Received: 3 September, 2012

UPDATE OF THE STANDARDIZED CPUE SERIES FOR MAJOR SHARK SPECIES CAUGHT BY THE PORTUGUESE PELAGIC LONGLINE FISHERY IN THE INDIAN OCEAN

Rui Coelho*, Miguel N. Santos & Pedro G. Lino

SUMMARY

Portuguese longliners targeting swordfish and operating in the Indian Ocean regularly capture elasmobranch fishes as bycatch. Of those, the blue shark (*Prionace glauca*) and the shortfin mako (*Isurus oxyrinchus*) constitute the two main shark species captured. A recent effort by *IPMA* (*Portuguese Sea and Atmospheric Institute*) has been recovering historical catch data on elasmobranchs captured since the late 1990’s to the present date in that fishery. Nominal CPUEs for these two major sharks were calculated as Kg /1000 hooks and standardized with Generalized Linear Models (GLMs). Several different modeling techniques were tested and compared, chosen depending on the specific proportion of zeros in the catch data for each species. The models tested included the delta method, tweedie, gamma and lognormal models. Model validation was carried out with residual analysis, and relative indexes of abundance for the two species were calculated. The results presented in this paper update a previous analysis on the trends of elasmobranch catch rates available from the Portuguese longline fishery operating in the Indian Ocean.

KEYWORDS: Bycatch, CPUE standardization, generalized linear models, longlines, Prionace glauca, Isurus oxyrinchus.

Authors Addresses: Instituto Português do Mar e Atmosfera (IPMA), Av. 5 de Outubro s/n, 8700-305 Olhão, PORTUGAL
*: Corresponding author e-mail: rpcoelho@ualg.pt

1. Introduction

The Portuguese pelagic longline fishery in the Indian Ocean started in the late 1990's and has traditionally targeted swordfish even though, in certain areas and seasons, it also catches relatively high quantities of sharks as bycatch. Of the pelagic shark species usually captured, the blue shark (BSH, *Prionace glauca*), and the shortfin mako (SMA, *Isurus oxyrinchus*) are the most frequent, and together can account for more than 90% of the total elasmobranch catch. The Portuguese fishing vessels operating in the IOTC area consist of pelagic longliners ranging in size from 35 to over 50m. The number of vessels licensed increased from the beginning of the fishery in 1998 (5 vessels) until 2009 (24 vessels), and the number of active vessels followed a similar trend, with a peak in 2006 (17 vessels). However, during the last 5 years, the active vessels in the convention area decreased to as low as three in 2009, with the reasons for such decrease mainly related with the increase of the exploitation costs (particularly oil prices in late 2000's), but also due to recent piracy related problems in the SW Indian Ocean.

In 2011 several papers were presented to the *IOTC Working Party on Ecosystems and Bycatch* describing the catches of elasmobranchs in that fishery and carrying out a preliminary analyzing. The papers presented included an analysis of the fin to body ratios for blue shark (Santos et al., 2011a), an overview of the fishery including shark bycatches (Santos, et. al., 2011b), an estimation of at-haulback mortalities for all elasmobranch species captured (Coelho et al., 2011a), and a preliminary CPUE standardization for the two major shark species (Coelho et al., 2011b).

The objectives of this paper are to present an update of the standardized CPUE time series for the two main shark species captured by the Portuguese pelagic longline fishery targeting swordfish in the Indian Ocean, specifically the blue shark and the shortfin mako.

2. Material and methods

2.1. Data collection

In a recent effort by the *Portuguese Sea and Atmosphere Institute* (*IPMA*, former INRB I.P./IPIMAR), the historical catch data from the Portuguese longliners targeting swordfish in the Indian Ocean started to be compiled and analyzed. Information on effort (number of hooks used per set) is available for most of the fishing sets, with the exception of the first year in the time series (1998) for which information on effort is not available. For this reason the time series analyzed in this paper refers to the years 1999-2011 (**Table 1**).

Table 1: Number of fishing sets with catch and effort information carried out by the Portuguese longline fleet in the Indian Ocean. The percentage of sets per year analyzed for this paper is indicated.

Year	Sets with catch information (N)	Sets with effort information (N)	% used for analysis
1998	113	0	0.0
1999	257	205	79.8
2000	340	333	97.9
2001	701	443	63.2
2002	877	578	65.9
2003	867	525	60.6
2004	756	495	65.5
2005	900	656	72.9
2006	2265	1931	85.3
2007	1739	1505	86.5
2008	360	360	100.0
2009	525	525	100.0
2010	630	623	98.9
2011	633	633	100.0

2.2. Data analysis

The CPUE analysis was carried out using the official fisheries statistics collected by the Portuguese Fisheries authorities, and the catch data refers to the total weight of the major shark species (blue shark and shortfin mako) captured per fishing set. The general location in the datasets use the FAO fishing Areas (47, 51 and 57) that is available for the entire time series, while starting in 2005 more detailed information (FAO Subareas) also started to be collected. Currently, *IPMA* is making an effort to also compile VMS (*Vessel Monitoring Systems*) and skippers' logbook data, aiming to further complete the analysis with more detailed catch, effort and geographical coverage of the fishery, particularly during the early years of the fishery.

For the CPUE standardization, the response variable considered for this study was Catch per Unit of Effort (CPUE), measured as biomass of live fish (kg) per 1,000 hooks deployed. The standardized CPUEs were estimated with Generalized Linear Models (GLM). For both the BSH and the SMA analysis, there were some fishing sets with zero catches, that results in a response variable of CPUE=0. As these zeros can cause mathematical problems for fitting the models, various methodologies were applied in a comparative way, and depending on the specific situation.

For the BSH, and because in this species the proportion of fishing sets with zero catches was relatively small, the different approaches tested and compared were tweedie, gamma and lognormal models. For the tweedie models the nominal CPUEs were used directly, given that this distribution can handle a certain proportion of zeros. For the gamma and lognormal models the response variable was defined as the nominal CPUE + constant (δ), with δ set to 10% of the overall mean catch rate. This value was recommended by Campbell (2004) as it seems to minimize the bias for this type of adjustments. Further, and in a comparative study, Shono (2008) showed that when the percentage of zeros in the dataset is relatively low (<10%), the method of adding a constant to the response variable performs relatively well.

For the SMA the proportion of fishing sets with zero catches was substantially higher, and it has been shown that in those situations simply adding a constant to the response variable may introduce significant bias in the analysis. Therefore, for this species the two approaches tested and compared were methods that can handle zeros in the response variable, namely tweedie models and the delta method. With the delta method two separate models were estimated and in this case the first model was a binomial model with a logit link function (that is used to model the proportion of fishing sets with positive catches), and the second model was a lognormal model for the nominal CPUEs of the positive sets. The final standardized CPUEs were estimated by least square means (LSMeans), calculated as the yearly probability of having a positive set multiplied by the expected catch rate conditional to the set being positive.

On all models tested, the explanatory variables that were considered were:

- Year (analyzed between 1999 and 2011);
- Quarter of the year (4 categories: 1 = January to March, 2 = April to June, 3 = July to September, 4 = October to December);
- Region (FAO Regions);
- Ratio (Based on the SWO/SWO+BSH ratio of captures).

The "ratio" was defined as the percentage of swordfish catches related to combined swordfish and blue shark catches. This ratio is in general considered as a good proxy indicator of target criteria more clearly directed at swordfish *vs.* a more diffuse fishing strategy aimed at the two main species (SWO and BSH). Moreover, it has been consistently applied to other fleets that have similar methods of operation, such as the Spanish fleet, with applications both to the Atlantic and the Indian Ocean (Ramos-Cartelle et al., 2011; Mejuto et al., 2012). The ratio was calculated for each set and then categorized into ten categories using the 10% percentiles.

The significance of the explanatory variables was assessed with likelihood ratio tests comparing each univariate model to the null model (considering a significance level of 5%), and by analyzing the deviance explained by each covariate. Goodness-of-fit and model validation was carried out with a residual analysis. The final estimated indexes of abundance were calculated by scaling the annual standardized CPUE values by the mean standardized CPUE in the time series.

Statistical analysis for this paper was carried out with the R Project for Statistical Computing version 2.14.1 (R Development Core Team, 2011; Fox and Weisberg, 2011; Dunn, 2011; Warnes, 2011; Højsgaard and Halekoh, 2012).

3. Results and Discussion

3.1. Catch and effort

The total effort of the Portuguese longline fleet in the Indian Ocean remained relatively constant between 1999 and 2004, followed by an increase during 2006-2007. For the more recent years of 2008 to 2011 the effort was again similar to the initial years of the early 2000's (**Figure 1**). The total nominal BSH and SMA catches tended to follow this general trend in the effort, with a peak in the catches of the two species also during those same years (**Figure 1**). In terms of ratios between SWO compared to the SWO + BSH catches, a generally increasing trend was observed during the time period (**Figure 1**).

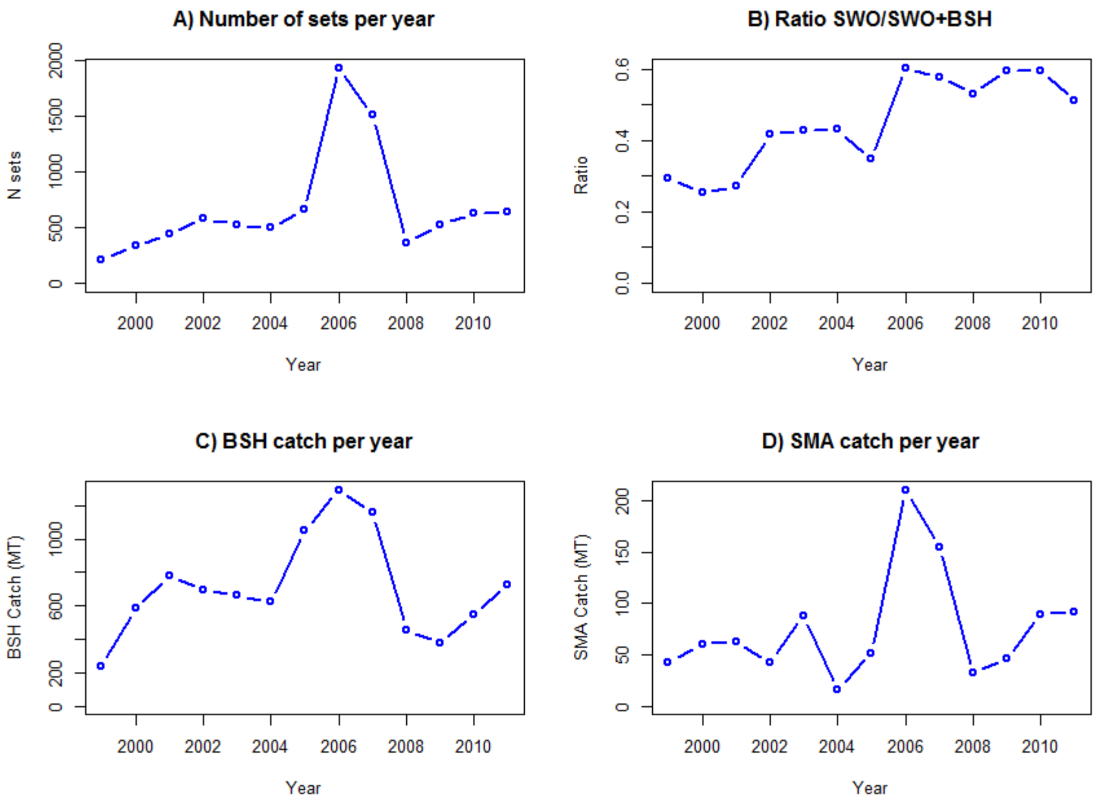

Figure 1: Descriptive plots of the total effort in sets (A), the ratio of swordfish compared to the swordfish and blue shark catches (B), the total nominal catch of blue shark (C) and the total nominal catch of shortfin mako (D), for the Portuguese longline fleet operating in the Indian Ocean.

3.2. CPUE analysis for blue shark

In terms of nominal CPUEs for the BSH, there were relatively large oscillations in the time series, with a general decreasing trend observed during the period (**Figure 2**).

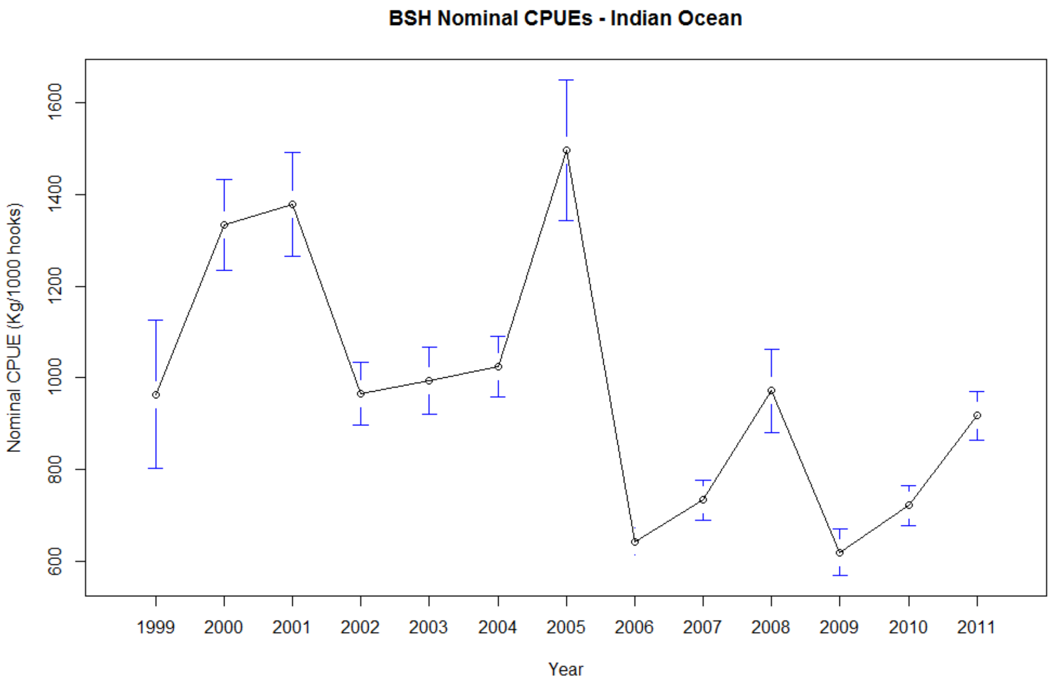

Figure 2. Nominal BSH CPUEs (Kg/1000 hooks) for the Portuguese pelagic longline fishery in the Indian Ocean, between 1999 and 2011.

The percentage of fishing sets with zero catches of BSH in the dataset was 2.9%, and this relatively low value of sets with zero catches seems to be common for this species, that appears with a relatively high frequency of occurrence in the catches. The distribution of the nominal BSH CPUE data was highly asymmetrical and skewed to the right (**Figure 3**).

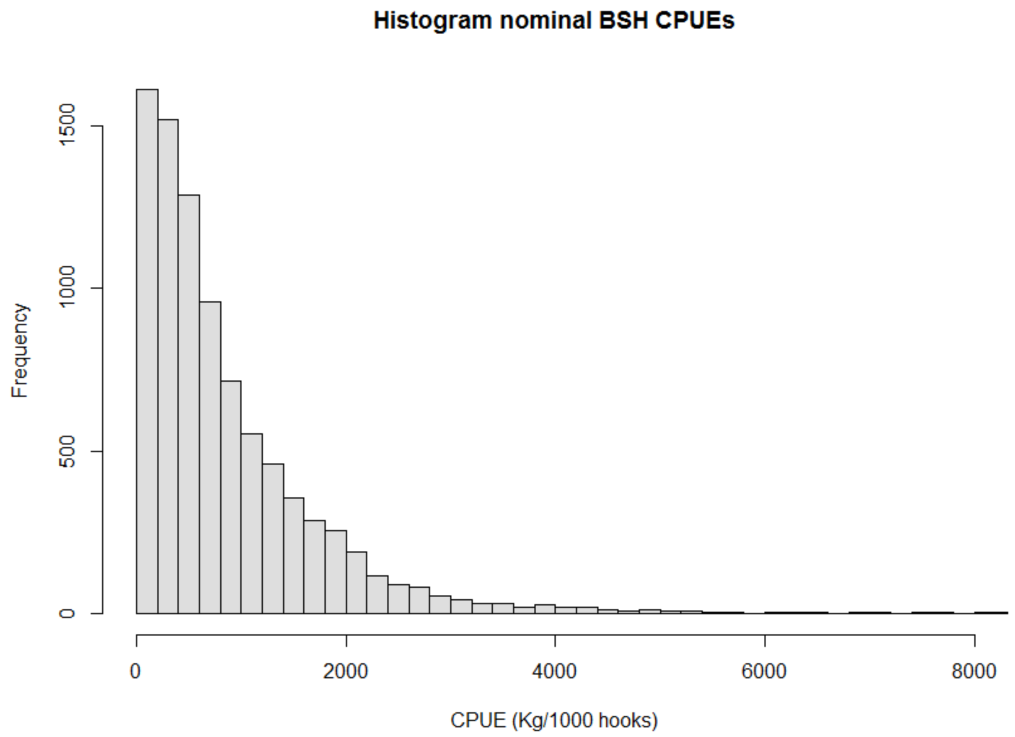

Figure 3: Distribution of the nominal blue shark nominal CPUE data from the Portuguese longline fishery in the Indian Ocean.

On all models tested, the explanatory variables initially used all contributed for explaining part of the deviance of the BSH CPUE variability, as well as the interaction between Area and Season, and therefore those variables were used in the models (**Table 2**).On all the models tested, the variables that contributed more for explaining the deviance were the Ratio, followed by the Year and the Quarter. The factors Area and the interaction between Area and Season were significant, but contributed less to the models (**Table 2**).

Table 2. Deviance of the parameters used for the different BSH CPUE standardization models (tweedie, gamma and lognormal). For each parameter it is indicated the degrees of freedom used, the deviance explained, the residual degrees of freedom and deviance after incorporating each parameter and the significance (*p-value*) of each parameter. For each model it is also indicated the coefficient of determination value (R^2).

Tweedie Model (R^2 = 69.4%)					
Parameter	Df	Deviance	Resid. Df.	Resid. deviance	Significance (*p-value*)
Null			8811	319623	
Year	12	29135	8799	290488	< 0.001
Quarter	3	35913	8796	254575	< 0.001
Area	2	431	8794	254144	< 0.001
Ratio	9	156467	8785	97677	< 0.001
Quarter:Area	6	265	8779	97412	< 0.001
Gamma Model (R^2 = 72.3%)					
Parameter	Df	Deviance	Resid. Df.	Resid. deviance	Significance (p-value)
Null			8811	6670	
Year	12	601	8799	6070	< 0.001
Quarter	3	747	8796	5323	< 0.001
Area	2	7	8794	5316	< 0.001
Ratio	9	3469	8785	1847	< 0.001
Quarter:Area	6	5	8779	1843	0.003
Log-Normal Model (R^2 = 72.2%)					
Parameter	Df	Deviance	Resid. Df.	Resid. deviance	Significance (p-value)
Null			8811	6986.6	
Year	12	586	8799	6400.6	< 0.001
Quarter	3	878	8796	5523	< 0.001
Area	2	9	8794	5514	< 0.001
Ratio	9	3578	8785	1936.4	< 0.001
Quarter:Area	6	4.2	8779	1932.3	0.004

In terms of model validation, the three tested models seemed relatively adequate for this particular situation with a relatively small quantity of zero values in the CPUEs. The residual analysis for the models tested, including the residuals distribution along the fitted values, the QQ plots and the residuals histograms did not identified any major problems in the residuals analysis (**Figure 4**). Some potential outliers were identified, but given the preliminary nature of these models those outliers were not excluded from the final models at this stage.

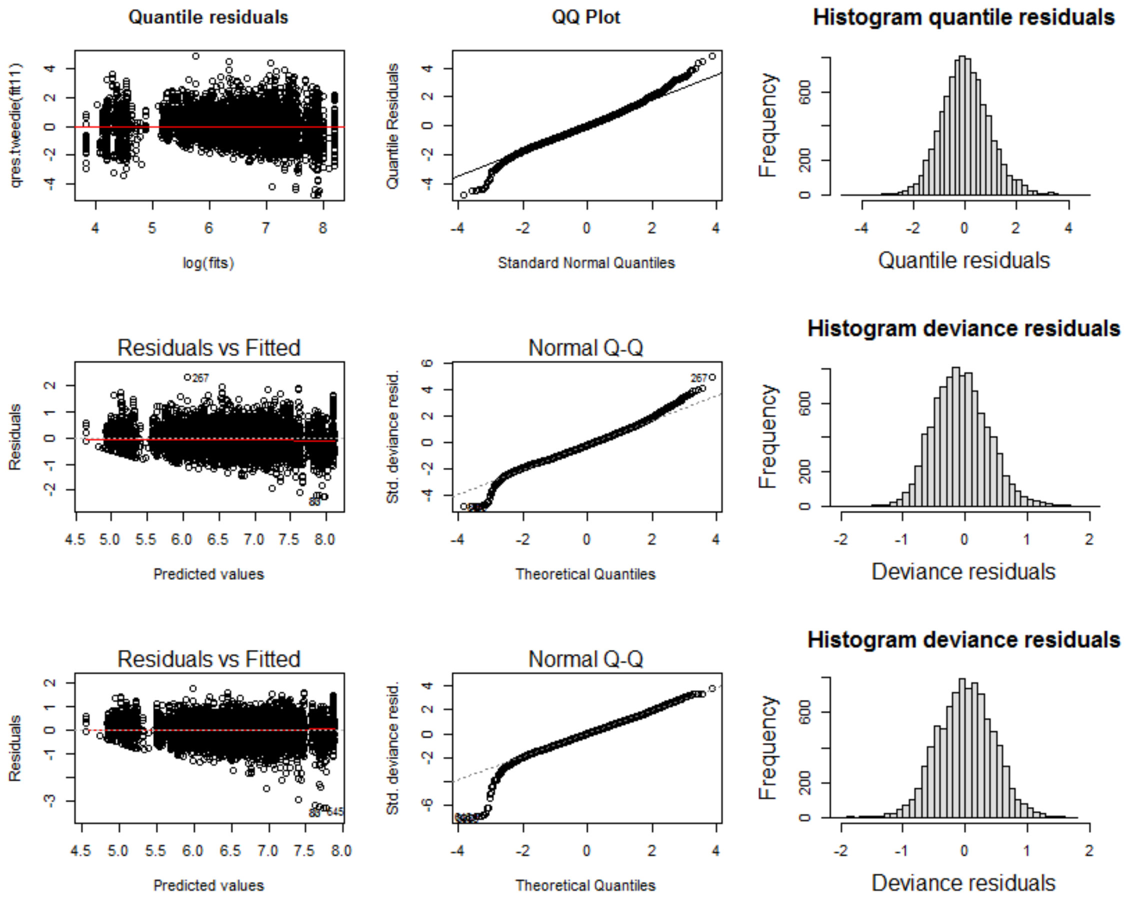

Figure 4. Residual analysis for the models tested for the BSH CPUE standardization, namely the tweedie model on the top, the gamma model in the middle, and the lognormal model on the bottom. For each model it is presented the residuals along the fitted values (graphics on the left), the QQPlot (graphics on the middle), and the histogram of the distribution of the residuals (graphics on the right).

In general, the standardized index of abundance of BSH in the Indian Ocean seems to have remained relatively stable between the time period considered (**Figure 5**). In terms of model comparisons, the results of the three tested models (tweedie, gamma and lognormal) produced on all cases very similar results, especially the tweedie and the gamma that followed almost exactly the same trend. The lognormal also followed a general trend very similar to the other two models, with only some minor differences (**Figure 5**).

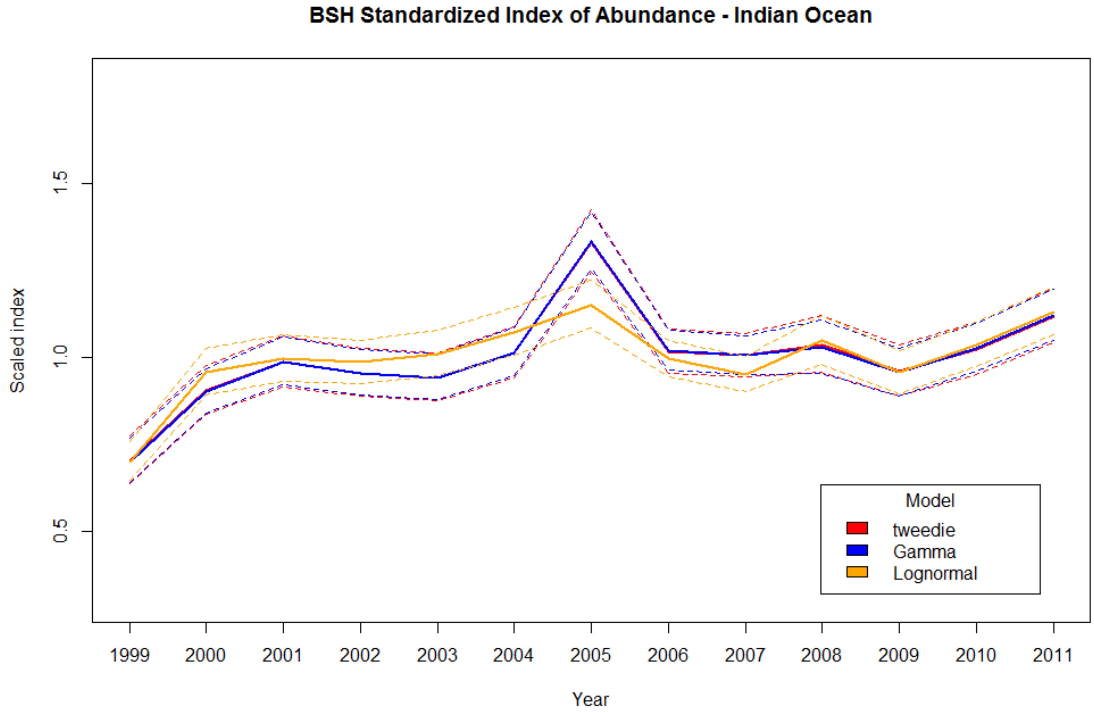

Figure 5. Scaled annual index of abundance for BSH captured by the Portuguese pelagic longline fleet in the Indian Ocean. The solid lines refer to the standardized series calculated with the different models, and the dotted lines refer to the respective 95% confidence intervals.

3.3. CPUE analysis for the shortfin mako

In terms of nominal CPUEs, a general decreasing trend was observed in the early years of the time series between 1999 and 2002, followed by a sequence of years with relatively large oscillations (2002-2004), and then followed by a period with a general increasing trend for the more recent years (2004-2011) (**Figure 6**).

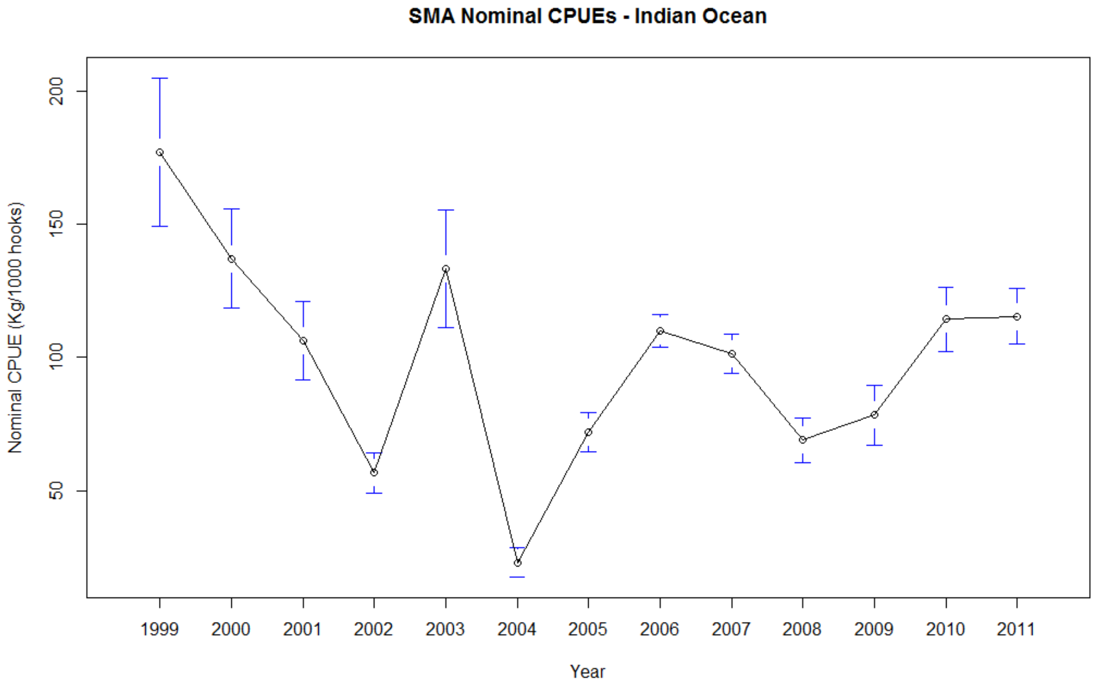

Figure 6. Nominal SMA CPUEs (Kg/1000 hooks) for the Portuguese pelagic longline fishery in the Indian Ocean, between 1999 and 2011.

The percentage of fishing sets with zero catches of SMA in the dataset was much higher than for the BSH, specifically 40.5%, and this relatively high value seems to be common for this species captured as bycatch in longline fisheries. Also for the Portuguese longline fleet but for the Atlantic Ocean, values of 60.1% and 59.4% were reported respectively for the North and South Atlantic regions (Coelho et al., 2012). The distribution of the nominal SMA CPUE data was also highly asymmetrical and skewed to the right (**Figure 7**).

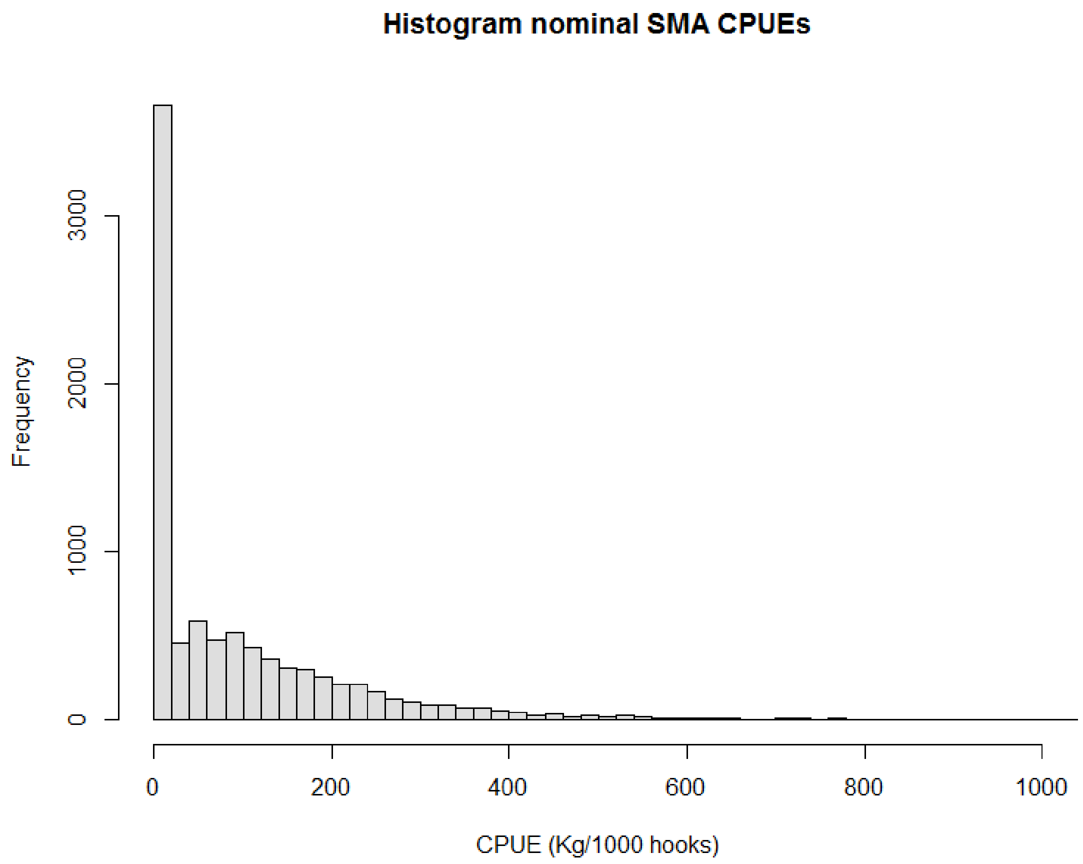

Figure 7: Distribution of the nominal shortfin mako CPUE data from the Portuguese longline fishery in the Indian Ocean.

On the SMA models tested, all the explanatory variables initially used contributed significantly for explaining part of the CPUEs variability, but the interaction between Area and Season was not significant. Therefore, for this species the full simple effects model was used (**Table 3**).On all the models tested, the variables that contributed more for explaining part of the deviance were the Ratio, followed by the Year and the Quarter (**Table 3**).

Table 3. Deviance of the parameters used for the different SMA CPUE standardization models (delta method and tweedie model). For each parameter it is indicated the degrees of freedom used, the deviance explained, the residual degrees of freedom and deviance after incorporating each parameter and the significance (*p-value*) of each parameter. For each model it is also indicated the coefficient of determination value (R^2).

1) DELTA LOGNORMAL					
1.1) Binomial model for the catch/no catch **(R^2=4.9%)**					
Parameter	Df	Deviance	Resid. Df.	Resid. deviance	Significance (*p-value*)
Null			8811	11894	
Year	12	733.06	8799	11161	< 0.001
Quarter	3	62.22	8796	11098	< 0.001
Region	2	37.83	8794	11061	< 0.001
Ratio	9	54.19	8785	11006	< 0.001
1.2) Lognormal model for the positive sets **(R^2=7.2%)**					
Parameter	Df	Deviance	Resid. Df.	Resid. deviance	Significance (*p-value*)
Null			5245	3527	
Year	12	85.444	5233	3441.5	< 0.001
Quarter	3	26.69	5230	3414.9	< 0.001
Region	2	4.844	5228	3410	0.023
Ratio	9	72.041	5219	3338	< 0.001
2) TWEEDIE MODEL (R^2=7.3%)					
Parameter	Df	Deviance	Resid. Df.	Resid. deviance	Significance (*p-value*)
Null			8811	327457	
Year	12	18805.4	8799	308652	< 0.001
Quarter	3	1461.1	8796	307191	< 0.001
Region	2	477.1	8794	306714	0.002
Ratio	9	3985	8785	302729	< 0.001

In terms of model validation, the two tested techniques (delta method and tweedie model) seemed relatively adequate for this particular situation with a relatively large quantity of zero values in the CPUEs. The residual analysis for the models tested, including the residuals distribution along the fitted values, the QQ plots and the residuals histograms did not identified any major problems in the residuals (**Figure 8**). Some potential outliers were identified, but given the preliminary nature of these models those outliers were not excluded from the final models at this stage.

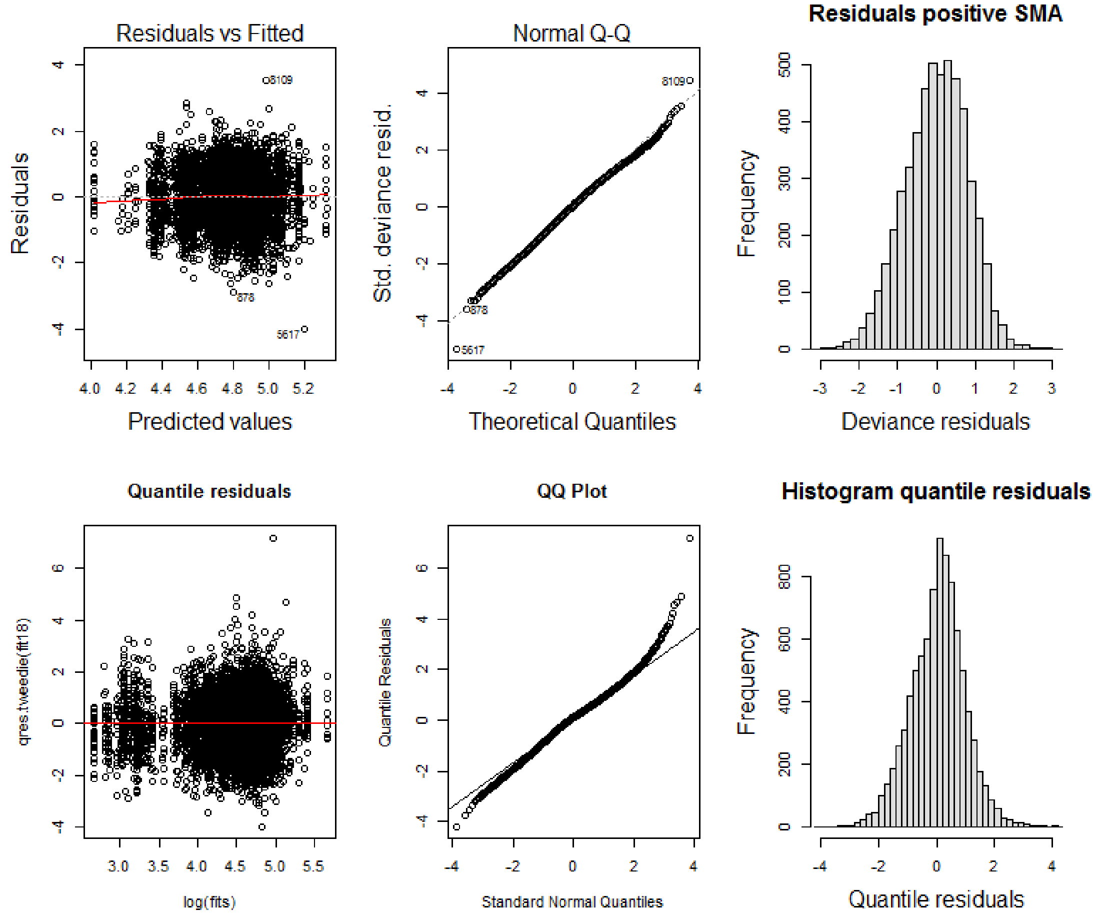

Figure 8. Residual analysis for the models tested for the SMA CPUE standardization, namely the delta lognormal (positive sets model) on the top, and the tweedie model on the bottom. For each model, it is presented the residuals along the fitted values (graphics on the left), the QQPlot (graphics on the middle), and the histogram of the distribution of the residuals (graphics on the right).

In general, the standardized index of abundance of SMA in the Indian Ocean shows some variability, with a general decreasing trends in the initial years of the time series (1999-2004), and then followed by a general increasing trend for the more recent years, until 2011 (**Figure 9**). In terms of model comparisons, the results of the delta lognormal approach and the tweedie model produced very similar results and trends in the series (**Figure 9**).

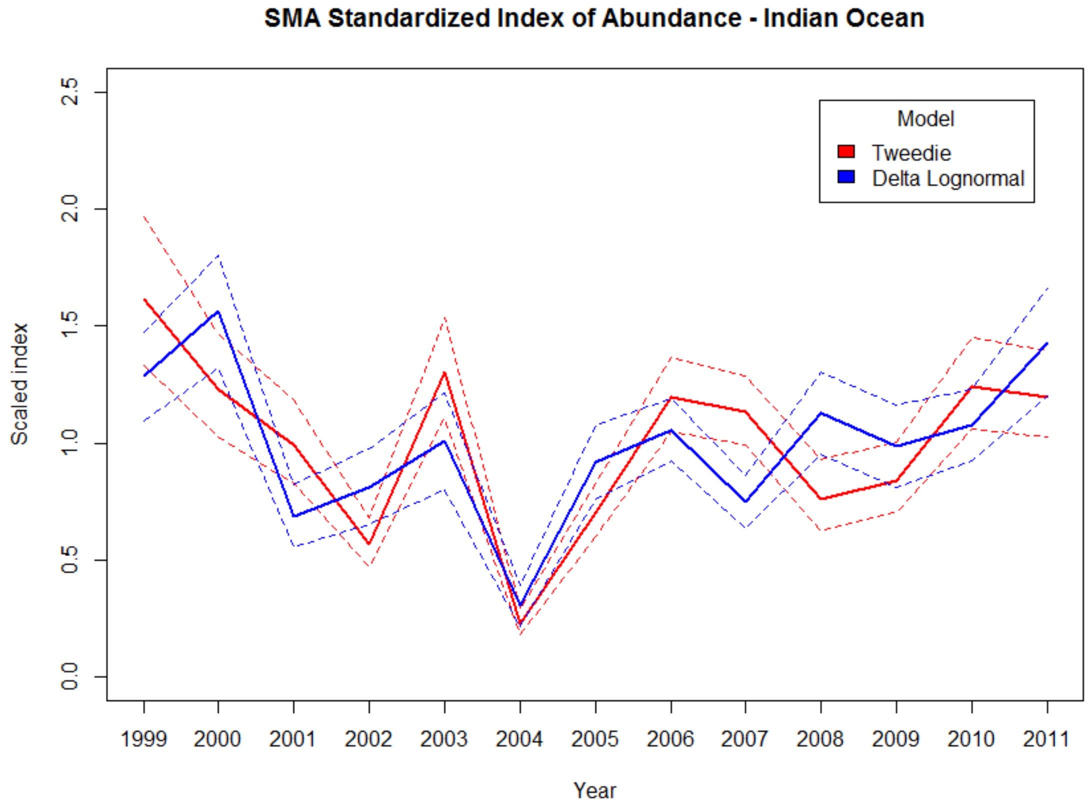

Figure 9. Scaled annual index of abundance for SMA captured by the Portuguese pelagic longline fleet in the Indian Ocean. The solid lines refer to the standardized series calculated with the two different models, and the dotted lines refer to the respective 95% confidence intervals.

4. Acknowledgments

All the data used was collected for within the EU Data Collection Framework. Rui Coelho was supported with a grant from FCT (Ref: BPD 40523 / 2007) co-funded by "POCI-2010 - Programa Operacional Ciência e Inovação 2010" and "FSE - Fundo Social Europeu". Special thanks are due to all skippers that voluntarily provided data from their fishing activities in the IOTC convention area.

5. References

Campbell, R.A. 2004. CPUE standardisation and the construction of indices of stock abundance in a spatially varying fishery using general linear models. *Fisheries Research*, 70: 209–227.

Coelho, R., Lino, P.G., Santos, M.N. 2011a. At-haulback mortality of elasmobranchs caught on the Portuguese longline swordfish fishery in the Indian Ocean. *IOTC Working Document*, IOTC–2011–WPEB07–31.

Coelho, R., Lino, P.G., Santos, M.N. 2011b. Standardized CPUE of major shark species caught by the Portuguese longline fishery in the Indian Ocean. *IOTC Working Document*, IOTC–2011–WPEB07–36.

Coelho, R., Lino, P.G., Santos, M.N. 2012. Standardized CPUE for the shortfin mako (*Isurus oxyrinchus*) caught by the Portuguese pelagic longline fishery. *ICCAT Standing Committee on Research and Statistics Document*, SCRS/2012/072.

Dunn, P.K. 2011). tweedie: Tweedie exponential family models. R package version 2.1.1.

Fox, J., Weisberg, S. 2011. An R Companion to Applied Regression, 2nd Edition. Sage, Thousand Oaks CA.

Højsgaard, S. and Halekoh, U. 2012. doBy: Groupwise summary statistics, general linear contrasts, population means (least-squares-means), and other utilities. R package version 4.5.0. http://CRAN.R-project.org/package=doBy.

Mejuto, J., García-Cortés, B., Ramos-Cartelle, A., De la Serna, J.M., González-González, I. 2012. Standardized catch rates of shortfin mako (*Isurus oxyrinchus*) caught by the Spanish surface longline fishery targeting swordfish in the Atlantic Ocean during the period 1990-2010. *ICCAT SCRS/2012/046*. 12pp.

Ramos-Cartelle, A., García-Cortés, B., Fernández-Costa, J., Mejuto, J. 2011. Standardized catch rates for the swordfish (*Xiphias gladius*) caught by the Spanish longline in the Indian Ocean during the period 2011-2012. *IOTC Working Document*. IOTC-2011-WPB09-23. 19pp.

R Development Core Team. 2011. R: A language and environment for statistical computing. Version 2.14.1. R Foundation for Statistical Computing, Vienna, Austria. http://www.R-project.org.

Santos, M.N., Coelho, R., Fernandez-Carvalho, J., Lino, P.G. 2011a. Observations on the ratio between fin and body weights for the blue shark caught by the Portuguese longline fleet in the Indian Ocean. *IOTC Working Document*, IOTC–2011–WPEB07–37.

Santos, M.N., Lino, P.G., Fernandez-Carvalho, J., Coelho, R. 2011b. Preliminary observations on the by-catch of elasmobranchs caught by the Portuguese longline fishery in the Indian Ocean: biology, ecology and fisheries. *IOTC Working Document*, IOTC–2011–WPEB07–30.

Shono, H. 2008. Application of the Tweedie distribution to zero-catch data in CPUE analysis. *Fish. Res.*, 93: 154–162

Warnes, G.R. 2011. gplots: various R programming tools for plotting data. R package version 2.10.1. http://CRAN.R-project.org/package=gplots.