Update on the CPUE standardization of the shortfin mako shark caught by the Taiwanese large-scale tuna longline fishery in the Indian Ocean

Xing-Han Wu[1] and Wen-Pei Tsai[2,*]

[1] Department of Oceanography, National Sun Yat-sen University, Kaohsiung, Taiwan

[2] Department of Fisheries Production and Management, National Kaohsiung University of Science and Technology, Kaohsiung, Taiwan

[*]Corresponding author. Email: wptsai@nkust.edu.tw

SUMMARY

In the present study, the shortfin mako shark catch and effort data from the logbook data of Taiwanese large longline fishing vessels operating in the Indian Ocean from 2005-2020 were analyzed. Based on the effort distribution, four areas, namely, (1) Northwest Indian Ocean (north of 10°S, east of 70°E); (2) Northeast Indian Ocean (north of 10°S, 70°E-120°E); (3) Southwest Indian Ocean (south of 10°S, 20°E-60°E); (4) Southeast Indian Ocean (south of 10°S, 60°E-120°E) were categorized. To cope with the large percentage of zero shark catch, the catch per unit effort (CPUE) of shortfin mako shark, as the number of fish caught per 1,000 hooks, was standardized using zero-inflated negative binomial model (ZINB) that allows for “extra” zeros. ZINB model includes the main variables Year, Quarter, Area, HPBF, CTNO, and Cluster. The standardized CPUE showed a stable and slightly increasing trend for shortfin mako sharks. The stable trend suggested that shortfin mako shark stocks in the Indian Ocean seems at the level of optimum utilization.

KEYWORDS

Shortfin mako sharks, Taiwanese longline fishery, standardized CPUE, by-catch, zero-inflated negative binomial model

1. Introduction

Shortfin mako shark, *Isurus oxyrinchus*, is one of the most commonly caught shark species in the Taiwanese commercial offshore longline fishery and the major by-catch of tuna longline fisheries in the far seas. Shortfin mako is a large apex predator that exhibits slow growth, low fecundity and late maturity, and is particularly susceptible to exploitation owing to its life-history characteristics. Clarke et al. (2006) mentioned that about half a million shortfin mako sharks were utilized in the global shark fin trade in 2000. Given the high fishing pressure on this species and declining population trends, the shortfin mako is currently listed as "Vulnerable" on the IUCN Red List of Threatened Species (Dulvy et al., 2008), but very little is known about the stock status of this species in the Indian Ocean. Since the International organizations and regional fisheries management organizations (RFMO's) have concerned on the conservation of elasmobranchs in recent years, it is necessary to examine the recent trend of shark species by examining the logbook of tuna fisheries. Shortfin mako and blue shark (*Prionace glauca*) are the major shark species for Taiwanese large-scale tuna longline (LSTL) fisheries. Reliable catch estimate for shortfin mako shark can be developed because the logbook records of shortfin mako sharks were representative of actual catches as all sharks were retained due to its high market value. Thus, the objectives of this study are to standardize the CPUE of shortfin mako sharks in the Indian Ocean based on the logbook data.

A large proportion of zero values is commonly found in by-catch data obtained from fisheries studies involving counts of abundance or CPUE standardization. The zero-inflated negative binomial modeling, which can account for a large proportion of zero values than expected, is an appropriate approach to model "extra" zero data. Such "extra" zero catches could be attributable to reporting error or misidentifications, survey error (in which sharks were present at the site of a longline set but were not observed because the gear deployment did not overlap with the depth distribution of sharks or did not attract sharks), or both (Brodziak and Walsh,2013). As sharks are common by-catch species in the tuna longline fishery, the zero-inflated negative binomial model (ZINB) is commonly used in CPUE standardization to address these excessive zero catch of sharks. In this study, the CPUEs of shortfin mako sharks in the Indian Ocean were standardized using zero-inflated negative binomial model based on logbook data and hopefully these CPUE series can be used in the shortfin mako shark stock assessment.

2. Material and methods

2.1. Source of data

The species-specific catch data including tunas, billfishes, and sharks from logbook data in 2005-2020 were used to standardize CPUE of shortfin mako shark of Taiwanese large-scale longline fishery in the Indian Ocean. The summary of these data were shown in **Table 1**. The catch rate of shortfin mako sharks might be affected by spatial and temporal factors. We used the following stratification in our models. For temporal factors, we separated the data into 4 quarters: the 1st quarter (January to March), the 2nd quarter (April to June), the 3rd quarter (July to September), and the 4th quarter (October to

December). For spatial stratification, based on the effort distribution and fishing grounds of the target species (**Fig. 1**), four areas, namely, (1) Northwest Indian Ocean (north of 10°S, east of 70°E); (2) Northeast Indian Ocean (north of 10°S, 70°E-120°E); (3) Southwest Indian Ocean (south of 10°S, 20°E-60°E); (4) Southeast Indian Ocean (south of 10°S, 60°E-120°E) were categorized. Fishing strategy and target species also played a vital role in the influential effects. The number of hooks between floats (NHBF, Sallow < 6 <= Middle < 10 <= Deep < 16 <= Ultra Deep) and vessel size (Vessel, CT5; CT6; CT7) were considered and the cluster results were also incorporated as effects into the CPUE standardization models. For standardization, CPUE was calculated by set of operations based on logbook data during the period of 2005-2020.

2.2. Cluster analysis

Cluster analysis was based on species composition from logbook data. These species were albacore (*Thunnus alalunga*), bigeye tuna (*Thunnus obesus*), yellowfin tuna (*Thunnus albacares*), southern bluefin tuna (*Thunnus maccoyii*), billfish, sharks, and others. A two-step method suggested by He *et al.* (1997) was applied to process the numerous longline data sets. The data were aggregated by week to avoid excessive noise caused by clustering operational data. The clusters were then merged with operational set-by-set data by using columns of vessel ID and operation date (year, month, and week) to identify the targeted fishing operations. For the two-step method, nonhierarchical cluster analysis (K-means method; Johnson and Wichern, 1988),) was first applied to group the datasets into 42 clusters based on catch composition (P_2^7 which means 2 species can be chosen with priority from 7 species). Ward's agglomerative hierarchical cluster analysis was applied to the dissimilarity matrix to calculate the squared Euclidean distances based on the mean species composition from the 42 nonhierarchical clusters. The clusters were defined as groupings such that the difference in the relative variance between groups and within group was >50% (Wang, 2019).

2.3. CPUE standardization

Between 2005 and 2020, data from a total of 510,141 longline sets were collected, which amounted to a total effort of 1,647,604,158 hooks and yielded 93,363 shortfin mako sharks. A large proportion of sets with zero catch of shortfin mako sharks (about 90%) in the Indian Ocean was found in the logbook data. Hence, to address these excessive zero catches, the zero-inflated negative binomial model (ZINB) (Lambert, 1992) was applied to the standardization of shortfin mako shark CPUE. This zero-inflated negative binomial model is comprised of a counts model that allows for overdispersion in both the zeros and positive catches and a binomial model that allows for "extra" zeros (Zuur *et al.*, 2009, 2012; Brodziak and Walsh, 2013), with the latter defined as a higher frequency of zeros than expected under the Poisson, negative binomial, or other count distributions (Zuur *et al.*, 2009). The model was fit using zeroinfl function of statistical computing language R (R Development Core and Team, 2020) to eliminate some biases by change of targeting species, fishing ground and fishing seasons.

Standardized CPUE series for the shortfin mako shark was constructed including main effects and interaction terms. The main effects chosen as input into the ZINB analyses were Year, Quarter, Area, number of number of hooks between floats (HPBF), vessel size (CTNO), and Cluster effects. The probability distribution of a zero-inflated negative binomial random variable *Y* is given by:
(Part 1: Binomial model; Part 2: Count model –Negative Binomial, link = logit)

$$\Pr(Y=0)=\omega+(1-\omega)(1+k\lambda)^{1/\mathrm{k}}$$

$$\Pr(Y>0)=(1-\omega)\frac{\Gamma(y+1/k)}{\Gamma(y+1)\Gamma(1/k)}\frac{(k\mu)^{y}}{(1+k\lambda)^{y+1/k}}$$

where *k* is the negative binomial dispersion parameter, *λ* is the mean of the underlying negative binomial distribution, and *ω* is the probability of an observation being drawn from the constant distribution that always generates zeros.

3. Results and discussion

The shortfin mako shark bycatch data are characterized by many zero values. Overall, 89.31% of the total sets in the Indian Ocean had zero bycatch of shortfin mako sharks (**Table 2**). As a result, the following models with many explanatory variables were finally selected. The best models for ZINB model chosen by BIC values in the Indian Ocean were

Catch = Year + Quarter + Area + HPBF + Cluster + CTNO

The best models were then used in the later analyses.

Standardized CPUE series of the shortfin mako shark in the Indian Ocean using the ZINB model were shown in **Figure 3**. The detail values for nominal and standardized CPUE were listed in **Tables 3.** The nominal CPUE of shortfin mako shark in the Indian Ocean showed an inter-annual fluctuation, particularly in year 2005 and 2011 (**Fig. 3**). However, this variability was slightly smoothed in the standardized CPUE series. In general, the standardized CPUE series of the shortfin mako sharks caught by Taiwanese large-scale longline fishery showed a stable increasing trend (**Fig. 3**). These stable trends suggested that the shortfin mako shark stock in the Indian Ocean seems at the level of optimum utilization during the period of 2005-2020.

The residual results from the ZINB model do not indicate severe departure from model assumptions (**Figs.** 4). The ANOVA tables for each model are given in **Table 1**. Most main effects and interaction terms tested were significant (mostly $P < 0.01$) and have been included in the final model. However, other factors may affect the standardization of CPUE trend. In addition to the temporal and spatial effects, environmental factors are important which may affect the representation of standardized CPUE of pelagic fish i.e., swordfish and blue shark in the North Pacific Ocean (Bigelow *et al.,* 1999), and big-eye tuna in the Indian Ocean (Okamoto *et al.,* 2001). In this report, environmental effects were not included in the model for standardization. The results obtained in this study can be improved if longer time series of logbook data are available and environmental factors were included in the model.

References

Bigelow, K.A., Boggs, C.H., and He, X., 1999. Environmental effects on swordfish and blue shark catch rates in the US North Pacific longline fishery. Fish. Oceanogr. 8(3): 178-198.

Brodziak, J., Walsh, W.A., 2013. Standardizing catch rates of bycatch species usingmultimodel inference: A case study of oceanic whitetip shark in the Hawaiilongline fishery. Can. J. Fish. Aquat. Sci. 70, 1723–1740. Cook, R. D., and Weisberg, S., 1982. Residuals and influence in regression, New York, NY: Chapman and Hall.

Clarke, S. C., McAllister, M. K., Milner-Gulland, E. J., Kirkwood, G. P., Michielsens, C. G. J., Agnew, D. J., Pikitch, E. K., Nakano, H., and Shivji, M. S., 2006. Global estimates of shark catches using trade records from commercial markets. Ecology Letters, 9: 1115-1126.

Dulvy, N. K., Baum, J. K., Clarke, S., Compagno, L. J. V., Cortés, E., Domingo, A., and Fordham, S., et al., 2008. You can swim but you can't hide: the global status and conservation of oceanic pelagic sharks and rays. Aquatic Conservation: Marine and Freshwater Ecosystem, 18: 459-482.

He, X., Bigelow, K.A., Boggs, C.H., 1997. Cluster analysis of longline sets and fishing strategies within the Hawaii-based fishery. Fisheries Research 31, 147–158.

Huang, H.-W., and K.-M. Liu., 2010. Bycatch and Discards by Taiwanese Large-Scale Tuna Longline Fleets in the Indian Ocean. Fisheries Research 106(3): 261-270.

Johnson, R., Wichern, K., 1988. Applied multivariate statistical analysis, second ed. Prentice Hall, New York.

Lambert, D., 1992. Zero-inflated Poisson regression, with an application to defects in manufacturing. Technometrics, 34(1), 1-14.

Okamoto, H., Miyabe, N., and Matsumoto, T., 2001. GLM analyses for standardization of Japanese longline CPUE for bigeye tuna in the Indian Ocean applying environmental factors. IOTC Proceedings 4: 491-522.

R Development Core Team., 2020. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.

Wang, S.P., 2019. CPUE Standardization of blue marlin caught by Taiwanese large scale longline fishery in the Indian ocean. IOTC–2019–WPB17–18.

Ward, J.H., 1963. Hierarchical grouping to optimise an objective function. J. Amer. Statist. Assoc. 58, 236–244.

Zuur, A.F., Ieno, E.N., Walker, N.J., Saveliev, A.A., Smith, G.M., 2009. Mixed EffectsModels and Extensions in Ecology with R. Springer Science and Business Media,New York, NY.

Zuur, A.F., Saveliev, A.A., Ieno, E.N., 2012. Zero Inflated Models and Generalized LinearMixed Models with R. Highland Statistics Ltd, Newburgh, United Kingdom.

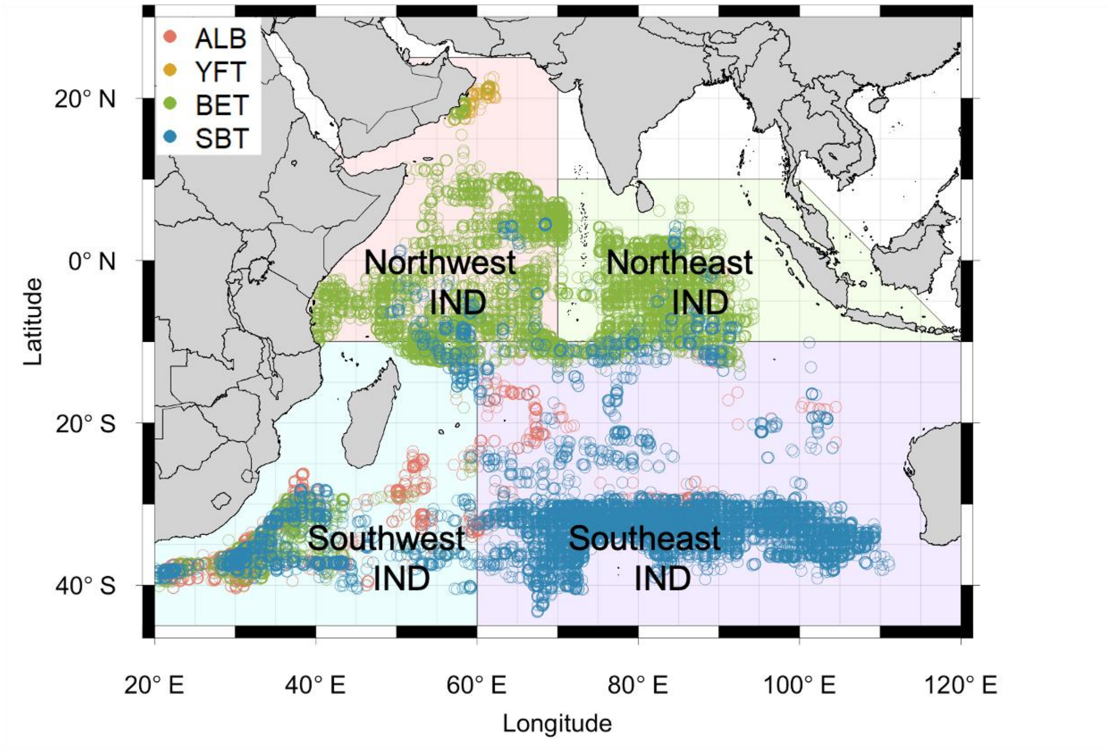


Figure 1. Area stratification used in this study based on Taiwanese large-scale tuna longline effort distribution and targeted species as recorded by observers (ALB= albacore; YFT= yellowfin tuna; BET= bigeye tuna; SBT= southern bluefin tuna).

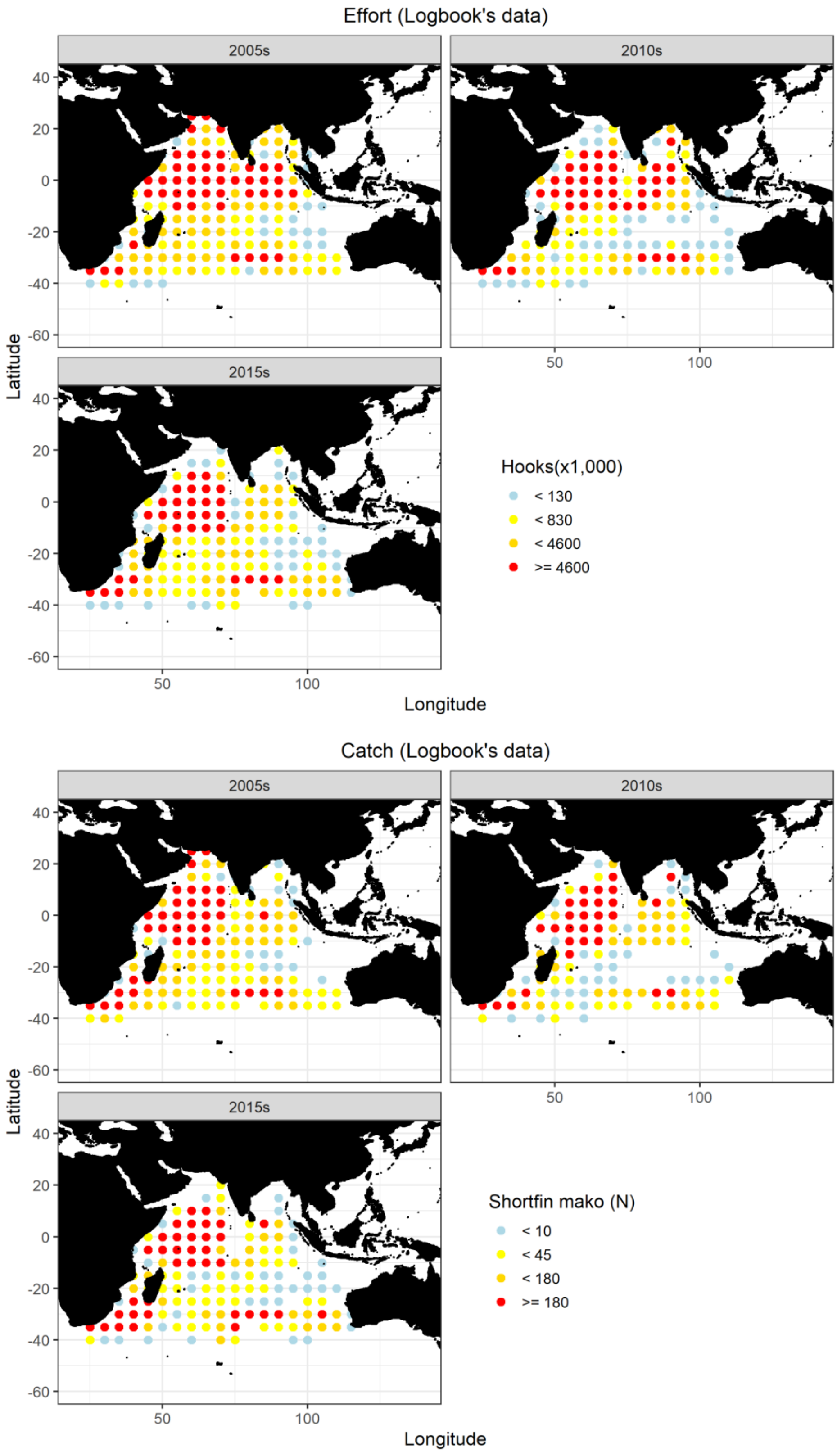


Figure 2. Logbook's efforts and shortfin mako catches distributions of the Taiwanese large-scale tuna longline fishery in the Indian Ocean from 2005 to 2020. (2005s: 2005-2009; 2010s: 2010-2014; 2015s: 2015-2020).

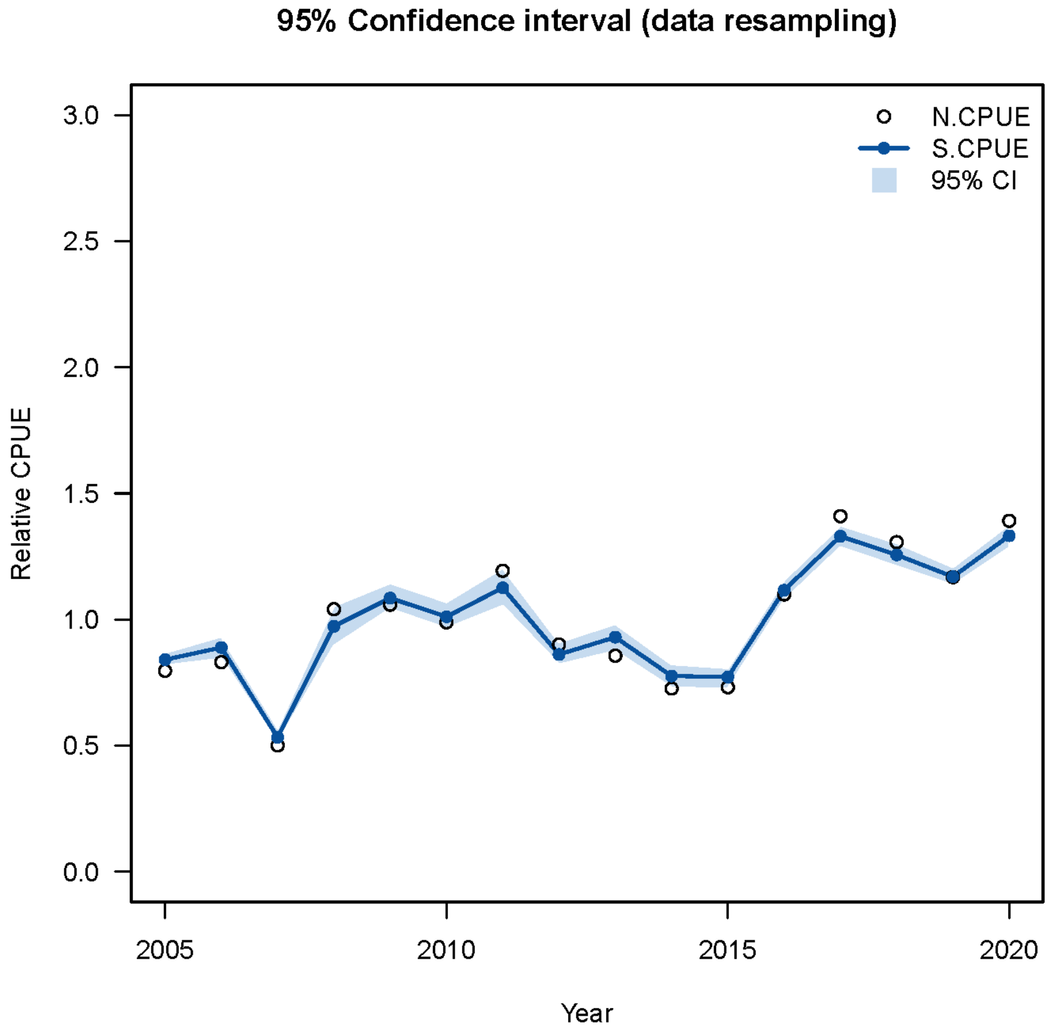


Figure 3. Logbook nominal and standardized CPUE with 95% CI of shortfin shark by Taiwanese longline vessels in the Indian Ocean from 2005 to 2020

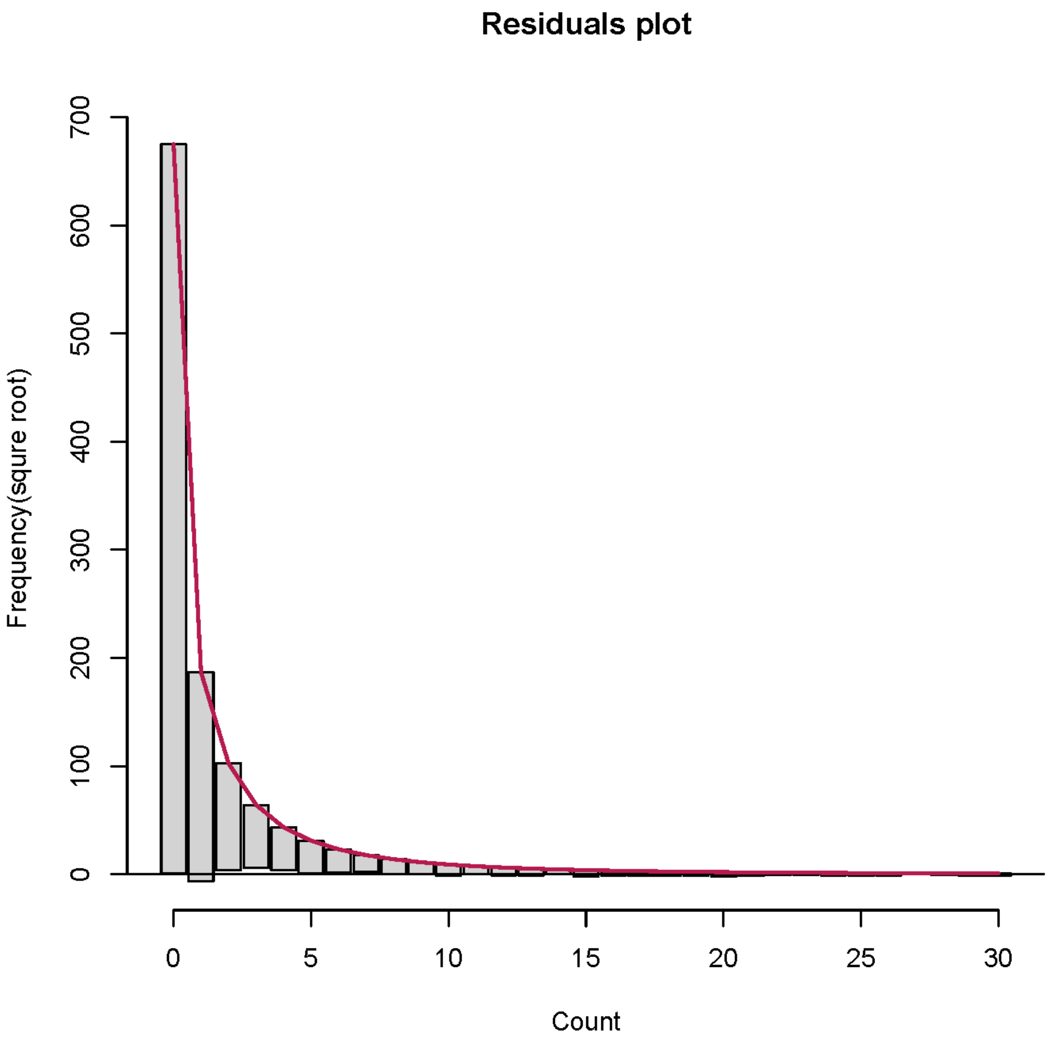


Figure 4. Residual plots for the ZINB model fit to the Indian Ocean longline shortfin mako shark bycatch data.

Table 1. Summary of information of the logbook data used in this study.

Year	Indian Ocean		
	No. of Hooks	No. of Sets	No. of Shortfin mako shark catches
2005	229,125,876	72,206	10,774
2006	111,539,175	34,699	5,398
2007	141,462,466	44,026	4,120
2008	102,533,017	31,810	6,189
2009	129,191,560	40,105	7,932
2010	97,638,819	29,863	5,533
2011	73,003,298	22,551	5,093
2012	76,970,711	25,285	4,398
2013	75,819,812	23,724	3,807
2014	58,376,963	18,475	2,514
2015	70,899,449	22,537	3,089
2016	101,592,087	31,567	6,506
2017	99,408,067	29,983	7,921
2018	93,070,520	28,034	6,864
2019	97,308,263	28,692	6,280
2020	89,664,075	26,584	6,945
Total	1,647,604,158	510,141	93,363

Table 2. The logbook percentage of zero-catch of shortfin mako shark for Taiwanese tuna longline vessels in the Indian Ocean from 2005 to 2020.

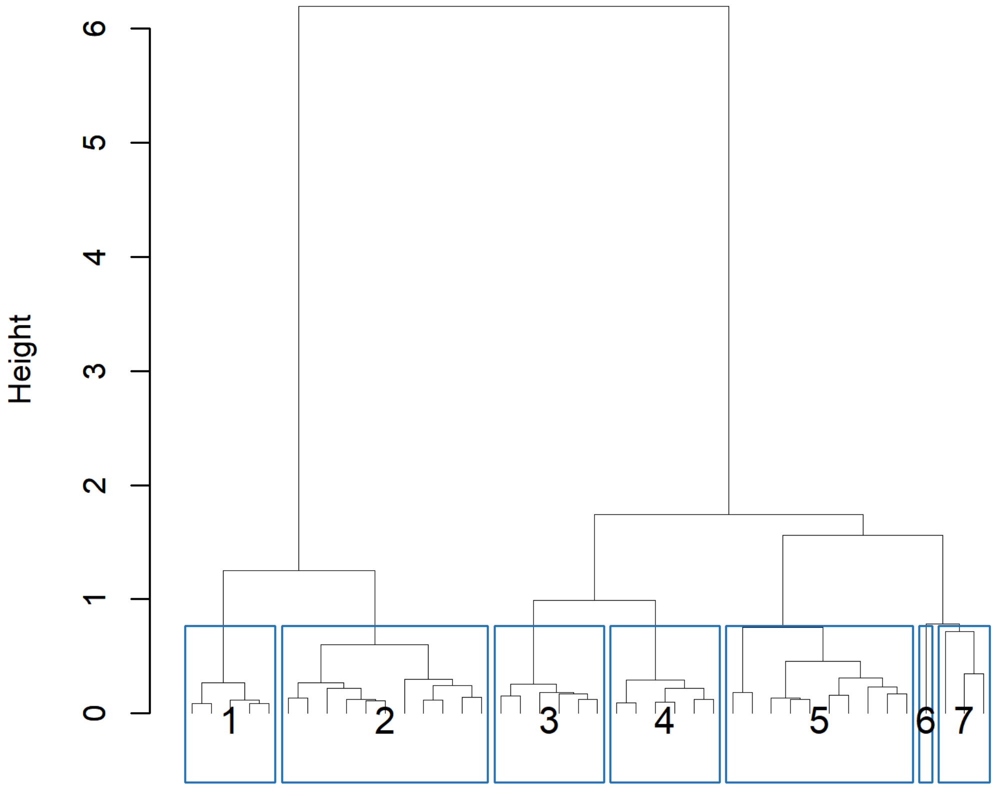
Year	Percentage of zero-catch
2005	90.44%
2006	91.51%
2007	94.66%
2008	93.38%
2009	90.88%
2010	91.15%
2011	90.17%
2012	90.21%
2013	89.52%
2014	90.48%
2015	90.36%
2016	86.12%
2017	83.56%
2018	83.97%
2019	84.03%
2020	83.30%
Average	89.31%

Table 3. Estimated nominal and standardized CPUE values for shortfin mako shark of the Taiwanese tuna longline fishery in the Indian Ocean.

Year	Original Values	
	Nominal	Standardized
2005	0.14921	0.15545
2006	0.15566	0.16439
2007	0.09377	0.09850
2008	0.19505	0.17996
2009	0.19833	0.20060
2010	0.18528	0.18693
2011	0.22359	0.20814
2012	0.16880	0.15927
2013	0.16050	0.17195
2014	0.13608	0.14350
2015	0.13708	0.14274
2016	0.20599	0.20633
2017	0.26411	0.24588
2018	0.24486	0.23229
2019	0.21885	0.21628
2020	0.26069	0.24630

Table 4 Deviance tables for the ZINB model of shortfin mako.

Zero-inflated negative binomial model				
Source	Df	X^2	Pr(>X^2)	
Year	15	2135.84	< 0.001	***
Quarter	3	640.66	< 0.001	***
Area	3	2192.78	< 0.001	***
Cluster	6	127.76	< 0.001	***
HPBF	2	790.01	< 0.001	***
CTNO	2	43.83	< 0.001	***
Signif. codes:	0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1			



Appendix Fig. 1. Dendrogram of 42 nonhierarchical clusters for 510,141 longline sets of the Taiwanese large-scale tuna longline fishery in the Indian Ocean from 2005 to 2020.