CCSBT-ESC/2008/34
(ESC Agenda Item 7)

CPUE standardization for southern bluefin tuna caught by Taiwanese longline fishery for 2002-2019

Ching-Ping Lu[1], Sheng-Ping Wang[1], Shu-Ting Chang[2], Ming-Hui Hish[3]

[1] Department of Environmental Biology and Fisheries Science,
National Taiwan Ocean University, Taiwan
[2] Overseas Fisheries Development Council, Taiwan
[3] Fisheries Agency, Council of Agriculture, Executive Yuan, Taiwan

ABSTRACT

The CPUE standardization analyses were conducted with the data of Taiwanese longline fleets operated in the waters of the south of 20°S of the Indian Ocean from 2002 to 2019. The purpose of the cluster analysis was explored for the targeting of fishing operations and also to produce the data filter for selecting the data for the CPUE standardizations. The targeting of fishing operation could be identified by the cluster analyses with the weekly-aggregated data instead of set-by-set data. For CPUE standardizations, a simple delta-lognormal model without interactions were adopted to avoid the confounding from interactions. We conducted the cluster analyses for central-eastern area (Area E) and western area (Area W) separately. The pattern of the CPUE trends remained similar as the past but slightly decreased in both areas with updated data in 2019.

1. INTRODUCTION

Before the 1990s, southern bluefin tuna (*Thunnus maccoyii*; SBT) was the bycatch species of Taiwanese tuna longline fishery while targeting albacore in the Indian Ocean. However, there were some fishing vessels equipped with deep-frozen freezers started targeting SBT seasonally in the Indian Ocean since 1990s. In order to improve the quality of SBT fishing information, Taiwanese SBT statistics system was reformed in 2002. Therefore, the reporting rate of SBT catch had substantially improved since then (Anon, 2014). Here, we aim to explore the temporal and spatial patterns based on the catch and effort data of Taiwanese longline fishery operated in the waters of the south of 20°S of the Indian Ocean. And the purpose of this study is to perform the CPUE standardization for SBT caught by Taiwanese longline fishery for the period from 2002 to 2019.

2. MATERIALS AND METHODS

2.1. Catch and Effort data

The operational catch and effort data of Taiwanese longline fisheries from 2002 to 2019 were provided by the Overseas Fisheries Development Council (OFDC) of Taiwan. The resolution of the data, which were compiled from Taiwanese longline vessels by 5×5 degree fishing location grids.

Following the findings of the previous studies (Wang et al., 2015;2017;2018) suggested, the SBT fishing ground could be divided into the central-eastern area (Area E) and western area (Area W) separated by the boundary at the 60°E (Fig. 1). Here, all of the analyses in this study were conducted based on this area stratification with the boundary at the 60°E.

2.2. Cluster analysis

According to the approach of the previous study (Wang et al. 2015) and the suggestions by the experts of CCSBT ESC meetings in 2015 and 2016, we executed the cluster analysis (He et al., 1997) to explore the targeting of fishing operations and to produce the data filter for selecting the data for the further CPUE standardization. Cluster analysis was conducted using the major species composition of the catches. The major species included albacore (ALB), bigeye tuna (BET), yellowfin tuna (YFT), swordfish (SWO), southern bluefin tuna (SBT) and other species (OTH, the majority of the catches is composed by the oilfish) (Fig. 2 and Fig. 3). Based on the suggestion and consideration of 2016 CCSBT ESC, the clustering operational set-by-set data might contain large amount noise that because most of SBT caught by Taiwanese vessels was bycatches and only part of vessels targeted SBT for some fishing operations during the SBT fishing seasons. Additionally, ESC suggested that the cluster analysis could be conducted using the aggregated data rather than the operational set-by-set data. Therefore, we conducted the cluster analyses with both monthly and weekly aggregated data and then merged the clusters with operational data sets to identify the SBT fishing operations. While using the monthly-aggregated data for running the cluster analysis, the proportion of SBT catches decreased substantially and it created more difficulties for the identification of the cluster contained SBT fishing operations (Wang et al., 2017). Therefore, we performed the cluster analyses using weekly-aggregated data in this study.

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 (R Core Team (2019) with functions hclust and cutree. The number of clusters was strongly influenced by the subjective choice (He et al. 1997). Here, there were at least two clusters (SBT sets and other tuna sets) as it expected. There were 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. Here, we kept the SBT catch proportions of a specific cluster as large as possible and the proportion of data sets of the smallest cluster was larger than 5%.

2.3. CPUE standardization

Because there was large amount of zero SBT catch occurred in the fishing data sets, we applied the delta-lognormal models for the CPUE standardization of SBT caught by Taiwanese longline fishery. Based on the suggestions in the previous ESC, the main effects of year, month, 5x5 grid and number of hooks between floats (NHBF) included in both of lognormal and delta models. To avoid the confounding resulted from interactions, the interactions between main effects were not considered in the models. The effects of latitude and longitude were replaced by the effect of 5x5 grid. Additionally, the effects of cluster and NHBF were included because various catch compositions can be observed in a cluster (Wang et al., 2017). The models were conducted as below:

$$\begin{array}{ll} \text{lognormal model:} & \log(CPUE) \\ \text{delta model:} & PA \end{array} = \mu + Y + M + G + C + NHBF + \varepsilon$$

where	$CPUE$	is the nominal CPUE of SBT (catch in number/1,000 hooks) from data sets with positive SBT catch,
	PA	is the presence and absence of SBT catch,
	μ	is the intercept,
	Y	is the effect of year,
	M	is the effect of month,
	G	is the effect of 5x5 grid,
	C	is the effect of cluster,
	$NHBF$	is the effect of number of hooks between floats,
	ε	is the error term, $\varepsilon \sim N(0, \sigma^2)$.

The effects of year, month, and 5x5 grid were treated as categorical variables. The effect of NHBF was treated as three categories with various hooks including regular (≤9 hooks), deep (10-14 hooks), and ultra-deep (≥15 hooks) (Wang and Nishida, 2011).

The standardized CPUE trends were estimated with the exponentiations of the adjusted means (least square means) of the effect of year (Butterworth, 1996; Maunder and Punt, 2004). The model was selected based on the value of Akaike information criterion (AIC) and the estimations of the models were performed using R (R Core Team (2019) with functions glm and lsmeans.

The standardized CPUE was calculated by the product of the CPUE of positive catch and the probability of positive catches:

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

where $CPUE$ is the least square means of the effect of year from the lognormal model,

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

3. RESULTS AND DISCUSSIONS

3.1. Cluster analysis

Based on the area stratification with the boundary at the 60°E., we conducted the cluster analyses for the Area E and Area W separately. First, for the Area E, there are four clusters were selected (Fig. 4). Cluster 1 was composed mainly of ALB and BET operations, and the rest of operations with less proportion including YFT, SBT, SWO and OTH were parts of components in the Cluster 1. The major operations in Cluster 2 was the ALB operations, also included the operations for BET, SBT and OTH. The operations grouping in Cluster 3 mainly belonged to the ALB operation. Cluster 4 was mainly contributed by the SBT operation (Figs. 5 and 6). Although the highest SBT catch proportion was occurred in Cluster 4, most of the SBT catches were contained in Cluster 2 (Fig. 7). For SBT Cluster (Cluster 4), several fishing characteristics were described by the main effects as bellowed: (1) the data mainly consisted of the data in the early 2000s; (2) the majority of fishing operations were occurred during June and September; (3) the NHBF concentrated at around 10 hooks; and (4) the operations also concentrated in the waters between 30°S and 35°S (Fig. 8). The spatial distribution of SBT catch proportion was illustrated that the SBT catch proportion of Cluster 4 was obviously higher than the rest of others clusters (Fig. 9).

Second, three clusters were selected in the Area W (Fig. 10). The ALB operations was the majority in Cluster 1. The ALB operations also contributed mostly in Cluster 2 and contained the other operations such as BET, YFT, SWO and OTH. The OTH operations was belonged to Cluster 3 where mainly consisted of oilfish (Figs. 11 and 12). Most of SBT catches were found in Cluster 2 and Cluster 3 in recent years. (Fig. 13). For the fishing characteristics with various factors were described: (1) For the factor of year in Clusters 2 and 3, Cluster 2 mainly consisted of the data before 2010, while the data of Cluster 3 were mainly after 2010; (2) NHBF of Cluster 1 was lower than that of Cluster 2 and Cluster 3; (3) Fishing areas by longitude and latitude were different among three Clusters (Fig. 14). The SBT catch proportion of Cluster 3 was lower than the other two Clusters by illustrating the SBT catch proportion for the spatial distribution. (Fig. 15).

3.2 CPUE standardization

For both of Areas E and W, the final models were selected with the models with the lowest value of AIC. The results of ANOVA analysis for the lognormal models are shown in Table 1. All of the effects were statistically significant for both areas. About 21% and 36% of CPUE variances were explained by the models for Area E and Area W, respectively. The distributions of standardized residuals and the Quantile-Quantile Plots indicated that the distributions of residuals fitted to the assumption of the normal distribution (Fig. 16). For delta models, all of the main effects were also statistically significant for both areas (Table 2) and about 36% and 19% of CPUE variances were explained by the models for Area E and Area W, respectively.

Standardized CPUE series demonstrated quite different patterns in Area E and Area W (Fig. 17). First, for Area E, the standardized CPUE series gradually increased from 2004 to 2007, after that revealed decreasing trend from 2007 to 2011, substantially increased in 2012 and then gradually decreased until 2015, and then remained higher pattern in recent four years. For Area W, the standardized CPUE series generally revealed a decreasing trend with a fluctuation since 2002 and after 2013 stayed stable low pattern until now. The pattern of CPUE trends in both area E and W were not changed greatly.

3.3 Retrospect analysis

We performed the retrospect analysis to understand the impact of including the updated data on the CPUE standardization. The analysis was conducted by removing the data from 2019 to 2012. The results indicated that the influence of including the updated data on the CPUE standardization was negligible for Area E. On the other Area W, applying updated data into CPUE standardization changed the standardized CPUE series slightly, however, the pattern of the CPUE trends remained similar (Fig. 18).

REFERENCES

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.

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

R Core Team (2019). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL http://www.R-project.org/

Wang, S.P., Chang, S.T., Huang, A.C., Lin, S.L., 2017. CPUE standardization for southern bluefin tuna caught by Taiwanese longline fishery for 2002-2016. CCSBT-ESC/1708/33.

Wang, S.P., Chang, S.T., Lai, I.L. Lin, S.L., 2015. CPUE analysis for southern bluefin tuna caught by Taiwanese longline fleet. CCSBT-ESC/1509/23.

Wang, S.P., Lu, C.P., Chang, S.T., Huang, A.C., 2018. CPUE standardization for southern bluefin tuna caught by Taiwanese longline fishery for 2002-2017. CCSBT-ESC/1809/39.

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.

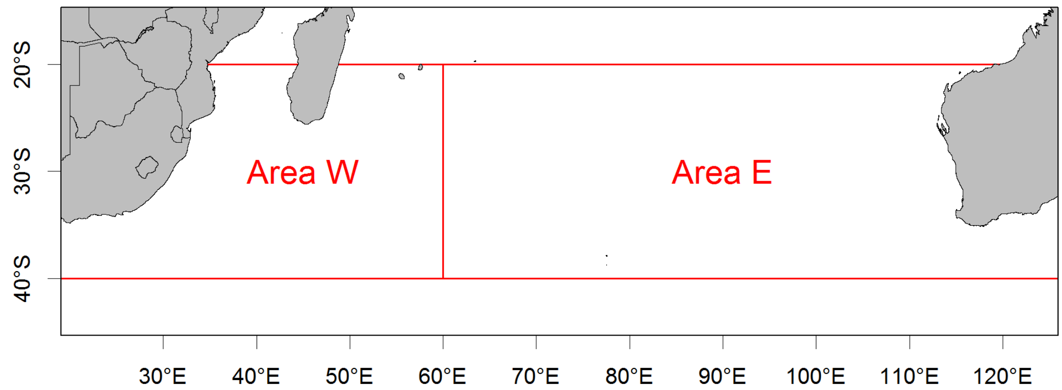


Fig. 1. Area stratification for southern bluefin tuna of Taiwanese large scale longline fishery in the Indian Ocean.

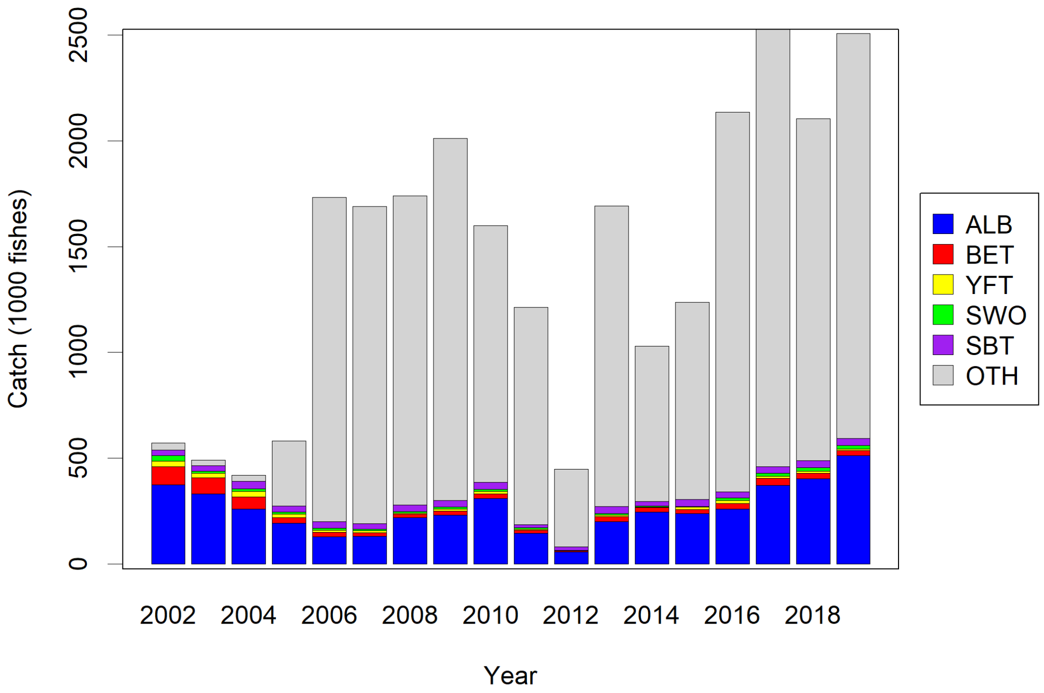


Fig. 2. Annual catch composition of the major species caught by Taiwanese longline fleets operated in the waters of south of 20°S from 2002 to 2019.

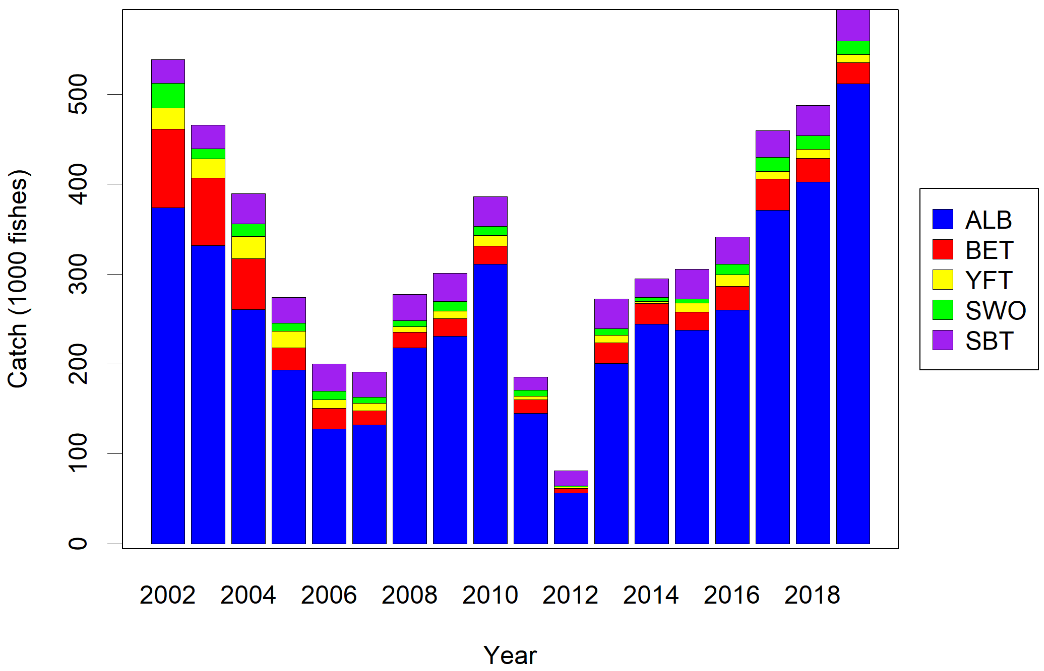


Fig. 3. Annual catch composition of the major species caught by Taiwanese longline fleets operated in the waters of south of 20°S from 2002 to 2019. The catches of OTH are excluded.

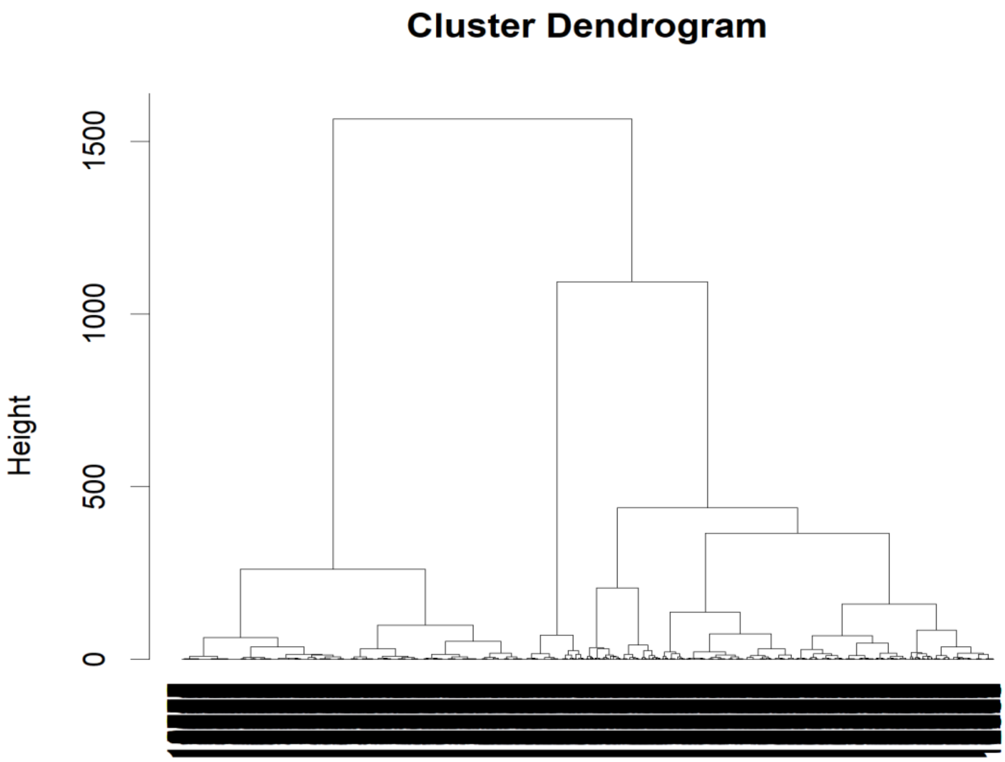


Fig. 4. The tree of cluster analysis using the data of Taiwanese large scale longline fishery in Southern Bluefin Tuna (SBT) Area E of the Indian Ocean.

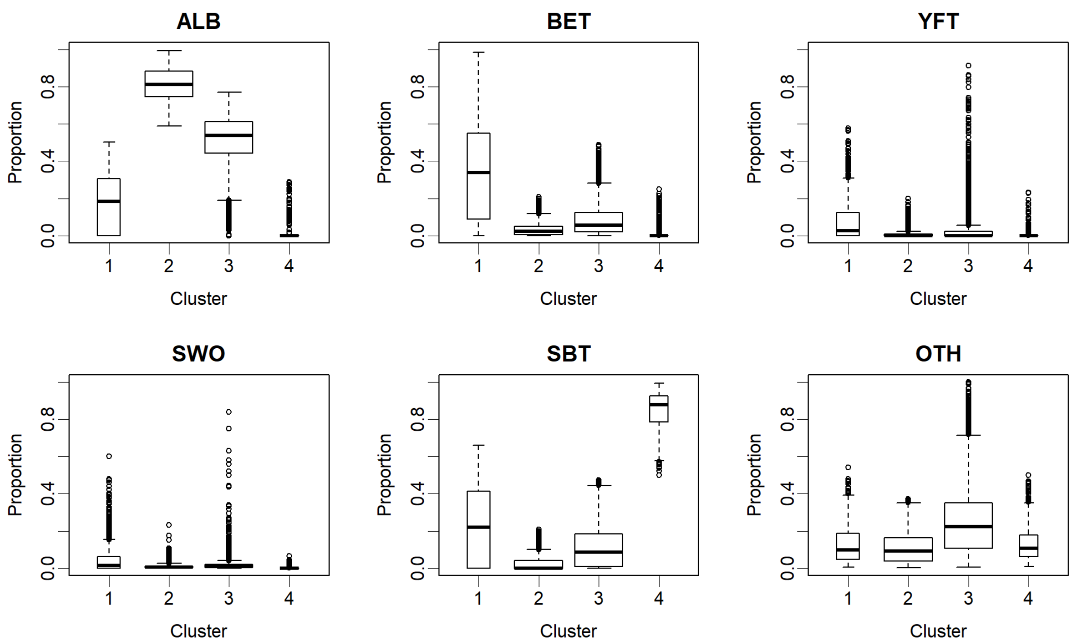


Fig. 5. Catch proportion by species for each cluster of Taiwanese large scale longline fishery in SBT Area E of the Indian Ocean.

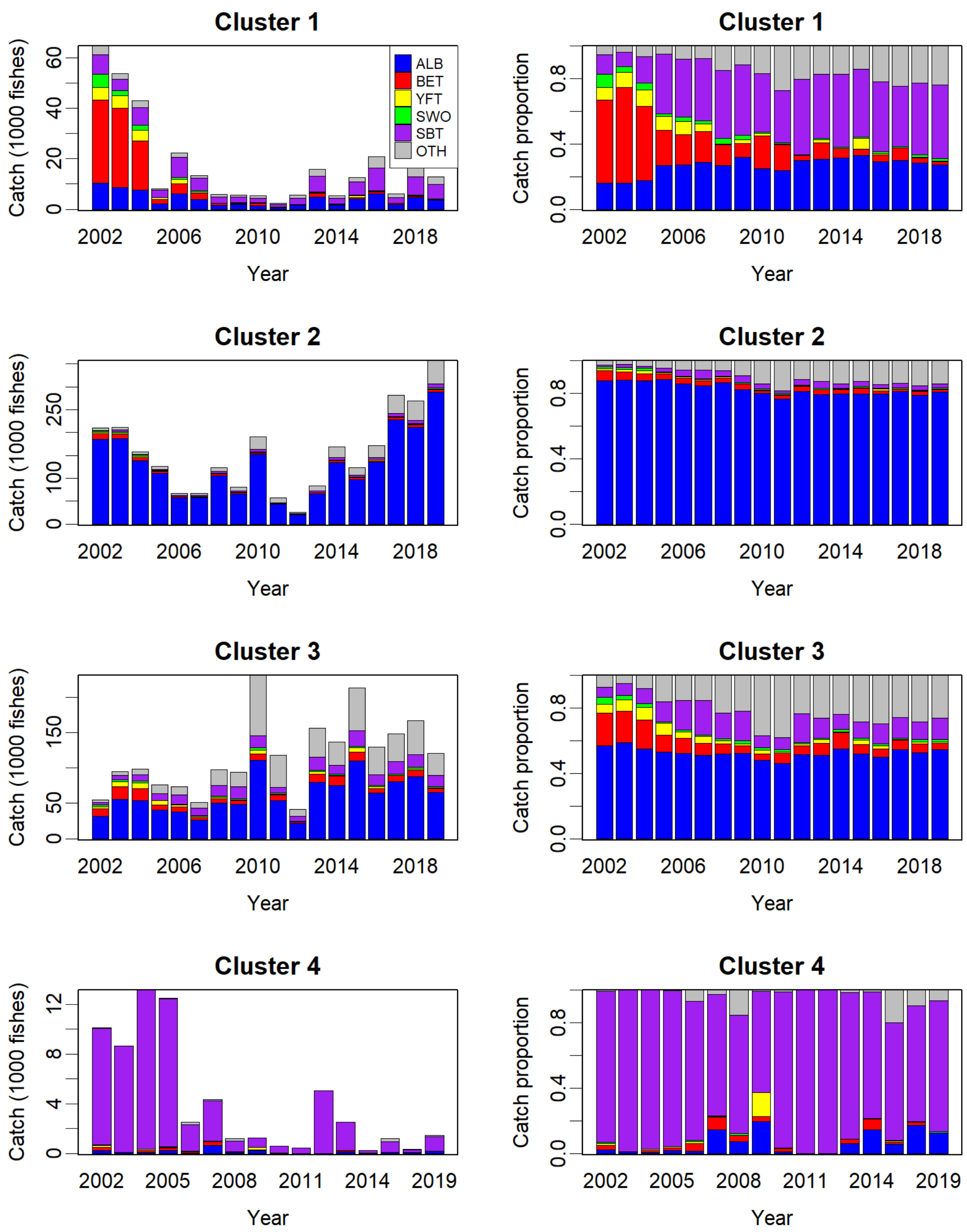
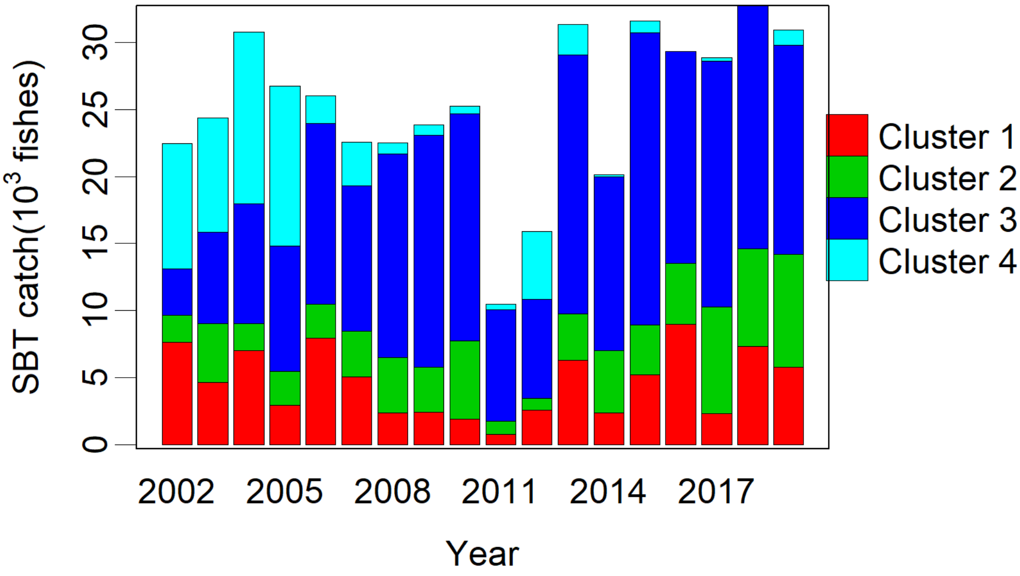


Fig. 6. Annual catch and catch proportion by species for each cluster of Taiwanese large scale longline fishery in SBT Area E of the Indian Ocean.



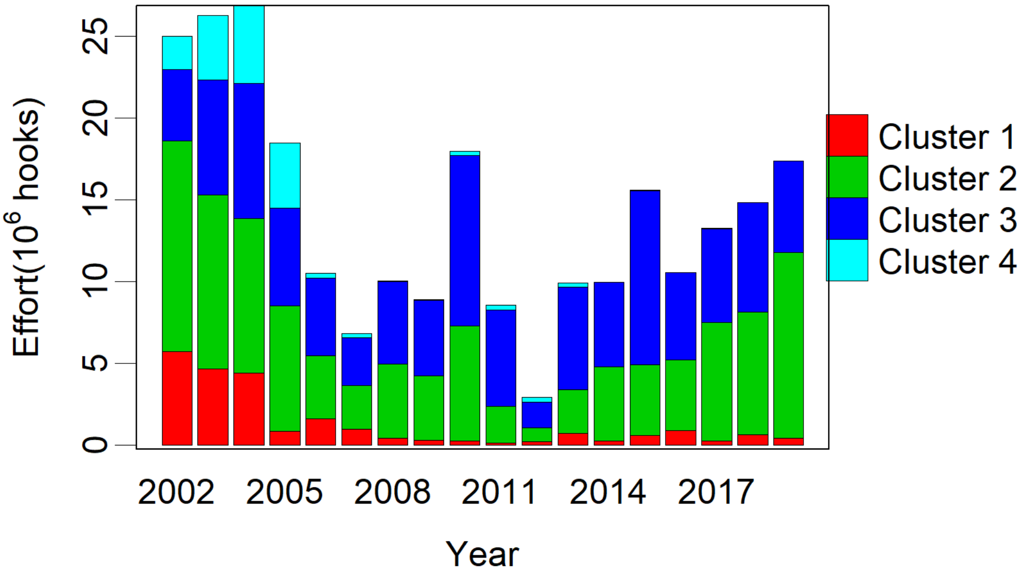


Fig. 7. Annual Southern Bluefin Tuna catches and efforts for each cluster of Taiwanese large scale longline fishery in Area E of the Indian Ocean.

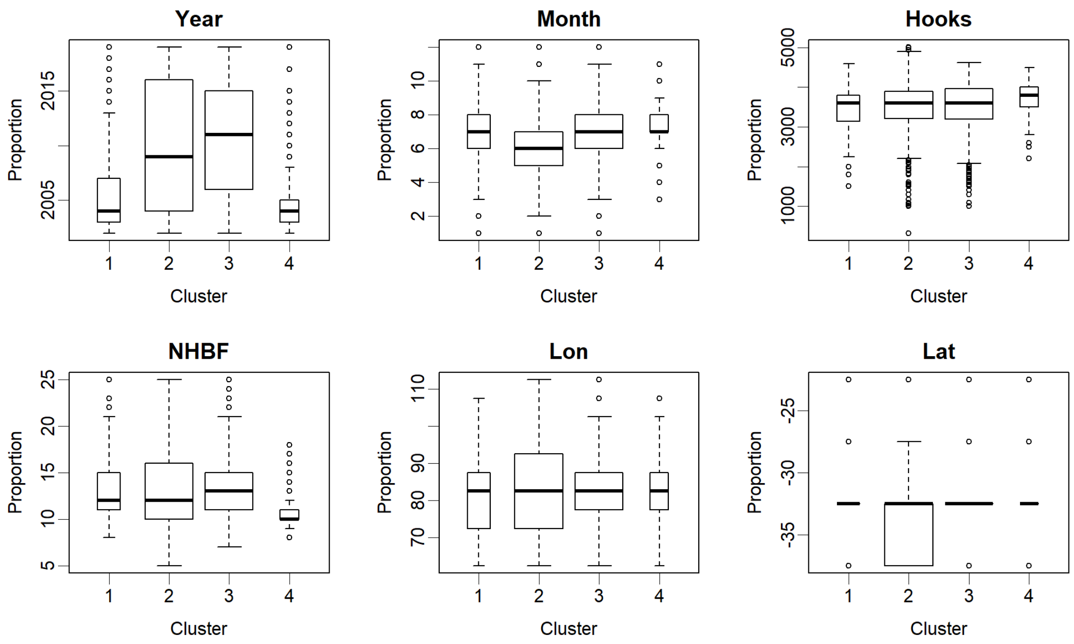


Fig. 8. Data composition by multiple factors for each cluster of Taiwanese large scale longline fishery in SBT Area E of the Indian Ocean.

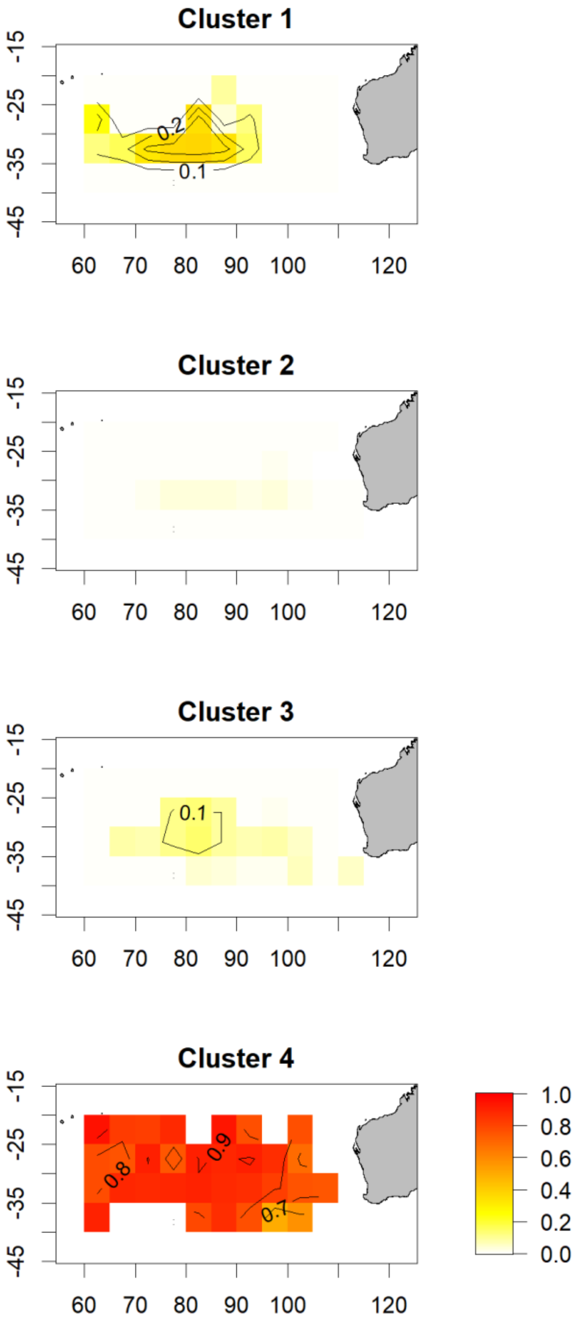


Fig. 9. Southern Bluefin Tuna catch distribution for each cluster of Taiwanese large scale longline fishery in Area E of the Indian Ocean. Red color represents high catch proportion and yellow color presents low catch proportion.

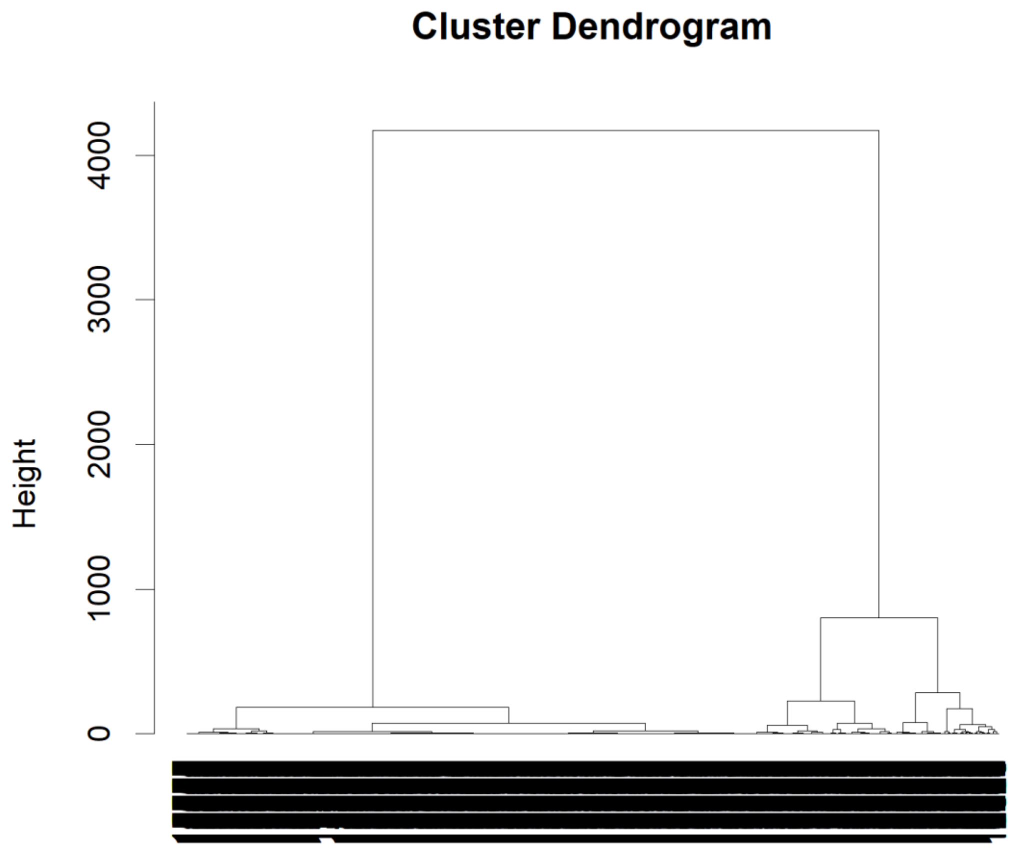


Fig. 10. The tree of cluster analysis using the data of Taiwanese large scale longline fishery in SBT Area W of the Indian Ocean.

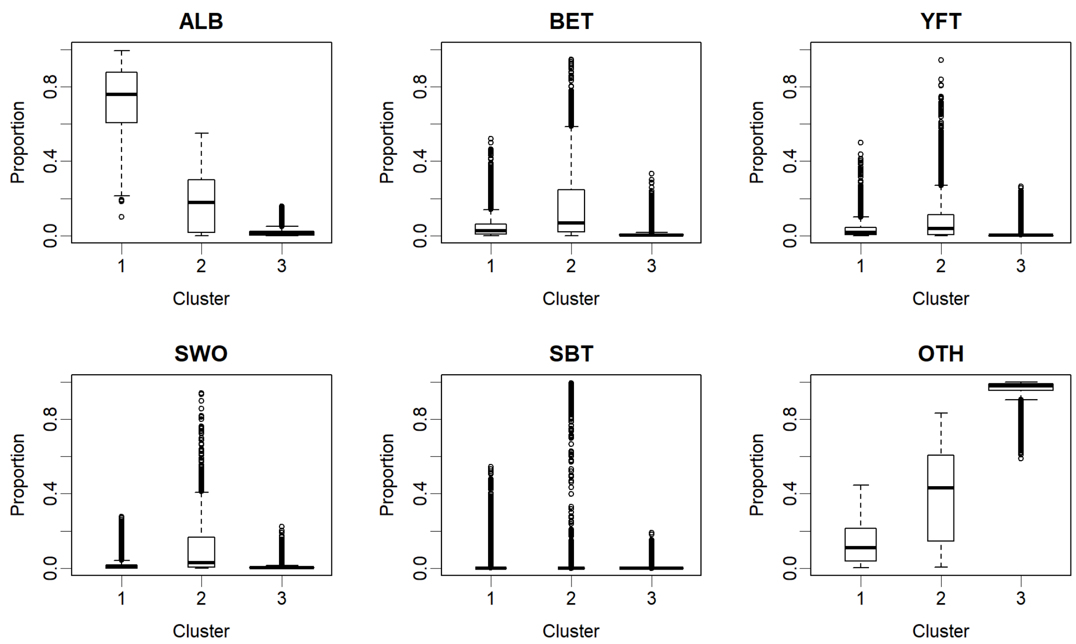


Fig. 11. Catch proportion by species for each cluster of Taiwanese large scale longline fishery in SBT Area W of the Indian Ocean.

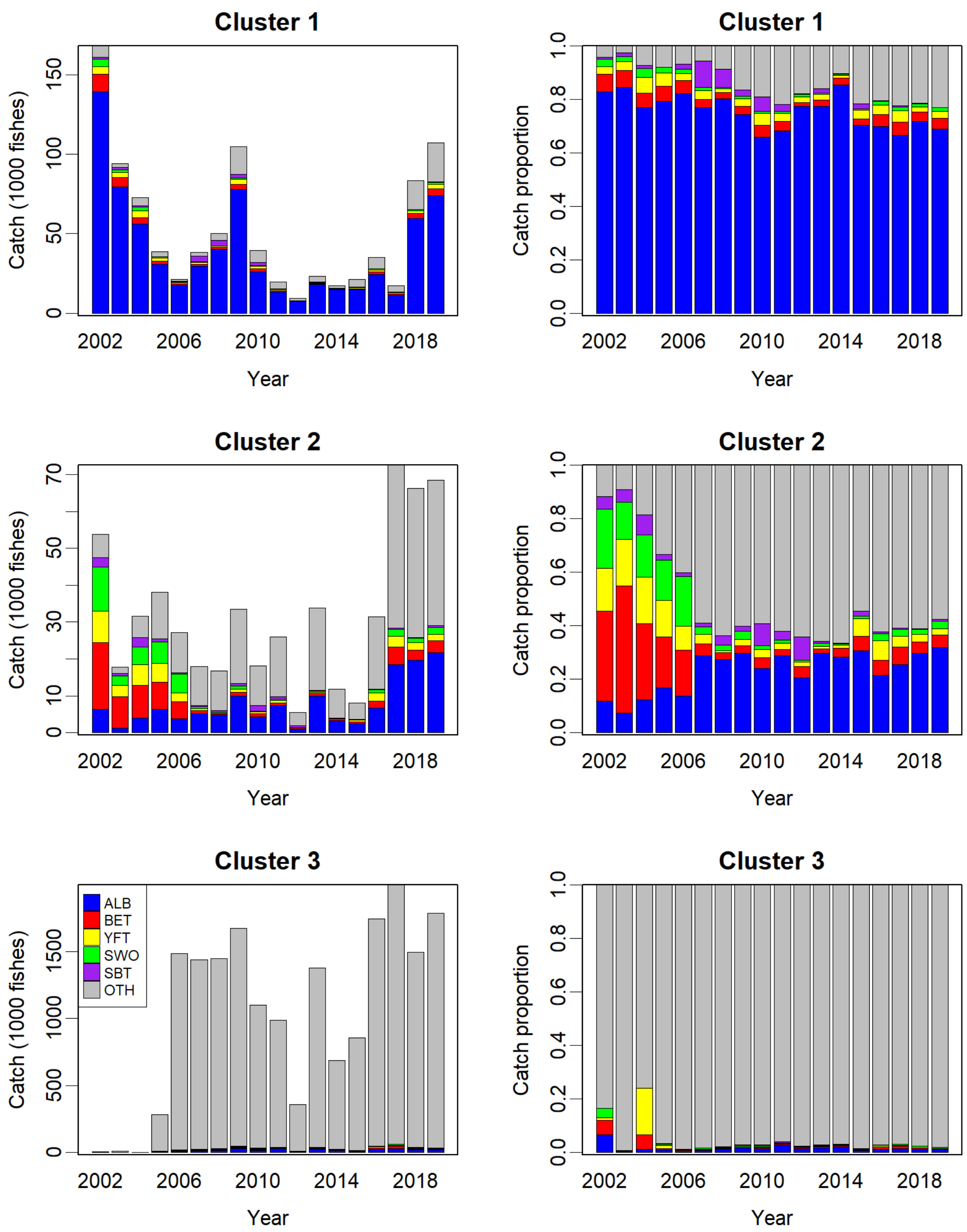
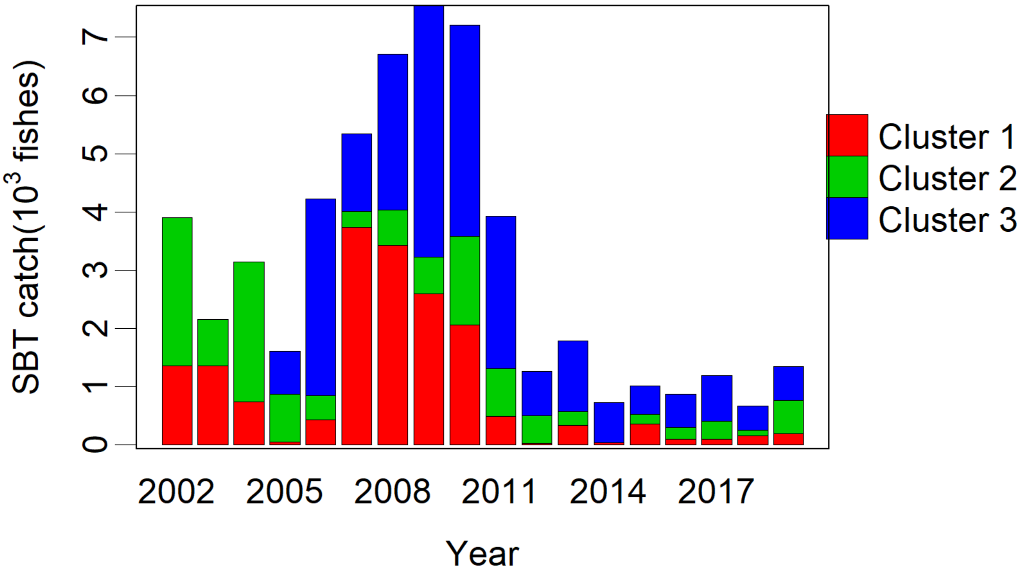


Fig. 12. Annual catch and catch proportion by species for each cluster of Taiwanese large scale longline fishery in SBT Area W of the Indian Ocean.



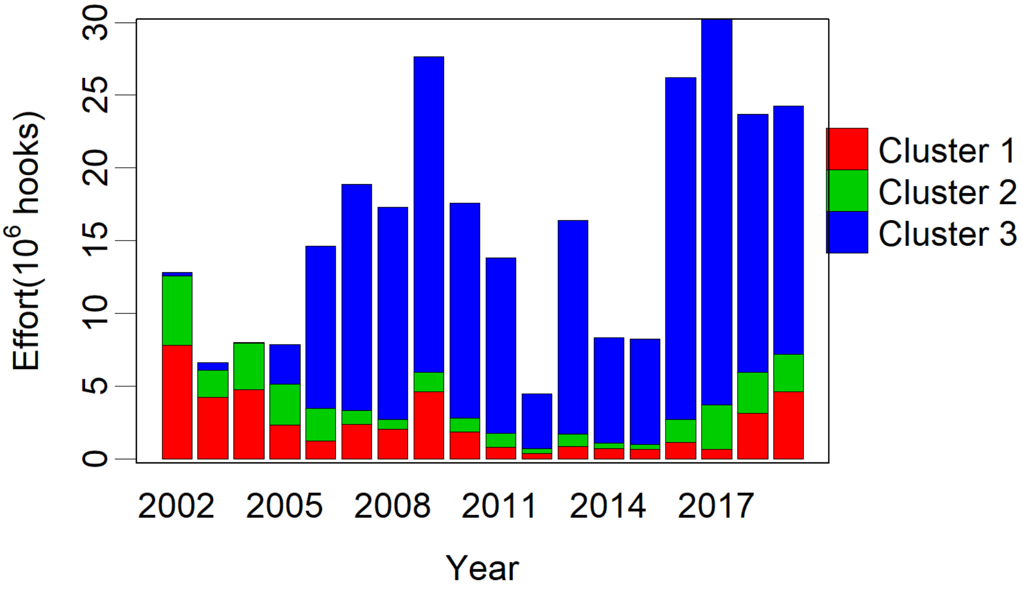


Fig. 13. Annual Southern Bluefin Tuna catches and efforts for each cluster of Taiwanese large scale longline fishery in Area W of the Indian Ocean.

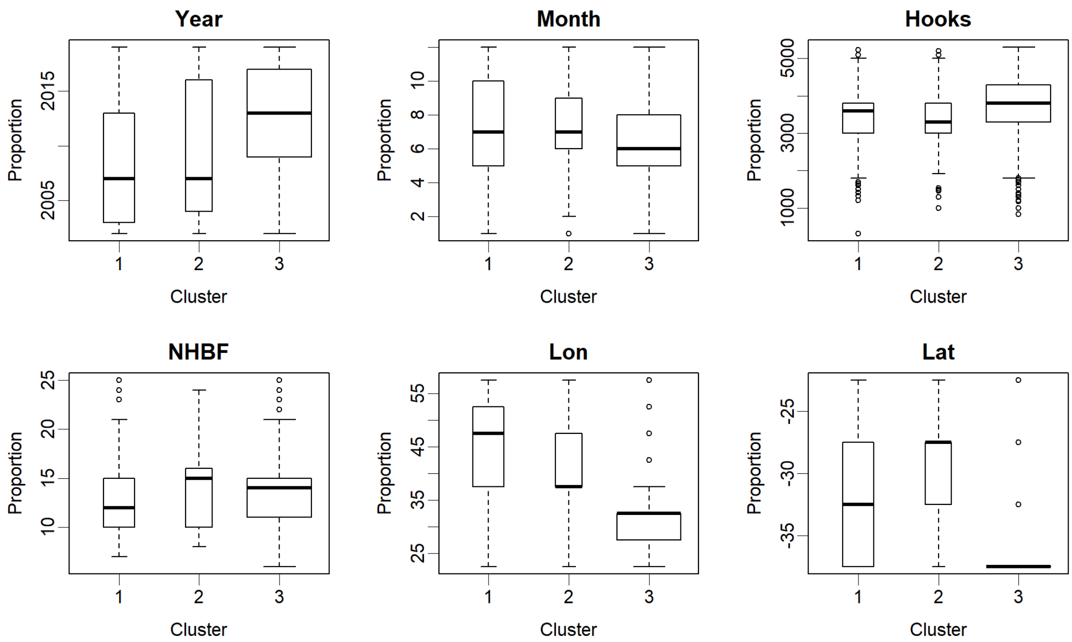


Fig. 14. Data composition by multiple factors for each cluster of Taiwanese large scale longline fishery in Southern Bluefin Tuna Area W of the Indian Ocean.

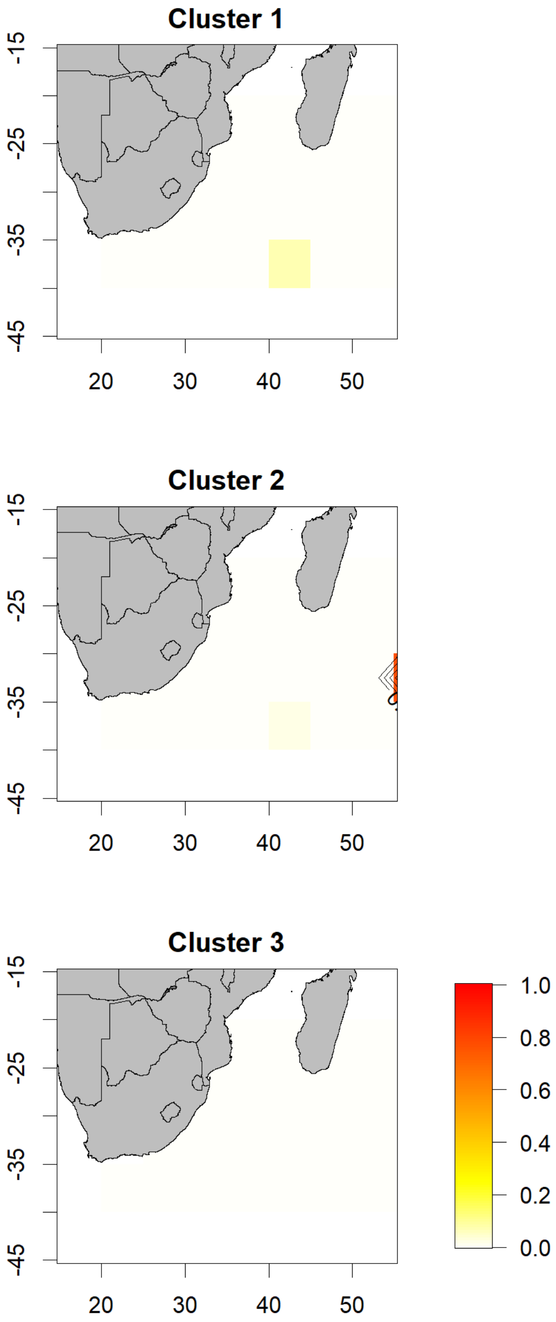
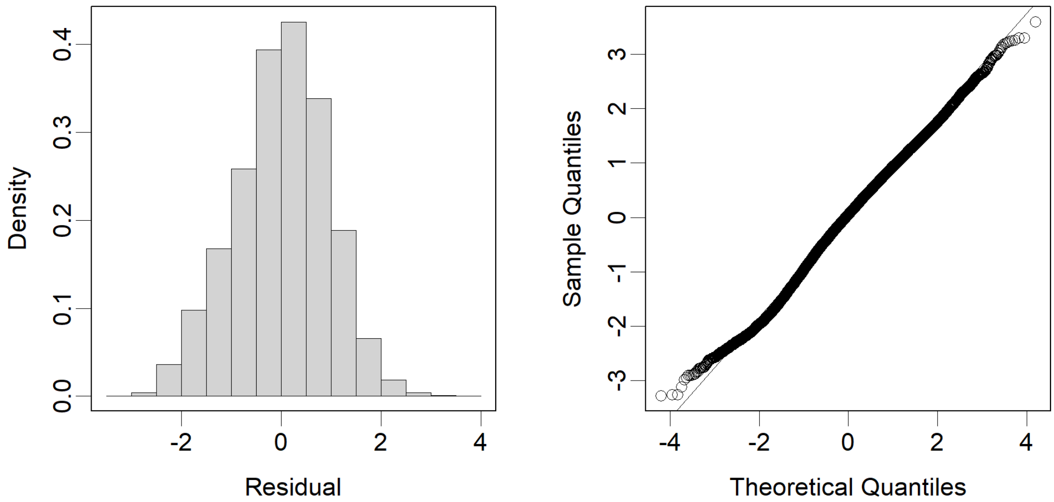


Fig. 15. Southern Bluefin Tuna catch distribution for each cluster of Taiwanese large scale longline fishery in Area W of the Indian Ocean. Red color represents high catch proportion and yellow color presents low catch proportion.

Area E



Area W

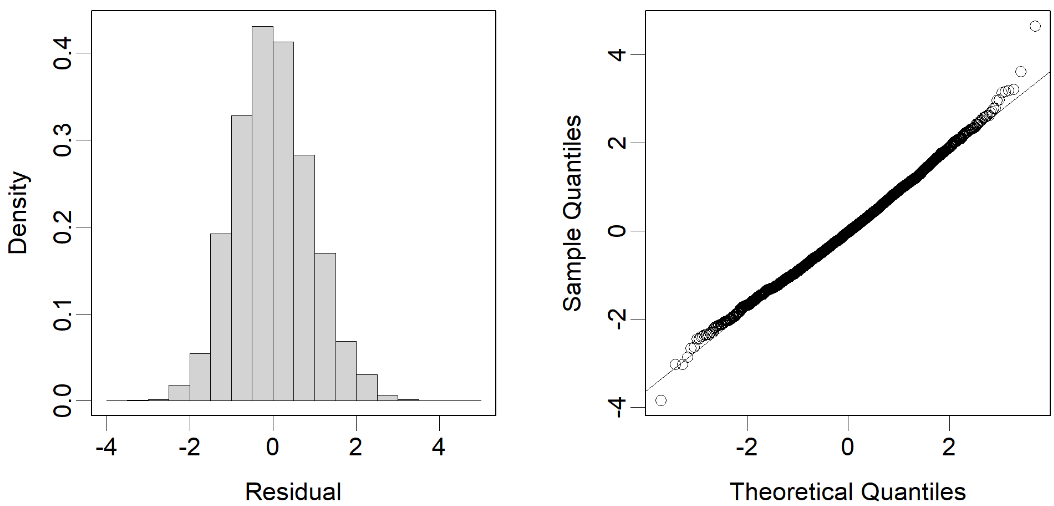


Fig. 16. The frequency distributions and Quantile-Quantile Plots for standardized residuals obtained from lognormal models for Area E and Area W.

Area E

Area W

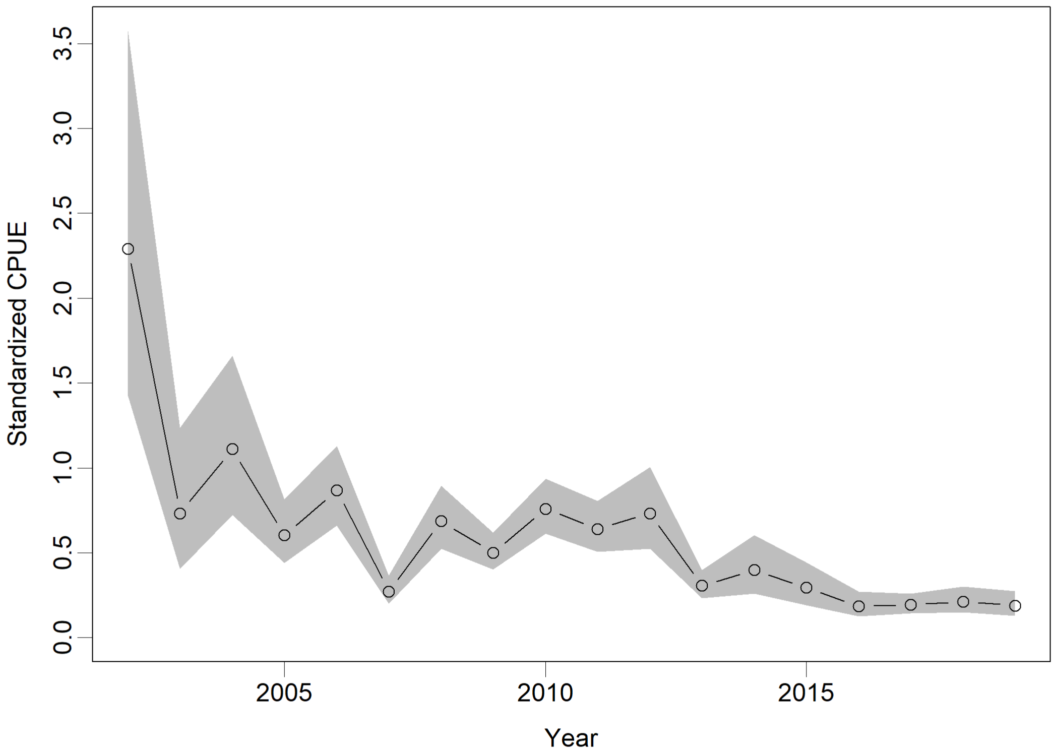


Fig. 14. Area-specific standardized CPUE of SBT caught by Taiwanese longline fishery. Shaded areas show the 95% confidence intervals.

Area E

Area W

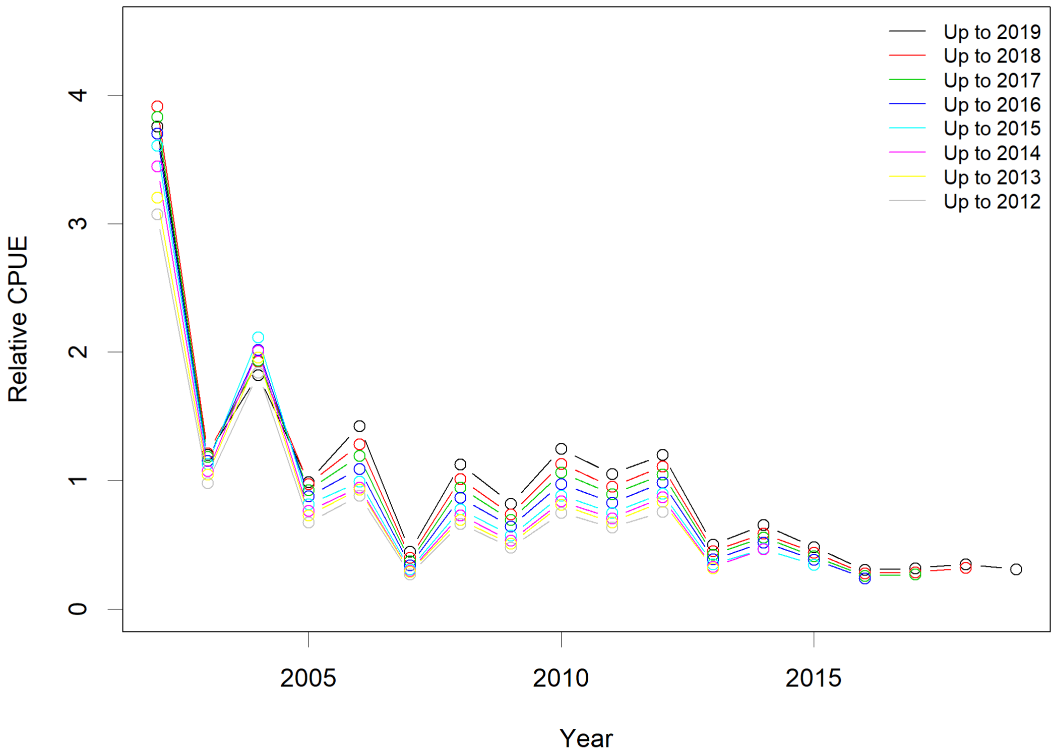


Fig. 18. The results of CPUE standardization based on including the updated data from different years.

Table 1. The results of ANOVA for the lognormal models for Area E and Area W.

Area E

Source of variance	SS	Df	F	Pr(>F)
Y	2334	17	154.891	< 2.2e-16 ***
M	498	9	62.441	< 2.2e-16 ***
G	782	37	23.858	< 2.2e-16 ***
C	2332	3	877.003	< 2.2e-16 ***
NHBF	43	2	24.527	2.265e-11 ***
Residuals	34174	38555		

Significant level: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Area W

Source of variance	SS	Df	F	Pr(>F)
Y	606.8	17	41.969	< 2.2e-16 ***
M	379.3	10	44.601	< 2.2e-16 ***
G	104.6	22	5.589	8.995e-16 ***
C	14.3	1	16.798	4.230e-05 ***
NHBF	38.9	2	22.873	1.306e-10 ***
Residuals	3865.3	4545		

Significant level: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Table 2. The results of ANOVA for the delta models for Area E and Area W.

Area E

Source of variance	LR Chisq	Df	Pr(>Chisq)
Y	6601.1	17	< 2.2e-16 ***
M	4808.5	11	< 2.2e-16 ***
G	5956.1	41	< 2.2e-16 ***
C	6829.0	3	< 2.2e-16 ***
NHBF	192.3	2	< 2.2e-16 ***

Significant level: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Area W

Source of variance	LR Chisq	Df	Pr(>Chisq)
Y	684.3	17	< 2.2e-16 ***
M	2989.9	11	< 2.2e-16 ***
G	2331.9	27	< 2.2e-16 ***
C	140.3	1	< 2.2e-16 ***
NHBF	16.9	2	0.0002104 ***

Significant level: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1