CPUE standardization of black marlin (*Makaira indica*) 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, Keelung, Taiwan.
* Corresponding author: wsp@mail.ntou.edu.tw

Paper submitted to the 19th meeting of the Working Party on Billfish, Indian Ocean Tuna Commission, 13-16 September, 2021.

ABSTRACT

In this study, the delta-linear models with different assumptions of error distribution were adopted to conduct the CPUE standardization of black marlin caught by the Taiwanese large-scale longline fishery in the Indian Ocean for 1979-2020 and 2005-2020. The groups of data sets derived from cluster analysis based on species compositions were incorporated in the models as a covariate for explaining the target. The results indicate that the targeting effects (clusters) provided most significant contributions to the explanation of the variance of CPUE for the models with positive catches, while the catch probability might be mainly influenced by the position of fishing operations. The standardized CPUE series obtained from different model assumptions revealed quite similar trends for all model except for delta-lognormal model. For 1979-2020, CPUE trends were similar for the northern areas (NW and NE) and they fluctuated before early 1990s, gradually declined until late 2000s, increased until mid-2010s, then substantially decreased again, and reveals an increasing trend in recent years. For 2005-2020, the trends of CPUE for the northern areas (NW and NE) also revealed similar patterns, CPUE increased from 2013 to 2016, decreased until 2018, and increased in recent years.

1. INTRODUCTION

Black marlin is considered to be a bycatch species of industrial and artisanal fisheries. Gillnet fisheries has increased year by year, account for more than 50% of total catches in the Indian Ocean, followed by troll and handlines (32%), with remaining catches recorded under longlines (12%). Catches have increased steadily since the 1990s, from 2,500 t in 1991 to over 13,000 t since 2004. In recent years catches have further increased sharply from 13,000 t in 2012 to over 22,000 t in 2016 – the highest catches recorded in the Indian Ocean – due to increases reported by the offshore gillnet fisheries of Iran. Sri Lanka has developed gillnet and longline fisheries since mid-1990s, and catches have continued to increase from 1,000 t to 3,000 t. The catches were mainly made by Iran (gillnet, 30%), India (gillnet and troll, 23%) and Sri Lanka (gillnet and fresh longline, 21%) from 2014 to 2018. Taiwan has made only about 2% of total catches of black marlin in the Indian Ocean (IOTC, 2020).

IOTC conducted a stock assessment for black marlin in the Indian Ocean in 2018 but the results had high uncertainty due to catches increased sharply while the opposite trend of CPUE (IOTC, 2018). Therefore, this study conducted CPUE standardization for black marlin to obtain 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) by 5x5 degree longitude and latitude grid for Taiwanese longline fishery during 1979-2020 were used. These data were provided by Oversea Fisheries Development Council of Taiwan (OFDC). It should be noted that the data in 2020 remains preliminary. For conducting the cluster analysis prior to the CPUE standardizations, the data were aggregated by 10-days duration (1st-10th, 11th-20th, and 21st~ in each month) (Kitakado et al., 2021).

2.2. CPUE Standardization

Because black marlin was bycatch species of Taiwanese longline fishery, a large amount of zero-catches was recorded in the operational catch and effort data sets. In previous study, 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-generalized linear models with different assumptions of error distribution were applied to conduct the CPUE standardization of black marlin in the Indian Ocean (Pennington, 1983; Lo et. al., 1992; Pennington, 1996; Andrade, 2008; Lauretta et al., 2016; Langley, 2019).

As the approach of Wang (2017), the models were simply conducted with the main effects of year, quarter, longitude, latitude, and fishing targeting (clusters derived from species compositions of data sets, Wang et al., 2021), while interactions between main effects were not incorporated into the models. The models for positive catches and presence/absence data were conducted as follows:

For CPUE of positive catches:

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

Delta model for presence and absence of catch:

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

where	*Catch*	is the nominal catch in number of positive catch of black marlin (catch in number/1,000 hooks),
	PA	is the nominal presence and 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,
	Lon	is the effect of longitude,
	Lat	is the effect of latitude,
	T	is the effect of targeting (cluster),
	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, poisson, gamma, negative-binomial and tweedie distributions for the error distribution of the model for the positive catches and specified “log” for the model link function. For the model with tweedie distribution, the index of power variance function was tested using values of 1.1-1.9. In addition, the models with negative-binomial and tweedie distributions were also performed by including all of positive and zero catches (catches were added 1 to avoid the problem for the logarithm of zero) to examine the model performance for the data overdispersed with excess of zero catches.

The models performed by stepwise search ("both" direction, i.e. "backward" and "forward") and selected based on the values of Akaike information criterion (AIC), the coefficient of determination (R^2) and Bayesian information criterion (BIC). The AIC and BIC, which were calculated based on the likelihoods with full constants obtained glm() and glm.nb(), were used to compare the models with different error distributions (e.g. Setyadji et al., 2019). In addition, dispersion statistics for Pearson residuals were calculated to check whether under- or overdispersions resulted from the models with an assumed error distribution.

The standardized CPUE were calculated based on the estimates of least square means of the interaction between the effects of year and area. The area-specific standardized CPUE trends were estimated based on the exponentiations of the adjust means (least square means) of the year effects (Butterworth, 1996; Maunder and Punt, 2004). The standardized relative abundance index was calculated by the product of the standardized CPUE of positive catches and the standardized probability of positive catches:

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

where	DL^{index}	is standardized CPUE
	$CPUE$	is the adjust means (least square means) of the year effect of the model for positive catches,
	PA	is the adjust means (least square means) of the year effect of the model for presence/absence of catches.

2.3. Time series of data for analysis

As the suggestion of WPTT (IOTC, 2021), Taiwanese data before 2005 were recommended not using to conduct cluster analysis for target species and CPUE standardization for tropical tunas due to the problem of data quality. Furthermore, the data problem might influence not only the catches of major tuna species but also the catch and effort data for other species. Therefore, CPUE standardizations were conducted using the data from 1979 to 2020 and from 2005 to 2020 for the use of further stock assessments.

3. RESULTS AND DISCUSSION

3.1. Historical fishing trends

Figs. 2 to 3 show the Taiwanese historical nominal catches and CPUE distribution of black marlin based on the logbook data of Taiwanese large-scale longline fishery. Black marlin were mainly caught in tropical area and coastal waters of the northern Indian Ocean. Although the catches increased as the fishing efforts increased in the southern Indian Ocean during 1990s and 2000s, high CPUE only occurred in the coastal waters of the northern Indian Ocean over the years.

Black marlin catches were mainly caught in the Area NE before the 1990s and most of the catches were made in the Area NW thereafter. Although substantial increase of the catches in Area NW from the early 2000s, substantial decrease thereafter (Fig. 4).

According to the analysis of length-frequency data by year, month, longitude and latitude (Fig. 5), length distribution showed obvious variation in both central tendency and dispersion before the early 2000. In recent year, central tendency has been significantly higher than in the past few years. There is no obvious trend of change in month. The degree of dispersion increased significantly in longitude of 55°E-90°E, i.e. more large and small fish were caught. Although the central tendency was more concentrated in latitude of 5°S-15°N, more large and small fish were caught.

3.2. CPUE standardization

CPUE standardizations were separately conducted for only northern areas (NW and NE, Fig. 1) since sparse catches of black marlin were made in the southern areas (Fig. 2).

3.2.1 Time series from 1979 to 2020

Based on the AIC model selection for the models for positive catches and delta model shows that some effects did not provide significant improvement to AIC and were excluded in different area. For the models for positive catches, the models with gamma error distribution would be the optimal models for all areas based on the values of AIC, BIC and Pearson dispersion statistics although R^2 may not be higher than other models (Table 1). Diagnostic plots for residuals also indicated that the models with gamma error distribution (Fig. 6) should be most appropriate than other models because there were less increasing or decreasing trends in the range of predicted values when assuming a gamma error distribution (plots for other models by areas were not shown here but the residuals revealed obvious patterns with predicted values). Therefore, the results obtained the delta-gamma 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 Table 2. The results indicate that the effects of *T* (clusters) provided most significant contributions to explanation of variance of CPUE for the models for positive catches for all areas, while in addition to the effects of *Y*, the effects of *Lat* and *Lon* were the most significant variable for the delta model. Thus, the catch rates derived from the positive catches of black marlin might be influenced by the targeting of the fishing operation, while the position of fishing operation might influence the opportunity of catching black marlin.

The area-specific standardized CPUE series are shown in Fig. 7. The CPUE series revealed quite similar trend for all model except for delta-lognormal model. The standardized CPUE series for the selection of model are shown in Fig. 8. The trend of CPUE in the northern areas (NW and NE) reveal relatively similar patterns and they fluctuated before early 1990s, reveal decreasing trends until late 2000s, increased until mid-2010s, then substantially decreased again, and increased in recent years.

3.2.2 Time series from 2005 to 2020

Based on the AIC model selection for the models for positive catches and delta model shows that some effects did not provide significant improvement to AIC and were excluded in different area. For the models for positive catches, the models with gamma error distribution would be the optimal model for all areas based on the values of AIC and BIC statistics although R^2 may not be higher than other models (Table 3). Diagnostic plots for residuals also indicated that the models with gamma error distribution (Fig. 9) should be most appropriate than other models because there were less increasing or decreasing trends in the range of predicted values when assuming a gamma error distribution (plots for other models by areas were not shown here but the residuals revealed obvious patterns with predicted values). Therefore, the results obtained the delta-gamma model were selected to produce the standardized CPUE series for further stock assessment.

The ANOVA tables for selected models are shown in Table 4. The results indicate that the effects of *T* (clusters) provided most significant contributions to explanation of variance of CPUE for the models for positive catches for all areas, while in addition to the effects of *Y*, the effects of *Lat* and *Lon* were the most significant variable for the delta model. Thus, the catch rates derived from the positive catches of black marlin might be influenced by the targeting of the fishing operation, while the position of fishing operation might influence the opportunity of catching black marlin.

The area-specific standardized CPUE series are shown in Fig. 10. The CPUE series revealed quite similar trend for all model except for delta-lognormal model. The standardized CPUE series for the selection of model are shown in Fig. 11. The trends of CPUE in the northern areas (NW and NE) reveal relatively similar patterns and increased from 2013 to 2016, decreased until 2018, and increased 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.

Butterworth, D.S., 1996. A possible alternative approach for generalized linear model analysis of tuna CPUE data. ICCAT Col. Vol. Sci. Pap., 45: 123-124.

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, 2018. Report of the 16th Session of the IOTC Working Party on Billfish. IOTC-2018-WPB16-R[E].

IOTC, 2020. Report of the 18th Session of the IOTC Working Party on Billfish. IOTC-2020-WPB18-R[E].

IOTC, 2021. Report of the 23rd Session of the IOTC Working Party on Tropical Tunas, Data Preparatory Meeting. IOTC-2021-WPTT23(DP)-R[E].

Kitakado, T., Wang, S.P., Satoh, K., Lee, S.I., Tsai, W.P., Matsumoto, T., Yokoi, H., Okamoto, K., Lee, M.Y., Lim, J.H., Kwon, Y., Su, N.J., Chang, S.T., Chang, F.C., 2021. Report of trilateral collaborative study among Japan, Korea and Taiwan for producing joint abundance indices for the yellowfin tunas in the Indian Ocean using longline fisheries data up to 2019. IOTC–2021-WPTT23(DP)-14.

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.

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.

Setyadji, B., Andrade, H.A., Proctor, C.H., 2019. Standardization of catch per unit effort with high proportion of zero catches: an application to black marlin *Istiompax indica* (Cuvier, 1832) caught by the Indonesian tuna longline fleet in the eastern Indian Ocean. Turk. J. Fish.& Aquat. Sci. 19(2), 119-129.

Wang, S.P., 2017. CPUE standardization of swordfish (*Xiphias gladius*) caught by Taiwanese longline fishery in the Indian Ocean. IOTC–2017–WPB15–17_Rev1.

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 targeting of fishing operation for Taiwanese large scale longline fishery in the Indian Ocean. IOTC-2021-WPB19-11.

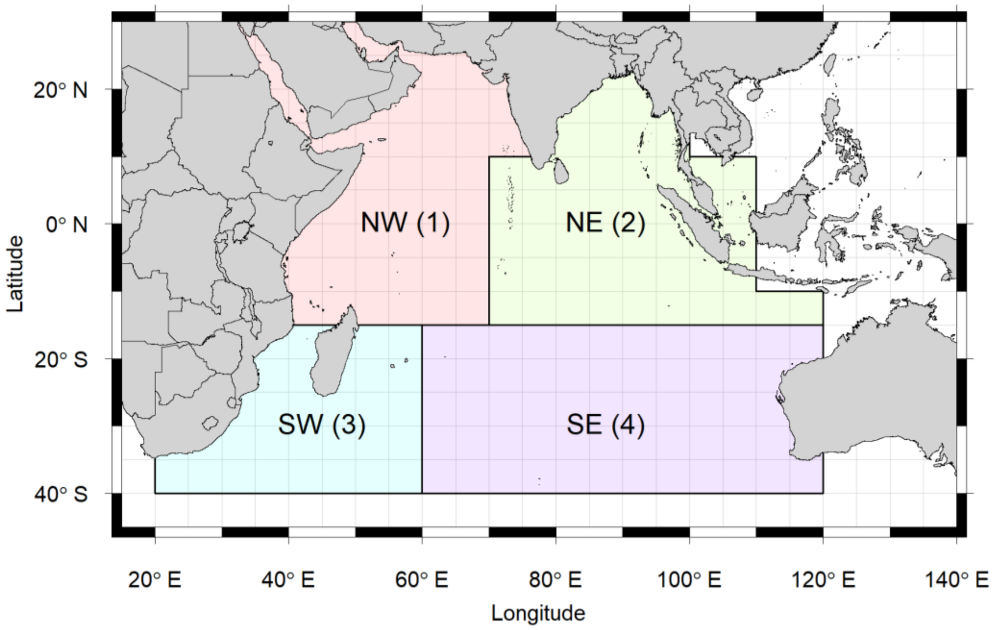


Fig. 1. Area stratification used for black marlin in the Indian Ocean.

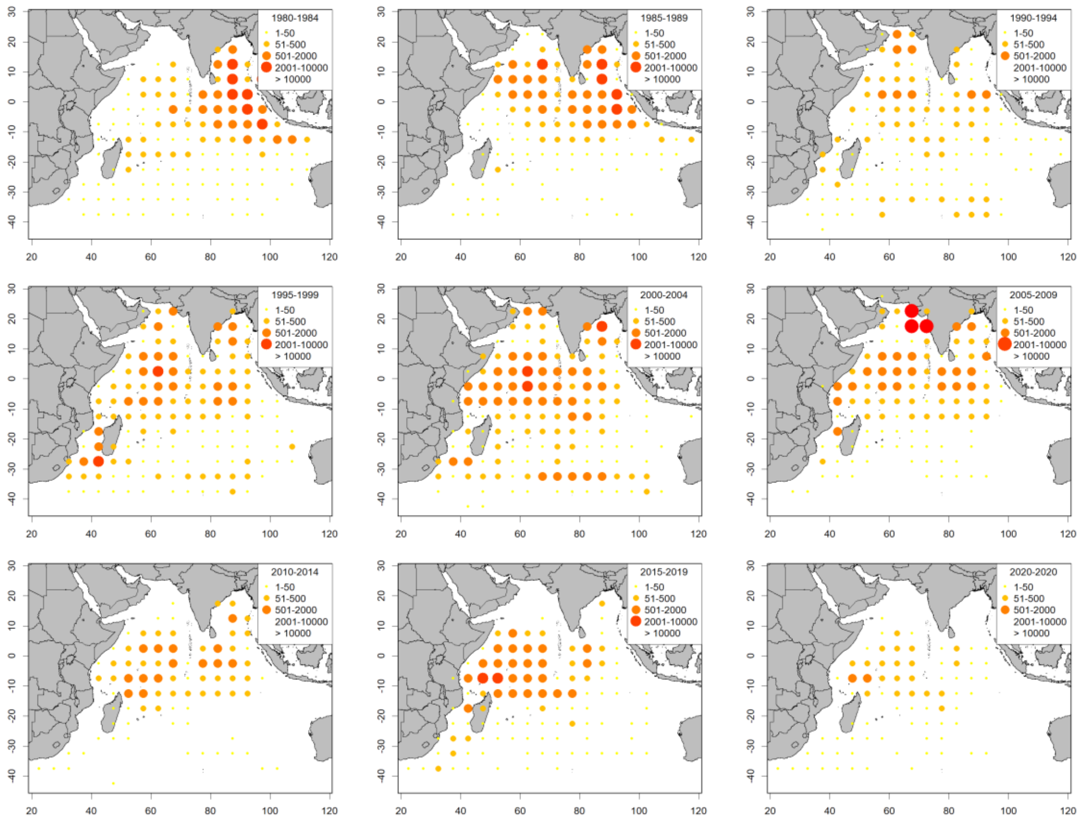


Fig. 2. Black marlin catch distribution of Taiwanese large-scale longline fishery in the Indian Ocean.

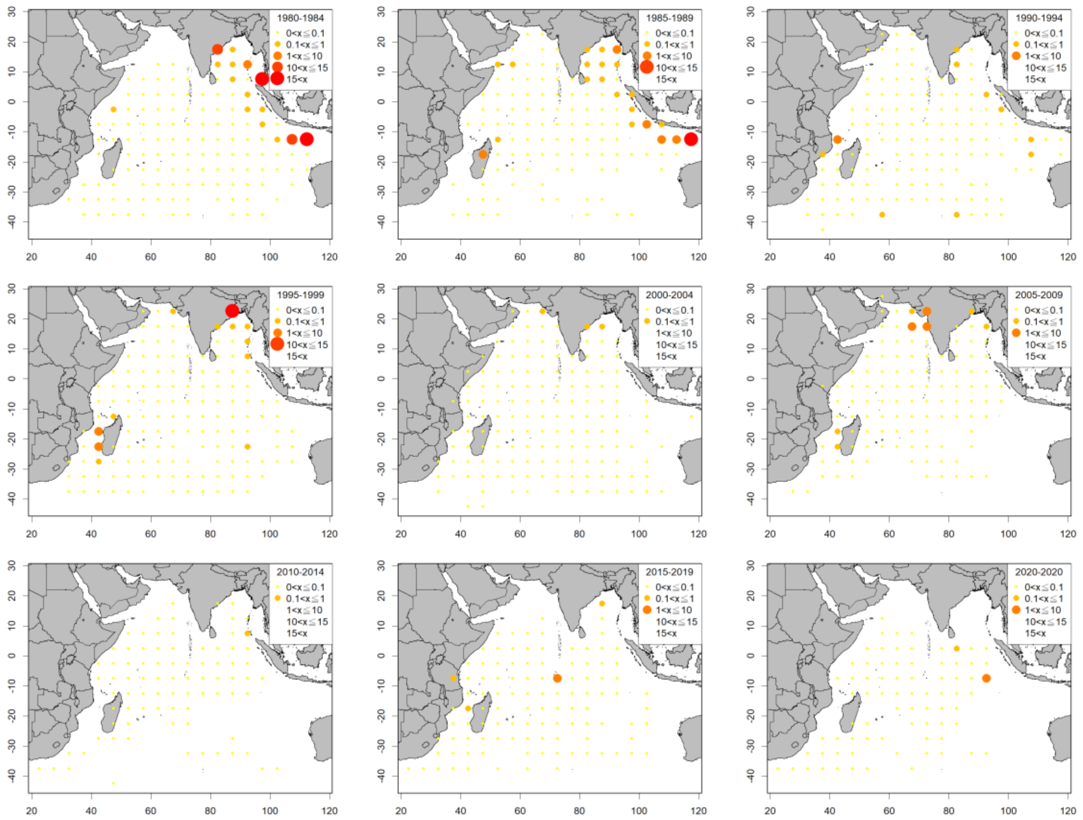


Fig. 3. Black marlin CPUE distribution of Taiwanese large-scale longline fishery in the Indian Ocean.

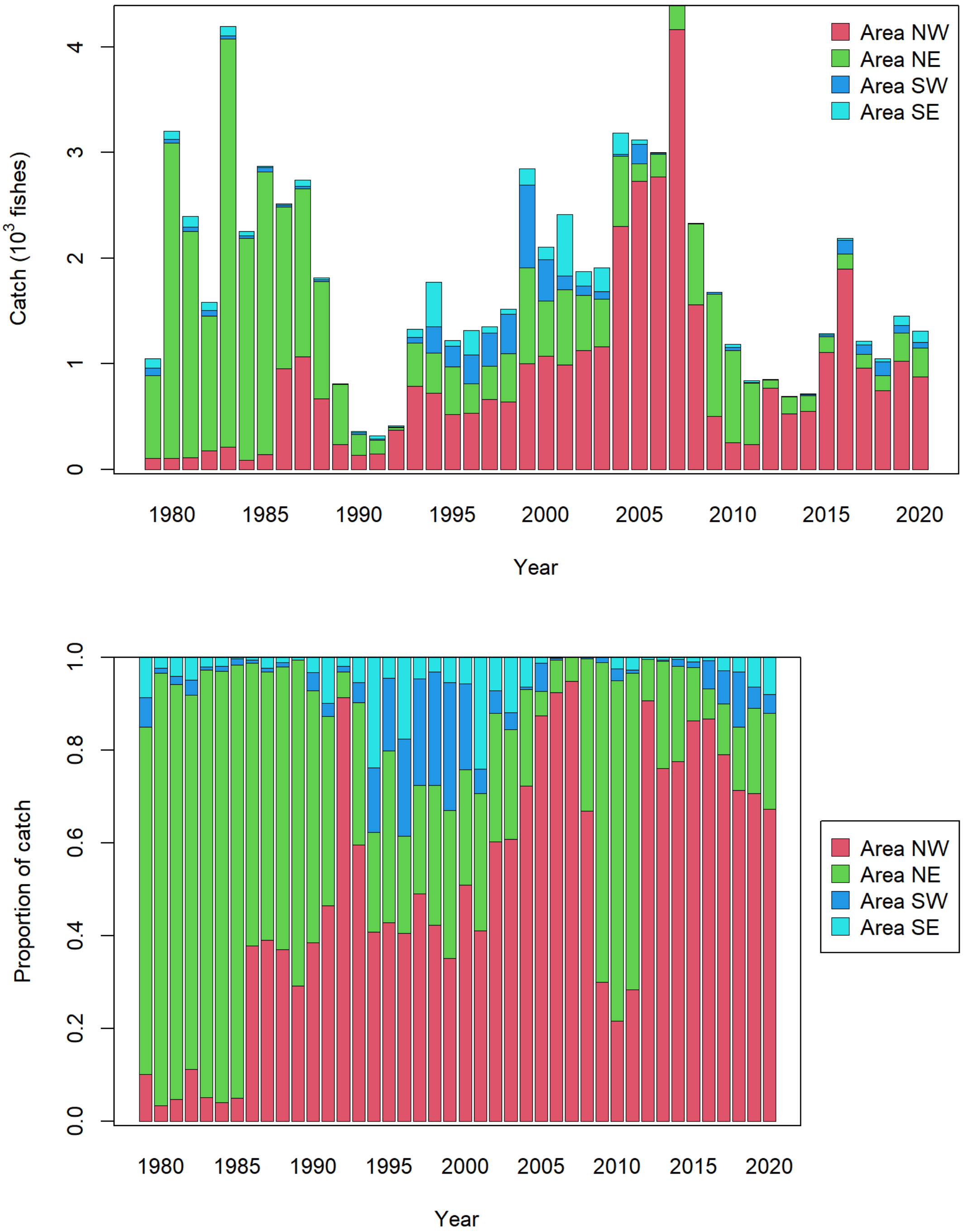


Fig. 4. Annual black marlin catches of Taiwanese large-scale longline fishery in the defined billfish area the Indian Ocean.

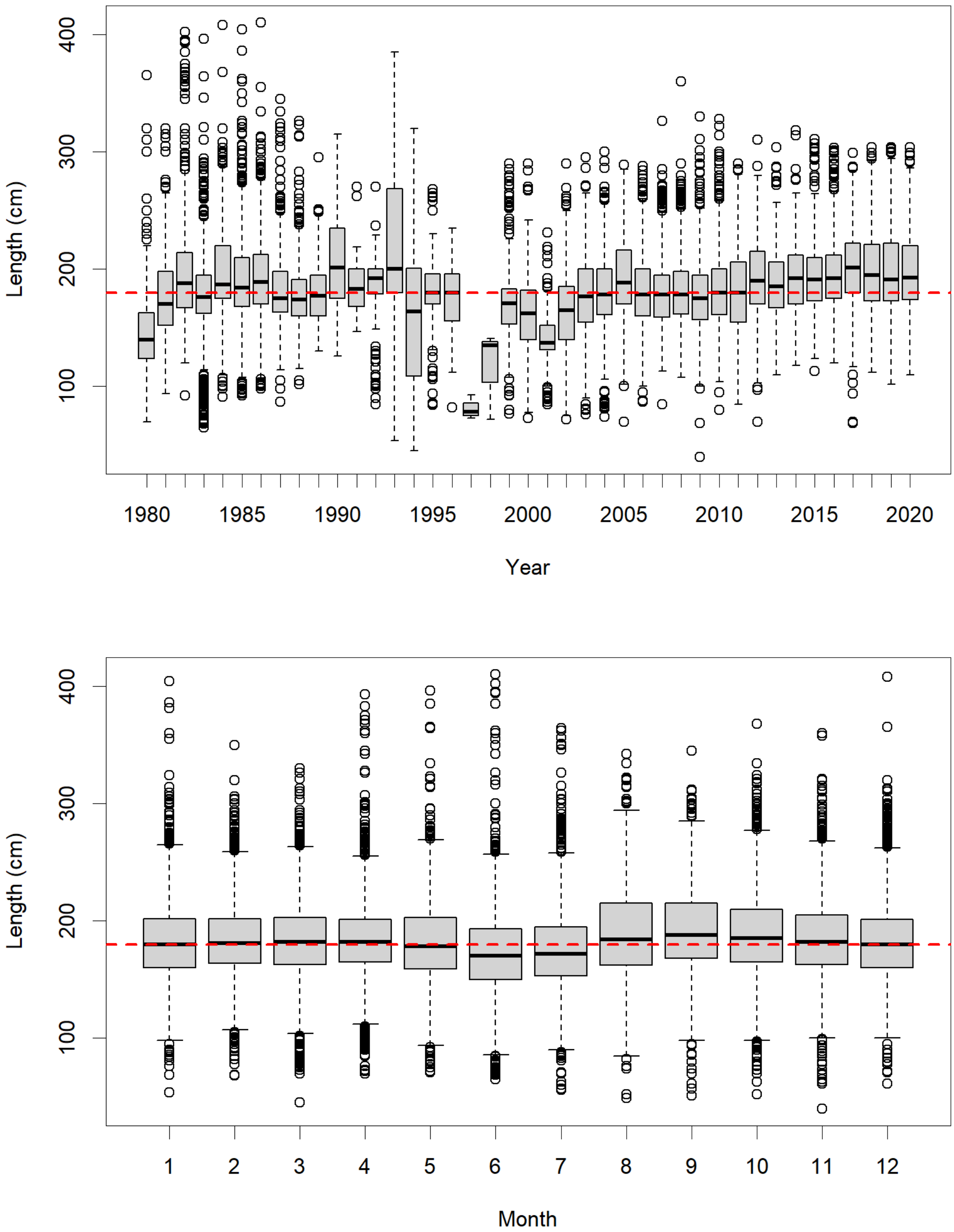
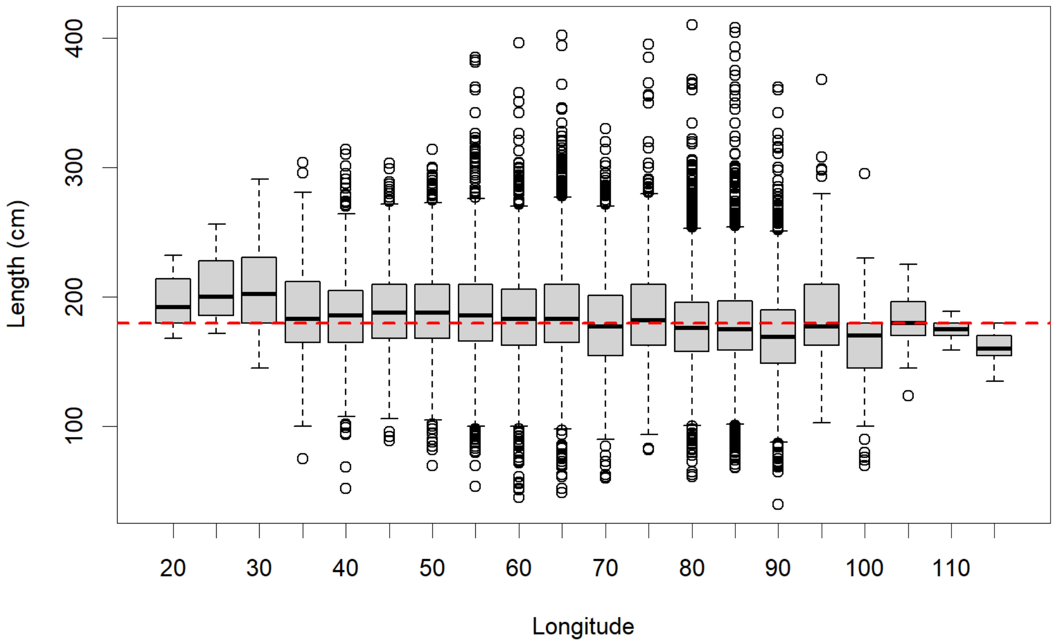


Fig. 5. The trend of the boxplot for the length data of black marlin of Taiwanese large-scale longline fishery in the Indian Ocean.



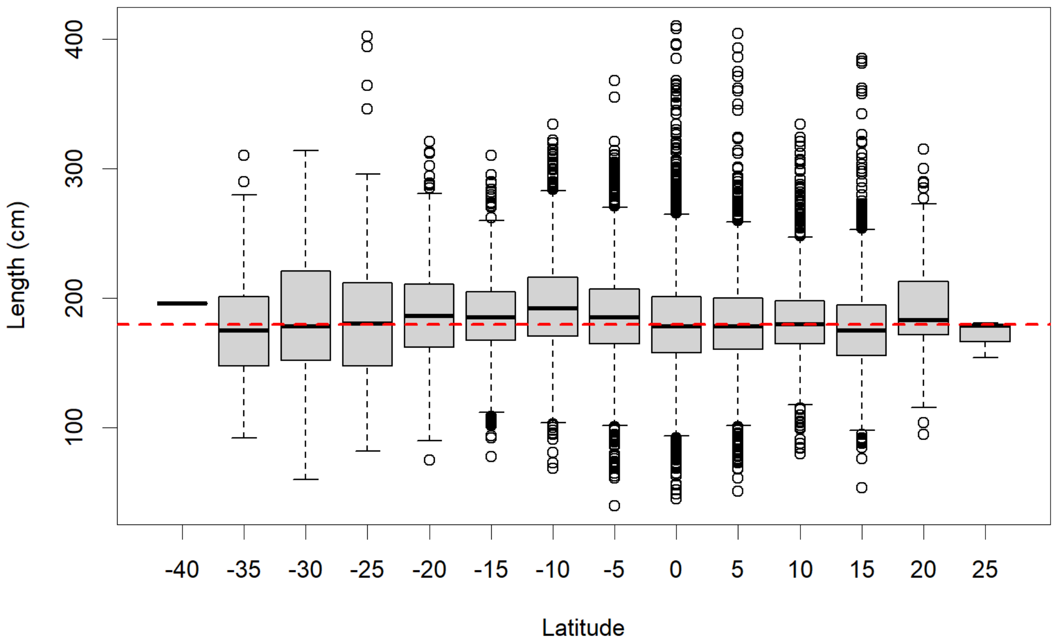


Fig. 5. (Continued).

Area NW

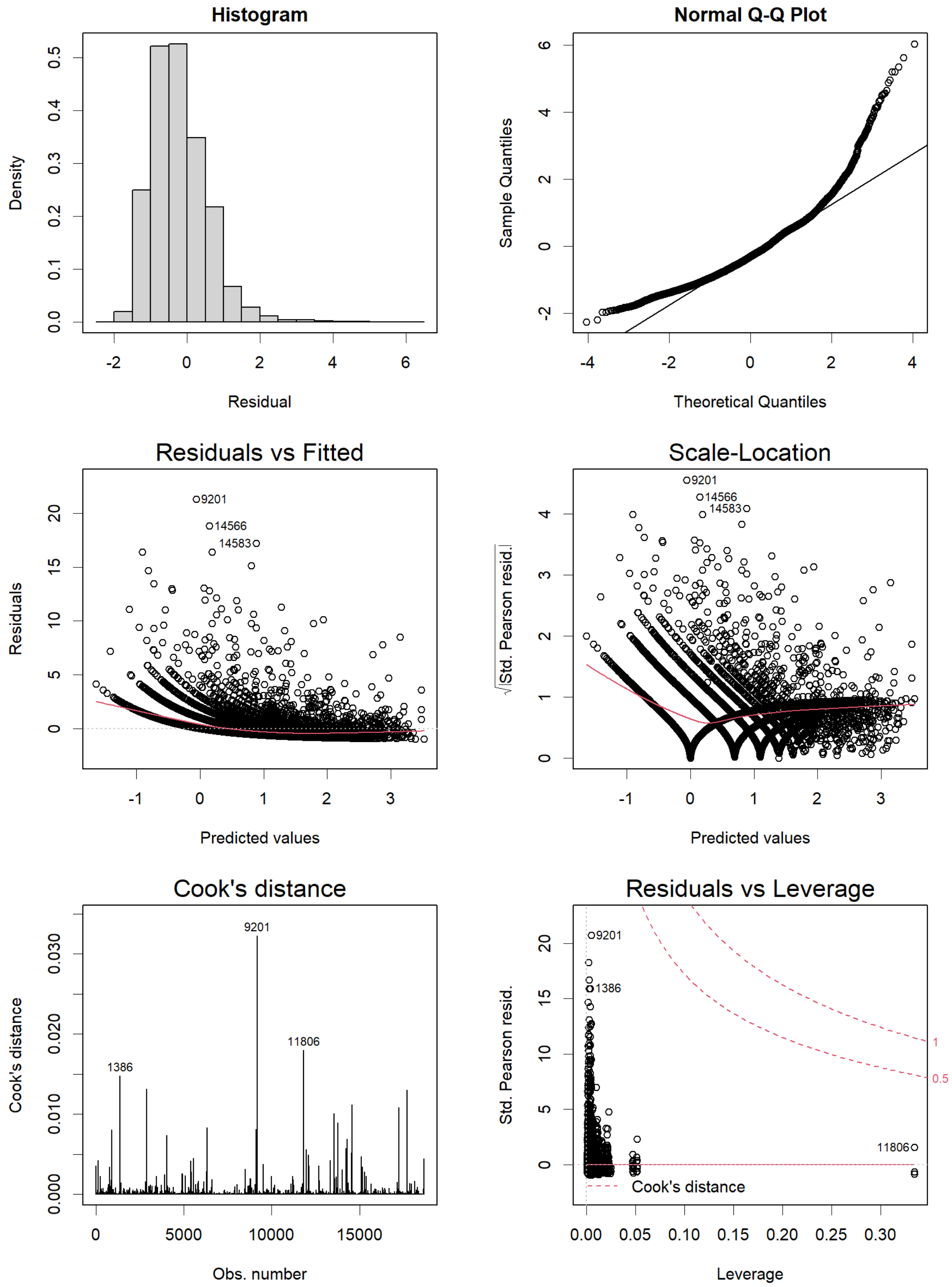


Fig. 6. Diagnostic plots for GLMs with gamma error distribution assumption for black marlin caught by Taiwanese large-scale longline fishery in the Indian Ocean from 1979 to 2020.

Area NE

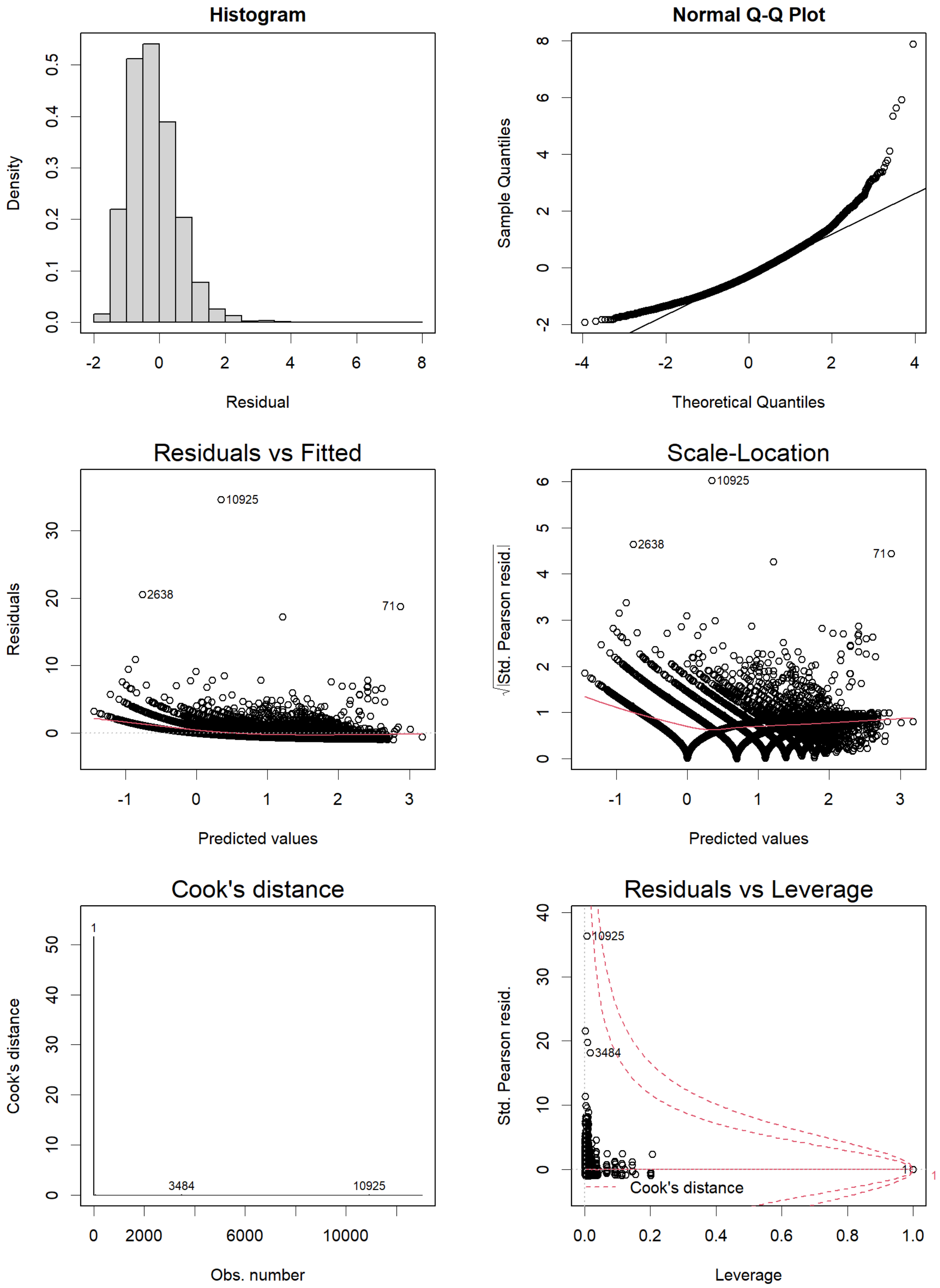


Fig. 6. (continued).

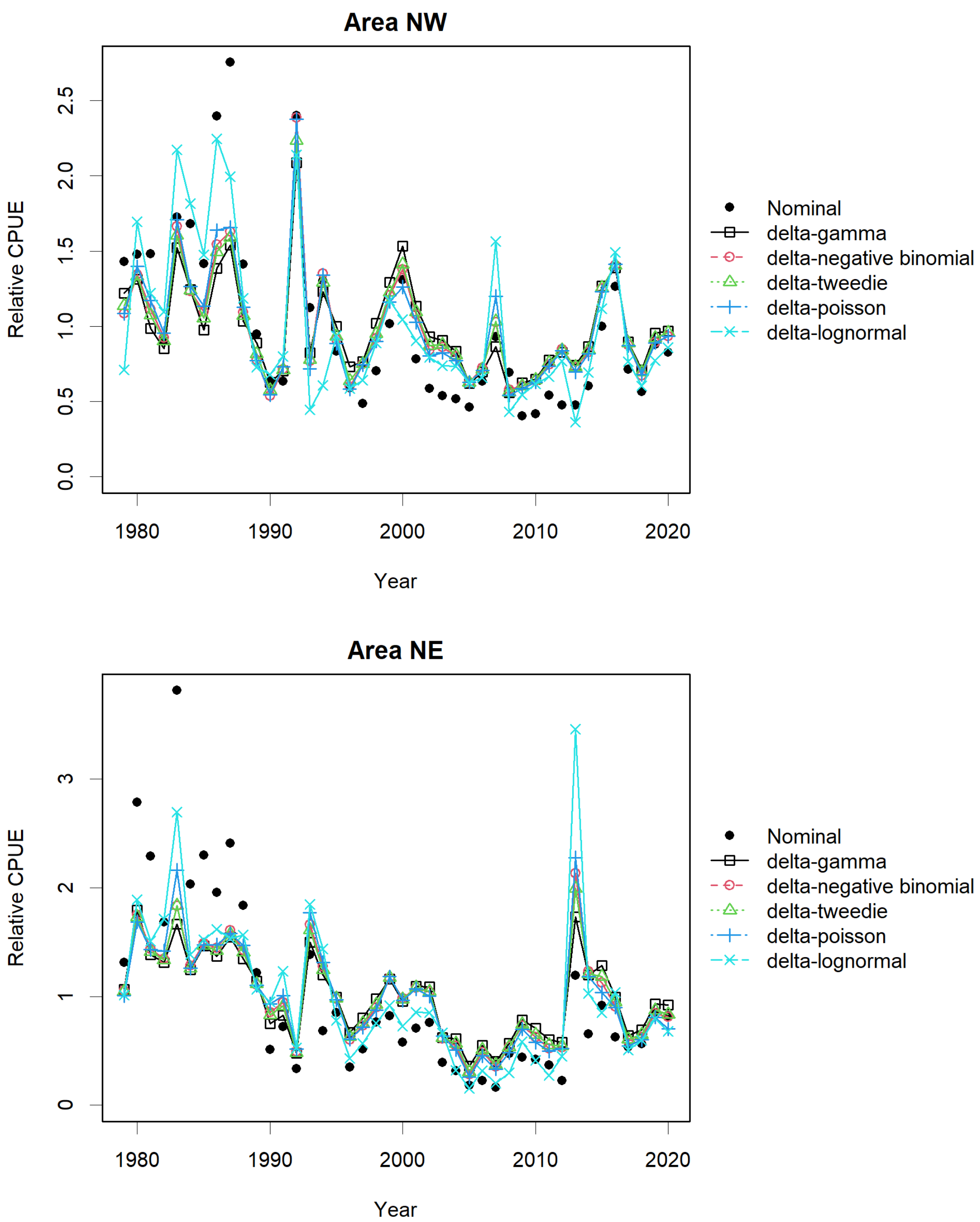
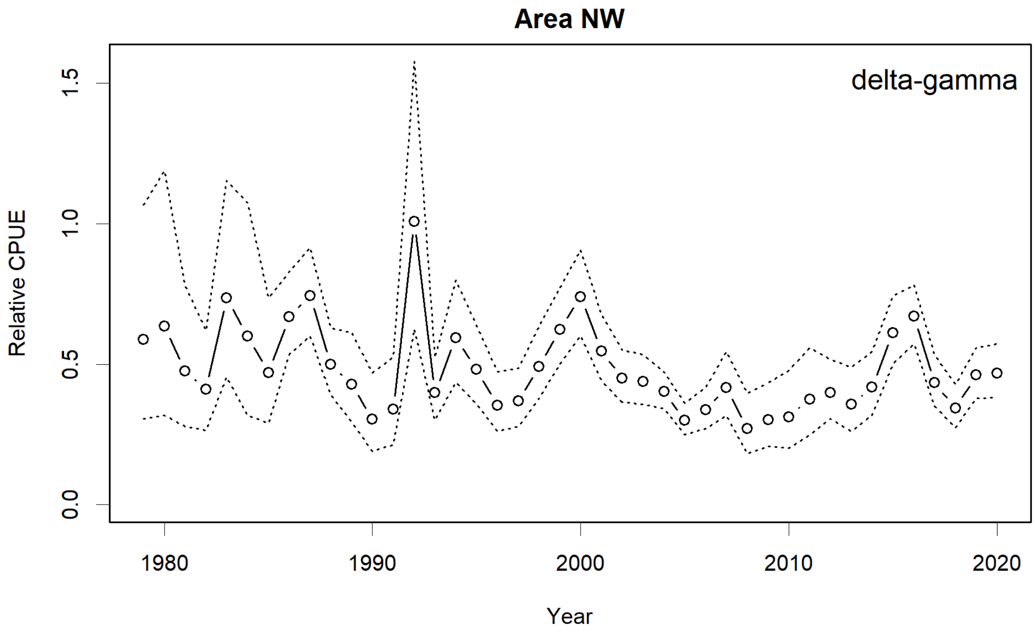


Fig. 7. Standardized CPUE series based on various GLMs for black marlin caught by Taiwanese large-scale longline fishery in the Indian Ocean from 1979 to 2020.



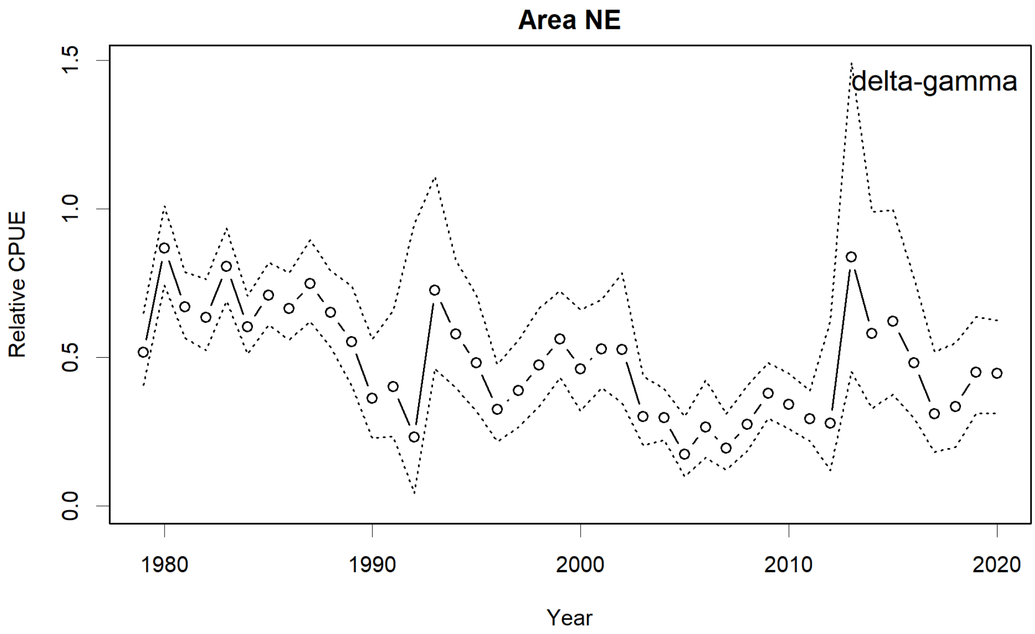


Fig. 8. Standardized CPUE series based on selected GLMs for striped marlin caught by Taiwanese large-scale longline fishery in the Indian Ocean from 1979 to 2020.

Area NW

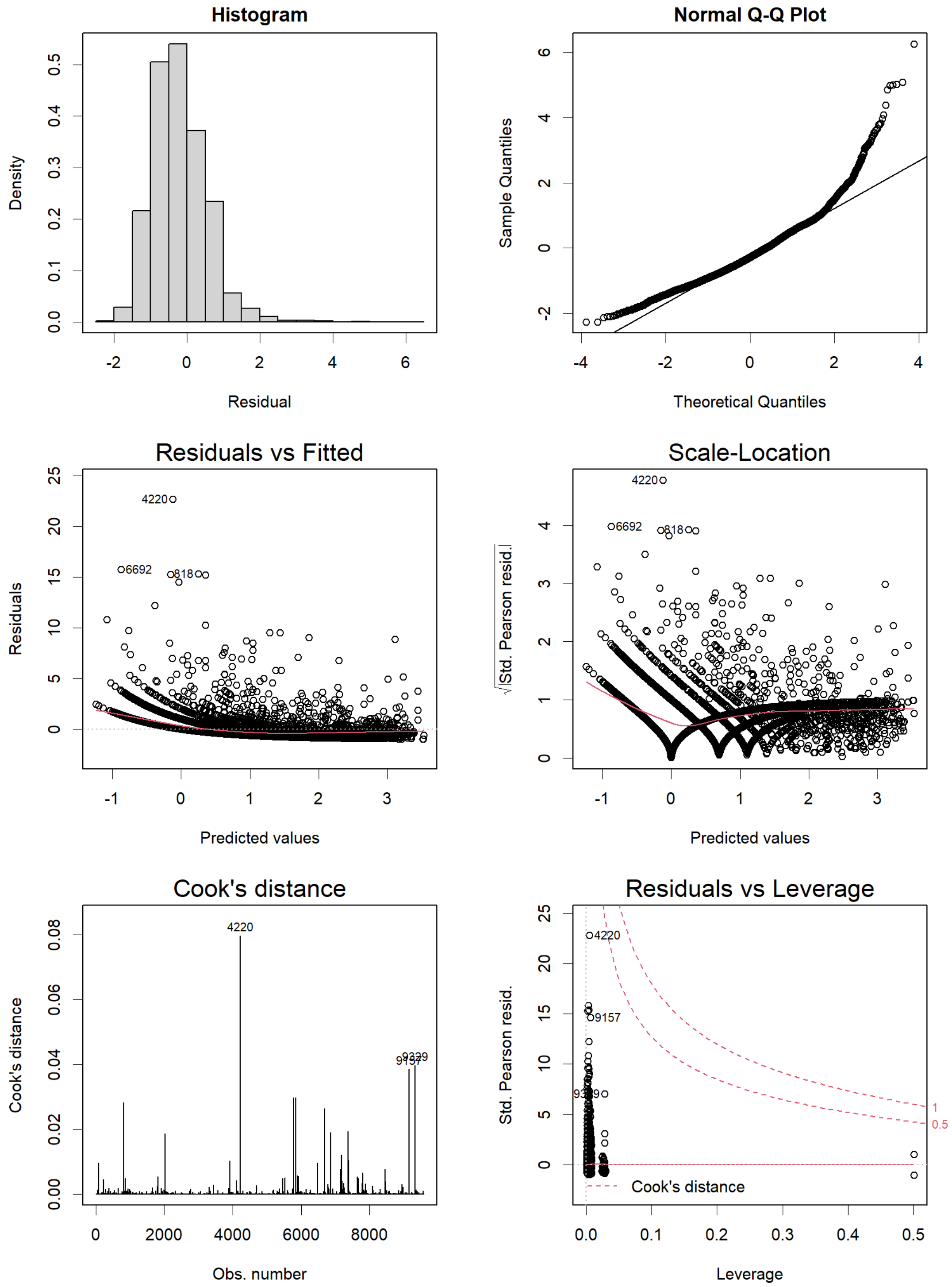


Fig. 9. Diagnostic plots for GLMs with gamma error distribution assumption for black marlin caught by Taiwanese large-scale longline fishery in the Indian Ocean from 2005 to 2020.

Area NE

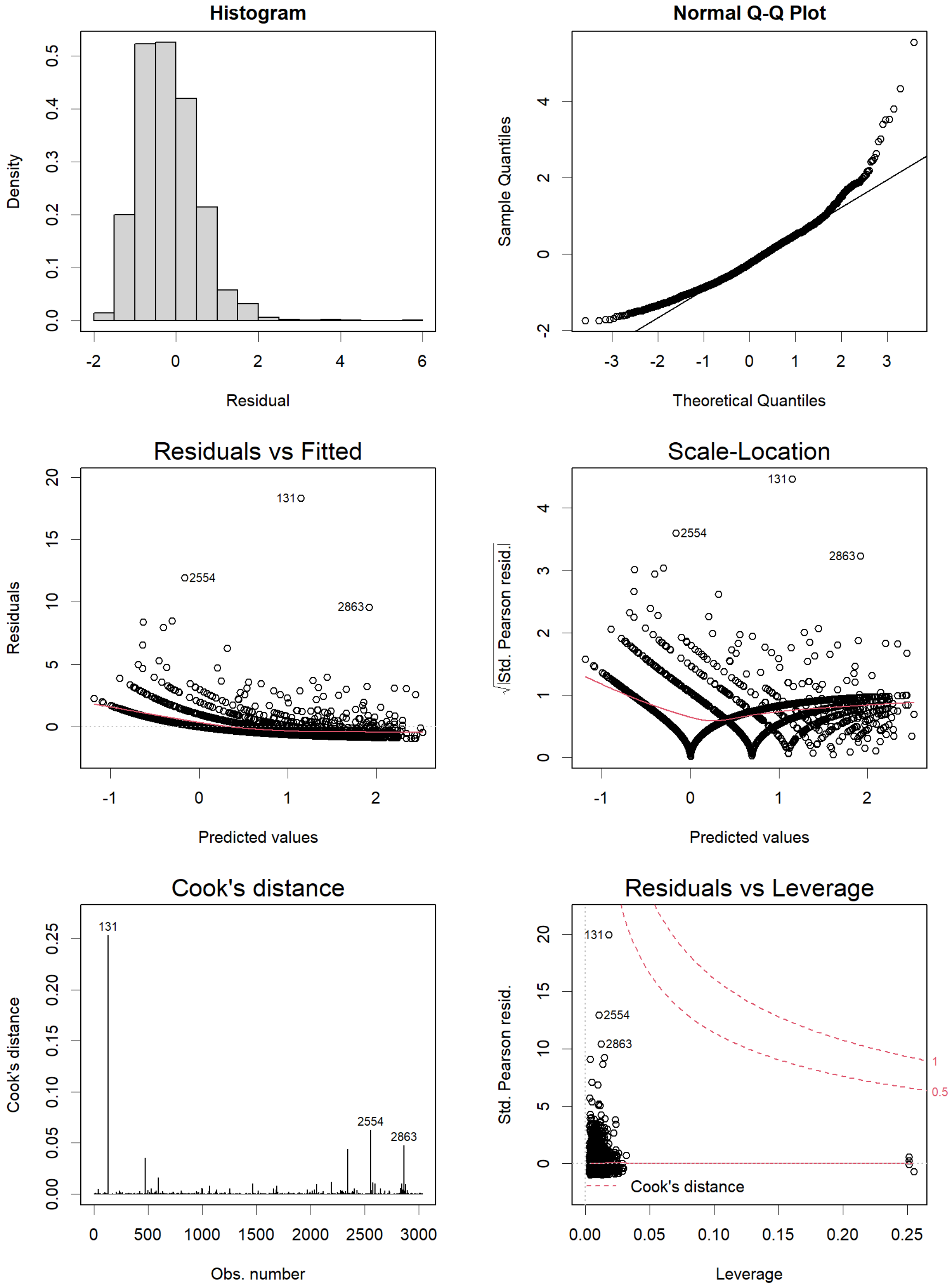


Fig. 9. (continued).

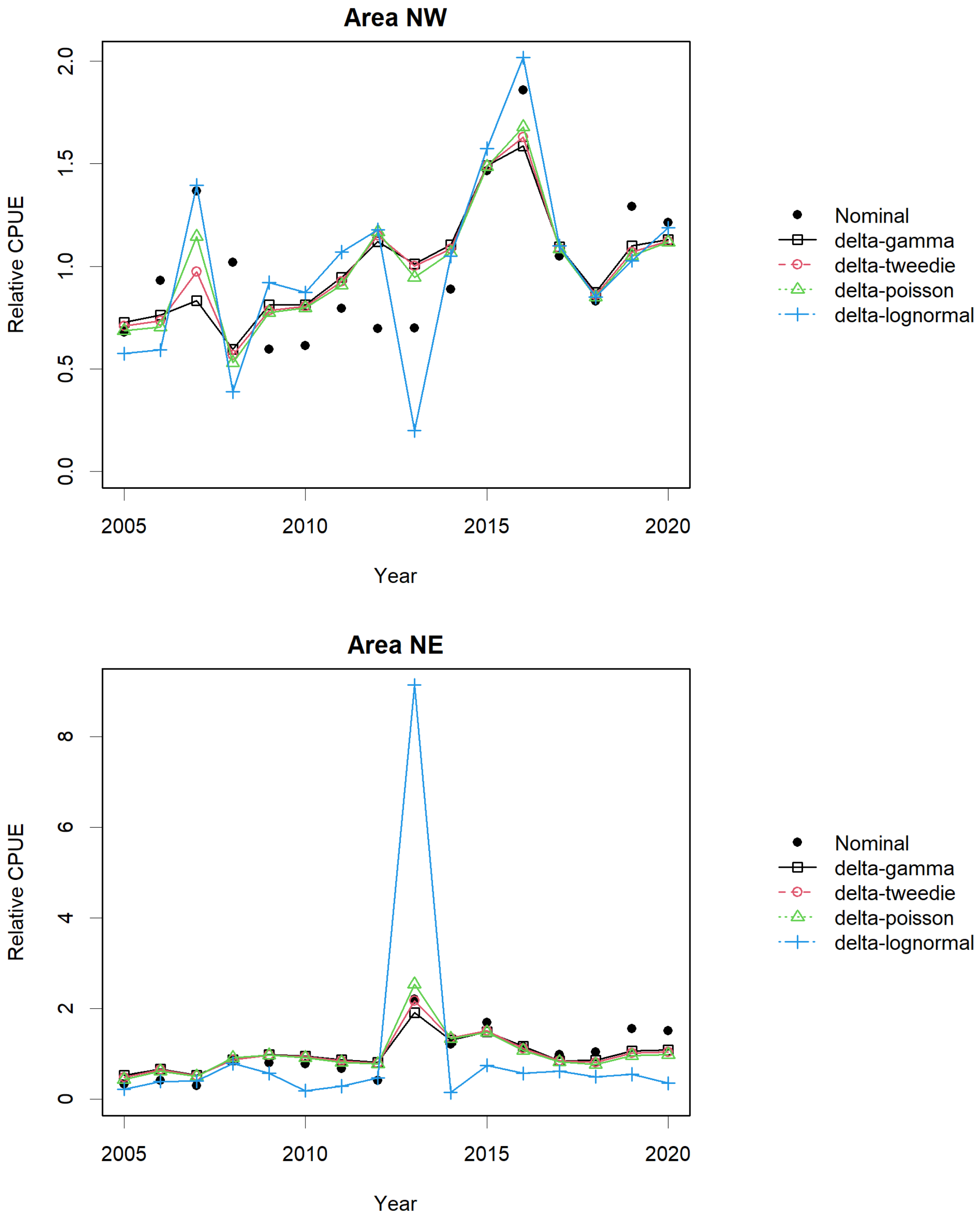


Fig. 10. Standardized CPUE series based on various GLMs for black marlin caught by Taiwanese large-scale longline fishery in the IndianOcean from 2005 to 2020.

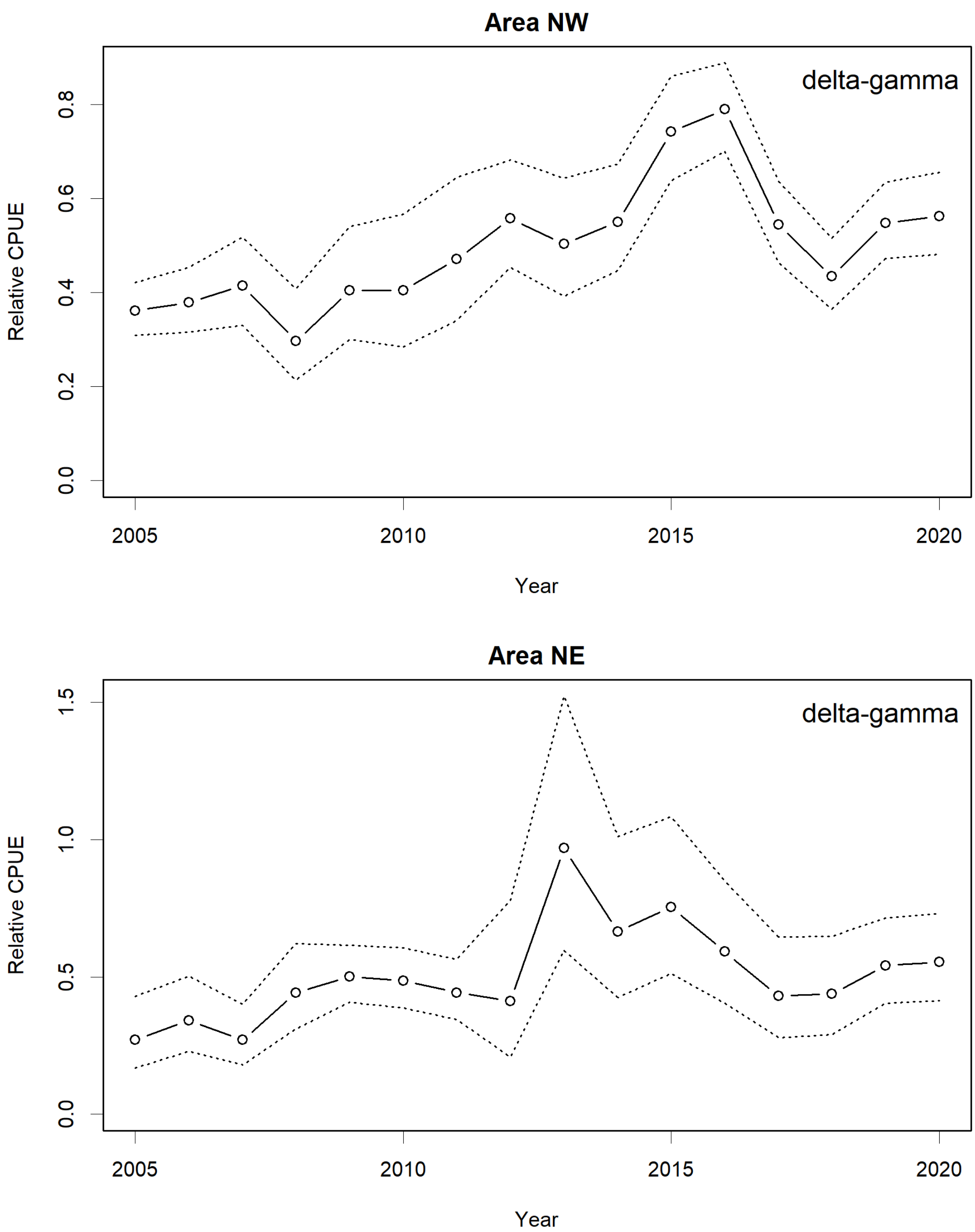


Fig. 11. Standardized CPUE series based on selected GLMs for striped marlin caught by Taiwanese large-scale longline fishery in the Indian Ocean from 2005 to 2020.

Table 1. Diagnostic statistics for standardized CPUE series based on various GLMs for black marlin caught by Taiwanese large-scale longline fishery in the Indian Ocean from 1979 to 2020.

Area	Model	R2	AIC	BIC	Dispersion
NE	gamma	0.242	45626	46150	0.913
NE	tweedie	0.284	48696	49219	1.485
NE	negative.binomial	0.304	48999	49522	1.318
NE	poisson	0.328	54020	54535	2.426
NE	lognormal	0.463	69804	70327	12.413
NW	gamma	0.191	59451	59976	1.065
NW	tweedie	0.235	64775	65300	1.700
NW	negative.binomial	0.251	65247	65772	1.414
NW	poisson	0.278	72495	73012	2.644
NW	lognormal	0.298	98639	99164	11.468

Table 2. ANOVA table for selected standardized CPUE series based on selected GLMs for black marlin caught by Taiwanese large-scale longline fishery in the Indian Ocean from 1979 to 2020.

NW

Positive catch model:

	Sum Sq	Df	F values	Pr(>F)
Y	830	41	19.0167	< 2.2e-16 ***
Q	13.2	3	4.1445	0.006047 ***
CT	32	3	10.0202	1.36E-06 ***
Lon	82.2	7	11.0334	5.41E-14 ***
Lat	68.2	8	8.0123	7.61E-11 ***
T	623.6	3	195.2465	< 2.2e-16 ***
Residuals	19813.9	18612		

Delta model

	LR Chisq	Df	Pr(>Chisq)
Y	5763.9	41	< 2.2e-16 ***
Q	166.3	3	< 2.2e-16 ***
CT	168.2	4	< 2.2e-16 ***
Lon	1008.9	8	< 2.2e-16 ***
Lat	464.7	8	< 2.2e-16 ***
T	116.3	3	< 2.2e-16 ***

Table 2. (continued).

NE

Positive catch model:

	Sum Sq	Df	F values	Pr(>F)
Y	437.1	41	11.679	< 2.2e-16 ***
Q	94.8	3	34.615	< 2.2e-16 ***
CT	93.4	5	20.474	< 2.2e-16 ***
Lon	285.1	9	34.703	< 2.2e-16 ***
Lat	89.3	7	13.979	< 2.2e-16 ***
T	501.2	3	183.03	< 2.2e-16 ***
Residuals	11820.2	12949		

Delta model

	LR Chisq	Df	Pr(>Chisq)
Y	6602.2	41	< 2.2e-16 ***
Q	146.0	3	< 2.2e-16 ***
CT	116.2	6	< 2.2e-16 ***
Lon	189.9	9	< 2.2e-16 ***
Lat	1111.8	7	< 2.2e-16 ***
T	63.9	3	8.54E-14 ***

Table 3. Diagnostic statistics for standardized CPUE series based on various GLMs for black marlin caught by Taiwanese large-scale longline fishery in the Indian Ocean from 2005 to 2020.

Area	Model	R2	AIC	BIC	Dispersion
NE	gamma	0.111	9104	9285	0.859
NE	tweedie	0.144	9870	10098	1.315
NE	negative.binomial	0.162	10161	10378	1.234
NE	poisson	0.165	10770	10980	1.968
NE	lognormal	0.284	13683	13911	5.255
NW	gamma	0.249	29847	30113	0.991
NW	tweedie	0.306	33256	33543	1.648
NW	negative.binomial	0.330	33447	33733	1.277
NW	poisson	0.351	39186	39465	2.865
NW	lognormal	0.341	54954	55226	17.986

Table 4. ANOVA table for selected standardized CPUE series based on selected GLMs for black marlin caught by Taiwanese large-scale longline fishery in the Indian Ocean from 2005 to 2020.

Area NW

Positive catch model:

	Sum Sq	Df	F values	Pr(>F)
Y	170.6	15	11.4774	< 2.2e-16 ***
CT	12.2	2	6.1624	0.002116 **
Lon	41.3	7	5.9586	6.13E-07 ***
Lat	209.8	8	26.4666	< 2.2e-16 ***
T	369.3	3	124.2283	< 2.2e-16 ***
Residuals	9464.3	9552		

Delta model

	LR Chisq	Df	Pr(>Chisq)
Y	1981.3	15	< 2.2e-16 ***
Q	103.2	3	< 2.2e-16 ***
CT	165.5	2	< 2.2e-16 ***
Lon	866.6	8	< 2.2e-16 ***
Lat	713.0	8	< 2.2e-16 ***
T	50.3	3	6.96E-11 ***

Table 4. (continued).

NE

Positive catch model:

	Sum Sq	Df	F values	Pr(>F)
Y	58.86	15	4.5684	1.05E-08 ***
Q	7.64	3	2.9634	0.03095 .
Lon				
Lat	18.69	7	3.1078	0.002868 ***
T	83.46	3	32.3875	< 2.2e-16 ***
Residuals	2581.28	3005		

Delta model

	LR Chisq	Df	Pr(>Chisq)
Y	551.9	15	< 2.2e-16 ***
Q	163.7	3	< 2.2e-16 ***
CT	54.5	2	1.50E-12 ***
Lon	20.7	7	0.0042263 **
Lat	372.1	7	< 2.2e-16 ***
T	17.0	3	0.0006955 ***