Food and Agriculture Organization of the United Nations

Indian Ocean Tuna Commission
Commission des Thons de l'Ocean Indien
iotc ctoi

[IOTC-2026-WPB24-26]

STOCK ASSESSMENT FOR SWORDFISH (*XIPHAS GLADIUS*) IN THE INDIAN OCEAN USING A STOCHASTIC SURPLUS PRODUCTION MODEL IN CONTINUOUS TIME (SPICT)

Sheng-Ping Wang*[1], Chih-Yu Lin, Wen-Qi Xu

Discussions: The paper is to inform meeting members of the first application of SPiCT to the stock assessment of Indian Ocean swordfish, including production-model evaluation, CPUE sensitivity analyses, stock status, and projections, as set out in Sections 3.1–3.4.

Feedback: The paper is to request review and comments on the SPiCT model configuration, treatment of CPUE indices, selection of Scenario j, and projection assumptions, as set out in Sections 2.2–3.4.

Conclusion: The paper is to seek confirmation that the SPiCT results provide useful complementary assessment information and support the conclusion that the stock was neither overfished nor subject to overfishing in 2024, as set out in Section 3.3.

Recommendation: The paper is to seek endorsement for continued evaluation of SPiCT as a complementary framework for future Indian Ocean swordfish assessments, as set out in Sections 3.3–3.4.

Decisions: No formal decision is requested from the Working Party; the paper is submitted for discussion, feedback, and consideration.

Department of Environmental Biology and Fisheries Science, National Taiwan Ocean University, Keelung, Taiwan.
*Corresponding author: wsp@mail.ntou.edu.tw

Contents

Abstract 2
1. Introduction 2
2. Materials and methods 3
2.1 Data inputs 3
2.2 SPiCT model specification and implementation 3
2.2.1 Model framework 3
2.2.2 Priors and model configuration 3
2.2.3 Model outputs and diagnostics 4
2.2.4 Kobe plot with joint uncertainty 4
2.2.5 Short-term projections and Kobe II strategy matrix 4
2.3 Prior and fitted parameter comparisons 5
2.4 Production model evaluation and base model selection 5
3. Results and discussion 5
3.1 Production model evaluation and base model selection 5
3.2 CPUE index sensitivity analysis 6
3.3 Stock status of the selected model 7
3.4 Projection analysis 7
Acknowledgements 8
References 8
Figures 9
Tables 14
Appendix 21

Abstract
This working paper presents the first application of the stochastic surplus production model in continuous time (SPiCT) to the stock assessment of Indian Ocean swordfish. Annual catch in weight from 1950 to 2024 and 13 standardized CPUE indices were analyzed. Alternative Fox, Schaefer, and generalized Pella Tomlinson production formulations were evaluated. The Pella Tomlinson shape parameter was poorly identified, particularly when its prior was removed, and the Fox formulation was selected as the base production model. Residual diagnostics identified persistent autocorrelation in CPUE indices I2, I6, I10, and I11. A final Fox model excluding these four indices (Scenario j) removed the major residual-autocorrelation signals while retaining stable retrospective behavior. Terminal biomass was estimated at 1.850 times the MSY-based biomass reference point (80% CI: 1.667 to 2.053), and fishing mortality at 0.333 times the MSY-based fishing mortality reference point (80% CI: 0.247 to 0.449), indicating that the stock was neither overfished nor subject to overfishing in 2024. Constant-catch projections from 60% to 140% of the recent catch baseline indicated low risk through 2036, although risk increased with catch level and projection horizon. The results support the use of SPiCT as a complementary assessment framework for Indian Ocean swordfish.

1. Introduction
Swordfish (*Xiphias gladius*) is a highly migratory species widely distributed throughout the Indian Ocean and supports important commercial fisheries. Swordfish has been taken by industrial longline fisheries since the 1950s, initially largely as bycatch, while targeted fisheries expanded substantially from the early 1990s. Total catches increased markedly during the 1990s and reached their highest levels around the late 1990s and early 2000s. More recently, annual catches averaged approximately 26,800 t during 2020–2024, with a reported catch of approximately 28,100 t in 2024 (IOTC, 2025). Longline fisheries continue to account for a substantial proportion of the total catch, although swordfish is also harvested by several other line and gillnet fisheries.

The most recent benchmark stock assessment for Indian Ocean swordfish was conducted in 2023 using data through 2021. Several assessment approaches were considered, including Stock Synthesis 3 (SS3) and surplus production models implemented with ASPIC and JABBA. The IOTC Scientific Committee adopted an ensemble of SS3 model configurations as the primary basis for management advice. The assessment indicated that the stock was neither overfished nor subject to overfishing, with the SS3 reference ensemble estimating biomass above the MSY-based reference level and fishing mortality below F_{MSY}. The ASPIC and JABBA assessments produced broadly consistent conclusions regarding stock status (Fu, 2023; Wang, 2023; Xu et al., 2023; IOTC, 2023, 2025).

Surplus production models remain useful complements to more data-intensive integrated assessments because they provide relatively simple and transparent estimates of stock productivity, biomass, fishing mortality, and MSY-based reference points. More recent state-space implementations additionally allow process variability and observation error to be represented separately. The stochastic surplus production model in continuous time (SPiCT; Pedersen and Berg, 2017) is one such approach. SPiCT formulates biomass and fishing mortality dynamics within a continuous-time state-space framework and provides model diagnostics, including one-step-ahead residual analyses and retrospective evaluation. These characteristics provide an alternative framework for examining the robustness of stock-status estimates derived from biomass dynamic models.

In this working paper, we provide the first application of SPiCT to the stock assessment of Indian Ocean swordfish. The analysis uses the updated historical catch series together with 13 standardized CPUE indices and evaluates alternative surplus production formulations and the sensitivity of model results to the configuration of abundance indices. The objective is not to replace the existing benchmark assessment, but to provide an additional and independent biomass-dynamic perspective on the current status of the stock and to evaluate the potential utility of SPiCT for future Indian Ocean swordfish assessments and discussions within the IOTC Working Party on Billfish.

2. Materials and methods

2.1 Data inputs

The original raised catch data were stratified by year, month, fishing ground, fleet, and fishing gear, and included catch in both weight and numbers. Because the surplus production model requires total removals in weight, catch records were summed across months, fishing grounds, fleets, and gears within each year to obtain a single annual total catch series covering 1950 to 2024 (Fig. 1).

Thirteen standardized catch-per-unit-effort (CPUE) indices representing different fleet-area combinations were used as relative abundance indices in the assessment. The temporal coverage differed among indices, with individual series beginning between 1994 and 2006 and ending between 2023 and 2025. The fleet labels and temporal coverage of the 13 CPUE indices are summarized in Table 1, and their trajectories are shown in Fig. 1. Catch and CPUE data were compiled into the standardized annual input format required for SPiCT prior to model fitting. These CPUE indices were standardized in the respective fleet-specific analyses [references to be added].

2.2 SPiCT model specification and implementation

2.2.1 Model framework

Stock status was assessed using the stochastic surplus production model in continuous time (SPiCT; Pedersen and Berg, 2017), implemented in R with the spict package (Mildenberger et al., 2026). SPiCT is a state-space biomass dynamics model formulated in continuous time and fitted within the Template Model Builder (TMB) framework by maximum likelihood, with the Laplace approximation used to integrate over the unobserved biomass and fishing mortality states. The production function follows the generalized Pella Tomlinson formulation:

$$\frac{dB(t)}{dt} = \frac{r}{n-1} B(t) \left[1 - \left(\frac{B(t)}{K} \right)^{n-1} \right] - C(t),$$

where $B(t)$ is biomass, K is carrying capacity, r is the intrinsic growth rate, n is the shape parameter of the production curve, and $C(t)$ is catch removed continuously from the population. This formulation reduces to the Schaefer model when $n = 2$ and approaches the Fox formulation as $n \to 1$. The shape parameter n was either fixed or estimated, with the treatment of prior information varying among scenarios as described in Section 2.4.

2.2.2 Priors and model configuration

Log-normal priors were specified for K, r, and the initial depletion ratio B_1/K (logbkfrac in SPiCT notation). The common priors used across the production-model scenarios were

$$\log(K) \sim N[\log(300000), 1^2],$$

$$\log(r) \sim N[\log(0.4), 0.5^2],$$

and

$$\log(B_1/K) \sim N[\log(0.99), 0.1^2].$$

The prior on B_1/K expressed the assumption that the stock was close to the unfished biomass level at the beginning of the 1950 catch series. Because the earliest CPUE indices begin several decades later, the early biomass trajectory is only weakly informed by abundance indices and is therefore particularly sensitive to the assumed initial depletion. The Euler integration time step was fixed at $1/16$ year for all model scenarios. Unless explicitly changed in a defined sensitivity analysis, all other numerical settings were kept identical among scenarios.

2.2.3 Model outputs and diagnostics

Convergence was assessed from the optimizer convergence status together with the Hessian and gradient of the optimized objective function. Parameter estimates, management reference points (B_{MSY}, F_{MSY}, and MSY), and annual absolute and relative biomass and fishing mortality trajectories were obtained with associated confidence intervals using standard spict utility functions. An 80% confidence interval was used throughout for consistency with IOTC reporting conventions.

Model fit was evaluated using one-step-ahead residual diagnostics for catch and each CPUE index. Diagnostics included tests of residual bias, serial correlation, and normality, together with graphical inspection of residual series, autocorrelation functions, and quantile plots. Ljung-Box P values for the CPUE indices were also evaluated using the Benjamini-Hochberg false discovery rate (FDR) adjustment because multiple index-specific residual tests were examined simultaneously. The FDR adjustment was used as supporting diagnostic evidence rather than as an automatic index-exclusion rule.

Retrospective analysis was performed using SPiCT's built-in retrospective routine by sequentially removing the most recent one to five years of data and refitting the model to each truncated dataset. Mohn's ρ was calculated for B/B_{MSY} and F/F_{MSY} using the package's standard implementation.

2.2.4 Kobe plot with joint uncertainty

Terminal stock status was summarized in Kobe space using B/B_{MSY} and F/F_{MSY}. To display multiple confidence levels and quadrant probabilities, a custom implementation was used. The implementation first attempts to retrieve the joint covariance matrix of $\log(B/B_{\mathrm{MSY}})$ and $\log(F/F_{\mathrm{MSY}})$ in the terminal assessment year from the fitted TMB object. The resulting marginal standard errors are cross-checked against the delta-method confidence intervals returned by SPiCT, and the joint covariance is used only when the marginal uncertainty is consistent. If the joint covariance cannot be recovered reliably, the joint distribution is approximated as bivariate log-normal using the fitted marginal standard errors and an assumed correlation coefficient of -0.6. The 50%, 80%, and 95% confidence ellipses and the probability of terminal stock status falling within each Kobe quadrant are then calculated from the resulting joint distribution. Results obtained using the assumed-correlation fallback are interpreted with this limitation in mind.

2.2.5 Short-term projections and Kobe II strategy matrix

For the final selected assessment scenario, constant-catch projections and the associated Kobe II strategy matrix were designed using a custom Monte Carlo routine based on the fitted production function. The reference catch level is the mean catch over the most recent three years. Catch is held at this recent mean for the first two projection years and is then set at fixed levels ranging from 60% to 140% of the reference catch in 10% increments. Projection results are summarized at three- and ten-year horizons and annually through the ten-year projection period.

Biomass is propagated forward using Euler integration with the same time step used in model fitting. Uncertainty in r and K is represented by draws from their fitted log-normal marginal distributions with an assumed correlation of -0.6, and annual process error is incorporated using the estimated process-error standard deviation. Each simulation is anchored to the terminal B/B_{MSY} estimate to avoid introducing excessive uncertainty by sampling absolute biomass and carrying capacity independently. For each projection year and catch scenario, the probabilities of $B < B_{\mathrm{MSY}}$, $F > F_{\mathrm{MSY}}$, and the stock remaining in the Kobe green quadrant are calculated from the Monte Carlo simulations. This custom projection routine approximates the uncertainty propagation of the fitted state-space model and is therefore interpreted as a management scenario analysis rather than an exact replication of SPiCT's internal forecasting machinery.

2.3 Prior and fitted parameter comparisons

Prior and fitted parameter distributions were compared graphically for the key production parameters K and r and the initial depletion ratio B_1/K. The specified priors were compared with marginal uncertainty reconstructed from the fitted parameter estimates and delta-method confidence intervals. Equivalent diagnostic plots were generated for fleet-specific catchability and observation-error parameters. These comparisons were used to evaluate how strongly the available data informed the specified model parameters. For the generalized Pella Tomlinson formulation, prior dependence and identifiability of the shape parameter n were evaluated directly from its estimate and uncertainty and from the no-prior sensitivity analysis described in Section 2.4.

2.4 Production model evaluation and base model selection

In the first stage of the analysis, alternative surplus production formulations were evaluated using the same catch and CPUE data and identical model settings, except for the specification of the production shape parameter n (Table 2). Scenario a represented the Fox formulation, with n fixed at 0.99 to approximate the Fox limit while avoiding numerical issues associated with setting n exactly equal to 1. Scenario b represented the Schaefer formulation, with n fixed at 2. Scenario c used the generalized Pella Tomlinson formulation, in which n was estimated using the broad SPiCT prior

$$\log(n) \sim N[\log(2), 2^2].$$

Scenario d used the same generalized formulation but removed the prior on $\log(n)$. Scenario d was included specifically as a prior-sensitivity and identifiability test and was not treated as a primary candidate base model. All four scenarios used the full catch time series and all 13 standardized CPUE indices.

The preferred production formulation for subsequent analyses was selected by comparing convergence, residual diagnostics, parameter identifiability and uncertainty, retrospective patterns, and the stability and precision of estimated management reference points and stock status among Scenarios a through d. After the production formulation was selected, it and all other model settings were held constant while sensitivity to alternative CPUE index configurations was evaluated. Candidate CPUE indices were identified from persistent residual diagnostic problems and potential influence on model estimation. Individual exclusion analyses were conducted first, followed by combined exclusion scenarios where appropriate. The specific CPUE sensitivity scenarios were therefore defined from the Stage 1 diagnostic results and are presented together with those results in Section 3.2.

3. Results and discussion

3.1 Production model evaluation and base model selection

All four Stage 1 model configurations achieved numerical convergence (Table 3). The three primary production formulations produced the same qualitative conclusion regarding terminal stock status, with biomass above B_{MSY} and fishing mortality below F_{MSY}. The Fox and Schaefer models produced relatively precise terminal stock-status estimates. Terminal B/B_{MSY} and F/F_{MSY} were estimated at 1.678 and 0.414 for the Fox model and 1.286 and 0.609 for the Schaefer model, respectively (Table 3).

Residual diagnostics were generally similar among the alternative production formulations. In particular, I2, I6, I10, and I11 showed persistent Ljung-Box residual autocorrelation across Scenarios a through d and remained significant after Benjamini-Hochberg FDR adjustment (Appendix Fig. A1). This consistency indicated that the serial-correlation patterns were associated primarily with these CPUE indices rather than with the choice of production function. I8 did not show the same persistent residual autocorrelation, but its estimated observation-error standard deviation was much smaller than those of the other indices, indicating that its influence on model estimation warranted separate evaluation in the subsequent CPUE sensitivity analysis.

Estimation of the generalized Pella Tomlinson production shape revealed substantial uncertainty in n (Table 3; Fig. 2). Under the broad prior, the estimated shape parameter was $n = 0.600$, with an 80% confidence interval of 0.137 to 2.622. This interval encompassed both the Fox approximation and the Schaefer value and therefore did not allow the production shape to be reliably distinguished. When the prior on n was removed, the estimate decreased to 0.143 and its 80% confidence interval expanded to 0.00056 to 36.79. Uncertainty in the associated management reference points and relative stock-status estimates increased correspondingly. Thus, although the no-prior model achieved numerical convergence, the production shape parameter was effectively unidentifiable without regularization.

Retrospective patterns were small for the Fox, Schaefer, and prior-constrained Pella Tomlinson models (Table 3; Appendix Fig. A2). Mohn's ρ for B/B_{MSY} was -0.027, -0.022, and -0.026, respectively, and the corresponding values for F/F_{MSY} were 0.051, 0.033, and 0.055. In contrast, removal of the prior on n resulted in larger retrospective patterns, with Mohn's $\rho = -0.131$ for B/B_{MSY} and 0.248 for F/F_{MSY}. These results provided further evidence that the additional flexibility in the generalized production function was not adequately supported by the available data.

The Fox formulation was therefore selected as the preferred base production model. Both the Fox and Schaefer formulations provided stable estimates and similar residual performance, and the slightly smaller retrospective statistics for the Schaefer model were not considered sufficient to outweigh the broader model-selection evidence. Estimating n in the generalized Pella Tomlinson formulation substantially increased uncertainty without improving residual diagnostics, while its point estimates tended toward lower values of n that were directionally more consistent with the Fox formulation than with the Schaefer formulation. The Fox model therefore provided a parsimonious fixed-shape formulation that avoided estimating a poorly identified shape parameter while retaining stable and relatively precise management quantities.

3.2 CPUE index sensitivity analysis

Following selection of the Fox production formulation, sensitivity to CPUE index configuration was evaluated using individual and combined index-exclusion scenarios (Table 4). All CPUE sensitivity scenarios converged successfully and produced the same qualitative terminal stock-status conclusion, with $B/B_{\text{MSY}} > 1$ and $F/F_{\text{MSY}} < 1$ (Table 5).

Individual exclusion of I10, I11, I2, or I6 resulted in relatively small changes in estimated production parameters, management reference points, and terminal stock status compared with the all-index Fox model (Table 5). However, removing any one of these indices did not resolve the persistent residual autocorrelation in the other problematic indices. After FDR adjustment, Scenario e retained significant residual autocorrelation in I2, I6, and I11; Scenario f in I2, I6, and I10; Scenario g in I6, I10, and I11; and Scenario h in I2, I10, and I11. When I2, I6, I10, and I11 were excluded simultaneously in Scenario j, none of the remaining CPUE indices showed significant Ljung-Box residual autocorrelation after FDR adjustment. Retrospective patterns also remained small, with Mohn's $\rho = -0.026$ for B/B_{MSY} and 0.063 for F/F_{MSY} (Table 5).

Exclusion of I8 produced substantially larger changes in the estimated production parameters. Relative to the all-index Fox model, removing I8 increased r from 0.280 to 0.425 and decreased K from approximately 378,000 t to 232,000 t (Table 5). In Scenario i, five indices, I2, I5, I6, I10, and I11, remained significant after FDR adjustment, so removal of I8 did not provide an overall improvement in residual behavior. A similar shift in production parameters occurred in Scenario k when I8 was removed together with I2, I6, I10, and I11. Scenario k removed the major residual-autocorrelation signals, but it also excluded an index that did not itself show persistent residual autocorrelation and produced a substantial shift in r and K. Scenarios i and k were therefore interpreted as influence sensitivities rather than preferred assessment configurations.

Scenario j was selected as the preferred final assessment configuration because it retained the Fox production formulation selected in Stage 1, excluded only the four CPUE indices with persistent residual autocorrelation across production-model formulations, removed the major residual-autocorrelation signals among the retained indices after FDR adjustment, and maintained stable retrospective behavior and relatively precise management quantities. The consistency of the general stock-status conclusion across all alternative CPUE configurations also indicated that the assessment conclusion was robust to CPUE index selection (Table 5).

3.3 Stock status of the selected model

Under the selected Scenario j, the intrinsic growth rate was estimated at $r = 0.286$ (80% CI: 0.236 to 0.347) and carrying capacity at $K = 417{,}820$ t (80% CI: 336,888 to 518,193 t). The estimated management reference points were $B_{\mathrm{MSY}} = 152{,}935$ t (80% CI: 123,312 to 189,676 t), $F_{\mathrm{MSY}} = 0.289\ \mathrm{yr}^{-1}$ (80% CI: 0.239 to 0.351 yr^{-1}), and MSY = 44,234 t (80% CI: 37,579 to 52,067 t).

The fitted biomass trajectory declined markedly from the early 1990s but remained above B_{MSY} throughout the assessment period, while estimated fishing mortality increased rapidly during the 1990s but remained below F_{MSY} (Fig. 3). In the terminal assessment year 2024, biomass was estimated at 282,966 t (80% CI: 212,328 to 377,104 t) and fishing mortality at $0.096\ \mathrm{yr}^{-1}$ (80% CI: 0.070 to $0.132\ \mathrm{yr}^{-1}$). These values corresponded to $B/B_{\mathrm{MSY}} = 1.850$ (80% CI: 1.667 to 2.053) and $F/F_{\mathrm{MSY}} = 0.333$ (80% CI: 0.247 to 0.449). The observed catch in 2024 was 28,113 t, equivalent to approximately 63.6% of the estimated MSY.

The terminal stock-status estimate was therefore located clearly within the Kobe green quadrant, indicating that the stock was neither overfished nor subject to overfishing in 2024 (Fig. 4). This conclusion was consistent with the 2023 benchmark and surplus production assessments, which reached the same qualitative stock-status conclusion (Fu, 2023; Wang, 2023; Xu et al., 2023; IOTC, 2023, 2025). Fits to the retained CPUE indices, final one-step-ahead residual diagnostics, and the retrospective analysis for Scenario j are provided in Appendix Figs. A3-A5.

3.4 Projection analysis

Short-term and medium-term projections were conducted for the selected Scenario j using constant-catch levels ranging from 60% to 140% of the recent catch baseline. The baseline catch was 26,761 t, calculated as the mean observed catch during 2022–2024. Catch was held at this baseline level during 2025 and 2026, after which the alternative catch levels were applied from 2027 onward. Projection performance was summarized for 2029 and 2036 in terms of the probabilities of biomass falling below B_{MSY} and fishing mortality exceeding F_{MSY} (Table 6). Annual probabilities of the stock remaining in the Kobe green quadrant ($B/B_{MSY} \geq 1$ and $F/F_{MSY} < 1$) from 2027 to 2036 are summarized in Table 7. Selected projected catch and relative biomass and fishing mortality trajectories are shown in Fig. 5.

Projected risks remained low across the range of catch levels considered, although risk increased with both catch level and projection horizon. At catch levels of 60–80% of the baseline, the probabilities of biomass falling below B_{MSY} and fishing mortality exceeding F_{MSY} were 0% at both the 2029 and 2036 horizons. Under the baseline catch scenario (100%; 26,761 t), the probability of biomass falling below B_{MSY} was 0% in 2029 and 0.1% in 2036, while the probability of fishing mortality exceeding F_{MSY} was 0.1% and 0.3%, respectively. At the highest catch level examined (140%; 37,466 t), these probabilities increased to 0.1% and 1.1% in 2029, and to 3.3% and 6.0% in 2036 for biomass and fishing mortality, respectively (Table 6).

The annual Kobe-quadrant probabilities showed the same pattern. The probability of remaining in the green quadrant in 2036 was 100% under the 60–80% catch scenarios, 99.7% under the baseline catch scenario, and declined progressively to 94.0% under the 140% catch scenario (Table 7). Consistent with these probabilities, lower catch levels produced increasing trajectories of B/B_{MSY} and decreasing trajectories of F/F_{MSY}, whereas higher catch levels produced the opposite pattern (Fig. 5).

Overall, the projections indicated a low probability of the stock moving outside the Kobe green quadrant over the evaluated period under all catch scenarios considered, but the risk increased as catch levels and the projection horizon increased. These results should be interpreted as comparative management-scenario projections conditional on the selected SPiCT model and the projection assumptions, rather than as direct catch-limit recommendations.

Acknowledgements

This study was financially supported by the Fisheries Agency, Ministry of Agriculture, Taiwan (Grant No. 115AS-6.4.2-FA-02(1) to Sheng-Ping Wang). The data used in this study were provided by the Indian Ocean Tuna Commission.

References

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

IOTC, 2023. Report of the 21st Session of the IOTC Working Party on Billfish, La Saline-les-Bains, Reunion, 6–9 September 2023. IOTC–2023–WPB21–R[E].

IOTC, 2025. Report of the 28th Session of the IOTC Scientific Committee. China, 1 – 5 December 2025. IOTC–2025–SC28–R[E].

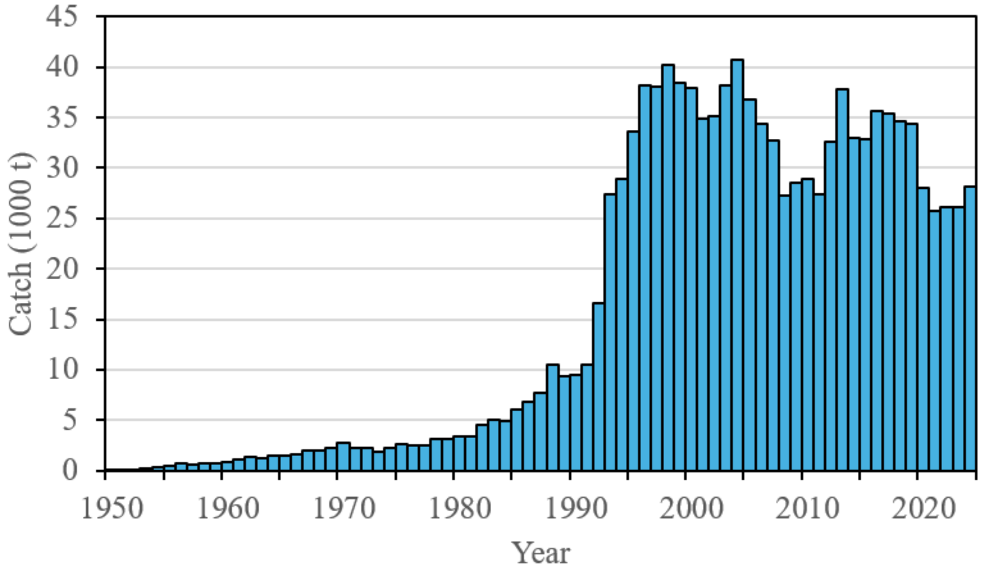
Mildenberger, T.K., Kokkalis, A., Berg, C.W., 2026. Guidelines for the stochastic production model in continuous time (SPiCT). DTU Aqua, Technical University of Denmark. https://github.com/dtuaqua/spict (accessed 24 August 2026).

Pedersen, M.W., Berg, C.W., 2017. A stochastic surplus production model in continuous time. Fish Fish. 18, 226–243. https://doi.org/10.1111/faf.12174.

Wang, S.-P., 2023. Updated stock assessment of swordfish (*Xiphias gladius*) in the Indian Ocean using the Bayesian state-space surplus production model (JABBA). IOTC–2023–WPB21–30.

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

Figures



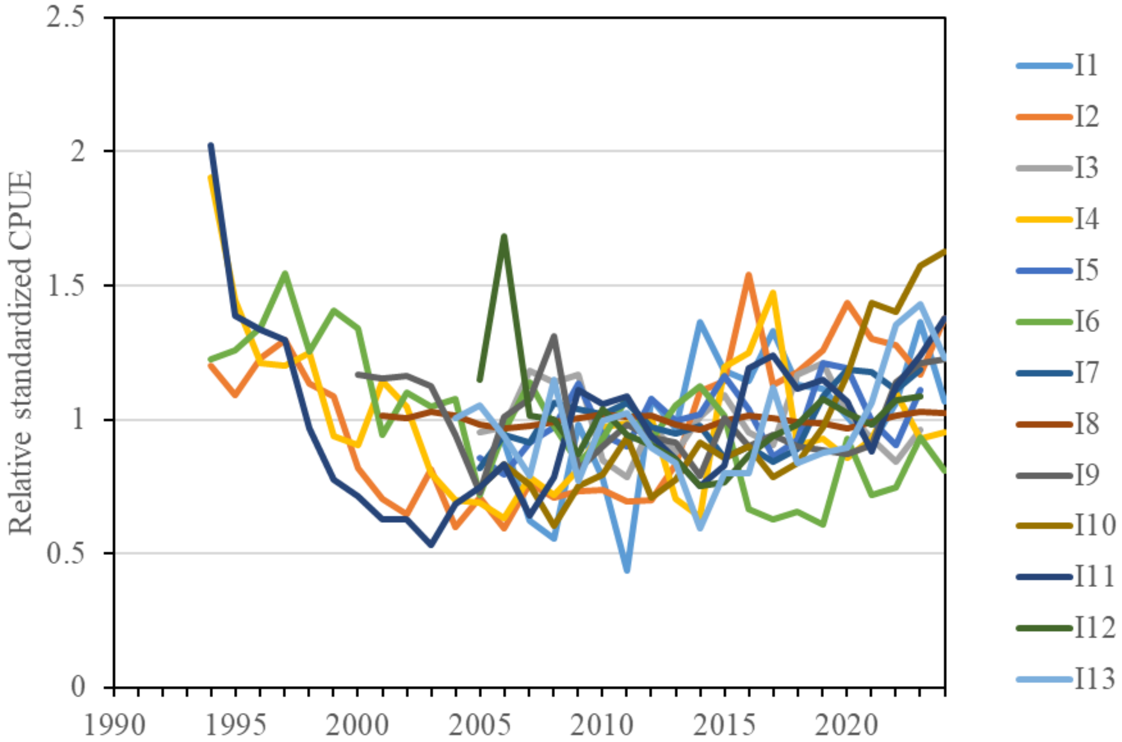


Fig. 1. Annual total catch in weight (upper panel) and standardized CPUE indices (lower panel) used in the SPiCT assessment.

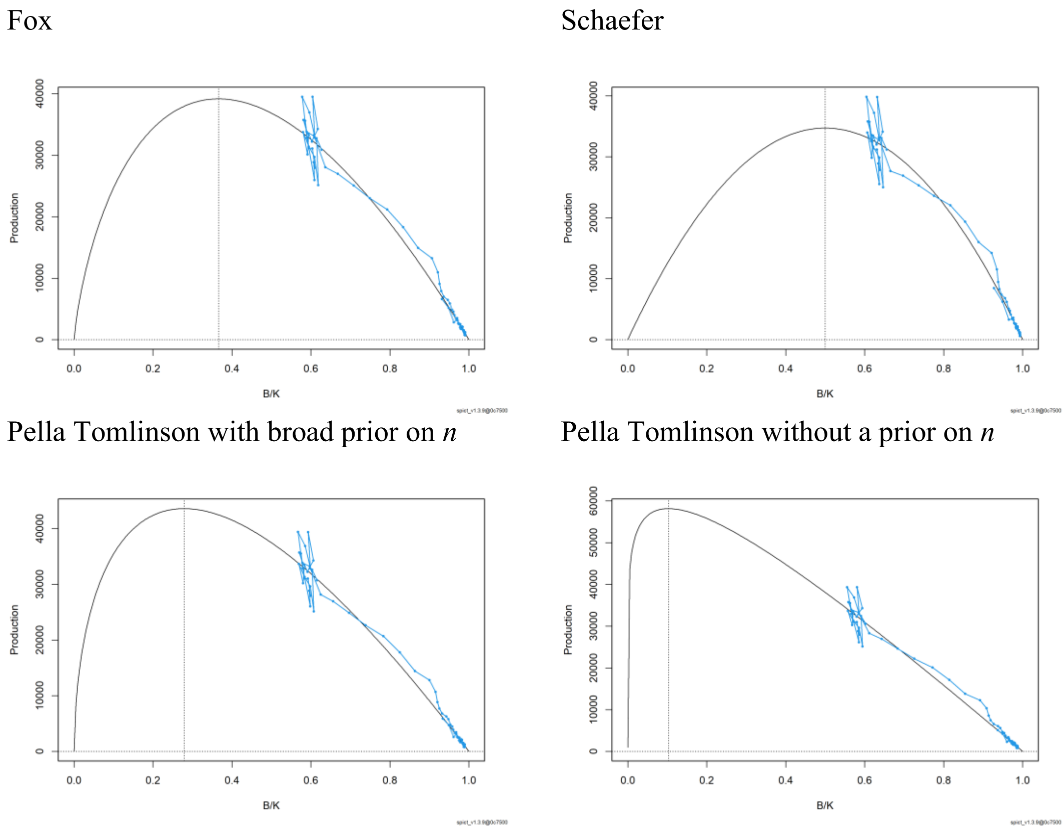
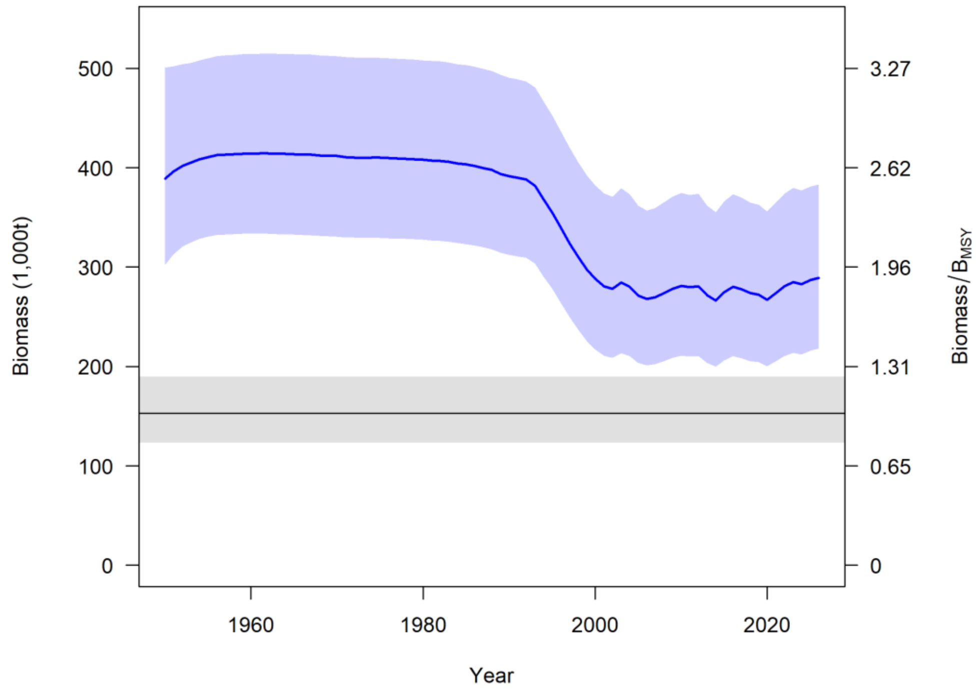


Fig. 2. Estimated surplus production curves and historical trajectories for the Stage 1 production model scenarios. Panels (a)–(d) correspond to the Fox, Schaefer, Pella Tomlinson with the broad prior on *n*, and Pella Tomlinson without a prior on *n*, respectively. The vertical dotted line indicates B_{MSY}/K.



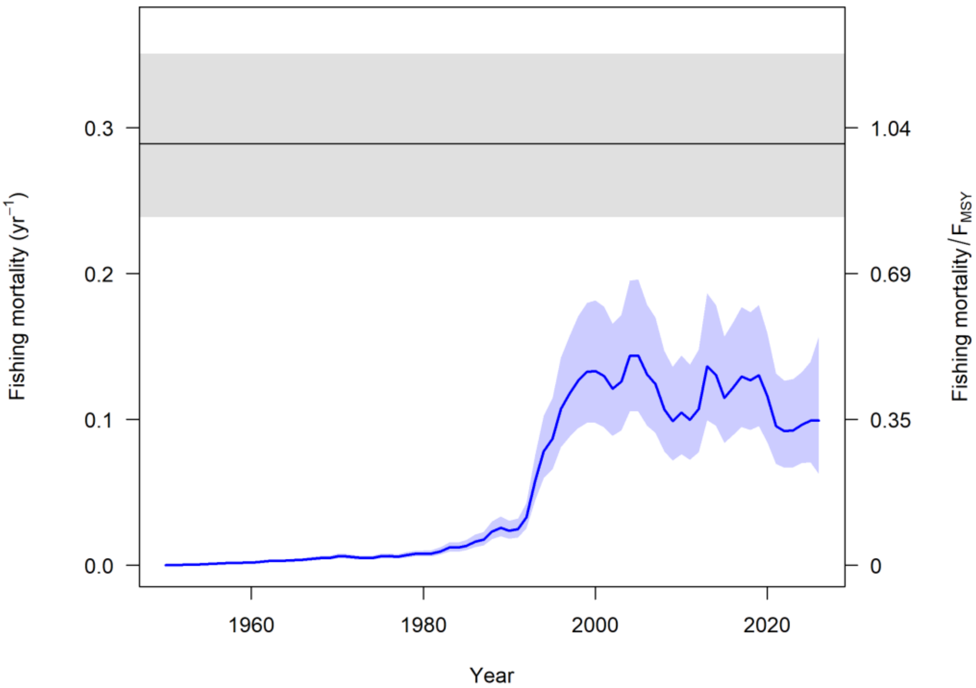


Fig. 3. Estimated biomass and fishing mortality trajectories from the selected Scenario j. Solid lines show fitted estimates and shaded regions show uncertainty. Horizontal reference lines and bands indicate B_{MSY} and F_{MSY} and their associated uncertainty.

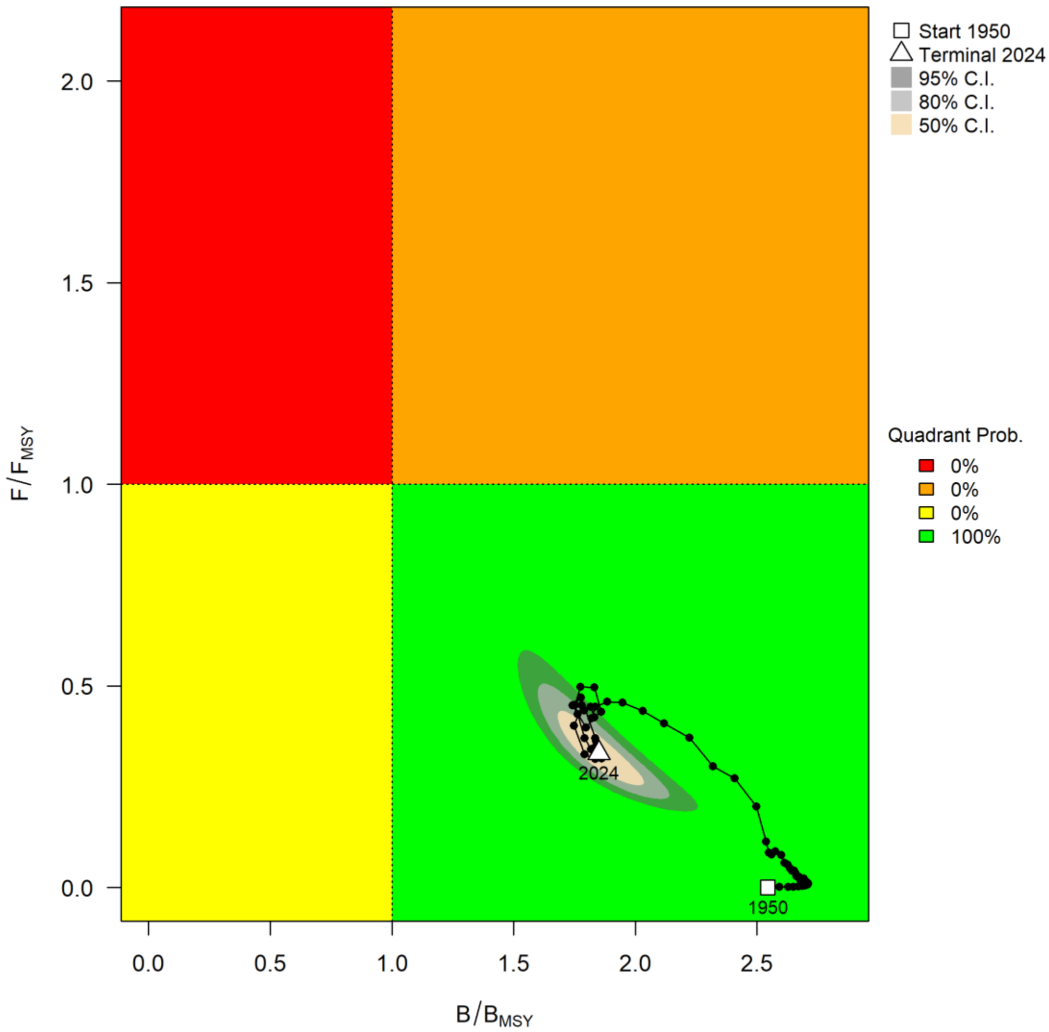


Fig. 4. Kobe plot for the selected Scenario j showing the estimated stock-status trajectory from 1950 to 2024 and the terminal 2024 estimate. Confidence regions of 50%, 80%, and 95% are shown for the terminal stock status.

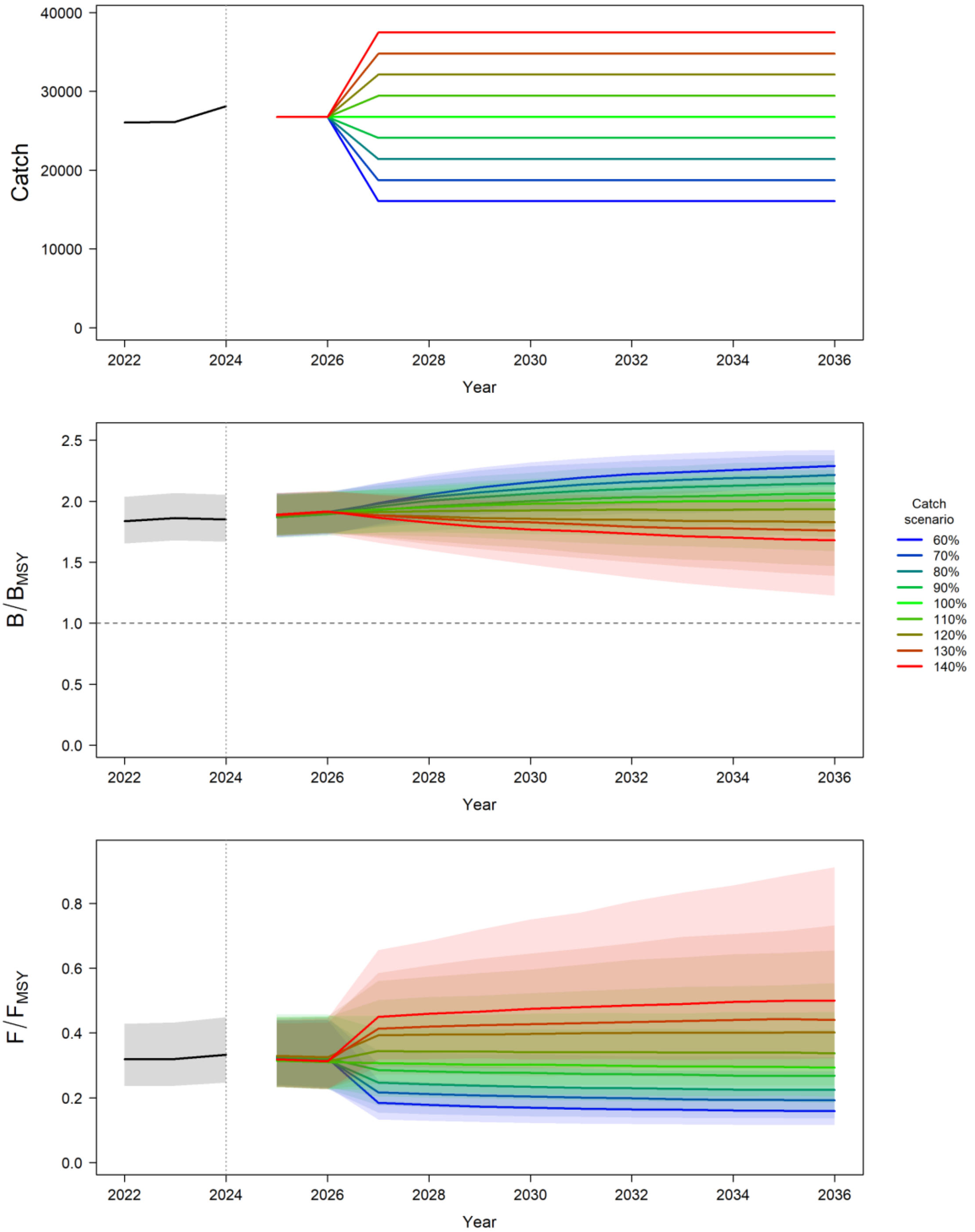


Fig. 5. Projected catch, *B*/*B*MSY, and F/F_{MSY} under alternative constant-catch scenarios for the selected Scenario j.

Tables

Table 1. Standardized CPUE indices used in the SPiCT assessment.

CPUE index	Fleet label	Time series
I1	ne_idn	2006–2024
I2	ne_jap	1994–2024
I3	ne_twn	2005–2023
I4	nw_jap	1994–2024
I5	nw_twn	2005–2023
I6	se_jap	1994–2024
I7	se_twn	2005–2023
I8	sw_eu_esp	2001–2024
I9	sw_eu_prt	2000–2024
I10	sw_eu_reu	2005–2025
I11	sw_jap	1994–2024
I12	sw_twn	2005–2023
I13	sw_zaf	2004–2025

Table 2. Stage 1 production model scenarios evaluated in the SPiCT assessment.

Scenario	Model	Key specifications / CPUE configuration	Description
a	Fox	*n*=0.99, fixed	Fixed Fox formulation evaluated as a candidate production model. The value 0.99 approximates the Fox limit while avoiding numerical issues at exactly *n*=1.
b	Schaefer	*n*=2, fixed	Standard Schaefer production formulation evaluated as an alternative fixed-shape model.
c	Pella Tomlinson	*n* estimated; $\log(n) \sim N[\log(2), 2^2]$	Generalized production formulation with *n* estimated using the broad SPiCT prior.
d	Pella Tomlinson, no *n* prior	*n* estimated; prior on log(*n*) disabled	Prior sensitivity and identifiability test for the Pella Tomlinson shape parameter. All other settings were identical to Scenario c. This scenario was not treated as a primary base-model candidate.

Table 3. Comparison of Stage 1 production model scenarios.

Scenario	Production formulation	*n* (80% CI)	B_{MSY}/K	MSY (80% CI)	Terminal B/B_{MSY} (80% CI)	Terminal F/F_{MSY} (80% CI)	Mohn's ρ_B	Mohn's ρ_F
a	Fox	0.99 fixed	0.366	39,146 (35,389–43,301)	1.678 (1.532–1.839)	0.414 (0.328–0.521)	-0.027	0.051
b	Schaefer	2 fixed	0.5	34,664 (32,168–37,355)	1.286 (1.184–1.395)	0.609 (0.499–0.745)	-0.022	0.033
c	Pella Tomlinson	0.600 (0.137–2.622)	0.279	43,590 (30,780–61,731)	2.163 (0.948–4.936)	0.288 (0.089–0.932)	-0.026	0.055
d	Pella Tomlinson, no prior on (n)	0.143 (0.00056–36.79)	0.103	58,152 (23,423–144,393)	5.709 (0.075–431.797)	0.082 (0.00044–15.299)	-0.131	0.248

Note: All scenarios achieved numerical convergence. Confidence intervals are 80%. Scenario d was used as an identifiability and prior sensitivity analysis rather than as a primary candidate base model.

Table 4. Stage 2 CPUE index sensitivity scenarios evaluated using the selected Fox production formulation.

Scenario	Model	Key specifications / CPUE configuration	Description
e	Fox	I10 excluded	Tests the influence of I10, which showed strong residual autocorrelation across alternative production model formulations.
f	Fox	I11 excluded	Tests the influence of I11, which showed strong residual autocorrelation across alternative production model formulations.
g	Fox	I2 excluded	Tests the influence of I2, which showed strong residual autocorrelation across alternative production model formulations.
h	Fox	I6 excluded	Tests the influence of I6, which showed significant residual autocorrelation across alternative production model formulations.
i	Fox	I8 excluded	Influence sensitivity for I8, which has a very small estimated observation-error standard deviation and may therefore receive disproportionate weight in the model.
j	Fox	I2, I6, I10, and I11 excluded	Combined exclusion of the four indices showing persistent residual autocorrelation across production model formulations.
k	Fox	I2, I6, I8, I10, and I11 excluded	More conservative stress test incorporating both residual diagnostics and potential index influence.

Table 5. Comparison of the all-index Fox model and alternative CPUE index sensitivity scenarios.

Scenario	Excluded indices	r	K	MSY	Terminal B/B_{MSY}	Terminal F/F_{MSY}	Mohn's ρ_B	Mohn's ρ_F	FDR-significant indices remaining
a	none	0.28	378,477	39,146	1.678	0.414	-0.027	0.051	I2, I6, I10, I11
e	I10	0.278	382,235	39,288	1.686	0.411	-0.029	0.054	I2, I6, I11
f	I11	0.278	395,371	40,653	1.737	0.385	-0.021	0.042	I2, I6, I10
g	I2	0.279	377,640	38,983	1.672	0.417	-0.025	0.045	I6, I10, I11
h	I6	0.285	388,457	40,914	1.745	0.381	-0.036	0.075	I2, I10, I11
i	I8	0.425	232,037	36,473	1.727	0.428	0.043	-0.053	I2, I5, I6, I10, I11
j	I2, I6, I10, I11	0.286	417,820	44,234	1.85	0.333	-0.026	0.063	None
k	I2, I6, I8, I10, I11	0.38	253,398	35,639	1.62	0.468	0.034	-0.042	None

Table 6. Swordfish: SPiCT 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 2022 to 2024 (26,761 t), ± 10%, ± 20%, ± 30%, ± 40% projected for 3 and 10 years.

Reference point and projection timeframe	**Alternative catch projections (relative to the average catch level from 2022 to 2024 of 26,761 t) and probability (%) of violating MSY-based target reference points (Btarg = B_{MSY}; Ftarg = F_{MSY})**								
	60% (16,057)	70% (18,733)	80% (21,409)	90% (24,085)	100% (26,761)	110% (29,437)	120% (32,113)	130% (34,790)	140% (37,466)
$B_{2029} < B_{MSY}$	0	0	0	0	0	0	0	0	0.1
$F_{2029} > F_{MSY}$	0	0	0	0	0.1	0.2	0.4	0.8	1.1
$B_{2036} < B_{MSY}$	0	0	0	0.1	0.1	0.5	0.9	2.0	3.3
$F_{2036} > F_{MSY}$	0	0	0	0.1	0.3	1.1	2.1	3.8	6.0

Table 7. Swordfish: SPiCT Indian Ocean assessment Kobe II Strategy Matrix. Probability (percentage) of achieving the green quadrant of the KOBE plot, for a range of constant catch projections. Catch in 2025 and 2026 are fixed at 26,761 t

Catch (t)\Year	2027	2028	2029	2030	2031	2032	2033	2034	2035	2036
16,057 (60%)	100	100	100	100	100	100	100	100	100	100
18,733 (70%)	100	100	100	100	100	100	100	100	100	100
21,409 (80%)	100	100	100	100	100	100	100	100	100	100
24,085 (90%)	100	100	100	100	100	99.9	99.9	99.9	99.9	99.9
26,761 (100%)	100	100	99.9	99.9	99.9	99.9	99.9	99.8	99.8	99.7
29,437 (110%)	99.9	99.8	99.8	99.8	99.6	99.6	99.6	99.4	99.3	98.9
32,113 (120%)	99.8	99.6	99.6	99.2	99.2	99.2	99.0	98.5	98.2	97.9
34,790 (130%)	99.8	99.7	99.2	99.2	98.7	97.8	97.7	97.0	96.3	96.2
37,466 (140%)	99.6	99.2	98.9	98.4	97.8	96.8	95.7	94.8	94.2	94.0

Appendix

Scenario a

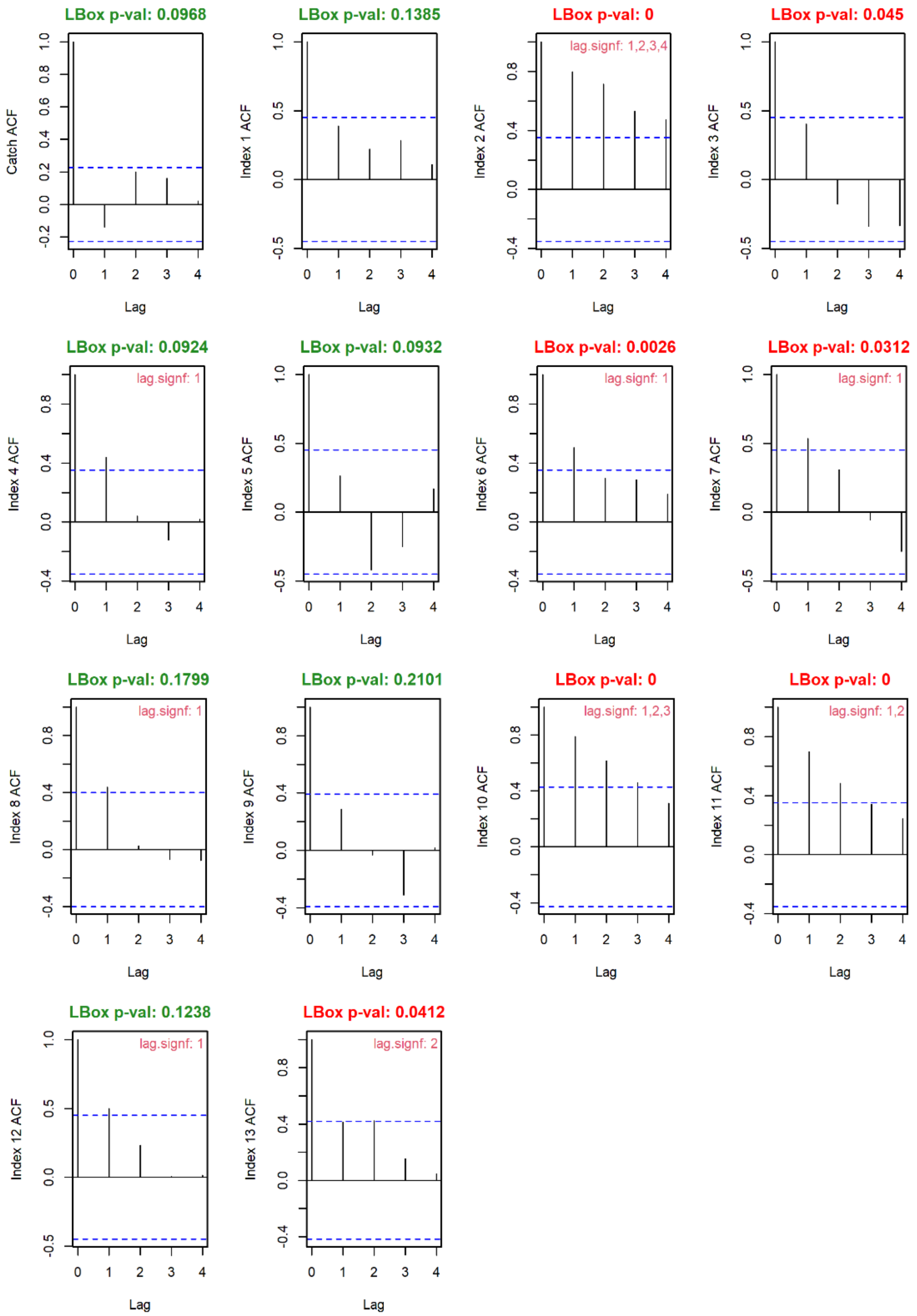
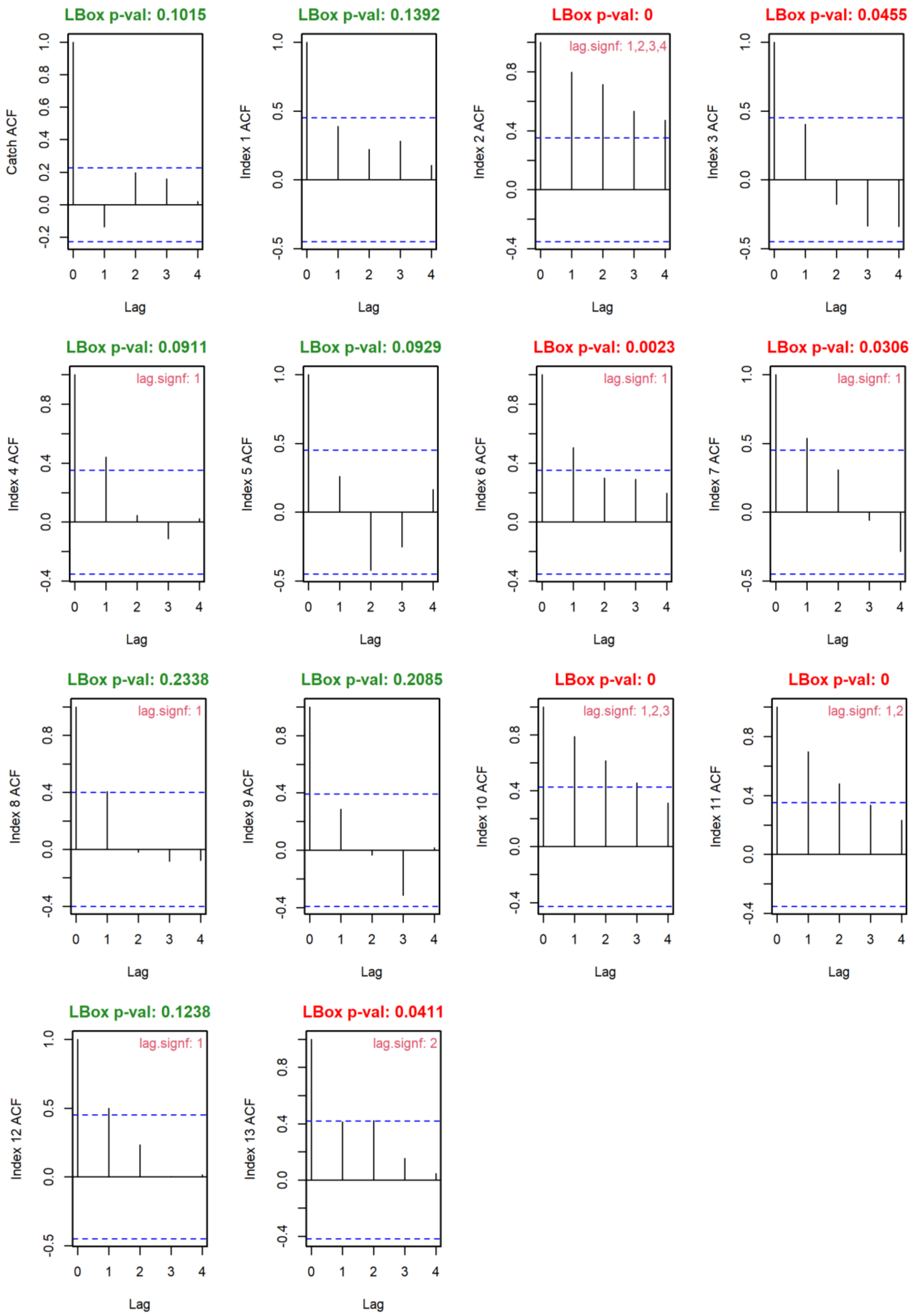
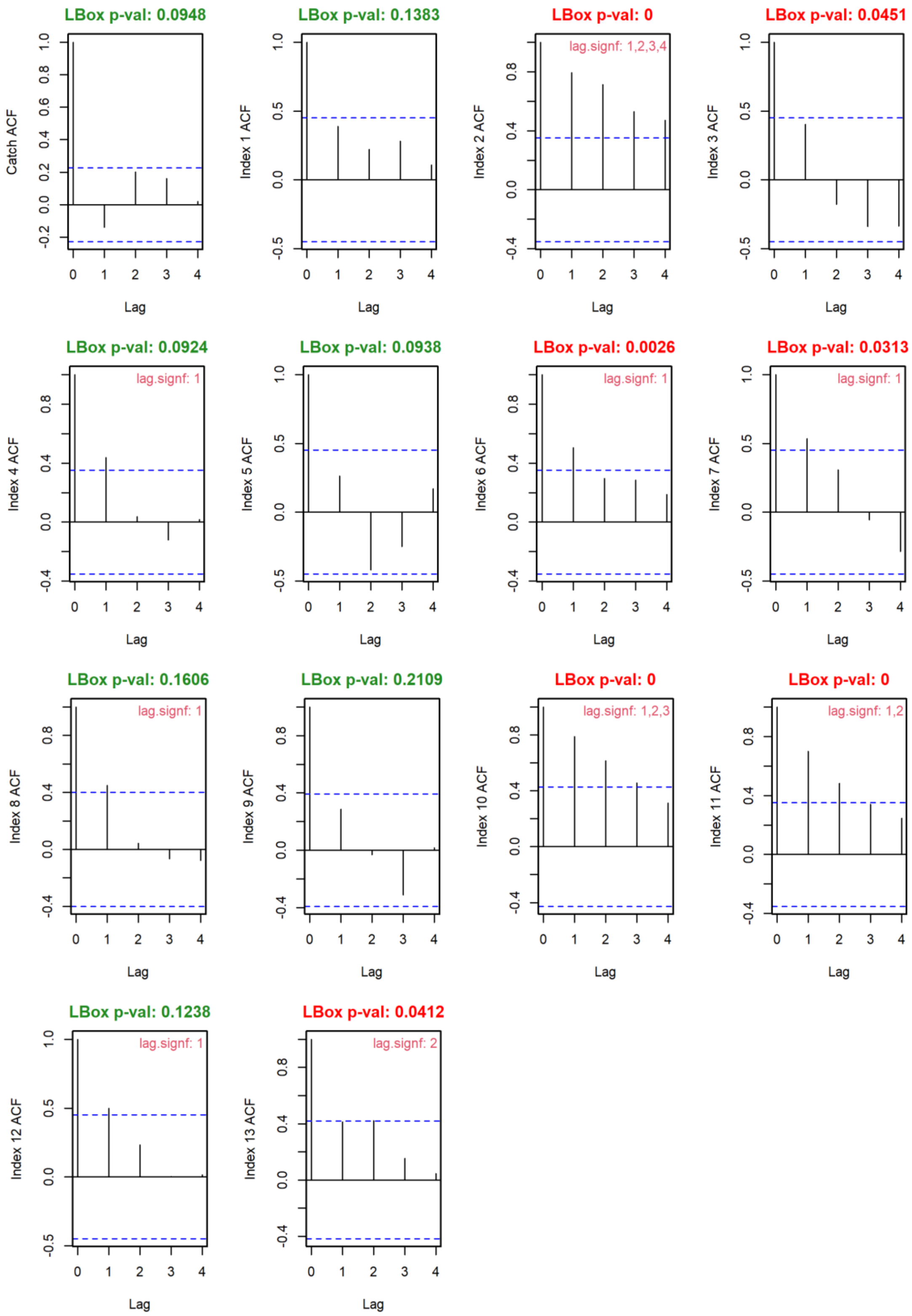


Fig. A1. One-step-ahead residual autocorrelation diagnostics for Stage 1 Scenarios a–d.

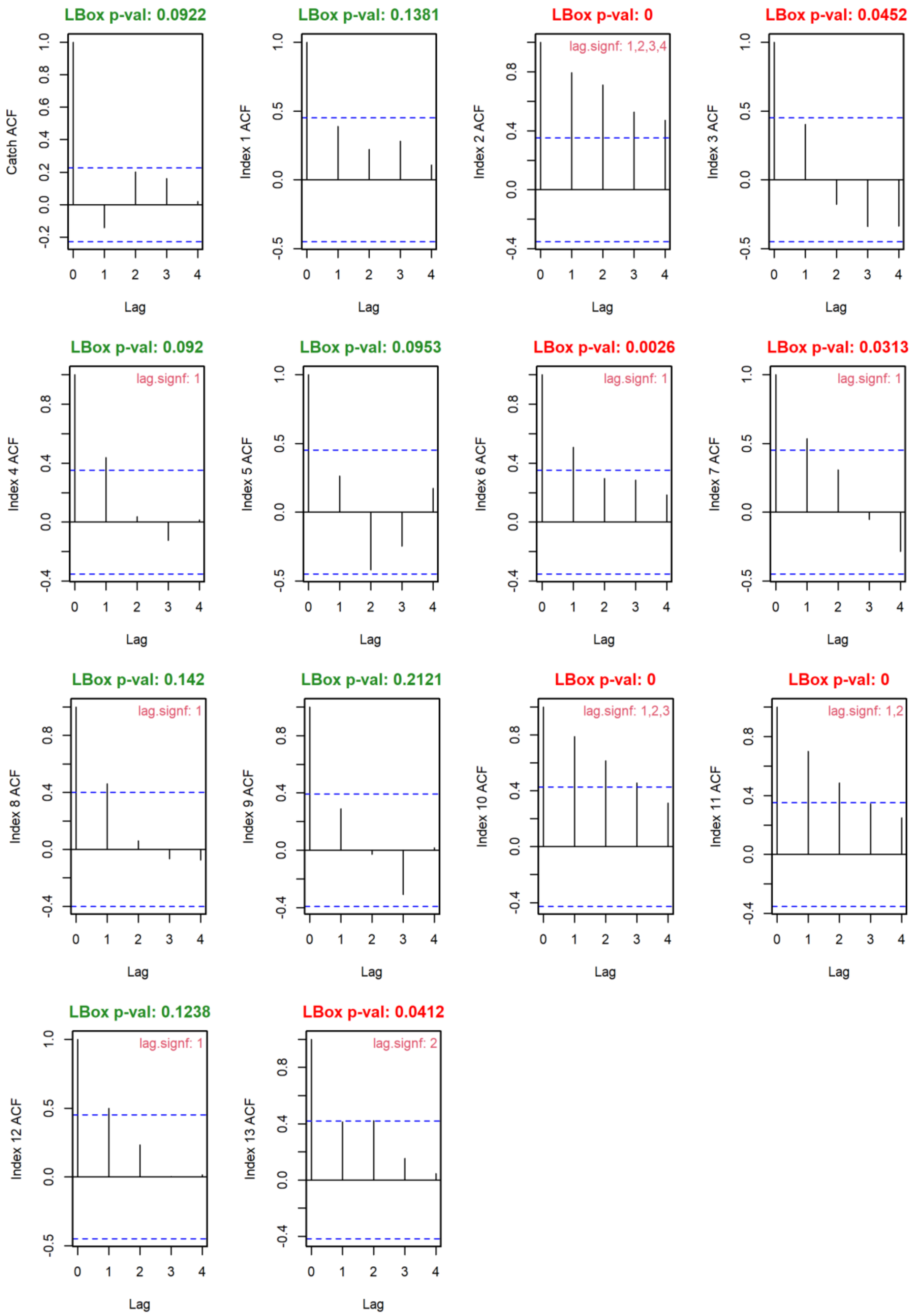
Scenario b


LBox p-val: 0.1015
Catch ACF
LBox p-val: 0.1392
Index 1 ACF
LBox p-val: 0
lag.signf: 1,2,3,4
Index 2 ACF
LBox p-val: 0.0455
Index 3 ACF
LBox p-val: 0.0911
lag.signf: 1
Index 4 ACF
LBox p-val: 0.0929
Index 5 ACF
LBox p-val: 0.0023
lag.signf: 1
Index 6 ACF
LBox p-val: 0.0306
lag.signf: 1
Index 7 ACF
LBox p-val: 0.2338
lag.signf: 1
Index 8 ACF
LBox p-val: 0.2085
Index 9 ACF
LBox p-val: 0
lag.signf: 1,2,3
Index 10 ACF
LBox p-val: 0
lag.signf: 1,2
Index 11 ACF
LBox p-val: 0.1238
lag.signf: 1
Index 12 ACF
LBox p-val: 0.0411
lag.signf: 2
Index 13 ACF
Lag

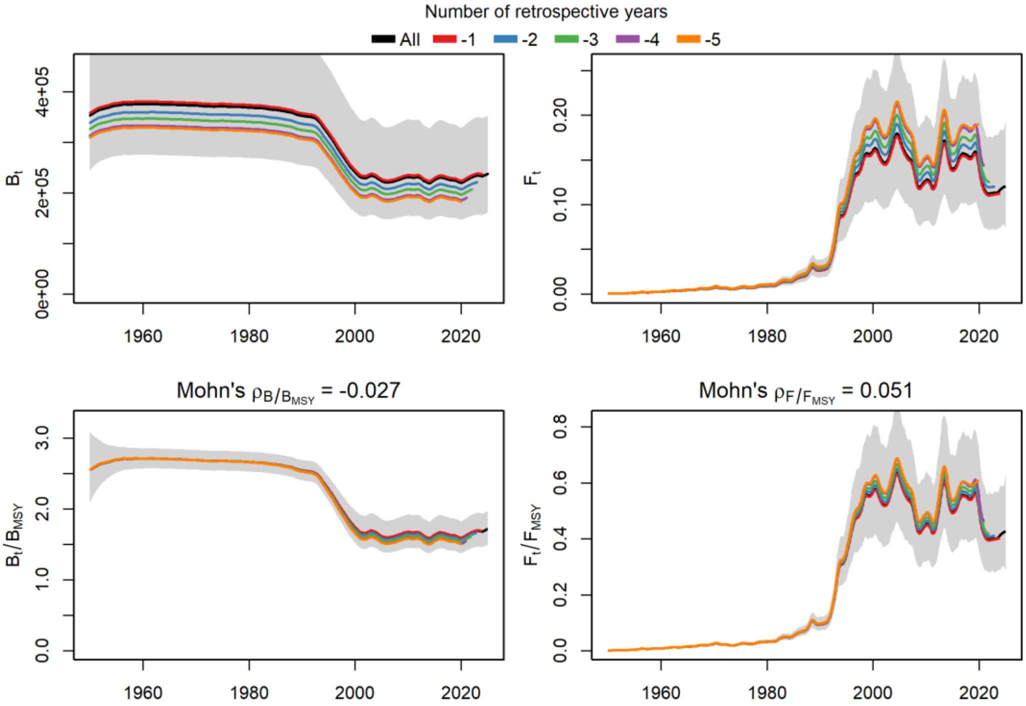
Scenario c

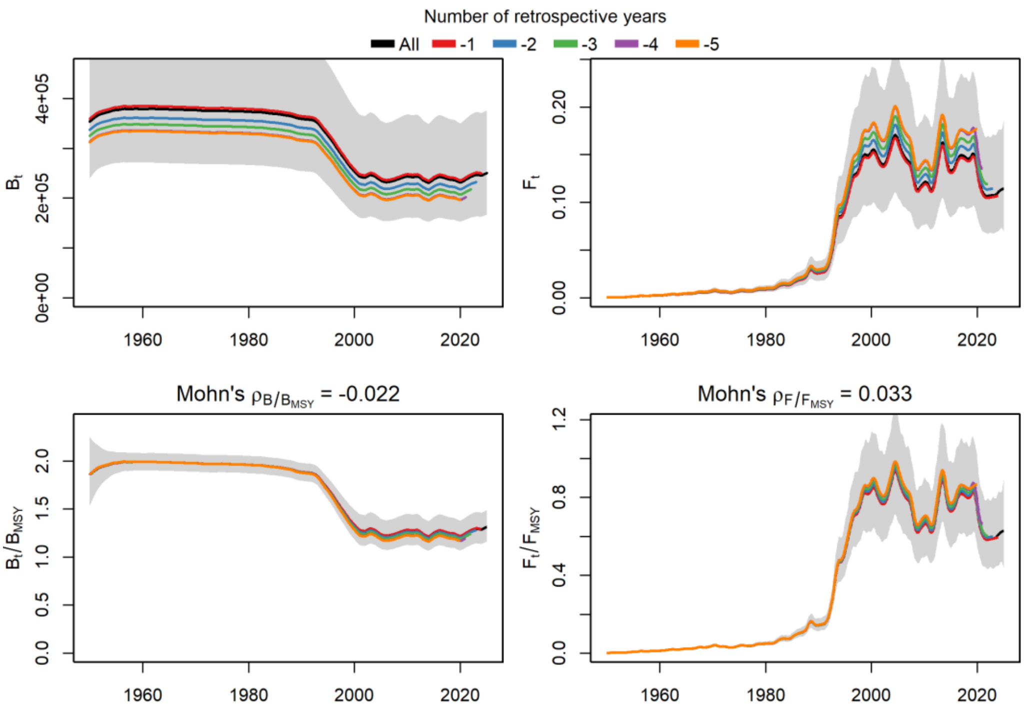

LBox p-val: 0.0948
Catch ACF
LBox p-val: 0.1383
Index 1 ACF
LBox p-val: 0
lag.signf: 1,2,3,4
Index 2 ACF
LBox p-val: 0.0451
Index 3 ACF
LBox p-val: 0.0924
lag.signf: 1
Index 4 ACF
LBox p-val: 0.0938
Index 5 ACF
LBox p-val: 0.0026
lag.signf: 1
Index 6 ACF
LBox p-val: 0.0313
lag.signf: 1
Index 7 ACF
LBox p-val: 0.1606
lag.signf: 1
Index 8 ACF
LBox p-val: 0.2109
Index 9 ACF
LBox p-val: 0
lag.signf: 1,2,3
Index 10 ACF
LBox p-val: 0
lag.signf: 1,2
Index 11 ACF
LBox p-val: 0.1238
lag.signf: 1
Index 12 ACF
LBox p-val: 0.0412
lag.signf: 2
Index 13 ACF
Lag

Scenario d


LBox p-val: 0.0922
Catch ACF
LBox p-val: 0.1381
Index 1 ACF
LBox p-val: 0
lag.signf: 1,2,3,4
Index 2 ACF
LBox p-val: 0.0452
Index 3 ACF
LBox p-val: 0.092
lag.signf: 1
Index 4 ACF
LBox p-val: 0.0953
Index 5 ACF
LBox p-val: 0.0026
lag.signf: 1
Index 6 ACF
LBox p-val: 0.0313
lag.signf: 1
Index 7 ACF
LBox p-val: 0.142
lag.signf: 1
Index 8 ACF
LBox p-val: 0.2121
Index 9 ACF
LBox p-val: 0
lag.signf: 1,2,3
Index 10 ACF
LBox p-val: 0
lag.signf: 1,2
Index 11 ACF
LBox p-val: 0.1238
lag.signf: 1
Index 12 ACF
LBox p-val: 0.0412
lag.signf: 2
Index 13 ACF
Lag

Scenario a

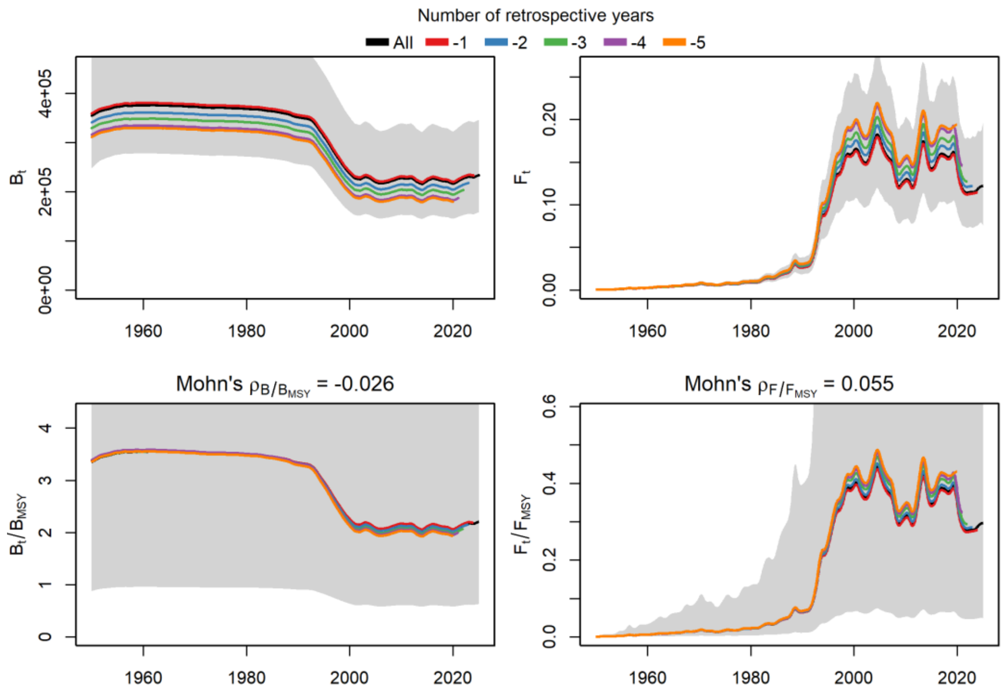




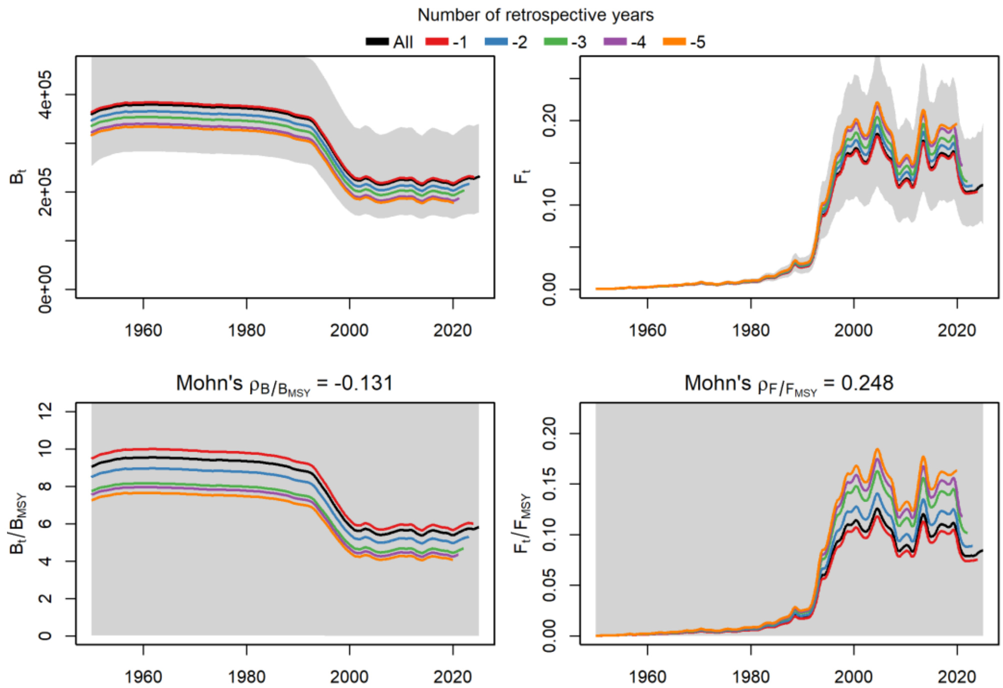
Scenario b

Fig. A2. Retrospective analyses for Stage 1 Scenarios a–d.

Scenario c


Number of retrospective years
All -1 -2 -3 -4 -5
B_t
F_t
Mohn's ρB/BMSY = -0.026
Mohn's ρF/FMSY = 0.055
B_t/B_MSY
F_t/F_MSY

Scenario d


Number of retrospective years
All -1 -2 -3 -4 -5
B_t
F_t
Mohn's ρB/BMSY = -0.131
Mohn's ρF/FMSY = 0.248
B_t/B_MSY
F_t/F_MSY

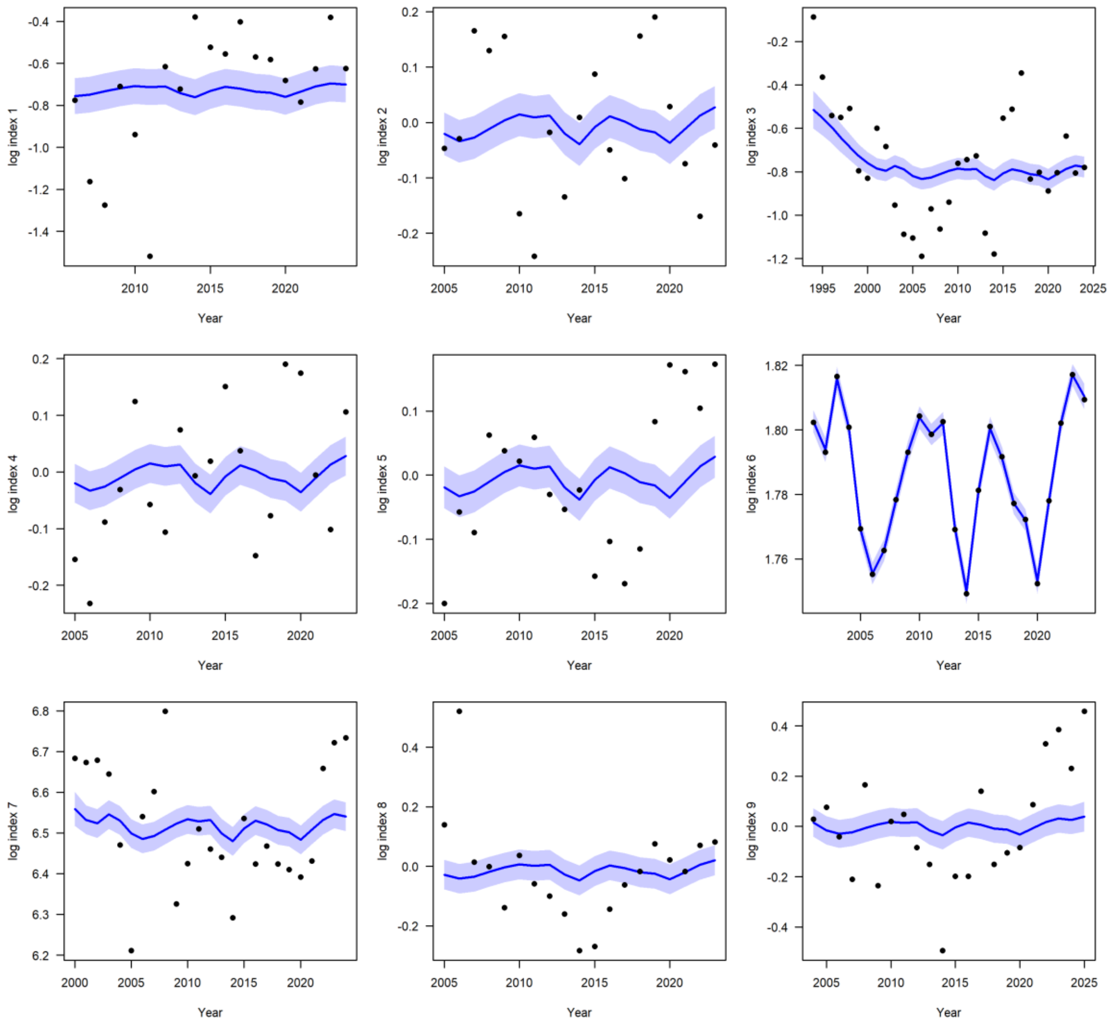


Fig. A3. Model fits to standardized CPUE indices for the selected Scenario j. Indices 1–9 correspond to original I1, I3, I4, I5, I7, I8, I9, I12, and I13, respectively.

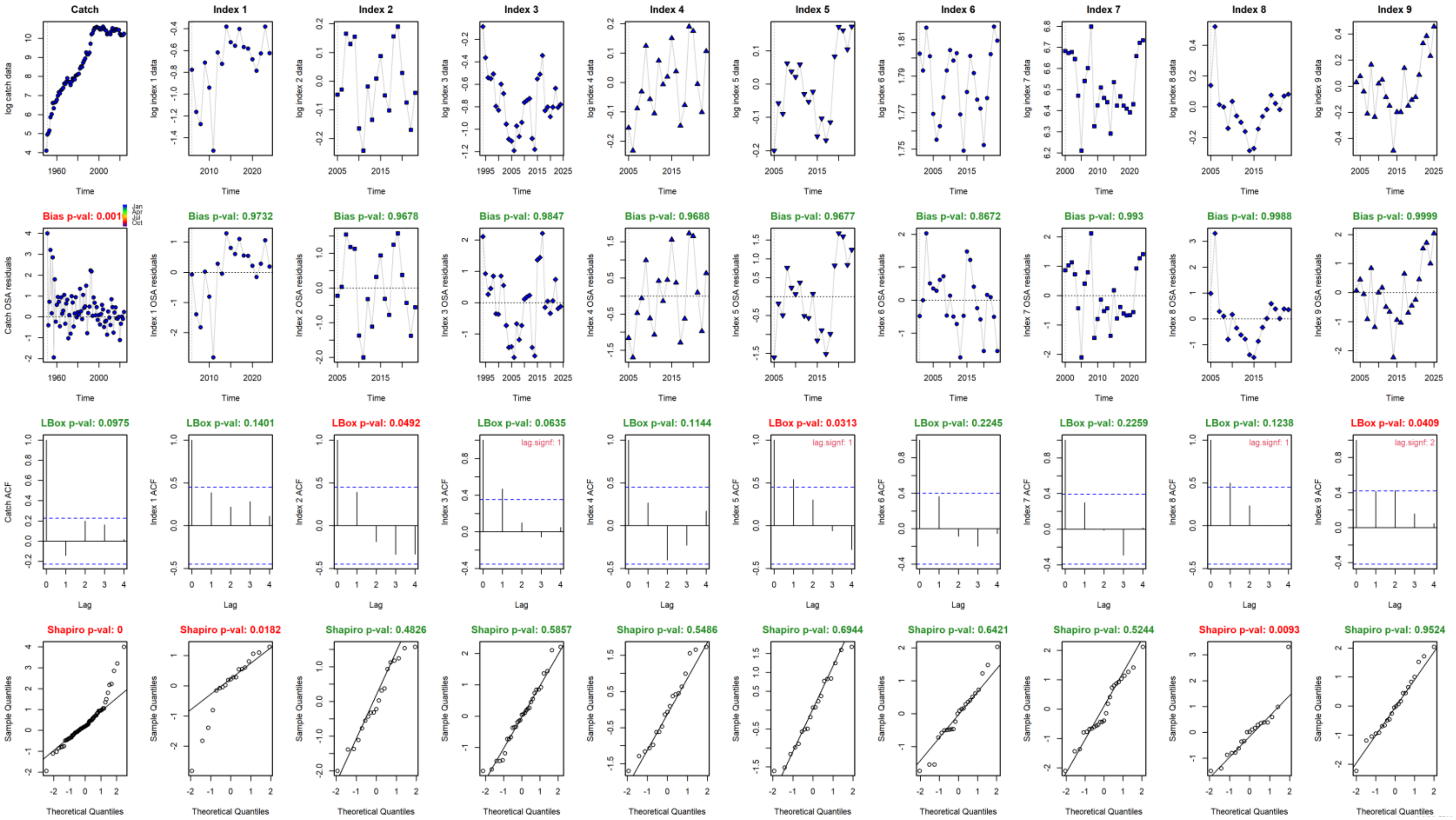


Fig. A4. One-step-ahead residual diagnostics for the selected Scenario j. Indices 1–9 correspond to original I1, I3, I4, I5, I7, I8, I9, I12, and I13, respectively.

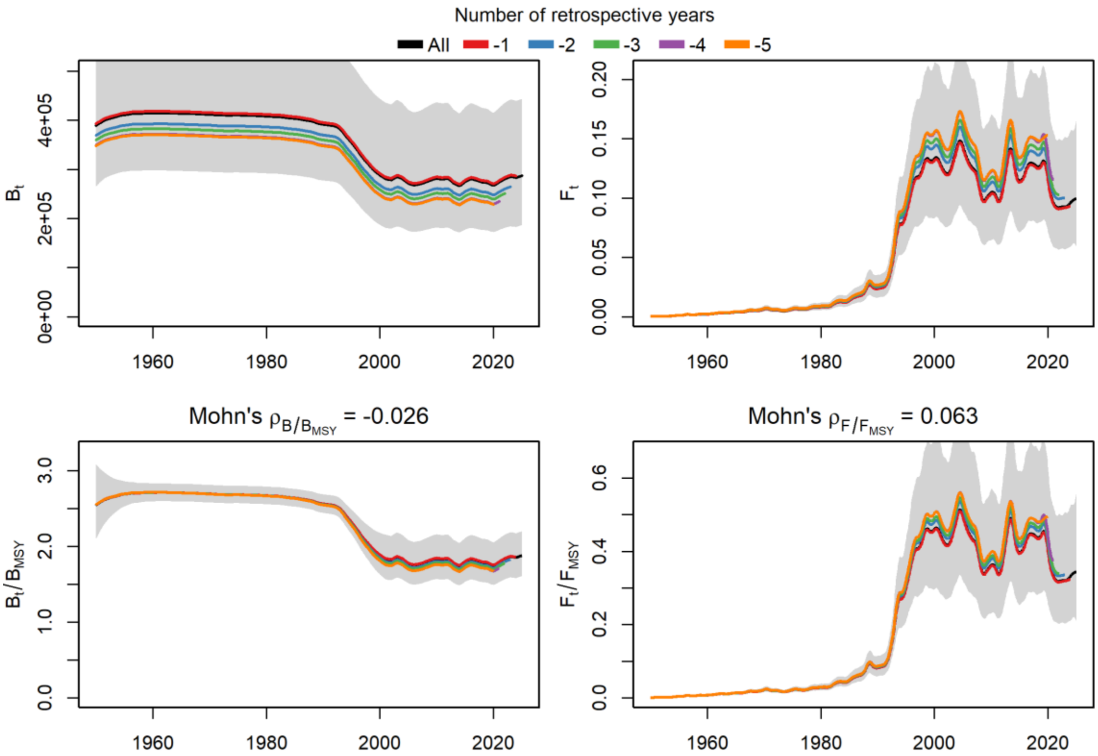


Fig. A5. Retrospective analysis for the selected Scenario j. Mohn's ρ was −0.026 for B/B_{MSY} and 0.063 for F/F_{MSY}.