Standardized CPUE of oceanic whitetip shark bycaught by the French Reunion-based pelagic longline fishery operating in the South West Indian Ocean (2007-2024)

Chloé Tellier[1,2], Philippe S. Sabarros[1,2,*]

[1]MARBEC, Univ Montpellier, CNRS, Ifremer, IRD, Sète, France
[2]Institut de Recherche pour le Développement (IRD), Observatoire des Écosystèmes Pélagiques Tropicaux Exploités (Ob7), Sète, France
*corresponding author: philippe.sabarros@ird.fr

Abstract

The oceanic whitetip shark *Carcharhinus longimanus* is a relatively common bycatch of the French swordfish-targeting longline fishery operating in the south-west Indian Ocean. Using observer and self-reported data collected aboard commercial longliners between 2007 and 2024, we present a standardized CPUE series for oceanic whitetip shark. The index was estimated using a generalized additive mixed model (GAMM) with a negative binomial distribution, which appropriately handled the high proportion of zero catches in the data. For the upcoming stock assessment, we recommend using the standardized CPUE for the period comprised between 2011 and 2024 where the monitoring effort has been consequent in comparison with previous years. Throughout this period, the standardized CPUE for the oceanic whitetip shark shows a slight but significant increasing trend.

Keywords

Oceanic whitetip shark | CPUE standardization | Longline | Western Indian Ocean

1. Introduction

Primary indices of abundance of target (e.g. tunas) and non-target species (e.g. sharks) are based on catch and effort data from commercial fisheries in the absence of fishery-independent abundance indicators. Fishery-based indices need to be standardized to remove the influence of various fishery-dependent factors such as the fishing effort variability, fishing strategy, habitat overlap, etc., so they can be used for stock assessment (Maunder and Punt, 2004).

The French longline fishery based in Reunion Island operates in the south-west Indian Ocean around Reunion Island, Mauritius, and Madagascar and mainly targets swordfish (*Xiphias gladius*) with relatively shallow night sets. The oceanic whitetip shark *Carcharhinus longimanus* is the third most common bycatch species of shark and represents about 2.4% of the total bycatch in number of individuals caught.

In this paper, we provide an index of abundance for this species based on observer and self-reported data of the French swordfish-targeting fishery based in Reunion Island for the period 2007-2024.

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 (2009-2024; Reg 199/2008 and 665/2008). The coverage in number of hooks monitored is presented in Figure 1. We retained a total of 5214 fishing operations monitored between 2007 and 2024 from the core fishing area that consists of 5°x5° squares with more than 120 fishing operations (Figure 2).

2.2. CPUE standardization

The response variable modeled was the number of sharks caught (*nb_caught*) per fishing operation. The fishing effort, defined as the number of hooks deployed (*total_hooks_count*), was included in the model as a log-transformed offset (*offset(log(total_hooks_count))*). This approach allows for the direct modeling of the catch per unit of effort (CPUE) using a distribution adapted to count data.

The proportion of zeros in the data is important (79%) and the catch distribution highly skewed, which justifies using models adapted for overdispersed count data. We estimated the standardized CPUE with a Generalized Additive Mixed Model (GAMM) using both the *glmmTMB* and the *gam* functions, respectively from *glmmTMB* and *mgcv* R packages (Brooks et al., 2017; Wood, 2017). We compared several approaches, including Poisson (P), Zero-Inflated Poisson (ZIP), Negative Binomial (NB) and Zero-Inflated Negative Binomial (ZINB) models.

The list of candidate covariates was determined based on previous work of CPUE standardization for oceanic whitetip shark and standardization for blue shark carried out with the same dataset (Sabarros et al., 2017; Sabarros et al., 2021; Sabarros et al., 2025). The potential non-linearity of continuous covariates was checked by performing univariate GAM models.

Candidate covariates were:

- Fixed effects:
    - *year* (factor): 2007 to 2024
    - *quarter* (factor): Q1 to Q4
    - *longitude* (continuous): longitude of the fishing operation, specifically the latitude where the line starts being retrieved (hauling).
    - *latitude* (continuous): latitude of the fishing operation, specifically the longitude where the line starts being retrieved (hauling).
    - *cwp55* (factor): 5°x5° square of the fishing operation, specifically the square where the line starts being retrieved (hauling).
    - *quarter:cwp55* (factor): interaction between quarter and 5°x5° square.
    - *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.
    - *hooks_per_basket* (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.).
- Random effects:
    - *vessel* (factor): the vessel name was 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.

We applied a forward stepwise model selection procedure, adding covariates sequentially and comparing models using the Akaike Information Criterion (AIC) to select for relevant and significant covariates.

The comparison between different P, ZIP, NB and ZINB approaches showed that ZIP and NB distributions had better results than Pand ZINB distributions. The best model with a ZIP distribution is called Mod 1; the best with a NB distribution is called Mod 2. Among those two, the best one, exhibiting the lowest AIC, is Mod 2, and is the one retained to estimate the CPUE.

The deviance tables (Type III ANOVA with Chi Square test) of Mod 1 and Mod 2 are provided in Table 1. The summary table of the retained model, Mod 2, is in Table 2, and the graphical analysis of residuals is presented in Figure 6. Finally, we present the yearly standardized CPUE series from the retained model computed (Table 3; Figure 6) as well as the scaled (by the mean) standardized CPUEs series (Figure 7).

3. Results

A NB distribution was selected as it provided the best fit for the data (lowest AIC score). The final retained model from the forward selection procedure was the following:

Mod 2: *nb_caught ~ year + quarter + s(longitude) + s(latitude) + quarter:cwp55 +* s*(hooks_per_basket) + s(percentage_circle_hooks) + s(percentage_squid_bait) + (1|vessel) + offset(log(total_hooks_count))*

The deviance analysis (Type III Wald Chi square test) of the final model indicates that *year* ($p < 2e-16$) and *quarter:cwp55* ($p < 2e-16$) have a significant effect on oceanic whitetip shark catch rates (Table 1). The smooth terms for *latitude* ($p = 0.000192$), *hooks_per_basket* ($p = 0.019725$), *percentage_circle_hooks* ($p = 0.026269$) and *percentage_squid_bait* ($p < 1.13e-05$) were also significant. However, the effects of *quarter* ($p = 0.25$) and the smooth term for *longitude* ($p = 0.322554$) were not statistically significant in the final model. These covariates were nevertheless retained in the final model as the forward selection procedure indicated their inclusion improved the model's overall fit based on AIC.

The random effect for *vessel* had a significative p-value ($p = 0.006788$), indicating considerable variability in catch rates among vessels. The dispersion parameter for the negative binomial family was 1.317, confirming significant overdispersion in the data.

Weights (Chi Square) of the model showed that *year* ($\chi^2 = 132.727$) and *quarter:cwp55* ($\chi^2 = 187.699$) were the most influential factors explaining the variability in CPUE. The smooth term for *percentage_squid_bait* ($\chi^2 = 32.619$) and the random effect of *vessel* ($\chi^2 = 20.678$) also had a relatively important explanatory power. The smooth terms for *latitude* ($\chi^2 = 16.952$), *hooks_per_basket* (χ^2 = 11.285) and *percentage_circle_hooks* (χ^2 = 7.608) had significant but lesser effects.

The standardized CPUE series is presented in Figure 1 and Table 3. Acknowledging the relatively low coverage rate in the first years of the program, we focus on the period from 2011 to 2024. The standardized CPUE series exhibits a period of fluctuation at low levels between 2011 and 2018, and then a slight increasing trend since 2018.

Overall, the retained standardized CPUE follows the nominal CPUE.

Throughout 2011-2024, the overall trend of the standardized CPUE is weak but significantly increasing (linear regression: b = 0.008656, p-value = 0.0231). Between 2011 and 2018, the evolution is not significant (b = -0.004729, p-value = 0.406). The increase is more pronounced but also not significant between 2018 and 2024 (b = 0.02314, p-value = 0.0731).

The residual analysis of the final model (Figure 6) confirms a good overall fit. The scaled residuals, analyzed using the *DHARMa* R package (Hartig et al., 2024), showed no significant deviation from the expected uniform distribution (Kolmogorov-Smirnov test, p = 0.6179), indicating that the negative binomial model structure is appropriate for the data. No residual patterns were detected against the predicted values, and specific tests confirmed the absence of significant remaining overdispersion (p = 0.236), too much outliers (p = 0.28) or zero-inflation (p = 0.904). Overall, the diagnostics suggest that the model is robust and appropriate for standardizing the CPUE of this species.

4. Discussion

The standardization of the oceanic whitetip shark CPUE series successfully accounted for the influence of several operational and environmental factors. The resulting index is considered a more reliable indicator of abundance trend than the nominal CPUE.

Significant effects on oceanic whitetip shark CPUE

The final model highlighted several key factors influencing catch rates.

The year, latitude and interaction between quarter and 5°x5° square have a significant effect on oceanic whitetip shark catch rates. Year and interaction between quarter and 5°x5° square were identified as the most influential variables. The year effect, which is the primary focus of standardization, showed significant inter-annual variability. The effect of latitude was non-linear (edf = 2.27859), with higher catch rates observed in the northern part of the fishing area (around 15°S). Although quarter and longitude were not significant in the final multivariate model, their inclusion, guided by the forward selection process, helps control for their potential confounding effects.

Fishing gear configuration also plays a role. The number of hooks per basket is a proxy of fishing depth. It had a significant non-linear effect (p = 2.04727), with lower catch rates generally associated with a higher number of hooks, suggesting that deeper sets are less likely to catch oceanic whitetip sharks. In Reunion Island longline fishery, hooks are generally set between 10 meters from the surface down to 120 meters for night fishing (Bach et al., 2014) but deeper sets, likely for targeting tunas during the day (with an increased number of hooks between floats), will reach deeper layers, which results in a lower oceanic whitetip shark CPUE. Such pattern would need to be further investigated.

The percentage of circle hooks had a significant, slightly non-linear effect. Our results indicate that higher proportions of circle hooks are associated with higher catch rates. A plausible explanation is that the feeding behavior and morphology of oceanic whitetip sharks make them more prone to being retained on circle hooks. Moreover, circle hooks may not always lodge directly in the mouth, which could make it more difficult for sharks to escape once hooked.

The percentage of squid bait also had a significant non-linear effect, with catch rates peaking when the proportion of squid was around 30–40%. Overall, catch rates is decreasing as the proportion of squid bait is increasing, suggesting that oceanic whitetip sharks may be less attracted to this type of bait and that its use could potentially reduce their capture.

Finally, the significant random effect of vessel confirms that there are consistent differences in fishing efficiency among vessels that are not explained by the other factors, validating its inclusion in the model to avoid biased estimates.

Relevance of the retained standardized CPUE series

The data considered in this standardization work only concern the core fishing area of the Reunion-based pelagic longline fishery (see Figure 2). This was a safer approach than considering the total dataset that includes scarce sets located in the northern Mozambique Channel and high seas that might exhibit different patterns in terms of oceanic whitetip shark catch rates than those in the core fishing area.

Compared to the nominal CPUE series, the retained standardized CPUE series is smoother but still shows variations over time. The robust diagnostics of the model's residuals provide confidence in the reliability of the standardization and the resulting CPUE series.

Acknowledging the relatively low coverage rate (< 3%) in number of hooks observed in the first years of implementation of the observation programs (Figure 1) we should consider discarding the early part of the standardized time series before 2011, and we recommend focusing on the 2011-2024 period for stock assessment purposes.

The standardized CPUE series exhibits a weak but significant increase throughout 2011-2024 with a rate of 0.9% per year (p = 0.0231), going from approximately 0.17 to 0.30 over this fourteen-year period.

5. Conclusion

According to the assessment of the retained standardization model, we believe that the standardized CPUE time series presented in this paper for oceanic whitetip sharks bycaught by the French pelagic longline fishery of the south-west Indian Ocean is reliable between 2011 and 2024 and can be used for stock assessment. Over this period, the standardized CPUE exhibits a really slight but significant increasing trend, mainly driven by an increase in recent years (2018-2024).

6. Acknowledgments

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

7. References

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.

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

8. Tables

Table 1. Deviance table (Type III ANOVA) of the covariates in GAMM Mod 1 (best model with a Zero-Inflated Poisson distribution) and Mod 2 (best model with a Negative Binomial Distribution). For each covariate, we indicate the degrees of freedom or effective degrees of freedom (Df or edf), the Chi Square test statistic (Chi.sq) and the significance (p-value).

Mod 1: *nb_caught ~ year* + *quarter* + *te(longitude, latitude)* + *quarter:cwp55* + *s(hooks_per_basket)* + *s(percentage_circle_hooks)* + *s(percentage_squid_bait)* + *(1|vessel)* + *offset(log(total_hooks_count))*

Mod 2: *nb_caught ~ year* + *quarter* + *s(longitude)* + *s(latitude)* + *quarter:cwp55* + s*(hooks_per_basket)* + *s(percentage_circle_hooks)* + *s(percentage_squid_bait)* + *(1|vessel)* + *offset(log(total_hooks_count))*

Model	Covariates	Df / edf	Chi.sq	p-value
Mod 1 (best model with a Zero-Inflated Poisson distribution) Random effect: *vessel* N = 5214 AIC = 6746.697	*year*	17	139.928	< 2e-16
	quarter	3	4.442	0.218
	te(longitude, latitude)	5.285	24.313	0.000329
	quarter:cwp55	28	134.691	6.57e-16
	s(hooks_per_basket)	1.938	11.189	0.017294
	s(percentage_circle_hooks)	1.314	7.985	0.027357
	s(percentage_squid_bait)	5.049	54.686	< 2e-16
Mod 2 (best model with a Negative Binomial Distribution) Random effect: *vessel* N = 5214 AIC = 6708.916	*year*	17	132.727	< 2e-16
	quarter	3	4.104	0.25
	s(longitude)	0.02761	0.027	0.322554
	s(latitude)	2.27859	16.952	0.000192
	quarter:cwp55	28	187.699	< 2e-16
	s(hooks_per_basket)	2.04727	11.285	0.019725
	s(percentage_circle_hooks)	1.10567	7.608	0.026269
	s(percentage_squid_bait)	4.51508	32.619	1.13e-05

Table 2. Summary table of the retained lognormal GAMM (Mod 2).

```
Family: Negative Binomial(1.317)
Link function: log

Formula:
nb_caught ~ year + quarter + s(longitude, bs = "cs") + s(latitude, bs = "cs") + quarter:cwp55 +
    s(branchlines_per_basket_count, bs = "cs") + s(percentage_circle_hooks, bs = "cs") +
    s(percentage_squid_bait, bs = "cs") +
 s(vessel, bs = "re") + offset(logtotalhooks)

Parametric coefficients:
                   Estimate Std. Error z value Pr(>|z|)
(Intercept)          -8.91893    0.68645 -12.993  < 2e-16 ***
year2008              0.19393    0.93511   0.207  0.83571
year2009              0.25129    0.62937   0.399  0.68969
year2010              0.78941    0.63881   1.236  0.21655
year2011              0.16435    0.50055   0.328  0.74266
year2012             -0.32604    0.48462  -0.673  0.50110
year2013             -0.79800    0.48978  -1.629  0.10325
year2014             -0.28342    0.48258  -0.587  0.55700
year2015             -0.41077    0.47975  -0.856  0.39188
year2016             -0.16056    0.48154  -0.333  0.73881
year2017             -0.11342    0.48252  -0.235  0.81417
year2018             -0.67094    0.49388  -1.359  0.17430
year2019             -0.01788    0.47943  -0.037  0.97025
year2020              0.29251    0.47816   0.612  0.54071
year2021              0.28405    0.47343   0.600  0.54852
year2022             -0.17448    0.48017  -0.363  0.71633
year2023              0.14689    0.47207   0.311  0.75568
year2024              0.70799    0.47265   1.498  0.13416
quarterQ2             0.56852    0.56814   1.001  0.31699
quarterQ3            -0.57546    0.71885  -0.801  0.42340
quarterQ4             0.34063    0.64926   0.525  0.59983
quarterQ1:cwp556215050  0.58764    0.54120   1.086  0.27757
quarterQ2:cwp556215050 -0.06011    0.32345  -0.186  0.85257
quarterQ3:cwp556215050  1.13431    0.58102   1.952  0.05091 .
quarterQ4:cwp556215050  0.21754    0.45024   0.483  0.62898
quarterQ1:cwp556215055 -0.21090    0.59416  -0.355  0.72263
quarterQ2:cwp556215055  0.22899    0.32644   0.701  0.48300
quarterQ3:cwp556215055  1.36754    0.56464   2.422  0.01544 *
quarterQ4:cwp556215055  0.44545    0.46309   0.962  0.33609
quarterQ1:cwp556220040  1.87600    0.57068   3.287  0.00101 **
quarterQ2:cwp556220040  1.20294    0.38900   3.092  0.00199 **
quarterQ3:cwp556220040  2.65748    0.57414   4.629 3.68e-06 ***
quarterQ4:cwp556220040  1.23016    0.48240   2.550  0.01077 *
quarterQ1:cwp556220045  0.37748    0.54890   0.688  0.49165
quarterQ2:cwp556220045 -0.13901    0.33834  -0.411  0.68118
quarterQ3:cwp556220045  0.01300    0.57381   0.023  0.98192
quarterQ4:cwp556220045  0.30919    0.50111   0.617  0.53723
quarterQ1:cwp556220050  0.48520    0.51311   0.946  0.34435
quarterQ2:cwp556220050 -0.09347    0.31444  -0.297  0.76628
quarterQ3:cwp556220050  0.22663    0.54526   0.416  0.67767
quarterQ4:cwp556220050 -0.18026    0.43987  -0.410  0.68196
```

```
quarterQ1:cwp556220055  0.61802   0.52992   1.166  0.24351
quarterQ2:cwp556220055  0.18508   0.33748   0.548  0.58340
quarterQ3:cwp556220055  0.04740   0.74735   0.063  0.94943
quarterQ4:cwp556220055 -0.17016   0.45632  -0.373  0.70923
quarterQ1:cwp556225045  0.36312   1.00286   0.362  0.71729
quarterQ2:cwp556225045 -0.90483   0.61813  -1.464  0.14324
quarterQ3:cwp556225045 -0.94757   0.84489  -1.122  0.26206
quarterQ4:cwp556225045 -0.42405   0.83602  -0.507  0.61199
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Approximate significance of smooth terms:
                                   edf Ref.df Chi.sq  p-value
s(longitude)                   0.02761      9  0.027 0.322554
s(latitude)                    2.27859      9 16.952 0.000192 ***
s(branchlines_per_basket_count)  2.04727      9 11.285 0.019725 *
s(percentage_circle_hooks)     1.10567      9  7.608 0.026269 *
s(percentage_squid_bait)       4.51508      9 32.619 1.13e-05 ***
s(vessel)                     11.14079     43 20.678 0.006788 **
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

R-sq.(adj) =  0.219   Deviance explained = 24.2%
-REML =   3343  Scale est. = 1        n = 5214
```

Table 3. Standardized CPUE (stdCPUE) time series for oceanic whitetip shark caught in the French longline fishery for the period 2007-2024. nCPUE designates the nominal CPUE. The stdCPUE is provided with 95% confidence interval (CI).

Year	nCPUE	stdCPUE	Lower CI	Upper CI
2007	0.2391	**0.1470**	0.0367	0.5891
2008	0.0928	**0.1784**	0.0267	1.1929
2009	0.1357	**0.1890**	0.0493	0.7247
2010	0.1631	**0.3237**	0.0834	1.2568
2011	0.2547	**0.1732**	0.0571	0.5252
2012	0.1129	**0.1061**	0.0361	0.3113
2013	0.0777	**0.0662**	0.0224	0.1954
2014	0.1560	**0.1107**	0.0391	0.3135
2015	0.1349	**0.0975**	0.0331	0.2868
2016	0.1687	**0.1252**	0.0430	0.3644
2017	0.1582	**0.1312**	0.0459	0.3753
2018	0.1690	**0.0751**	0.0252	0.2236
2019	0.2404	**0.1444**	0.0490	0.4257
2020	0.2920	**0.1969**	0.0673	0.5764
2021	0.3440	**0.1953**	0.0674	0.5657
2022	0.2453	**0.1234**	0.0421	0.3618
2023	0.3935	**0.1702**	0.0588	0.4929
2024	0.5721	**0.2983**	0.1040	0.8557

9. Figures

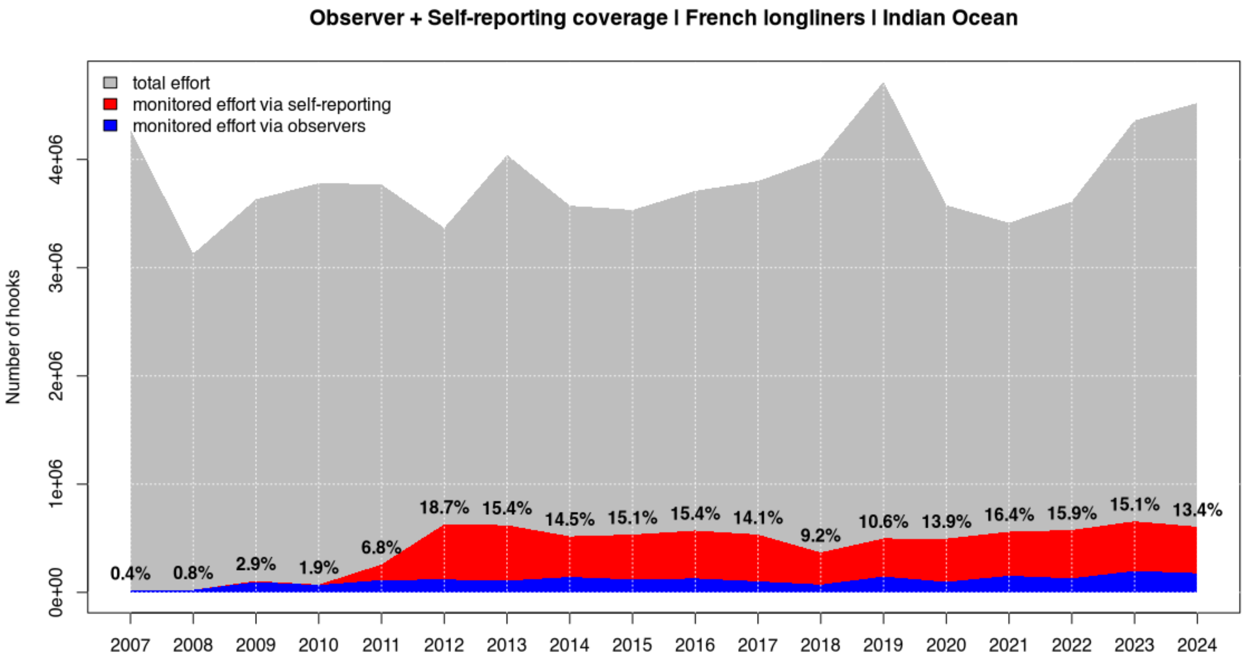

Figure 1. Observer and self-reporting effort coverage in number of hooks deployed in the French longline fishery operating in the south-west Indian Ocean between 2007 and 2024.

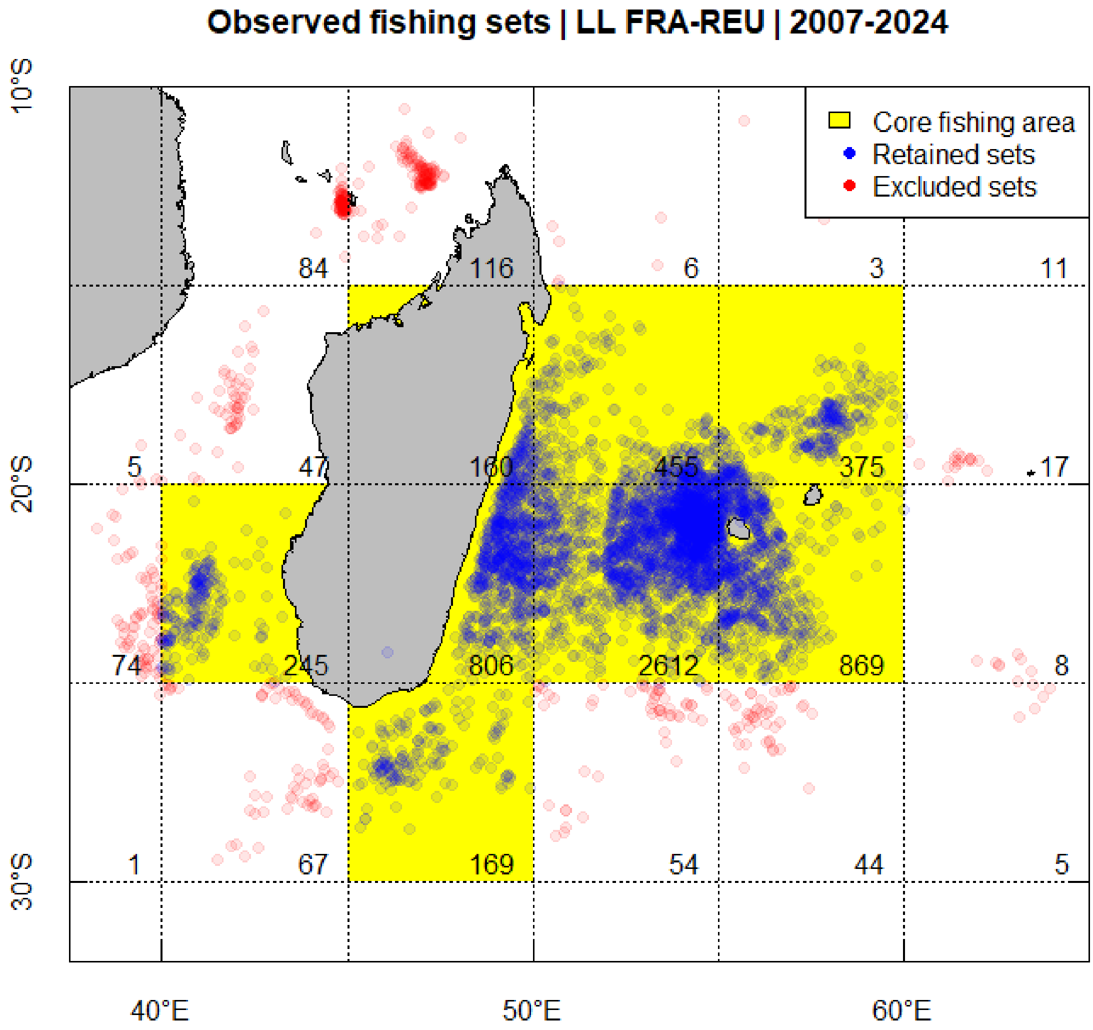

Figure 2. Distribution of fishing sets (hauling start position) between 2007 and 2024. The yellow area represents the core fishing area with retained sets in blue. Excluded sets are shown in red. Numbers in the corners of 5°x5° squares are the number of sets.

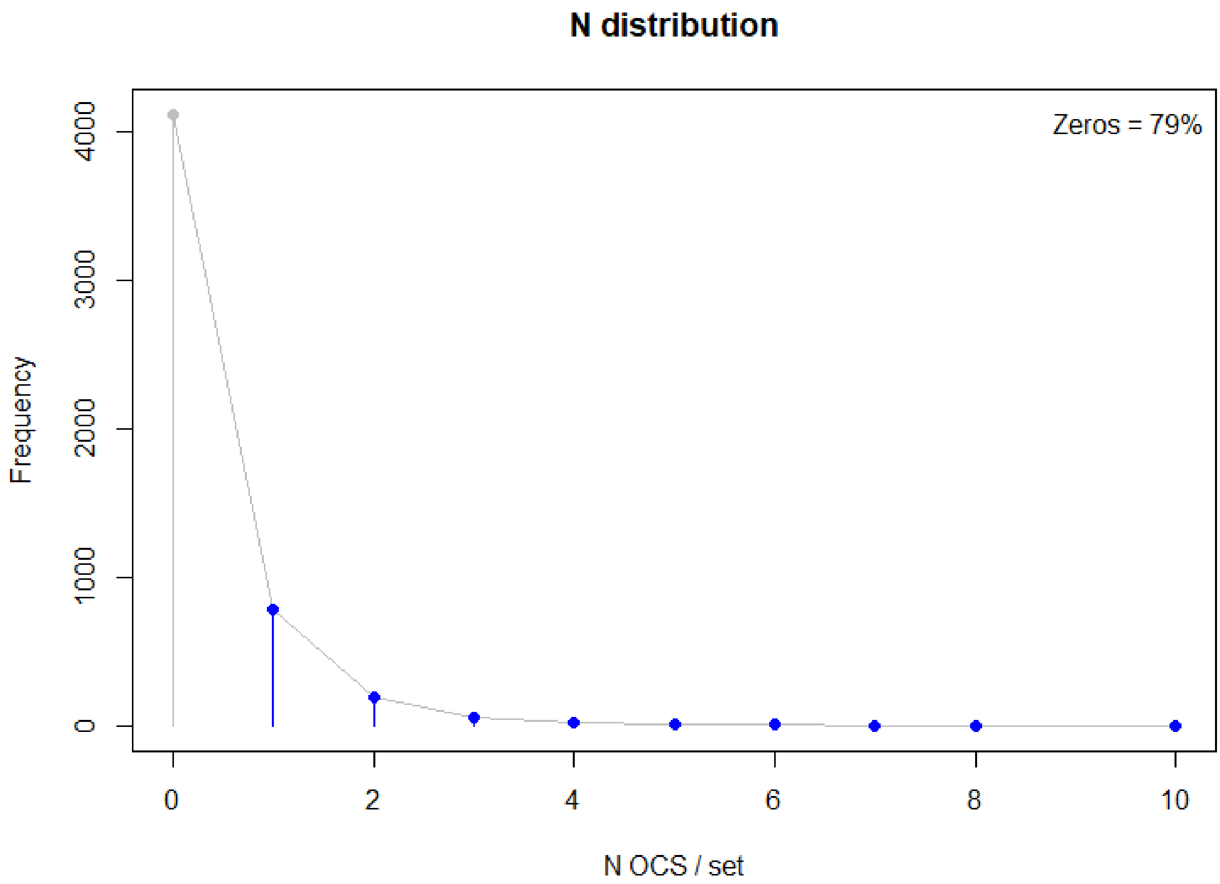

Figure 3. Frequency distribution of oceanic whitetip shark catches per set, showing a high proportion of zeros (79%).

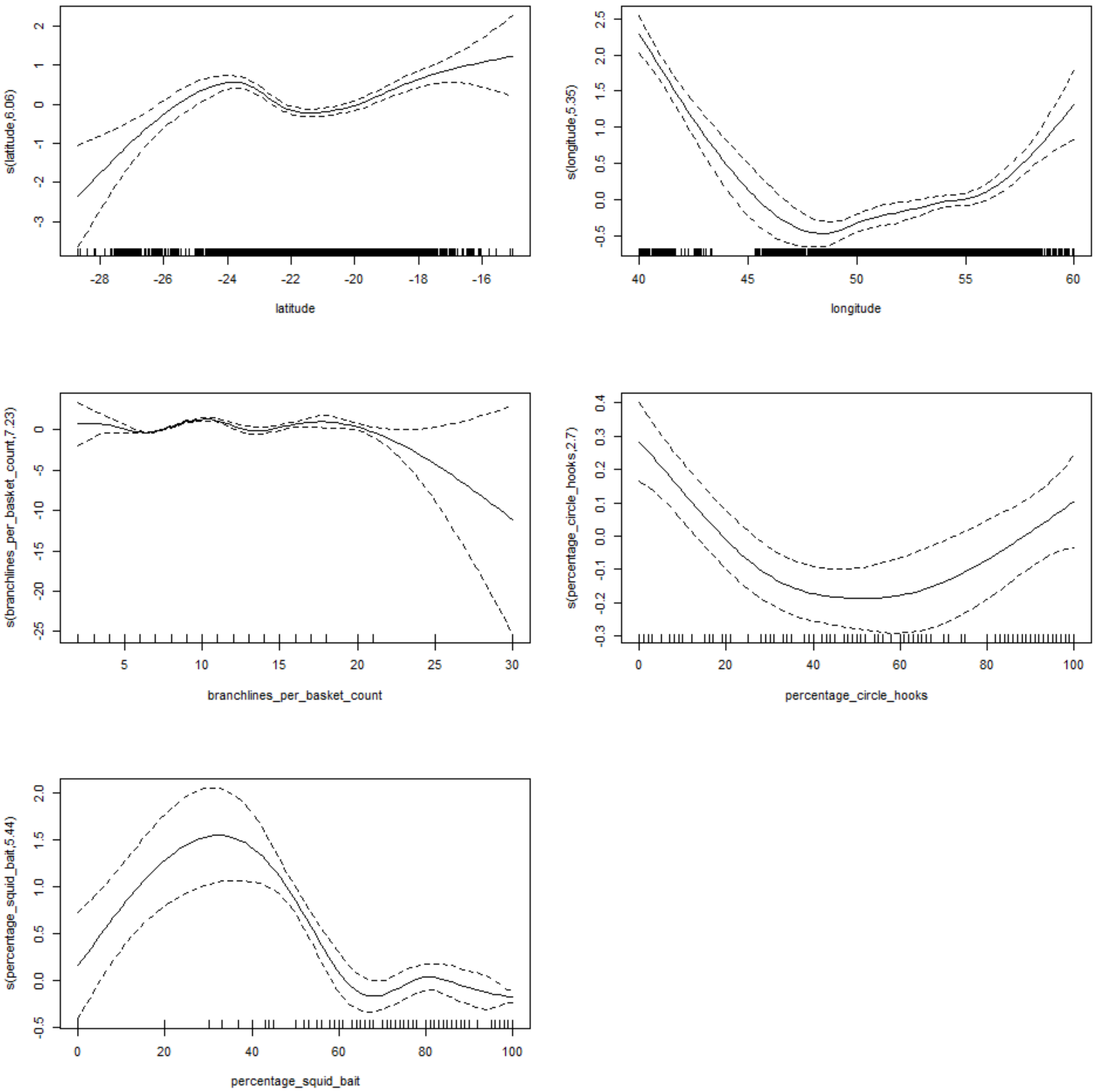

Figure 4. Individual univariate GAMMs for each continuous covariate used to explain CPUE in the retained model (Mod 2).

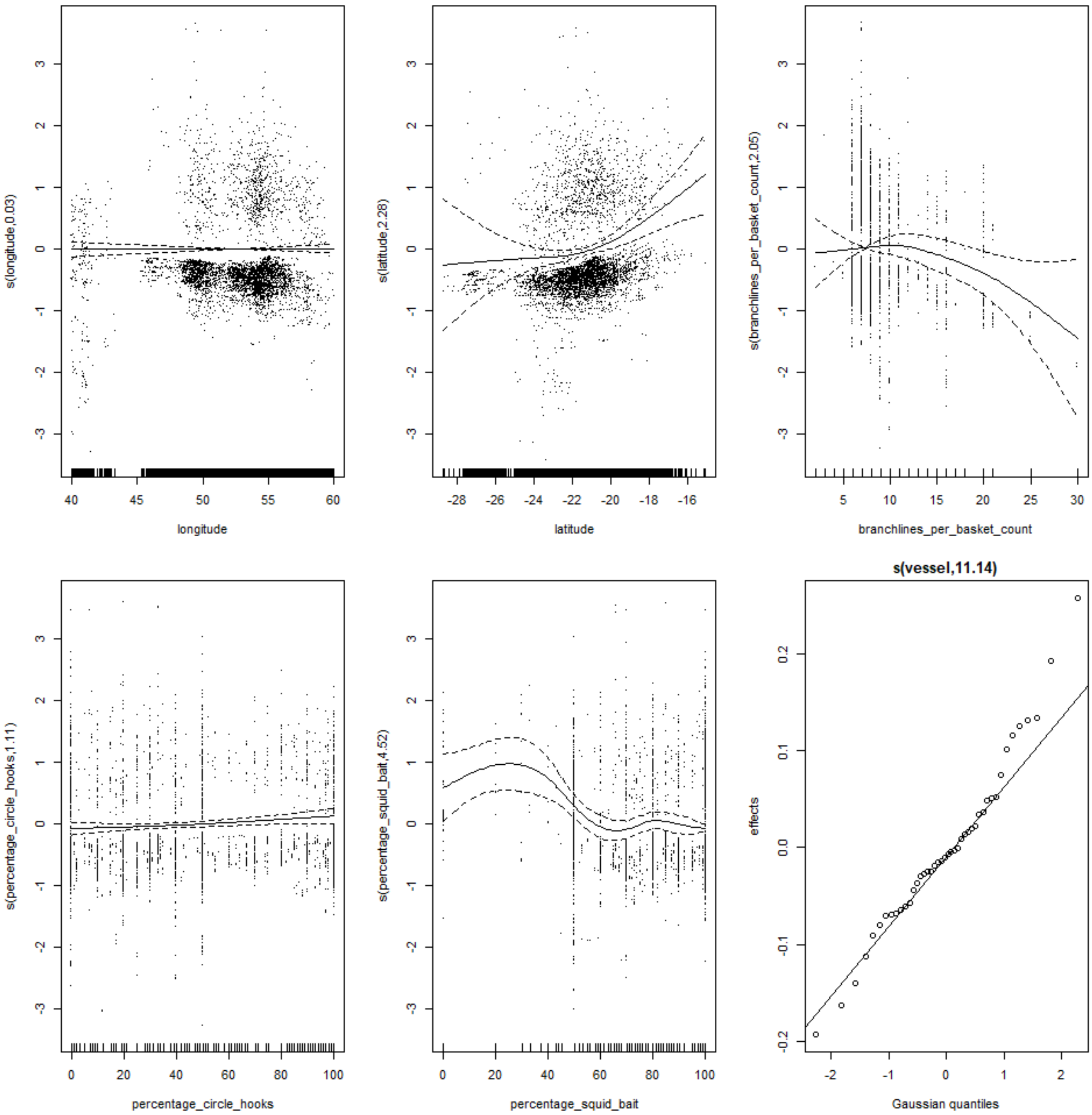

Figure 5. Splines of the continuous covariates used to explain CPUE in the retained model (Mod 2).

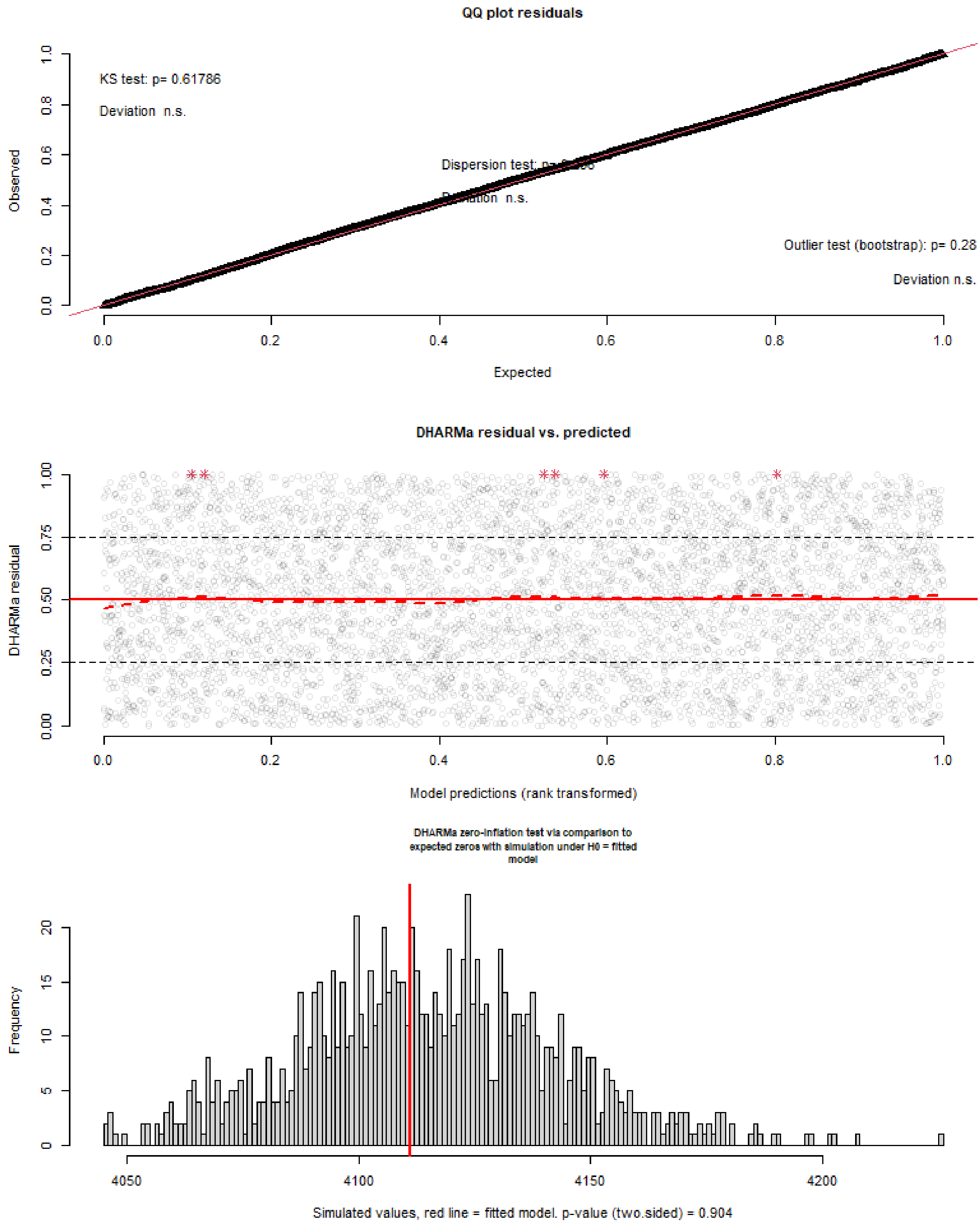

Figure 6. Residual analysis of model selected for oceanic whitetip shark CPUE standardization including the covariates selected by the forward-stepwise model selection.

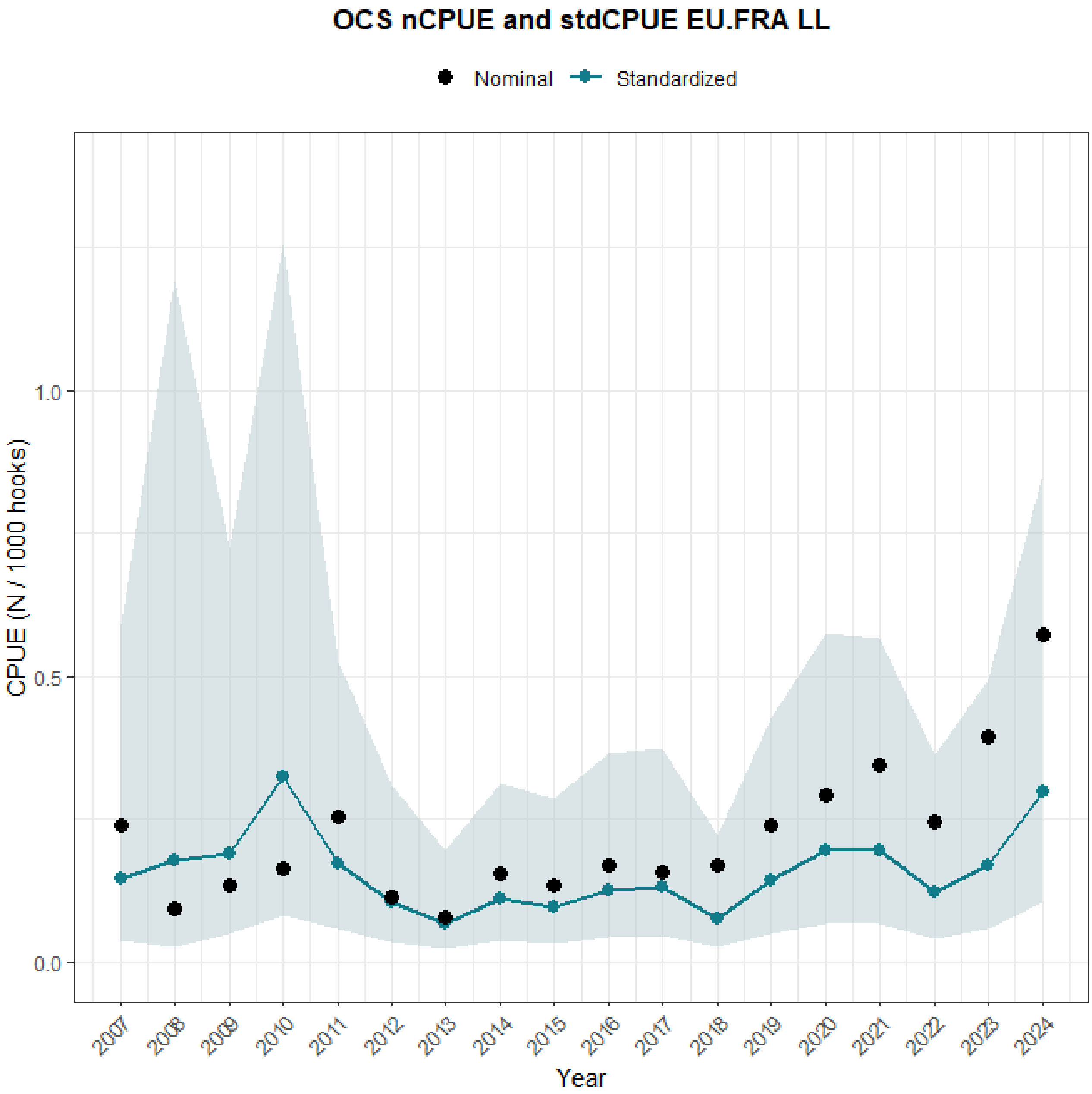

Figure 7. Nominal and standardized CPUE (N/1000 hooks) time series for the oceanic whitetip shark (OCS) caught by the French longline fishery based in Reunion Island (EU.FRA LL) for the period 2007-2024. The standardized index is derived from the final negative binomial GAMM. Error bars represent the 95% confidence intervals.

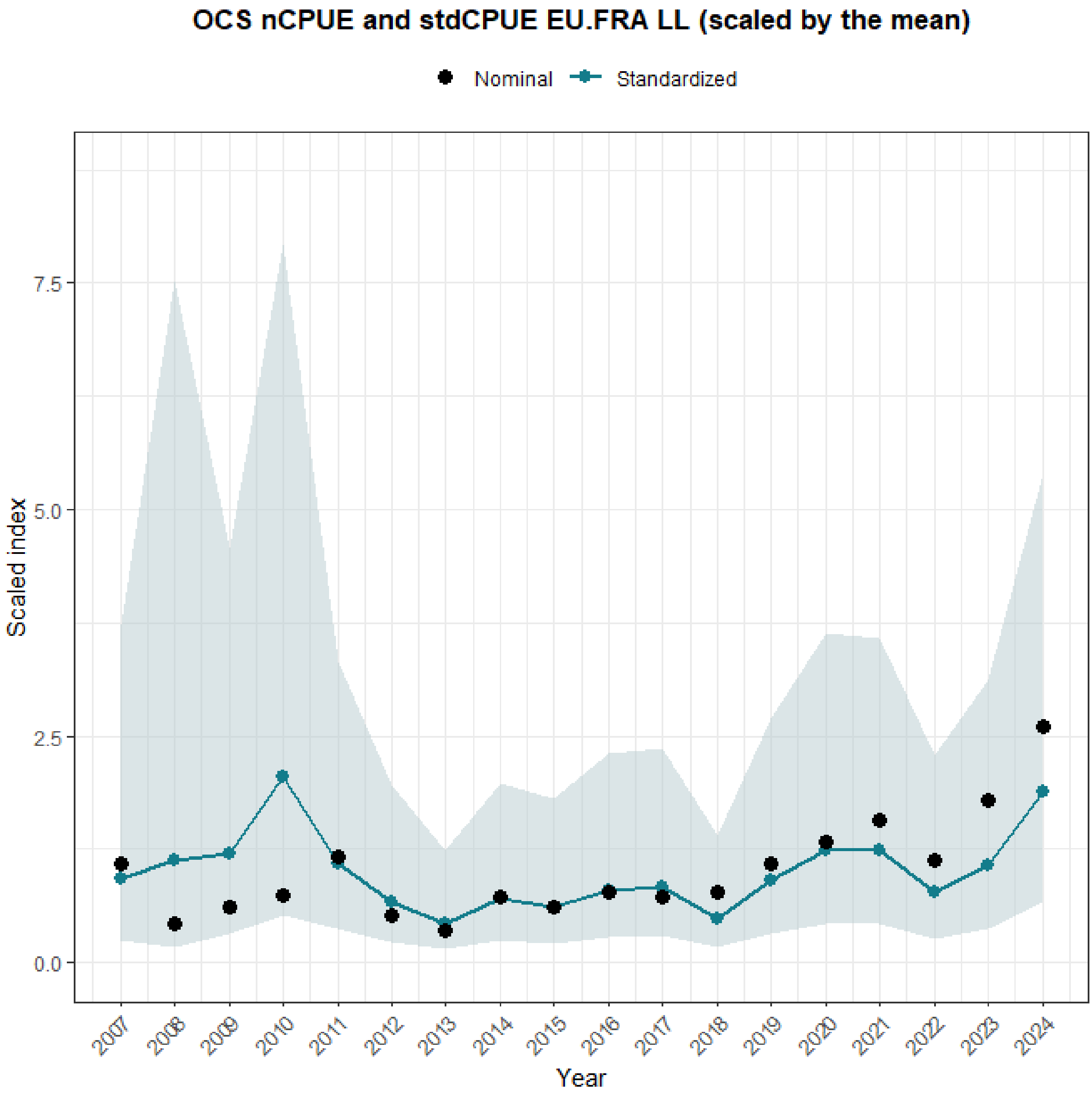

Figure 8. Scaled (by the mean) nominal and standardized CPUE (N/1000 hooks) time series for the oceanic whitetip shark (OCS) caught by the French longline fishery based in Reunion Island (EU.FRA LL) for the period 2007-2024. The standardized index is derived from the final negative binomial GAMM. Error bars represent the 95% confidence intervals.