SCRS/2009/066 Collect. Vol. Sci. Pap. ICCAT, 65(5): 1762--1771 (2010)

STANDARDIZED CATCH RATE OF SAILFISH (*ISTIOPHORUS PLATYPTERUS*) CAUGHT BY BRAZILIAN LONGLINERS IN THE ATLANTIC OCEAN (1978-2008)

Catarina Wor[1], Bruno L. Mourato[1,2], Humberto G. Hazin[1], Fábio H. V. Hazin[1],
Paulo Travassos[1] and Humber Andrade[1]

SUMMARY

In this paper generalized linear models were used to generate two standardized sailfish CPUE series from the Brazilian longline fishery between 1978 and 2008. The first standardized series included year, quarter and area as explanatory variables, while the second series included all those factors plus a cluster variable, as an indicative of the targeting strategy. A delta lognormal error distribution was assumed in both cases, due to the high proportion of sets with zero catches (close to 80%). In both models the variable "Year" was the one which accounted for the greatest part of the explained variance, followed by its interactions with Target and Quarter, for the positive catches of the model with the Target factor, and followed by Year interactions with Quarter and Area in the model without the target factor.It was not clear, though, which of the standardized CPUE series reflected more accurately the actual abundance of the stock.

RÉSUMÉ

Le présent document a utilisé des modèles linéaires généralisés pour créer deux séries standardisées de la CPUE pour le voilier de la pêcherie palangrière brésilienne entre 1978 et 2008. La première série standardisée incluait l'année, le trimestre et la zone comme variables explicatives, tandis que la seconde série incluait tous ces facteurs, plus une variable de groupement, comme indication de la stratégie de ciblage. Une distribution d'erreur delta lognormale a été postulée dans les deux cas, en raison de la forte proportion des opérations avec capture nulle (près de 80%). Dans les deux modèles, la variable « Année » tenait compte de la plus grande partie de la variance expliquée, suivie de ses interactions avec « Cible » et « Trimestre », pour les prises positives du modèle avec le facteur « Cible », et suivie des interactions « Année » avec « Trimestre » et « Zone » dans le modèle sans le facteur « cible ». Il ne s'est toutefois pas dégagé clairement laquelle des séries standardisées de CPUE reflétait plus exactement l'abondance réelle du stock.

RESUMEN

En este documento se utilizaron modelos lineales generalizados para generar dos series estandarizadas de la CPUE del pez vela de la pesquería de palangre brasileña entre 1978 y 2008. La primera serie estandarizada incluía año, trimestre y área como variables explicativas, mientras que la segunda serie incluía todos estos factores más una variable de conglomerado como indicativo de la estrategia de pesca dirigida a las especies. En ambos casos se asumió una distribución de error delta lognormal debido a la elevada proporción de lances con capturas cero (aproximadamente el 80%). En ambos modelos, la variable "año" era la que respondía de la mayor parte de la varianza explicada, seguida por sus interacciones con "objetivo" y "trimestre" para las capturas positivas del modelo con el factor "objetivo", y seguida por las interacciones de "año" con "trimestre" y "área" en el modelo sin el factor "objetivo". Sin embargo, no quedó claro cuál de las series de CPUE estandarizada reflejaba de forma más precisa la abundancia real del stock.

KEYWORDS

Catch/effort, abundance, regression analysis, Atlantic sailfish

---

[1] Universidade Federal Rural de Pernambuco/Departamento de Pesca e Aqüicultura, Recife- PE, Brasil.
[2] PhD student, Departamento de Oceanografia-Universidade Federal de Pernambuco-UFPE

1. Introduction

The sailfish (*Istiophorus platypterus*) is a highly valued species for the recreational fisheries, as well as for small-scale and artisanal fisheries, particularly in developing countries, where it represents an important source of food and income. It is also often caught, as by-catch, by the pelagic longline fishery, directed to tunas and swordfish, which accounts, by far, for the main impact on its stocks in the Atlantic Ocean (Restrepo *et. al.*, 2003), even though it is not targeted.

Due to its relevance in the total fishing mortality of sailfish and its relatively greater availability, sailfish CPUE series from tuna longline series is generally one of the main abundance index used as a basis for stock assessments. However, for such use, the CPUE time series should be indeed proportional to the actual abundance of the stock, which implies the need to remove the bias caused by other effects that might affect nominal CPUE but are not related with the real abundance (e.g.: fishing area, seasonal changes and targeting strategy). This process of removal, commonly named CPUE standardization, is often done by the use of generalized linear models (GLMs).

Among the several variables that may influence the CPUE in the tuna longline fishery, the targeting strategy is certainly one of the most important ones, although accurate methods to include it as a factor in the standardization process have been so far elusive (Hazin *et al.* 2007; Schirripa and Goodyear, in press). Recently, clustering methods (*e.g.* cluster analysis) have been applied in the analysis of fishing data, aiming at categorizing fishing efforts based on the proportion of the several species present in the catches. The objective of this paper was to assess the influence of the targeting strategy, as reflected in the catch composition grouped by a cluster analysis, on the CPUE of sailfish caught by Brazilian tuna longline fishery, from 1978 to 2008. For this purpose, 2 sailfish CPUE series were generated by GLM, with and without the use of the targeting strategy, as reflected by the cluster analysis, as a factor.

2. Material and methods

2.1Database

Catch and effort data from 56,504 sets done by the Brazilian tuna longline fleet (chartered and national vessels), carried out from 1978 to 2008 (**Figure 1),** were analyzed in order to obtain standardized indices. Longline sets were widely distributed in the western Equatorial and South Atlantic Ocean, ranging from 015ºW to 050ºW and from 08ºN to 52ºS, but only the kernel area of the effort distribution, comprised between 10° N, 30 S, 50° W and 20°W, was considered in the analysis (**Figure 1**). The resolution of 1º latitude x 1º longitude, taken from the position in the ending of each longline set, was used to characterize the geographical distribution of the fishing effort.

2.2 Analysis

GLMs using set by set data and assuming a delta lognormal error distribution were used to generate two standardized CPUE series. The first standardized series included year, quarter and area as explanatory variables. The variable area had two factors (“North” and “South”) divided by the limit of 15°S (**Figure 1)**. The second standardized series included all factors used in the first series and also the cluster variable, represented by 6 different levels.

In order to generate the clustering variable, the k-means clustering analysis was employed, with the number of clusters being selected by the most important species. Six levels were selected for the clustering factor, named as SWO_BSH, YFT, FAL, ALB, BET and OTHER SHK, according to the dominating species or group of species in the cluster. The multivariate means of each group are shown in **Table 1**.

The response variable used was CPUE as the number of fish caught per 1,000 hooks. The delta-lognormal generalized linear model was chosen due to the large proportion of sets with zero catches (close to 80%), stemming from its by-catch nature in this fishery (**Figure 2**). In both models, for positive and for the proportion of positive catches, we started with full models, considering all explanatory variables as main factors and also the first order interactions among them. In order to select the factors that were important to explain the CPUE variability we relied both in analysis of variance tables and AIC (Akaike Information Criterion).

3. Results and discussion

In both models, with and without the target factor, all included variables and their first order interactions were considered relevant and kept in the models. The analysis of deviance table **(Table 2)** shows that in both models the variable "Year" was the one which accounted for the greatest part of the explained variance, followed by its interactions with Target and Quarter, for the positive catches of the model with the Target factor, and followed by Year interactions with Quarter and Area in the model without the cluster as a factor. Differently from previous work (Hazin *et al.* 2008), where the target factor accounted for more than half of the explained variance, in the case of the sailfish, it responded for only 12 % and 8 % of the explained variance, for the positive sets and for the proportion of positives, respectively. Such a difference is certainly related to the fact that the sailfish is not a target species, a fact reflected in its rather low percentage in all clusters, ranging from 1.4 to 3.5%.

Standard diagnostic plots are shown in **Figures 3 and 4**. QQ-normal plots (second panel from left to right) suggest that normal distribution fits the residuals well in both models. Despite the high variance commonly associated to CPUE series generated without data aggregation, the confidence intervals for both generated series were quite narrow during the analyzed period (**Table 3 and Figure 5**).

Figure 6 shows scaled nominal and standardized CPUE series. In spite of a generally similar trend between both standardized series, they showed some important differences. The cluster-excluded series showed a much more pronounced peak in 1993 and a more consistent upward trend from 1994 on, in spite of a strong oscillation, from year to year. The cluster-included series, in turn, showed a peak in 1996, and a more stable trend in recent years, particularly in the last 3 included in the model (2006-2008). Although it is not clear, from the present data, which of the series reflects more accurately the actual abundance of the stock, the model with the target factor seemed to fare a little better in comparison with the model without target factor, since it showed a little lower mean variation coefficient (model with target CV=6.8%; and model without target CV= 10.1%).

4. Acknowledgements

The authors acknowledge the Secretaria Especial de Aqüicultura e Pesca da Presidência da Republica do Brasil (SEAP), which made this study possible.

5. References

Hazin, F.H.V., Hazin, H.G. Travassos, P. and Carvalho, F.C. 2008, A comparison of bigeye tuna, *Thunnus obesus*, CPUE series, for Brazilian tuna longline fisheries, from 1978 to 2005, with and without target species as a factor in GLM analysis. Collect. Vol. Sci. Pap. ICCAT, 62(2): 404-416.

ICCAT. 2006, International Commission for the Conservation of Atlantic Tunas: Report for biennial period, 2006-07. Part I, Vol. 2, 244p.

Restrepo, V., Prince, E.D., Scott, G.P., Uozumi, Y. 2003, ICCAT stock assessments of Atlantic billfish. Marine and Freshwater Research 54: 361-367.

Schirripa, M. J. and Goodyear, C.P. *In press.* CPUE standardizations of species targeting in longline fisheries: Initial investigation of the utility of simulated data for method verification. Collect. Vol. Sci. Pap. ICCAT, 65.

Table 1. Distribution of longline sets done by the Brazilian tuna longline fishery in the Atlantic Ocean, from 1978 to 2008, by cluster.

Species	*SWO_BSH*	*YFT*	*FAL*	*ALB*	*BET*	*OTHER SHK*
Sets (%)	**38.0%**	**24.8%**	**1.1%**	**17.8%**	**15.5%**	**2.8%**
Other tunas	10.6%	0.2%	0.0%	0.8%	0.1%	2.3%
Yellowfin tuna	5.0%	**43.9%**	4.2%	6.2%	11.9%	8.2%
Albacore	4.4%	8.5%	0.3%	**73.8%**	5.3%	7.1%
Bigeye tuna	4.3%	8.0%	0.7%	4.5%	**56.3%**	5.9%
Swordfish	**36.7%**	7.8%	8.2%	3.2%	13.2%	9.5%
Sailfish	1.9%	3.5%	1.4%	1.4%	1.4%	2.6%
White Marlin	0.9%	1.6%	0.6%	0.7%	1.0%	2.0%
Blue Marlin	1.4%	1.5%	0.4%	0.6%	1.3%	1.1%
Other billfish	1.1%	0.1%	0.0%	0.4%	0.1%	0.7%
Wahoo	1.0%	5.2%	0.2%	0.7%	0.7%	1.1%
Dolphin fish	2.5%	1.5%	0.2%	0.4%	0.6%	2.0%
Blue Shark	**21.2%**	3.7%	4.5%	1.8%	3.2%	4.0%
Hammerhead sharks	1.2%	0.3%	0.7%	0.1%	0.1%	0.5%
Bigeye Thresher shark	0.2%	0.1%	0.2%	0.0%	0.0%	0.1%
Shortfin mako	1.8%	0.5%	0.4%	0.5%	0.2%	0.5%
Silky shark	0.5%	0.3%	**75.6%**	0.0%	0.2%	0.3%
Oceanic whitetip shark	0.1%	0.0%	0.1%	0.0%	0.0%	0.0%
Other Sharks	1.3%	1.7%	1.3%	1.7%	1.9%	**49.3%**
Other Teleosts	3.9%	11.7%	0.9%	3.1%	2.5%	2.7%

Table 2. Deviance analysis table of explanatory variables for Delta-lognormal models with and without cluster among its explanatory variables.

Models without Target

Positive sets

	Df	*Deviance*	*Resid. Df*	*Resid. Dev*	*Pr(Chi)*	*Explained Deviance*	*Explained Model*
NULL			13405	11503			
Year	30	823.668811	13375	10680	7.51E-207	47.3%	7.2%
Area	1	4.19022035	13374	10675	0.0170741	0.2%	7.2%
Quarter	3	35.5153396	13371	10640	1.99E-010	2.0%	7.5%
Year:Area	24	251.242727	13347	10389	5.17E-057	14.4%	9.7%
Year:Quarter	90	610.042054	13257	9778.6	1.58E-115	35.0%	15.0%
Area:Quarter	3	18.3678463	13254	9760.2	1.60E-005	1.1%	15.2%

Proportion of positive sets

	Df	*Deviance*	*Resid. Df*	*Resid. Dev*	*Pr(Chi)*	*Explained Deviance*	*Explained Model*
NULL			218	6512.7			
Year	30	2976.92043	188	3535.8	0	50.3%	45.7%
Area	1	807.028402	187	2728.8	1.60E-177	13.6%	58.1%
Quarter	3	26.0024158	184	2702.8	9.53E-006	0.4%	58.5%
Year:Area	25	992.659787	159	1710.1	2.10E-193	16.8%	73.7%
Year:Quarter	90	1077.46206	69	632.65	6.68E-169	18.2%	90.3%
Area:Quarter	3	37.4342881	66	595.22	3.72E-008	0.6%	90.9%

Models With Target

Positive sets

	Df	*Deviance*	*Resid. Df*	*Resid. Dev*	*Pr(Chi)*	*Explained Deviance*	*Explained Model*
NULL			13405	11503.2			
Year	30	823.7	13375	10679.6	0.0000	33%	7%
Quarter	3	35.1	13372	10644.4	0.0000	1%	7%
Area	1	4.6	13371	10639.8	0.0324	0%	8%
Target	5	289.8	13366	10350.0	0.0000	12%	10%
Year:Quarter	90	579.8	13276	9770.2	0.0000	23%	15%
Quarter:Area	3	32.8	13273	9737.4	0.0000	1%	15%
Year:Target	133	638.2	13140	9099.3	0.0000	25%	21%
Year:Area	24	113.7	13116	8985.5	0.0000	5%	22%

Proportion of positive sets

	Df	*Deviance*	*Resid. Df*	*Resid. Dev*	*Pr(Chi)*	*Explained Deviance*	*Explained Model*
NULL			1064	10903.0			
Year	30	2976.9	1034	7926.0	0.0000	38%	27%
Quarter	3	96.2	1031	7829.8	0.0000	1%	28%
Area	1	736.8	1030	7093.0	0.0000	9%	35%
Target	5	608.8	1025	6484.2	0.0000	8%	41%
Year:Quarter	90	1133.7	935	5350.6	0.0000	15%	51%
Quarter:Area	3	59.9	932	5290.7	0.0000	1%	51%
Year:Target	143	1597.1	789	3693.6	0.0000	20%	66%
Year:Area	25	598.9	764	3094.7	0.0000	8%	72%

Table 3. Nominal and standardized (not scaled) CPUE values. The standard errors (SE) and coefficient of variation (CV) are provided.

Year	*CPUE STD Target*	*SE STD Target*	*CV_Target*	*CPUE STD_No-Target*	*SE STD_No-Target*	*CV_No-Target*	*CPUE Nominal*
1978	0.630	0.021	3.4%	0.432	0.055	12.7%	0.492
1979	0.401	0.027	6.6%	0.378	0.052	13.8%	0.391
1980	0.542	0.042	7.7%	0.373	0.052	13.9%	0.344
1981	0.805	0.225	27.9%	0.330	0.056	16.9%	0.262
1982	0.452	0.067	14.8%	0.163	0.026	16.0%	0.136
1983	0.350	0.063	18.0%	0.126	0.017	13.3%	0.323
1984	0.128	0.007	5.7%	0.174	0.034	19.7%	0.228
1985	0.098	0.023	23.2%	0.069	0.015	22.4%	0.082
1986	0.111	0.001	1.0%	0.164	0.014	8.4%	0.214
1987	0.340	0.012	3.5%	0.263	0.029	11.1%	0.438
1988	0.541	0.019	3.5%	0.253	0.027	10.8%	0.266
1989	1.177	0.105	8.9%	0.493	0.045	9.2%	0.558
1990	0.201	0.032	16.1%	0.107	0.034	31.6%	0.134
1991	0.502	0.030	6.1%	0.312	0.031	10.0%	0.355
1992	0.464	0.013	2.9%	0.339	0.041	12.0%	0.364
1993	0.728	0.080	10.9%	0.774	0.190	24.6%	0.575
1994	0.169	0.005	3.0%	0.081	0.012	14.4%	0.123
1995	0.578	0.020	3.5%	0.252	0.021	8.4%	0.327
1996	1.903	0.226	11.9%	0.381	0.039	10.2%	0.695
1997	0.872	0.033	3.8%	0.331	0.022	6.8%	0.610
1998	0.165	0.002	0.9%	0.295	0.015	5.1%	0.521
1999	0.220	0.001	0.3%	0.185	0.009	4.7%	0.268
2000	0.243	0.000	0.2%	0.246	0.008	3.4%	0.344
2001	0.327	0.001	0.2%	0.430	0.014	3.3%	0.506
2002	0.915	0.006	0.7%	0.307	0.016	5.2%	0.413
2003	0.204	0.004	2.0%	0.233	0.019	8.3%	0.326
2004	0.608	0.003	0.5%	0.359	0.010	2.8%	0.888
2005	1.224	0.002	0.1%	0.522	0.013	2.6%	1.049
2006	0.377	0.001	0.3%	0.343	0.014	4.0%	0.455
2007	0.482	0.018	3.8%	0.439	0.028	6.3%	0.708
2008	0.339	0.002	0.5%	0.657	0.032	4.8%	0.738
Means	**0.519**	**0.035**	**6.8%**	**0.316**	**0.032**	**10.1%**	**0.478**

Year	*CPUE Target*	*SE Target*	*CV Target*	*CPUE No-Target*	*SE No-Target*	*CV_No-Target*	*CPUE Nominal*
1978	0.630	0.021	3.4%	0.432	0.055	12.7%	2.034
1979	0.401	0.027	6.6%	0.378	0.052	13.8%	2.559
1980	0.542	0.042	7.7%	0.373	0.052	13.9%	2.903
1981	0.805	0.225	27.9%	0.330	0.056	16.9%	3.812
1982	0.452	0.067	14.8%	0.163	0.026	16.0%	7.341
1983	0.350	0.063	18.0%	0.126	0.017	13.3%	3.093
1984	0.128	0.007	5.7%	0.174	0.034	19.7%	4.379
1985	0.098	0.023	23.2%	0.069	0.015	22.4%	12.128
1986	0.111	0.001	1.0%	0.164	0.014	8.4%	4.668
1987	0.340	0.012	3.5%	0.263	0.029	11.1%	2.281
1988	0.541	0.019	3.5%	0.253	0.027	10.8%	3.765
1989	1.177	0.105	8.9%	0.493	0.045	9.2%	1.793
1990	0.201	0.032	16.1%	0.107	0.034	31.6%	7.472
1991	0.502	0.030	6.1%	0.312	0.031	10.0%	2.816
1992	0.464	0.013	2.9%	0.339	0.041	12.0%	2.749
1993	0.728	0.080	10.9%	0.774	0.190	24.6%	1.740
1994	0.169	0.005	3.0%	0.081	0.012	14.4%	8.121
1995	0.578	0.020	3.5%	0.252	0.021	8.4%	3.057
1996	1.903	0.226	11.9%	0.381	0.039	10.2%	1.438
1997	0.872	0.033	3.8%	0.331	0.022	6.8%	1.638
1998	0.165	0.002	0.9%	0.295	0.015	5.1%	1.919
1999	0.220	0.001	0.3%	0.185	0.009	4.7%	3.736
2000	0.243	0.000	0.2%	0.246	0.008	3.4%	2.906
2001	0.327	0.001	0.2%	0.430	0.014	3.3%	1.977
2002	0.915	0.006	0.7%	0.307	0.016	5.2%	2.421
2003	0.204	0.004	2.0%	0.233	0.019	8.3%	3.064
2004	0.608	0.003	0.5%	0.359	0.010	2.8%	1.126
2005	1.224	0.002	0.1%	0.522	0.013	2.6%	0.954
2006	0.377	0.001	0.3%	0.343	0.014	4.0%	2.198
2007	0.482	0.018	3.8%	0.439	0.028	6.3%	1.412
2008	0.339	0.002	0.5%	0.657	0.032	4.8%	1.356

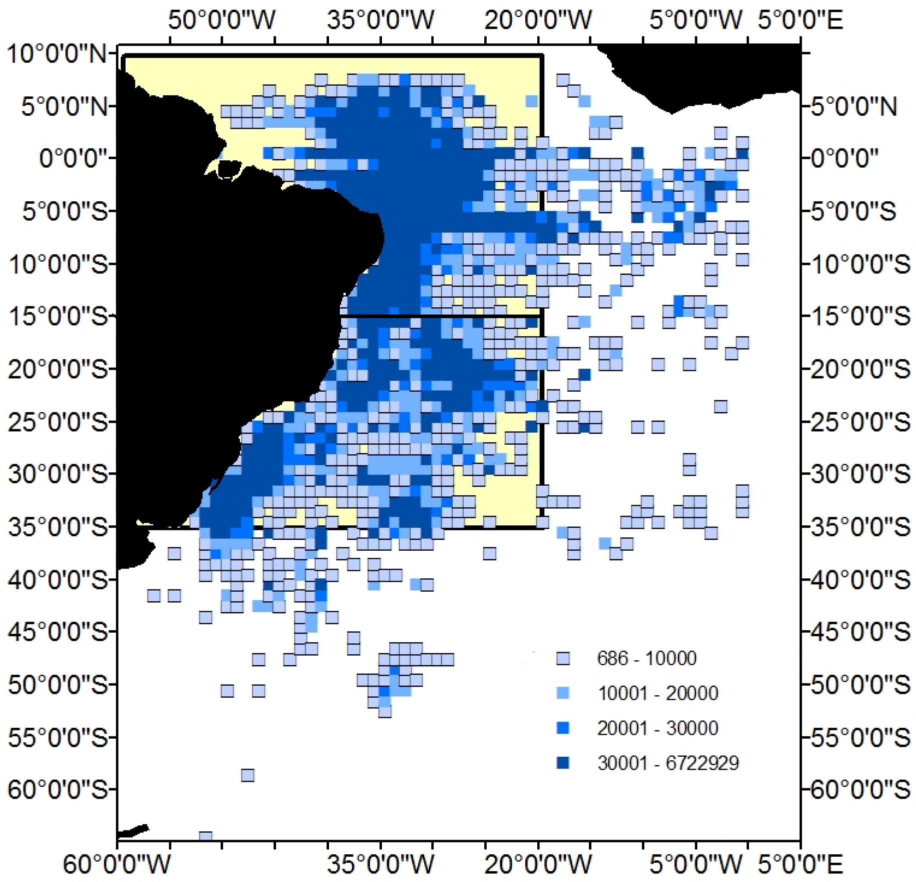

Figure 1. Distribution of effort, in number of sets. from Brazilian tuna longliners (national and chartered vessels) from 1978 to 2008.

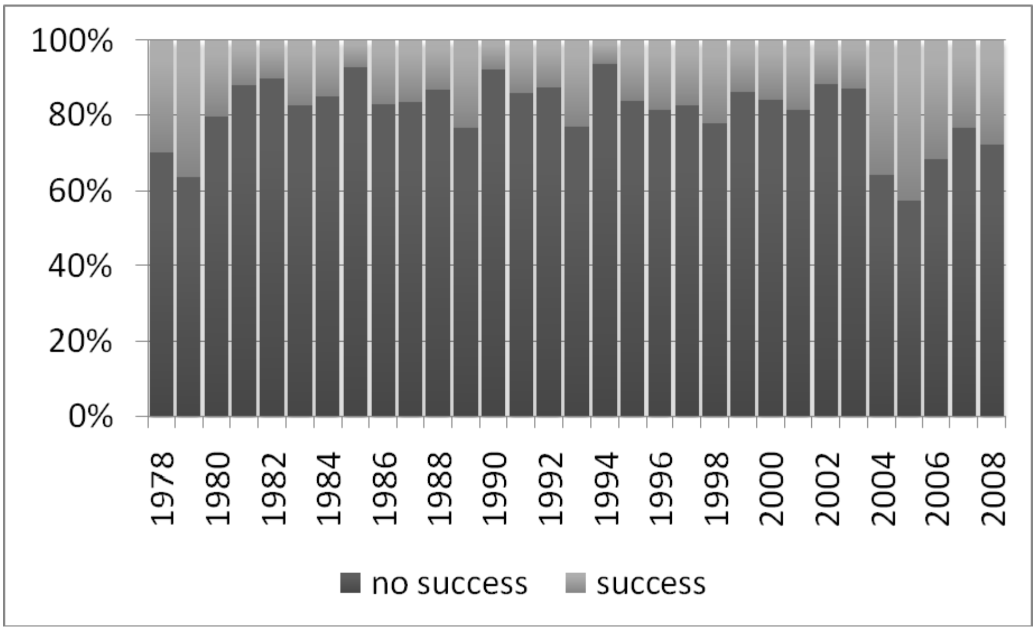

Figure 2. Annual proportion of sets with and without sailfish catches between 1978 and 2008.

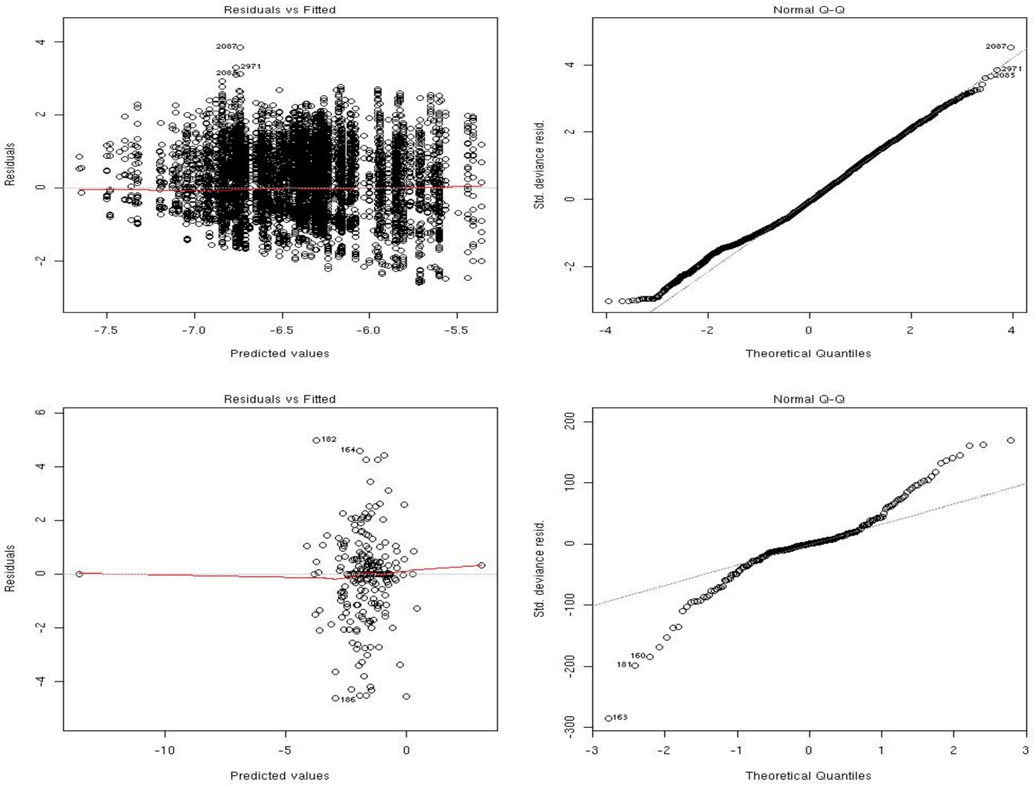

Figure 3. Residual vs. Predicted values and residuals vs. quartiles of standard normal for the model without the cluster factor (positive sets on top and proportion of positives on bottom).

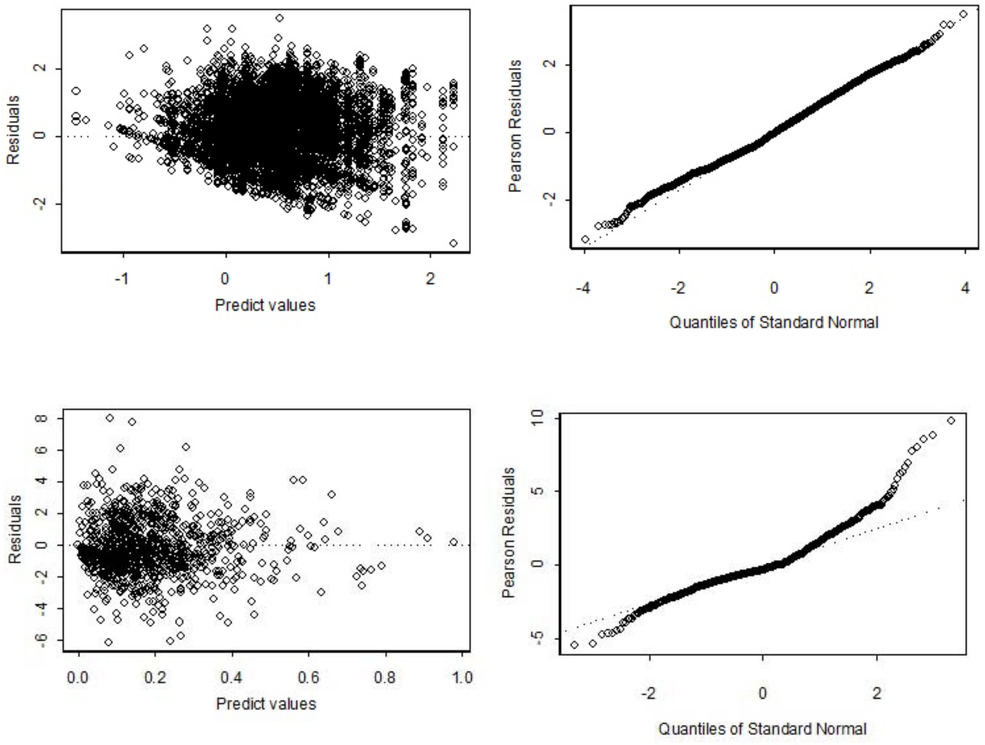

Figure 4. Residual vs. Predicted values and residuals vs. quartiles of standard normal for the model with the Target factor (positive sets on top and proportion of positives on bottom).

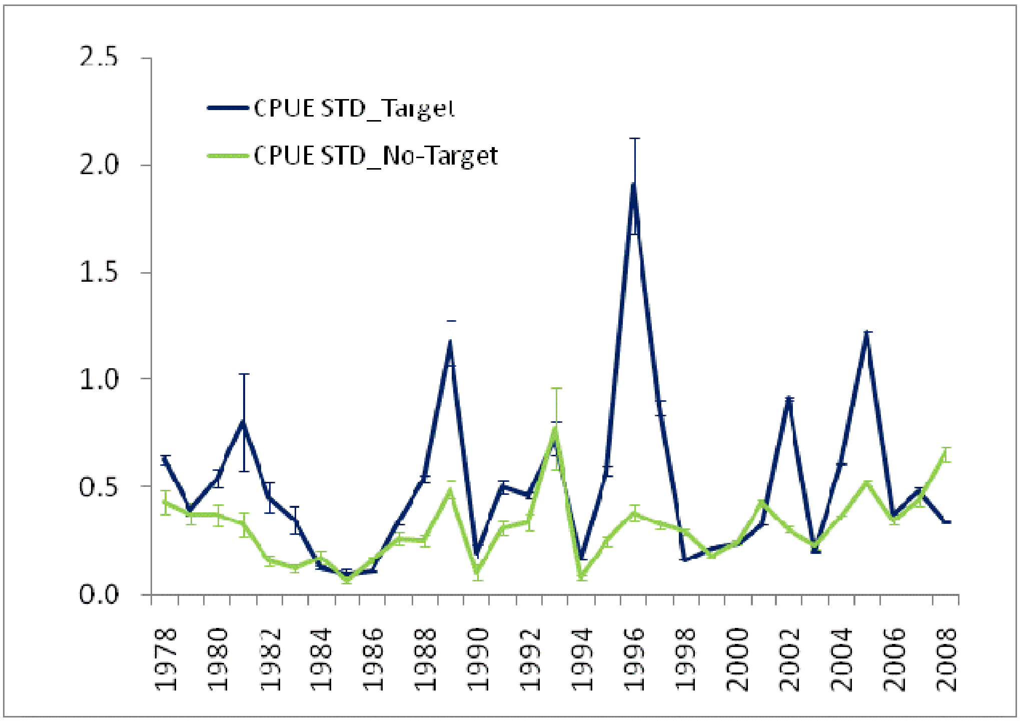

Figure 5. Standardized CPUE for Sailfish for both models with and without Target factor. Vertical bars are 95%confidence intervals for Standardized series.

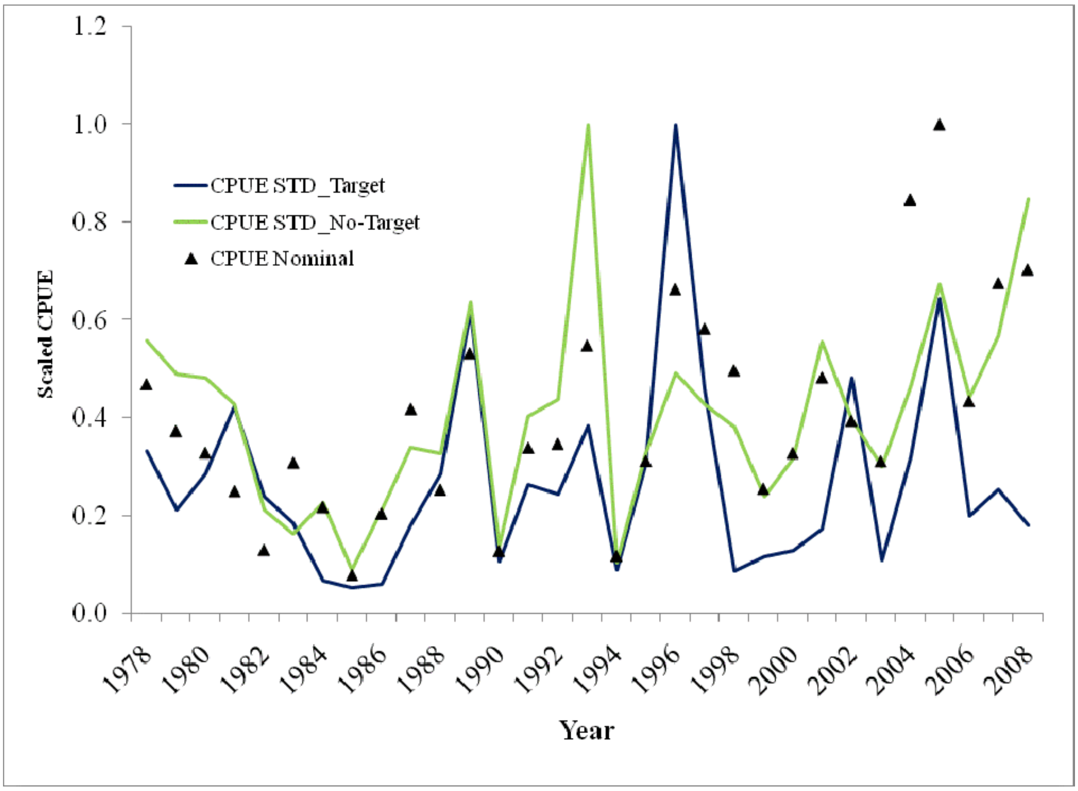

Figure 6. Scaled nominal (diamonds) and standardized (lines) sailfish CPUE series for the model with and without the inclusion of the cluster factor.