IOTC-2026-WPB24-16

STANDARDIZED SWORDFISH CPUE BASED ON OBSERVER DATA FROM THE FRENCH REUNION-BASED PELAGIC LONGLINE FISHERY IN THE SOUTH-WEST INDIAN OCEAN (2007-2025)

Chloe Tellier[1,2] and Philippe S. Sabarros[1,2,*]

[1]MARBEC, Univ Montpellier, CNRS, Ifremer, IRD, Sete, France

[2]Institut de Recherche pour le Developpement (IRD), Observatoire des Ecosystemes Pelagiques Tropicaux Exploites (Ob7), Sete, France

*corresponding author: philippe.sabarros@ird.fr

Purpose, expectation and action intended (select one or more and include corresponding text)

Discussions (for information and open discussion):

This paper presents a standardized CPUE series for swordfish, for the upcoming stock assessment of this species. This standardization is based on observer and self-reported data collected between 2007 and 2025, for the French Reunion-based pelagic longline fishery operating in the southwestern Indian Ocean. It will be presented in Section 5.2.

Abstract

For the upcoming stock assessment of swordfish (*Xiphias gladius*), we developed a standardized catch-per-unit-effort (CPUE) series for this species, based on the French Reunion-based pelagic longline fishery operating in the southwestern Indian Ocean. We used observer and self-reported data collected between 2007 and 2025. This observer-based series is intended to be compared with a second standardized CPUE derived from logbook data. Observer data provide more detailed information for standardization, whereas logbook data provide broader fishery coverage. Comparing the trends derived from these two data sources will provide an opportunity to assess the consistency of the resulting indices of relative abundance. Temporal, spatial, operational and environmental covariates were considered during model selection. Three modelling approaches were evaluated: conventional Generalized Linear Mixed Models (GLMMs), GLMMs incorporating a spatial random field, and a GLMM incorporating a spatiotemporal random field, with the latter two fitted using the *sdmTMB* R package. Poisson and Negative Binomial distributions were also compared. The retained index was estimated using a GLMM with a spatiotemporal random effect and a Negative Binomial distribution, which provided the best fit to the overdispersed catch data. For the upcoming swordfish stock assessment, we recommend using the standardized CPUE series for the period between 2011 and 2025, during which monitoring coverage is substantially higher than in the earlier years. Over this period, the standardized swordfish CPUE shows a slight but statistically significant increasing trend.

Keywords

Swordfish | CPUE standardization | Longline | Observer data | Western Indian Ocean | sdmTMB | GLMM

1. Introduction

Fishery-dependent CPUE indices are commonly used as indices of relative abundance for both target species, such as tunas and billfishes, and non-target species, such as sharks, particularly when fishery-independent abundance indicators are unavailable. However, observed catch rates may be influenced by a range of factors unrelated to changes in population abundance, including changes in fishing effort, fishing practices and strategies, fishing gear, vessel characteristics, spatial and temporal fishing patterns, and the overlap between fishing operations and the distribution of the species. CPUE data therefore need to be standardized to account for these sources of variability before being used as indices of relative abundance in stock assessments (Maunder and Punt, 2004).

The French Reunion-based pelagic longline fishery operates mainly in the southwestern Indian Ocean, including waters around Reunion Island, Madagascar and the Mozambique Channel. Swordfish (*Xiphias gladius*) is the main target species of this fishery, although substantial catches of tuna species are also recorded. Fishing operations mainly consist of relatively shallow longline sets deployed in the evening, left to fish overnight, and retrieved the following morning.

The spatial distribution of both fishing effort and swordfish catch rates may vary over time. Consequently, changes in the spatial distribution of the fleet may affect observed catch rates independently of changes in swordfish abundance. In addition to conventional generalized linear modelling approaches, spatial and spatiotemporal models provide a framework for accounting for spatial structure in catch rates and for potential changes in this structure through time.

The objective of this study is to develop a standardized CPUE series for swordfish caught by the French Reunion-based pelagic longline fleet over the period 2007–2025. Observer and self-reported data collected from this fishery were used to evaluate the effects of temporal, spatial, operational and environmental covariates on swordfish catch rates. Several modelling approaches, including conventional GLMMs, GLMMs with spatial random effect, and GLMM with spatiotemporal random effect, were compared. The resulting standardized CPUE series is intended to provide an index of relative abundance for consideration in the upcoming stock assessment of swordfish in the Indian Ocean.

2. Material and methods

2.1. Data

We used data collected by sea-going observers on French longline vessels (Bach et al., 2008) as well as data collected by fishermen themselves called "self-reported data" (Bach et al., 2013). Data were collected through CAPPER (2007-2008) and EU Data Collection Framework (2009-2025; Reg 199/2008 and 665/2008). The coverage in number of hooks monitored is presented in Figure 1. We retained a total of 5784 fishing operations monitored between 2007 and 2025 from the core fishing area that consists of 5°x5° squares with more than 120 fishing operations (Figure 2).

2.2. Covariate selection

The list of candidate covariates was determined based on previous work of CPUE standardization for swordfish (Da Silva et al., 2025) and standardization for other species carried out with the same dataset (Sabarros et al., 2017; Sabarros et al., 2021; Sabarros et al., 2025; Tellier and Sabarros, 2025). Candidate covariates included temporal, spatial, operational and environmental variables.

Candidate covariates are:

- Fixed effects:
    - *year* (factor): year of the fishing operation, from 2007 to 2025. Year is included in all models to estimate annual relative abundance indices.
    - *quarter* (factor): calendar quarter, from Q1 to Q4.
    - *month* (factor): calendar month, from January to December. Month and quarter are considered as alternative temporal covariates and are therefore mutually exclusive during model selection.
    - *longitude* (continuous): longitude of the fishing operation, specifically the longitude where the line starts being retrieved (hauling).
    - *latitude* (continuous): latitude of the fishing operation, specifically the latitude where the line starts being retrieved (hauling).
    - *soaking_time* (continuous): time in hours from when the first hook is deployed to when the last hook retrieved.
    - *setting_start_time* (continuous): time (hh:mm) when the first buoy is deployed.
    - *hauling_end_time* (continuous): time (hh:mm) when the last buoy is retrieved.
    - *branchlines_per_basket_count* (continuous): number of hooks per basket as a relative index of fishing depth range/targeting.

- *percentage_circle_hooks* (continuous): relative proportion of circle hooks to other types of hooks (J-hooks, tuna hooks, Teracima hooks).
- *percentage_squid_bait* (continuous): proportion of squid bait relatively to other bait used (mackerel, etc.).
- *moon_illumination* (continuous): lunar illumination, expressed as an index ranging from 0 to 100. A value of 0 corresponds to minimum lunar illumination, approximately corresponding to the new moon, whereas a value of 100 corresponds to maximum lunar illumination, approximately corresponding to the full moon. Lunar illumination was calculated using the *lunar.illumination* function from the *lunar* R package.

- Random effects:
  - *vessel* (factor): the vessel name is used as a random effect given that we wanted to incorporate the vessel effect variability in the model but without estimating specific parameters for each vessel.

A forward stepwise model selection procedure was applied to identify covariates improving model fit. At each step, each remaining candidate covariate was added separately to the current model, and the resulting models were compared using the Akaike Information Criterion (AIC). For continuous covariates, alternative functional forms were evaluated, including a linear effect and spline representations with different numbers of degrees of freedom. The representation associated with the lowest AIC was retained for each candidate covariate.

The candidate producing the greatest reduction in AIC was selected at each step. A variable was retained only when its inclusion reduced the AIC by at least two units compared with the current model. The procedure was repeated until no remaining candidate covariate resulted in an AIC reduction of at least two units.

2.3. Comparison of modelling approaches

The response variable is the number of swordfish caught per fishing operation (*nb_caught*). Fishing effort, defined as the total number of hooks deployed during each fishing operation (*total_hooks_count*), is included as a log-transformed offset. This allowed swordfish catches to be modelled while accounting for differences in fishing effort among operations, thereby providing standardized estimates of catch rates, using a distribution adapted to count data.

We compared several modelling approaches: conventional GLMMs, GLMMs with a spatial random effect, and GLMM with a spatiotemporal random effect.

2.3.1. Conventional GLMMs

Standardized swordfish CPUE was first estimated using GLMMs fitted with the *glmmTMB* R package (Brooks et al., 2017). Candidate covariates were selected using the forward stepwise procedure described above. The covariate *year* was included as a categorical fixed effect in all models to estimate annual changes in standardized catch rates, while *vessel* was included as a random intercept.

Continuous covariates were evaluated as either linear or non-linear (represented using splines) effects. For spline representations, several degrees of freedom ranging from 2 to 5 were evaluated during model selection. For each selected continuous covariate showing a non-linear relationship with catch rates, the degree of freedom associated with the lowest AIC was retained, and will be used in the models adjusted with *sdmTMB*.

An exception was made for *percentage_circle_hooks*, for which only a linear representation was considered. This choice is based on the expected nature of the relationship between the proportion of circle hooks and catch rates and avoided fitting an unnecessarily complex non-linear relationship.

Poisson and Negative Binomial distributions were compared. A zero-inflated component was not considered because only approximately 3% of the fishing operations had zero swordfish catches (Figure 3). The distribution which provided the better fit was retained for subsequent analyses.

The best model fitted using the conventional GLMM approach is hereafter referred to as Mod 1.

2.3.2. GLMMs with a spatial random effect

Spatial GLMMs were fitted using *sdmTMB* R package (Anderson et al., 2022), by incorporating a spatial random field. The spatial effect accounts for residual spatial structure in swordfish catch rates that is not explained by the fixed effects included in the model.

In *sdmTMB*, the spatial field is approximated using a triangular mesh covering the study area. Rather than estimating a separate spatial effect at every fishing location, the spatial process is estimated at the mesh knots and interpolated between them. This approach substantially reduces the computational complexity of the spatial model while retaining the ability to represent spatial variation in catch rates.

The influence of mesh resolution was evaluated by fitting models using meshes containing 50, 80, 100 and 150 knots. Coarser meshes reduce computational requirements but may fail to capture fine-scale spatial variation, whereas finer meshes provide a more detailed representation of the spatial field at the expense of increased computational time. The mesh therefore influences the fit of the model, and consequently its AIC.

The different meshes evaluated are presented in Figure 5, and the standardized CPUE series obtained using each mesh are compared in Figure 6. Based on this sensitivity analysis, the mesh associated with the lowest AIC was retained for the GLMM with a spatiotemporal effect.

The best GLMM including a spatial effect is hereafter referred to as Mod 2.

2.3.3. GLMM with a spatiotemporal random effect

Spatiotemporal GLMM was also fitted using *sdmTMB* R package, by adding a spatiotemporal random field. While the spatial field represents persistent spatial variation in swordfish catch rates, the spatiotemporal field allows the spatial structure of catch rates to vary among years.

This approach can account for temporal changes in the spatial distribution of swordfish, including potential shifts in areas of relatively high and low catch rates through time. It can also help account for changes in the spatial distribution of fishing effort over the study period that may otherwise influence observed catch rates.

The GLMM including a spatiotemporal effect is hereafter referred to as Mod 3.

2.3.4. Selection of the final model

Mod 1, Mod 2 and Mod 3 were compared using AIC. The standardized annual CPUE series estimated by the three modelling approaches were also compared and are presented in Figure 7. The model with the lowest AIC was retained as the final model for standardization of CPUE.

Diagnostic and model evaluation results for Mod 1, Mod 2 and Mod 3 are presented in Table 1. A summary of the retained model is provided in Table 2. The estimated partial effects of the selected continuous covariates are presented in Figure 8, the estimated effects of the selected quantitative covariate are presented in Figure 9, and the graphical analysis of scaled residuals obtained using the *DHARMa* R package (Hartig et al., 2024) is shown in Figure 10, Figure 11, Figure 12 and Figure 13.

2. 4. Standardized CPUE predictions

Annual standardized CPUE estimates were derived from predictions of the retained model. A prediction dataset was created containing one observation for each year of the study period (2007 to 2025). The year factor was varied across predictions, while all selected continuous covariates were held constant at their median values calculated over the complete dataset. The month effect was fixed to January (month = 1), corresponding to the reference conditions used for the standardized predictions. Fishing effort was standardized to 1000 hooks by setting the log-transformed offset to log(1000).

Predictions were obtained on the link scale using the *predict* function from the *sdmTMB* R package. Random effects and the spatial and spatiotemporal random fields were excluded from the predictions in order to estimate annual standardized CPUE under common fishing conditions. Predicted values were then exponentiated to obtain standardized CPUE estimates expressed as the expected number of swordfish caught per 1000 hooks.

Confidence intervals of 95% were calculated on the link scale using the estimated standard errors:

$$CI_{lower} = \exp(\eta - 1.96 \times SE_{\eta})$$

$$CI_{upper} = \exp(\eta + 1.96 \times SE_{\eta})$$

where η is the predicted value on the link scale and SE_{η} is its associated standard error.

For comparison, nominal annual CPUE is calculated as:

$$\text{Nominal } CPUE_y = \frac{\sum Swordfish\ caught_y}{\frac{\sum Hooks\ deployed_y}{1000}}$$

where the sums were calculated across all fishing operations monitored during year y.

The annual standardized CPUE estimates from the retained model are presented in Table 3 and Figure 14. Standardized annual CPUE values scaled by their overall mean are also presented in Table 3 and Figure 15.

3. Results

3.1. Model selection and comparison

The results of the model selection procedure and the characteristics of the three final models are presented in Table 1.

The Negative Binomial distribution provided a better fit than the Poisson distribution (results not shown), consistent with the presence of overdispersion in the catch data (variance >> expected value), and was therefore retained.

For the conventional GLMM approach (Mod 1), the selected covariates include *month*, *hauling_end_time*, *latitude*, *longitude*, *moon_illumination*, *setting_start_time*, *branchlines_per_basket_count*, *soaking_time* and *percentage_circle_hooks*. The GLMMs with spatial or spatiotemporal random effects (Mod 2 and Mod 3) retain the same covariates, except for latitude and longitude, as spatial variation is accounted for by spatial and spatiotemporal random fields.

For non-linear effects, the degrees of freedom retained are 3 for *latitude*, 5 for *longitude* (for conventional GLMM approach only), 3 for *hauling_end_time*, 3 for *moon_illumination*, 5 for *setting_start_time*, 3 for *soaking_time* and 2 for *branchlines_per_basket_count*. A linear effect was imposed for *percentage_circle_hooks*.

For the GLMMs with a spatial random effect, mesh sensitivity was evaluated using 50, 80, 100 and 150 knots. The corresponding AIC values are 32419.40, 32403.92, 32391.46 and 32391.81, respectively. The mesh containing 100 knots provides the lowest AIC and was therefore retained for GLMMs including spatial or spatiotemporal random effect. The standardized CPUE series obtained using the different mesh resolutions are quite similar (Figure 6).

The three approaches (Mod 1, Mod 2, Mod 3) produced also quite similar standardized CPUE trends (Figure 7), indicating that the estimated annual indices are relatively robust to the choice of modelling framework. The final models have AIC values of 32458.62 for Mod 1, 32391.46 for Mod 2 and 32194.77 for Mod 3 (Table 1). The GLMM with a spatiotemporal random effect therefore provides the best fit according to AIC and is retained as the final model. The final retained model (Mod 3) is therefore:

nb_caught ~ *year* + *month* + s(*hauling_end_time*) + s(*moon_illumination*) + s(*setting_start_time*) + s(*soaking_time*) + s(*branchlines_per_basket_count*) + *percentage_circle_hooks* + (1|*vessel*) + offset(log(*total_hooks_count*))

where the notation s() represents the spline terms described above, and a spatiotemporal random field was included using *sdmTMB* R package.

3.2. Characteristics of the retained model

A summary of the retained model, a GLMM with a spatiotemporal random field, is presented in Table 2.

The estimated standard deviation associated with the vessel random effect is 0.21 on the log scale, indicating variability in swordfish catch rates among vessels after accounting for the other covariates included in the model.

The estimated month effect shows seasonal differences in standardized catch rates relative to January, which is used as the reference month. For example, the coefficient estimated for February is -0.08, corresponding to an expected catch rate of exp(-0.08) = 0.92 times that one estimated for January, all other covariates being held constant.

The annual coefficients also represent differences in catch rates relative to 2007, which is used as the reference year. For example, the coefficient estimated for 2025 is 0.21, corresponding to an expected catch rate of exp(0.21) = 1.23 times that one estimated for 2007 under otherwise identical conditions.

The dispersion parameter of the Negative Binomial model is estimated at 6.48. The catch data are overdispersed relative to the assumptions of a Poisson distribution, supporting the use of the Negative Binomial distribution selected during model comparison.

The estimated spatiotemporal AR1 correlation is 0.56, indicating moderate temporal persistence in the spatiotemporal random field: the residual spatial structure in swordfish catch rates showed partial continuity between successive years, rather than changing independently from one year to the next.

The estimated partial effects of the six selected continuous covariates are presented in Figure 8. The figures show the estimated spline effects and their approximate 95% confidence intervals, together with histograms representing the distribution of observations in the original dataset. The estimated *month* effects are presented in Figure 9.

3.3. CPUE trends

The annual nominal and standardized swordfish CPUE series are presented in Table 3 and Figure 14. Overall, the standardized CPUE series follows a similar temporal pattern to the nominal CPUE series, although differences are observed in the magnitude of the annual indices.

For comparison of relative annual changes, the nominal and standardized CPUE series scaled by their respective overall means are presented in Figure 15.

Acknowledging the relatively low coverage rate in the first years of the program (Figure 1), we focused on the period between 2011 and 2025. The standardized CPUE series exhibits an increasing trend during this period: standardized CPUE increased from 6.69 swordfish per 1000 hooks in 2011 to 8.72 swordfish per 1000 hooks in 2025. The highest standardized CPUE during this period is observed in 2024, with 9.48 swordfish per 1000 hooks. Throughout 2011-2025, the overall linear trend of the standardized CPUE is weak but significantly increasing (linear regression: b = 0.25214, p-value = 1.11e-05).

3.4. Residual analysis

Residual diagnostics for the final model were performed using the *DHARMa* R package. The diagnostic plots and associated tests are presented in Figure 10, Figure 11, Figure 12 and Figure 13. Overall, the graphical diagnostics indicate no major systematic patterns, although several formal tests detect small but statistically significant deviations from the assumptions of the fitted model.

The QQ plot of the simulated scaled residuals (Figure 10A) shows an overall close agreement with the expected uniform distribution. However, the Kolmogorov–Smirnov test detects a statistically significant deviation from uniformity ($D = 0.0257$, p-value < 0.05). The magnitude of this deviation is nevertheless really small.

The residuals plotted against the fitted values (Figure 10B) show no strong systematic structure. Residuals are distributed across the range of fitted values, and the fitted quantile regression curve remains close to the expected value of 0.5, suggesting no pronounced bias related to the magnitude of the model predictions.

The *DHARMa* dispersion test indicates significant underdispersion (dispersion = 0.82, p-value < 0.05), indicating that the variability of the observed residuals is lower than that expected from simulations under the fitted model (Figure 11A). This suggests that the fitted model does not reproduce all aspects of the variability present in the data.

The zero-inflation test also detects a significant deviation between the observed and simulated numbers of zero catches (ratioObsSim = 1.36, p-value < 0.05; Figure 11B). The ratio greater than one indicates that the observed dataset contains more zero catches than expected under the fitted model. However, the absolute difference remains limited: among the 5784 observations, the number of zero catches simulated under the fitted model ranges from 104 to 186, compared with 193 observed zero catches in the original dataset.

Finally, the bootstrapped outlier test detects a significant excess of extreme residuals (Figure 11C). Fifteen observations, corresponding to approximately 0.26% of the dataset, are classified as outliers, compared with an expected frequency of approximately 0.15% (p-value < 0.05). Although statistically significant, the absolute number of observations concerned remains low.

Residuals were also examined in relation to the covariates retained in the final model. No major systematic patterns are observed for the continuous covariates (Figure 12), with residual quantile regression curves generally remaining close to the expected value of 0.5. For the categorical covariates (Figure 13), no significant within-group deviations from uniformity are detected for either *month* or *year*. Residual variance is homogeneous among years but differed significantly among months according to Levene's test, indicating some remaining heterogeneity in residual dispersion across months.

Overall, the residual diagnostics reveals small but statistically significant departures from the fitted model assumptions. However, the graphical diagnostics shows no major systematic patterns, and the magnitudes of the detected deviations are generally limited. The retained model is therefore considered adequate for estimating the standardized CPUE series, while acknowledging that it does not fully reproduce all features of the observed catch distribution.

4. Discussion

The standardized CPUE series obtained accounts for spatial, temporal, environmental, operational and vessel-related sources of variability in catch rates. The resulting index followed a temporal pattern quite similar to the nominal CPUE (Figure 14), but is considered a more appropriate index of relative abundance than nominal catch rates for use in stock assessment.

4.1. Significant effects on swordfish CPUE

The final standardization model identified several covariates influencing swordfish catch rates. The estimated partial effects of the selected covariates are presented in Figure 8 and Figure 9. Their inclusion in the final model indicates that accounting for these factors improved model fit.

The vessel random effect indicates the presence of inter-vessel differences in fishing efficiency. Such differences may reflect variations in vessel characteristics, fishing practices, gear configuration, fishermen experiences or other vessel-specific factors not explicitly represented by the available covariates. Accounting for this variability is essential to avoid biased estimates and to ensure that the observed effects are not confounded by vessel-specific fishing practices.

Month is the most important temporal covariate retained during model selection and reveals a clear seasonal pattern in swordfish catch rates (Figure 9). Catch rates are generally higher between November and January and lower between April and August. This pattern indicates substantial seasonal variability in the availability and/or catchability of swordfish in the fishing area. Seasonal changes in catch rates could result from intra-annual changes in the spatial distribution of swordfish and from the seasonal patterns in fishing operations spatial distribution.

The proportion of circle hooks (Figure 8D) influences negatively swordfish catches. Although the estimated effect is relatively low across the observed range, the model suggests a gradual decrease in catch rates as the proportion of circle hooks increased. This result indicates that changes in hook configuration may influence swordfish catchability. The relatively wide confidence intervals, particularly at high proportions of circle hooks, should nevertheless be considered when interpreting the magnitude of this effect.

The timing of fishing operations influences also swordfish catches. The effect of *setting_start_time* (Figure 8B) increased progressively over most of the observed range and reached its maximum around 19:00. Relatively few observations are available at the extremes of the distribution, and the confidence intervals are consequently wider in these areas. The effect of *hauling_end_time* (Figure 8A) is comparatively less pronounced over the range containing most observations. Catch rates are relatively stable until approximately 12:00–13:00, followed by a slight increase around 14:00–15:00 and a marked decrease at later hauling times. Together, these results suggest that the temporal configuration of fishing operations may influence swordfish catchability.

Soaking time (Figure 8C) shows a positive effect on swordfish catch rates over the range represented by the majority of observations. This relationship was expected, as a longer period during which the hooks remain available to fish may increase the probability to catch swordfish. Although the total number of hooks is already accounted for through the offset, soaking time represents an additional dimension of fishing effort and catchability that may vary among operations. Its inclusion in the model therefore helps account for differences in effective fishing conditions that are not captured by hook numbers alone.

The number of branchlines per basket (Figure 8E) is retained as a non-linear operational covariate. In the range containing most observations, the estimated effect is negative, with catches decreasing as the number of branchlines per basket increased. The number of branchlines per basket can be considered as a proxy for the depth at which hooks are deployed, which suggests lower catch rates in deeper sets. In the Reunion Island fishery, hooks are typically deployed from 10 m below the surface down to ~120 m during night sets (Bach et al., 2014). The increase in the estimated effect at the highest values should be interpreted cautiously because these configurations are relatively uncommon and are associated with greater uncertainty.

Moon illumination (Figure 8F) is also retained as an important environmental covariate and shows a non-linear relationship with swordfish catch rates. According to the estimated spline effect, catch rates are highest at intermediate to relatively high levels of lunar illumination, with a maximum around 70% illumination, whereas lower catch rates are estimated around the new and full moon (Bach et al., 2014). These periods also correspond to the ranges of lunar illumination with the highest number of observed fishing operations. Lunar illumination may affect swordfish catchability through changes in their vertical distribution and behavior, thereby modifying the overlap between the depth range of the longline and that occupied by the fish. Similar relationships between lunar conditions and swordfish behavior or catchability have been reported in previous studies (Sabarros et al., 2014; Guyomard et al., 2004; Bach et al., 2014; Poisson et al., 2010).

These results highlight a limitation of the present model structure. The effects of the selected covariates were evaluated individually during the forward selection procedure, and potentially relevant interactions between covariates were not tested. In particular, an interaction between *branchlines_per_basket_count* and *moon_illumination* could be biologically meaningful, as previous studies have suggested that the vertical distribution of swordfish may vary according to lunar conditions (Bach et al., 2014). If fishing depth is not adjusted accordingly, the effect of fishing depth on catch rates may therefore depend on lunar illumination. Similarly, an interaction between *setting_start_time* and *hauling_end_time* could potentially account for combined effects of the temporal configuration of fishing operations.

Such interactions should be represented using tensor-product smooths in a Generalized Additive Modelling framework. In *sdmTMB*, however, spline and tensor-product terms must be constructed manually before model fitting, making the evaluation of non-linear interactions less straightforward. Consequently, only the main effects of the candidate covariates were evaluated in the present analysis. Future analyses could investigate whether the inclusion of biologically motivated interactions improves model fit and further reduces unexplained variability.

4.2. Relevance of the retained standardized CPUE series

This standardization is based exclusively on the core fishing grounds of the Reunion Island longline fleet (Figure 2). Restricting the analysis to this area avoided the potential biases associated with the inclusion of a small number of sets conducted in peripheral regions such as the northern Mozambique Channel or the high seas, where different environmental conditions and fishing strategies may influence catch rates. This conservative approach, which was also adopted in previous CPUE standardization analyses conducted for other species caught by the same fleet (Sabarros et al., 2025; Tellier and Sabarros, 2025), ensures that the resulting index is more representative and reliable.

The comparison of the three modelling approaches further supports the robustness of the resulting annual CPUE series. Although the best conventional GLMM (Mod 1), the best GLMM with spatial random field (Mod 2) and the GLMM with spatiotemporal random field (Mod 3) differed in their ability to account for residual spatial structure, the standardized CPUE trends obtained from these approaches are quite similar (Figure 7). Similarly, the standardized indices obtained using different mesh resolutions showed only limited differences (Figure 6). This indicates that the main temporal pattern of the standardized CPUE is relatively robust to the choice of modelling framework and mesh resolution. The spatiotemporal model nevertheless provided the lowest AIC and was therefore retained as the final model. Its improved fit is consistent with the presence of spatial variation in swordfish catch rates that changes through time and that is not fully accounted for by the fixed effects.

Residual diagnostics identified several statistically significant departures from the assumptions of the fitted model. The tests of residual uniformity, dispersion, zero inflation and outliers all detected significant deviations. However, the graphical diagnostics showed no major systematic patterns, and the magnitude of the detected deviations is limited. In particular, the deviation from residual uniformity is small, with a Kolmogorov-Smirnov statistic of only 0.0257. Similarly, although the zero-inflation test indicated more observed zero catches than expected under the fitted model, the absolute difference is limited: 193 zero catches were observed among 5784 fishing operations, compared with between 104 and 186 zero catches in simulations from the fitted model. The excess of outliers is also small in absolute terms, with 15 observations classified as outliers, corresponding to approximately 0.26% of the dataset, compared with an expected frequency of approximately 0.15%. These diagnostics support the use of the retained model for deriving the standardized CPUE index while acknowledging that further improvements to the model structure could be investigated in future analyses.

However, some caution is warranted for the early years of the series. Observer coverage was relatively low (<3% of effort in hooks observed) between 2007 and 2010 (Figure 1), raising concerns about data representativeness. For this reason, we recommend discarding this initial period and focusing on the 2011–2025 series for stock assessment and management purposes.

Within this period, the standardized CPUE shows an overall increasing trend. The linear regression fitted to the annual standardized indices indicated a positive and statistically significant slope. This result suggests an increase in the abundance of swordfish.

This series based on observer and self-collected data is intended to be compared with another standardized CPUE series, derived from logbook data. Comparing the trends derived from these two data sources will allow to assess the consistency of the increase in the relative abundance of swordfish.

5. Conclusion

This paper proposes a standardized swordfish CPUE series for the French Reunion-based pelagic longline fishery in the southwestern Indian Ocean, based on observer and self-reported data collected between 2007 and 2025. This series will be compared to another CPUE series, based on logbook data from the same fishery. Given the limited monitoring coverage during 2007–2010, we recommend that only the 2011–2025 portion of the series be considered for stock assessment purposes. Among the modelling approaches evaluated, the Negative Binomial GLMM with a spatiotemporal random effect provided the best fit and is retained to derive the annual abundance index. The standardized CPUE series is robust to the choice of modelling framework and mesh resolution, and residual diagnostics indicated that the retained model is adequate despite minor departures from some model assumptions. Over this period, the standardized CPUE showed a slight but statistically significant increasing trend.

Acknowledgements

We thank the observers and captains that collected data through CAPPER and EU DCF data collection programs.

References

Anderson, S.C., Ward, E.J., English, P.A., Barnett, L.A.K., Thorson, J.T., 2022. sdmTMB: An R Package for Fast, Flexible, and User-Friendly Generalized Linear Mixed Effects Models with Spatial and Spatiotemporal Random Fields. https://doi.org/10.1101/2022.03.24.485545

Bach, P., Rabearisoa, N., Filippi, T., Hubas, S., 2008. The first year of SEALOR: Database of SEA-going observer surveys monitoring the local pelagic longline fishery based in La Reunion, IOTC-2008-WPEB08-13. 8th IOTC Working Party on Ecosystems and Bycatch – Bangkok, Thailand.

Bach, P., Sabarros, P.S., Le Foulgoc, L., Richard, E., Lamoureux, J.P., Romanov, E., 2013. Self-reporting data collection project for the pelagic longline fishery based in La Reunion, IOTC-2013-WPEB09-42. 9th IOTC Working Party on Ecosystems and Bycatch – La Réunion, France.

Bach, P., Sabarros, P.S., Romanov, E., Puech, A., Capello, M., Lucas, V., 2014. Patterns of swordfish capture in relation to fishing time, moon illumination and fishing depth, IOTC-2014-WPB12-14. 12th IOTC Working Party on Billfish – Yokohama, Japan.

Brooks, M.E., Kristensen, K., Van Benthem, K.J., Magnusson, A., Berg, C.W., Nielsen, A., Skaug, H.J., Mächler, M., Bolker, B.M., 2017. Modeling zero-inflated count data with glmmTMB. https://doi.org/10.1101/132753

Da Silva, C., West, W., Kerwath, S., 2025. CPUE Standardisation of Swordfish caught in South African waters, IOTC-2025-WPB23-22. 23th Session of the IOTC Working Party on Billfish – Sète, France.

Guyomard, D., Desruisseaux, M., Poisson, F., Petit, M., Taquet, M., 2004. GAM analysis of operational and environmental factors affecting swordfish (Xiphias gladius) catch and CPUE of the Reunion Island longline fishery, in the South Western Indian Ocean, IOTC-2004-WPB-04. 4th Session of the IOTC Working Party on Billfish – Albion Fisheries Research Centre, Mauritius.

Hartig, F., 2020. Package "DHARMa."

Maunder, M.N., Punt, A.E., 2004. Standardizing catch and effort data: a review of recent approaches. Fisheries Research 70, 141–159. https://doi.org/10.1016/j.fishres.2004.08.002

Poisson François, Gaertner Jean-Claude, Taquet Marc, Durbec Jean-Pierre, Bigelow Keith, 2010. Effects of lunar cycle and fishing operations on longline-caught pelagic fish: fishing performance, capture time, and survival of fish. Fishery Bulletin. 108 (3). 268-281. https://archimer.ifremer.fr/doc/00011/12237/

Sabarros, P.S., Romanov, E., Dagorne, D., Foulgoc, L.L., Ternon, J.-F., Bach, P., 2014. Environmental drivers of swordfish local abundance in the south-west Indian Ocean, IOTC-2014-WPB12-15. 12th Session of the IOTC Working Party on Billfish – Yokohama, Japan.

Sabarros, P.S., Coelho, R., Bach, P., 2017. Standardized CPUE of blue shark caught by the French swordfish fishery in the south-west Indian Ocean (2007-2016), IOTC-2017-WPEB13-27. 13th Session of the IOTC Working Party on Ecosystems and Bycatch – San Sabastián, Spain.

Sabarros, P.S., Coelho, R., Romanov, E.V., Guillon, N., Bach, P., 2021. Updated standardized CPUE of blue shark bycaught by the French Reunion-based pelagic longline fishery (2007-2020), IOTC-2021-WPEB17(DP)-08_Rev1. 17th Session of the IOTC Working Party on Ecosystems and Bycatch – Data Preparatory, Online.

Sabarros, P.S., Tellier, C., Bach, P., Coelho, R., Romanov, E.V., 2025. Updated standardized CPUE of blue shark bycaught by the French Reunion-based pelagic longline fishery (2007-2024), IOTC-2025-WPEB21(DP)-05. 21th Session of the IOTC Working Party on Ecosystems and Bycatch – Data Preparatory, Online.

Tellier, C., Sabarros, P.S., 2025. Standardized CPUE of oceanic whitetip shark bycaught by the French Reunion-based pelagic longline fishery operating in the South West Indian Ocean (2007-2024), IOTC-2025-WPEB21(AS)-18. 21th Session of the IOTC Working Party on Ecosystems and Bycatch – Sète, France.

Tables

Table 1. Comparison between three models: the best conventional GLMM (Mod 1), the best one with a spatial effect (Mod 2), and the best one with a spatiotemporal effect (Mod 3).

Model	**Covariates selected (in order of selection)**	**Degrees of freedom (for quantitative covariates)**	**Mesh**	**AIC**
Mod 1: **Best GLMM** Random effect: *vessel* N = 5784 Distribution: Negative Binomiale	*month*	11	/	32458.62
	hauling_end_time	3		
	latitude	3		
	longitude	5		
	moon_illumination	3		
	setting_start_time	5		
	branchlines_per_basket_count	2		
	soaking_time	3		
	percentage_circle_hooks	1		
Mod 2: **Best GLMM with a spatial random effect** Random effect: *vessel* N = 5784 Distribution: Negative Binomiale	*month*	11	100 knots	32391.46
	hauling_end_time	3		
	moon_illumination	3		
	setting_start_time	5		
	soaking_time	3		
	branchlines_per_basket_count	2		
	percentage_circle_hooks	1		
Mod 3: **GLMM with a spatiotemporal random effect** Random effect: *vessel* N = 5784 Distribution: Negative Binomiale	*month*	11	100 knots	32194.77
	hauling_end_time	3		
	moon_illumination	3		
	setting_start_time	5		
	soaking_time	3		
	branchlines_per_basket_count	2		
	percentage_circle_hooks	1		

Table 2. Summary of the retained model (Mod 3).

```
Spatiotemporal model fit by ML ['sdmTMB']
Formula:
nb_caught ~ 1 + factor(month) + hauling_end_time_spl1 + hauling_end_time_spl2 +
   hauling_end_time_spl3 + moon_illumination_spl1 + moon_illumination_spl2 +
   moon_illumination_spl3 + setting_start_time_spl1 + setting_start_time_spl2 +
   setting_start_time_spl3 + setting_start_time_spl4 + setting_start_time_spl5 +
   soaking_time_spl1 + soaking_time_spl2 + soaking_time_spl3 +
   branchlines_per_basket_count_spl1 + branchlines_per_basket_count_spl2 +
   percentage_circle_hooks_spl1 + year + (1 | vessel)
Mesh: mesh (isotropic covariance)
Time column: character
Data: data
Family: nbinom2(link = 'log')

Random intercepts and/or slopes:

Conditional model:
    Groups      Name    Variance   Std.Dev.
    vessel (Intercept)      0.04       0.21

Conditional model:
                          coef.est coef.se
(Intercept)                   -4.90    0.42
factor(month)2                -0.08    0.04
factor(month)3                -0.09    0.04
factor(month)4                -0.23    0.05
factor(month)5                -0.44    0.04
factor(month)6                -0.54    0.04
factor(month)7                -0.44    0.04
factor(month)8                -0.28    0.05
factor(month)9                -0.09    0.04
factor(month)10               -0.09    0.04
factor(month)11               -0.03    0.04
factor(month)12                0.04    0.04
hauling_end_time_spl1          0.26    0.16
hauling_end_time_spl2         -0.55    0.46
hauling_end_time_spl3         -0.80    0.21
moon_illumination_spl1         0.21    0.04
moon_illumination_spl2         0.10    0.06
moon_illumination_spl3        -0.02    0.02
setting_start_time_spl1        0.71    0.22
setting_start_time_spl2        0.70    0.23
setting_start_time_spl3        0.89    0.17
setting_start_time_spl4        0.70    0.48
setting_start_time_spl5        0.44    0.24
soaking_time_spl1              1.05    0.26
soaking_time_spl2             -0.62    0.64
soaking_time_spl3              0.20    0.22
branchlines_per_basket_count_spl1   -0.92    0.21
```

```
branchlines_per_basket_count_spl2   -0.05   0.16
percentage_circle_hooks_spl1       -0.09   0.05
year2008                     -0.13   0.25
year2009                     -0.28   0.25
year2010                      0.36   0.26
year2011                     -0.06   0.22
year2012                     -0.18   0.22
year2013                     -0.23   0.22
year2014                     -0.34   0.22
year2015                     -0.16   0.22
year2016                     -0.04   0.22
year2017                      0.01   0.23
year2018                     -0.10   0.22
year2019                     -0.01   0.23
year2020                      0.04   0.23
year2021                      0.10   0.23
year2022                      0.18   0.23
year2023                      0.19   0.22
year2024                      0.29   0.23
year2025                      0.21   0.22

Dispersion parameter: 6.48
Spatiotemporal AR1 correlation (rho): 0.56
Matérn range: 4.03
Spatial SD: 0.25
Spatiotemporal marginal AR1 SD: 0.24
ML criterion at convergence: 16044.384

See ?tidy.sdmTMB to extract these values as a data frame.
```

Table 3. Standardized CPUE (stdCPUE) time series for swordfish caught in the French longline fishery for the period 2007-2025. nCPUE designates the nominal CPUE. The stdCPUE is provided with 95% confidence interval (CI).

Year	nCPUE	stdCPUE	Lower CI	Upper CI
2007	4.9750	**7.1000**	4.5980	10.9634
2008	7.6667	**6.2545**	3.9549	9.8909
2009	6.3829	**5.3415**	3.6211	7.8793
2010	13.6694	**10.2245**	6.8096	15.3521
2011	7.0905	**6.6910**	4.9339	9.0737
2012	5.3682	**5.9402**	4.3932	8.0318
2013	4.7521	**5.6310**	4.1334	7.6710
2014	4.2576	**5.0577**	3.7160	6.8838
2015	5.0111	**6.0241**	4.4585	8.1395
2016	5.9489	**6.8233**	5.1093	9.1124
2017	5.1648	**7.1914**	5.3271	9.7081
2018	4.4568	**6.4183**	4.7922	8.5962
2019	5.2494	**7.0122**	5.1882	9.4775
2020	6.3366	**7.4218**	5.4704	10.0694
2021	7.0798	**7.8811**	5.7911	10.7255
2022	8.7639	**8.4859**	6.2796	11.4674
2023	9.3295	**8.6046**	6.4383	11.4998
2024	8.2596	**9.4795**	6.9877	12.8598
2025	7.4010	**8.7184**	6.5585	11.5897

Figures

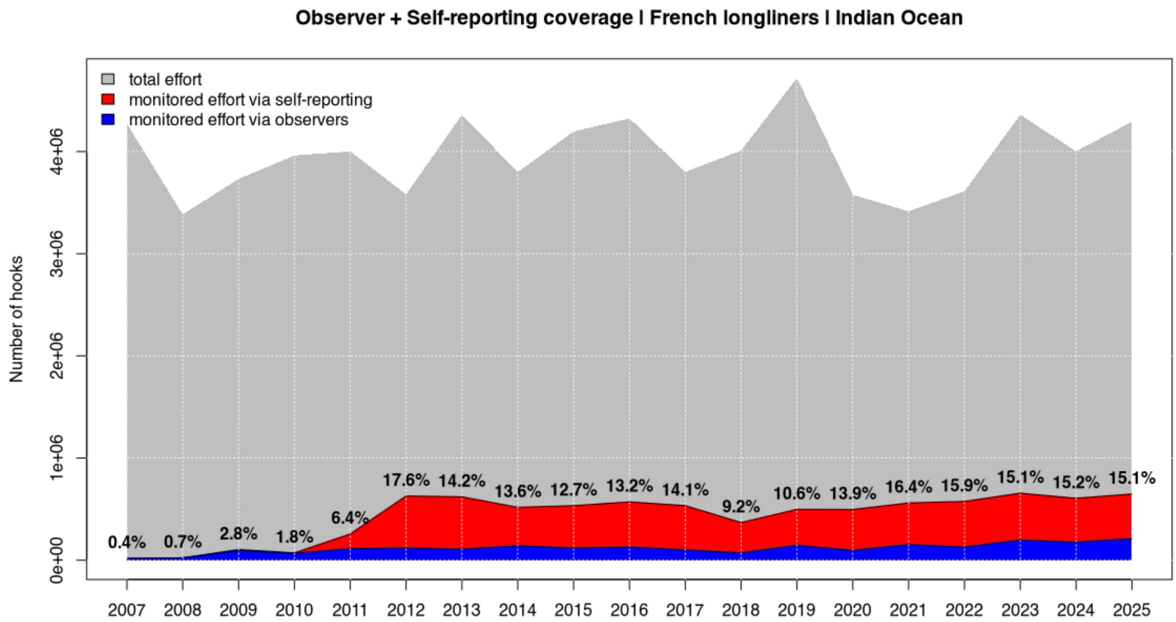

Figure 1. Observer and self-reporting effort coverage in number of hooks deployed in the French longline fishery operating in the south-west Indian Ocean between 2007 and 2025.

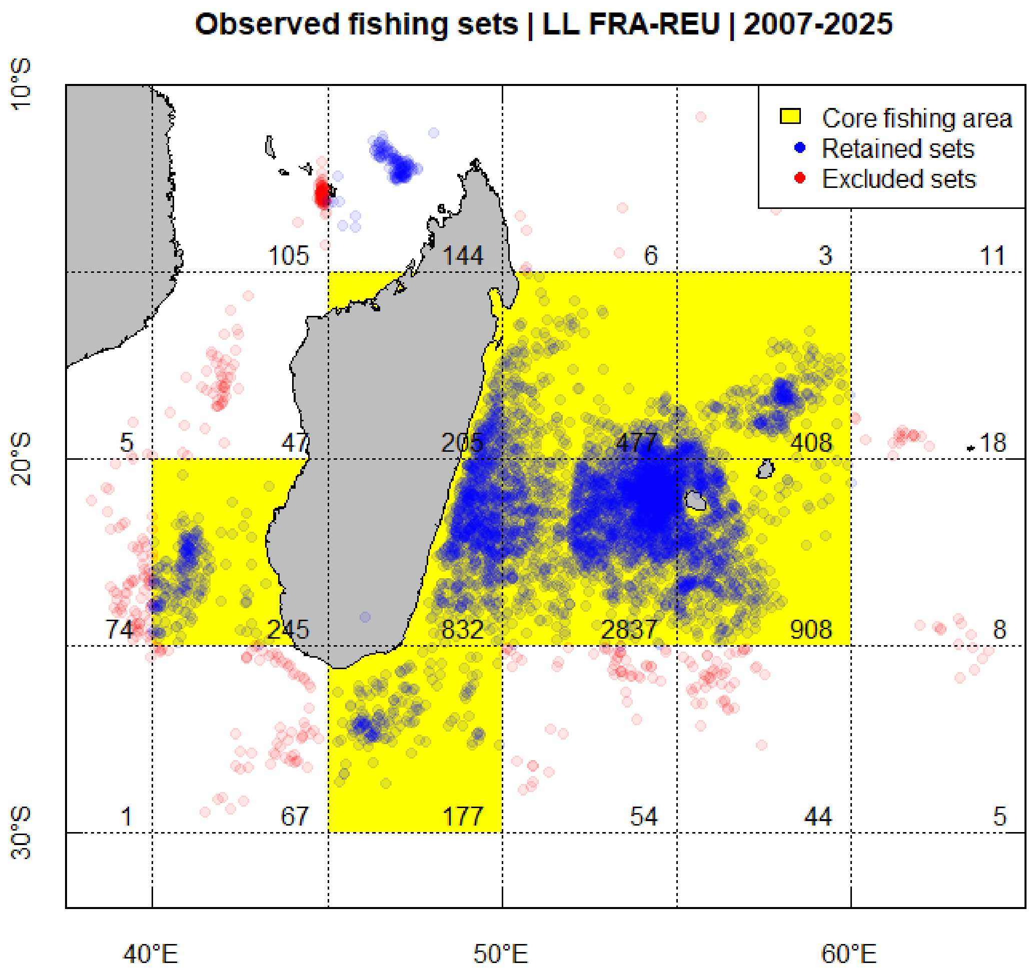

Figure 2. Distribution of fishing sets (hauling start position) between 2007 and 2025. The yellow area represents the core fishing area with retained sets in blue. Excluded sets are shown in red. Numbers in the corners of 5°x5° squares are the number of sets.

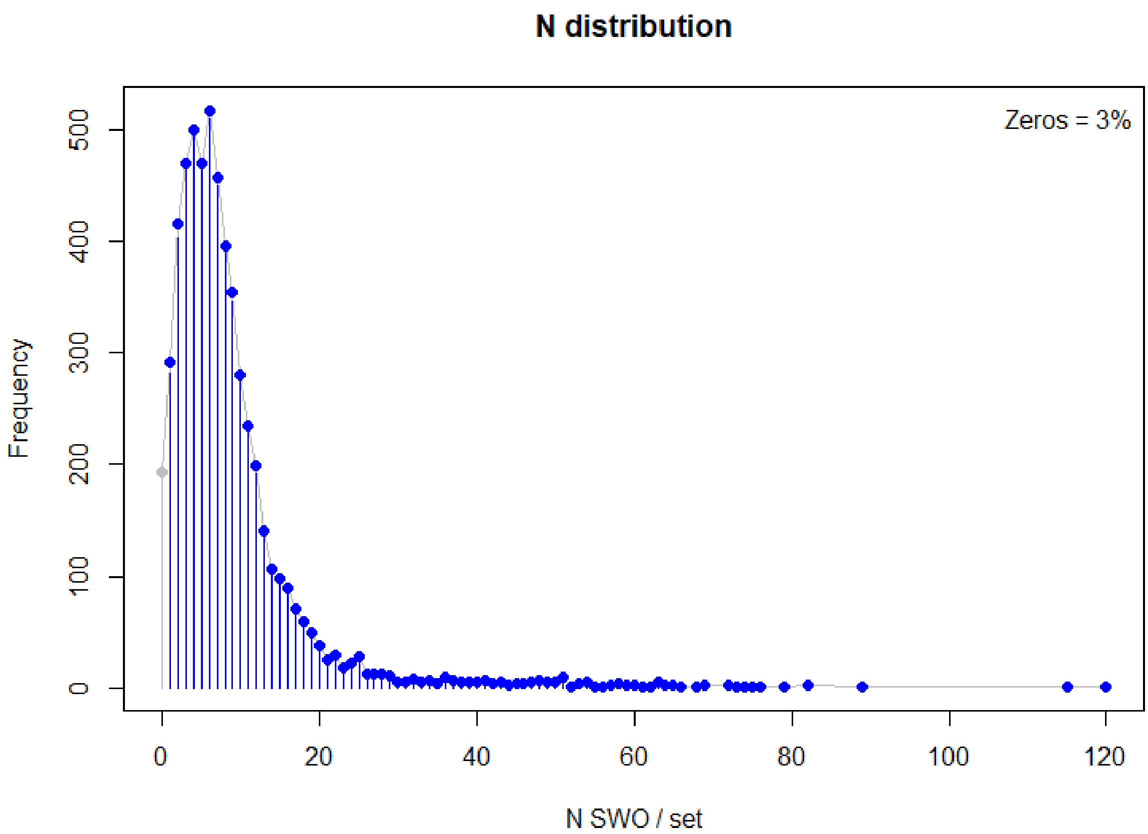

Figure 3. Frequency distribution of swordfish catches per set, showing a low proportion of zeros (3%).

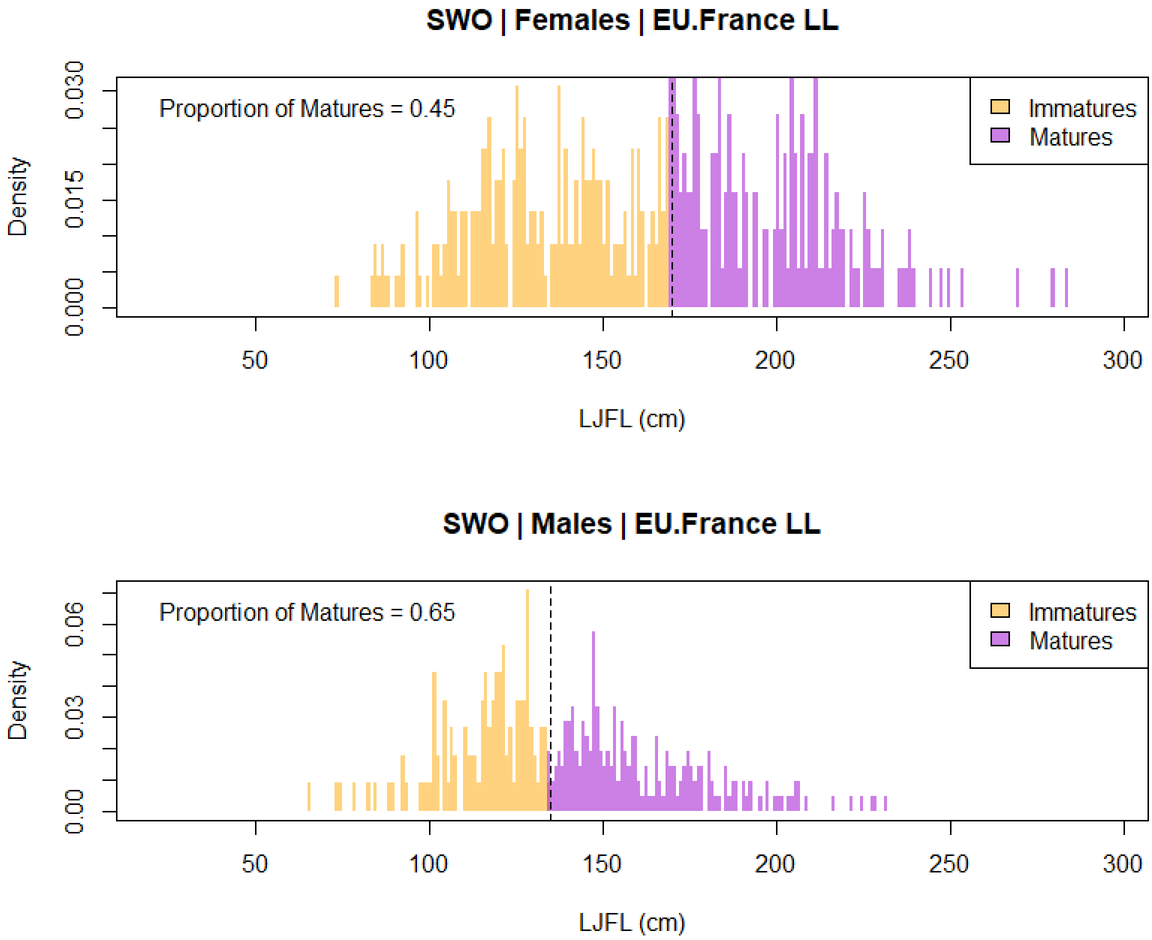

Figure 4. Length-frequency distribution of swordfish by sex, showing immature and mature individuals relative to maturity thresholds (170 cm LJFL for females, 135 cm LJFL for males).

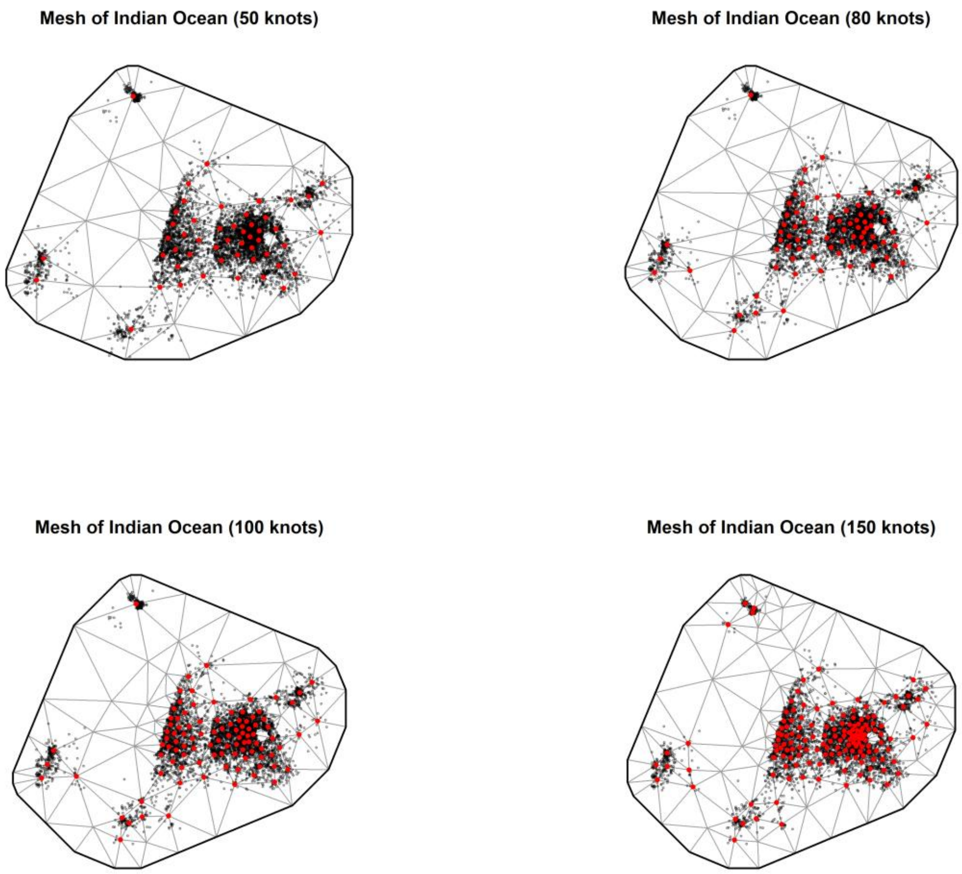

Figure 5. Spatial meshes (50, 80, 100, and 150 knots) tested for GLMMs with a spatial random field.

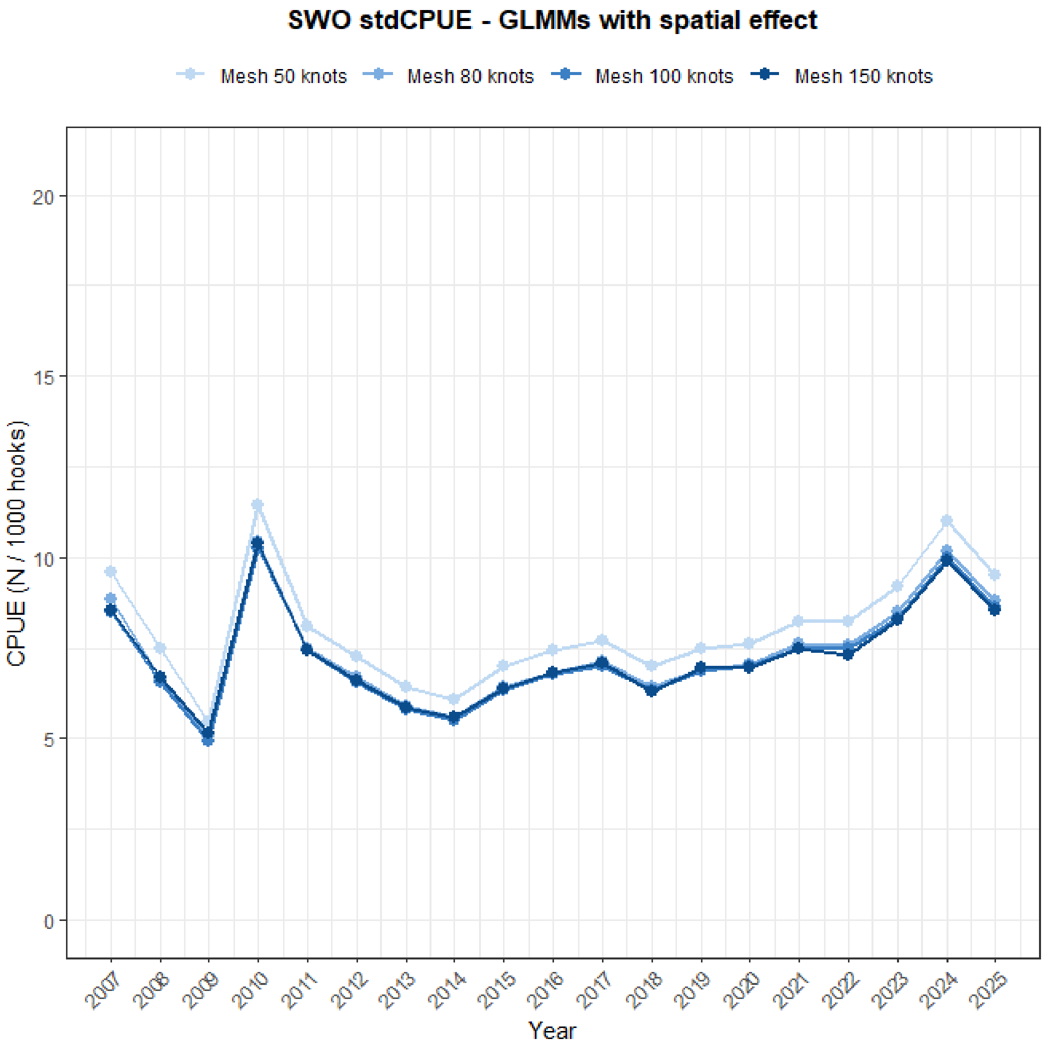

Figure 6. Swordfish standardized CPUE time series, estimated from GLMMs with a spatial random effect, using four mesh resolutions: 50, 80, 100, and 150 knots.

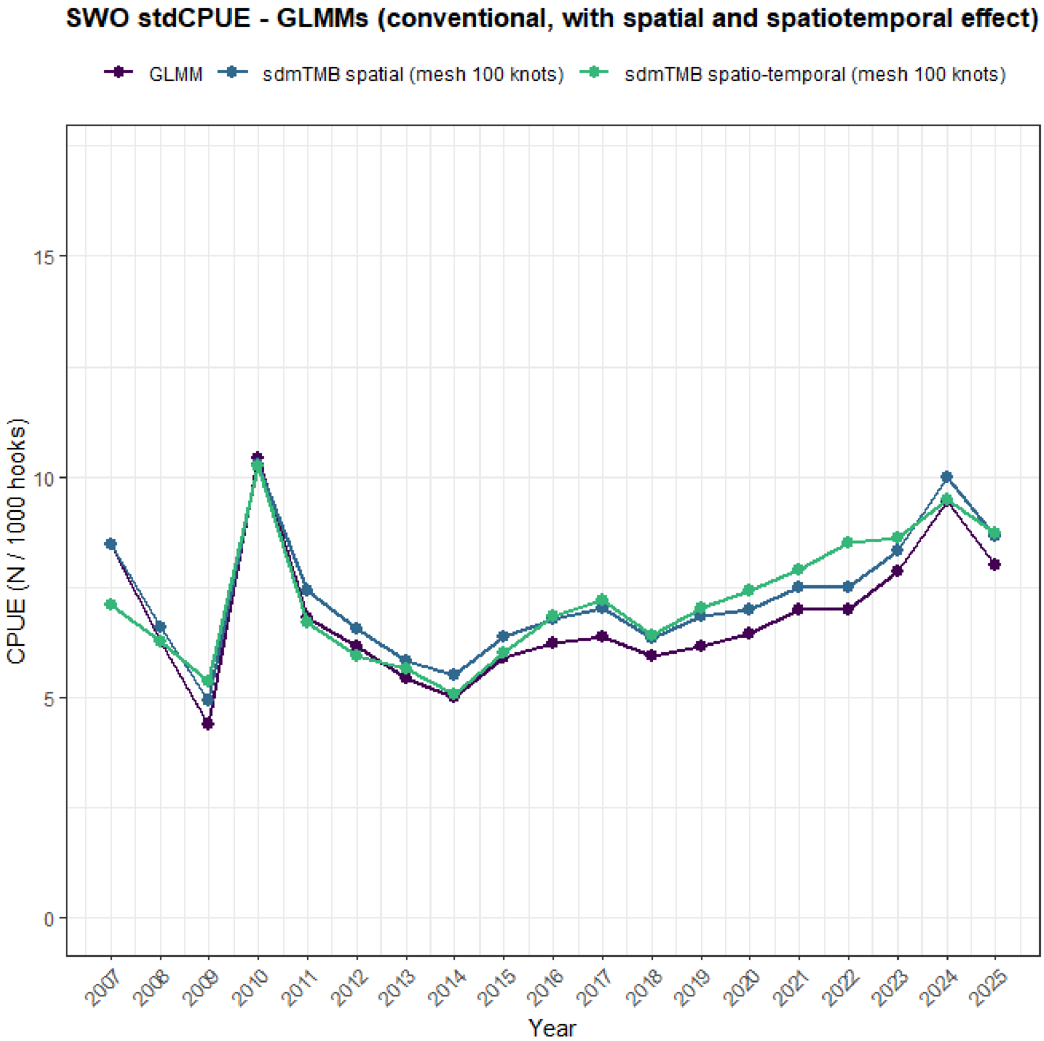

Figure 7. Swordfish standardized CPUE time series, estimated with the best conventional GLMM, the best GLMM with a spatial random field, and the GLMM with a spatiotemporal random field.

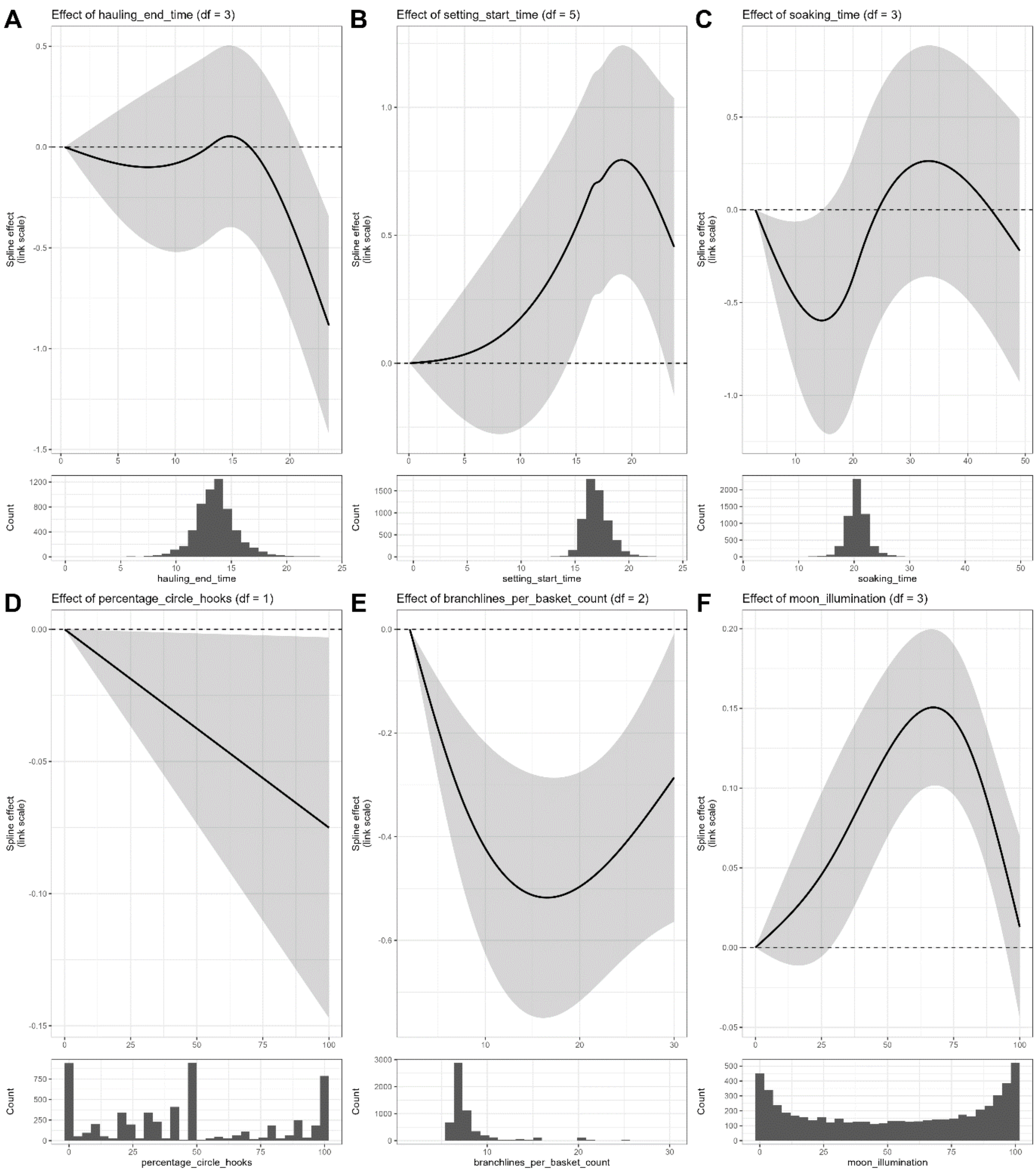

Figure 8. Partial effects (splines) of the continuous covariates used to explain CPUE in the retained model (Mod 3): *hauling_end_time* (A), *setting_start_time* (B), *soaking_time* (C), *percentage_circle_hooks* (D), *branchlines_per_basket_count* (E) and *moon_illumination* (F).

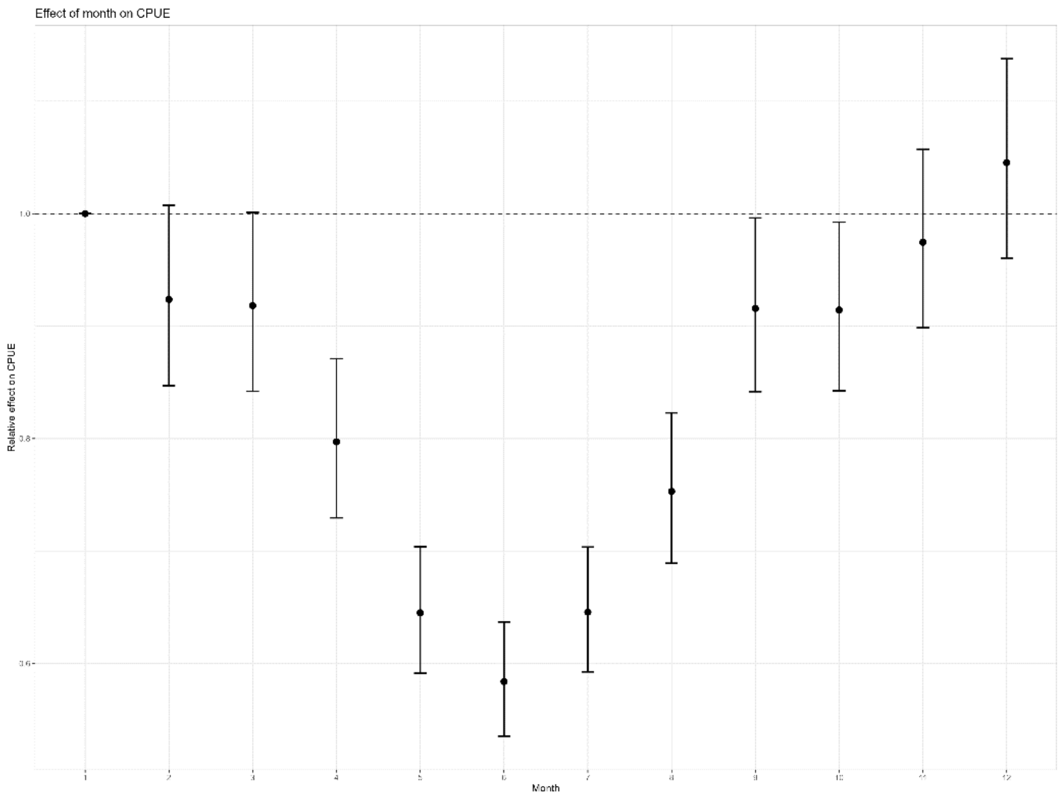

Figure 9. Effect of the factor covariate month used to explain CPUE in the retained Negative Binomial GAMM (Mod 3).

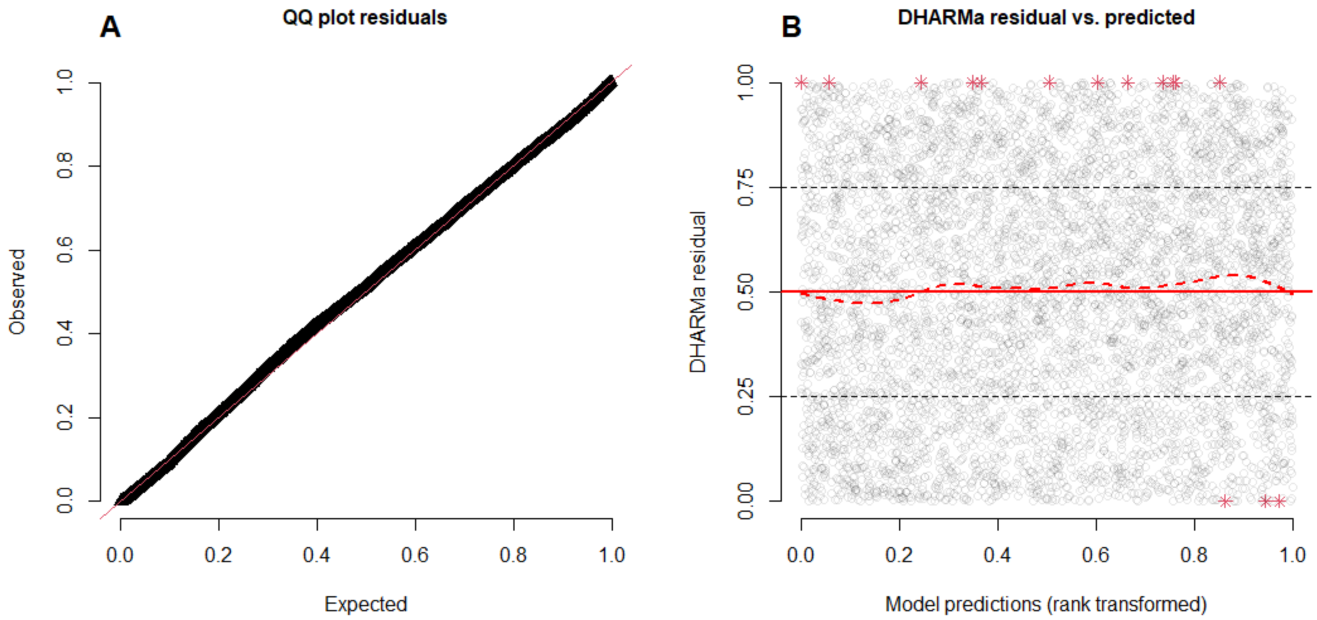

Figure 10. Residual analysis of model selected for swordfish CPUE standardization: QQ plot of the simulated scaled residuals (A) and residuals plotted against fitted values (B).

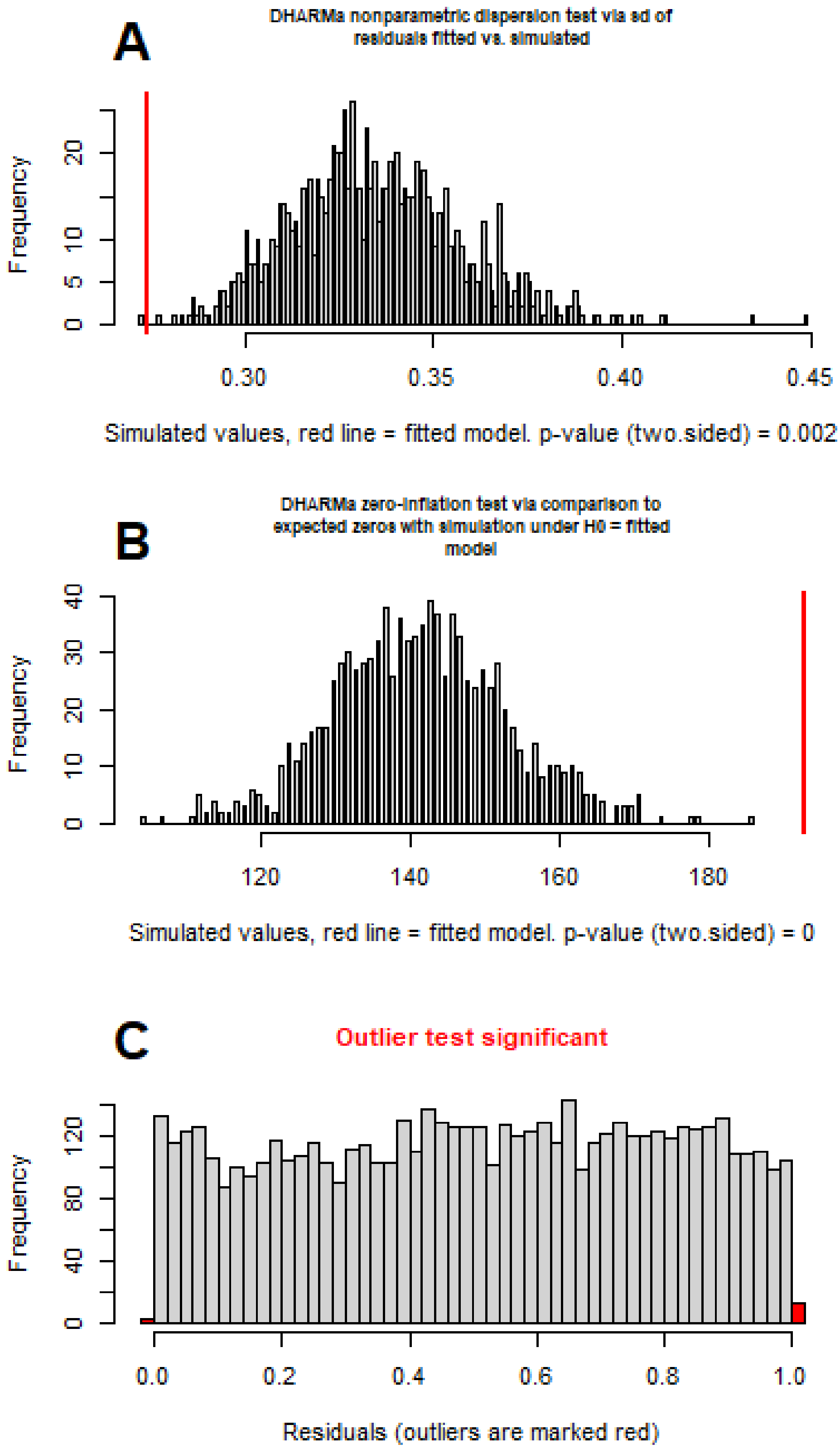

Figure 11. Residual analysis of model selected for swordfish CPUE standardization: dispersion test (A), zero-inflation test (B) and outliers test (C).

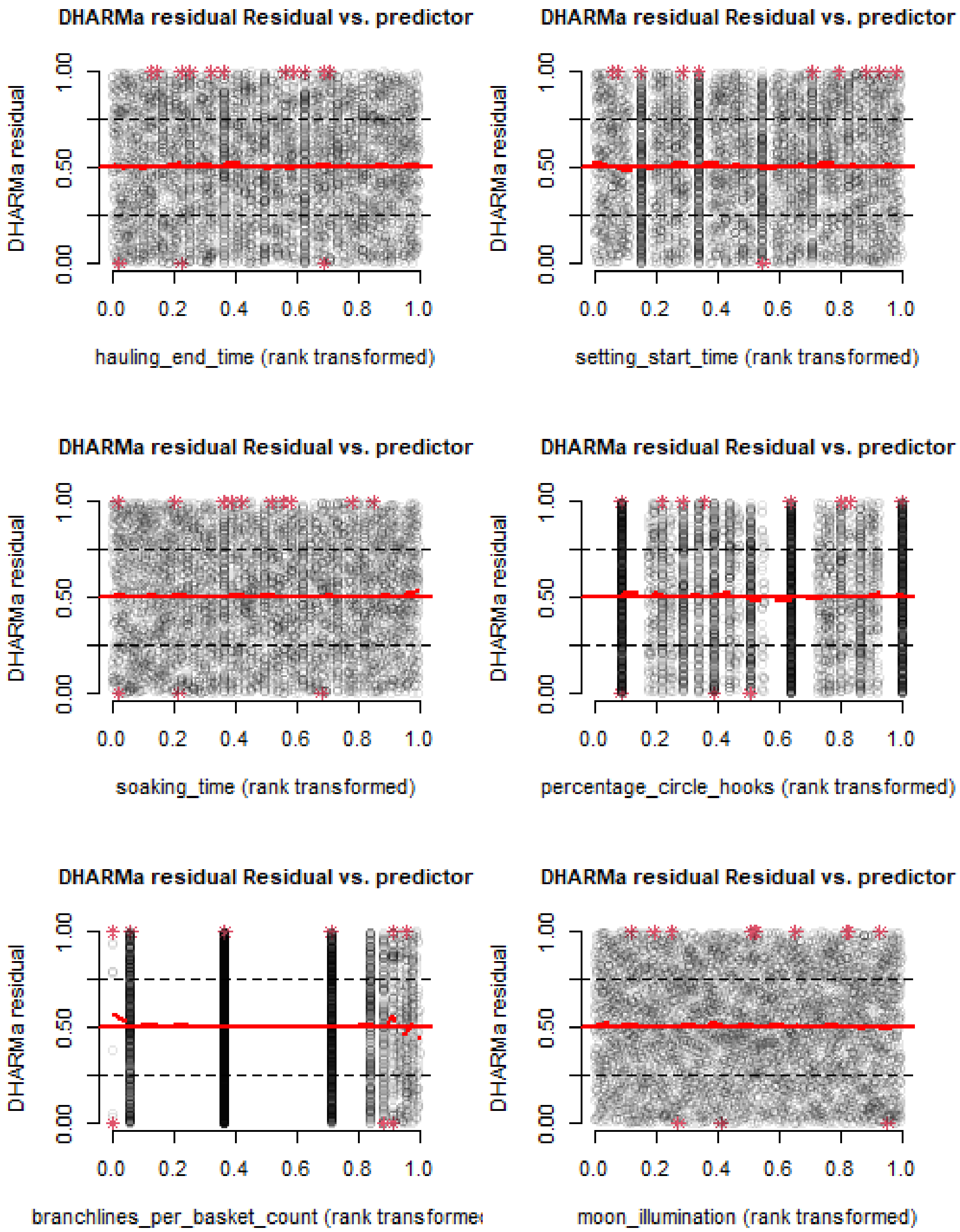

Figure 12. Residuals of model selected for swordfish CPUE standardization (Mod 3), for each quantitative covariate selected.

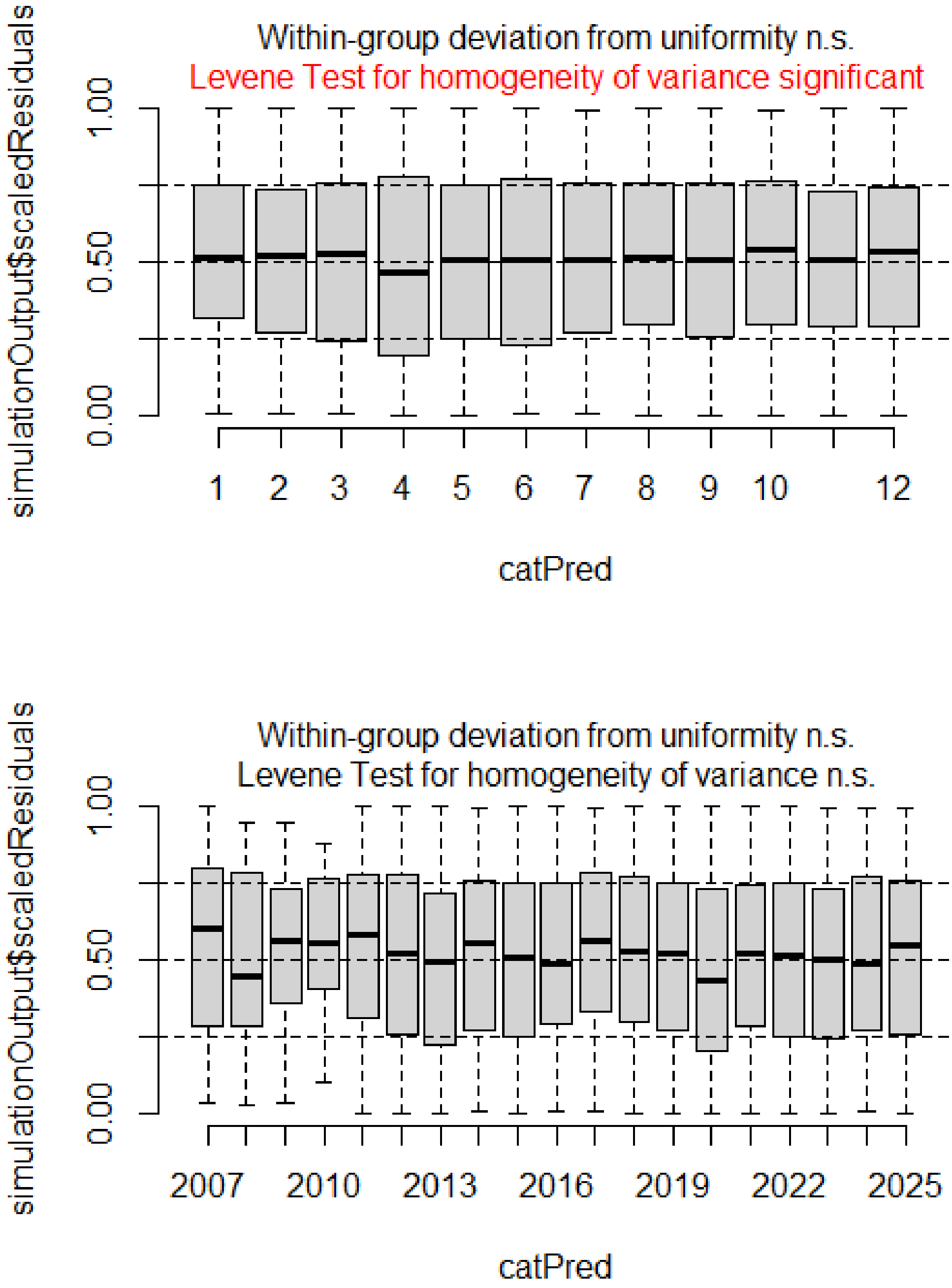

Figure 13. Residuals of model selected for swordfish CPUE standardization (Mod 3), for covariates *month* and *year*.

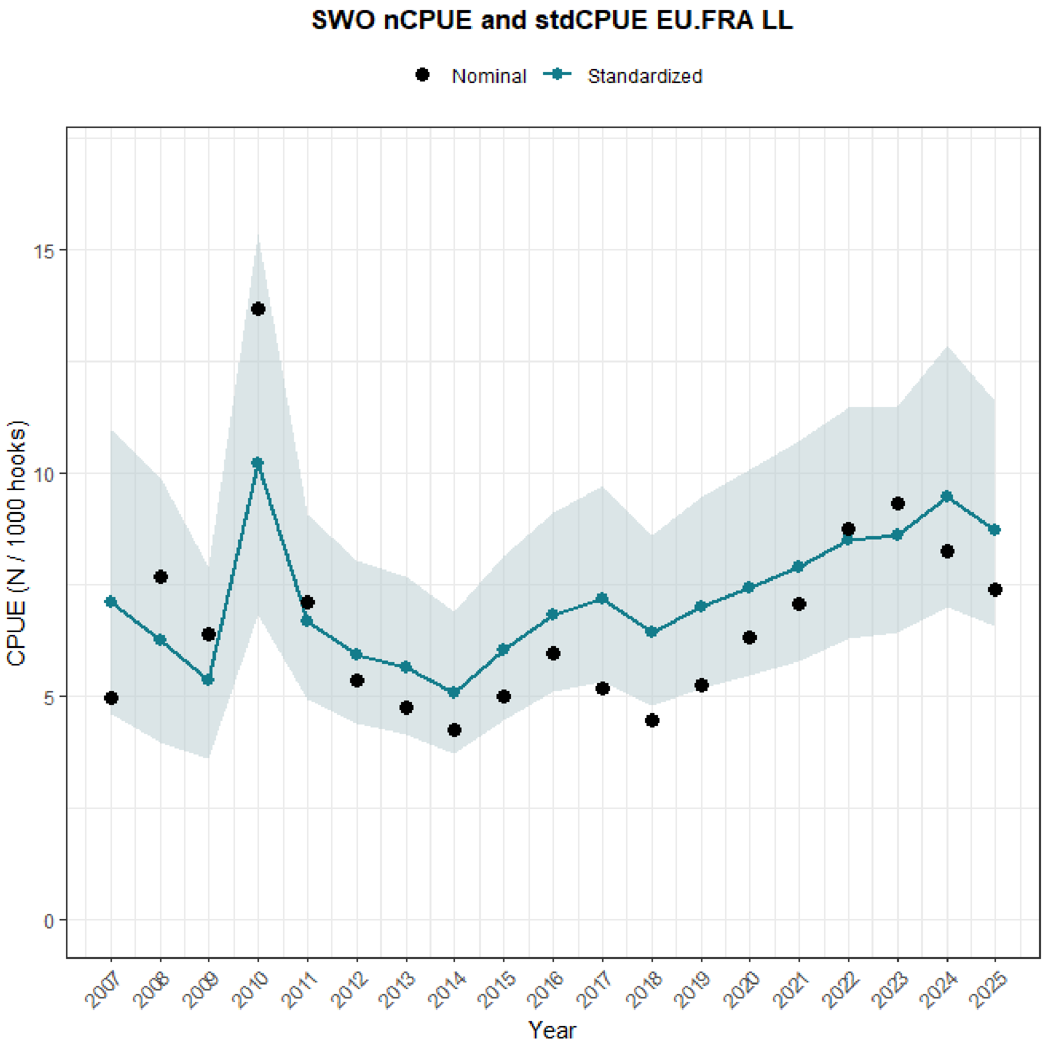

Figure 14. Nominal and standardized CPUE (N/1000 hooks) time series for the swordfish (SWO) caught by the French longline fishery based in Reunion Island (EU.FRA LL) for the period 2007-2025. The standardized index is derived from the final GLMM with a spatiotemporal random field, fitted with *sdmTMB*. The gray envelope represents the 95% confidence intervals.

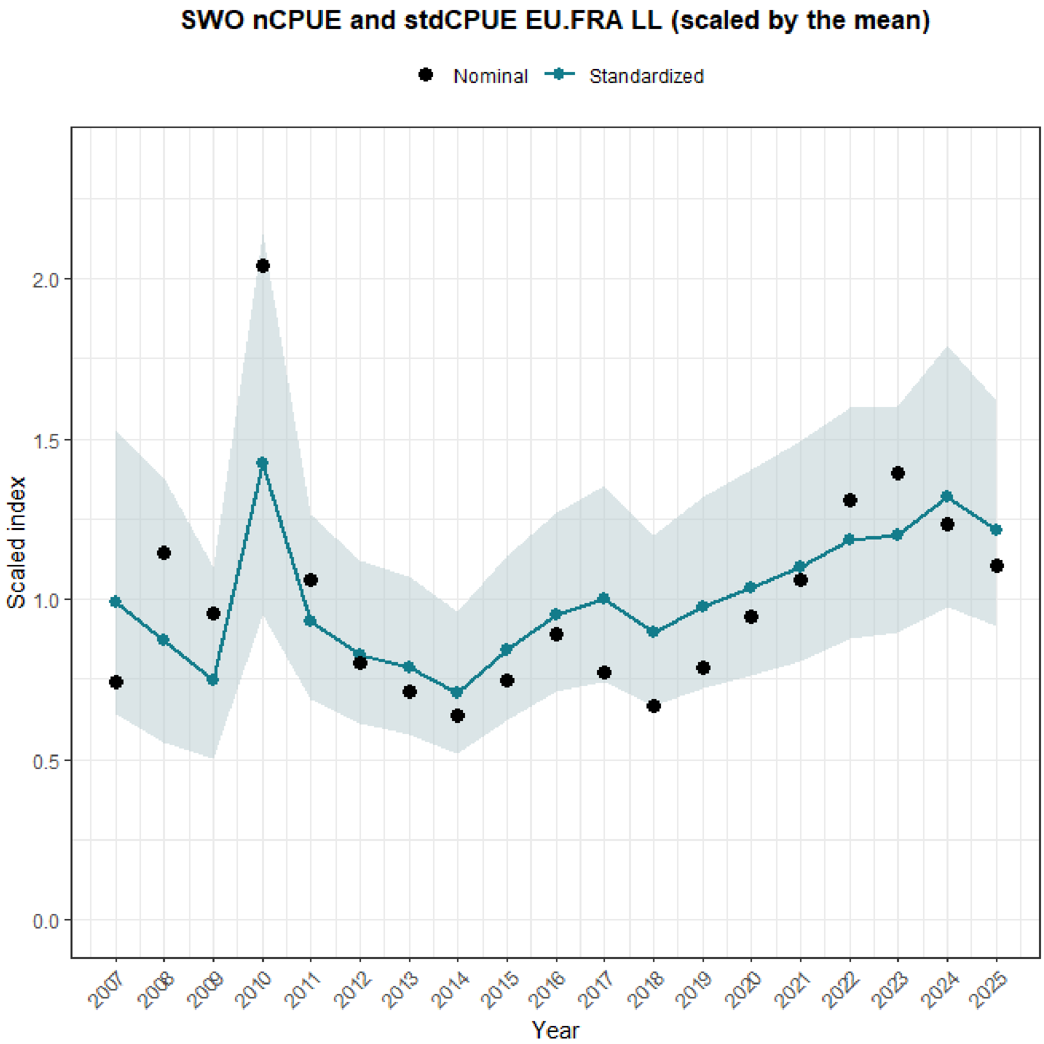

Figure 15. Scaled (by the mean) nominal and standardized CPUE time series for the swordfish (SWO) caught by the French longline fishery based in Reunion Island (EU.FRA LL) for the period 2007-2025. The standardized index is derived from the final GLMM with a spatiotemporal random field, fitted with *sdmTMB*. The gray envelope represents the 95% confidence intervals.