STANDARDIZED CPUE INDICES FOR SWORDFISH (*Xiphias gladius*) FROM THE INDONESIAN TUNA LONGLINE FISHERY

Bram Setyadji[1], Zulkarnaen Fahmi[1], & Humber Andrade[2]

[1]Research Institute for Tuna Fisheries, Bali, Indonesia
[2]Federal Rural Universityof Pernambuco, Recife, Brazil

ABSTRACT

Most of the billfish caught in the Indian Ocean are swordfish or Indo Pacific sailfish. Some fleets (e.g. Portuguese and Spanish longline fleets) aim at swordfish, while it is a bycatch for fleets targeting tunas (e.g. Japanese, Taiwanese and Indonesian longline fleets). Relative abundance indices are the input data for stock assessment analyses that provide useful information for decision making and fishery management. In this paper, a Generalized Linear Model (GLM) was used to standardize the catch per unit effort (CPUE) and to estimate relative abundance indices based on the Indonesian longline dataset. Data was collected by scientific observers from August 2005 to December 2016. Most of the vessels monitored were based at Benoa Fishing Port, Bali. Conventional models for counting data were used, but also zero inflated and hurdle models, because the catches were often zero. Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC) were used to select the best models among all those evaluated. Catches are often equal to zero because swordfish is a bycatch for the Indonesian fleet. Both AIC and BIC suggested that the simple negative binomial (NB) model is the best option. The trends were relatively similar to the nominal series, but with smoother peaks. In general, there were tendency of slightly increasing catch trends in the last decade, with the series varying along the period.
KEYWORDS: Swordfish, standardized CPUE, GLM, Eastern Indian Ocean.

INTRODUCTION

Swordfish (*Xiphias gladius*) is a large oceanic apex predator that inhabits all the world's oceans. It has been known as subject for exploitation worldwide, mainly in in Pacific Ocean (Brodziak & Ishimura, 2010), Atlantic Ocean and Meditteranean Sea (Tserpes & Tsimenides, 1995). Throughout the Indian Ocean, swordfish are primarily taken by longline fisheries, and commercial harvest was first recorded by the Japanese in the early 1950's as a bycatch of their tuna longline fisheries (IOTC, 2016). Since 1990s the catches of swordfish increased sharply to a peak of 35 000 t in 1998 (IOTC, 2016) as a result of changes in targeting from tunas to swordfish by Taiwanese longline fleets, and the development of other longline fleets operating under various flags (e.g. Indonesia, Australia, La Réunion, Seychelles and Mauritius) and arrival of longline fleets from the Atlantic Ocean (e.g. Portugal, Spain and UK).

In recent years (2012-2015), Indonesian fleets responsible for approximately 20% of total catch of swordfish in the Indian Ocean (~8,000 mt), followed by Taiwan (17%), Sri Lanka (12%) and Spain (12%) (IOTC, 2016). However, the catch might be lower in the upcoming years due to the issuance of Ministerial Regulation No. 56/2014 and No. 57/2014 about moratorium on ex-foreign fishing vessels and prohibition of transshipment at sea within Indonesia national jurisdiction, which resulted in enormous longline vessel reduction from 584 in 2015 to 271 in 2016. In spite of the relatively high catches, swordfish is considered as a bycatch of the commercial Indonesian tuna longline fishery (Setyadji *et al.*, 2012).

Current stock status of swordfish in Indian Ocean is considered not overfished and not a subject to overfishing (IOTC, 2016). However, the most recent catches (41,760 t in 2015) are 2,360 t above the Maximum Sustainable Yield (MSY) level (39,400 t). Hence, catches in 2017 should be reduced below the MSY level.

Estimations of relative abundance indices (e.g. standardized CPUE) convey important information concerning the status of fisheries stocks. Furthermore, those indices are necessary to run simple models and they are also used as auxiliary data in more detailed stock assessment models (Rodriguez-Marin, 2003). Maunder & Punt (2004) suggested that standardizing catch per unit of effort (CPUE) for species with low number of catch with substantial proportion of zero catch such as billfishes, are need to consider inclusion of models that analyze the proportions of zeros and the positive catch rates separately (i.e., Delta distribution models) or use of zero-inflated models. However, in many cases longline data are best fitted by two part or hurdle models, because they require more flexibility than can accommodated by single-part distributions, particularly for rare and non-targeted species (Zuur *et al*., 2012).

Our analytical objective was to standardize CPUE using six types of GLMs (see below) for swordfish from Indonesian scientific observer data in eastern Indian Ocean. We were interested in how the data-limited of swordfish fishery can construct a fairly robust relative abundance indices amid the "spatial gap" of existing dataset for standardized CPUE in eastern Indian Ocean (e.g. Japanese and Taiwanese longline dataset). We believe the results are valuable as an auxiliary information to assess the status of swordfish in the Indian Ocean.

MATERIALS AND METHODS

Data collection

In this paper, we have analyzed the data gathered by the scientific observers onboard Indonesian commercial tuna longline vessels, which are mainly based in Benoa Fishing Port, Bali. The observer program started in 2005 as an Indonesia-Australia collaboration (Project FIS/2002/074 of Australian Centre for International Agricultural Research), and since 2010 it has been conducted by the Research Institute for Tuna Fisheries (RITF Indonesia). Database include information about 100 fishing trips and 2,565 longline fishing sets from August 2005 to December 2016 with over 3.3 millions of total hooks deployed. Cumulative days at sea covered by observer per year varied between 111 and 404 days with 280 days in average (Table 1). Main fishing grounds cover from west to southern part of Indonesian waters, stretched from 0º-35ºS and 75º-130ºE (Figure 1).

Dataset includes information concerning the number of fish caught by species, total number of hooks, number of hooks between floats (HBF), start time of the set, soak time, and geographic position (latitude and longitude) where the longlines were deployed into the water. The response variable in the models was the catch of swordfish in number (N). Year and quarter were used as categorical (factor) explanatory variables. Additional information was used as explanatory variables as follows:

- Year: analyzed between 2005 and 2016;
- Quarter of the year: 4 categories: 1 = January to March, 2 = April to June, 3 = July to September, 4 = October to December;
- Area: treated as categorical variable, describing spatial catch within and outside the Indonesian EEZ. We also use latitude and longitude as additional quantitative variables.
- Start time of the set: treated as quantitative variable, the values were rounded to the nearest integer;
- Soak time: calculated as the time elapsed between the start of the fishing setting and the start of hauling of the longline. Soak time in the model was treated as continuous variable, thus the values were rounded to the nearest integer;
- Moon phase (29.5 days): categorized into two periods, as light and dark, and assumed the demilunes (first/last quarters), and waning gibbous and full moon as light period; new moon and crescent as dark period (Akyol, 2013)
- Number of hooks between floats: treated as quantitative variable instead of factor.

CPUE standardization

We considered six GLM models for modeling the number of swordfish for modelling the nominal catch (number of fish) as response variable while effort was included in the models as an offset caught. These models are Poisson (P) and negative binomial (NB), which we refer to as the standard models, zero-inflated Poisson (ZIP), zero-inflated negative binomial (ZINB), Poisson hurdle (PH), and negative binomial hurdle (NBH) models.

We used a forward approach to select the explanatory variables and the order they were included in the full model. The first step was to fit simple models with one variable at a time. The variable included in the model with lowest residual deviance was selected first. As second step the model with the selected variable then received other variables one at a time, and the model with lowest residual deviance was again selected. This procedure continued until residual deviance did not decrease as new variables were added to the previous selected model. Finally, all main effects and first order interactions were considered and a backward procedure based on Akaike Information Criterion (AIC) (Akaike, 1974) and Bayesian Information Criterion (BIC) (Schwarz, 1978) were used to select the final models for the six approaches. We also rely in AIC and BIC to compare these models.

The qualities of the fittings were assessed by comparing the observed frequency distributions of the number of fish caught to the predicted frequency distribution, as calculated using the selected models. Kolmogorov-Smirnov test was used to assess if the difference of the two distributions (observed and predicted) were significant. Maps were produced using QGIS version 2.14 (QGIS Developer Team, 2009) and the statistical analyses were carried out using R software version 3.3.3 (R Core Team, 2016), particularly the package *pscl* (Zeileis *et al*., 2008), *lsmeans* (Lenth, 2016), *MASS* (Venables & Ripley, 2002), *Hmisc* (Harrell Jr, 2017), and *statmod* (Giner & Smyth, 2016).

RESULTS

Descriptive Catch Statistic

RITF scientific observers recorded catch and operational data at sea following Indonesian tuna longline commercial vessels from 2005-2016. The dataset contained 100 trips, 2565 sets, 2797 days-at-sea, and more than 3.3 hooks deployed, respectively (Table 1). The spatial data distributed mainly in eastern Indian Ocean with most of the observation were conducted in the area south of Indonesian waters, between 0º-35º S and 75º-125º E (Figure 1).

CPUE data characteristics

SWO nominal CPUE series is presented in Figure 2. In general, the catches of SWO during the last decade were highly variable, but showing an increasing trend. The lowest CPUE recorded was in 2011 (0.12±0.04), as the highest was in 2009 (0.53±0.06). On the other hand, the proportion of zero catch per set for SWO was also high, around 71.07%. Varying annually between a minimum of 0.61±0.03 in 2012 and a maximum of 0.82±0.04 in 2005 (Figure 3).

CPUE standardization

The number of parameters (k), AIC, BIC, logarithm of the likelihood (logLik), number of predicted zero catches, and *p* values of Kolmogorov-Smirnov test as calculated using six model structures (P, NB, ZIP, ZINB, HP and HNB are shown in Table 2. Overall the logarithm of likelihood of zero-inflated and particularly hurdle models were high but they are more complex with large number of parameters. The number of zero catches in the database was 1823. Negative binomial model was selected because it has the lowest value of both AIC and BIC (4330.40 and 4599.47, respectively).

Several explanatory variables tested for the SWO CPUE standardization were significant and contributed significantly for explaining part of the deviance. Some interactions were also significant, and were therefore included in the final model. On the final model, the factors that contributed most for the deviance were the Start_Set, followed by Year, HBF, Moon, Quarter, Area, Long, Lat and the interactions (Table 3). In terms of model validation, the residual analysis, including the residuals distribution along the fitted values, the QQ plots and the residuals histograms, showed that the model was adequate with no major outliers or trends in the residuals (Figure 4).

The final standardized SWO CPUE index (N/1000 hooks) from Indonesian data in the eastern Indian Ocean between 2005 and 2016 is shown in Figure 5 and Table 4. The trends were relatively similar to the nominal series, but with smoother peaks. In general, there were no noticeable trends, with the series varying along the period.

ACKNOWLEDGEMENT

The authors wish to thank the scientific observers who performed invaluable roles in fisheries data collection as part of Australian Centre for International Agricultural Research Project (ACIAR) FIS/2002/074, and subsequently as observers of Research Institute for Tuna Fisheries (2011 and onwards). We also acknowledge the funding support from ACIAR for the aforementioned project.

REFERENCES

Akyol, O. 2013. The Influence of the Moon Phase on the CPUEs of Swordfish Gillnet Fishery in the Aegean Sea, Turkey. Turkish Journal of Fisheries and Aquatic Sciences. 13: 355-358. doi: 0.4194/1303-2712-v13_2_18

Akaike, H. 1974. A new look at the statistical model identification. IEEE Transactions on Automatic Control AC. 19: 716-723.

Brodziak, J. & Ishimura, G. 2010. Stock assessment of North Pacific swordfish (Xiphias gladius) in 2009. Pacific Islands Fish. Sci. Cent., Natl. Mar. Fish. Serv., NOAA, Honolulu, HI 96822-2396. Pacific Islands Fish. Sci. Cent. Admin. Rep. H-10-01, 37 p.

Giner, G & Smyth, GK (2016). statmod: probability calculations for the inverse Gaussian distribution. R Journal. 8(1): 339-351

Harrell Jr, F.E. 2017. Hmisc: Harrell Miscellaneous. R package version 4.0-3. https://CRAN.R-project.org/package=Hmisc

IOTC (Indian Ocean Tuna Commission). 2016. Report of the Fourteenth Session of the IOTC Working Party on Billfish. Victoria, Seychelles, 6-10 September 2016. IOTC–2016–WPB14–R[E]: 94 pp.

Lenth, R.V. 2016. Least-Squares Means: The R Package lsmeans. Journal of Statistical Software. 69(1): 1-33. doi:10.18637/jss.v069.i01

Maunder, M.N. & Punt, A.E. 2004. Standardizing catch and effort data: a review of recent approaches. Fisheries Research. 70: 141–159. doi:10.1016/j.fishres.2004.08.002.

Rodriguez-Marin, E., Arrizabalaga, H., Ortiz, M., Rodriguez-Cabello, C., Moreno, G. & Kell, L. T. 2003. Standardization of bluefin tuna, Thunnus thynnus, catch per unit effort in the baitboat fishery of the Bay of Biscay (Eastern Atlantic). ICES Journal of Marine Science, 60: 1216–1231. doi:10.1016/S1054–3139(03)00139-5.

Schwarz, G. 1978. Estimating the dimension of a model. Annals of Statistics. 6, 461-464.

Tserpes, G. & Tsimenides, N. 1995. Determination of age and growth of swordfish, Xiphias gladius L., 1758, in the Eastern Mediterranean using anal-fin spines. Fishery Bulletin. 93: 594-602

Venables, W. N. & Ripley, B. D. (2002) Modern Applied Statistics with S. Fourth Edition. Springer, New York. ISBN 0-387-95457-0

Zeileis, A., Kleiber, C. & Jackman, S. 2008. Regression Models for Count Data in R. Journal of Statistical Software 27(8). URL http://www.jstatsoft.org/v27/i08/.

Zuur, A.F., Saveliev, A.A., Ieno, E.N. 2012. Zero inflated models and generalized linear mixed models with R. Newburgh: Highland Statistics Limited. 324 pp.

Table 1. Summary of observed fishing effort from Indonesian tuna longline fishery during 2005–2016. Results are pooled and also presented by year of observation. Operational parameters are means (upper entries) and standard deviations (lower parenthetical entries).

Year	Trips	Sets	Days at Sea	Total Hooks	Hooks per Set		Hooks per Float		Mean Latitude		Mean Longitude	
2005	9	108	117	157,065	1,454.31	(151.8)	18.6	(1.5)	14.3ºS	(1.0º)	111.8ºE	(2.1º)
2006	13	401	401	577,243	1,439.51	(214.9)	11.2	(3.9)	16.9ºS	(6.0º)	113.4ºE	(5.4º)
2007	13	265	258	406,135	1,532.58	(326.5)	14.0	(4.4)	17.0ºS	(6.4º)	103.5ºE	(13.3º)
2008	15	370	404	483,662	1,307.19	(385.9)	13.0	(4.5)	14.2ºS	(2.6º)	107.3ºE	(14.1º)
2009	13	283	288	323,042	1,141.49	(234.7)	12.1	(4.9)	11.4ºS	(3.3º)	113.2ºE	(5.6º)
2010	6	165	152	220,394	1,335.72	(457.5)	13.6	(5.2)	12.0ºS	(3.3º)	113.3ºE	(6.0º)
2011	3	105	111	110,384	1,051.28	(173.9)	12.0	-	13.7ºS	(0.9º)	117.4ºE	(1.3º)
2012	8	198	192	290,265	1,465.98	(559.1)	14.1	(2.3)	18.9ºS	(7.8º)	104.5ºE	(10.8º)
2013	7	225	198	252,919	1,124.08	(210.4)	12.7	(2.1)	12.4ºS	(1.1º)	114.6ºE	(6.6º)
2014	5	167	265	193,740	1,160.12	(176.9)	15.0	(2.0)	11.0ºS	(1.7º)	105.7ºE	(7.5º)
2015	5	148	241	172,463	1,165.29	(145.2)	14.1	(3.2)	10.8ºS	(2.7º)	103.8ºE	(8.1º)
2016	3	130	170	175,868	1,352.83	(209.0)	11.3	(3.3)	9.2ºS	(4.3º)	106.0ºE	(7.4º)

Table 2. Summary of indicators as calculated using six model structures: Poisson (P), Negative Binomial (NB), Zero-inflated with Poisson (ZIP), Zero-inflated with Negative Binomial (ZINB), Hurdle with Poisson (HP), and Hurdle with Negative Binomial (HNB). The terms in the column at left indicate: number of parameters (k), Akaike (AIC) and Bayesian (BIC) Information Criteria, logarithm of the likelihood (logLik), number of predicted zero catches (zero), and *p* values as calculated using a Kolmogorov-Smirnov test.

Model Structure	P	NB	ZIP	ZINB	HP	HNB
k	75	45	44	46	90	88
AIC	4470.02	4330.40	4475.18	4400.26	4353.23	4340.54
BIC	4908.72	4599.47	4732.55	4669.33	4879.67	4855.28
logLIk	-2160.01	-2119.20	-2193.59	-2153.13	-2086.62	-2081.27
zero	1712	1780	1815	1830	1823	1823
p.value	0.016	0.864	0.999	1.000	1.000	1.000

Table 3. Deviance table of the parameters used for SWO CPUE standardization, using a negative binomial model (NB). Each parameter indicated the degrees of freedom (Df), the deviance (Dev), the residual degrees of freedom (Resid Df), the residual deviance (Resid. Dev), the Chi-square test statistic and the significance (p-value).

	Df	Deviance	Resid. Df	Resid. Dev	Pr(>Chi)	
NULL	2563	2420.8				
Year	11	98.711	2552	2322.1	3.21E-16	***
Start_Set	1	129.246	2551	2192.8	<2.20E-16	***
HBF	1	35.854	2550	2156.9	2.13E-09	***
Moon	3	35.166	2547	2121.8	1.12E-07	***
Soak_Time	1	11.575	2546	2110.2	0.0006686	***
Quarter	3	17.351	2543	2092.9	0.0005985	***
Latt	1	0.005	2542	2092.8	0.9430643	
Long	1	0.025	2541	2092.8	0.8747718	
Area	1	0.783	2540	2092	0.3762931	
Start_Set:Soak_Time	1	5.347	2539	2086.7	0.0207573	*
Start_Set:Long	1	4.118	2538	2082.6	0.0424273	*
HBF:Moon	3	15.644	2535	2066.9	0.0013415	**
HBF:Latt	1	0.807	2534	2066.1	0.3690882	
HBF:Long	1	10.096	2533	2056	0.0014857	**
HBF:Area	1	2.321	2532	2053.7	0.1276281	
Soak_Time:Quarter	3	9.955	2529	2043.8	0.0189515	*
Soak_Time:Long	1	16.614	2528	2027.1	4.58E-05	***
Quarter:Latt	3	27.117	2525	2000	5.56E-06	***
Quarter:Long	3	38.143	2522	1961.9	2.64E-08	***
Latt:Long	1	3.424	2521	1958.5	0.0642577	.
Latt:Area	1	7.825	2520	1950.6	0.0051539	**
Long:Area	1	3.31	2519	1947.3	0.0688788	.

Signif. codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Table 4. Nominal and standardized CPUEs (N/1000 hooks) of SWO using the Indonesia data in the eastern Indian Ocean. The point estimates, 95% confidence intervals and the standard deviation (SD) of the standardized index are presented, as well as the nominal CPUE values.

Year	Nominal CPUE	Standardized CPUE Index (N/1000 Hooks)			
		Estimate	sd	Lower CI (95%)	Upper CI (95%)
2005	0.154	0.185	0.054	0.104	0.328
2006	0.372	0.267	0.048	0.187	0.381
2007	0.314	0.190	0.043	0.122	0.294
2008	0.251	0.193	0.039	0.130	0.286
2009	0.531	0.365	0.073	0.247	0.540
2010	0.365	0.272	0.058	0.179	0.412
2011	0.122	0.146	0.046	0.079	0.271
2012	0.464	0.363	0.078	0.238	0.554
2013	0.367	0.353	0.065	0.246	0.506
2014	0.436	0.371	0.075	0.250	0.550
2015	0.261	0.175	0.042	0.110	0.279
2016	0.539	0.354	0.071	0.239	0.525

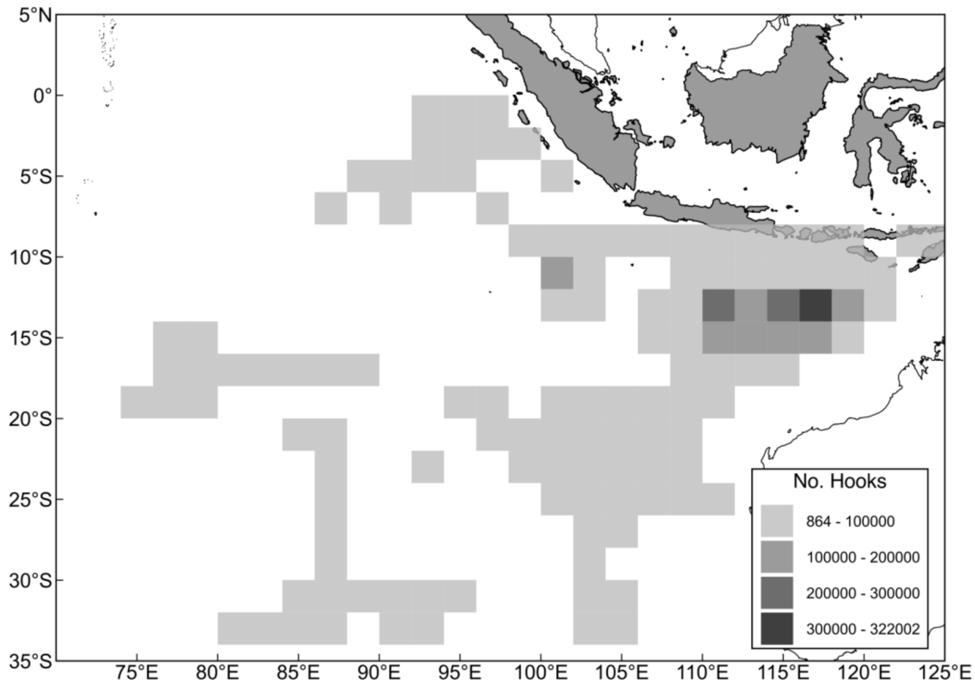

Figure 1. Distribution of the Indonesia observer data used in this SWO CPUE standardization. The effort is represented in 2*2 degree grids with darker and lighter colors representing respectively to areas with more and less effort in number of hooks.

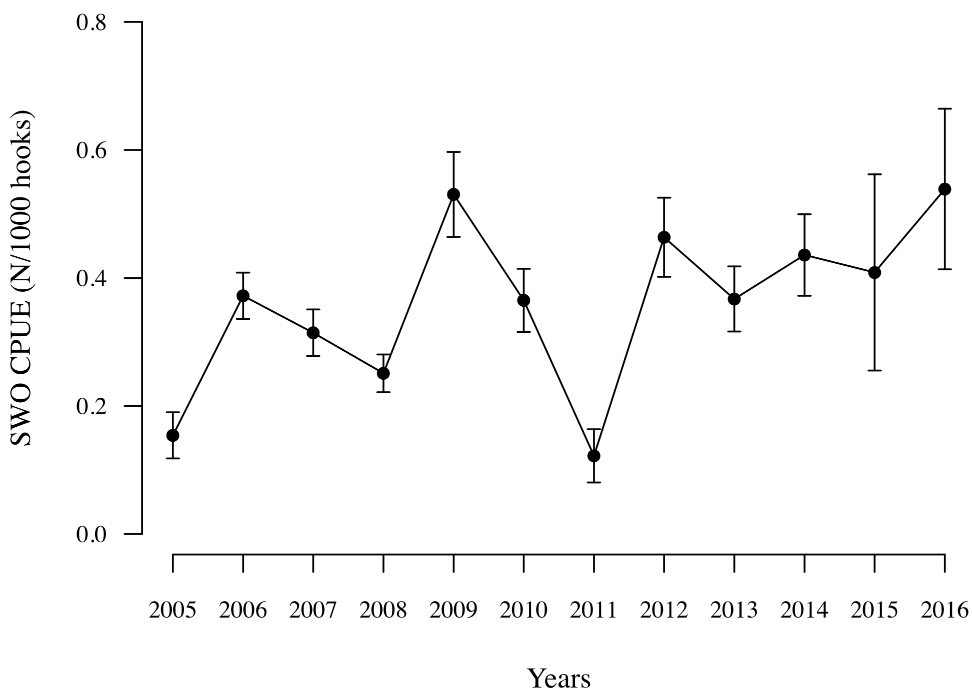

Figure 2. Nominal CPUE series (N/1000 hooks) for SWO in the Indonesia data, between 2005 and 2016. The error bars refer to the standard errors.

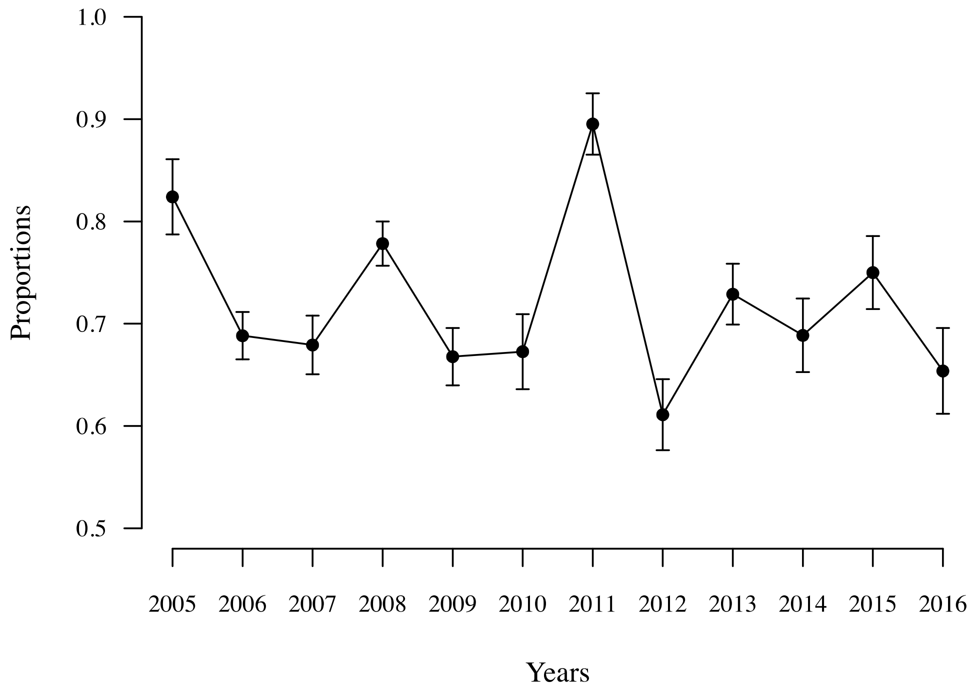

Figure 3. Proportion of zero SWO catches per set and per year, in the Indonesia data, between 2005 and 2016. The error bars refer to the standard errors.

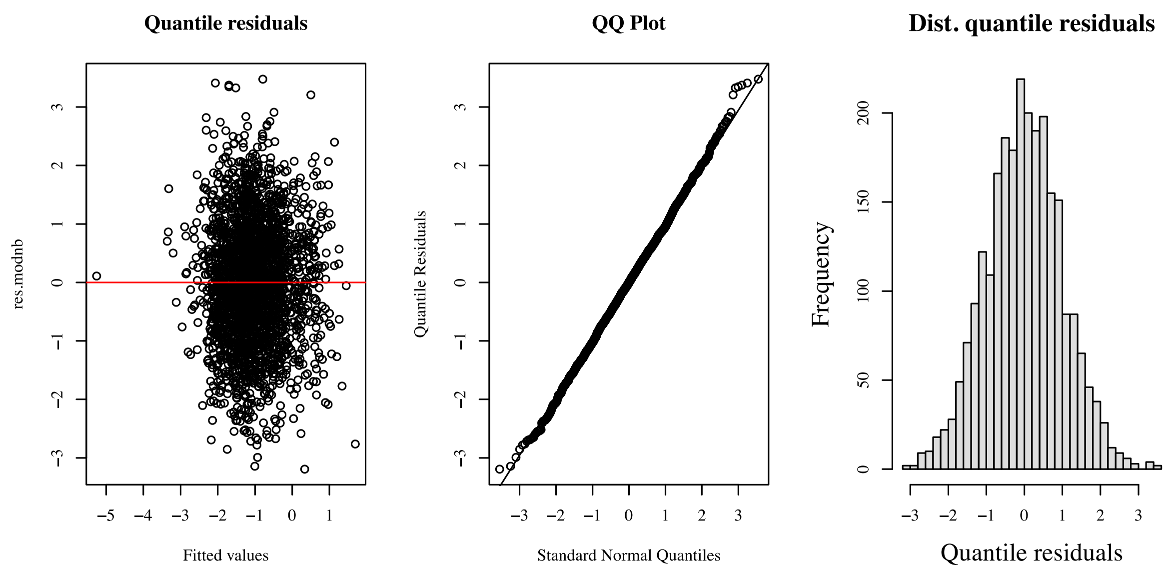

Figure 4. Residual analysis for the final SWO CPUE standardization model for the Indonesia data, between 2005 and 2016. In the plot, it is presented the histogram of the distribution of the residuals (right), the QQPlot (middle) and the residuals along the fitted values on the log scale (left).

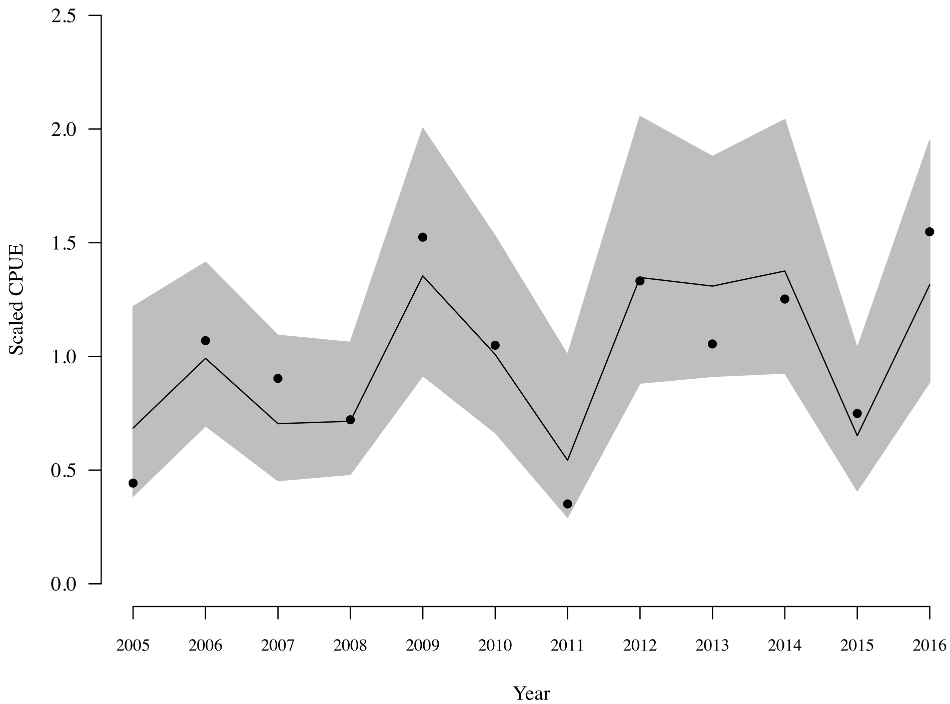

Figure 5. Standardized CPUE series for SWO from Indonesia data using a negative binomial model, between 2005 and 2016. The solid lines refer to the standardized index with the 95% confidence intervals, and the dots represent the nominal CPUE series. Both series are scaled by their means.