CPUE standardization of striped marlin (*Tetrapturus audax*) caught by Taiwanese large scale longline fishery in the Indian Ocean

Sheng-Ping Wang

Department of Environmental Biology and Fisheries Science, National Taiwan Ocean University, Keelung, Taiwan.

ABSTRACT

In this study, the delta-gamma general linear models with the targeting effect derived from cluster analysis were used to conduct the CPUE standardization of striped marlin caught by the Taiwanese longline fishery in the Indian Ocean for 1979-2017. The trends of CPUE series were obviously different for northern and southern Indian Ocean, while the area-aggregated CPUE series generally revealed decreasing trends since 1980s and fluctuated in recent years.

1. INTRODUCTION

Striped marlin are largely considered to be a non-target species of industrial fisheries. Longlines account for around 69% of total catches in the Indian Ocean, followed by gillnets (24%), with remaining catches recorded under troll and handlines. In recent years, the catches of striped marlin were mainly made by Indonesia (drifting longline and coastal longline, 36%), Taiwan (longline, 24%), Iran (gillnet, 14%), and Pakistan (gillnet, 8%). The catches reported under longlines are highly variable, with lower catch levels between 2009 and 2011 largely due to declining catches reported by Taiwan, deep-freezing and fresh-tuna longliners. Catches of striped marlin have since increased in 2012 and 2013, as longline vessels have resumed operations in the north-west Indian Ocean (IOTC, 2017).

The annual proportion of striped marlin caught by Taiwanese large scale longline fishery was generally less 3% of total catches except for the years before the late 1980s, and revealed a decreasing trend since 1980s (Wang, 2017). Based on the areas defined by Wang and Nishida (2011) (Fig. 1), the catches of striped marlin are mainly made by Taiwanese large scale longline operated in the northern areas (Fig. 2). The nominal CPUE distribution of striped marlin of Taiwanese large scale longline fishery also indicated that high values of CPUE occurred in tropical and subtropical areas in the 1980s and 1990s; CPUE substantially decreased since 2000s and high CPUE occurred in the offshore area in the norther area in 2000s and in the western area in 2010s (Wang, 2017).

Because striped marlin was bycatch species of Taiwanese lognline fishery, large amount of zero-catches was recorded in the operational catch and effort data sets of Taiwanese longline fishery. In recent decades, the annual proportions of zero-catch were about 70-90% of total data sets. In previous study (Wang, 2015), the delta-lognormal GLM (Pennington, 1983; Lo et. al., 1992; Pennington, 1996) was applied to conduct CPUE standardization of striped marlin in the Indian Ocean but the model with lognormal assumption for the residuals might not appropriate for fitting to the data. Therefore, a delta-gamma GLM was adopted in this study. In addition, the targeting of fishing operation was identified from the cluster analyses as recommended by the Fifth IOTC CPUE Workshop.

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 1980-2017 were provided by Oversea Fisheries Development Council of Taiwan (OFDC). It should be noted that the data in 2017 is preliminary.

The data of number of hooks between float (NHBF) were available since 1994 and the collection of NHBF data were more complete since 1995. Therefore, the data of NHBF may not be applicable to conduct the long-term CPUE standardization for fishes caught by Taiwanese longline fishery in the Indian Ocean.

2.2. Cluster analysis

The cluster analysis (He et al., 1997) was adopted to conduct to explore the targeting of fishing operations and to produce the data filter for selecting the data for CPUE standardization. Cluster analysis was performed based on species composition of the catches of albacore (ALB), bigeye tuna (BET), yellowfin tuna (YFT), swordfish (SWO), blue marlin (BUM), striped marlin (MLS), black marlin (BLM) and other species (OTH). However, clustering operational set-by-set data might include large amount noise because most of billfishes were caught by Taiwanese vessels as bycatches. Therefore, the cluster analysis was performed based weekly-aggregated data and then merged the clusters with operational data sets to identify the targeting fishing operations.

The hierarchical cluster analysis with Ward minimum variance method was applied to the squared Euclidean distances calculated from the aggregated data sets. The analyses were performed using R functions hclust and cutree (The R Foundation for Statistical Computing Platform, 2018).

He et al. (1997) indicated that the choice for the number of clusters to produce was largely subjective. At least two clusters were expected. More than two clusters were produced to allow other possible categories to emerge. Additional clusters were considered until the smallest cluster contained very few efforts. In this study, we kept the proportion of data sets of the smallest cluster was about 5-10%. In addition, cluster analyses were performed by four fishing areas separately.

2.3. CPUE Standardization

A delta-gamma GLM was applied to standardize the CPUE. As the approach of Wang (2017), the models were simply conducted with the main effects considered in this analysis were year, month, 5x5 longitude-latitude grid, and the effects related to the fishing configurations (clusters), while interactions between main effects were not incorporated into the models. In addition, CPUE standardizations were also performed by four fishing areas separately. The gamma and delta models were conducted as follows:

Gamma model for CPUE of positive catch:

$$\log(CPUE) = \mu + Y + M + G + T + \varepsilon^{\mathrm{gamma}}$$

Delta model for presence and absence of catch:

$$PA = \mu + Y + M + CT + G + T + \varepsilon^{del}$$

where	$CPUE$	is the nominal CPUE of positive catch of striped marlin (catch in number/1,000 hooks),
	PA	is the nominal presence and absence of catch,
	μ	is the intercept,
	Y	is the effect of year,
	M	is the effect of month,
	CT	is the effect of vessel scale,
	G	is the effect of 5x5 longitude-latitude grid,
	T	is the effect of targeting (principal component scores (PC_i)

derived from the ith principle component),

ε^{gamma} is the error term, ε^{gamma} ~ Gamma distribution with log link function,

ε^{del} is the error term, ε^{del} ~ Binomial distribution.

The models performed by stepwise search ("both" direction, i.e. "backward" and "forward") and selected based on the values of the coefficient of determination (R^2), Akaike information criterion (AIC) and Bayesian information criterion (BIC). 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:

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

where $CPUE$ is the adjust means (least square means) of the year effect of the gamma model,

$\tilde{P}$ is the adjust means (least square means) of the year effect of the delta model.

2.4. Area-aggregated CPUE series

The estimation of annual standardized CPUE was calculated from the weighted average of the area indices (Punt et al., 2000):

$$U_y = \sum_a S_a^1 U_{y,a}$$

Where U_y is CPUE for year y,

$U_{y,a}$ is CPUE for year y and area a,

S_a^1 is the relative size of the area a.

The relative sizes of nine IOTC statistics areas for swordfish in the Indian Ocean (Nishida and Wang et al., 2006) were used to be aggregated into four areas used in this study.

Area	NW	NE	SW	SE
Relative area size	0.2478	0.2577	0.1638	0.3307

In addition, area-specific standardized CPUE was also aggregated by the proportions of annual area-specific catch and effort data:

$$U_y = \sum_a S_{y,a}^2 U_{y,a}$$

Where $S_{y,a}^2$ is the proportion of the catch or hooks in year y and area a.

3. RESULTS AND DISCUSSION

3.1. Cluster analysis

Four clusters were selected for all of four areas (Fig. 3). For Area NW, Cluster 1 mainly belonged to the BET operations; Cluster 2 consisted of mixture of the BET and YFT operations; Cluster 3 contained the operations for OTH; operations of Cluster 4 were mainly for ALB (Figs. 4 and 5). Cluster 4 contained more shallow operations and the fishing ground mainly located in the norther waters, while differences in other data categories were relatively minor (Fig. 6). Based on the spatial distribution of MLS catch proportion, high catches generally occurred in the coastal waters (Fig. 7). MLS catches were made from all of four clusters but catches of Cluster 1 obviously increased after early 1990s (Fig. 8).

For Area NE, Cluster 1 also belonged to the BET operations; Cluster 2 mainly consisted of the BET operations but also contained the operations for OTH; Cluster 3 and 4 were YFT and ALB operations, respectively (Figs. 9 and 10). Cluster 2 mainly consisted of the data after the late 2000s; the data of Cluster 3 were from the years and distributed in tropical area; operations of Cluster 4 concentrated in the waters around 10°S (Fig. 11). High MLS catches occurred in the waters of Bay of Bengal (Fig. 12). Cluster 1 contained more MLS catches but large amount of catches was also made from the operations of Clusters 2 and 3 in early years (Fig. 13).

For Area SW, operations of Cluster 1 were mainly for YFT; Cluster 2 consisted of mixture of the BET and YFT operations; Cluster 3 contained the operations for OTH; Cluster 4 mainly consisted of operations for ALB (Figs. 14 and 15). The data of Cluster 3 were mainly from the years after the late 2000s and operations with deep sets deployed in the western waters in the beginning of the year; Cluster 4 contained the data in the years before 2000s and operations concentrated in the waters around 50°E (Fig. 16). Few MLS catches were made in the area but relative higher catches occurred in the waters around South Africa (Fig. 17). Most of MLS catches were from Cluster 4 before the early 1990s, changed to be made by Cluster 2 during 1990s and 2000s, and MLS catches substantially decreased for all of four clusters (Fig. 18).

For Area SE, Cluster 1 consisted of mixture of the ALB and OTH operations; Cluster 2 mainly contained operations for ALB; operations of Cluster 3 were mainly for OTH; Cluster 4 belonged to the mixture of the BET and YFT operations (Figs. 19 and 20). Cluster 2 contained more data than other clusters and mainly from the years before early 2000s; operations of Cluster 4 concentrated in the waters with lower latitude (Fig. 21). High catches occurred in the waters around the central Indian Ocean and the coastal area of Australia (Fig. 22). Most of MLS catches were made by the operations of Cluster 2 but Clusters 1 and 4 also contained large proportion of MLS catches after early 1990s (Fig. 23).

3.2. CPUE standardization

Based on the model selections for the gamma models incorporated clusters as the effects related to targeting of operations, all of main effects were statistically significant and remained in the models. The ANOVA tables for selected gamma models are shown in the Table 1. The results indicate that the effects of *T* (clusters) provided significant contributions to explanation of variance for the models for all of four areas. Thus, the targeting of fishing operation might influence the CPUE derived from the positive catch of striped marlin.

For the delta models, all of the effects were statistically significant and remained in the models. The ANOVA tables for selected delta models are shown in the Table 2. Comparing to the gamma models for positive catches, the effect of *T* (clusters) were less influential for the catch probability although this effect still significant in the models for all of four areas. The results indicated that the catch probability of striped marlin in the Indian Ocean might be mainly influenced by spatial effect.

The area-specific standardized CPUE series are shown in Fig. 24. The trends of CPUE series in the northern areas (NW and NE) reveal similar trends, substantially decreased since 1980s, and increased in some of recent years. In the southern areas (SW and SE), the CPUE fluctuated before the early 2000s, substantially decreased until the late 2000s, and slightly increased in recent years. In 2017, the values of standardized CPUE decreased for all of four areas.

Since very few MLS catches were made in southern areas, the area-aggregated CPUE series were calculated based on the results from northern areas. Although the CPUE series aggregated by various weightings were slightly different, the trends generally revealed decreasing trends since 1980s and fluctuated in recent years (Fig. 25).

3.3. Retrospect analysis

The retrospect analysis was conducted to test the influence of including the updated data on the CPUE standardization. The analysis was performed by removing the data from 2017 to 2012. The results indicated that the influence of including the updated data on the CPUE standardization was negligible for all of four areas (Fig. 26).

REFERENCE

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.

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

IOTC, 2017. Report of the 15th Session of the IOTC Working Party on Billfish. IOTC-2017-WPB15-R[E].

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.

Punt, A. E., Walker, T.I., Taylor, B.L., Pribac, F., 2000. Standardization of catch and effort data in a spatially-structured shark fishery. Fish. Res. 45: 129-145.

Wang, S.P., 2015. CPUE standardization of striped marlin (*Kajikia audax*) caught by Taiwanese longline fishery in the Indian Ocean using targeting effect derived from cluster and principle component analyses. IOTC–2015–WPB13–31 Rev_1.

Wang, S.P., 2017. CPUE standardization of striped marlin (*Tetrapturus audax*) caught by Taiwanese longline fishery in the Indian Ocean. IOTC–2017–WPB15–29_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.

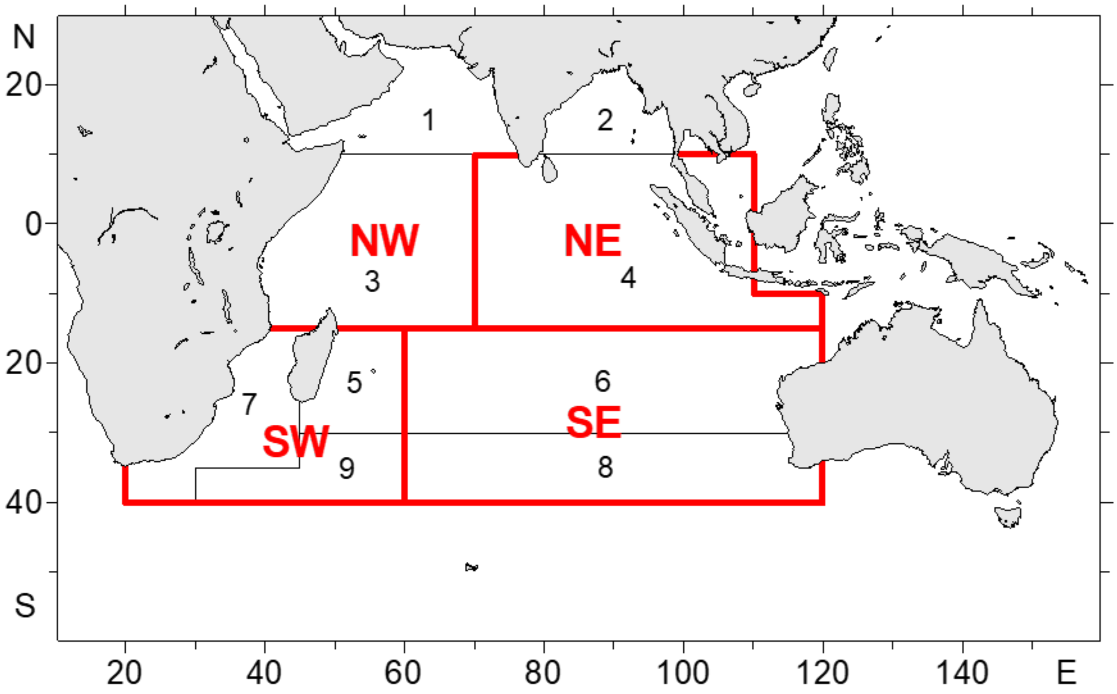


Fig. 1. Area stratification used for striped marlin in the Indian Ocean.

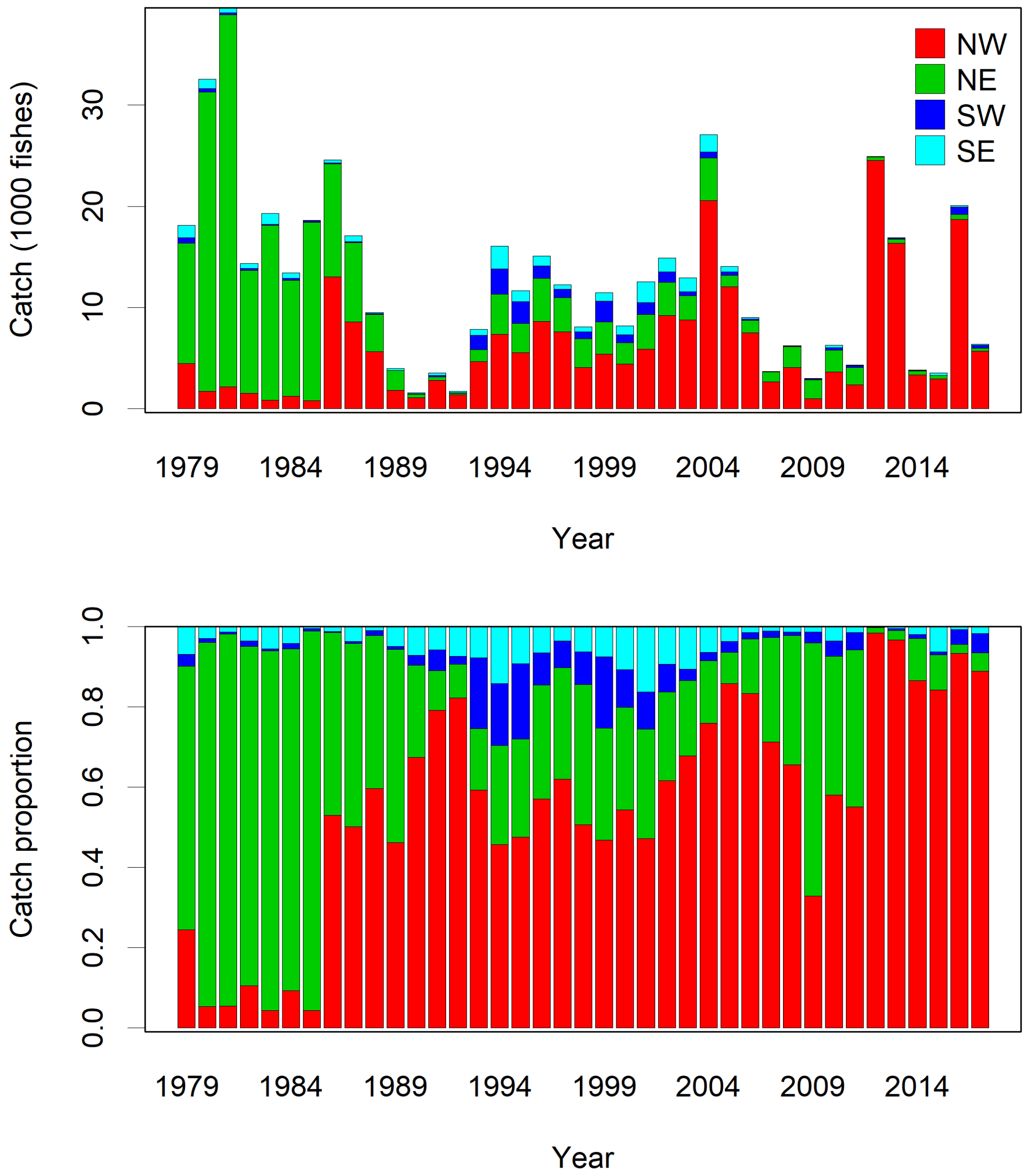


Fig. 2. The catches and catch proportions by areas of striped marlin caught by Taiwanese large scale longline operated in the Indian Ocean.

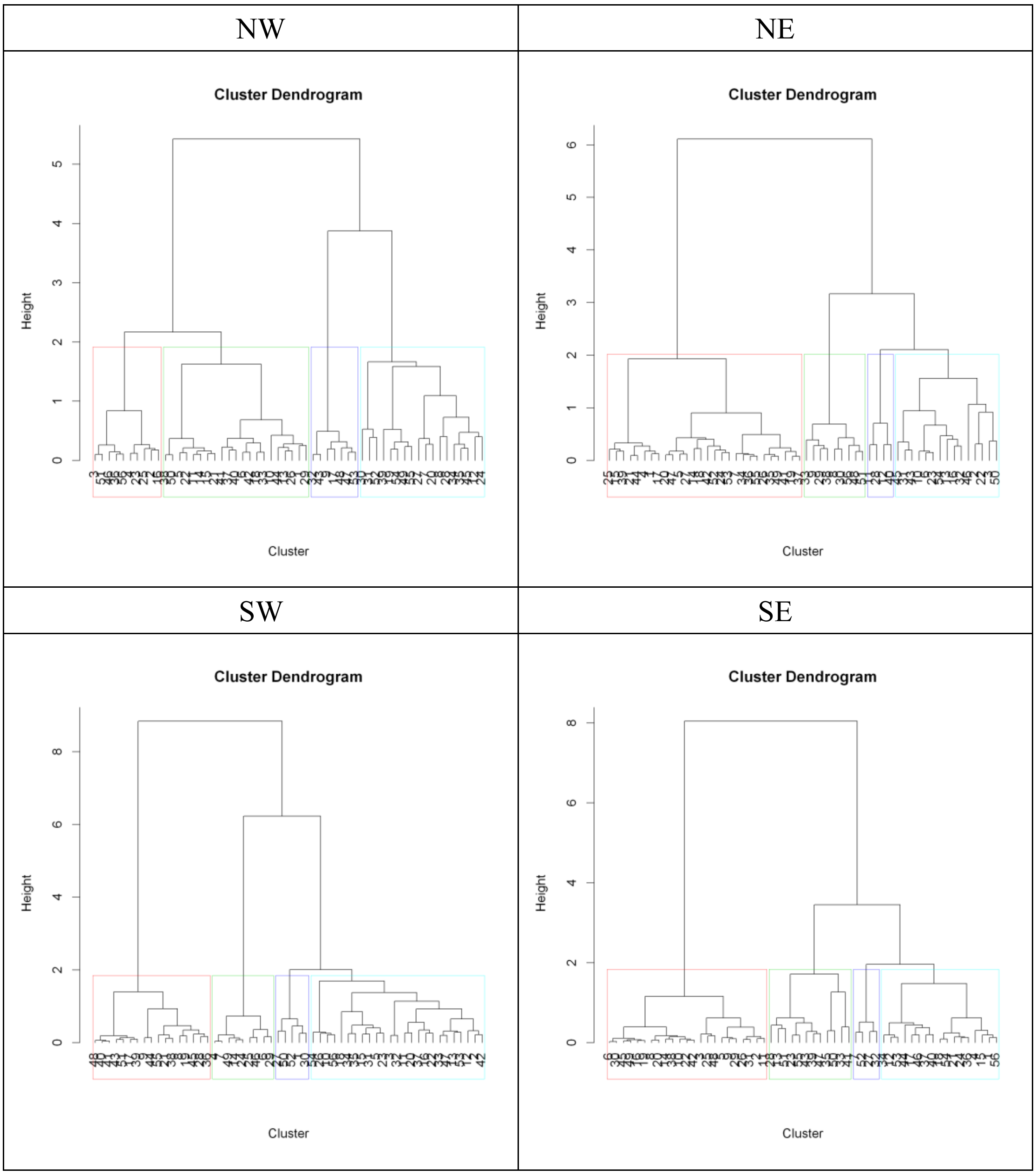


Fig. 3. Cluster trees for the data of Taiwanese large scale longline fishery in the Indian Ocean.

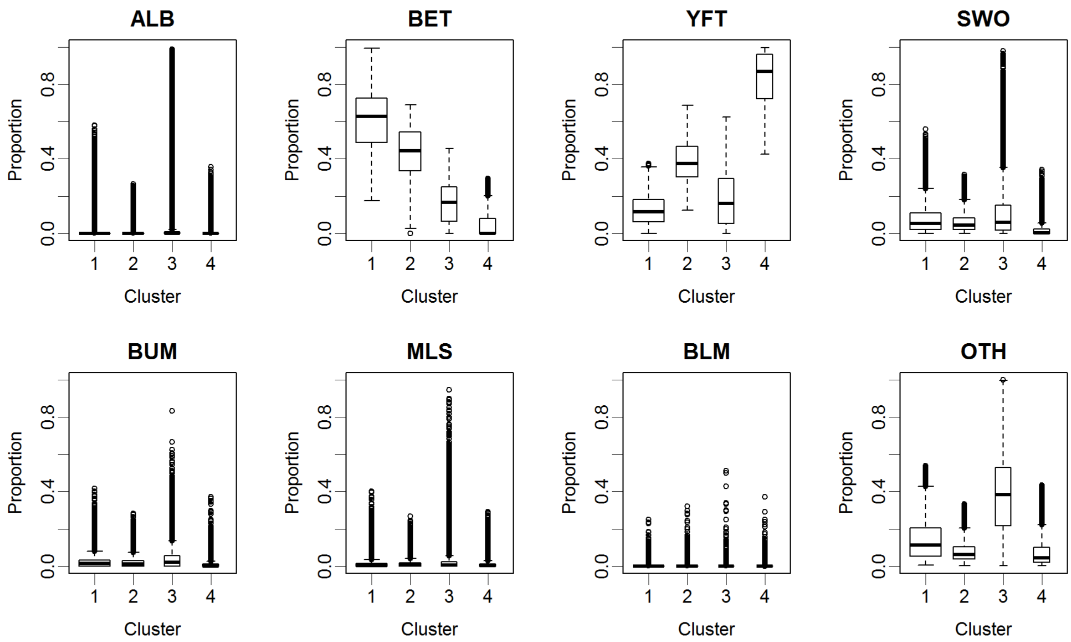


Fig. 4. Catch proportion by species for each cluster of Taiwanese large scale longline fishery in Area NW of the Indian Ocean.

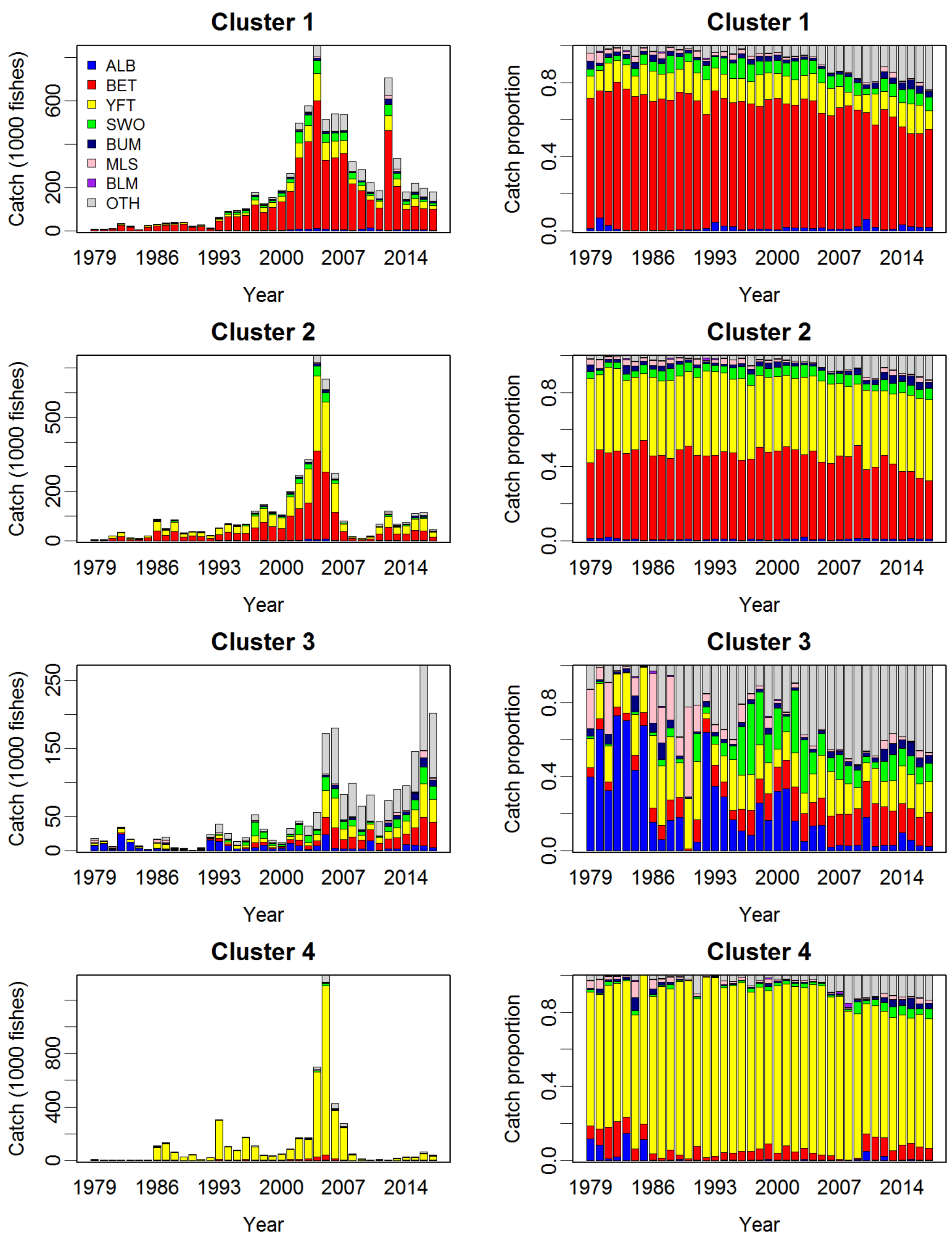


Fig. 5. Annual and catch proportion by species for each cluster of Taiwanese large scale longline fishery in Area NW of the Indian Ocean.

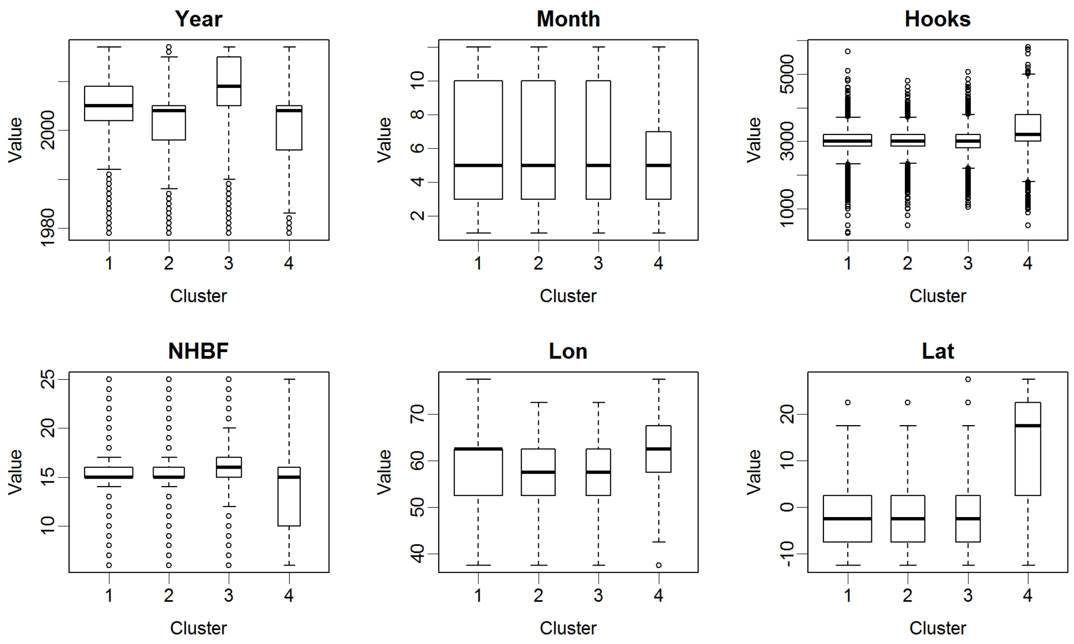


Fig. 6. Data composition by factors for each cluster of Taiwanese large scale longline fishery in Area NW of the Indian Ocean.

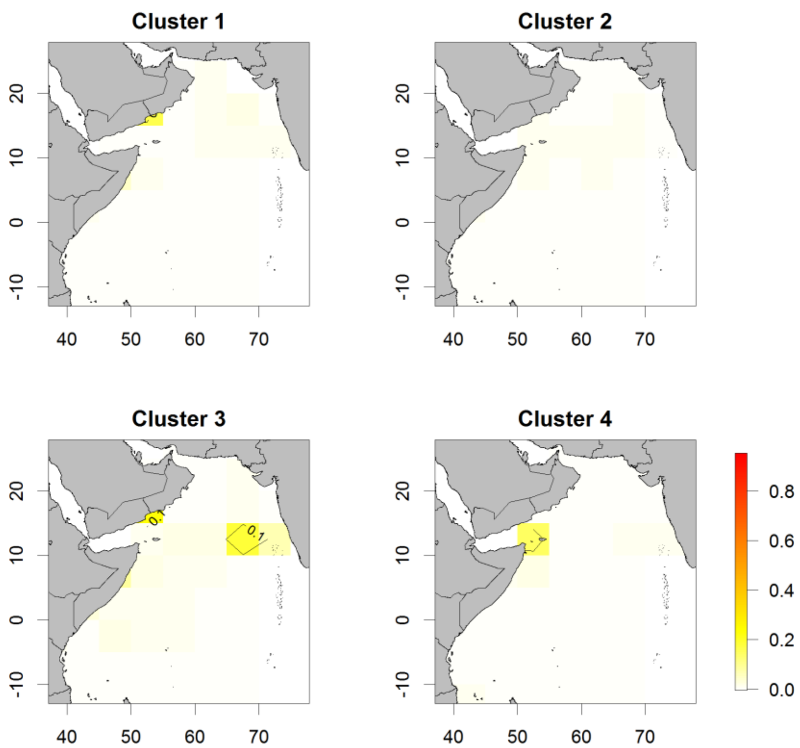


Fig. 7. Striped marlin catch distribution for each cluster of Taiwanese large scale longline fishery in Area NW of the Indian Ocean. Yellow is high catch and red is low catch.

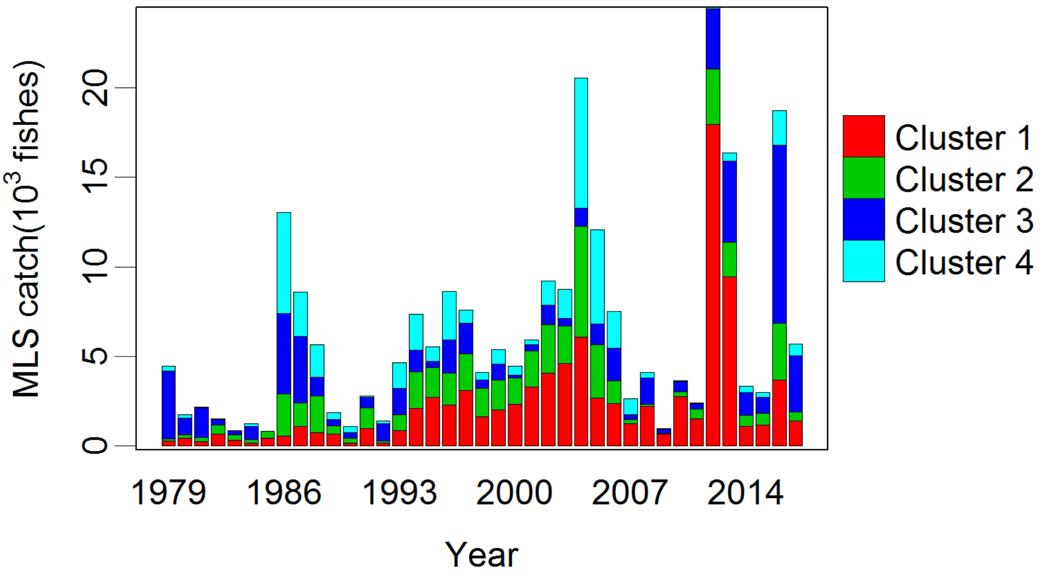


Fig. 8. Annual striped marlin catches for each cluster of Taiwanese large scale longline fishery in Area NW of the Indian Ocean.

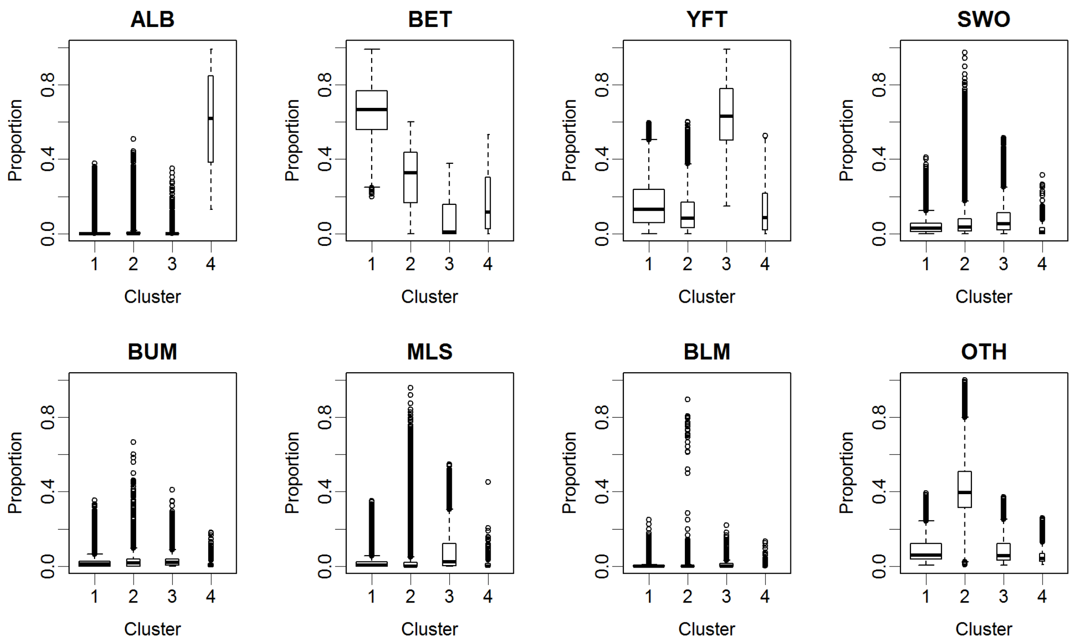


Fig. 9. Catch proportion by species for each cluster of Taiwanese large scale longline fishery in Area NE of the Indian Ocean.

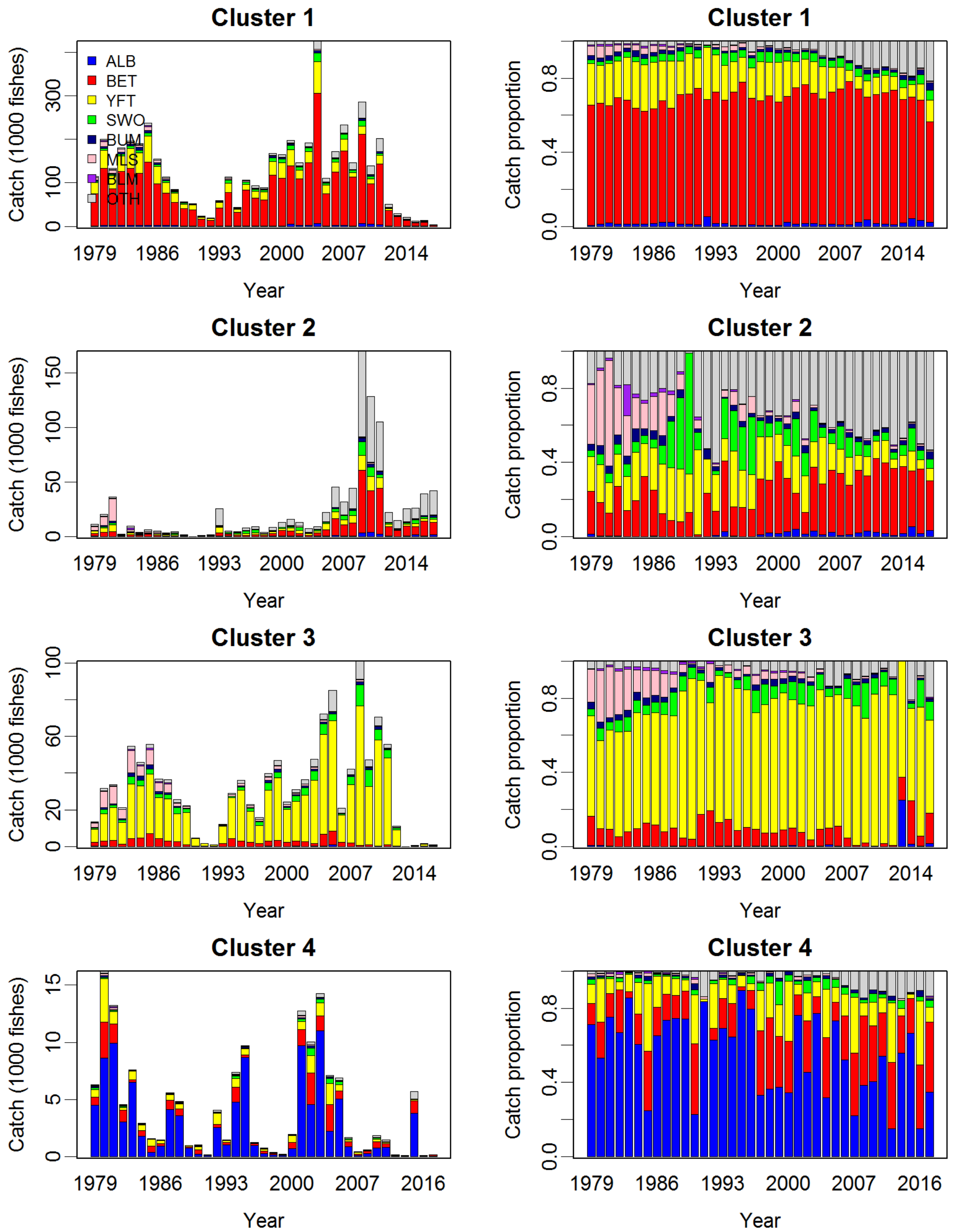


Fig. 10. Annual catch and catch proportion by species for each cluster of Taiwanese large scale longline fishery in Area NE of the Indian Ocean.

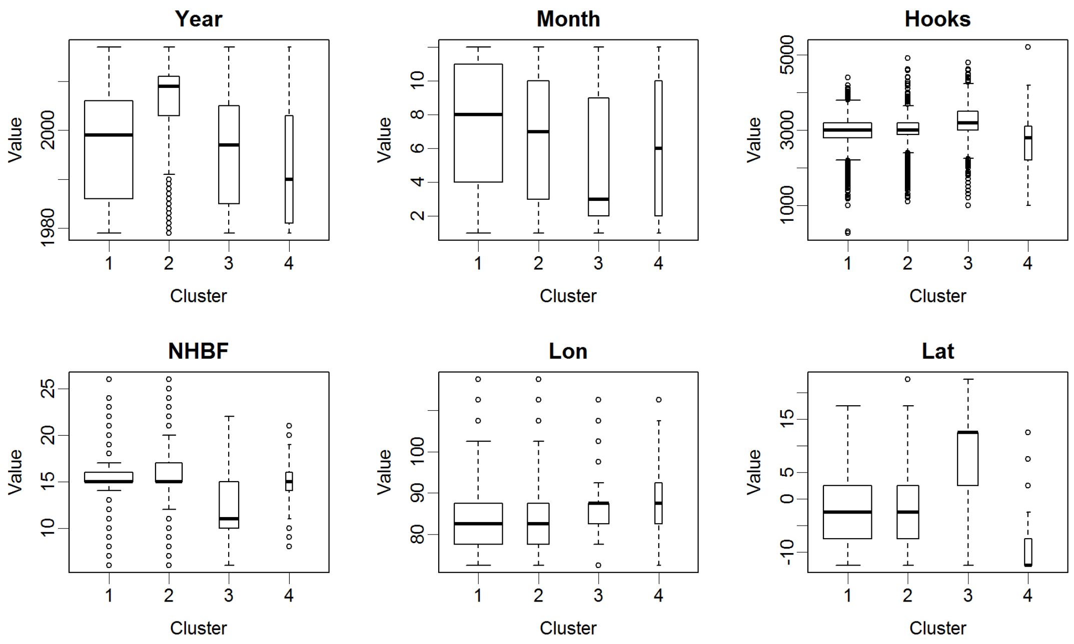


Fig. 11. Data composition by factors for each cluster of Taiwanese large scale longline fishery in Area NE of the Indian Ocean.

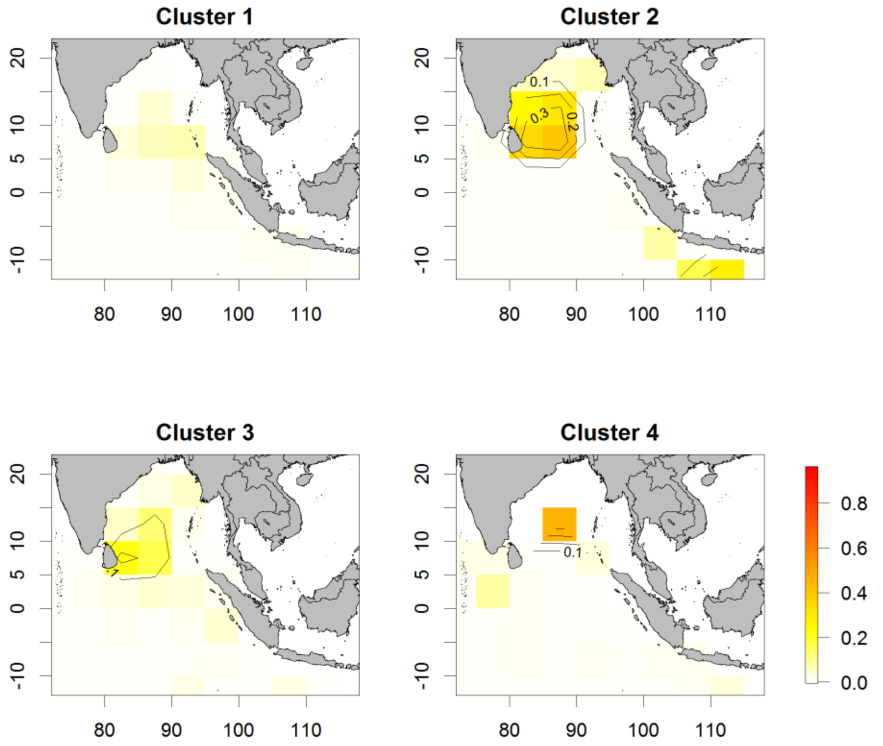


Fig. 12. Striped marlin catch distribution for each cluster of Taiwanese large scale longline fishery in Area NE of the Indian Ocean. Yellow is high catch and red is low catch.

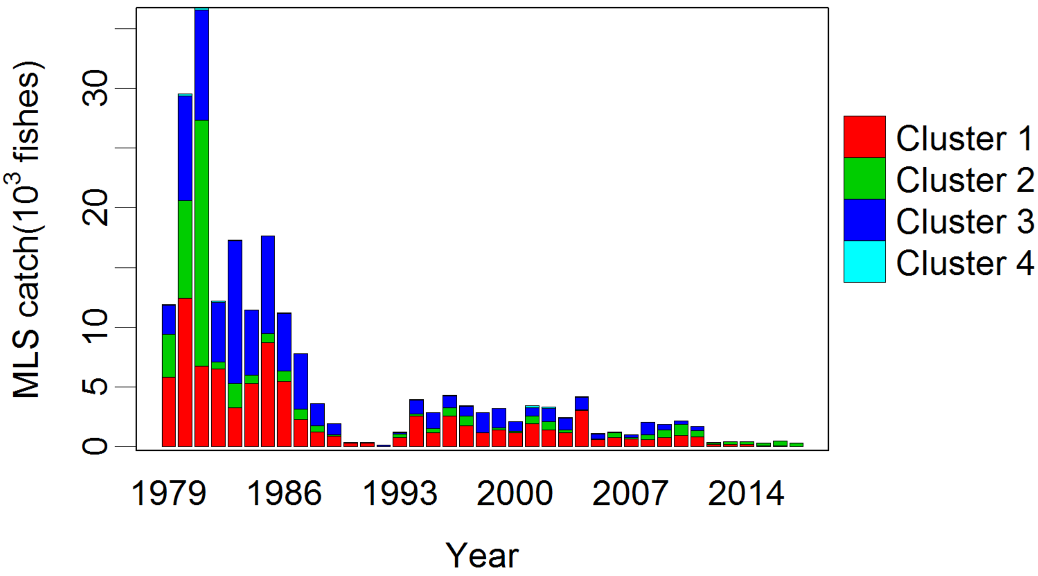


Fig. 13. Annual striped marlin catches for each cluster of Taiwanese large scale longline fishery in Area NE of the Indian Ocean.

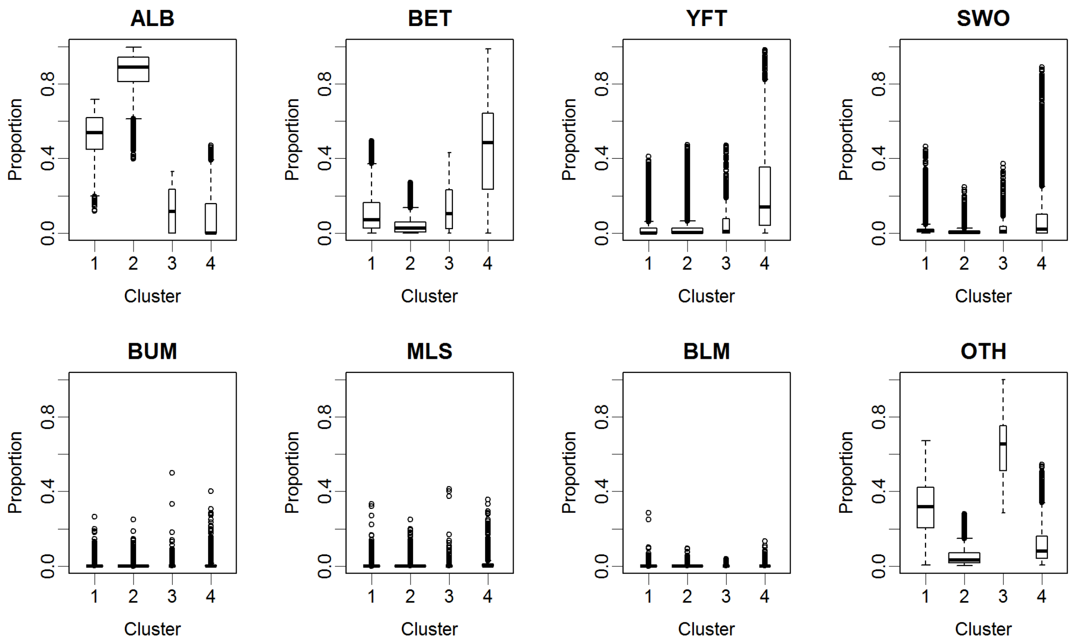


Fig. 14. Catch proportion by species for each cluster of Taiwanese large scale longline fishery in Area SW of the Indian Ocean.

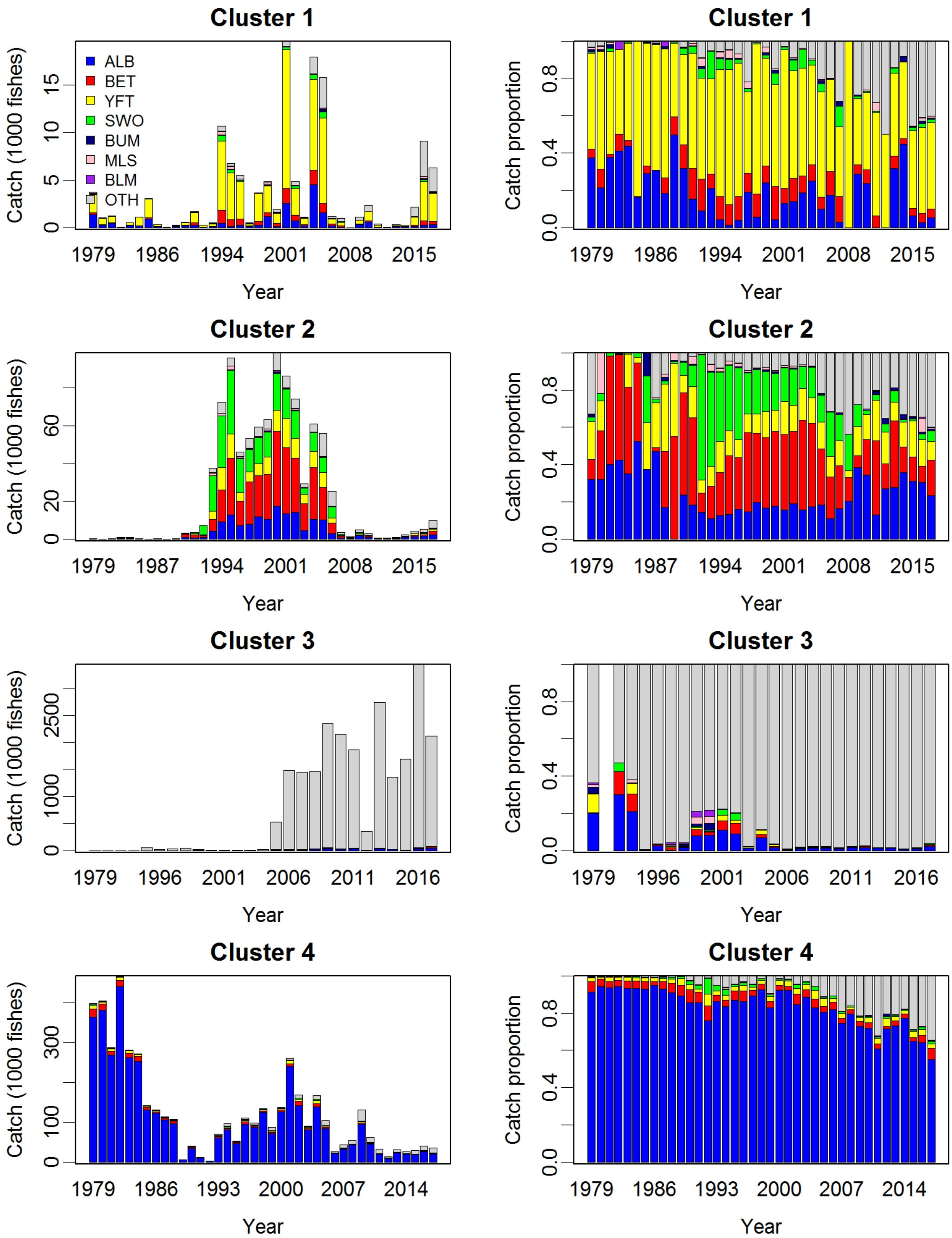


Fig. 15. Annual and catch proportion by species for each cluster of Taiwanese large scale longline fishery in Area SW of the Indian Ocean.

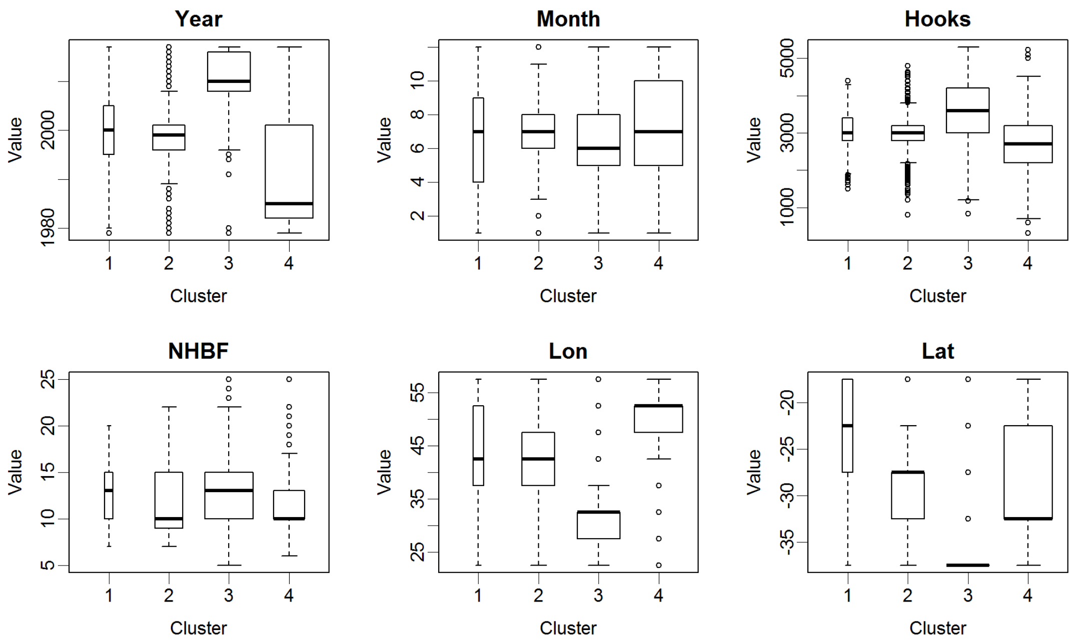


Fig. 16. Data composition by factors for each cluster of Taiwanese large scale longline fishery in Area SW of the Indian Ocean.

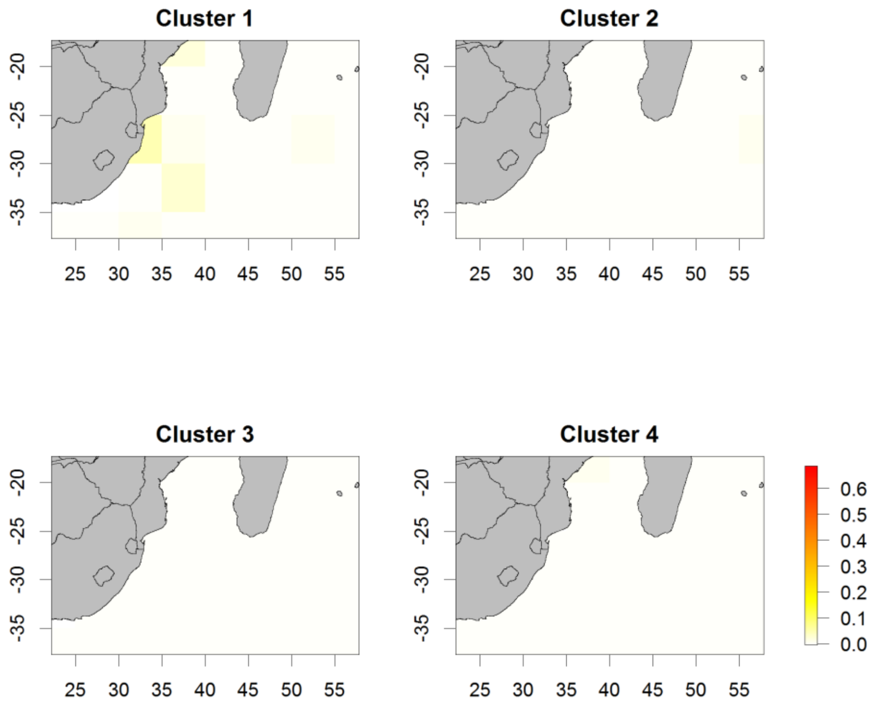


Fig. 17. Striped marlin catch distribution for each cluster of Taiwanese large scale longline fishery in Area SW of the Indian Ocean. Yellow is high catch and red is low catch.

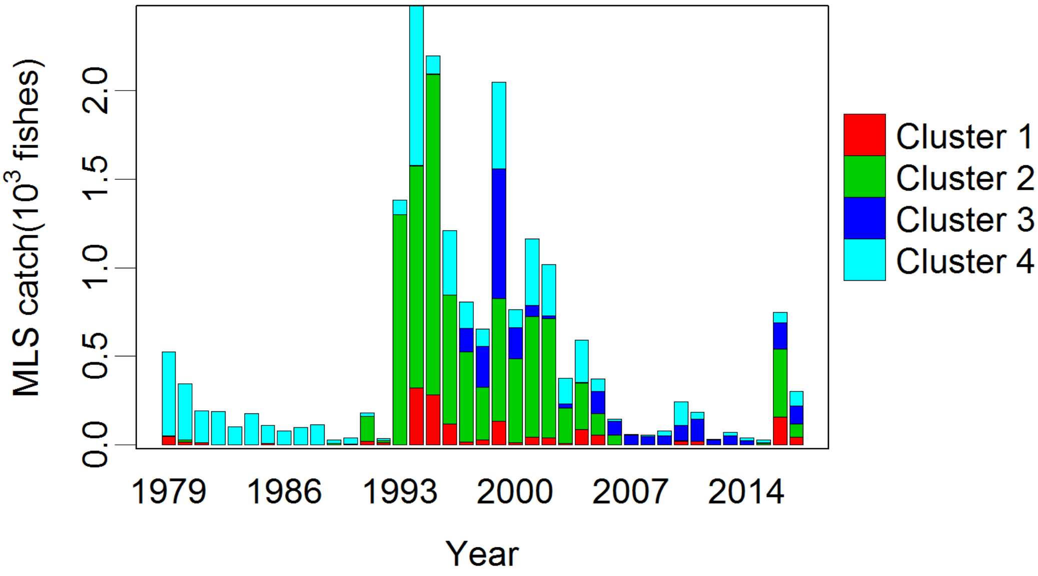


Fig. 18. Annual striped marlin catches for each cluster of Taiwanese large scale longline fishery in Area SW of the Indian Ocean.

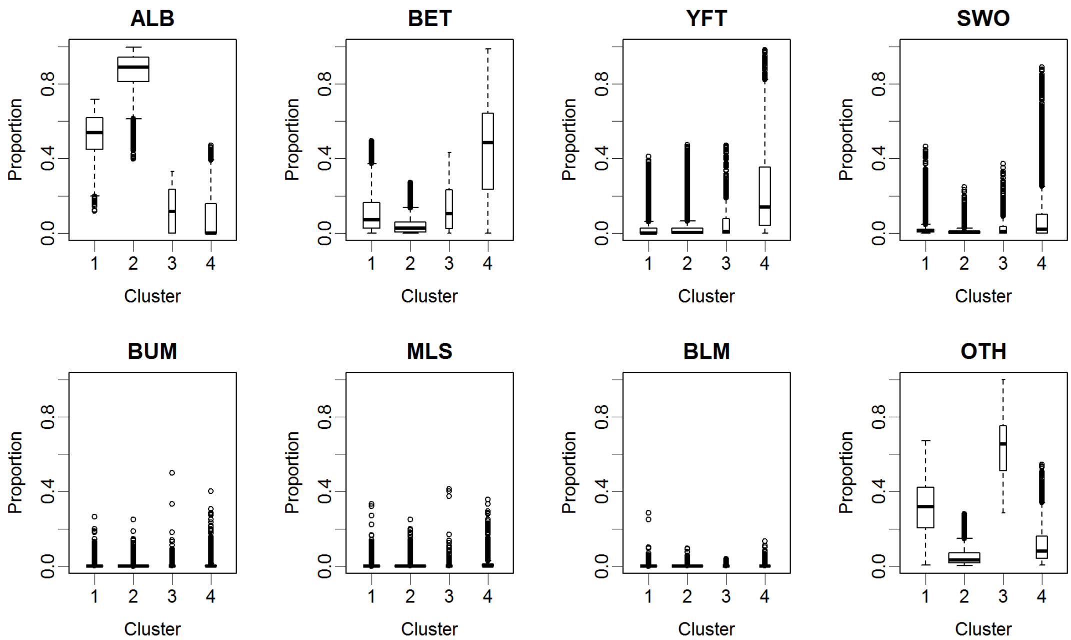


Fig. 19. Catch proportion by species for each cluster of Taiwanese large scale longline fishery in Area SE of the Indian Ocean.

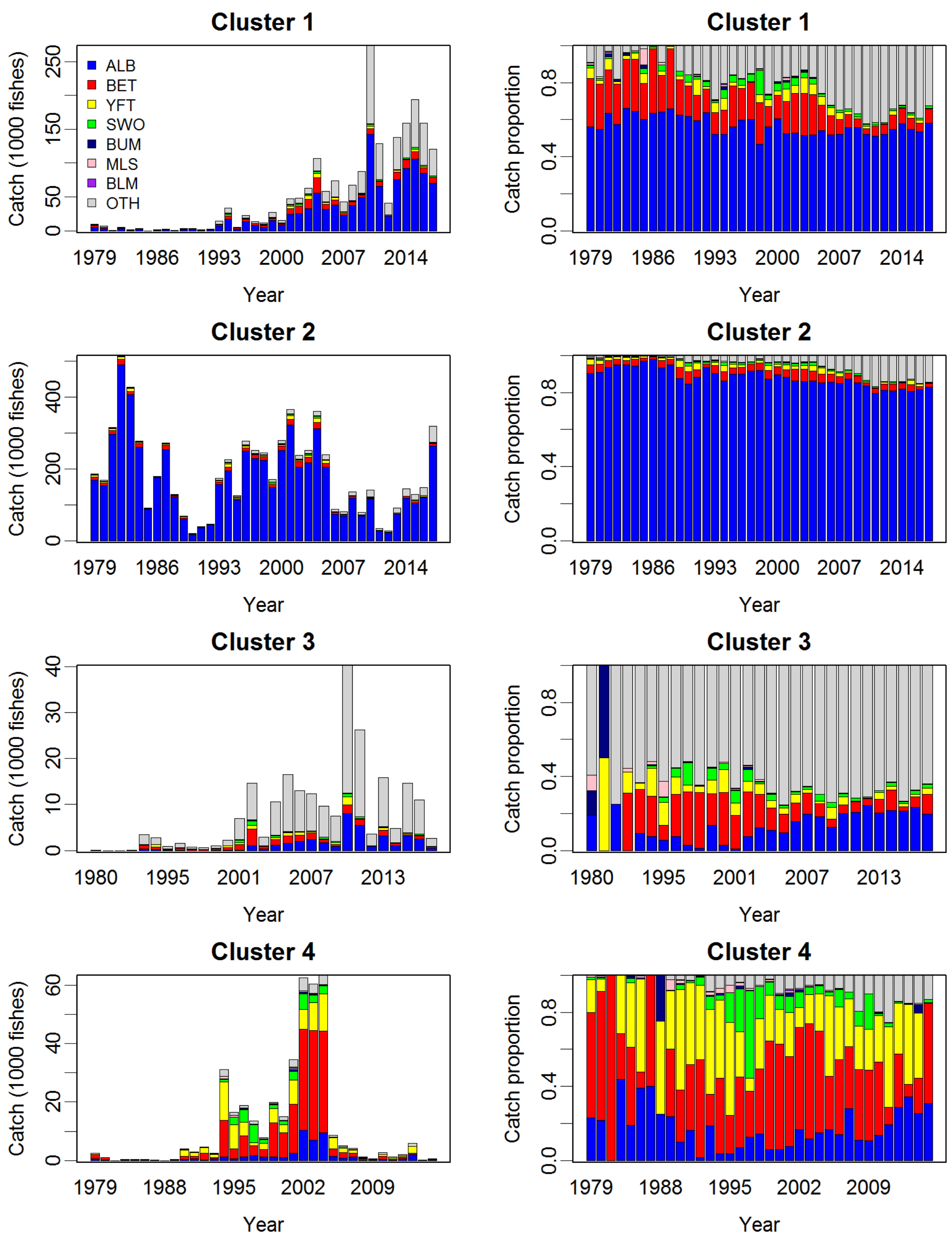


Fig. 20. Annual and catch proportion by species for each cluster of Taiwanese large scale longline fishery in Area SE of the Indian Ocean.

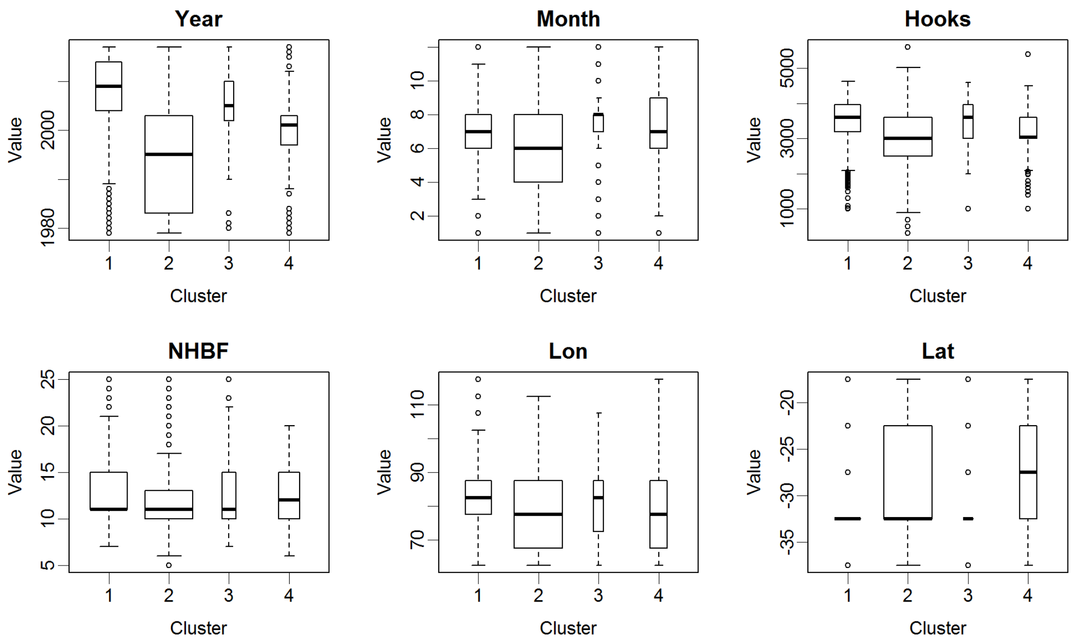


Fig. 21. Data composition by factors for each cluster of Taiwanese large scale longline fishery in Area SE of the Indian Ocean.

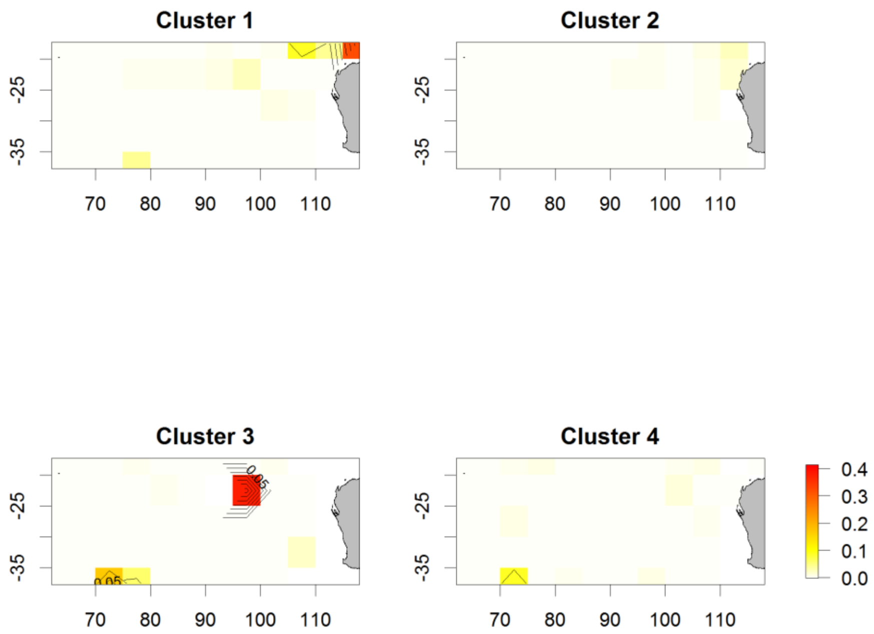


Fig. 22. Striped marlin catch distribution for each cluster of Taiwanese large scale longline fishery in Area SE of the Indian Ocean. Yellow is high catch and red is low catch.

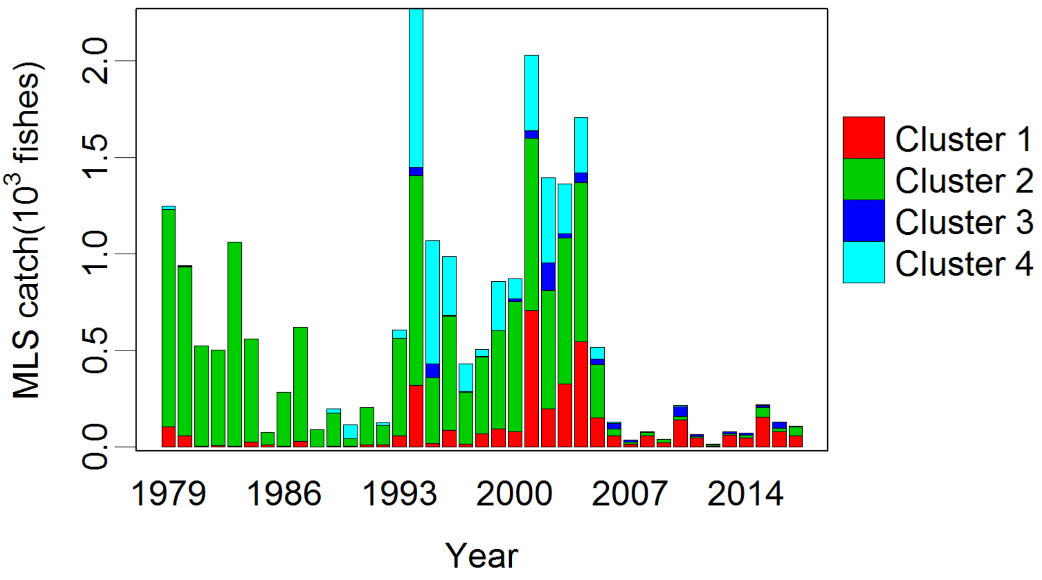


Fig. 23. Annual striped marlin catches for each cluster of Taiwanese large scale longline fishery in Area SE of the Indian Ocean.

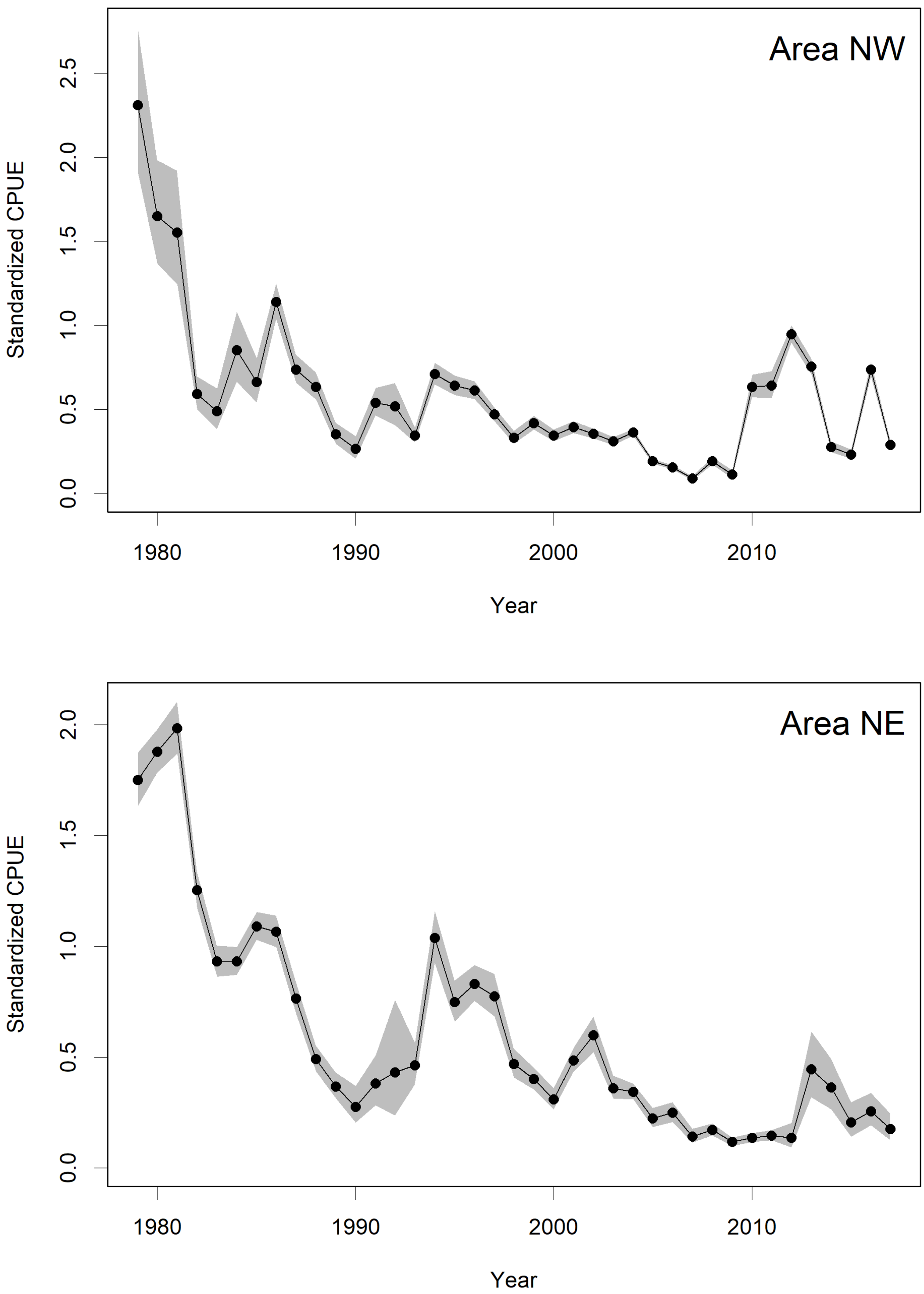
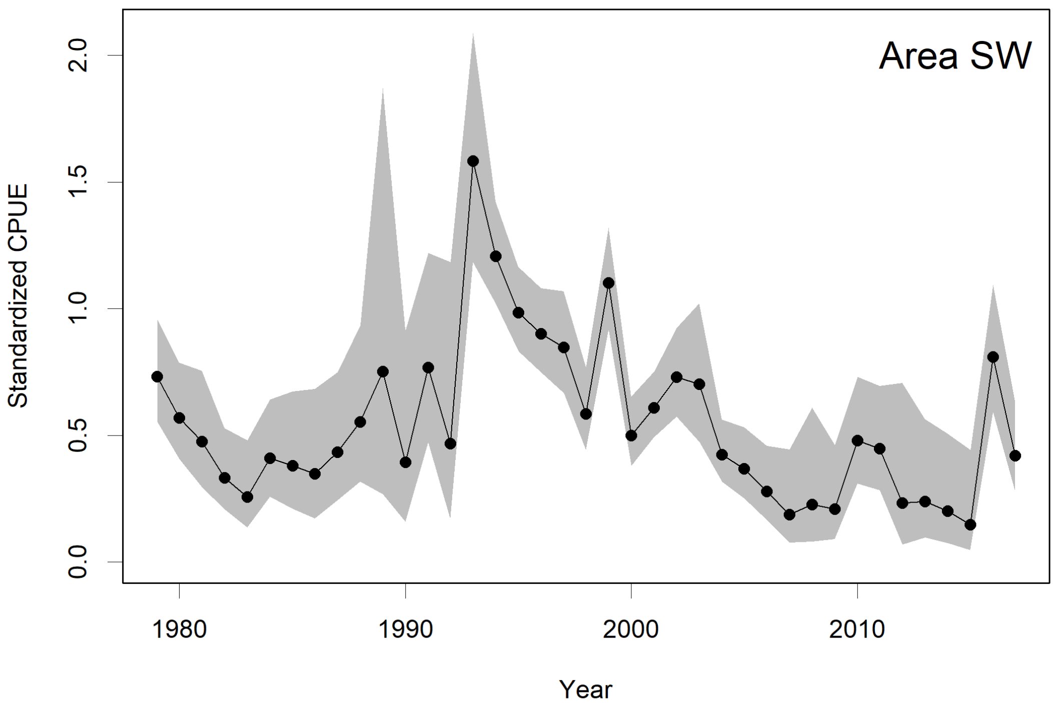


Fig. 24. The trajectory of area-specific standardized CPUE with 95% confidence interval for striped marlin caught by Taiwanese large scale longline fishery in the Indian Ocean.



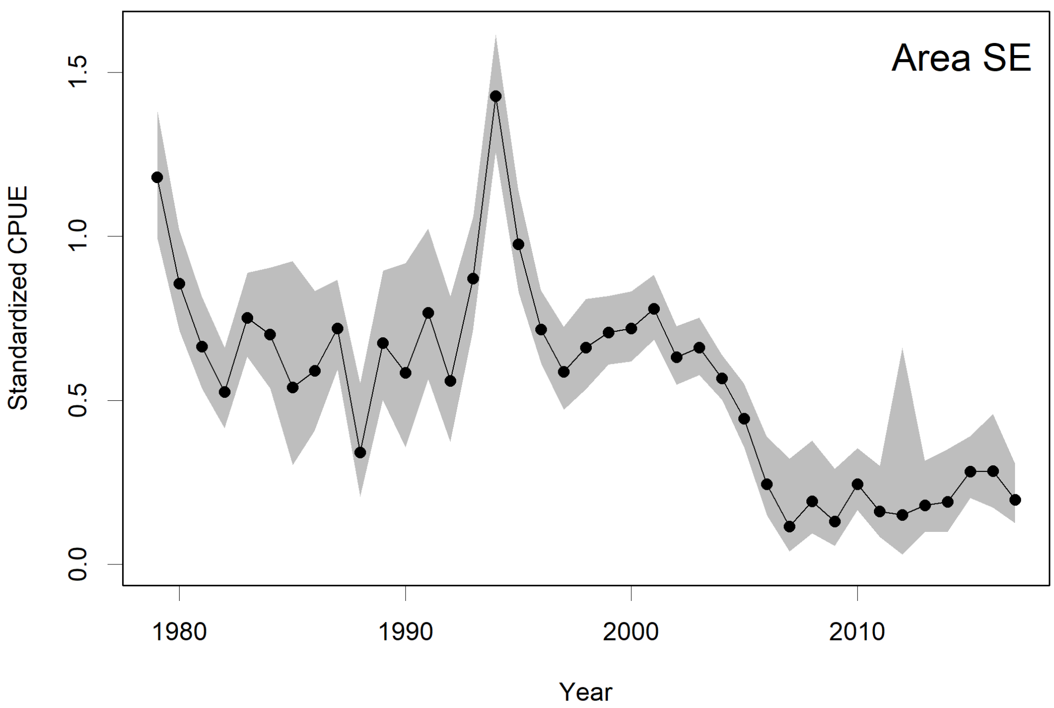


Fig. 24. (Continued).

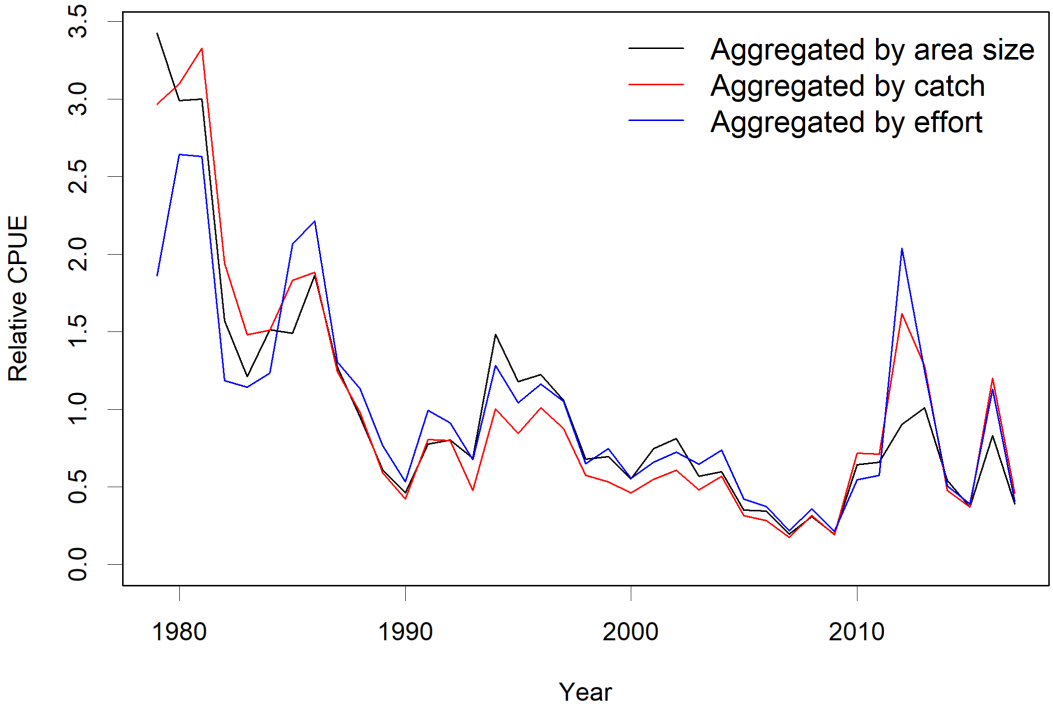
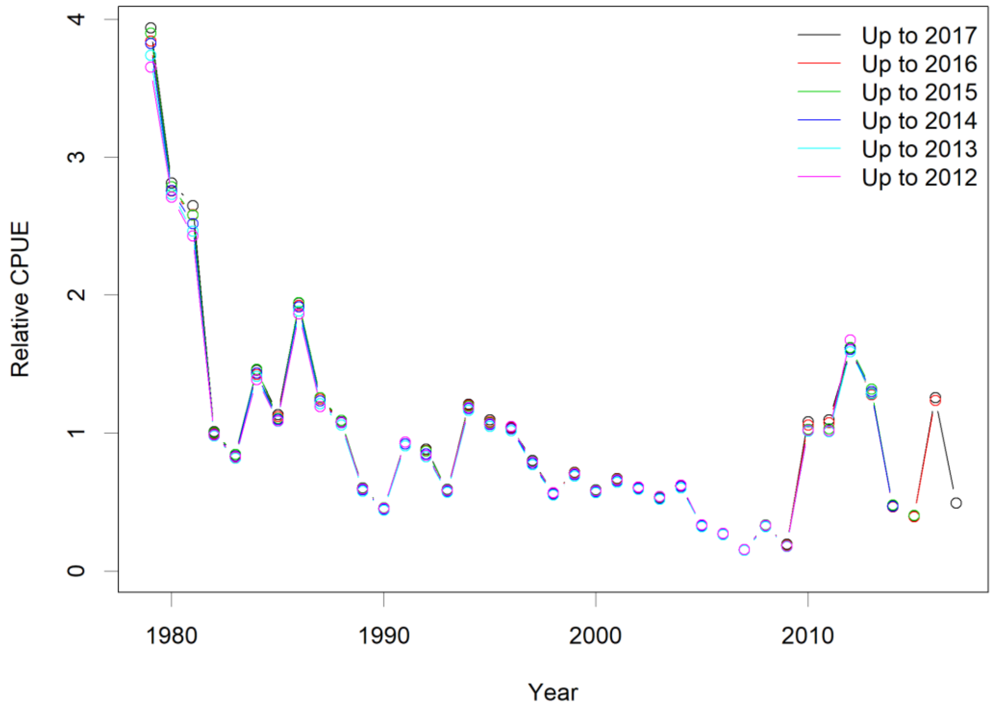


Fig. 25. The trajectory of area-aggregated standardized CPUE with 95% confidence interval for striped marlin caught by Taiwanese large scale longline fishery in the Indian Ocean.

Area NW



Area NE

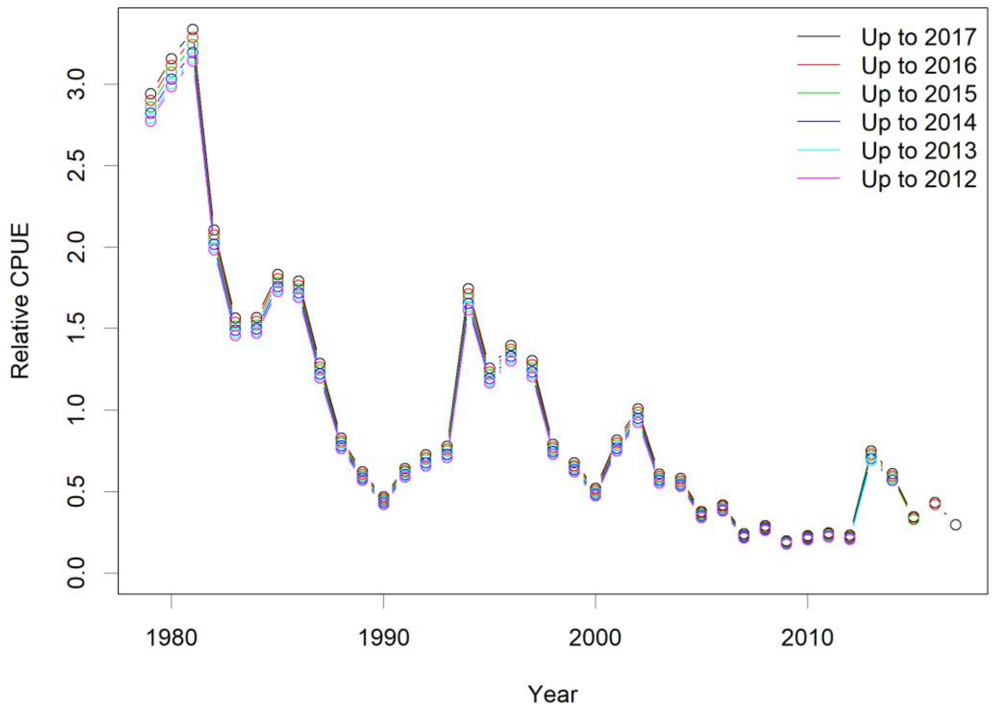
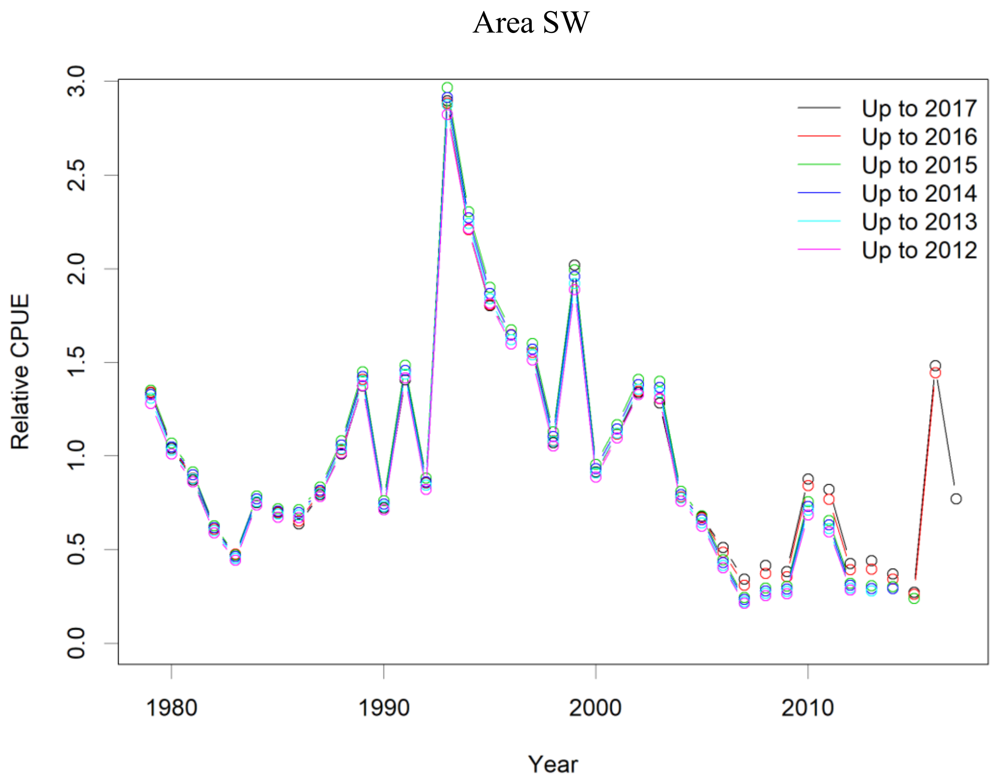


Fig. 26. CPUE standardization with retrospective analysis for striped marlin of Taiwan large scale longline fishery in the Indian Ocean.



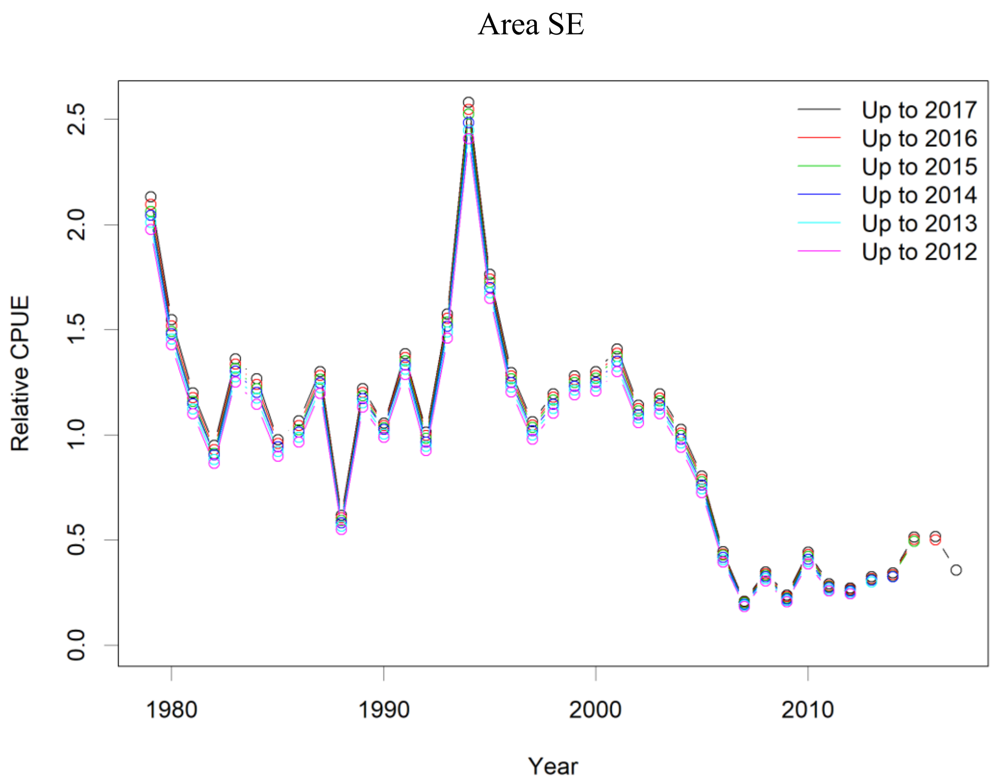


Fig. 26. (Continued).

Table 1. The ANOVA tables for selected gamma models.

Area NW

Variable	SS	Df	F	Pr(>F)
Y	6351	38	163.0936	< 2.2e-16 ***
M	515	11	45.6949	< 2.2e-16 ***
CT	23	3	7.3965	5.96E-05 ***
G	3378	45	73.2523	< 2.2e-16 ***
T	2119	3	689.179	< 2.2e-16 ***
Residuals	111189	108507		

Area NE

Variable	SS	Df	F	Pr(>F)
Y	6273	38	230.589	< 2.2e-16 ***
M	524	11	66.533	< 2.2e-16 ***
CT	91	3	42.606	< 2.2e-16 ***
G	1964	42	65.31	< 2.2e-16 ***
T	1293	3	602.036	< 2.2e-16 ***
Residuals	47000	65655		

Area SW

Variable	SS	Df	F	Pr(>F)
Y	407.4	38	9.3033	< 2.2e-16 ***
M	44.5	11	3.5115	6.27E-05 ***
CT	44.3	2	19.2034	4.75E-09 ***
G	556.5	29	16.6519	< 2.2e-16 ***
T	120.9	3	34.9863	< 2.2e-16 ***
Residuals	11314.7	9819		

Area SE

Variable	SS	Df	F	Pr(>F)
Y	664	38	32.0728	< 2.2e-16 ***
M	6.5	11	1.0899	0.3646
CT	16.8	3	10.2692	9.49E-07 ***
G	150.8	53	5.2222	< 2.2e-16 ***
T	132.2	3	80.8833	< 2.2e-16 ***
Residuals	7185	13189		

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

Table 2. The ANOVA tables for selected delta models.

Area NW

Variable	LR Chisq	Df	Pr(>Chisq)
Y	31935	38	< 2.2e-16 ***
M	2602	11	< 2.2e-16 ***
CT	270	4	< 2.2e-16 ***
G	9482	49	< 2.2e-16 ***
T	538	3	< 2.2e-16 ***

Area NE

Variable	LR Chisq	Df	Pr(>Chisq)
Y	27116.3	38	< 2.2e-16 ***
M	777.1	11	< 2.2e-16 ***
CT	401.6	4	< 2.2e-16 ***
G	6610.9	42	< 2.2e-16 ***
T	440.9	3	< 2.2e-16 ***

Area SW

Variable	LR Chisq	Df	Pr(>Chisq)
Y	2966.12	38	< 2.2e-16 ***
M	412.5	11	< 2.2e-16 ***
CT	19.18	3	0.0002507 ***
G	1084	32	< 2.2e-16 ***
T	464.35	3	< 2.2e-16 ***

Area SE

Variable	LR Chisq	Df	Pr(>Chisq)
Y	3123.11	38	< 2.2e-16 ***
M	778.3	11	< 2.2e-16 ***
CT	71.43	4	1.13E-14 ***
G	2599.26	54	< 2.2e-16 ***
T	119.73	3	< 2.2e-16 ***

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