STANDARDIZED CATCH PER UNIT EFFORT (CPUE) OF SHORTFIN MAKO (*ISURUS OXYRINCHUS*) FOR THE MOROCCAN LONGLINE FISHERY

M. Serghini[1]*, H. Moustahfid[2], H. Habiba[1], L. Aziza[1], N. Abid[1], S. Baibat[1]

SUMMARY

Shortfin mako shark, Isurus oxyrinchus*, is harvested as bycatch by the Moroccan longliners targeting swordfish,* Xiphias gladius*, in the south of Moroccan Atlantic waters. A time series of standardized catch per unit effort (CPUE) for shortfin mako was estimated by first analyzing the fleet dynamic and identification of fishing tactics using multi-table method, and then using two statistical models, including Generalized Linear Models (GLM) and Boosted Regression Trees model (BRT) with main effects and two-way interactions. BRT with two-way interactions was selected as the best model to estimate CPUE with less RMSE and high PDE. The standardized CPUE analysis indicates a declining trend since the early years and slight increase and stability in the last four years of the time series.*

RÉSUMÉ

*Le requin-taupe bleu (*Isurus oxyrinchus*) est capturé comme prise accessoire par les palangriers marocains ciblant l'espadon (*Xiphias gladius*) au sud des eaux marocaines. Une série temporelle de capture par unité d'effort standardisée (CPUE) pour le requin-taupe bleu a été estimée en analysant d'abord la dynamique de la flottille et l'identification des tactiques de pêche en utilisant une méthode à tableaux multiples, puis en utilisant deux modèles statistiques, y compris des modèles linéaires généralisés (GLM) et un modèle à arbre de régression augmentée (BRT) avec des effets principaux et des interactions à double sens. Le BRT avec des interactions à double sens a été sélectionné comme le meilleur modèle pour estimer la CPUE avec moins de RMSE et une PDE élevée. L'analyse de la CPUE standardisée indique une tendance à la baisse depuis les premières années et une légère augmentation et stabilité au cours des quatre dernières années de la série temporelle.*

RESUMEN

*El marrajo dientuso (*Isurus oxyrinchus*) es capturado de forma fortuita por los palangreros marroquíes que dirigen su actividad al pez espada (*Xiphias gladius*) en la parte meridional de las aguas marroquíes. Se estimó una serie de captura por unidad de esfuerzo (CPUE) estandarizada para el marrajo dientuso, analizando en primer lugar la dinámica de la flota e identificando las tácticas de pesca utilizando un método de tablas múltiples y usando después dos modelos estadísticos, modelos lineales generalizados (GLM) y el modelo de árbol de regresión potenciado (BRT) con efectos principales e interacciones en dos sentidos. Se eligió BRT con interacciones en dos sentidos como el mejor modelo para estimar la CPUE con menos RMSE y PDE elevada. El análisis de la CPUE estandarizada indica una tendencia descendente desde los primeros años y un ligero incremento y estabilidad en los cuatro últimos años de la serie temporal.*

KEYWORDS

Shortfin mako, Isurus oxyrinchus*, catch per unit effort, CPUE, fleet dynamic, fishing tactics, multi-table analysis, boosted regression trees, generalized linear model, Longline fisheries, Morocco*

[1] National Institute of Fisheries Research, Casablanca, Morocco
* Corresponding author: serghinimansour@gmail.com
[2] NOAA, 1315 East-West Highway, Rm. 2620, Silver Spring, MD 20910

1. Introduction

Shortfin mako *Isurus oxyrinchus*, is distributed worldwide in temperate and tropical waters. It has an important role in the equilibrium of the marine ecosystem (Bonhommeau *et al*., 2009). This species is harvested as bycatch by the Moroccan longliners targeting swordfish *Xiphias gladius* between the latitudes 20 and 26 N° (**Figure 1**). The average catch of shortfin mako reached 680 tons during the period 2012-2016. During the same period, this species represented 22% of the total catch of the longliners operating in the North Atlantic.

Because of the growing demand for shortfin mako, the catch of this species have increased during the last two decades as estimated by the SCRS. In 2017, the North Atlantic shortfin mako was assessed by the SCRS as being overexploited, however this assessment has not used the data from the Moroccan longline fishery that harvest a significant amount of this species. In this paper, we provide for the first time a standardized CPUE of shortfin mako using the most recent catch and effort data available from the Moroccan longline fishery.

2. Material and methods

2.1. Description of data source

Three types of data were used: i) data for the period 2010-2018 from the Office National des Pêches (ONP) to identify different trips, ii) data from the Moroccan Fishery Department to get technical characteristics of each longliner, and iii) data from port survey conducted by the Institute National de Recherche Halieutique (INRH) at the different ports. These survey data served to identify the types of gears, fishing areas, target and bycatch species and fishing season by gear.

The first type of data was aggregated to get the total catch by vessel, species and trip at the national level. From this database, a new dataset was constructed by aggregating for each vessel the number of trips by year and by port. Then this matrix was analyzed to cluster the vessels by landing port. Given that the VMS data series do not cover the whole study period (DPM), we assumed that each landing port represents one particular fishing area. This hypothesis was confirmed by interviewing the fishermen. Each cluster of vessels was identified as having the maximum of catch landed in a group of ports. As a result, six groups of ports were determined (**Table 1**).

With the purpose of standardizing the CPUE for the shortfin mako, we focused our analysis on the group of ports that accounted for more 50% of the total catches of this species (the group Dakhla-Tantan). Then, to identify the different tactics of this fishery that changes inter-annually, we used the tactics analysis following the approach described in Serghini *et al*., In preparation. The results of this analysis showed that the group of vessels within the selected group of ports (mainly Dakhla) practice several tactics of which fishing for swordfish is the main tactic in association with the large pelagic sharks including shortfin mako. This tactic contains 1089 trips that were used for standardizing CPUE for this species.

2.2. Estimation of the fishing effort

The catch data by fishing vessel and by daily sales of each species landed were used to estimate the fishing effort (number of days at sea). First, the commercial data (ONP) of longliners by their corresponding port were compiled with catches by species (in kg), registration numbers, and date of landing for the period 2010 to 2018. The daily sales of each species for one trip were added together by vessel and by port. Then, we built a data matrix with daily sales, registration number, port, date and catch per species as variables. According to the survey conducted in each port, generally a sale of the catch landed by one vessel corresponds to one trip (in general sales are one to two days). Then, to estimate the successive daily sales that correspond to one trip we aggregated by registration number, port and species. Finally, a new matrix was created with rows correspondent to the different trips.

To estimate the number of fishing days elapsed between two successive trips we subtracted the number of days that the vessel spent to navigate to fishing grounds and returning back to the port (2 days on average) and the days staying at the port (2 days). To calculate the effort, we multiplied the number of hooks deployed during one fishing day (estimated to 1200 hooks) by the number of fishing days.

2.3. Size structure of catches

The sizes comprised between 120 and 240 cm dominated the size structure of the catches (**Figure 2**). This is similar to the size structure of the fish targeted by longliners in North Atlantic (Natanson *et al*., 2006).

2.4. Impact of the management regulation

Since its development in 2003, the Moroccan longline fishery targeting swordfish and catching the shortfin mako as bycatch remained stable in terms of both target species and the geographic extent of the fishing area. The fishing activity takes place almost the whole year except for the month of May to repair the vessels and the gears. In addition, there was no big changes in the ICCAT management regulation for the target species; the TAC for swordfish has remained the same over the period considered for the analysis (2010-2017) (e.g. 13700 tonnes). Therefore, there is no impact on our CPUE standardization.

2.5. Standardization of the CPUE

2.5.1. Selecting of explanatory variables and trips for the standardization of CPUEs

Analysis of fleet dynamic and identification of fishing tactics is particularly important for standardizing CPUE. This approach helps to select a sample of the fleet to be used for CPUE standardization and to remove all fishing trips that are directed to other species. Given that fishing tactics can be described by a combination of many characteristics variables, the approach is to extract all fishing trips that contributed to the constitution of each fishing tactics and use its significant characteristics variables as explanatory variables in the model to standardized CPUE. This approach allowed us to select the following explanatory variables to construct a full model: species, year, REGISTRATION.No and month. The choice of longliner registration number, among the explanatory variables, is justified by taking on account the variability of different longliners characteristics, the crew and the fishing areas frequented. Species were log(x+1) transformed for use in the final statistical model using Boosted Regression Trees (BRT) and Generalized Linear Models (GLM):

LOG(Shortfin.mako+1) ~YEAR+REGISTRATION.No+MONTH+
LOG(Swordfish+1)+LOG(Blue.shark+1)+LOG(Yellowfin.tuna+1)+LOG(Bigeye.tuna+1)

2.5.2. Boosted Regression Trees (BRT) and Generalized Linear Models (GLM)

The BRT is an ensemble of method that combine statistical models and machine learning through two algorithms : regression trees are from the classification and regression tree (decision tree) group of models, and boosting build (Machine Learning) (Elith *et al*., 2008). Handling exploratory variables of different types, treatment of missing data, handling outliers and insensitive to data distribution, dealing with complex nonlinear relations. Moreover, interactions are easily implemented, without concern for potentially complicated calculations of the standardized year effect (De'ath, 2007; Elith *et al*., 2008; Hastie *et al*., 2009).

In the current study, shortfin mako CPUE was standardized using BRT main effect (tree complexity tc=1) and two way interactions (tc=2). The Catch and effort data (1089 longline sets) were subdivided on the training and test data sets (70% , 30%) for residual analyses and test prediction model with test data. After the validation of the model, the global data (train data + test data) were used to estimate final standardized CPUE. We could justify this choice by the number of rows of data (1089), close to the requested rows number (1000). The small size of the training and test data sets degrades prediction error model (PE; accuracy) (De'ath, 2007). The reader is referred to (Albeare, 2009; Mateo and Hanselman, 2014) for more descriptions of R script to performed BRT.

A Generalized Linear Models (GLM) approach (Mccullagh and Nelder, 1989) was applied with the same data used for BRT models, under a log normal error distribution assumption. The step AIC analysis was performed to select the statistically significant factors in the final model.

3. Results and discussions

The number of observations (trips) analyzed by variables combination is summarized in **Table 2**. A total of 1089 trips, which represent an average of 121 observations per year, were used to compute the standardized CPUE of shortfin mako for the period 2010-2018.

The GLM ANOVA results showed that the variables Month, REGISTRATION.No and Year, are significant at 1% level and contribute with 50.5%, 34.65% and 13.9%, respectively (**Table 3**). However, in BRT models, the relative influence of each explanatory variables are ordered according to importance as follows: month (47.17%), REGISTRATION.No (24.91%), Swordfish (12.31%), Year (11.22%) and Blue shark (3.87%) (**Table 4).**

The diagnostic plots showed that both BRT and GLM models fit well the data given the normal pattern of the residuals. The fitted values vs the residuals did not show any particular trend (**Figure 3, 4, 5 and 6**). The computed annual standardized CPUE using the BRT and GLM models are summarized in **Tables 5 and 6**. The yearly trend of the CPUE is illustrated by the **Figures 7 and 8**. The comparison between the different models shows that the BRT two-way interactions (RMSE= 0.792 and PED = 56%) fit better the data than BRT main effect model (RMSE= 0.841 and PED = 43%) and GLM (RMSE= 0.807 and PED = 39%) (**Tables 7 and 3**).

The standardized CPUE obtained from the best fit model (BRT two-way interactions) indicates a declining trend since the early years and slight increase and stability in the last four years of the time series (**Figure 7)**.

Finally, to evaluate the changes in patterns of yearly-standardized CPUE we used step plots technique (Bentley *et al*., 2012). The results are shown is **Table 8.**

These results show that when we added the variable "Months" in the model, the effect appears positive (green color) in 2010, 2011, 2013 and negative (pink color) in 2012 and 2014. The variable REGISTRATION.No has a positive effect in 2016, 2017 and negative effect in 2010, 2011 and 2018. The swordfish had a negative effect on standardized CPUE in 2010 and 2017 and a positive effect in 2011. The Blue shark had a negative effect during the period 2010 and 2011. The interaction analysis (**Figure 9**) of the explanatory variables shows that there is interaction between pairs of factors in model fitting: (REGISTRATION.No, Swordfish), (Swordfish, YEAR) and (MONTH, YEAR).

These changes in patterns and interaction analysis indicates that further research and analysis are necessarily to understand the contribution of each explanatory variables.

Literature listed

Albeare, S.M., 2009. Comparisons of Boosted Regression Tree, GLM And GAM Performance In The Standardization Of Yellowfin Tuna Catch-Rate Data From The Gulf Of Mexico Lonline Fishery 85.

Bentley, N., Kendrick, T.H., Starr, P.J., Breen, P.A., 2012. Influence plots and metrics: tools for better understanding fisheries catch-per-unit-effort standardizations. ICES J. Mar. Sci 69, 84–88. doi:doi:10.1093/icesjms/fsr174

Bonhommeau, S., Fromentin, J., Internationale, C., 2009. Les populations ichtyologiques de grands pélagiques / SRM MMN. IFREMER- UMR EME. 7p.

De'ath, G., 2007. Boosted trees for ecological modeling and prediction. Ecology 88, 243–251. doi:10.1890/0012-9658(2007)88[243:BTFEMA]2.0.CO;2

Elith, J., Leathwick, J.R., Hastie, T., 2008. A working guide to boosted regression trees. J. Anim. Ecol. 77, 802–813. doi:10.1111/j.1365-2656.2008.01390.x

Hastie, T., Tibshirani, R., Friedman, J.H., 2009. The Elements of Statistical Learning: Data Mining, Inference, and Prediction. Springer-Verlag, New York 1, 337–387. doi:10.1007/b94608

Mateo, I., Hanselman, D., 2014. A Comparison of Statistical Methods to Standardize Catch-Per-Unit-Effort of the Alaska Longline Sablefish Fishery. U.S.Dep.Commer.,NOAA Tech. Memo. NMFS-AFSC-269 1–71.

Mccullagh, P., Nelder, J., 1989. Generalized linear models. Second Edition. Chapman & Hall. University Press. Cambridge. 511 pp.

Natanson, L.J., Kohler, N.E., Ardizzone, D., Cailliet, G.M., Wintner, S.P., Mollet, H.F., 2006. Validated age and growth estimates for the shortfin mako , Isurus oxyrinchus , in the North Atlantic Ocean 367–383. doi:10.1007/s10641-006-9127-z

Table 1. Fleet dynamic; percentage of trips of vessel in and out of cluster.

Cluster name	Ports	% trips in cluster	% trips out cluster	number of vessel	number of trips in cluster	number of trips out cluster
Cluster 1	DAKHLA	92,000	8,000	270	20384	1800
	TANTAN					
Cluster 2	AGADIR	97,000	3,000	290	19777	703
	SAFI					
	LAAYOUNE					
	TANTAN					
	BOUJDOUR					
	JORF AL ASFAR					
	ESSAOUIRA					
	TARFAYA					
	JADIDA					
	SIDI IFNI					
Cluster 3	LARACHE	70,000	30,000	77	1264	529
	CASABLANCA					
	JADIDA					
	MEHDIA					
Cluster 4	TANGER	99,000	1,000	67	58851	544
	AL HOCEIMA					
	NADOR					
	MDIQ					
	RAS KABDANA					
Cluster 5	MOHAMEDIA	97	3,000	59	4936	132
	MEHDIA					
Cluster 6	ASILAH	96	4	17	1727	38

Table 2. Number of observations by year and month.

ear/Month	1	2	3	4	5	6	7	8	9	10	11	12	Total
2010	0	0	14	0	0	0	0	0	0	0	13	6	33
2011	10	16	15	1	0	12	15	14	9	16	17	7	132
2012	11	12	5	10	7	13	6	18	15	17	13	6	133
2013	12	15	19	16	6	2	17	17	11	15	14	8	152
2014	10	13	16	19	3	9	13	12	15	19	16	10	155
2015	18	18	16	17	3	2	13	13	8	16	17	10	151
2016	16	7	17	7	0	10	8	8	12	11	2	1	99
2017	18	16	13	8	0	1	12	7	10	14	18	7	124
2018	12	9	16	6	5	6	14	1	16	11	12	2	110
Total	107	106	131	84	24	55	98	90	96	119	122	57	1089

Table 3. GLM: Summary of the results of the ANOVA.

	Df	Deviance	% deviance	Resid. Df	Resid. Dev	F	Pr(>F)
NULL	1088	946.45		1088	946.45		
YEAR	8	54.423	13.90	1080	892.03	12.8325	< 2e-16 ***
REGISTRATION.No	17	135.629	34.65	1063	756.4	15.0494	< 2e-16 ***
MONTH	11	197.673	50.50	1052	558.73	33.8977	< 2e-16 ***
Swordfish	1	1.857	0.47	1051	556.87	3.502	0.06157 .
Blue.shark	1	0.811	0.21	1050	556.06	1.53	0.21640
Yellowfin.tuna	1	0.237	0.06	1049	555.82	0.4465	0.50417
Bigeye.tuna	1	0.501	0.13	1048	555.32	0.9452	0.33118
Bullet.tuna	1	0.275	0.07	1047	555.05	0.5189	0.47146

Signif. codes : 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

R^2 =0.41, (Adjusted-R^2 = 0.39) and RMSE= 0.807

Table 4. Model 1-Shortfin mako-tc2-BRT-Two-way Interaction: Relative influence of model terms calculated by the contribution of each term in reducing overall model deviance.

Variable	Relative influence
MONTH	47.17
REGISTRATION.No	24.91
Swordfish	12.32
YEAR	11.22
Blue.shark	3.88
Bigeye.tuna	0.34
Yellowfin.tuna	0.17

Table 5. Nominal and standardized CPUE using the BRT Two-way interactions.

YEAR	Nom.CPUE	Stand.CPUE
2010	112	113
2011	113	107
2012	86	83
2013	109	115
2014	56	64
2015	85	89
2016	90	88
2017	108	95
2018	103	99

Table 6. Nominal and standardized CPUE using the GLM approach.

YEAR	Nom.CPUE	Stand.CPUE	LowerLim	UpperLim	CV
2010	112	134	98	184	3.27
2011	113	102	81	128	2.49
2012	86	84	67	105	2.56
2013	109	118	96	145	2.17
2014	56	62	50	77	2.62
2015	85	93	75	116	2.45
2016	90	84	67	107	2.67
2017	108	101	82	126	2.34
2018	103	95	76	119	2.51

Table 7. Brief model summary statistics: Model 1-Shortfin mako-tc2-BRT-Two-way Interaction and Model 1-Shortfin mako-tc1-BRT- main effects model.

Model 1-Shortfin mako-tc2-BRT-Two-way Interaction	fitting final gbm model with a fixed number of 2500 trees for Shortfin.mako
	mean total deviance = 0.869
	mean residual deviance = 0.381
	estimated cv deviance = 0.683 ; se = 0.049
	training data correlation = 0.756
	cv correlation = 0.481 ; se = 0.024
Percentage of explained deviance	**56%**
RMSE	**0.7922744**
Model 1-Shortfin mako-tc1-BRT- main effects model	fitting final gbm model with a fixed number of 2000 trees for Shortfin.mako
	mean total deviance = 0.869
	mean residual deviance = 0.499
	estimated cv deviance = 0.683 ; se = 0.024
	training data correlation = 0.657
	cv correlation = 0.466 ; se = 0.019
Percentage of explained deviance	**43%**
RMSE	**0.8419857**

Table 8. Changes in patterns of yearly-standardized CPUE of Shortfin mako removing each explanatory variables one at a time. Cell with the pink color: negative influence, Cell with the green color: positive influence. Unit is kg/1200 hook*day.

YEAR	MONTH	REGISTRATION.No	Swordfish	Blue.shark	Bigeye.tuna	Yellowfin.tuna	Bullet.tuna
2010	19	-16	-15	-6	-5	-7	-7
2011	8	-13	4	-3	-5	-7	-5
2012	-5	-4	0	3	1	1	1
2013	15	0	-1	0	1	-1	-1
2014	-13	1	3	-1	0	0	0
2015	-1	-3	-1	-3	-1	-1	-1
2016	-2	6	1	-3	0	2	1
2017	-2	5	-5	2	0	0	0
2018	4	-10	0	2	-2	-5	-2

Figure 1. Geographical delimitation of the fishing area of the swordfish longline fishery operating south of the Moroccan Atlantic waters.

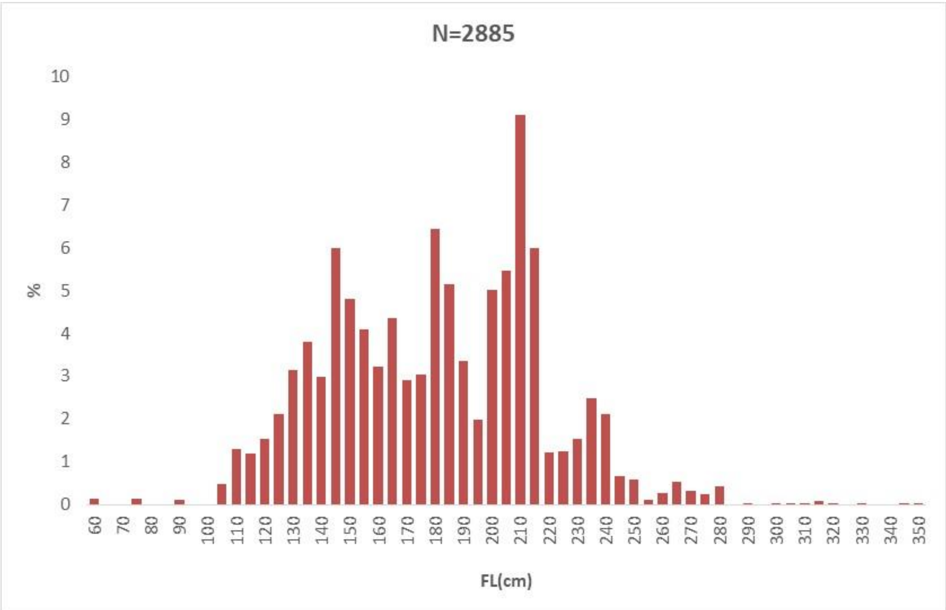


Figure 2. Size frequency of shortfin mako catches by the Moroccan longline fishery in the Atlantic for the period 2014-2017.

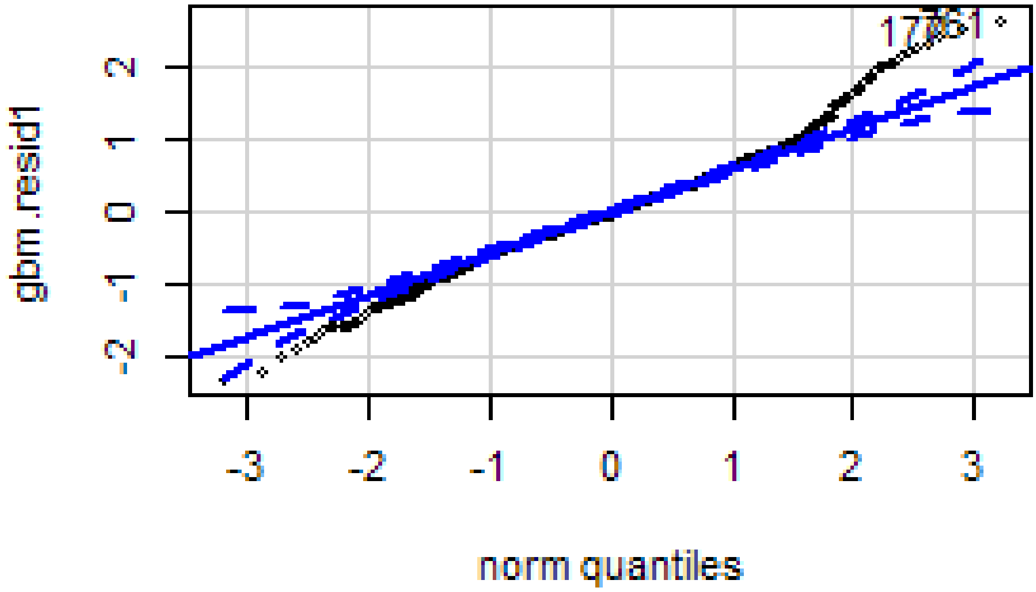


Figure 3. Shortfin mako-tc2-BRT-Two-way Interaction. QQ plots.

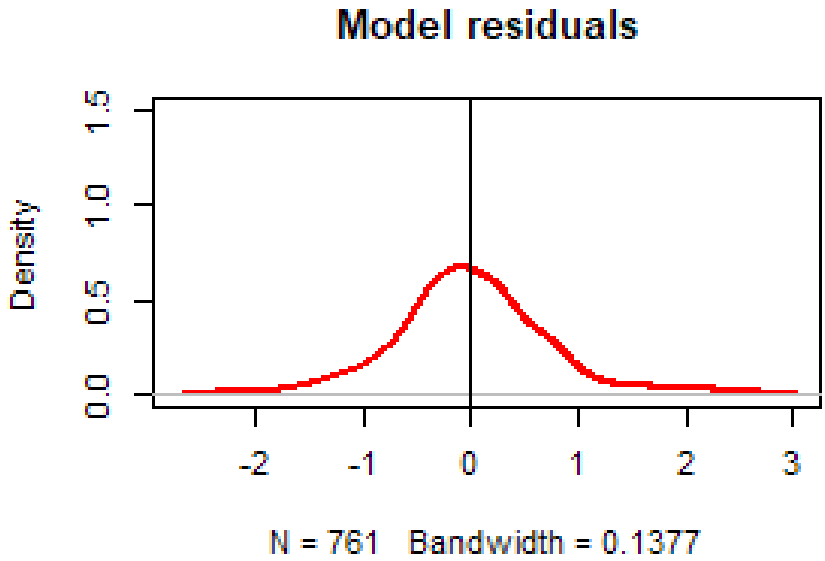


Figure 4. Shortfin mako-tc2-BRT-Two-way Interaction. Density functions for the BRT model residuals; values calculated using the default Gaussian kernel.

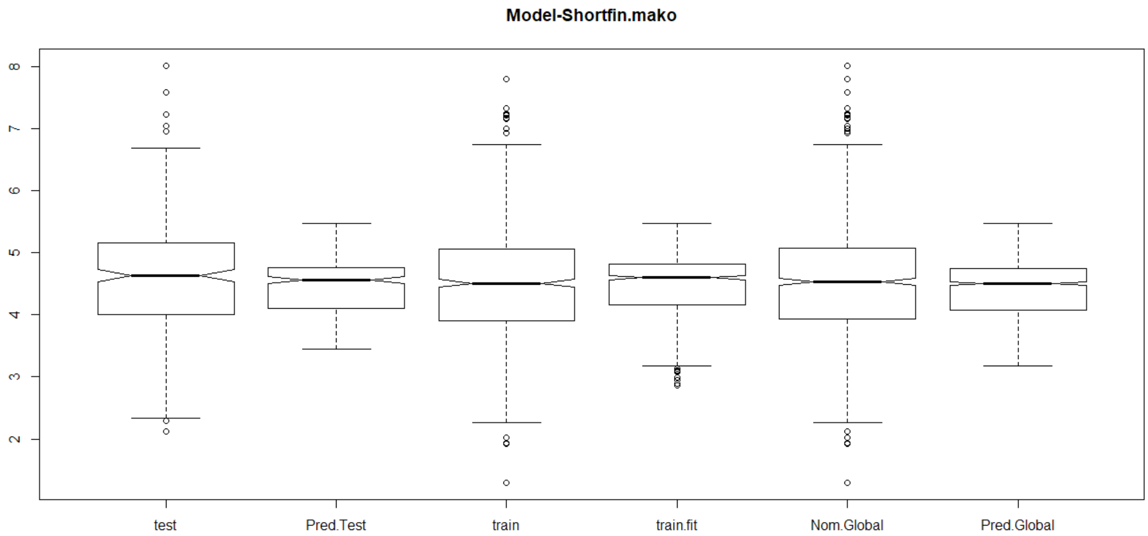


Figure 5. Boxplots of actual log(CPUE+1) data versus model predictions, using the test dataset, train dataset and global data; notches indicate robust estimates of the medians.

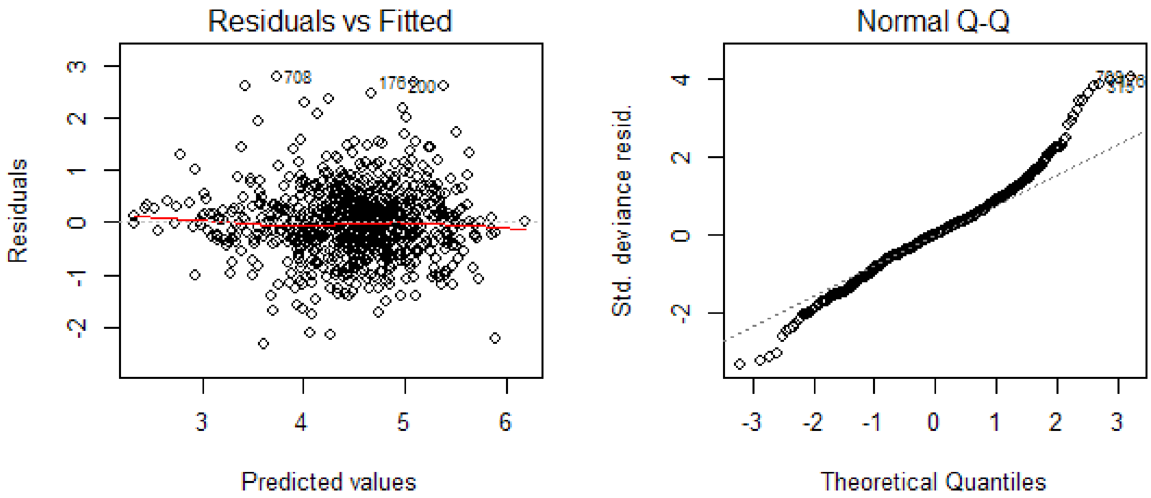
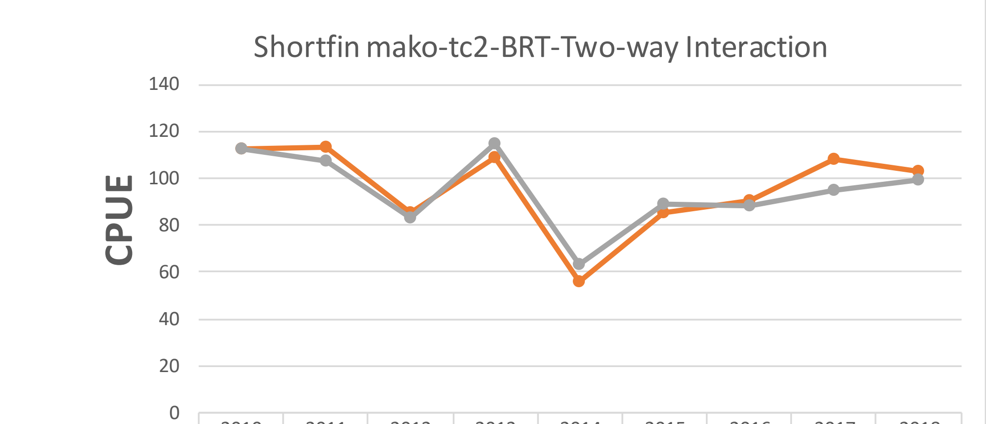
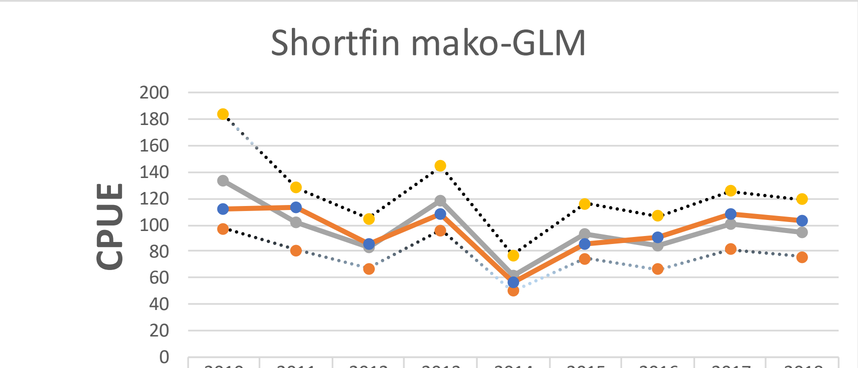


Figure 6. Shortfin mako-GLM, Residuals *vs* predicted and QQ plots.



	2010	2011	2012	2013	2014	2015	2016	2017	2018
CPUE.Nom	112	113	86	109	56	85	90	108	103
CPUE.Stand	113	107	83	115	64	89	88	95	99

Figure 7. Comparisons of the standardized CPUE indices and nominal indices of Shortfin mako from the Moroccan longline fishery, 2010-2018.



	2010	2011	2012	2013	2014	2015	2016	2017	2018
LowerLim	98	81	67	96	50	75	67	82	76
Stad.CPUE	134	102	84	118	62	93	84	101	95
UpperLim	184	128	105	145	77	116	107	126	119
Nom.CPUE	112	113	86	109	56	85	90	108	103

Figure 8. Comparisons of the standardized and nominal index of Shortfin mako from the Moroccan longline fishery, 2010-2018.

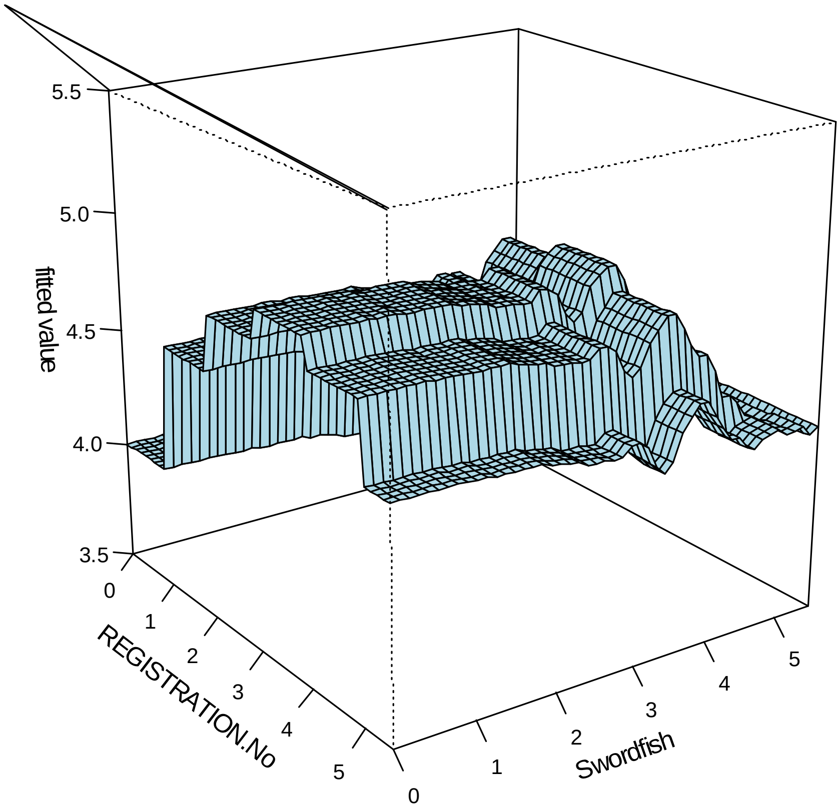


Figure 9. Shortfin mako-tc2-BRT-Two-way Interaction: Interaction plot for the REGISTRATION.No x Swordfish cross-term using global data; interaction size= 10.11.