CPUE standardization of blue marlin (*Makaira nigricans*) caught by Taiwanese large-scale longline fishery in the Indian Ocean

Chih-Yu Lin, Sheng-Ping Wang*, Wen-Qi Xu

Department of Environmental Biology and Fisheries Science, National Taiwan Ocean University, Taiwan.
* Corresponding author: wsp@mail.ntou.edu.tw

Paper submitted to the 20th meeting of the Working Party on Billfish, Indian Ocean Tuna Commission, 12-15 September, 2022.

ABSTRACT

This paper briefly described the historical patterns of blue marlin catches caught by Taiwanese large-scale longline fishery in the Indian Ocean. The cluster analysis was adopted to explore the targeting of fishing operations. In addition, the CPUE standardizations were conducted using delta-general linear models with different assumptions of the error distributions. Based on the diagnostic statistics and trend of model fits, the standardized CPUE series obtained based on the delta GLM with inverse gaussian error distribution for positive catches would be recommended by this study. The results indicate that the effect of latitude provided the most significant contributions to the explanation of the variance of CPUE for positive catches and delta models for both northern areas (NW and NE), except for the year effects. The standardized CPUE series in both northern areas revealed decreasing trends in recent years.

1. INTRODUCTION

Blue marlin is largely considered to be a non-target species of industrial and artisanal fisheries. Longline catches account for around 68% of total catches in the Indian Ocean, followed by gillnets (15%), with remaining catches recorded under troll and handlines. Based on the catches data from 2012 to 2021, main fleets consisted of Taiwan (longline, 43%), Sri Lanka (gillnet, handline and longline, 21%), Indonesia (gillnet, handline and longline, 7%). Catches reported by drifting longliners were more or less stable until the late-70's, at around 3,000 t to 4,000 t, and have steadily increased since then to reach values between 8,000 t and to over 10,000 t since the early 1990's. The highest catches reported by longliners have been recorded between

2012 and 2016, and are likely to be the consequence of higher catch rates by some longline fleets which appear to have resumed operations in the western tropical Indian Ocean. (IOTC, 2021).

IOTC conducted a stock assessment for blue marlin in the Indian Ocean in 2019 and the results indicated that the current stock status of blue marlin have been overfished and overfishing (IOTC, 2021). Therefore, this study conducted CPUE standardization for blue marlin in the Indian Ocean for providing the relative abundance indices for further stock assessment.

2. MATERIALS AND METHODS

2.1. Catch and Effort data

In this study, daily operational catch and effort data (logbook) with 5x5 degree longitude and latitude grid for Taiwanese longline fishery during 1979-2021 were provided by Oversea Fisheries Development Council of Taiwan (OFDC). For the area stratification, this study adopted the four areas stratification for swordfish by Wang and Nishida (2011) (Fig. 1).

As the discussions and suggestions from previous IOTC meetings (IOTC, 2021), Taiwanese data before 2005 were recommended not using to analyze the targeting of fishing operations and conduct the CPUE standardization for billfish due to the problem of data quality. However, the data problem might not only influence the misreport for the catches of major tropical tunas but also lead to uncertainties in the catch and effort data for other species. Therefore, CPUE standardizations were conducted using the data from 2005 to 2021 as suggested in previous meetings.

2.2. Cluster analysis

The details of the procedures of cluster analysis were described in Wang et al. (2021). This study adopted a direct hierarchical clustering with agglomerative algorithm, which brings a fast and efficient implementation through features of memory-saving routines in hierarchical clustering of vector data (Müllner, 2013). The trials conducted using R function “hculst.vector” of package “fastcluster” (Müllner 2021) with Ward's minimum variance linkage methods (“ward.D” for the argument “method” in “hclust.vector” of R function) applied to the squared Euclidean distances between data points calculated based on the species composition.

In this study, the number of clusters was selected based on the elbow method, i.e. the change in deviance between/within clusters against different numbers of clusters. The number of clusters was determined when the improvement in the sum of within-cluster variations was less than 10%.

2.3. CPUE Standardization

Because blue marlin was bycatch species of Taiwanese longline fishery, large amount of zero-catches was recorded in the operational catch and effort data sets. In previous studies, ignoring zero observations or replacing them by a constant was the most common approach. Nowadays, an alternative and popular way to deal with zeros was through the delta approach (Hinton and Maunder, 2004; Maunder and Punt, 2004). IOTC (2016) also noted that the use of the delta approach is appropriate for high proportion of zero catches. Therefore, the delta-general linear models with different assumption of error distribution were adopted in this study (Pennington, 1983; Lo et. al., 1992; Pennington, 1996; Andrade, 2008; Lauretta et al., 2016; Langley, 2019).

As the approach of Xu et al. (2021), the main effects, including year, quarter, longitude, latitude and fishing targeting (clusters), were first considered in the model, and then the interactions between main effects with significance were incorporated into the models. Hinton and Maunder (2004) indicated that interactions with the year effect would invalidate the year effect as an abundance index. Therefore, the interactions associated with the year effect were not considered in the model. The collinearity diagnostics were also conducted for all of the main effects and interactions using generalized variance-inflation factors (GVIF, Fox and Monette, 1992; Fox and Weisberg, 2019). In this study, the main effects or interactions with the value of GVIF^1/2df less than 5 were retained in the model.

CPUE standardizations were performed by areas separately (Fig. 1). The models for positive catches and delta models were conducted as follows:

For CPUE of positive catches:

$$Catch = \mu + Y + Q + CT + Lon + Lat + T + offset(\log(Hooks)) + \text{interactions} + \varepsilon^{pos}$$

Delta model for presence and absence of catches:

$$PA = \mu + Y + Q + CT + Lon + Lat + T + Hook + \text{interactions} + \varepsilon^{del}$$

where	*Catch*	is the catch in number/1,000 hooks,
	PA	is the presence/absence of catch,
	Hooks	is the effort of 1,000 hooks,
	μ	is the intercept,

Y	is the effect of year,
Q	is the effect of quarter,
CT	is the effect of vessel scale,
T	is the effect of targeting (cluster),
Hook	is the effect of number of hooks,
ε^{pos}	is the error term assumed based on various distribution,
ε^{del}	is the error term, ε^{del} ~ Binomial distribution.

To examine the appropriateness of the assumption of error distribution, this study applied normal, gamma and inverse gaussian distributions to the error distribution of the model for the positive catches and specified “log” for the model link function. The stepwise search ("both" direction, i.e. "backward" and "forward") based on the Akaike information criterion (AIC) were performed for selecting the explanatory variables for each model. Then, the coefficient of determination (R^2), and Bayesian information criterion (BIC) were calculated for the models with selected explanatory variables.

The standardized CPUE series were calculated based on the estimates of least square means of the interaction between the effects of year and area, and calculated by the product of the standardized CPUE of positive catches and the delta model:

$$DL^{index} = e^{\log(CPUE)} \times \left(\frac{e^{PA}}{1+e^{PA}} \right)$$

where DL^{index} is the standardized CPUE

3. RESULTS AND DISCUSSION

3.1. Fishing trends

Figs. 2 and 3 show the distributions of catch (numbers) and CPUE of blue marlin based on the logbook data of Taiwanese large-scale longline fishery in the Indian Ocean. Blue marlin was mainly caught in tropical and subtropical of Indian Ocean. High CPUE occurred in the tropical and subtropical of Indian Ocean.

The blue marlin catches were mainly caught with high effort in northern waters, especially for the area NW. Although the catches in the northwestern Indian Ocean increased significantly around 2012, 2015 and 2016, the catches substantially decreased int the following years (Fig. 4 and Fig. 5).

3.2. Cluster analysis

Based on the patterns from the elbow method, the determined numbers of clusters were 4 for all sub-areas (Figs. 6 and 7). For each sub-area, the species compositions revealed different patterns by clusters (Fig. 8). The main target species, such as yellowfin tuna, bigeye tuna and albacore, can be grouped by clusters and some other species can be also grouped in a particular area (e.g. other species and sharks in the area SE and SW). In general, clusters mainly consisted of yellowfin tuna and bigeye tuna in the northern areas (NW and NE), while albacore and other species were the major species in the southern areas (SW and SE).

Fig. 9 shows the annual trends of catches of blue marlin and fishing efforts. Because blue marlin was caught as bycatch, their catches were grouped into different clusters when the levels of fishing efforts changed with time periods. In addition, the proportion of zero-catch during 2005 to 2021 was very high for blue marlin in all areas. Even though the catch of blue marlin was relatively high in northern area, the proportion of zero-catch was more than 60% (Fig. 10).

3.3. CPUE standardization

CPUE standardizations were separately conducted for only northern areas (NW and NE, Fig. 1) since the catches and CPUE of blue marlin in the southern areas were much lower than those in the northern areas (Figs. 2-4).

Based on the model selection of AIC for positive catches and delta models, some of main effects or interactions were excluded by GVIF or AIC. For the models for positive catches, the models with inverse gaussian error distribution would be the optimal models for all areas based on the values of AIC and BIC although the values of R^2 were not higher than other models (Table 1). In addition, residual diagnostic plots also show that the models with inverse gaussian error distribution (Fig. 11) should be most appropriate than other models because there were less increasing or decreasing trends in the range of predicted values (plots for other models by areas were not shown here but the residuals revealed obvious patterns with predicted values). Therefore, the delta-inverse gaussian model were selected to produce the standardized CPUE series for further stock assessment.

The ANOVA tables for selected models for each area are shown in the Table 2. Except for the effect of year, the effects of *Lat* provided the most significant contributions to the explanation of the variance of the CPUE for both positive catch and delta models and for both NW and NE areas. This implied that the catch amount and opportunity might be mainly determined by the latitude that blue marlin and fishing vessels distributed.

The area-specific standardized CPUE series obtained from various models are shown in Fig. 12 and the CPUE series revealed similar trends for all model. The standardized CPUE of positive catches and catch probability obtained from the selected model are shown in Fig. 13 and CPUE of positive catches and catch probability revealed similar trends for both areas.

The standardized CPUE series with 95% confidence intervals obtained from the selected model are shown in Fig. 14. The CPUE series in area NW revealed increasing trends since 2005, decreasing around 2012-2014 and substantially decreased after a peak occurred in 2015. The CPUE series in area NE gradually increased since 2005 and decreased in recent years.

REFERENCE

Andrade, H.A., 2008. Using delta-gamma generalized linear models to standardize catch rates of yellowfin tuna caught by Brazilian bait-boats. ICCAT SCRS/2008/166.

Fox, J., Monette, G., 1992. Generalized collinearity diagnostics. J. Am. Stat. Assoc., 87: 178–183.

Fox, J., Weisberg, S., 2019. An R companion to applied regression, Third Edition. Thousand Oaks CA: Sage.

Hinton, M.G., Maunder, M.N., 2004. Methods for standardizing CPUE and how to select among them. Col. Vol. Sci. Pap. ICCAT, 56: 169-177.

IOTC, 2016. Report of the 14th Session of the IOTC Working Party on Billfish. IOTC-2016-WPB14-R[E].

IOTC, 2021. Report of the 19th Session of the IOTC Working Party on Billfish. 13–16 September 2021, Microsoft Teams Online. IOTC–2021–WPB19–R[E].

Langley, A.D., 2019. An investigation of the performance of CPUE modelling approaches – a simulation study. New Zealand Fisheries Assessment Report 2019/57.

Lauretta, M.V., Walter, J.F., Christman, M.C., 2016. Some considerations for CPUE standardization; variance estimation and distributional considerations. ICCAT Collect. Vol. Sci. Pap. ICCAT, 72(9): 2304-2312.

Lo, N.C.H., Jacobson, L.D., Squire, J.L., 1992. Indices of relative abundance from fish spotter data based on delta-lognormal models. Can. J. Fish. Aquat. Sci., 49: 2515-2526.

Maunder, N.M., Punt, A.E., 2004. Standardizing catch and effort data: a review of recent approaches. Fish. Res., 70: 141-159.

Müllner, D., 2013. fastcluster: Fast Hierarchical, Agglomerative Clustering Routines for R and Python. Journal of Statistical Software, 53(9): 1-18.

Müllner, D., 2021. The fastcluster package: User’s manual, Version 1.2.3. https://cran.r-project.org/web/packages/fastcluster/vignettes/fastcluster.pdf

Pennington, M., 1983. Efficient estimation of abundance, for fish and plankton surveys. Biometrics, 39: 281-286.

Pennington, M., 1996. Estimating the mean and variance from highly skewed marine data. Can. J. Fish. Aquat. Sci., 94: 498-505.

Wang, S.P., Nishida, T., 2011. CPUE standardization of swordfish (*Xiphias gladius*) caught by Taiwanese longline fishery in the Indian Ocean. IOTC–2011–WPB09–12.

Wang, S.P., Xu, W.Q., Lin, C.Y., Kitakado, T., 2021. Analysis on fishing strategy for target species for Taiwanese large-scale longline fishery in the Indian Ocean. OTC–2021–WPB19–11.

Xu, W.Q., Wang, S.P., Lin, C.Y., 2021. CPUE standardization of striped marlin (*Tetrapturus audax*) caught by Taiwanese large-scale longline fishery in the Indian Ocean. IOTC–2021–WPB19–13_Rev1.

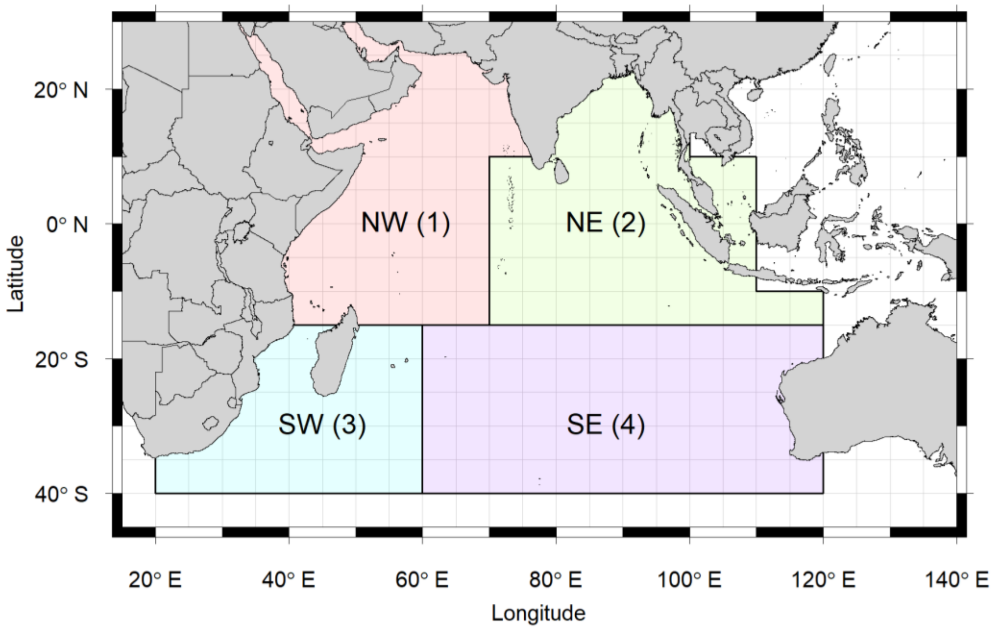


Fig. 1. Area stratification for swordfish in the Indian Ocean.

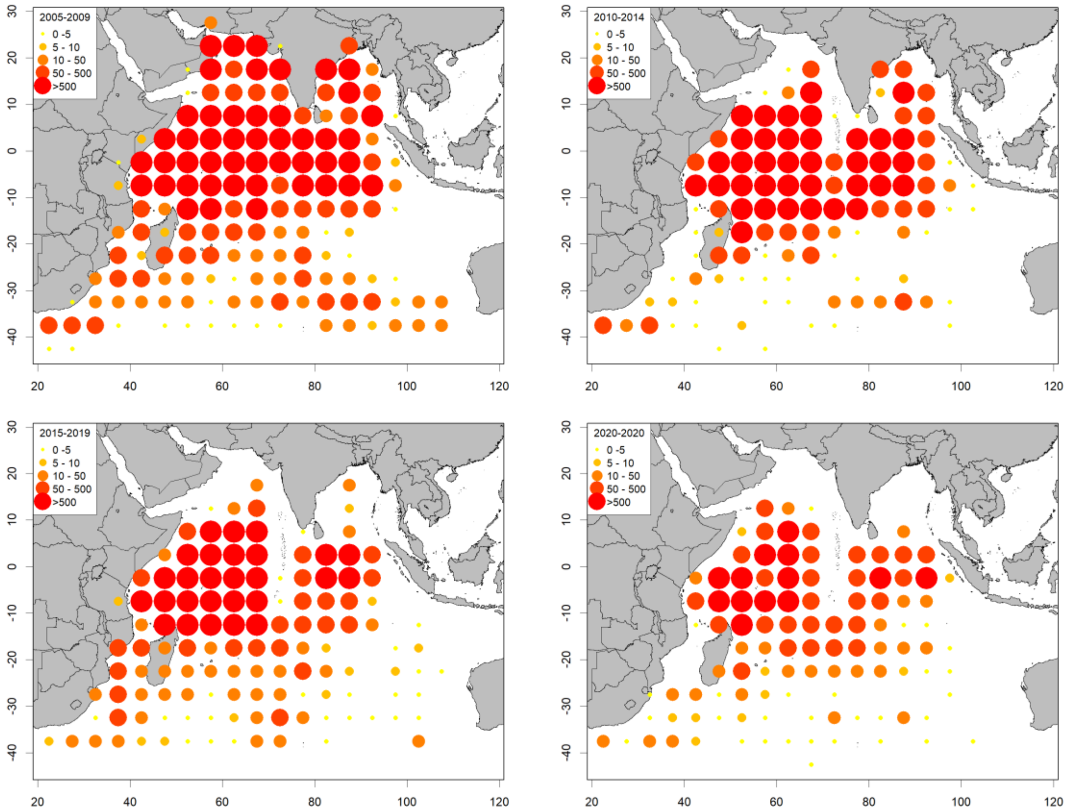


Fig. 2. Blue marlin catch distribution of Taiwanese large-scale longline fishery in the Indian Ocean.

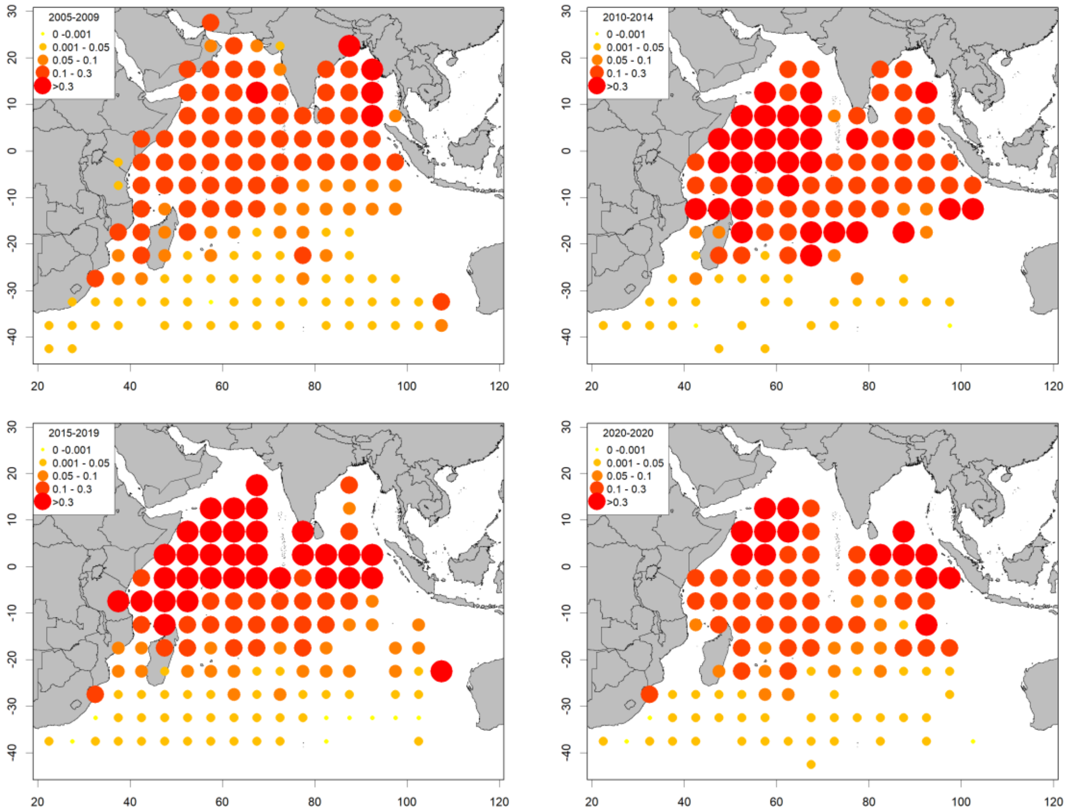


Fig. 3. Blue marlin CPUE distribution of Taiwanese large-scale longline fishery in the Indian Ocean.

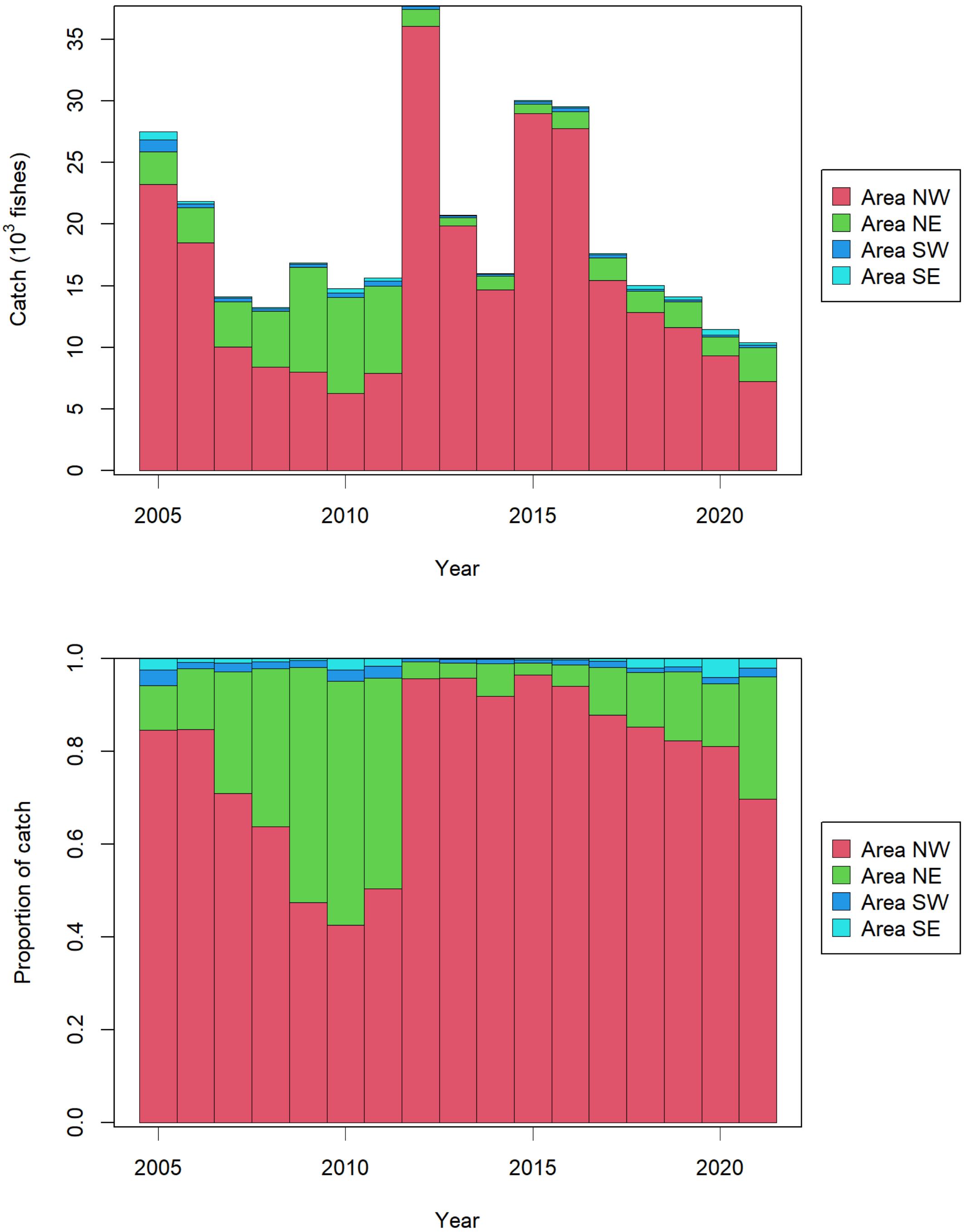


Fig. 4. Annual blue marlin catches of Taiwanese large-scale longline fishery in the Indian Ocean.

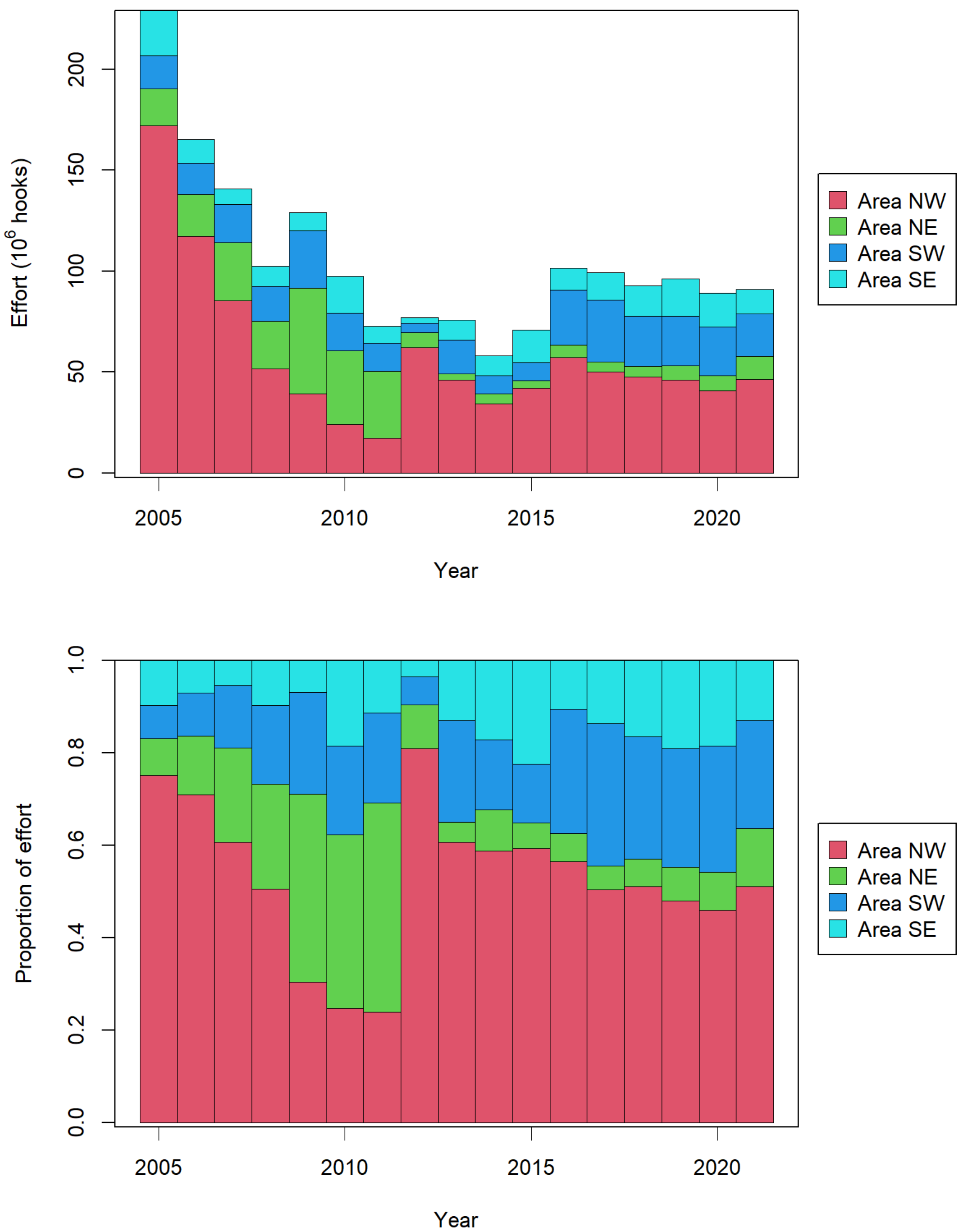
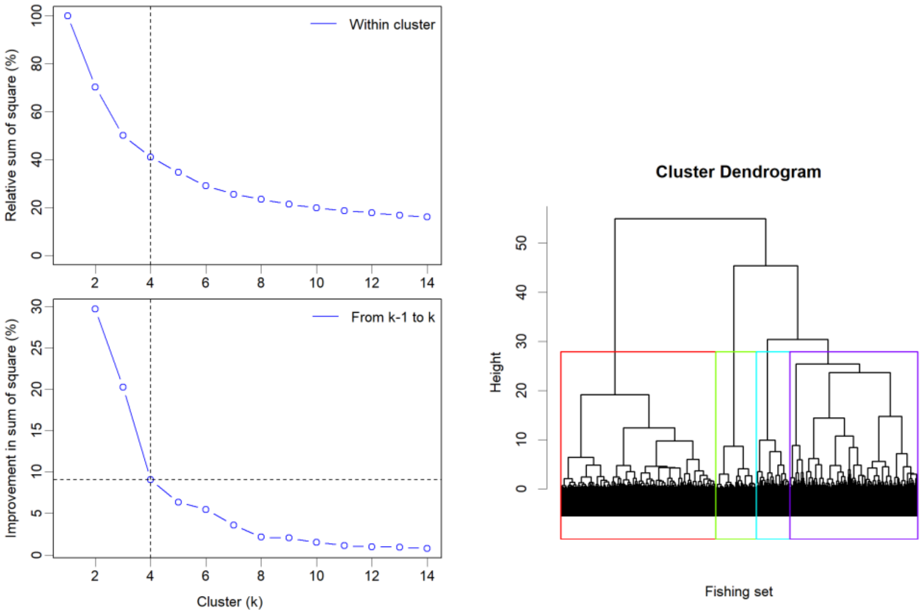


Fig. 5. Annual efforts (number of hooks) of Taiwanese large-scale longline fishery in the Indian Ocean.

NW



NE

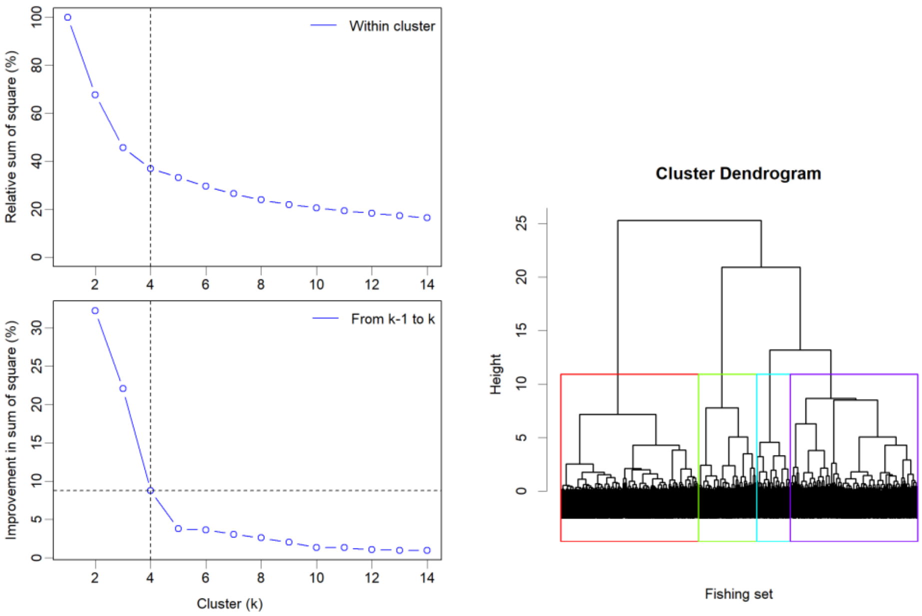
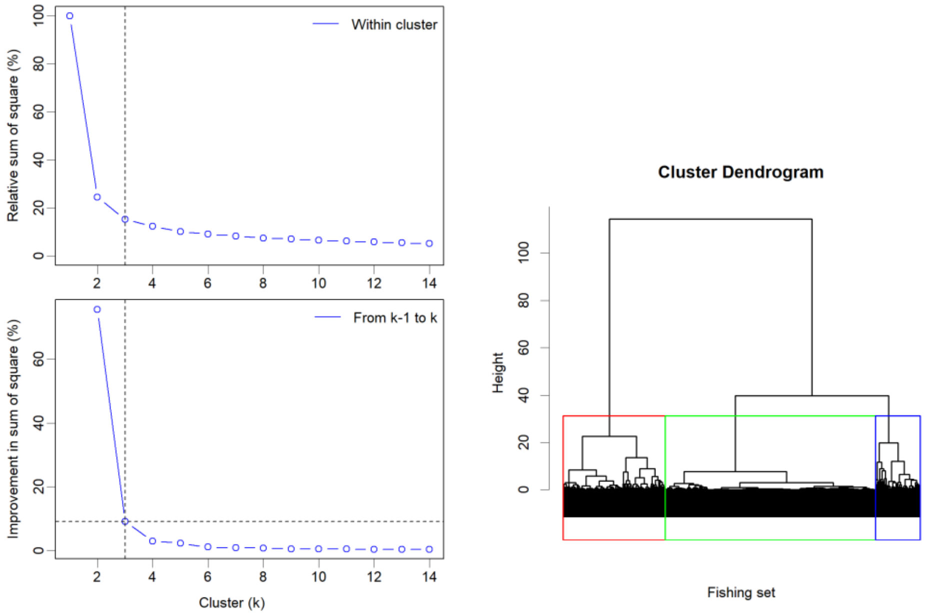


Fig. 6. Sum of squares within clusters for the data of Taiwanese large-scale longline fishery in billfish area of the Indian Ocean.

SW



SE

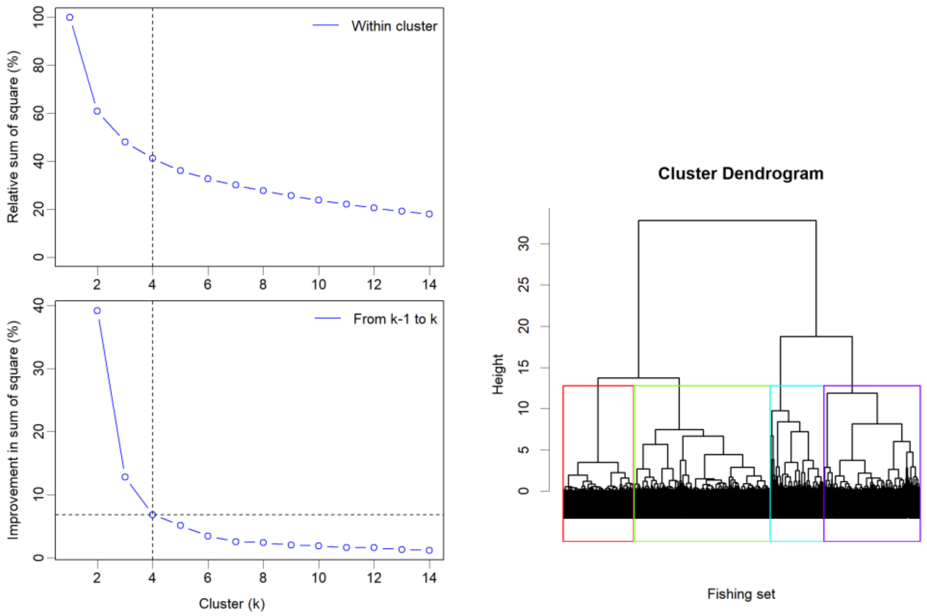


Fig. 6. (Continued).

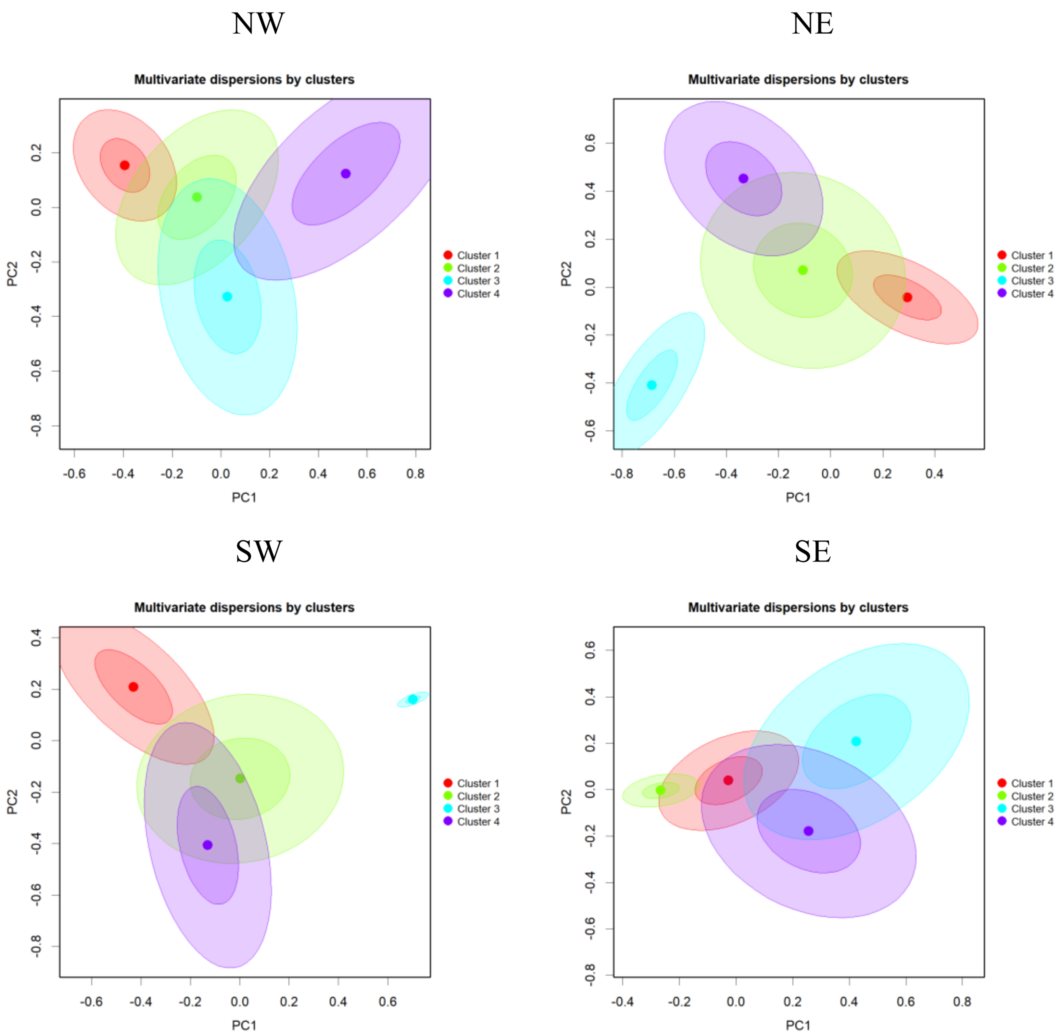


Fig. 7. Multivariate dispersions of the centroids by clusters derived from PCA for the data of Taiwanese large-scale longline fishery in billfish area of the Indian Ocean.

NW

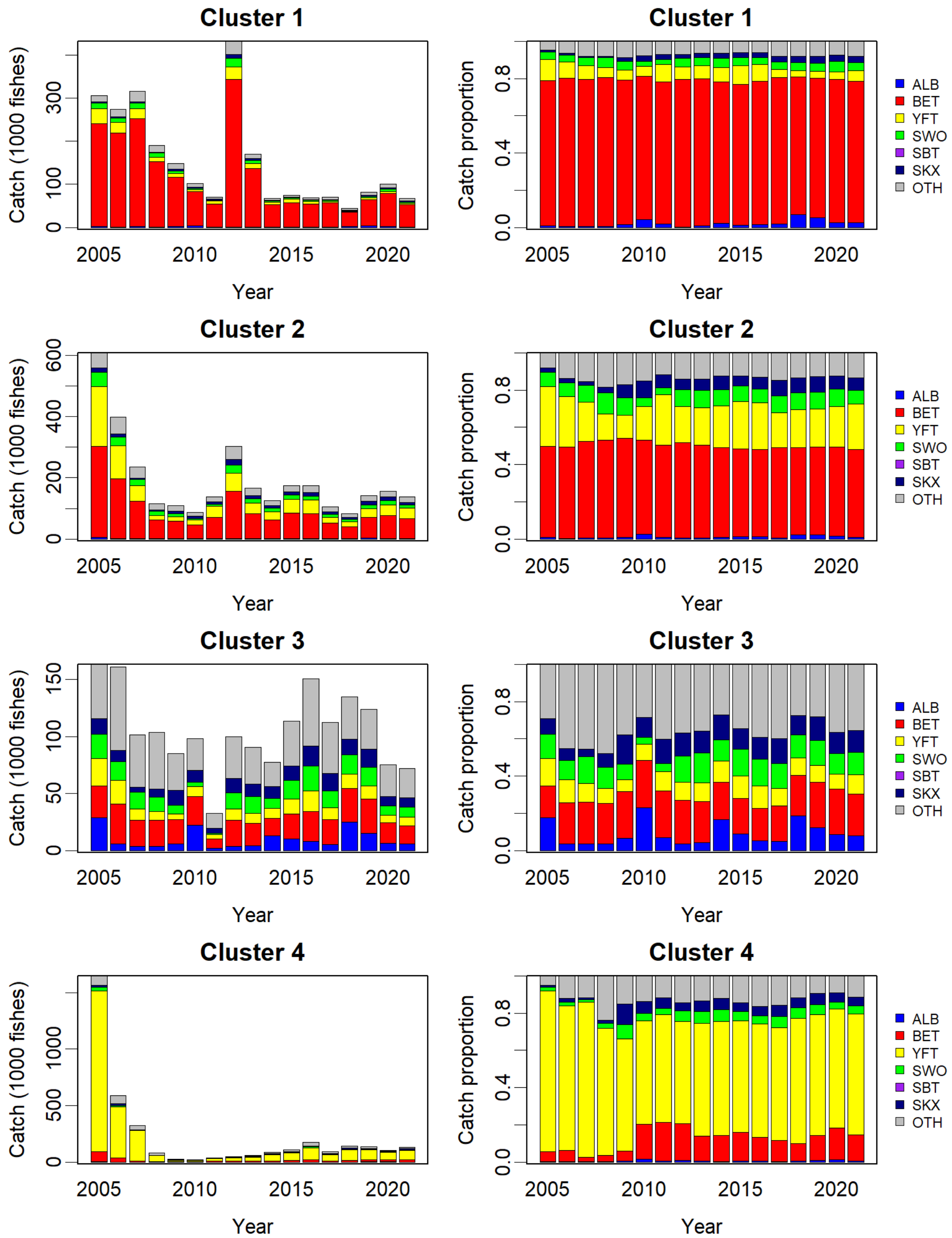


Fig. 8. Annual catches and compositions by species for each cluster of Taiwanese large-scale longline fishery in billfish area of the Indian Ocean.

NE

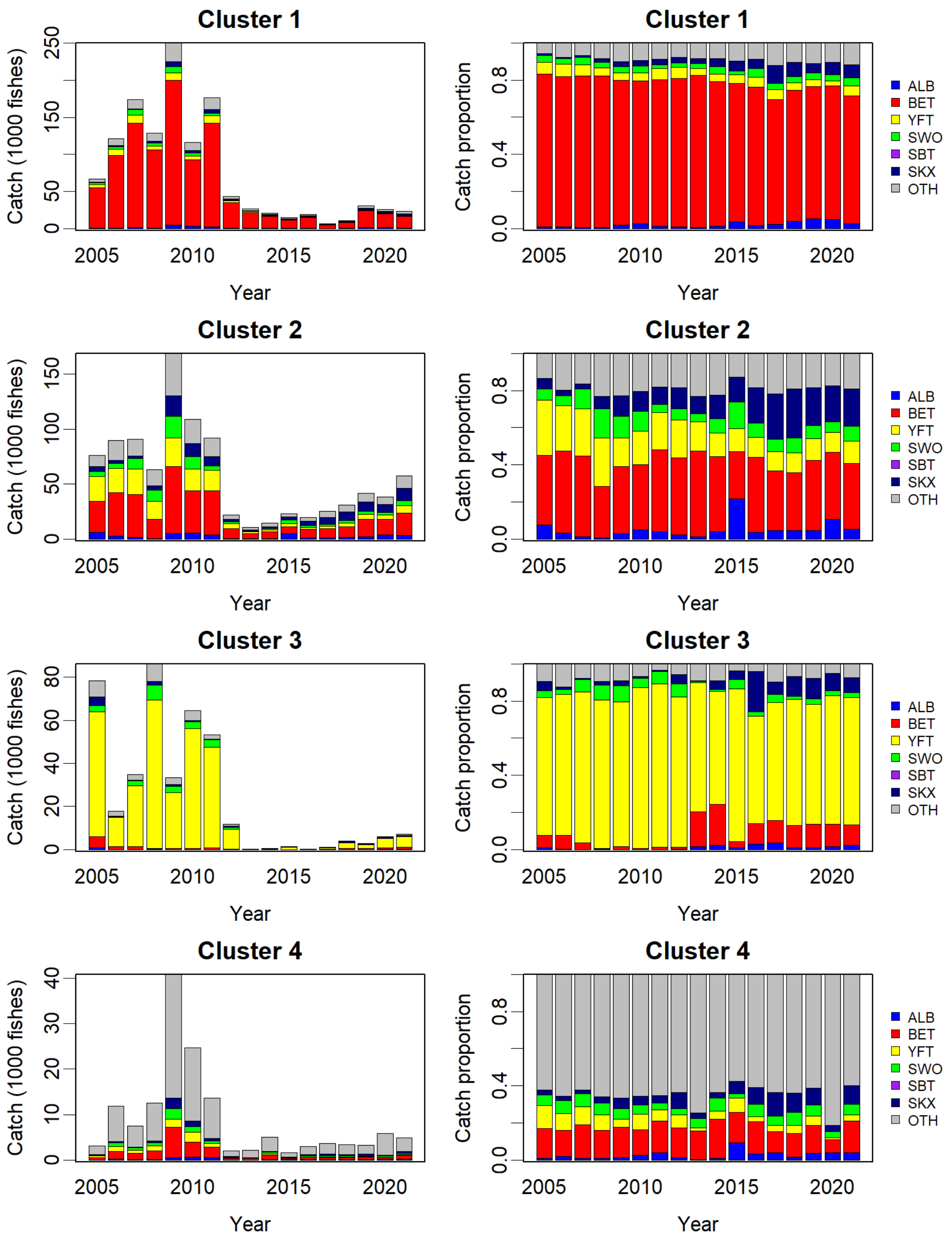


Fig. 8. (Continued).

SW

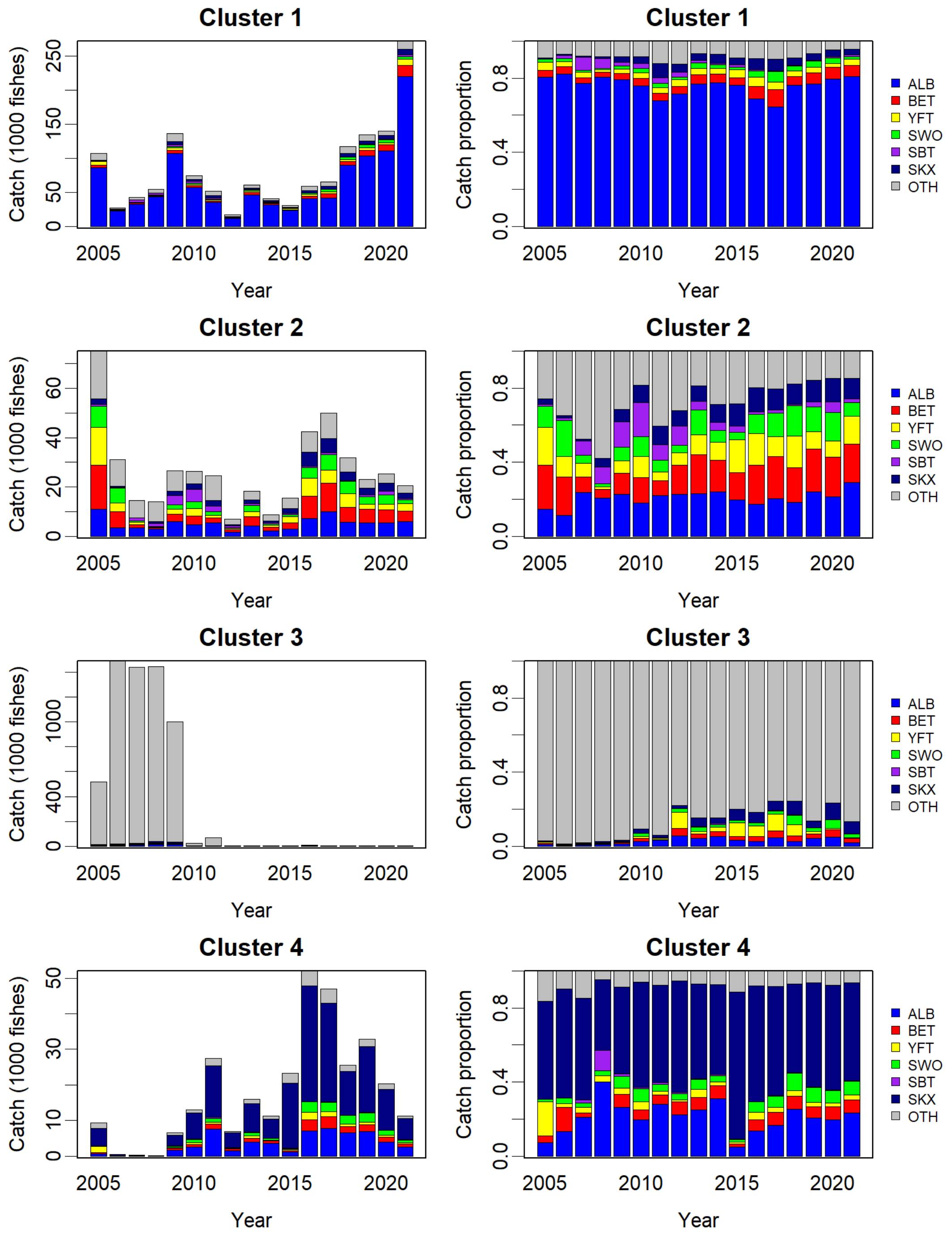


Fig. 8. (Continued).

SE

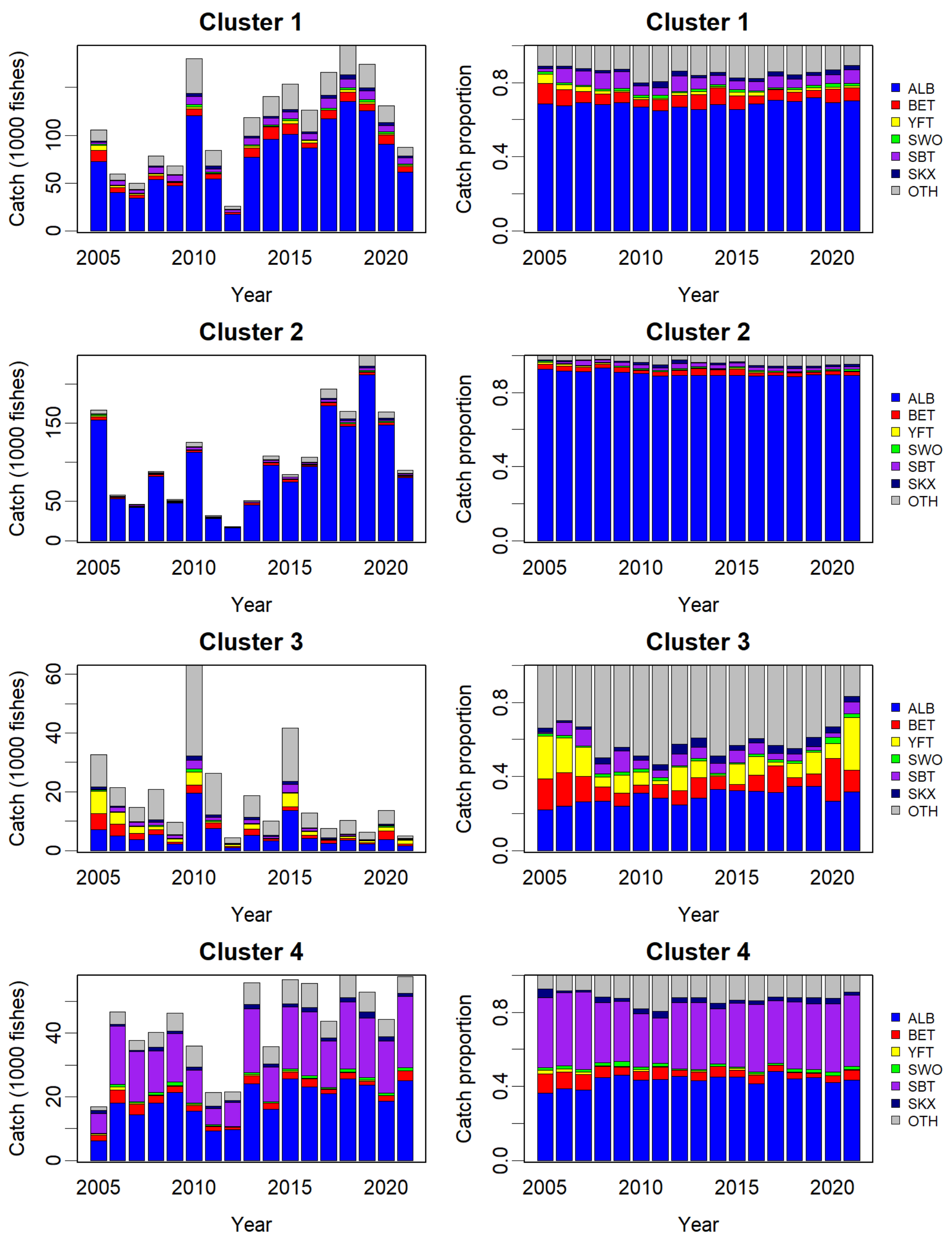


Fig. 8. (Continued).

NW

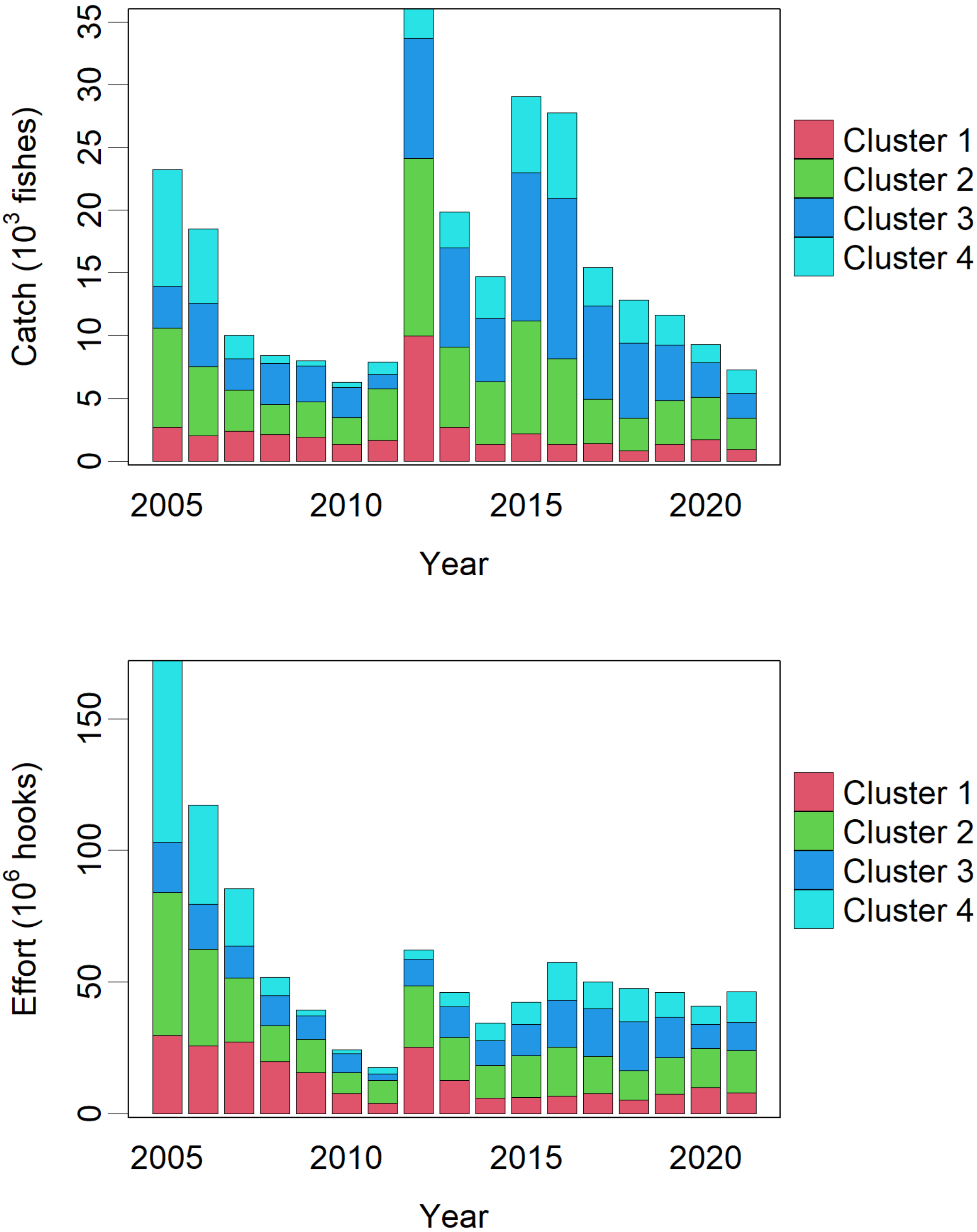
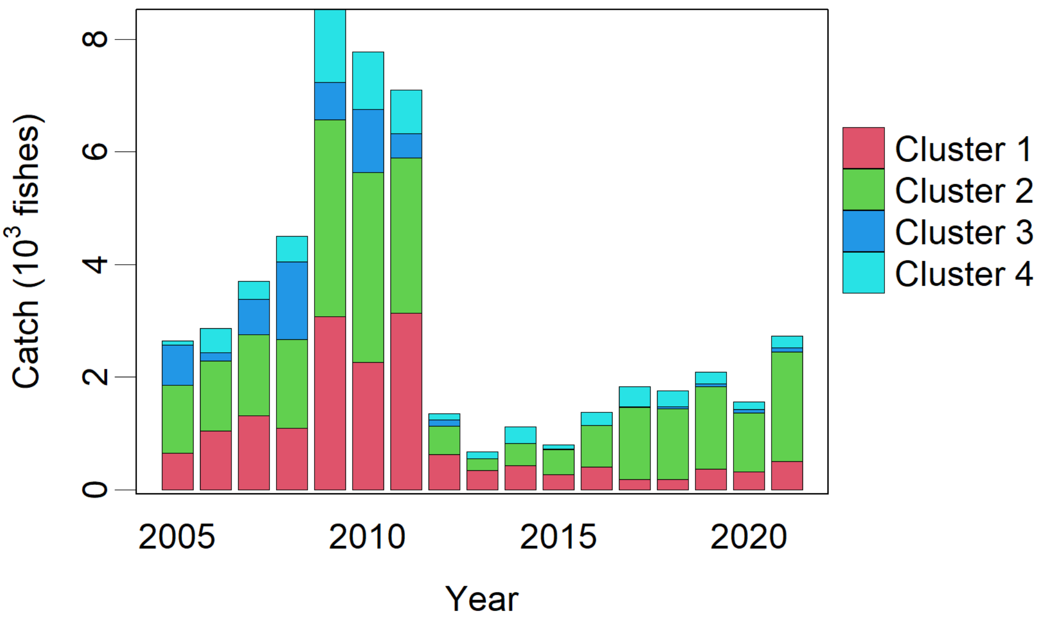


Fig. 9. Annual blue marlin catches and efforts for each cluster of Taiwanese large-scale longline fishery in billfish area of the Indian Ocean.

NE



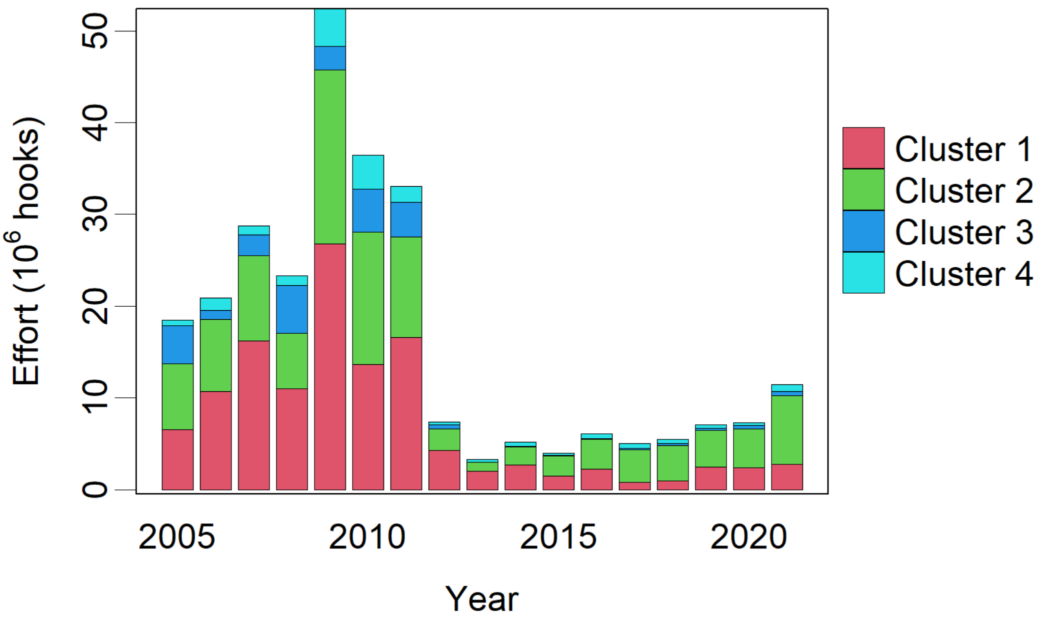
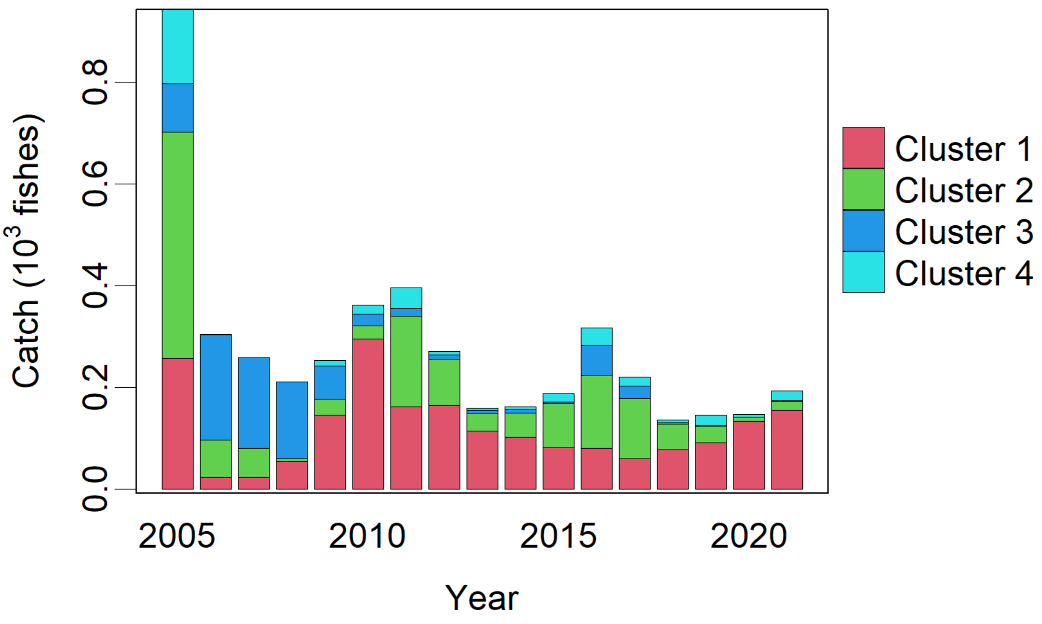


Fig. 9. (Continued).

SW



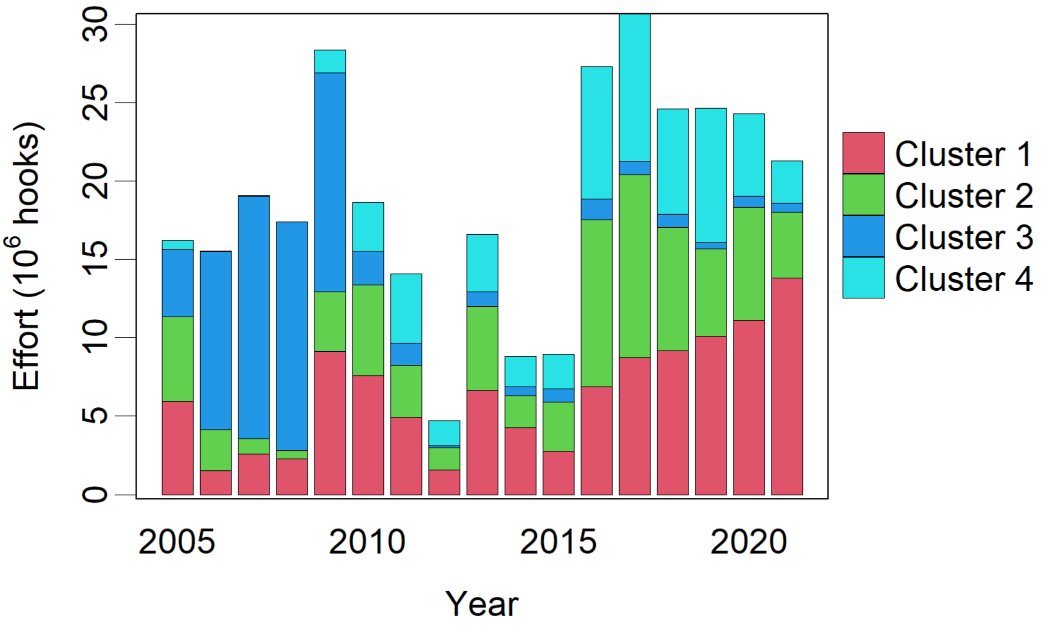


Fig. 9. (Continued).

SE

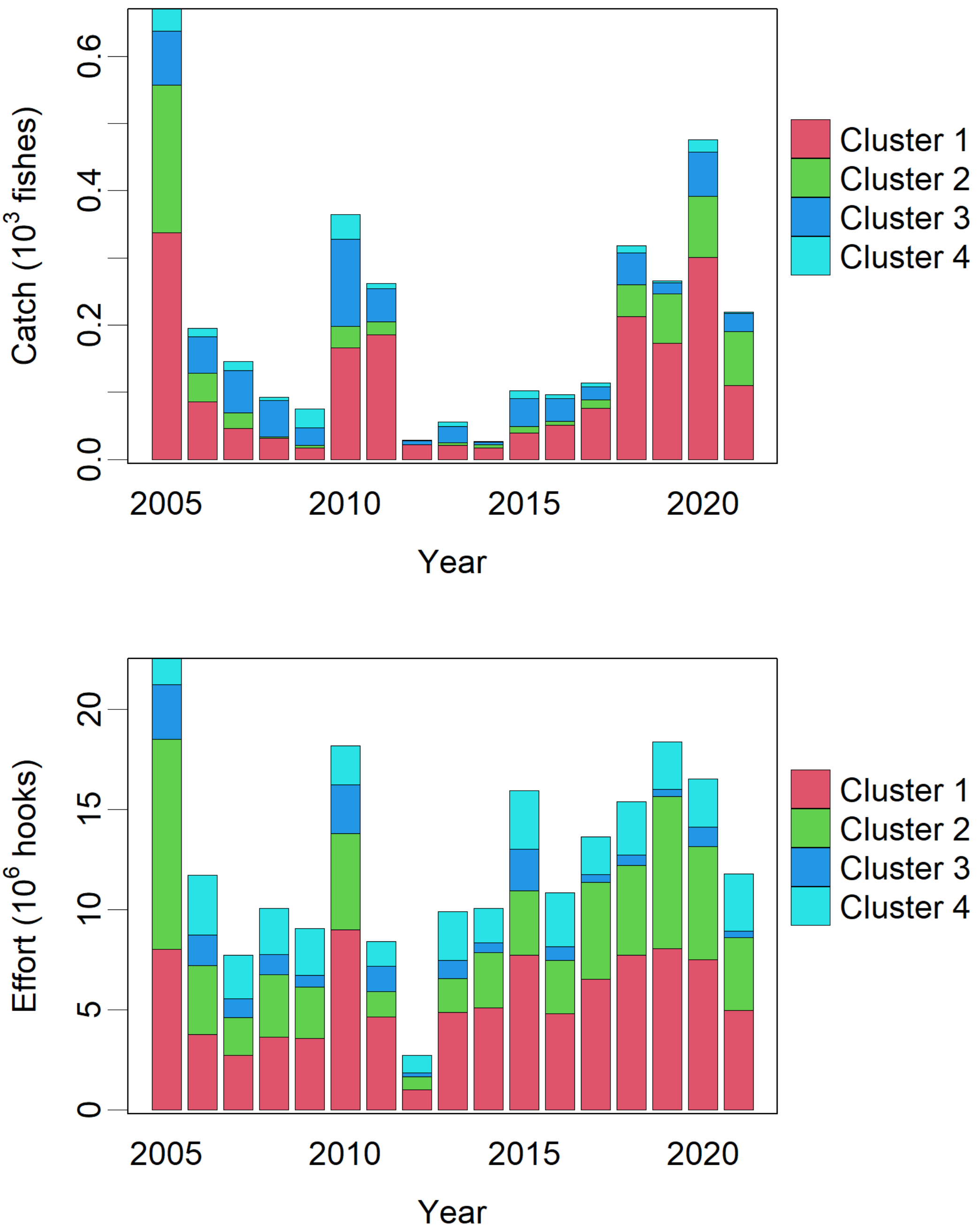


Fig. 9. (Continued).

NW

NE

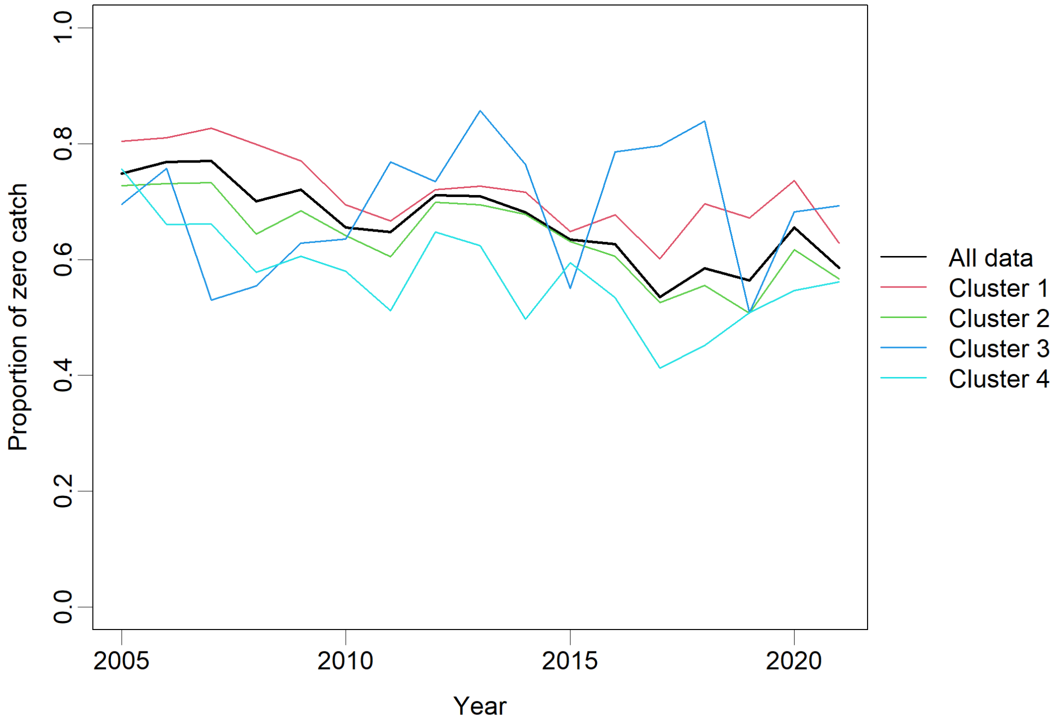
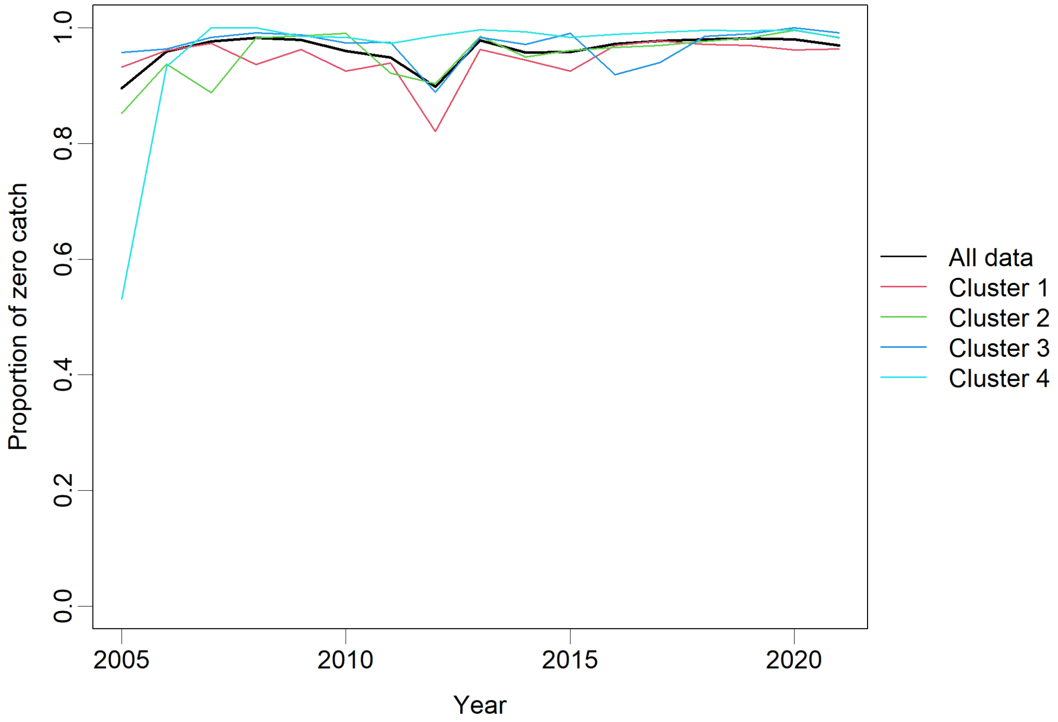


Fig. 10. Annual zero proportion of blue marlin catches for each cluster of Taiwanese large-scale longline fishery in billfish area of the Indian Ocean.

SW



SE

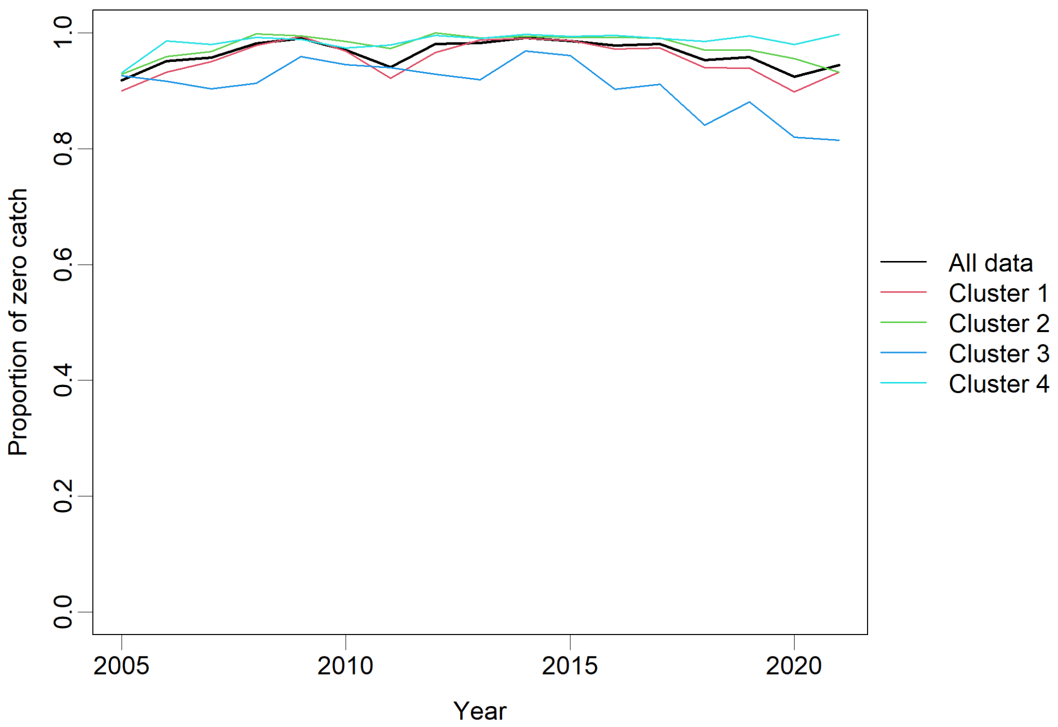


Fig. 10. (Continued).

NW

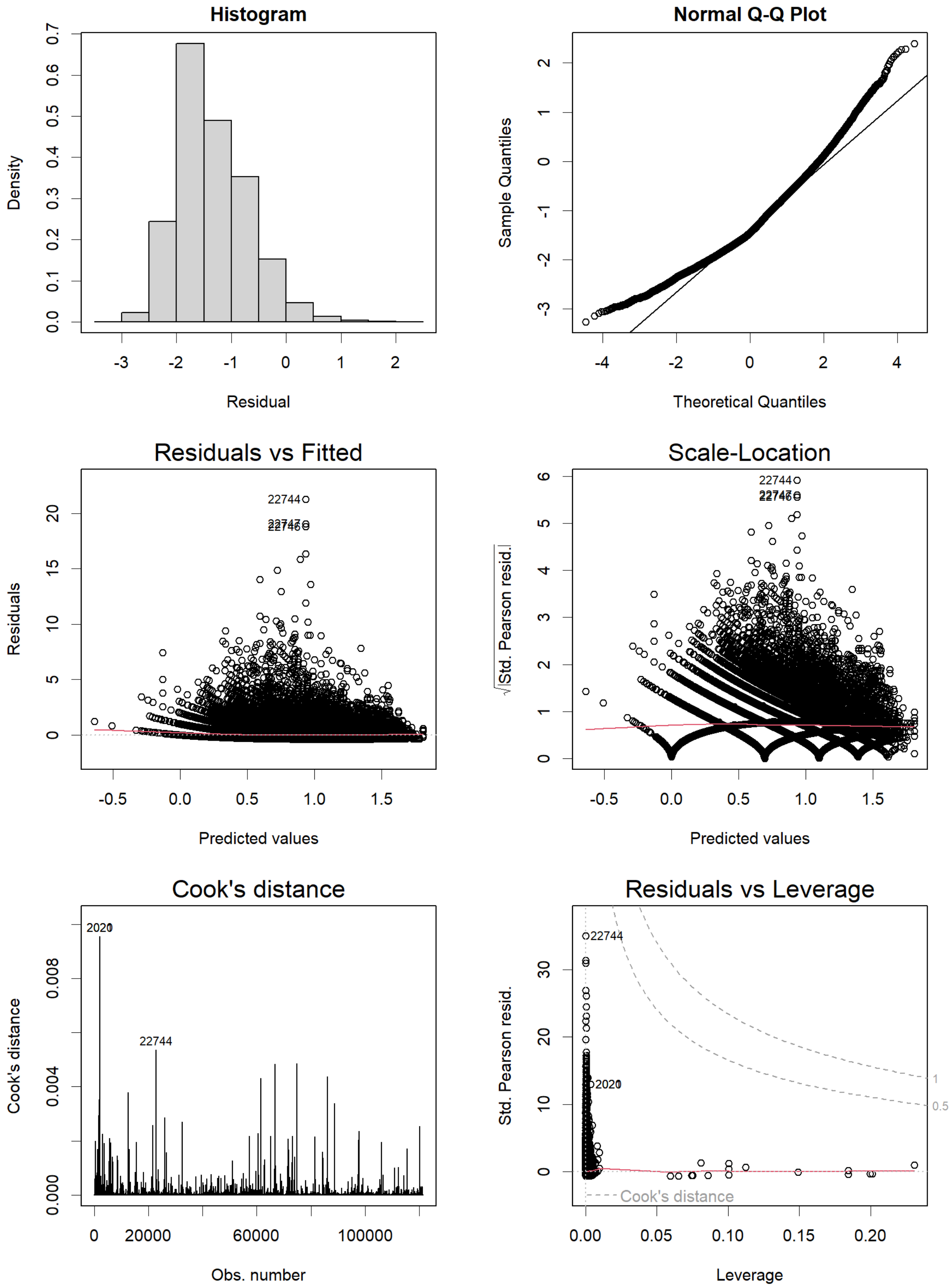


Fig. 11. Diagnostic plots for GLMs with inverse gaussian error distribution assumption for blue marlin caught by Taiwanese large-scale longline fishery in the Indian Ocean from 2005 to 2021.

NE

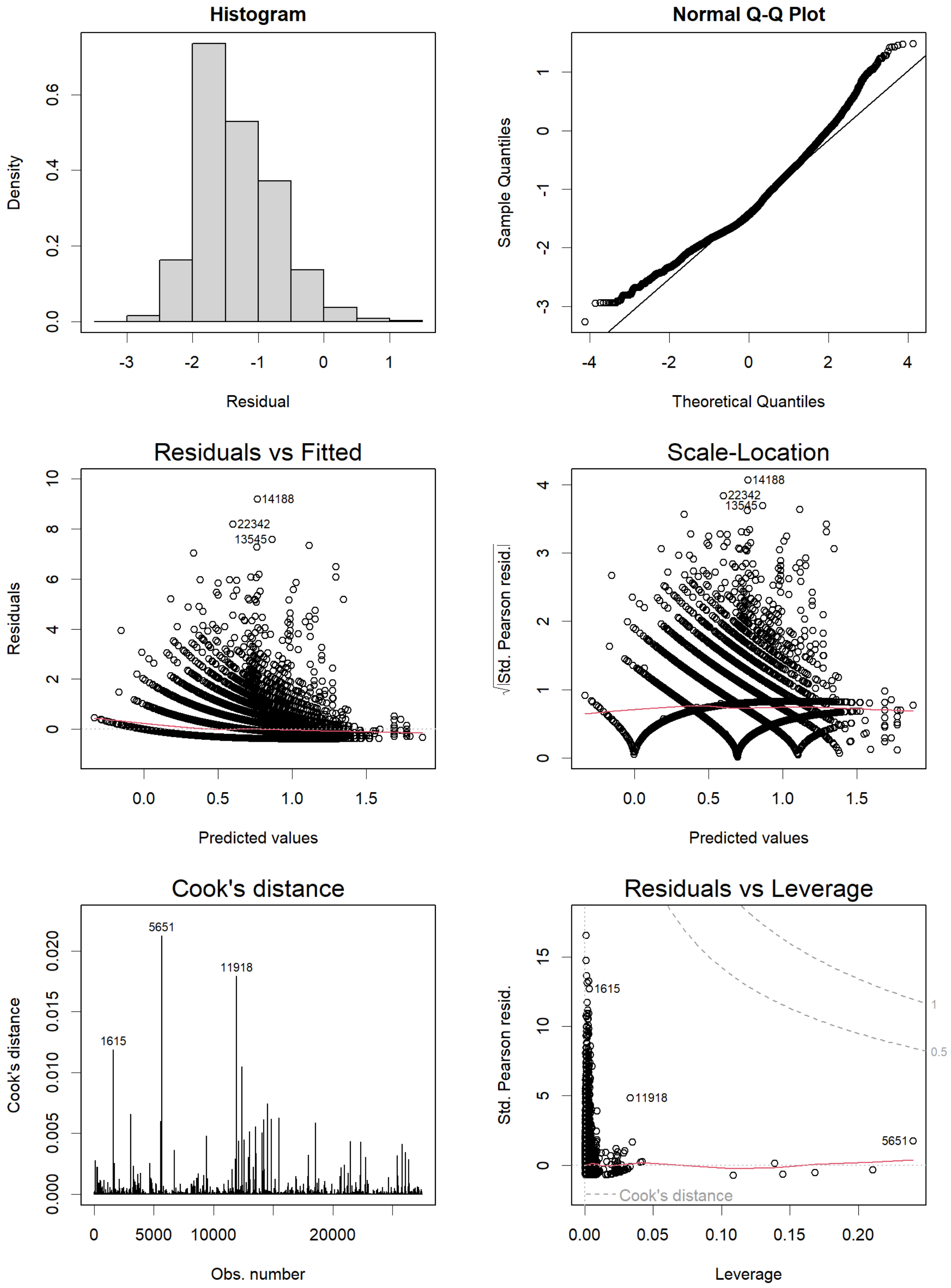


Fig. 11. (Continued).

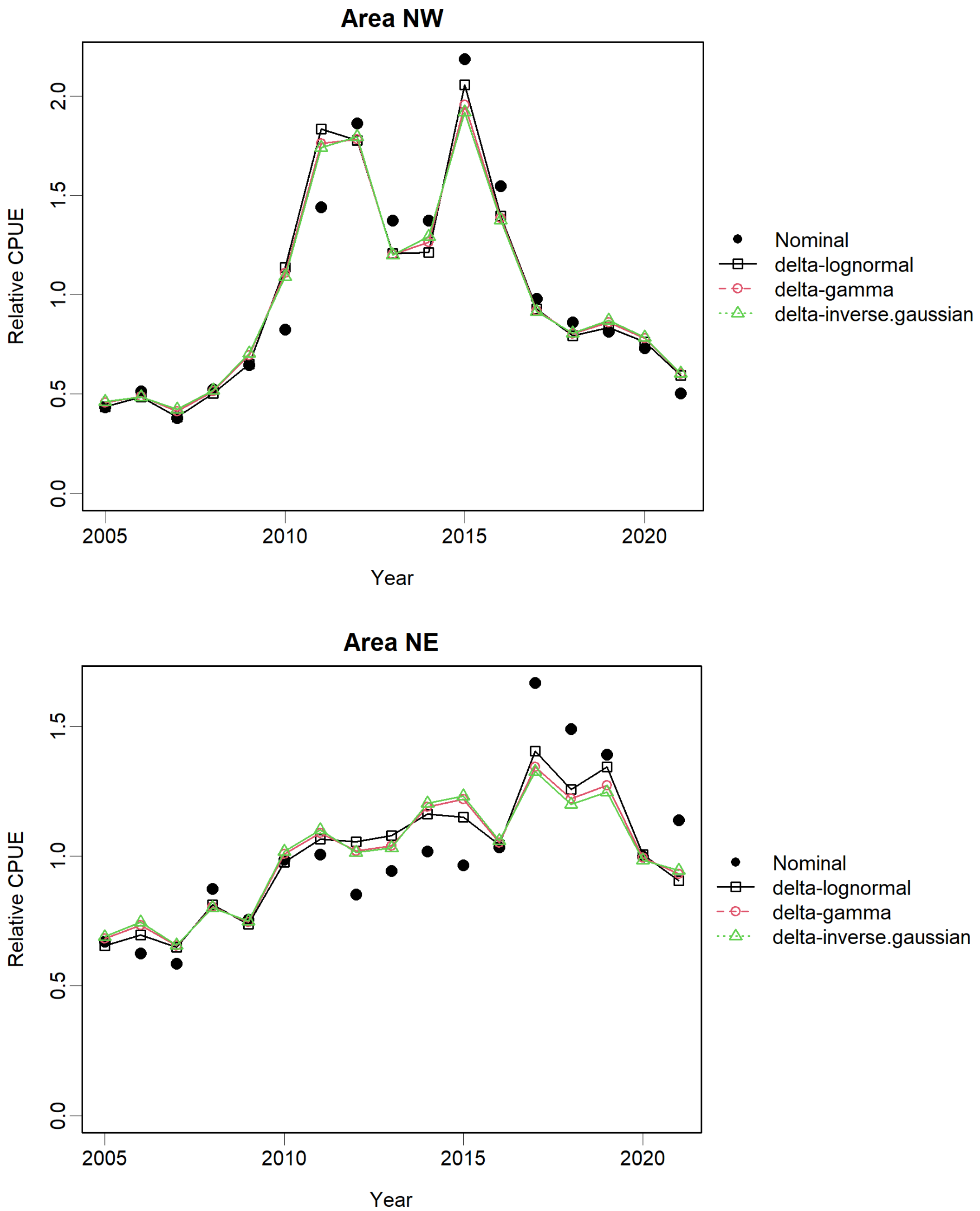


Fig. 12. Standardized CPUE series based on various GLMs for blue marlin caught by Taiwanese large-scale longline fishery in the Indian Ocean from 2005 to 2021.

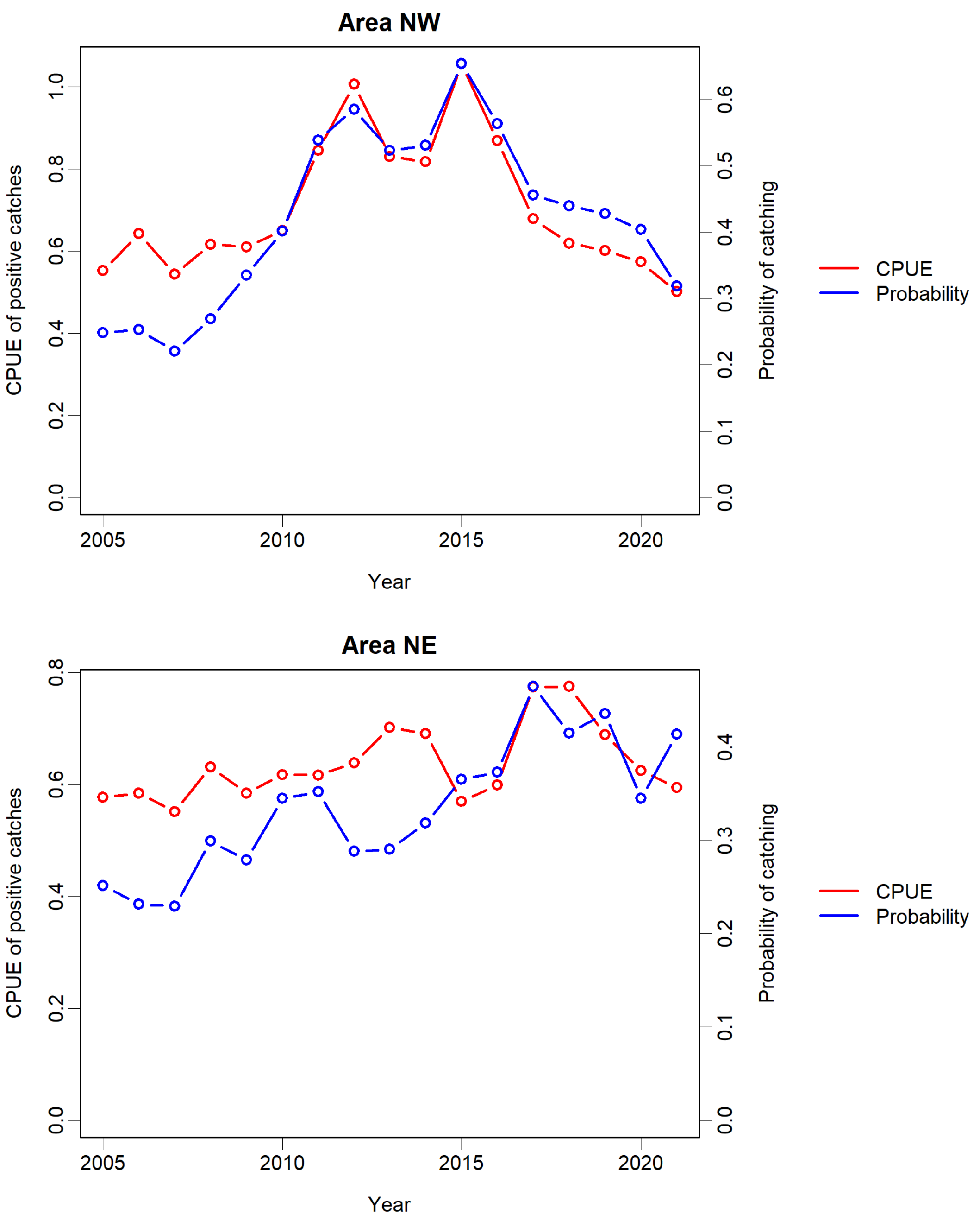


Fig. 13. Standardized CPUE of positive catches and catch probability based on selected model for blue marlin caught by Taiwanese large-scale longline fishery in the Indian Ocean from 2005 to 2021.

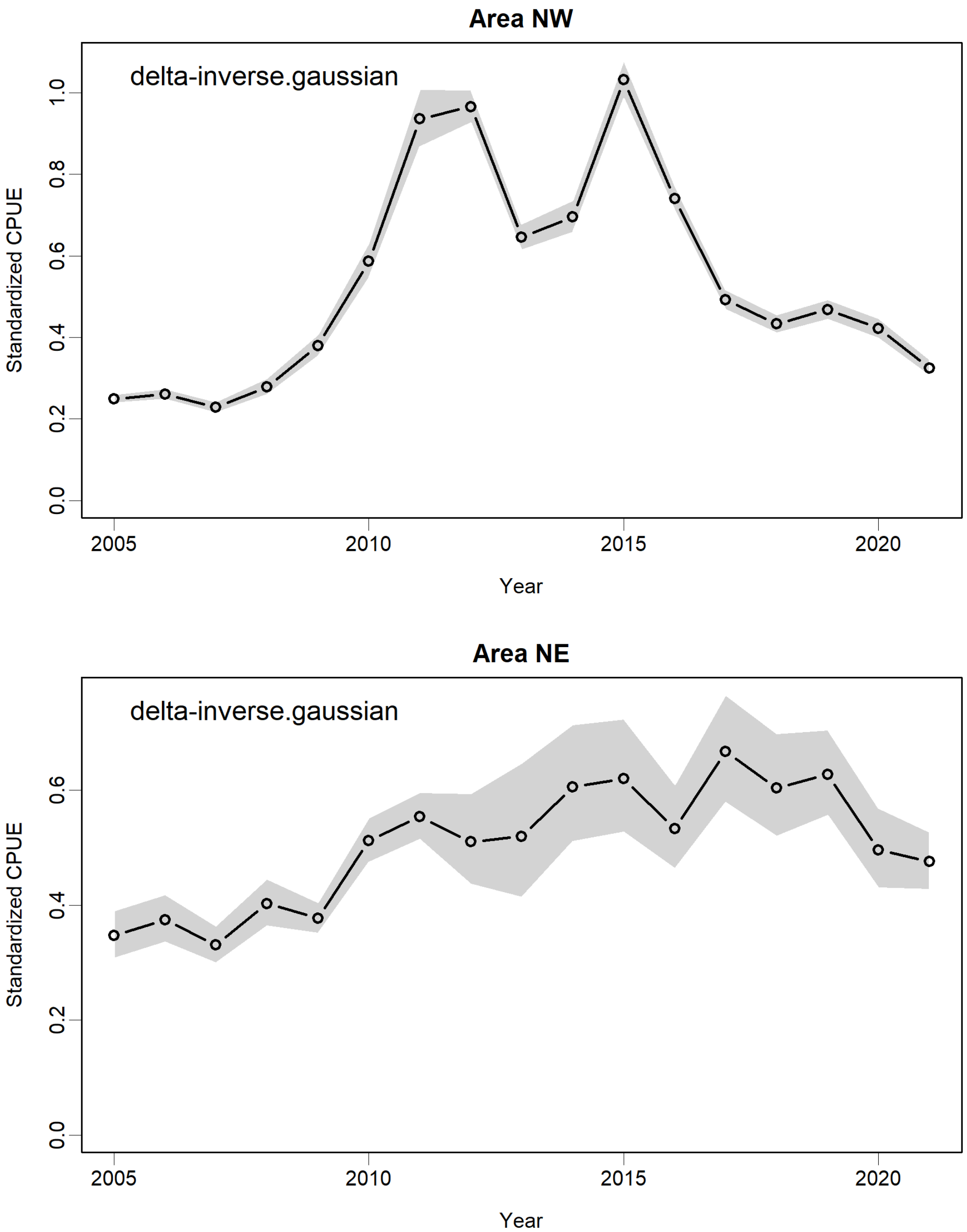


Fig. 14. Standardized CPUE series with 95% confidence intervals based on selected model for blue marlin caught by Taiwanese large-scale longline fishery in the Indian Ocean from 2005 to 2021.

Table 1. Diagnostic statistics for standardized CPUE series based on various models for positive catches of blue marlin caught by Taiwanese large-scale longline fishery in the Indian Ocean from 2005 to 2021.

Area	Model	R^2	AIC	BIC
NW	lognormal	0.105	539,935	540,537
NW	gamma	0.183	382,089	382,691
NW	inverse.gaussian	0.175	343,859	344,461
NE	lognormal	0.071	107,255	107,698
NE	gamma	0.112	76,436	76,814
NE	inverse.gaussian	0.112	68,512	68,890

Table 2. ANOVA table for selected standardized CPUE series based on selected GLMs for blue marlin caught by Taiwanese large-scale longline fishery in the Indian Ocean from 2005 to 2021.

Area NW

Positive catch model:

	Sum Sq	Df	F values	Pr(>F)
Y	2,629.0	16	445.6	< 2.2e-16 ***
Q	6.0	3	5.2	0.0014 **
CT	4.0	2	5.7	0.0034 **
Lon	79.0	7	30.7	< 2.2e-16 ***
Lat	502.0	8	170.2	< 2.2e-16 ***
T	56.0	3	50.4	< 2.2e-16 ***
Q:CT	12.0	6	5.4	0.0000 ***
Q:T	28.0	9	8.4	0.0000 ***
CT:T	21.0	6	9.6	0.0000 ***
Residuals	44,766.0	121385		

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

Delta model

	LR Chisq	Df	Pr(>Chisq)
Y	21641.9	16	< 2.2e-16 ***
Q	117.9	3	< 2.2e-16 ***
CT	186.1	2	< 2.2e-16 ***
Lon	185.3	8	< 2.2e-16 ***
Lat	4927.5	8	< 2.2e-16 ***
T	174.0	3	< 2.2e-16 ***
hook	82.0	1	< 2.2e-16 ***
Q:CT	209.2	6	< 2.2e-16 ***
Q:T	227.4	9	< 2.2e-16 ***
CT:T	30.8	6	2.83E-05 ***

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

Table 2. (Continued).

Area NE

Positive catch model:

	Sum Sq	Df	F values	Pr(>F)
Y	58.5	16	11.8	< 2.2e-16 ***
Q	7.9	3	8.5	0.0000 ***
Lon	6.1	6	3.3	0.0033 **
Lat	210.1	7	97.1	< 2.2e-16 ***
T	46.8	3	50.5	< 2.2e-16 ***
Q:T	20.5	9	7.4	0.0000 ***
Residuals	8,484.8	27457		

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

Delta model

	LR Chisq	Df	Pr(>Chisq)
Y	1558.6	16	< 2.2e-16 ***
Q	1.0	3	0.7922079
CT	22.2	2	1.55E-05 ***
Lon	155.7	7	< 2.2e-16 ***
Lat	3509.5	7	< 2.2e-16 ***
T	11.0	3	0.0118136 *
hook	12.8	1	0.0003535 ***
Q:CT	38.1	6	1.07E-06 ***
Q:T	36.4	9	3.42E-05 ***
CT:T	44.8	6	5.10E-08 ***

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