Standardisation of yellowfin tuna CPUE for the EU purse seine fleet operating in the Indian Ocean.

By Isidora Katara[1], Daniel Gaertner[2], Francis Marsac[2], Maitane Grande[3], David Kaplan[2], Agurtzane Urtizberea[3], Loreleï Guéry[2], Mathieu Depetris[2], Antoine Duparc[2], Laurent Floch[2], Jon Lopez[4] and Francisco Abascal[5]

[1] Centre for Environment, Fisheries and Aquaculture Science (CEFAS), Pakefield Road, Lowestoft, Suffolk, NR33 0HT, UK

[2] Institut de Recherche pour le Développement (IRD), UMR MARBEC, CS 30171, Av. Jean Monnet, 34203 Sete Cedex, France

[3] AZTI Tecnalia, Herrera Kaia, Portualde z/g, 20110 Gipuzkoa, Spain

[4] Inter-American Tropical Tuna Commission, 8901 La Jolla Shores Dr., 92037, La Jolla, United States

[5] Instituto Español de Oceanografía (IEO), Centro Oceanográfico de Canarias. Vía Espaldón, dársena pesquera, Parcela 8 38180 Santa Cruz de Tenerife, Spain

Abstract

The EU purse seine fleet catches of yellowfin tuna (*Thunnus albacares*) from the Indian Ocean were standardized using the framework described in Katara *et al*. (2016, 2017) with a Delta-lognormal generalised linear mixed model developed specifically for the standardisation of tropical tuna catch per unit effort (CPUE) time series. With the aim to depict the trend in abundance for adults and for juveniles yellowfin separately, the CPUE time series were treated by fishing mode: free school (FSC) sets and sets associated with floating objects (FOBs).

CPUE for FSC was defined as the catch per hour of large yellowfin tuna (> 10 kg). For FOBs sets (i.e., dFAD and log sets), CPUE was defined as the catch per positive set of small yellowfin tuna (< 10 kg) – a positive set defined as a set with small yellowfin catches > 0 . Due to the availability of covariates that likely affect them, the time series considered were 1986-2017 and 2010-2017 for FSC and FOB, respectively. . In both cases, the least absolute shrinkage and selection operator method (LASSO) was applied for model selection.

A step forward compared to previous years was the inclusion of environmental variables known to affect catchability. Another improvement is the availability of information on dFAD densities, i.e. densities of FOBs with transmitting buoys. This standardization of yellowfin tunas CPUE for the European purse seiners in the Indian Ocean, therefore, represents a significant advance over previous efforts, having used the most recently available data on nontraditional explanatory factors, particularly on dFAD density. Nevertheless, several avenues for future progress are noted in the discussion, such as further improvements in dFAD density estimates and inclusion of additional or different explanatory variables to best represent the impacts of fishery change on CPUE.

Keywords: CPUE standardization; purse seine fishery; dFADs; FOBs; FSC; mixed models; yellowfin tuna

Introduction

This paper is the result of the *Workshop for the development of yellowfin indices of abundance for the EU tropical tuna purse seine fishery operating in the Indian Ocean*, hosted at the IRD MARBEC laboratory in Sète, France in September 2018 as part of the EU funded project Cecofad II[1]. The workshop aimed at developing standardised CPUE time series to be provided to IOTC as an input for the upcoming stock assessment of yellowfin tuna (YFT, *Thunnus albacares*). Catch per unit effort (CPUE) time series were standardized by fishing mode: sets on free school (FSC) and sets associated with floating objects (FOBs).

We followed the recommendations of the 2016 workshop for *the development of indices of abundance for the EU tropical tuna purse seine fishery* (Gaertner *et al*., 2017), as well as the framework described in Katara *et al*. (2016, 2017). Delta-lognormal generalised linear mixed models (Delta-lognormal GLMMs) with LASSO component were developed for the standardisation of the time series. Along with the commonly used covariates relating to vessel characteristics and spatiotemporal variability, we also considered three environmental variables (chlorophyll-a, the vertical current shear and the depth of 20°C isotherm) and information on drifting fish aggregating devices (hereafter dFADs) density that may affect catchability (Escalle *et al*., 2018).

Material and Methods

CONVENTIONAL FISHING DATA

Logbook data for the French and Spanish purse seine fleets targeting tropical tuna in the Indian Ocean from 1986 to 2017 were analysed to derive the standardised CPUEs. The logbook databases are managed by the Tuna Observatory (Ob7) and the IEO for the French and the Spanish fleets, respectively. The raw logbook data (Level 0) produced by the skippers were corrected in terms of total catch per set (to account for the difference between reported catch at sea and landed catch) and species composition (based on port size sampling and the T3 methodology – see Pallarès and Hallier 1997, Duparc *et al*., 2018) to generate the Level 1 logbook database used in this paper.

The database was split into 2 datasets: (i) free-school sets (FSC), i.e. non-associated school sets and whales' sets and (ii) FOB-related sets, including dFAD, logs and whale-sharks' sets. The FSC dataset was used to derive CPUE for the adult fraction of the yellowfin stock, by selecting the size categories 2 and 3 (10-30 kg and >30 kg respectively). The FOB sets dataset was used to derive CPUE for the juvenile fraction of the yellowfin stock, based on the size category 1 (< 10kg) in the logbook records.

The analysis was restricted to:

- the period 1986-2017 for FSC sets because the fishery only reached its full spatial distribution after 1985, and the period 2007-2017 for FOB related sets due to dFAD density data availability;

- the area defined by all grid cells that were fished for at least 5 years over a period of no less than 15 years, to avoid areas that are not routinely fished;

- high-seas and all EEZs except for the Somali EEZ due to the effects of piracy (Okamoto, 2011; Chassot *et al*., 2012; Guillotreau *et al*., 2012).

dFAD AND BUOYS DATA

We assumed that the density of the surrounding floating objects (FOBs) can affect the size of the school aggregated under a floatting object (Fonteneau and Marsac, 2016). However, it remains

[1] Catch, effort, and ecosystem impacts of tropical tuna fisheries (CECOFAD2); EASME/EMFF/2016/008

difficult to estimate the total number of FOBs, i.e. drifting Fish Aggregating Devices (dFADs) and instrumented and non-instrumented logs, by 1°x1° grid cell - month. Previous studies showed that the vast majority of FOBs in the Indian Ocean were currently dFADs with buoys (Maufroy *et al*. 2015). Consequently, for the sake of simplicity we considered that the density of transmitting buoys (on dFADs or logs, but hereafter referred to as dFADs) can be used as a proxy of the total FOB density. No information on dFADs with inactive transmitting buoys or natural FOBs (logs) without transmitting buoys were available.

For each month over the period 2010-2017, the average number of transmitting buoys found in each 1ºx1º cell of the Indian Ocean was calculated. For French dFAD trajectory data, individual dFAD water trajectories were linearly interpolated at midnight GMT on each day. These daily French interpolated water positions were then assigned to 1ºx1º grid cells, summed for each month and then these sums were divided by the number of days in the month. Spanish data consisted of one position per day, so no interpolation was necessary. French data coverage was 100% over the period 2010-2017, whereas Spanish data coverage was partial, progressively increasing from a mean of 32% in 2010 to over 70% in 2017 (estimated by the fraction of the fleet and the types of transmitting buoys for which data were available; the monthly variation of the coverage rate was not quantified in this study). To correct for the partial coverage, total Spanish buoy densities were extrapolated from available data by dividing the initial Spanish dFAD density values in each grid cell-month strata by the fraction data coverage for the corresponding month (i.e., the number of vessels sharing the information and availability of information by buoy model, assuming the same deployment strategy for all Spanish vessels). The total number of Spanish dFADs was estimated for grid cells in which the vessels regularly operate (> 40ºS and longitude < 90ºE).

The dFAD datasets of the 2 fleets were then combined and three indicators were calculated:

- distance of the set from the edge of the nearest dFAD hotspot. dFAD hotspots were derived using the ArcGIS algorithm for hotspots (Getis and Ord, 1992), positive hotspots with p-values < 0.05 were selected,
- dFAD density in an area of a 1ºx1º buffer around the fishing set, and (iii) dFAD density in an area of a 2ºx2º buffer around the fishing set.

ENVIRONMENTAL DATA

Three environmental datasets were considered:

1. chlorophyll-a concentration derived from SeaWiFS and MODIS (O'Reilly *et al*., 1998) over the period September 1997 to December 2017. High Chl-a values indicate areas with high productivity and potentially high density of micronekton organisms preyed upon by tuna. For instance, the record catches of yellowfin in 2004-2005 were associated with anomalously high levels of Chl-a (Marsac 2008, Fonteneau *et al*., 2008) and an outburst of the stomatopod *Natosquilla investigatoris* found in abundance in tuna stomachs (Potier *et al*. 2007). At a monthly timescale, 1°x1° squares with high levels of Chl-a can thus be indicative of foraging aggregations of tuna and thus increased catchability.

2. The vertical current shear between depths 5 and 145 m; the depth range refers to the major part of the water column sampled during a purse seine set. The vertical current shear may affect the tension on the net with possible consequences on the depth reached by the seine. We used the method developed by Bigelow *et al*. (2006), simplifying the equation by taking only two levels instead of integrating through all depth levels. The vertical current shear calculated between two levels z1 and z2 (5 and 145 m respectively) is obtained by:

S = Log(W / Δ z)

$W = \left|\underline{U}_{z1} - \underline{U}_{z2}\right|$ where $\underline{U}_z$ is the horizontal velocity vector at depth level z.

3. Changes in depth of the mixed layer are known to affect the catchability of purse seines, as surface-dwelling tunas mostly gather above the thermocline (Green, 1967; Cayré and Marsac 1993; Bertrand *et al*., 2002). A deep thermocline may, therefore, decrease the vulnerability of tuna schools to purse seining. We used the depth of 20°C isotherm as a proxy of the thermocline depth, and subsequently, of the mixed layer depth.

Chlorophyll is obtained by direct (satellite) measurements (SeaWiFS for 1997-2002 and Modis for 2003-2017). The vertical current shear and the depth of the mixed layer were computed from ocean model outputs. We used the Global Ocean Data Assimilation System (GODAS) of the NCEP/NOAA which is an assimilated model incorporating continuous real-time data from the Global Ocean Observing System. The grid resolution is 1° in longitude and 1/3° in latitude, and we used the monthly time steps of GODAS which cover the period 1980 to the present.

MODELLING APPROACH

We followed the modelling approach developed by Katara *et al*. (2016, 2017). As the CPUE data for free schools followed a zero-inflated lognormal distribution, delta-lognormal GLMMs were used which comprised of two sub-models (a binomial GLMM that standardises the probability of a positive set, and a lognormal GLMM that standardises catch conditional to positive set). We performed the binomial GLMM where the full model included the following fixed effects: fleet country, vessel capacity category, year that the vessel started its activity, year, month, quarter, fishing set duration, mixed layer depth, vertical current shear, latitude, longitude, vessel length, vessel horsepower, vessel storage capacity, the proportion of FOB sets per trip. The last parameter was included as a proxy for vessels' fishing strategy changes across time due to the increase of dFADs. We also tested for the interactions between year and month and between geographical coordinates. The random structure of the model includes the Exclusive Economic Zone, the T3 area used for logbook catch correction, a vessel unique identifier, a trip unique identifier and the interaction between year and 1ºx1º square.

The full model for the lognormal GLMM included the following fixed effects: fleet country, vessel capacity category, year that the vessel started its activity, year, month, quarter, fishing set duration, mixed layer depth, vertical current shear, latitude, longitude, vessel length, vessel horsepower, vessel storage capacity, the proportion of FOB sets per trip, interaction between year and month and interaction between geographical coordinates. The random structure of the model includes the Exclusive Economic Zone, the T3 area used for logbook catch correction, a vessel unique identifier, a trip unique identifier and the interaction between year and 1ºx1º grid cell.

For both models, if the residuals indicated non-linear relationships, polynomials were used to describe those relationships.

Table 1 available variables for the calculation of CPUE and the development of the standardisation models.

Variable	**description**
fleet country	France; Spain
Vessel ID	Unique vessel identifier
vessel capacity category	Vessel category related with vessel length and capacity
the year that the vessel started its activity	
vessel length	In meters
vessel horsepower	In kws

vessel storage capacity	In m^3
quarter	A quarter of the year at which the fishing set took place
year	Year at which the fishing set took place
month	Month at which the fishing set took place
Time at sea	Duration of the fishing trip
Time fishing	Cumulated time dedicated to fishing
Fishing set duration	In h
Searching time	In h
Mixed layer depth	Details in the main document
Vertical current shear	Details in the main document
chlorophyll-a concentration	Details in main document
latitude	Fishing set location coordinates
longitude	Fishing set location coordinates
distance from a dFAD hotspot	The distance of the set from the edge of a FAD hotspot. Temporal Resolution: monthly
dFAD density at a 1º buffer	Density around the set. Temporal Resolution: monthly
dFAD density at a 2º buffer	Density around the set. Temporal Resolution: monthly
the proportion of FOB sets per trip	The number of FOB sets divided by the total number of sets
Exclusive Economic Zone	Identifiers of EEZs and the offshore area
T3 area used for logbook catch correction	T3 areas used for correcting the species composition of the catch reported in the logbook
trip ID	trip unique identifier
1ºx1º grid cell	Reference grid of the fishing area at a 1ºx1º resolution

The CPUE for FOB related sets was defined as catch per positive set of small YFT (size < 10 kg) – positive set being every set with small YFT catch > 0. Because the ratio of positive sets remains practically stable and greater than 90% throughout the time series, the CPUE was simplified to catch per set conditional to catch > 0 (i.e. catch per positive set, the second sub-model of the Delta-lognormal GLMM approach). The full model included the following fixed effects: fleet country, vessel capacity category, year that the vessel started its activity, year, month, quarter, time at sea, time fishing, fishing set duration, mixed layer depth, vertical current shear, chlorophyll-a concentration, latitude, longitude, distance from a dFAD hotspot, dFAD density at a 1º buffer, dFAD density at a 2º buffer, searching time, vessel length, vessel horsepower, vessel storage capacity, proportion of FOB sets per trip, interaction between year and month and interaction between geographical coordinates. If the residuals indicated non-linear relationships, polynomials were used to describe them. The random structure of the model includes the Exclusive Economic Zone, the T3 area used for logbook catch correction, a vessel unique identifier, a trip unique identifier and the interaction between year and 1ºx1º grid cell.

Model selection involved the use of the LASSO regression (Tibshirani 1996, 2011), using algorithms that handle continuous explanatory variables (R package: glmnet; Friedman *et al*. 2009, 2010) and grouped covariates (R package: grpreg; Breheny and Breheny, 2018). Given a linear regression with standardized predictors *xi* and centred response values y_i for i=1,2, …, N and j=1,2, …, p, the glmnet algorithm estimates the regression coefficients $b=\{b_j\}$ to minimize:

$$\frac{1}{N}\sum_{i=p}^{N} w_i\, l(y_i, b_0 + b^T x_i) + \lambda\left[\frac{(1-\alpha)||b||_2^2}{2} + \alpha||b||_1\right]$$

where λ covers a range of values, l(y,η) is the negative log-likelihood contribution for observation i and a controls the elastic-net penalty (for lasso α=1). The tuning parameter λ is chosen through cross-validation.

The LASSO procedure was followed by backward model selection for both the random and fixed effects of the mixed models using AIC and BIC. Finally, the selected model was refitted as an unrestricted GLMM (R-package: lme4; Bates et al., 2014) but not with LASSO, as LASSO estimated coefficients are known to be biased (Friedman *et al.*, 2001). Finally, the standardized CPUEs were fitted using estimated marginal means (R package: emmeans; Lenth, 2018).

Residuals were tested for patterns including spatial/temporal autocorrelation (R package: DHARMa; Hartig, 2017). All the statistical analyses were computed using the software R (v3.4.3; R Core Team, 2017).

Results

FSC SETS (1986-2017 PERIOD)

Binomial GLMM (probability of large-size YFT catch > 0)

The model selection methods (Figs 1-2) gave the following model: i. the fixed effects included vessel capacity category (p-value < 0.001), year (p-value < 0.001), month (p-value < 0.001), longitude (p-value < 0.001) and a 2nd degree polynomial for latitude (p-value < 0.001) and ii. the random effects included a trip unique identifier and the interaction between year and 1ºx1º square. All fixed and random effects were statistically significant with 99% confidence. There was no obvious trend in residuals (Figs 3-4). Estimated marginal means time series of probability catch > 0 at an annual scale is shown in Fig 5. GLMM tables and results are presented in appendices.

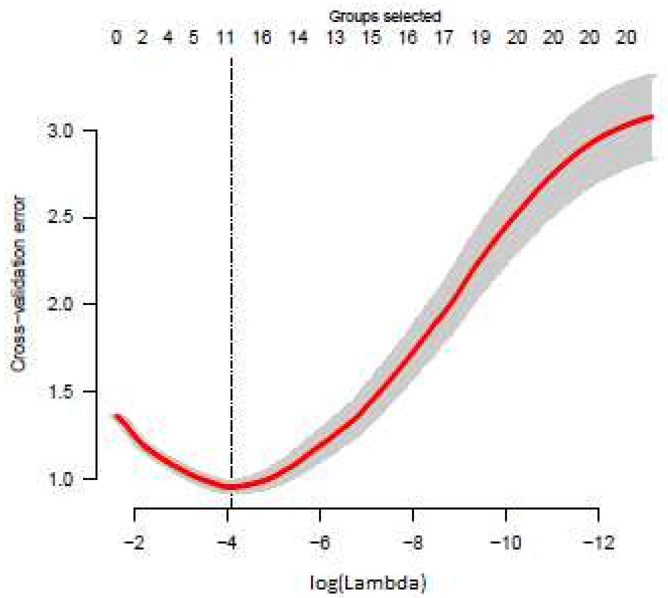

Figure 1 FSC sets – probability (catch > 0): cross-validation estimation of lambda for group Lasso (grpreg).

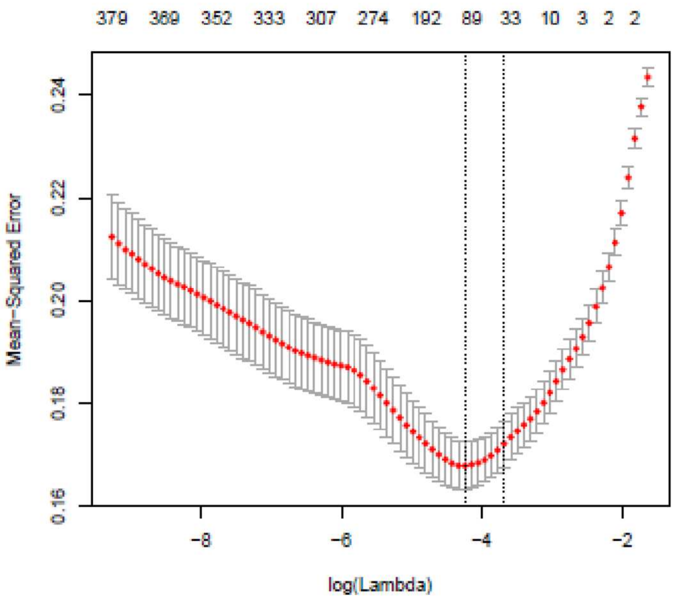

Figure 2 FSC sets – probability (catch > 0): cross-validation estimation of lambda for Lasso with glmnet.

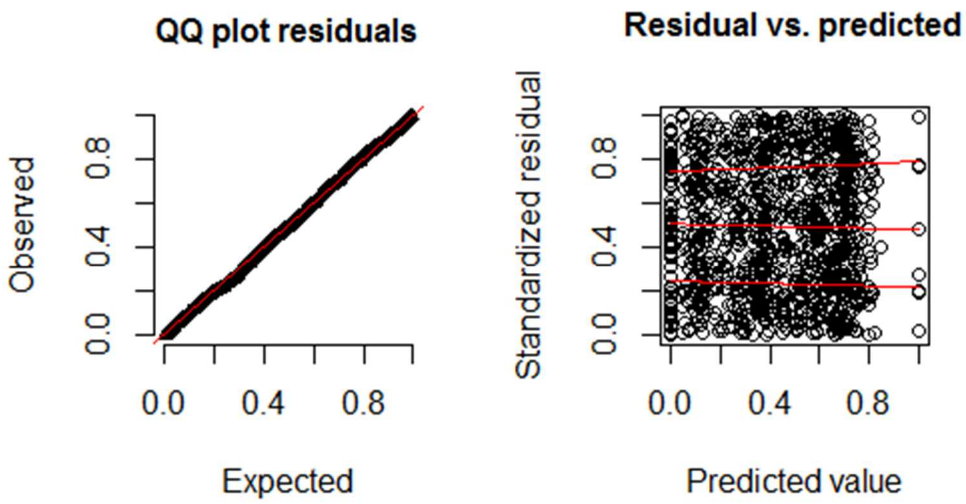

Figure 3 FSC sets – probability (catch > 0): testing residuals for normality (left) and homogeneity (right)

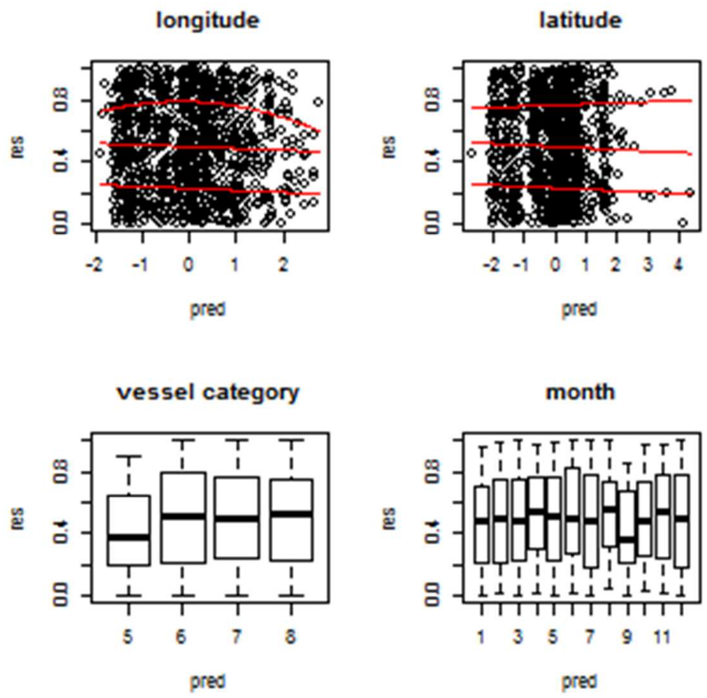

Figure 4 FSC sets – probability (catch > 0): Residuals versus fixed effects.

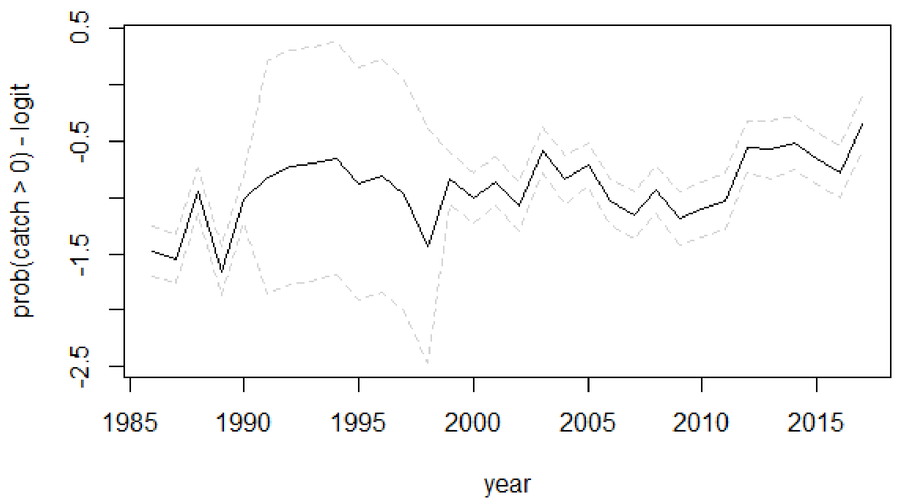

Figure 5 FSC sets – predicted probability (catch > 0): standardised time series (by year) with 95% confidence intervals.

Log-Normal GLMM (catch per hour conditional to YFT catch > 0)

The model selection methods (Figs 6-8) gave the following model: i. the fixed effects included fleet country (p-value = 0.16), vessel capacity category (p-value = 0.09), the interaction between year and month (p-value < 0.001) , mixed layer depth (p-value = 0.005), vertical current shear (p-value = 0.1) , year that the vessel started its activity (p-value < 0.001) and ii. the random effects included the

Exclusive Economic Zone, the T3 area of logbook catch correction, a vessel unique identifier, a trip unique identifier and the interaction between year and 1ºx1º square. The residuals (Figs 9-10) show a reasonable fit of the model (Fig 11) with a negligible divergence from normality and homogeneity. GLMM tables and results are presented in appendices.

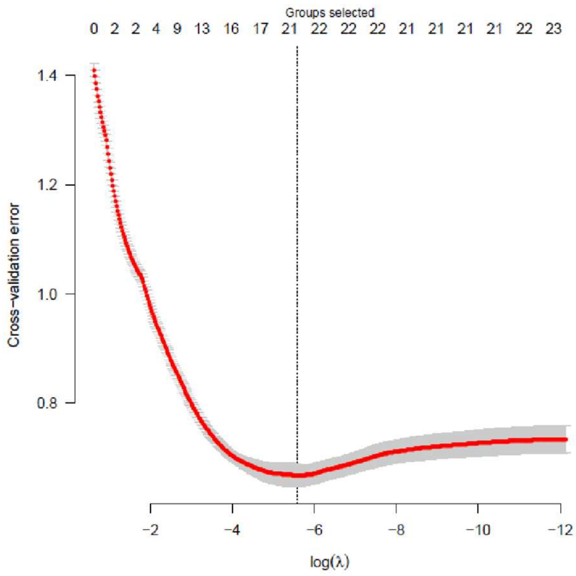

Figure 6 FSC sets – catch per hour | catch > 0: cross-validation estimation of lambda for group Lasso (grpreg)

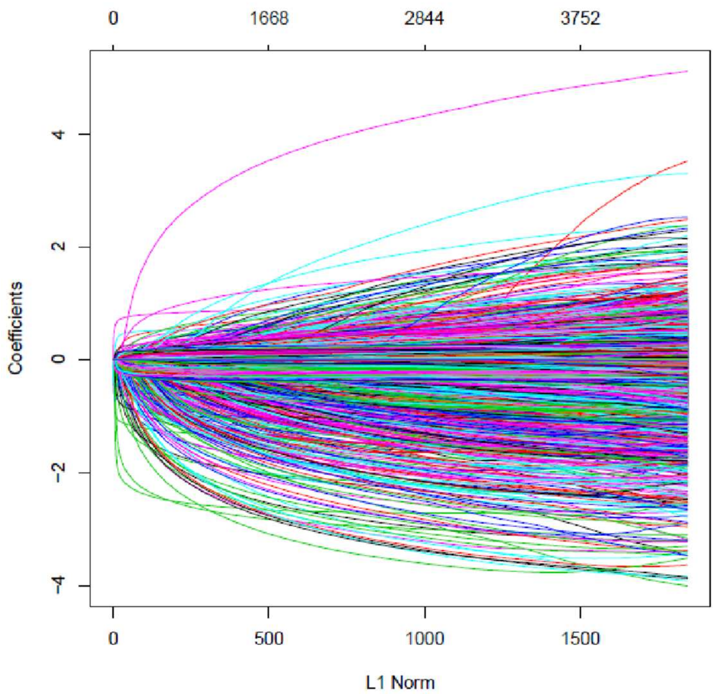

Figure 7 FSC sets – catch per hour | catch > 0: visualisation of the path of the coefficients against l_1-norm for glmnet

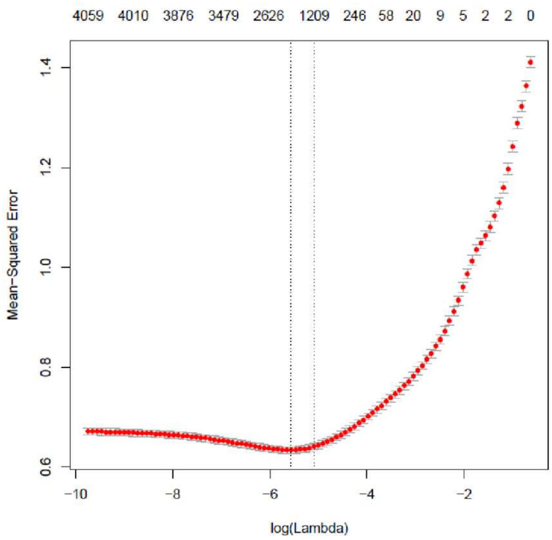

Figure 8 FSC sets – catch per hour | catch > 0: cross-validation estimation of lambda for glmnet

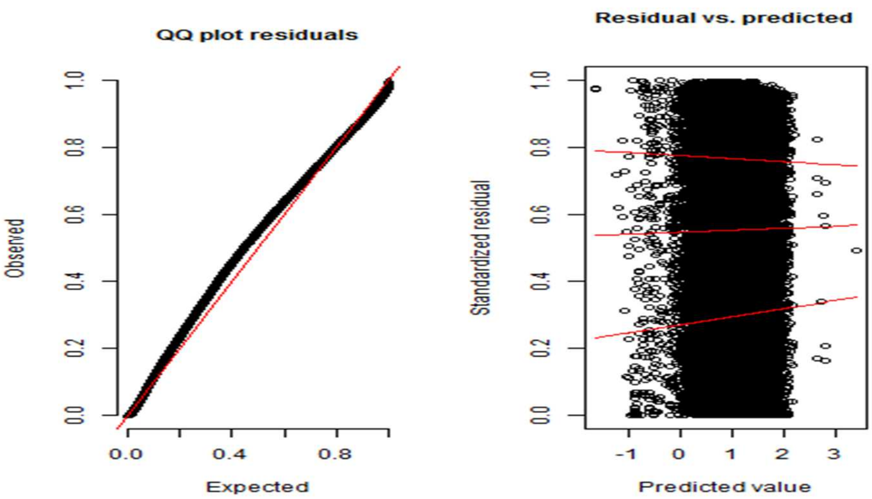

Figure 9 FSC sets – catch per set | catch > 0: testing residuals for normality (left) and homogeneity (right)

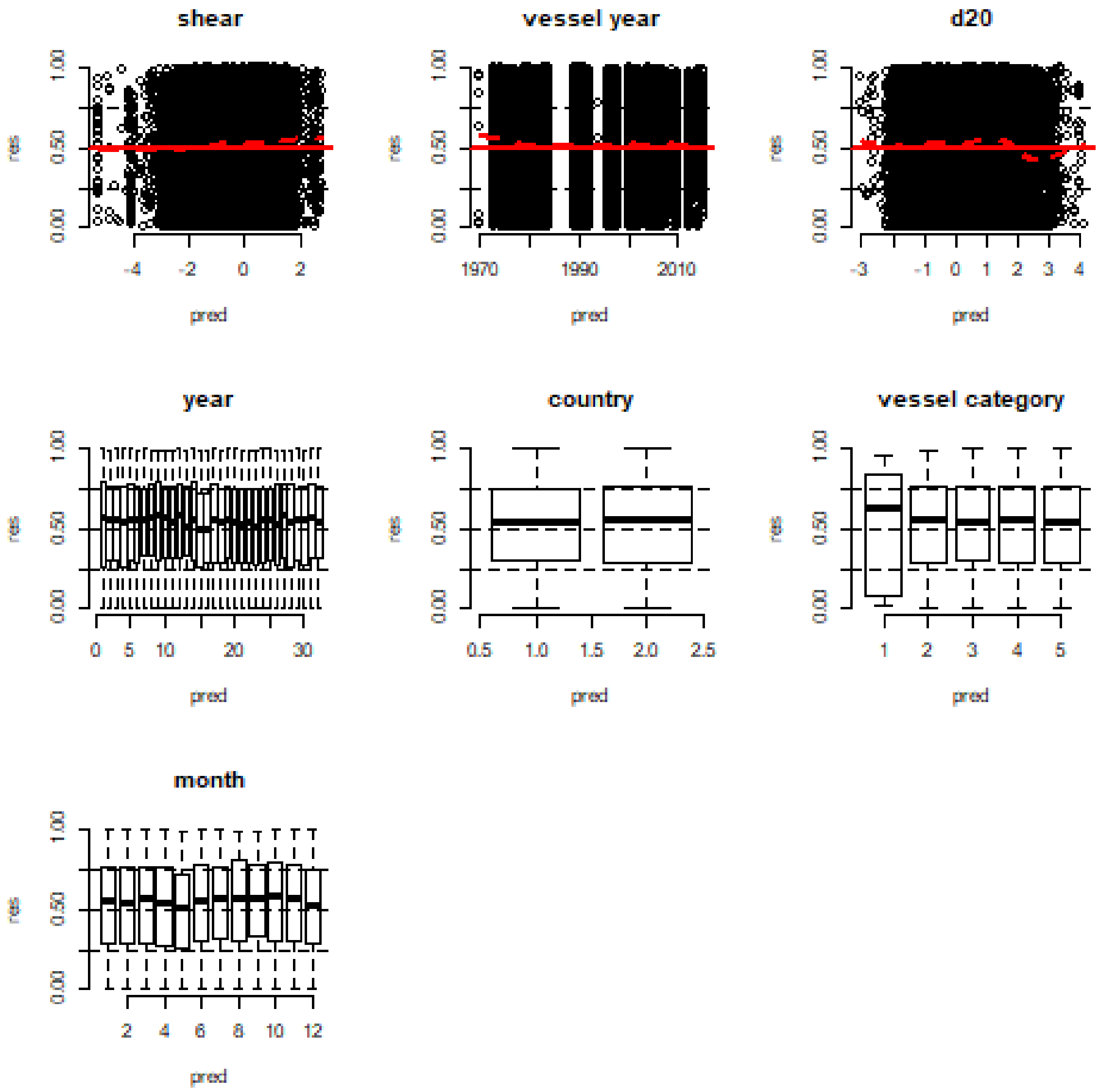

Figure 10 FSC sets – catch per set | catch > 0: Residuals versus fixed effects.

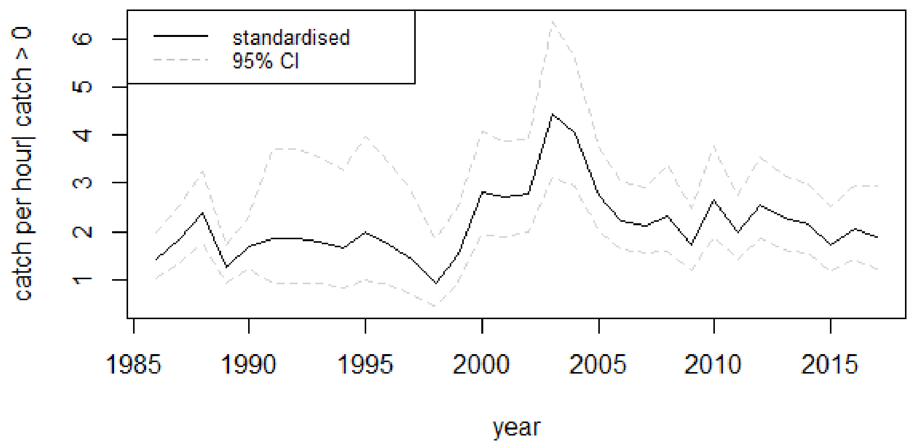

Figure 11 FSC sets – catch per set | catch > 0: standardised time series (by year) with 95% confidence intervals.

Delta lognormal GLMM approach

The product of the two sub-models described above provided the standardised CPUE time series for free school sets (Fig 12).

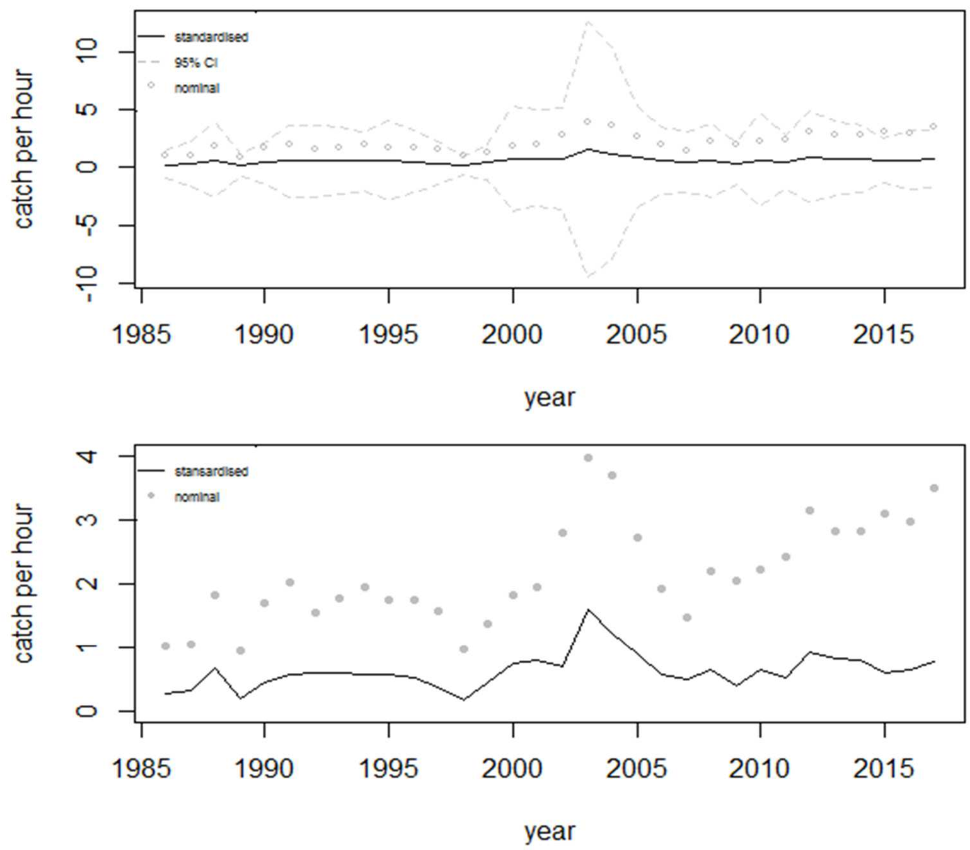

Figure 12 standardised CPUE (catch per hour) for free school sets with 95% CIs (top) and compared to nominal CPUE (bottom). Time series on an annual basis.

FOB-RELATED SETS (2007-2017 PERIOD)

Log-Normal GLMM (catch per set conditional to YFT catch > 0)

The model selection methods (Figs 13-15) gave the following final model: i. the fixed effects included vessel capacity category, the interaction between year and month (p-value < 0.0001), mixed layer depth (p-value < 0.0001), vertical current shear (p-value = 0.08), the interaction between latitude and longitude (p-value < 0.0001), vessel horsepower (p-value = 0.005), vessel storage capacity (p-value = 0.2), proportion of FOB sets per trip (p-value = 0.03), fleet country (p-value = 0.0005) and ii. the random effects included the Exclusive Economic Zone, the T3 area of logbook catch correction, a vessel unique identifier, a trip unique identifier and the interaction between year and 1ºx1º square. The residuals (Figs 16-17) show a reasonable fit of the model (Fig 18) with a slight divergence from normality and homogeneity that can be considered negligible due to the robustness of the model. GLMM tables and results are presented in appendices

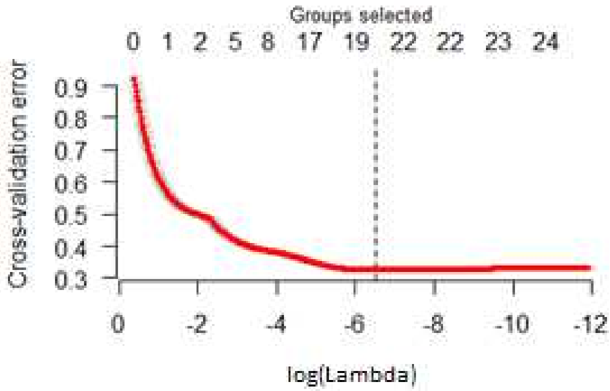

Figure 13 FOB-related catch per set| catch > 0: cross-validation estimation of lambda for group LASSO (grpreg).

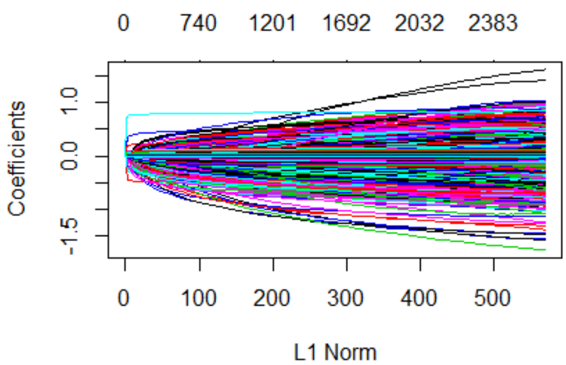

Figure 14 FOB-related catch per set| catch > 0: visualisation of the path of the coefficients against l_1-norm for the Lasso regression (glmnet).

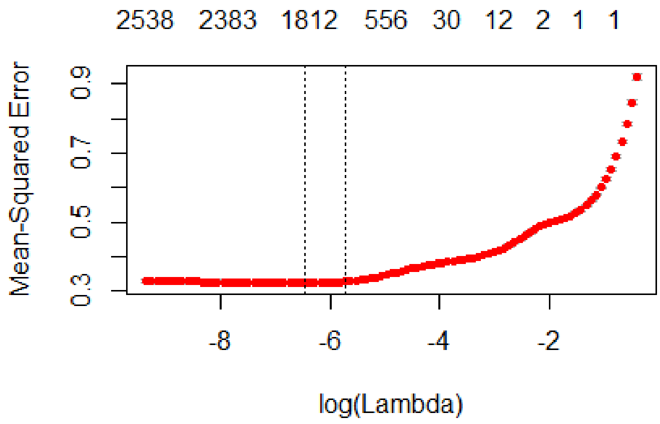

Figure 15 FOB-related catch per set| catch > 0: cross-validation estimation of lambda for Lasso with glmnet.

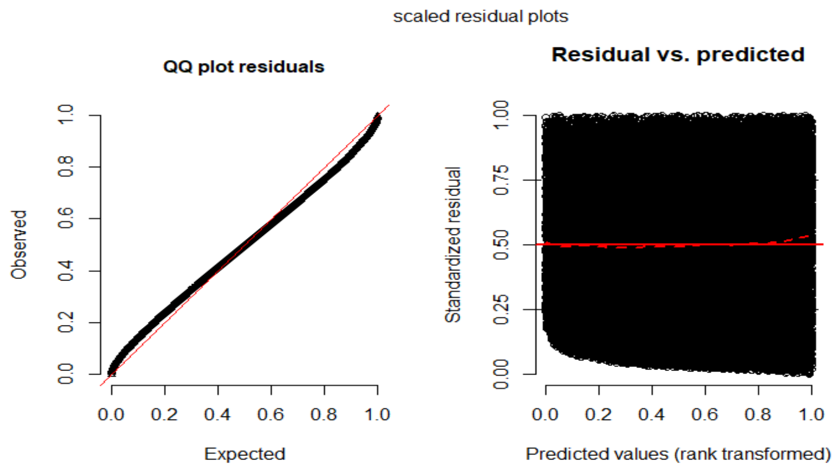

Figure 16 FOB-related catch per set| catch > 0: testing residuals for normality (left) and homogeneity (right)

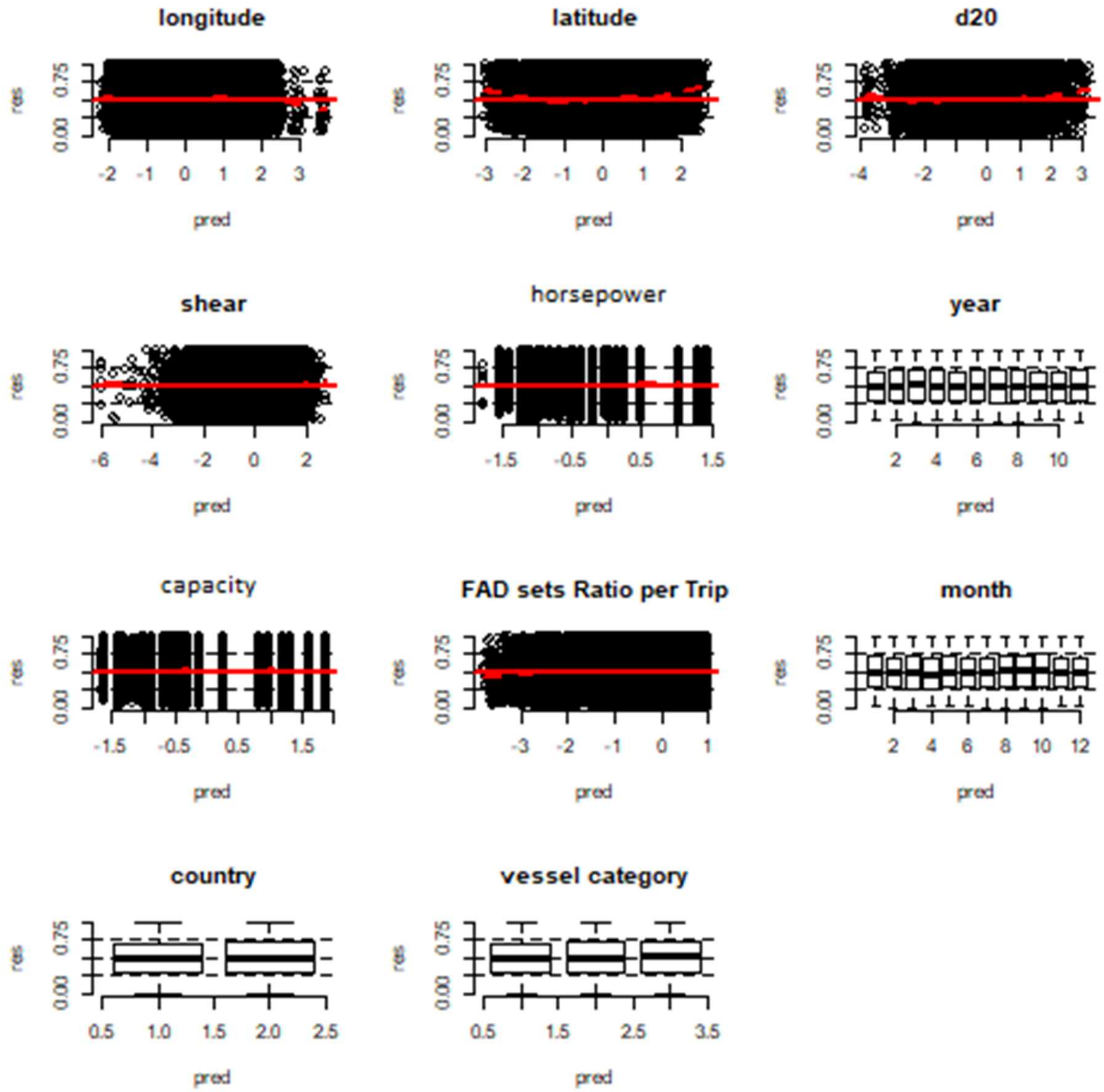

Figure 17 FOB-related catch per set| catch > 0: Residuals versus fixed effects

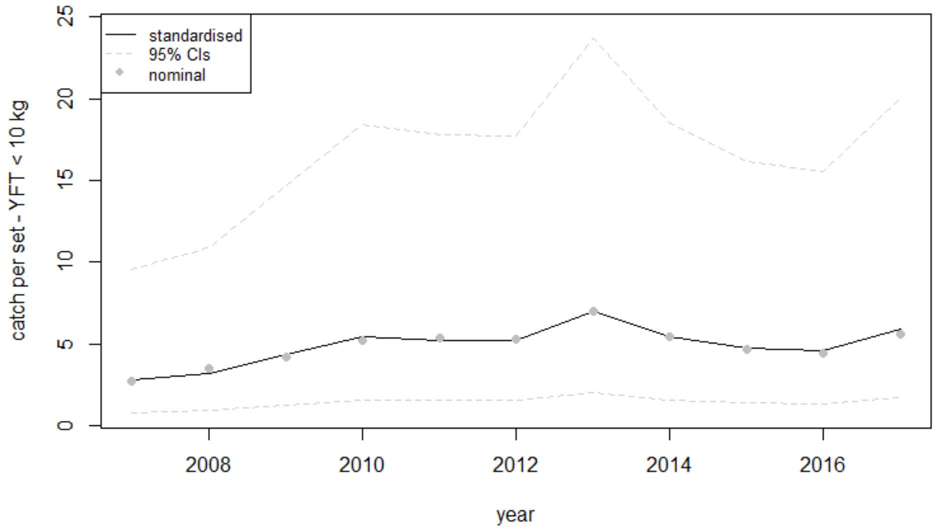

Figure 18 Catch of YFT < 10 kg per set for FOB related sets. Nominal and standardised CPUE with 95% confidence intervals. Time series on an annual basis.

Discussion

As mentioned in the Method section, we followed the framework for CPUE standardization for the tropical tuna purse seine fisheries described in Katara *et al*. (2016) to account for the hierarchical structure of the data, for the non-randomised sampling and the numerous candidate variables linked to technological developments and evolving fishing strategies. A step forward compared to previous years was the inclusion of environmental variables known to affect catchability. Another major improvement is the availability of information on dFAD densities, i.e. densities of FOBs with transmitting buoys. Indeed, this study of the standardization of yellowfin tunas CPUE for the European purse seiners operating in the Indian Ocean represents to our knowledge the most extensive effort to include data nontraditional explanatory factors, particularly on dFAD densities.

Environmental variables, primarily mixed layer and, to a lesser extent, vertical current shear, were important variables for predicting catch per hour on free schools and catch per FOB.

It has been theorized that dFAD densities should affect catch per FOB set via, for example, disruption of tuna schooling behaviour at high FOBs densities (Fonteneau and Marsac, 2016). In this preliminary analysis, dFAD densities were not informative for models of catch on FOB sets. There are a number of possible explanations for this. Though to our knowledge this study represents the most thorough effort to collect dFAD position and density information in any tropical ocean, incomplete data coverage is a potential explanatory factor for the lack of an observed dFAD density effect. Whereas complete data on French buoys have been used in this analysis, Spanish buoy densities were estimated from partial data for which the coverage rate over time (i.e. by month) has not been quantified yet. As buoy densities are estimated on the relatively fine scale of 1° months, this could lead to bias in buoy density estimates over space and time. One indication that this might be the case is that the proportion of EU dFADs that are French as estimated from our buoy density data is considerably higher than a previous estimate in the Indian Ocean of 10.4% for 2013 based on random encounters with dFADs noted by observers aboard EU purse seiners (Maufroy *et al.*, 2017; Figure 19). It is also noteworthy that although Spanish and French data look very similar after ~2014, suggesting that the Spanish data is a reasonable representation of the true spatial distribution of dFADs after this time, the larger differences observed between data from the two fleets prior to 2014 may be indicative of potential problems requiring further exploration (Figure 20; other explanations, such as differences in fishing strategy, are also quite possible). As data coverage has greatly improved over time and continues to improve, it is likely that these concerns regarding data coverage will diminish in future studies.

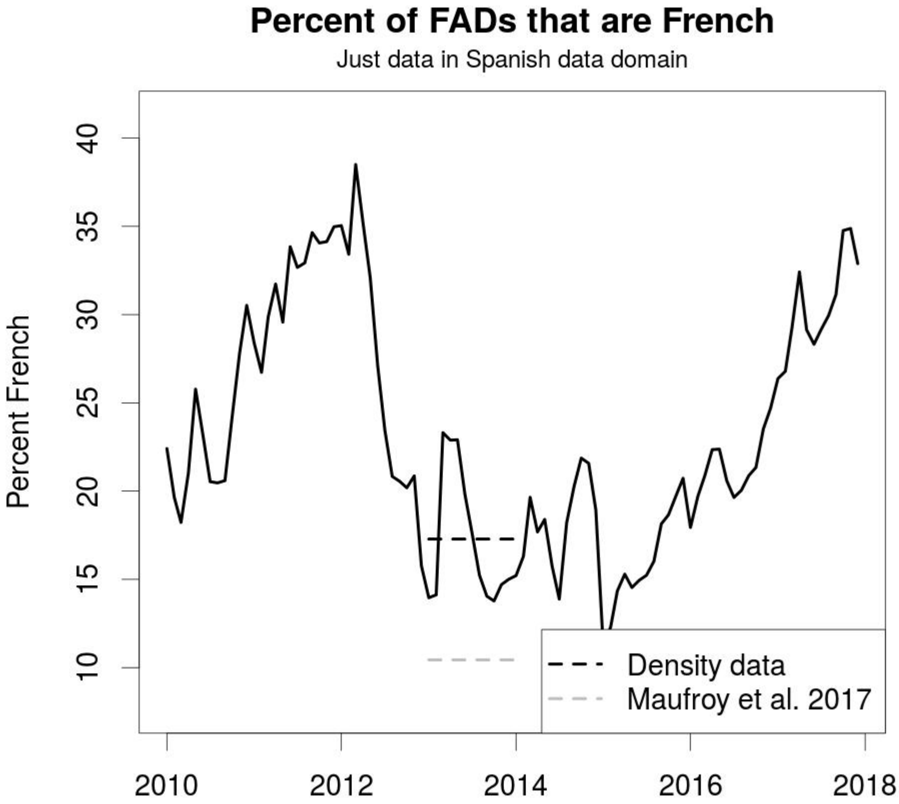

Figure 19: Proportion of EU (French and Spanish) dFADs that are French as estimated from our buoy density data. French data have been limited to the same data domain as Spanish data. Mean proportions for 2013 for our data and from Maufroy et al. (2017) are shown be horizontal dashed black and gray lines, respectively.

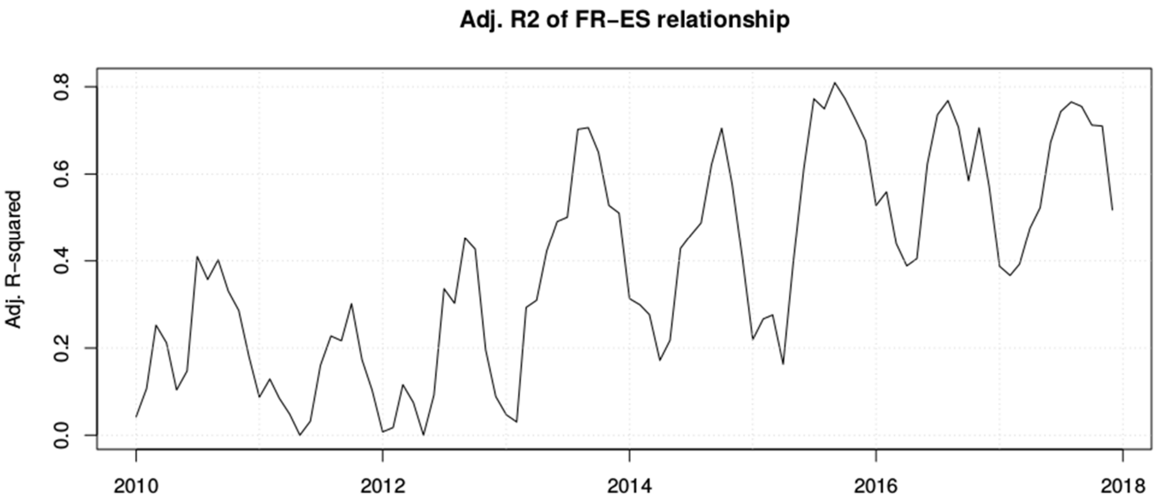

Figure 20: Adjusted R^2 of the linear relationship between French and Spanish buoy densities. Models were calculated by month using ordinary least squares based on the data for all cells for which either the Spanish or the French buoy density estimates were non-zero.

In the western central Pacific, dFAD density was found to significantly impact catch rates on FOBs despite the fraction of dFAD trajectory data coverage being only 30-40% (Escalle *et al.*, 2018). Nevertheless, dFAD density explained less deviance than the coordinate variables with the overall model explaining only 6-17% of the deviance (Escalle *et al*., 2018). Furthermore, the range of dFAD densities observed in Escalle et al. (2018) was far superior to that in our study (>3000 per 1º month maximum density in their study versus a maximum of 170 per 1º month in ours) and the authors noted that decreases in CPUE were only observed above a dFAD density of 250 per 1º month. It is, therefore, plausible that dFAD densities observed in our study may be too low for disruption of schooling behaviour (the theorized mechanism by which total FOBs densities affect FOB CPUE) to be occurring or measureable.

Numerous other explanations for the lack of an impact of dFAD densities on CPUE rates are possible. Transmitting buoy density estimates may not accurately represent FOB densities because they do not include FOBs without a buoy or with an inactive buoy. However, it has been estimated that the vast majority of FOBs in the Indian Ocean are currently dFADs with buoys (Maufroy et al. 2015) and inactive transmitting buoys were considered a relatively rare phenomenon until the implementation of dFAD management plans in 2016. Furthermore, it may be that the 1ºx1º grid cell size and the 1-month temporal resolution used for dFAD densities may be too large to accurately represent the impact of local dFAD densities on smaller spatial and temporal scales on tuna behaviour and skipper decision making in the Indian Ocean. In addition, the decision of a skipper to set on a particular dFAD does not directly depend on the overall density of that strata, but rather depends on the distribution and predicted biomass associated with the boats "own" dFADs in the region (owned density/aggregation). In other words, skipper choice is driven by information available to the skipper at the time of fishing, not the total dFAD density, which is not generally known by fishers. Though this does not address the potential of total dFAD density to disrupt schooling behaviour, it may explain fishers setting on schools in areas with a variety of total dFAD densities. Furthermore, based on the high rate of buoy ownership change (Snouck-Hurgronje *et al*. 2017), skippers may make sets on accessible dFADs with low aggregated biomass underneath to avoid losing fishing opportunities (this has been already described in 2017 WPTT report), independently of the dFAD density. Buoy density may also be impacting the decision of whether or not to fish in a zone at a given time (particularly since the advent of echosounder buoys that remotely estimate fishable biomass), as opposed to impacting how much is caught in a given set. Ideally, future studies should include variables representing the local density of dFADs owned by the individual fishing boat conducting the set (i.e., the dFAD density information directly available to the skipper) and examine the impact of dFAD density on the decision to fish in addition to the amount of fish caught to more fully explore the suite of potential mechanisms impacting purse-seine CPUE rates on FOBs.

Acknowledgements

The authors are thankful to the French purse seiner association (ORTHONGEL) and tuna owner compagnies (Compagnie Française du Thon Océanique, Saupiquet, Sapmer) and Spanish tuna purse seiner association (OPAGAC) for providing the tracking and the identification of the buoys. The workshop and this study were supported by the EU project CECOFAD2 (EASME/EMFF/2016/008). Also, thanks to the Fmort project of IATTC, who supported the attendance of one of their scientists to the workshop.

References

Bates, D., Maechler, M., Bolker, B. and Walker, S., 2014. lme4: Linear mixed-effects models using Eigen and S4. R package version, 1(7), pp.1-23.

Bertrand, A., Josse, E., Bach, P., Gros, P., and Dagorn, L.. 2002. Hydrological and trophic characteristics of tuna habitat: consequences on tuna distribution and longline catchability. *Can J Fish Aquat Sci*, 59: 1002–1013

Bigelow, K., Musyl, M.K., Poisson, F., Kleiber, P. 2006. Pelagic longline gear depth and shoaling. *Fish. Res.* 77: 173-183

Breheny, P. and Breheny, M.P., 2018. Package 'grpreg'.

Cayré, P., Marsac, F., 1993. Modelling the yellowfin tuna (Thunnus albacares) vertical distribution using sonic tagging results and local environmental parameters. *Aquat Living Resour*, 6: 1–14.

Chassot E, Guillotreau P, Kaplan DM, Vallée T (2012) The tuna fishery and piracy. In: Norchi CH, Proutière-Maulion G (eds) Piracy in comparative perspective: Problems, strategies, law. Editions A. Pedone & Hart Publishing, Oxford, United Kingdom, p 51–72

Duparc A., Cauquil P., Depetris M., Floch L., Gaertner D., Lebranchu J., Marsac F., Bach P., 2018. Assessment of accuracy in processing purse seine tropical tuna catches with the T3 methodology. IOTC-2018-WPTT-16

Escalle, L., Brouwer, S., Pilling, G., (2018) Estimates of the number of FADs active and FAD deployments per vessel in the WCPO. ECPFC-SC14-2018/MI-WP-10

Fonteneau, A., Lucas, V., Tew-Kai, E., Delgado, A., Demarcq, H., 2008. Mesoscale exploitation of a major tuna concentration in the Indian Ocean. *Aquat. Living Resour.*, 21: 109-121c

Fonteneau, A., Marsac, F. (2016). Fishery indicators suggest symptoms of overfishing for the Indian Ocean skipjack stock. IOTC-2016-WPTT18-INF02, 15p.

Friedman, J., Hastie, T. and Tibshirani, R., 2009. glmnet: Lasso and elastic-net regularized generalized linear models. *R package version*, *1*(4).

Friedman, J., Hastie, T. and Tibshirani, R., 2010. Regularization paths for generalized linear models via coordinate descent. *Journal of statistical software*, *33*(1), p.1.

Gaertner, D., Katara, I., Billet, N., Fonteneau, A., Lopez, J., Murua, H. and Daniel, P., 2017. Workshop for the development of Skipjack indices of abundance for the EU tropical tuna purse seine fishery operating in the Indian Ocean.

Getis, A. and Ord, J.K., 1992. The analysis of spatial association by use of distance statistics. *Geographical analysis*, *24*(3), pp.189-206.

Green, R.E., 1967. Relationship of the thermocline to success of purse seining for tuna. *Trans. Am. Fish. Soc*. 96(2): 126-130

Guillotreau P, Vallée T, Chassot E, Kaplan DM (2012) The economic impact of piracy on the EU purse-seine tuna fishery in the West Indian Ocean. International Workshop on "The impacts of piracy on fisheries in the Indian Ocean", 28-29 February. European Bureau for Conservation & Development, Mahé, Republic of Seychelles. February 28

Hartig, F., 2017. DHARMa: residual diagnostics for hierarchical (multi-level/mixed) regression models. R package version 0.1. 5.

Katara, I., Gaertner, D., Chassot, E., Soto, M., Abascal, F., Fonteneau, A., Floch, L., Lopez, J. and Cervantes, A., 2016. A framework for the standardisation of tropical tuna purse seine CPUE: application to the yellowfin tuna in the Indian Ocean. IOTC-2016-WPTT18-24.

Katara, I., Gaertner, D., Billet, N., Lopez, J., Fonteneau, A., Murua, H., Daniel, P. and Báez, J.C., 2017. *Standardisation of skipjack tuna CPUE for the EU purse seine fleet operating in the Indian Ocean*. IOTC-2017-WPTT19.

Lenth, R., 2018. Emmeans: Estimated marginal means, aka least-squares means. *R Package Version*, *1*(2).

Marsac, F, 2008. Outlook of ocean climate variability in the West tropical Indian Ocean, 1997-2008. IOTC-2008-WPTT-27, 9p.

Maufroy A, Chassot E, Joo R, Kaplan DM (2015) Large-Scale Examination of Spatio-Temporal Patterns of Drifting Fish Aggregating Devices (dFADs) from Tropical Tuna Fisheries of the Indian and Atlantic Oceans. PLoS ONE 10:e0128023. doi:10.1371/journal.pone.0128023

Maufroy A, Kaplan DM, Bez N, Molina D, Delgado A, Murua H, Floch L, Chassot E (2017) Massive increase in the use of drifting Fish Aggregating Devices (dFADs) by tropical tuna purse seine fisheries in the Atlantic and Indian oceans. ICES J Mar Sci 74:215–225. doi:10.1093/icesjms/fsw175

Okamoto H (2011) Preliminary analysis of the effect of piracy activity in the northwestern Indian Ocean on the CPUE trend on bigeye and yellowfin. IOTC–2011–WPTT13–44. Indian Ocean Tuna Commission, Malé, Maldives Available from http://www.iotc.org/files/proceedings/2011/wptt/IOTC-2011-WPTT13-44.pdf

Pallarès P., and Hallier J.P. 1997. Analyse du schéma d'échantillonnage multi-spécifiques des thonidés tropicaux. Rapport scientifique. IEO/ORSTOM, Programme N°95/37.

Potier, M., Marsac, F., Cherel, Y., Lucas, V., Sabatié, R., Maury, O., Ménard, F., 2007. Forage fauna in the diet of three large pelagic fishes (lancetfish, swordfish and yellowfin tuna) in the western equatorial Indian Ocean. Fish. Res 83: 60-72.

R Core Team (2017). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

Tibshirani, R., 2011. Regression shrinkage and selection via the lasso: a retrospective. J. R. Stat. Soc. Ser. B Stat. Methodol. 73, 273–282.

Tibshirani, R., 1996. Regression shrinkage and selection via the lasso. J. R. Stat. Soc. Ser. B Methodol. 267–288.

APPENDICES

FSC SETS (1986-2017 PERIOD)

Binomial GLMM (probability of large-size YFT catch > 0)

```
Generalized linear mixed model fit by maximum likelihood (Laplace Approximation) ['glmerMod']
 Family: binomial  ( logit )
Formula: cpue01 ~ annee_de_peche + mois_de_peche + CAT + lon + poly(lat,
    2, raw = TRUE) + (1 | annee_de_peche/cwpll) + (1 | tripID)
   Data: D

     AIC      BIC   logLik deviance df.resid
109408.8 109909.2 -54651.4 109302.8    93172

Scaled residuals:
    Min      1Q  Median      3Q     Max
-2.7837 -0.7220 -0.3553  0.8373 10.2110

Random effects:
 Groups               Name        Variance  Std.Dev.
 cwpll:annee_de_peche (Intercept) 4.568e-01 0.675875
 tripID               (Intercept) 2.575e-01 0.507444
 annee_de_peche       (Intercept) 9.703e-05 0.009851
Number of obs: 93225, groups:  cwpll:annee_de_peche, 9214; tripID, 4267; annee_de_peche, 32

Fixed effects:
                           Estimate Std. Error z value Pr(>|z|)
(Intercept)                -2.61066    0.36759  -7.102 1.23e-12 ***
annee_de_peche1987         -0.06150    0.11944  -0.515 0.606595
annee_de_peche1988          0.53789    0.11741   4.581 4.62e-06 ***
annee_de_peche1989         -0.17245    0.12091  -1.426 0.153804
annee_de_peche1990          0.46742    0.11795   3.963 7.40e-05 ***
annee_de_peche1991          0.65815    0.53202   1.237 0.216063
annee_de_peche1992          0.74905    0.53102   1.411 0.158368
annee_de_peche1993          0.77757    0.53084   1.465 0.142980
annee_de_peche1994          0.82191    0.53094   1.548 0.121619
annee_de_peche1995          0.60256    0.53116   1.134 0.256611
annee_de_peche1996          0.66828    0.52978   1.261 0.207152
annee_de_peche1997          0.50815    0.53023   0.958 0.337881
annee_de_peche1998          0.04966    0.53440   0.093 0.925960
annee_de_peche1999          0.64637    0.12785   5.056 4.28e-07 ***
annee_de_peche2000          0.47225    0.12450   3.793 0.000149 ***
annee_de_peche2001          0.62330    0.11872   5.250 1.52e-07 ***
annee_de_peche2002          0.40358    0.12223   3.302 0.000960 ***
annee_de_peche2003          0.90215    0.11838   7.621 2.52e-14 ***
annee_de_peche2004          0.63829    0.11963   5.336 9.52e-08 ***
annee_de_peche2005          0.76822    0.11422   6.726 1.75e-11 ***
annee_de_peche2006          0.44311    0.11625   3.812 0.000138 ***
annee_de_peche2007          0.32965    0.11881   2.775 0.005528 **
```

```
Analysis of Deviance Table (Type II Wald chisquare tests)

Response: cpue01
                          Chisq Df Pr(>Chisq)
annee_de_peche           352.99 31  < 2.2e-16 ***
mois_de_peche            611.75 11  < 2.2e-16 ***
CAT                      232.47  4  < 2.2e-16 ***
lon                      204.76  1  < 2.2e-16 ***
poly(lat, 2, raw = TRUE) 752.42  2  < 2.2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Log-Normal GLMM (catch per hour conditional to YFT catch > 0)

```
Linear mixed model fit by REML ['lmerMod']
Formula: log.cpue ~ annee_de_peche + mois_de_peche + AN_SERV + CAT + shear +
    d20 + (1 | see) + (1 | set) + (1 | cwpll:annee_de_peche) +
    (1 | vessel) + (1 | tripID) + pays + annee_de_peche:mois_de_peche
   Data: D

REML criterion at convergence: 104700.2

Scaled residuals:
    Min      1Q  Median      3Q     Max
-4.3115 -0.6007  0.0807  0.6832  3.0057

Random effects:
 Groups               Name        Variance Std.Dev.
 cwpll:annee_de_peche (Intercept) 0.11702  0.3421
 tripID               (Intercept) 0.06011  0.2452
 vessel               (Intercept) 0.01392  0.1180
 see                  (Intercept) 0.03365  0.1834
 set                  (Intercept) 0.10248  0.3201
 Residual                         1.06252  1.0308
Number of obs: 34855, groups:  cwpll:annee_de_peche, 5066; tripID, 3365; vessel, 95; see, 14; set, 13

Fixed effects:
                                 Estimate Std. Error t value
(Intercept)                    -21.728741   4.169222  -5.212
annee_de_peche1987               0.468057   0.185744   2.520
annee_de_peche1988               0.135540   0.159701   0.849
annee_de_peche1989               0.002461   0.163312   0.015
annee_de_peche1990              -0.360393   0.162818  -2.213
annee_de_peche1991               0.481150   0.361972   1.329
annee_de_peche1992              -0.277295   0.361400  -0.767
annee_de_peche1993              -0.173070   0.358080  -0.483
annee_de_peche1994               0.284982   0.351275   0.811
annee_de_peche1995              -0.337722   0.355908  -0.949
annee_de_peche1996              -0.255039   0.370414  -0.689
annee_de_peche1997              -0.100106   0.355209  -0.282
annee_de_peche1998              -0.375733   0.642752  -0.585
annee_de_peche1999              -0.054711   0.250811  -0.218
annee_de_peche2000               0.382561   0.155963   2.453
annee_de_peche2001               0.448783   0.160118   2.803
annee_de_peche2002               0.075246   0.161442   0.466
annee_de_peche2003               0.705791   0.149674   4.716
annee_de_peche2004               0.924286   0.156634   5.901
annee_de_peche2005               0.326007   0.149756   2.177
annee_de_peche2006               0.443984   0.144460   3.073
annee_de_peche2007              -0.102179   0.163136  -0.626
annee_de_peche2008               0.085084   0.143578   0.593
annee_de_peche2009               0.447653   0.170668   2.623
annee_de_peche2010               0.124287   0.201613   0.616
annee_de_peche2011               0.389679   0.187388   2.080
annee_de_peche2012               0.528521   0.177191   2.983
annee_de_peche2013              -0.147042   0.184279  -0.798
```

```
Analysis of Deviance Table (Type II Wald chisquare tests)

Response: log.cpue
                                Chisq  Df Pr(>Chisq)
annee_de_peche               342.1307  31  < 2.2e-16 ***
mois_de_peche                193.9904  11  < 2.2e-16 ***
AN_SERV                       27.0932   1  1.939e-07 ***
CAT                            8.1288   4   0.086971 .
shear                          2.3336   1   0.126605
d20                            7.8359   1   0.005122 **
pays                           1.9038   1   0.167649
annee_de_peche:mois_de_peche 1345.3326 337  < 2.2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

FOB-RELATED SETS (2007-2017 PERIOD)

Log-Normal GLMM (catch per set conditional to YFT catch > 0)

```
Linear mixed model fit by REML ['lmerMod']
Formula: log.cpue ~ categorie + annee_de_peche * mois_de_peche + d20 +
    shear + lat * lon + PUIS.CV + CAP.M3 + FADFSCHRatioTrip +      pays + (1 | zee) + (1 | zet) + (1 | annee_de_peche/cwp11) +
    (1 | vessel) + (1 | tripID)
   Data: dp

REML criterion at convergence: 114908

Scaled residuals:
    Min      1Q  Median      3Q     Max
-3.0785 -0.6858  0.0082  0.6789  3.6216

Random effects:
 Groups               Name        Variance Std.Dev.
 cwp11:annee_de_peche (Intercept) 0.027289 0.16520
 tripID               (Intercept) 0.028721 0.16947
 vessel               (Intercept) 0.005669 0.07529
 zet                  (Intercept) 0.053751 0.23184
 zee                  (Intercept) 0.022865 0.15121
 annee_de_peche       (Intercept) 0.379156 0.61576
 Residual                         0.700385 0.83689
Number of obs: 45094, groups:  cwp11:annee_de_peche, 4787; tripID, 1933; vessel, 50; zet, 13; zee, 13; annee_de_peche, 11

Fixed effects:
                                     Estimate Std. Error t value
(Intercept)                          0.799010   0.628938   1.270
categorie7                          -0.071644   0.042233  -1.696
categorie8                          -0.157346   0.082581  -1.905
annee_de_peche2008                  -0.118680   0.883540  -0.134
annee_de_peche2009                   0.830490   0.875550   0.949
annee_de_peche2010                   0.883017   0.875178   1.009
annee_de_peche2011                   1.266789   0.875750   1.447
annee_de_peche2012                   0.922854   0.875453   1.054
annee_de_peche2013                   1.221385   0.875485   1.395
annee_de_peche2014                   1.152869   0.875083   1.317
annee_de_peche2015                   0.848171   0.876244   0.968
annee_de_peche2016                   0.874539   0.874775   1.000
annee_de_peche2017                   0.585293   0.874704   0.669
mois_de_peche2                       0.246091   0.088950   2.767
mois_de_peche3                       0.103069   0.088973   1.158
mois_de_peche4                       0.340215   0.086597   3.929
mois_de_peche5                       0.185893   0.091441   2.033
mois_de_peche6                       0.243143   0.110418   2.202
mois_de_peche7                       0.462283   0.085793   5.388
mois_de_peche8                       0.350419   0.080847   4.334
mois_de_peche9                       0.329926   0.080442   4.101
mois_de_peche10                      0.276652   0.084519   3.273
mois_de_peche11                      0.162962   0.088419   1.843
mois_de_peche12                      0.164741   0.090300   1.824
d20                                  0.047676   0.010478   4.550
shear                                0.010352   0.005915   1.750
lat                                  0.096747   0.019656   4.922
lon                                 -0.070577   0.010058  -7.017
PUIS.CV                              0.081956   0.029514   2.777
CAP.M3                               0.051359   0.039361   1.305
FADFSCHRatioTrip                     0.014527   0.006749   2.153
pays4                                0.125268   0.036187   3.462
annee_de_peche2008:mois_de_peche2   -0.123022   0.208660  -0.590
annee_de_peche2009:mois_de_peche2   -0.448797   0.123098  -3.646
annee_de_peche2010:mois_de_peche2   -0.289527   0.119080  -2.431
annee_de_peche2011:mois_de_peche2   -0.451104   0.126188  -3.575
annee_de_peche2012:mois_de_peche2   -0.335417   0.128459  -2.611
```

```
Analysis of Deviance Table (Type II Wald chisquare tests)

Response: log.cpue
                                 Chisq  Df Pr(>Chisq)
categorie                       3.7924   2  0.1501385
annee_de_peche                  1.5823  10  0.9986553
mois_de_peche                 182.8495  11  < 2.2e-16 ***
d20                            20.7038   1  5.361e-06 ***
shear                           3.0631   1  0.0800885 .
lat                            21.3516   1  3.823e-06 ***
lon                            64.1955   1  1.127e-15 ***
PUIS.CV                         7.7109   1  0.0054889 **
CAP.M3                          1.7025   1  0.1919559
FADFSCHRatioTrip                4.6334   1  0.0313556 *
pays                           11.9834   1  0.0005368 ***
annee_de_peche:mois_de_peche 1018.1658 110  < 2.2e-16 ***
lat:lon                        39.1262   1  3.973e-10 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```