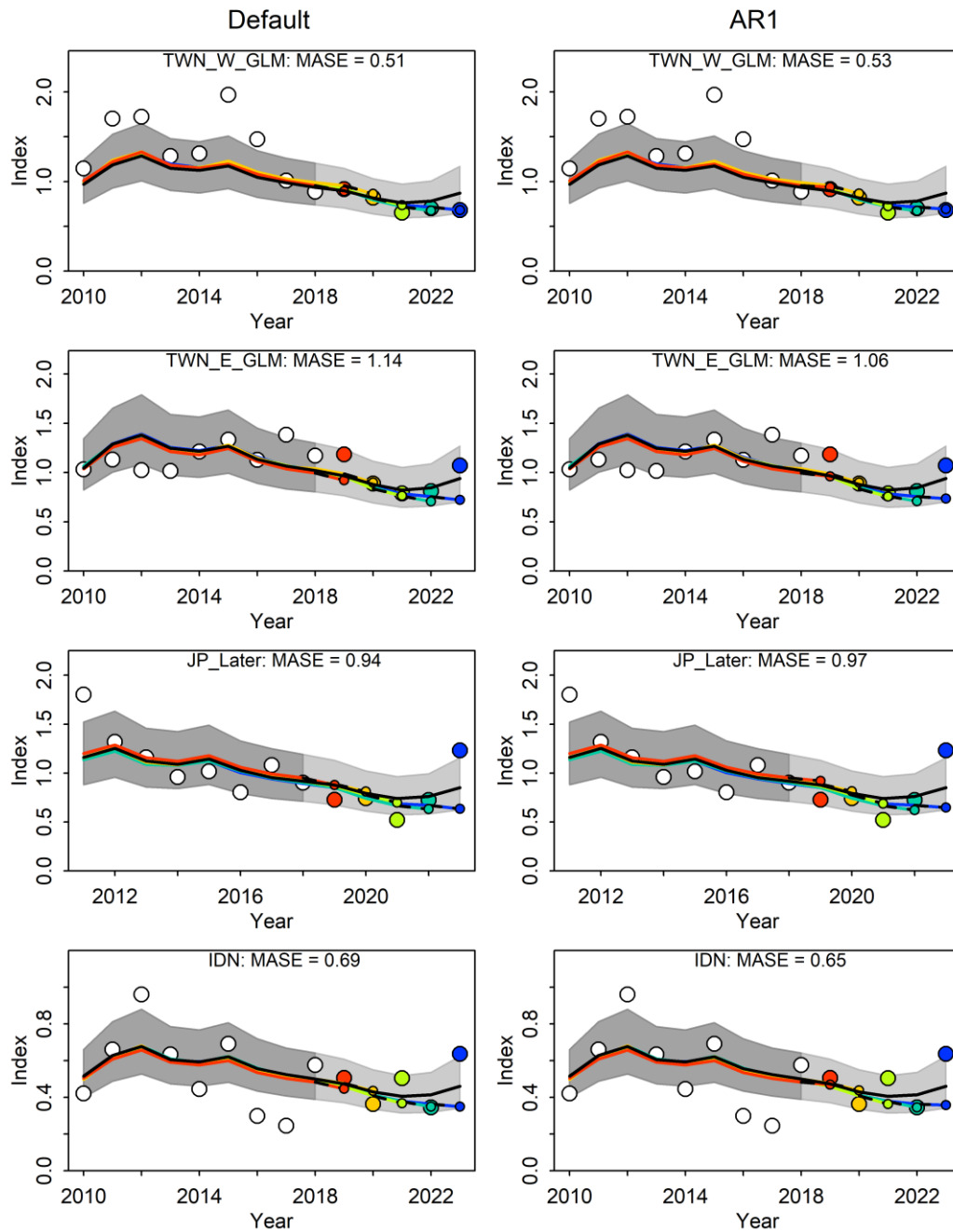


Fig. 9. Hindcasting cross-validation (HCxval) of JABBA for blue marlin in the Indian Ocean under scenarios S1, S6, and S10, showing one-year-ahead forecasts of CPUE values (2010–2023) with observed CPUE indices indicated as color-coded points and 95% confidence intervals.

S6

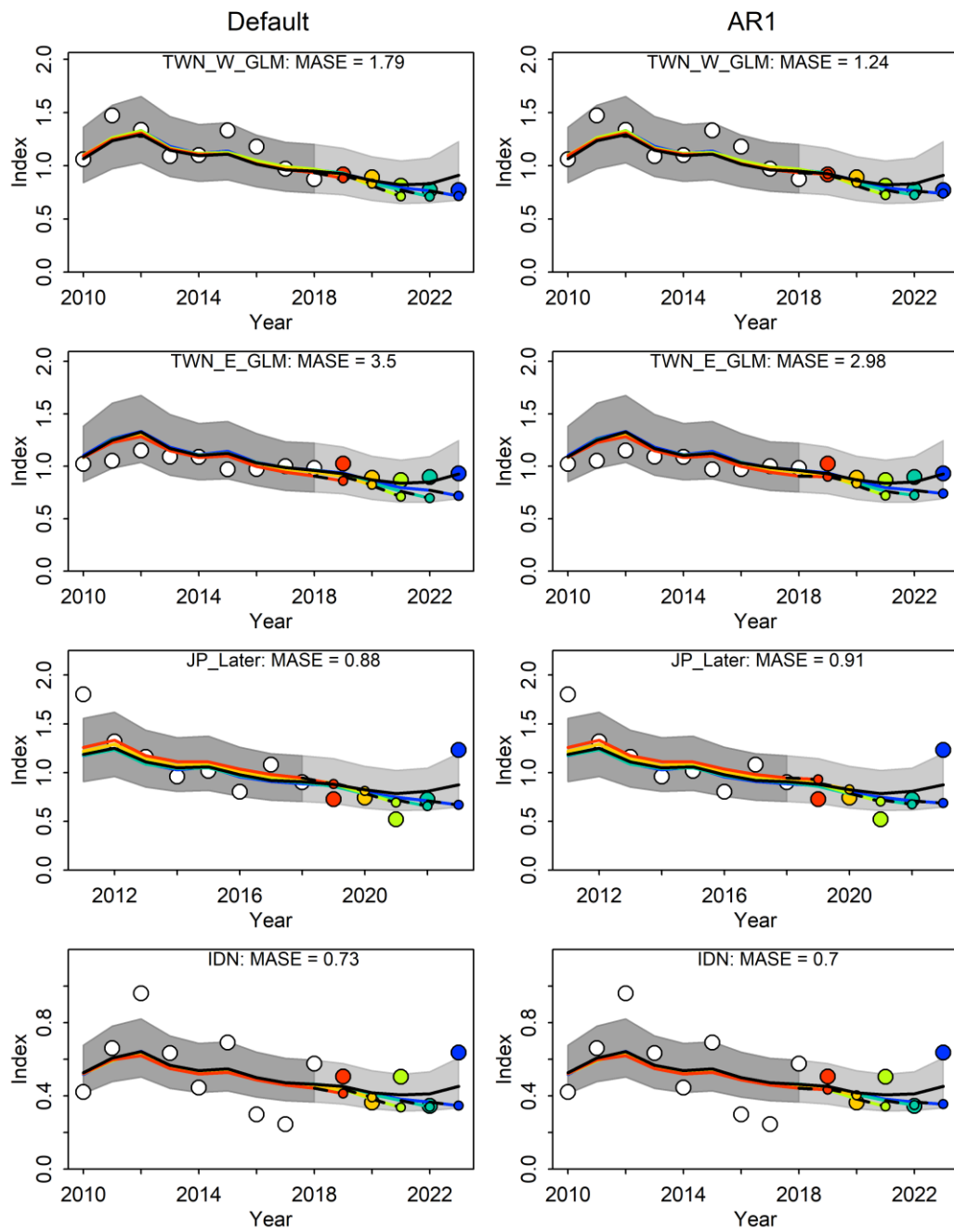


Fig. 9. (Continued).

S10

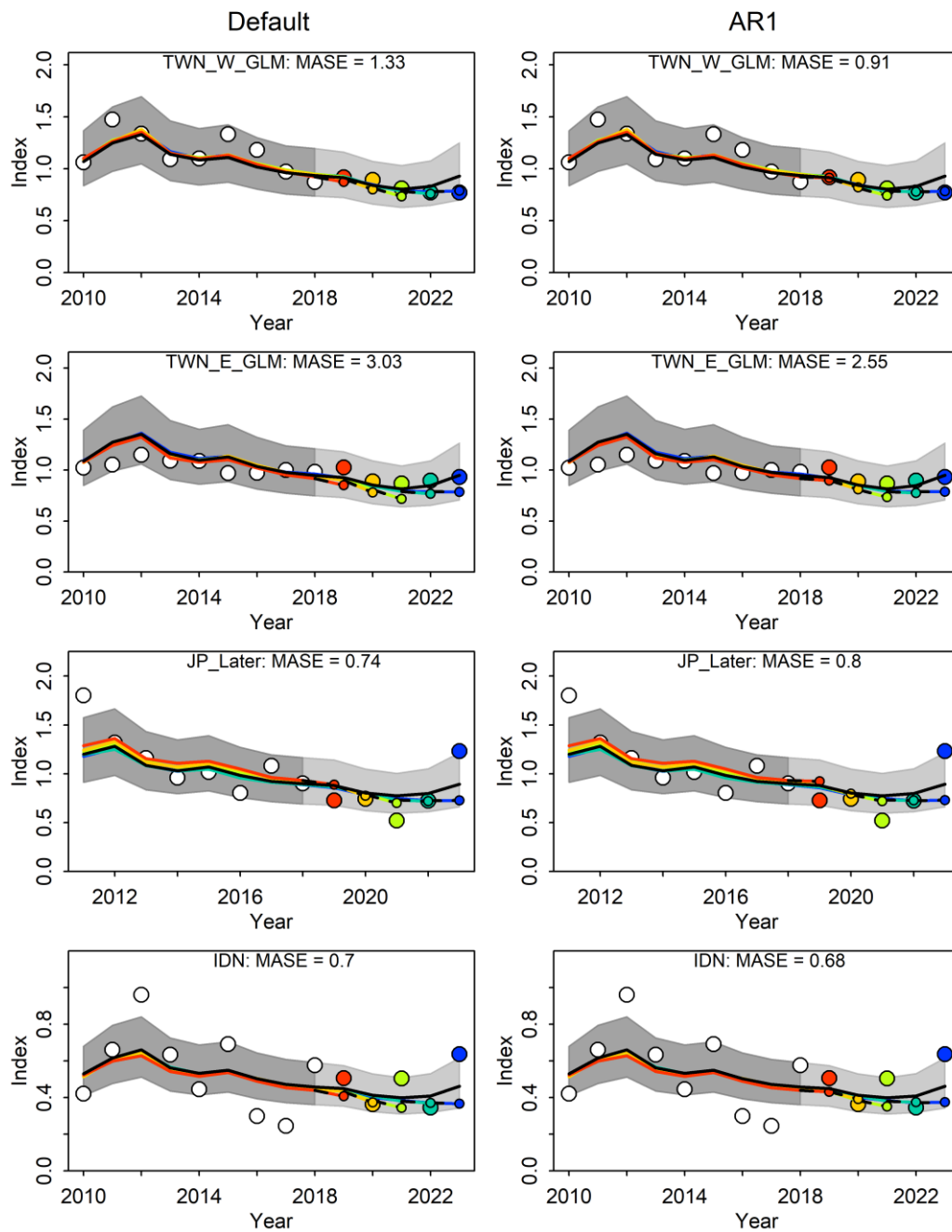


Fig. 9. (Continued).

S1

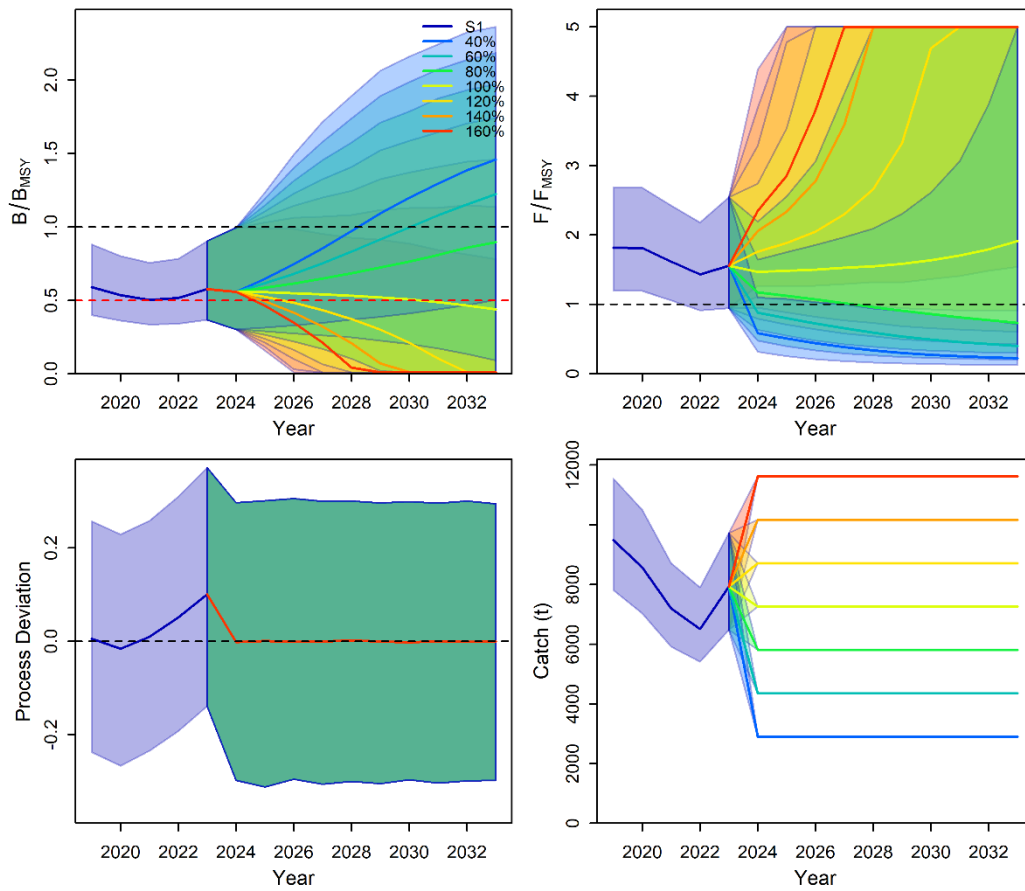


Fig. 10. Projections with 95% confidence intervals of JABBA based on the future catch set at constant levels from 40% to 160% for blue marlin in the Indian Ocean from scenarios S1, S6 and S10.

S6

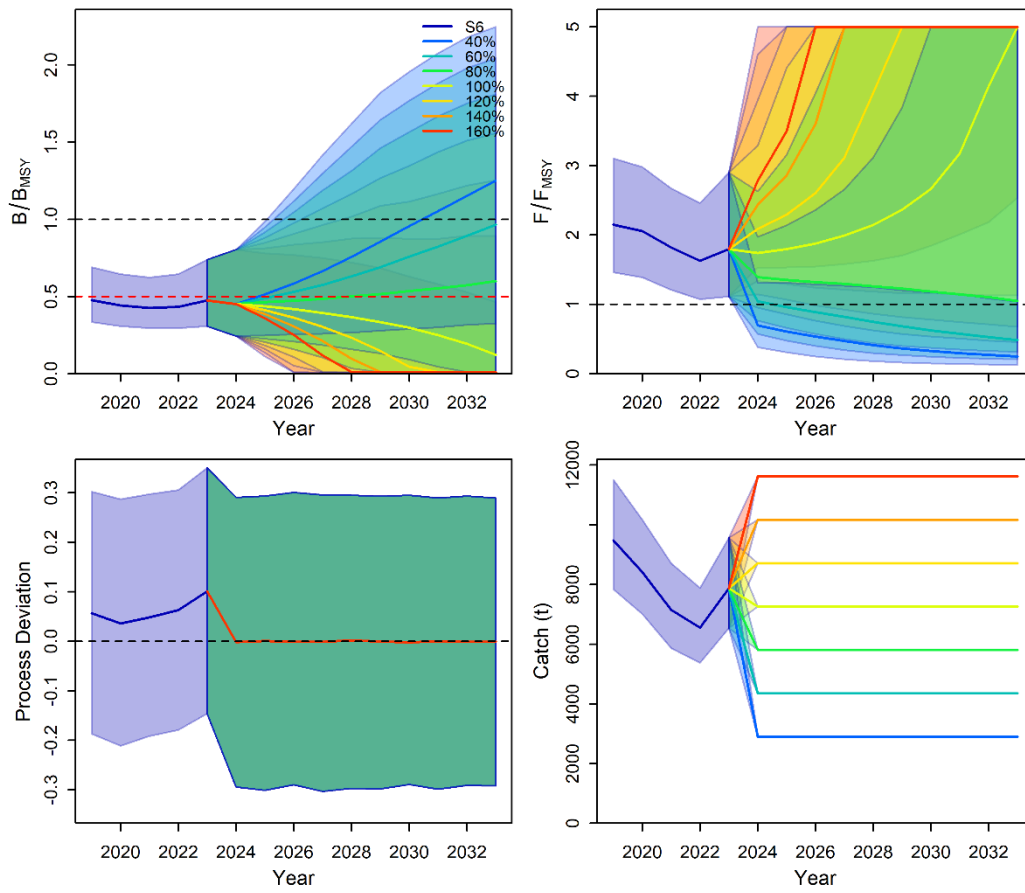


Fig. 10. (Continued).

S10

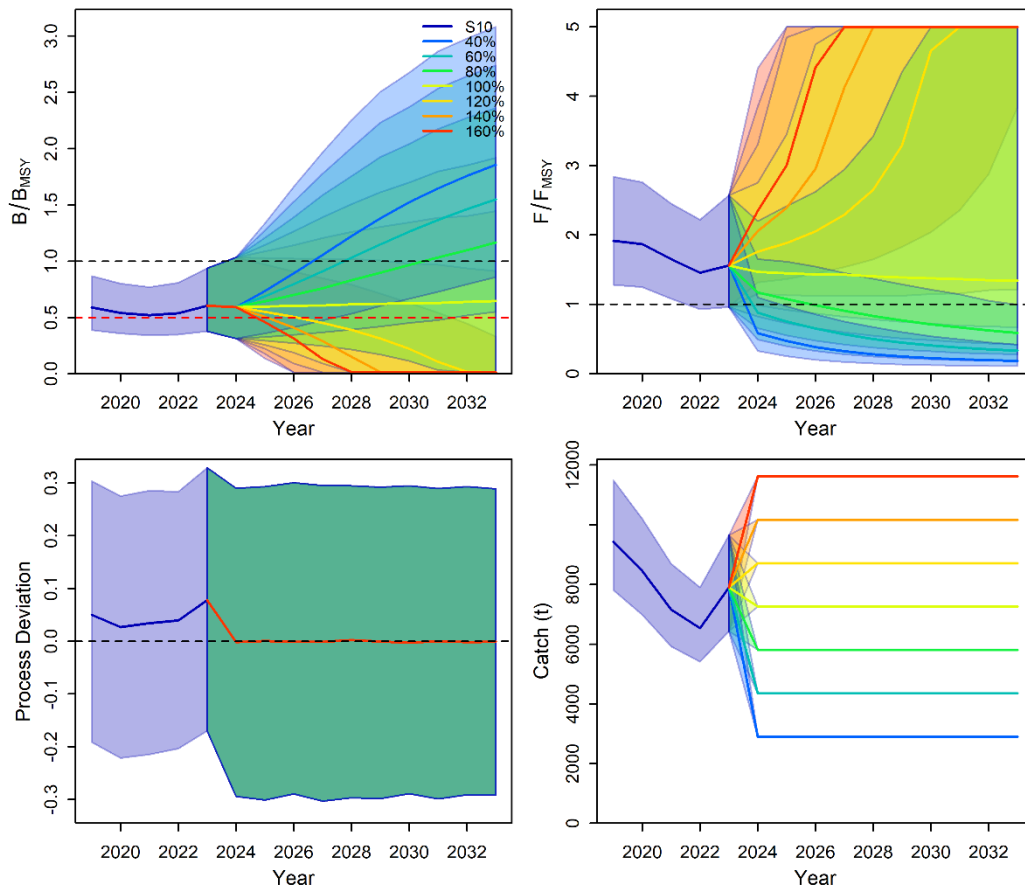
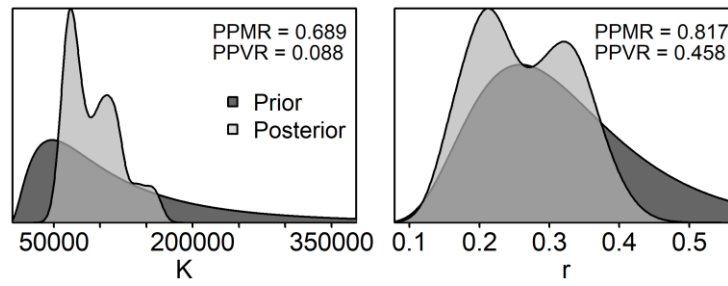


Fig. 10. (Continued).

S10



S10c

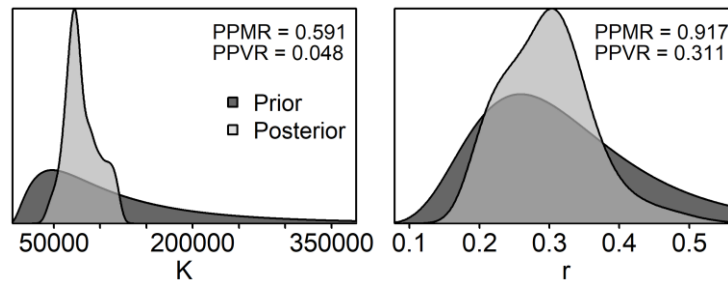


Fig. 11. Prior (dark) and posterior (light) distributions of K and r for scenarios S10 and S10c, with PPMR and PPVR showing stronger data signal and precision in S10c.

Table 1. Summary of standardized CPUE indices used in the JABBA assessment of blue marlin in the Indian Ocean.

Fleet and area	Period	Abbreviation
Taiwan, North-West Indian Ocean (GLM)	2005-2023	TW_W_GLM
Taiwan, North-East Indian Ocean (GLM)	2005-2023	TW_E_GLM
Taiwan, North-West Indian Ocean (sdmTMB)	2005-2023	TW_W_sdmTMB
Taiwan, North-East Indian Ocean (sdmTMB)	2005-2023	TW_N_sdmTMB
Japan	1979-2010	JP_early
Japan	2011-2023	JP_later
Indonesia	2006-2023	IDN

Table 2. Prior distributions for key parameters (K , r , ψ , and observation error variance) applied in the JABBA assessment of blue marlin in the Indian Ocean.

Parameter	Distribution	Specification
Carrying capacity (K)	Lognormal	$\mu = 106,557$; $CV = 300\%$
Intrinsic growth rate (r)	Lognormal	$\mu = 0.3$; $CV = 40\%$ (equivalent to $\log.sd = 0.385$)
Initial depletion ($\psi = B_1/K$)	Beta	$\mu = 0.95$; $CV = 5\%$
Observation error variance	Fixed	$\sigma^2 = 0.25$
Process error variance	Inverse-gamma	$\alpha = 0.001$; $\beta = 0.001$ (estimated in 2022)

Table 3. Summary of model scenarios (S1–S10) combining alternative CPUE indices, production model types, and treatments of process error standard deviation (σ_{proc}) in the JABBA assessment of blue marlin in the Indian Ocean.

Scenario	Alternative Taiwanese CPUE	Model type	σ_{proc}
1	TWN-GLM	Schaefer	Estimated ($igamma=c(0.001,0.001)$)
2	TWN-GLM	Schaefer	Fixed (0.15)
3	TWN-GLM	Schaefer	Fixed (0.20)
4	TWN-GLM	Schaefer	Fixed (0.10)
5	TWN-sdmTMB	Schaefer	Estimated ($igamma=c(0.001,0.001)$)
6	TWN-sdmTMB	Schaefer	Fixed (0.15)
7	TWN-GLM	Fox	Estimated ($igamma=c(0.001,0.001)$)
8	TWN-GLM	Fox	Fixed (0.15)
9	TWN-sdmTMB	Fox	Estimated ($igamma=c(0.001,0.001)$)
10	TWN-sdmTMB	Fox	Fixed (0.15)

Table 4. Estimates of key management quantities from JABBA all scenarios for blue marlin in the Indian Ocean.

Scenario	MSY	F_{2023}	F_{MSY}	B_0	B_{2023}	B_{MSY}	B/B_0	B_{2023}/B_{MSY}	F_{2023}/F_{MSY}
S1	8,849	0.283	0.178	98,666	27,978	49,333	0.288	0.577	1.557
S2	8,958	0.280	0.176	107,266	28,100	53,633	0.282	0.564	1.566
S3	10,629	0.293	0.177	126,149	27,001	63,075	0.226	0.452	1.662
S4	8,292	0.239	0.147	115,557	32,904	57,778	0.295	0.589	1.620
S5	8,836	0.307	0.184	92,038	25,712	46,019	0.257	0.515	1.749
S6	9,275	0.289	0.158	123,236	27,288	61,618	0.237	0.474	1.798
S7	8,161	0.420	0.315	68,514	18,858	25,218	0.258	0.702	1.390
S8	8,166	0.470	0.351	63,216	16,880	23,268	0.271	0.737	1.317
S9	8,095	0.525	0.377	58,315	15,044	21,464	0.256	0.696	1.408
S10	8,380	0.403	0.259	87,763	19,563	32,302	0.222	0.604	1.561

Table 5. The statistics related to CPUE fit (standzfdized RMSE, Std_RMSE; Bayesian.p) and prediction skill (MASE, MASE.adj and MAE.PR) for scenarios.

Scenario	B	F	Bmsy	Fmsy	procB	MSY
S1	-0.106	0.128	-0.079	0.123	-0.003	-0.017
S2	-0.083	0.091	-0.032	0.072	-0.001	-0.029
S3	-0.045	0.068	0.023	0.029	0.001	-0.045
S4	-0.087	0.111	-0.092	0.14	-0.005	-0.017
S5	-0.111	0.143	-0.097	0.151	-0.011	-0.029
S6	-0.089	0.103	-0.083	0.134	-0.011	-0.028
S7	-0.045	0.067	-0.025	0.056	-0.005	-0.015
S8	-0.106	0.152	-0.042	0.063	-0.001	-0.01
S9	-0.073	0.09	-0.084	0.115	-0.01	-0.016
S10	-0.054	0.072	-0.074	0.111	-0.006	-0.017
S10a	-0.129	0.176	-0.122	0.183	-0.01	-0.029
S10b	-0.089	0.106	-0.021	0.036	-0.002	-0.008
S10c	-0.111	0.143	-0.015	0.052	-0.008	-0.025
S10d	-0.086	0.103	-0.018	0.063	-0.007	-0.024
S10e	-0.115	0.136	-0.028	0.062	-0.007	-0.022

Table 6. Mohn's p from retrospective analysis showing the influence of adding new data.

Scenario	B	F	Bmsy	Fmsy	procB	MSY
S1	-0.106	0.128	-0.079	0.123	-0.003	-0.017
S2	-0.083	0.091	-0.032	0.072	-0.001	-0.029
S3	-0.045	0.068	0.023	0.029	0.001	-0.045
S4	-0.087	0.111	-0.092	0.14	-0.005	-0.017
S5	-0.111	0.143	-0.097	0.151	-0.011	-0.029
S6	-0.089	0.103	-0.083	0.134	-0.011	-0.028
S7	-0.045	0.067	-0.025	0.056	-0.005	-0.015
S8	-0.106	0.152	-0.042	0.063	-0.001	-0.01
S9	-0.073	0.09	-0.084	0.115	-0.01	-0.016
S10	-0.054	0.072	-0.074	0.111	-0.006	-0.017
S10a	-0.129	0.176	-0.122	0.183	-0.01	-0.029
S10b	-0.089	0.106	-0.021	0.036	-0.002	-0.008
S10c	-0.111	0.143	-0.015	0.052	-0.008	-0.025
S10d	-0.086	0.103	-0.018	0.063	-0.007	-0.024
S10e	-0.115	0.136	-0.028	0.062	-0.007	-0.022

Table 7. Additional sensitivity scenarios requested by the IOTC WPB, showing alternative CPUE combinations, model types, and process-error settings.

Scenario	CPUE used	Model type	Process error sigma
10	All	Fox (est. phi)	Fixed (0.15)
10a	Exclude TWN	Fox (est. phi)	Fixed (0.15)
10b	Exclude JPN_late & IDN	Fox (est. phi)	Fixed (0.15)
10c	Use estimated CV	Fox (est. phi)	Fixed (0.15)
10d	All	Fox (phi=0.99)	Fixed (0.15)
10e	All	Pella-m (est. m)	Fixed (0.15)

Table 8. Stock status summary table for the blue assessment (JABBA) obtained from the case “S10c” (CI = 80% confidence interval).

Catch (1,000 t) in 2023	7.905
Average catch (1,000 t) 2021–2023	7.262
MSY (1,000 t)	8.351 (7.516,9.232)
F_{MSY}	0.30 (0.21,0.38)
B_0 (1,000 t)	75.863 (60.557,108.310)
B_{2023} (1,000 t)	17.512 (13.062,24.148)
B_{MSY}	27.922 (22.289,39.865)
B_{2023} / B_0	0.23 (0.18,0.29)
B_{2023} / B_{MSY}	0.62 (0.48,0.78)
F_{2023} / F_{MSY}	1.54 (1.16,2.06)

Table 9. Blue marlin: JABBA Indian Ocean assessment Kobe II Strategy Matrix. Probability (percentage) of violating the MSY-based target reference points for nine constant catch projections relative to the average catch level from 2021 to 2023 (7,262 t)*, $\pm 10\%$, $\pm 20\%$, $\pm 30\%$ $\pm 40\%$ projected for 3 and 10 years.

Reference point and projection timeframe	Alternative catch projections (relative to the average catch level from 2021 to 2023 of 7,262 t) and probability (%) of violating MSY-based target reference points ($B_{targ} = B_{MSY}$; $F_{targ} = F_{MSY}$)								
	60%	70%	80%	90%	100%	110%	120%	130%	140%
	(4,357)	(5,083)	(5,809)	(6,536)	(7,262)	(7,988)	(8,714)	(9,440)	(10,167)
$B_{2028} < B_{MSY}$	65	69	74	78	82	85	88	90	92
$F_{2028} > F_{MSY}$	19	31	45	60	72	81	88	93	96
$B_{2035} < B_{MSY}$	14	23	36	50	64	77	87	93	97
$F_{2035} > F_{MSY}$	5	12	23	40	58	75	87	94	98

Table 10. Blue marlin: Probability (percentage) of achieving the KOBE green quadrat from 2024-2033 for a range of constant catch projections (JABBA). Catch in 2024 and 2025 are fixed at 7,262 t.

Catch (t)\Year	2026	2027	2028	2029	2030	2031	2032	2033	2034	2035
4,357 (60%)	10	22	35	48	59	67	74	80	83	86
5,083 (70%)	10	20	31	41	50	58	64	69	73	77
5,809 (80%)	10	18	26	34	41	47	53	57	61	64
6,536 (90%)	10	16	22	27	33	37	41	44	47	50
7,262 (100%)	10	14	18	22	25	27	30	32	34	35
7,988 (110%)	10	12	15	16	18	19	20	21	22	22
8,714 (120%)	8	9	11	11	12	12	12	13	12	12
9,440 (130%)	6	6	6	7	6	6	6	6	6	6
10,167 (140%)	4	4	4	4	3	3	3	3	3	2