CPUE standardization analyses for southern bluefin tuna based on the Taiwanese longline fishery data from 2002 to 2022

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 statistical information of Taiwanese longline fleets from 2002 to 2022. The operating waters of the large scale Taiwanese longliner were mainly the south of 20°S of the Indian Ocean. While conducting the CPUE standardization, for the first step, the cluster analysis was processed to explore the targeting of fishing operations and to produce the data filter for selecting the data for the CPUE standardizations. In order to clarify multiple styles of the targeting of fishing operations, the cluster analyses were conducted with the weekly-aggregated data instead of set-by-set data. Second, the simple delta-lognormal model without interactions was adopted to avoid the confounding from interactions for the CPUE standardizations analyses. The cluster analyses were applied for main fishing area (central-eastern area: Area E) and secondary fishing area (western area: Area W) separately. The pattern of the CPUE trends in both fishing areas (Area E and Area W) remained similar as the past. With updated data in 2022, the CPUE trend of Area E increased and the CPUE trend of Area W decreased slightly.

1. INTRODUCTION

Before the 1990s, Taiwanese tuna longline fishery targeting albacore also explore Southern bluefin tuna (*Thunnus maccoyii*; SBT) as the bycatch species in the Indian Ocean. Since the 1990s, some Taiwanese fishing vessels equipped with deep-frozen freezers started to the seasonal targeting SBT fishery in the Indian Ocean. In order to improve the quality of fishing information, Taiwanese SBT statistics system was reformed for the collection of the related SBT fishing information in 2002. And the reporting rate of SBT catch data had also substantially improved since then (Anon, 2014). It would be helpful for better understanding of the SBT fishery indicators with the improved Taiwanese SBT statistics fishing information. Therefore, the main goal of this study was to explore the variation of temporal and spatial patterns of SBT using the major species of catch and effort data caught by Taiwanese longline fishery operated in the waters of the south of 20°S of the Indian Ocean. Here, the temporal pattern and analyses of the CPUE standardization for SBT caught by Taiwanese longline fishery were conducted with the updated information from 2002 to 2022.

2. MATERIALS AND METHODS

2.1. Catch and Effort data

In this study, we applied the operational catch and effort data of Taiwanese longline fisheries from 2002 to 2022 provided by the Overseas Fisheries Development Council (OFDC) of Taiwan for the CPUE standardization. The resolution of the dataset was compiled by 5×5 degree fishing location grids from Taiwanese longline vessels fishing information.

Following the suggestions of the previous studies (Wang et al., 2015;2017;2018), the SBT fishing ground of Taiwanese longline fisheries was divided into the central-eastern area (Area E) and western area (Area W) by the boundary at the 60°E (Fig. 1). And, it is suitable for the CPUE standardization analyses of the SBT caught by Taiwanese longline fishery. 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

Based on the methodology of the previous study (Wang et al. 2015) and the suggestions by the experts of CCSBT ESC meetings in 2015 and 2016, we carried out 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 of the central-eastern area (Area E) and western area (Area W). Cluster analyses were conducted with the major species composition of the catches in both areas. 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). According to the recommendation of CCSBT ESC meeting in 2016, 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. Therefore, ESC suggested that the cluster analysis could be processed using the aggregated data rather than the operational set-by-set data. With all the suggestions and considerations, 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 applying the cluster analysis with the monthly-aggregated data, the proportion of SBT catches decreased substantially and the difficulties were increased to identify 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's minimum variance method was performed to the squared Euclidean distances calculated from the aggregated data sets. The analyses were conducted 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) were included in both of lognormal and delta models. The interactions between main effects were not considered in the models to avoid the confounding resulted from interactions. 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:

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

where

$\tilde{C}PUE$	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

Cluster analysis was performed using the major species composition of the catches. The annually composition of the catches of six major species from 2002 to 2022 were shown in Fig. 2 and Fig. 3. The ALB catch was the majority of the catch of Taiwanese large scale longline fisheries excluded other species(OTH). The cluster analyses were conducted for the Area E and Area W separately with the boundary at the 60°E of Taiwanese SBT fishing area. For the first section, Area E, there are four clusters were selected (Fig. 4). According to the catch proportion by species of each cluster, ALB, BET and SBT operations were the main components in the Cluster 1, and the rest of operations with less proportion of YFT, SWO and OTH were composed in this Cluster. The ALB operations took the majority of this Cluster 2 and the catch proportion of ALB were over 0.8 each year, also included relatively fewer operations for BET, SBT and OTH. The operations grouping in Cluster 3 showed the similar composition with Cluster 2, the larger part of operations belonged to the ALB operation, also included relatively fewer operations for BET, SBT and OTH. And in the Cluster 4, it was mainly and clearly contributed by the SBT operations (Figs. 5 and 6). Although the highest

SBT catch proportion was occurred in Cluster 4, most of the SBT catches were contained in Cluster 1 and Cluster 3 (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 concentrated and 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).

For the second section, the Area W, the cluster analysis was conducted and three clusters were grouped in this Area (Fig. 10). According to the result, the ALB operations was found as the majority parts of composition in Cluster 1. The ALB operations also contributed mostly in Cluster 2 and contained the other operations such as BET, YFT, SBT, SWO and OTH. The OTH operations was belonged to Cluster 3 where mostly composed by oilfish (Figs. 11 and 12). In recent years, most of SBT catches were found in both Cluster 2 and Cluster 3. (Fig. 13). For the fishing characteristics with various factors were described: (1) For the factor of year, Cluster 1 and 2 mainly consisted of the data period between 2005 and 2020, while the majority data of Cluster 3 were after 2010; (2) NHBF of Cluster 1 was distributed wider than that of Cluster 2 and Cluster 3; And the NHBF of Cluster 3 was more concentrated than the others; (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 23% and 34% 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 35% and 18% of CPUE variances were explained by the models for Area E and Area W, respectively.

Standardized CPUE series of Area E and Area W presented quite different patterns (Fig. 17). In the Area E, the standardized CPUE series gradually showed the increasing trend from 2004 to 2007, after that revealed decreasing trend from 2007 to 2011, substantially increased in 2012. After that the trend gradually decreased until 2015, and then remained higher and relatively stable increasing pattern in recent two years with the updated data to 2022. And 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 2019, after that it presented increasing trend. And with the updated data of 2022, it revealed slightly decreased and back to the similar level of 2020. The pattern of CPUE trends in both area E and W were relatively stable increasing trend since 2019.

3.3 Retrospect analysis

The retrospect analysis was performed to understand the influence while using the updated data to 2022 on the CPUE standardization. The analysis was performed by removing the data from 2022 to 2012. The results indicated that the influence of adding the updated data on the CPUE standardization was negligible for both Area E and Area W. Although the index values of standardized CPUE series slightly changed while applying updated data into CPUE standardization, however, the pattern of the CPUE trends remained high similarity of variation patterns (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.

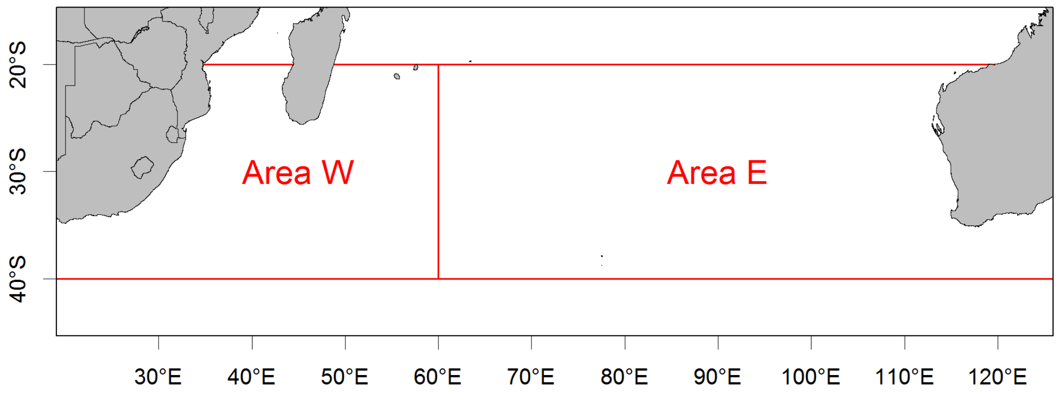


Fig. 1. Fishing area stratification for southern bluefin tuna of Taiwanese large scale longline fishery in the Indian Ocean.

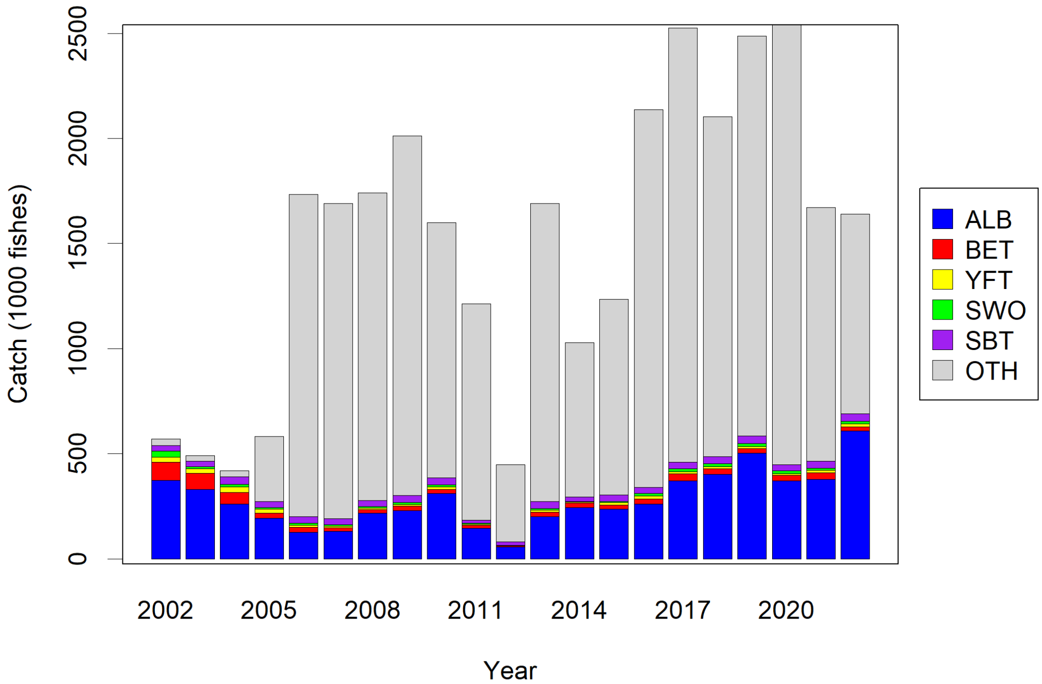


Fig. 2. Annual catch composition of the major species caught by Taiwanese longline fleets operated in the waters of south of 20°S in the Indian Ocean from 2002 to 2022.

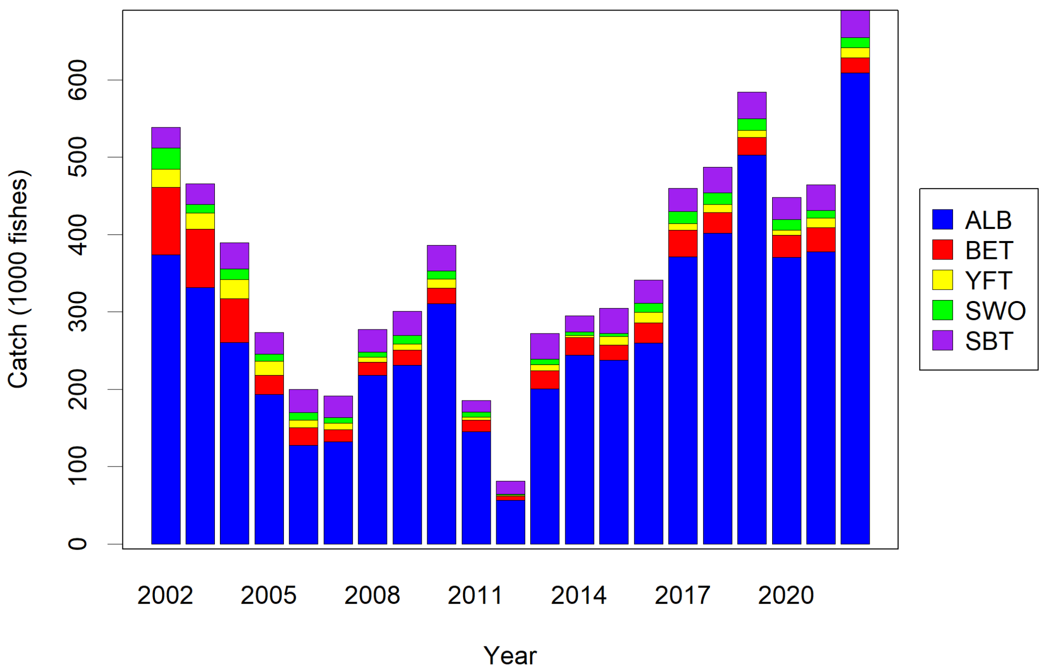


Fig. 3. Annual catch composition of the major species caught by Taiwanese longline fleets operated in the waters of south of 20°S in the Indian Ocean from 2002 to 2022. The catches of OTH are excluded.

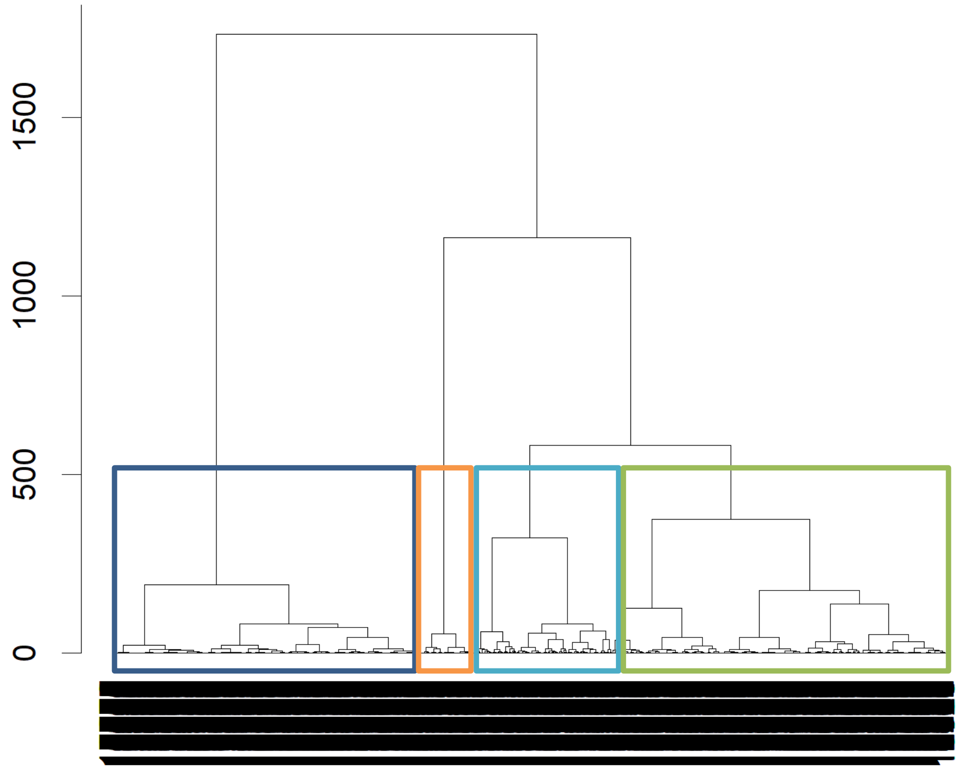


Fig. 4. The tree of cluster analysis using the data of Taiwanese large scale longline fishery in Southern Bluefin Tuna (SBT) Area E of the Indian Ocean.

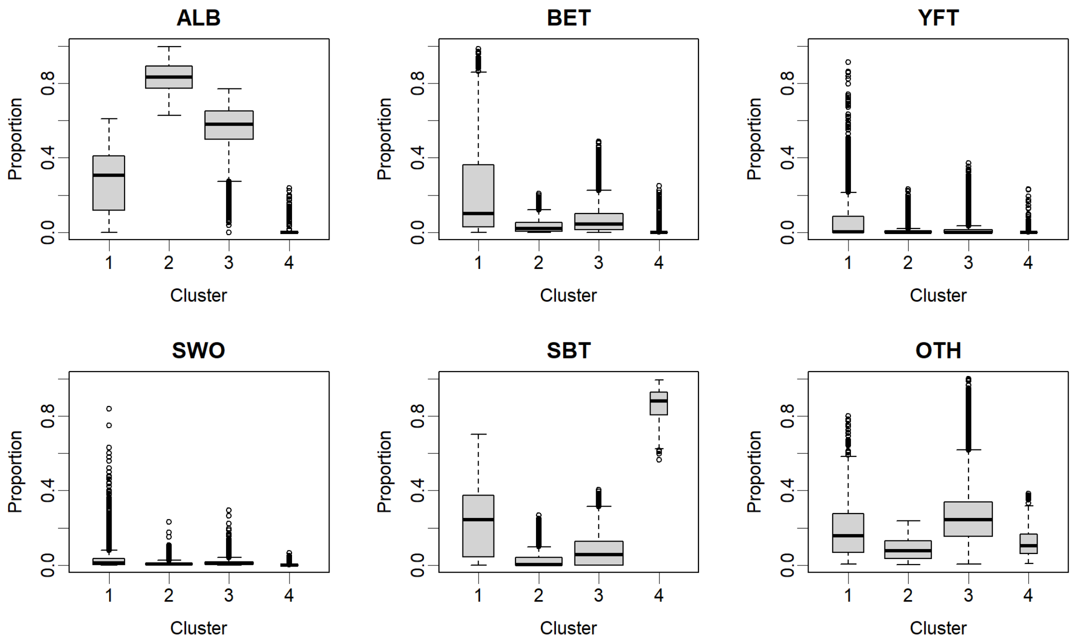


Fig. 5. Catch proportion by species for each cluster of Taiwanese large scale longline fishery in SBT Area E of the Indian Ocean.

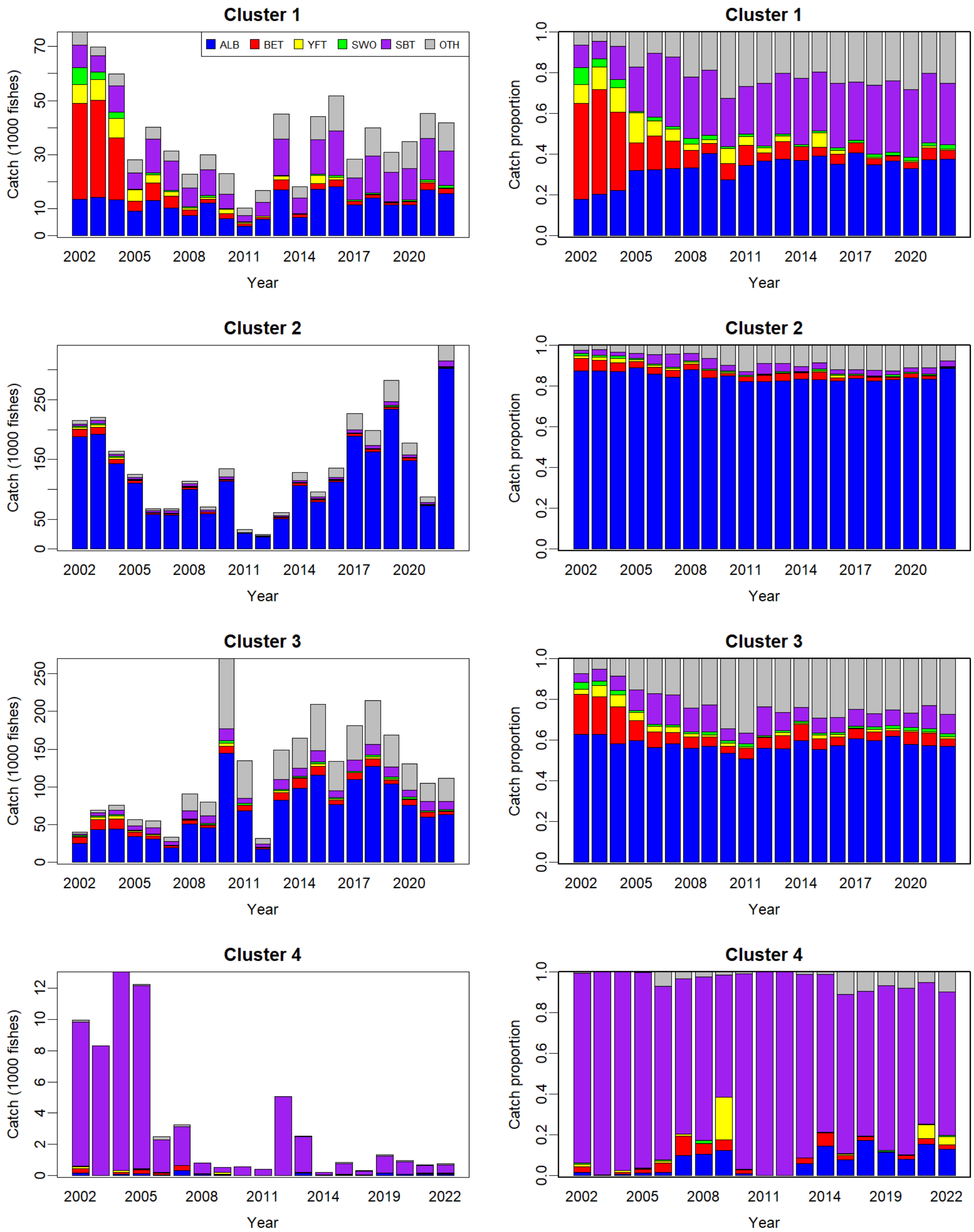


Fig. 6. Annual catch and catch proportion by species for each cluster of Taiwanese large scale longline fishery in SBT Area E of the Indian Ocean.

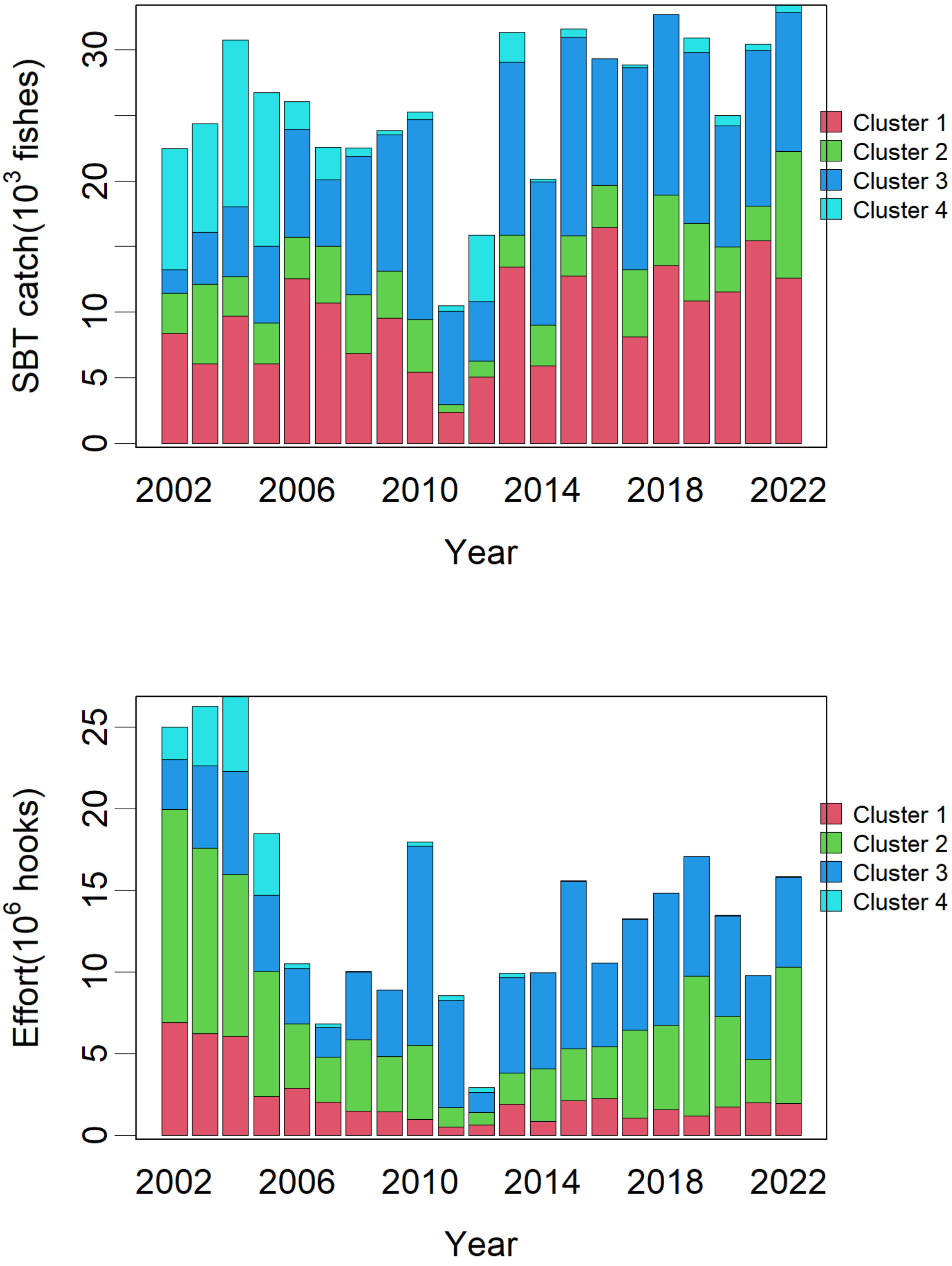


Fig. 7. Annual Southern Bluefin Tuna catches and efforts for each cluster of Taiwanese large scale longline fishery in Area E of the Indian Ocean.

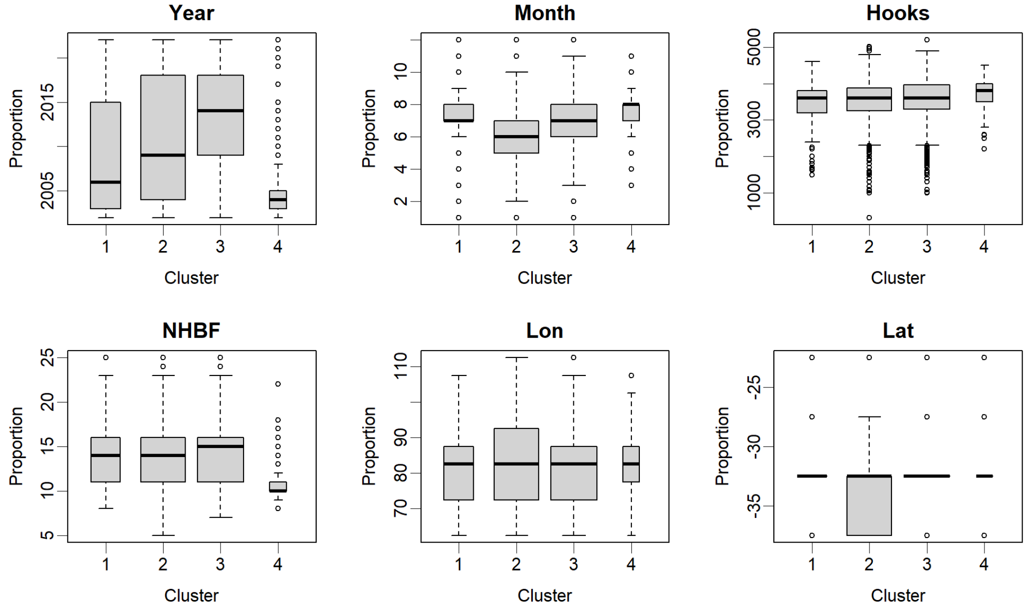


Fig. 8. Data composition by multiple factors for each cluster of Taiwanese large scale longline fishery in SBT Area E of the Indian Ocean.

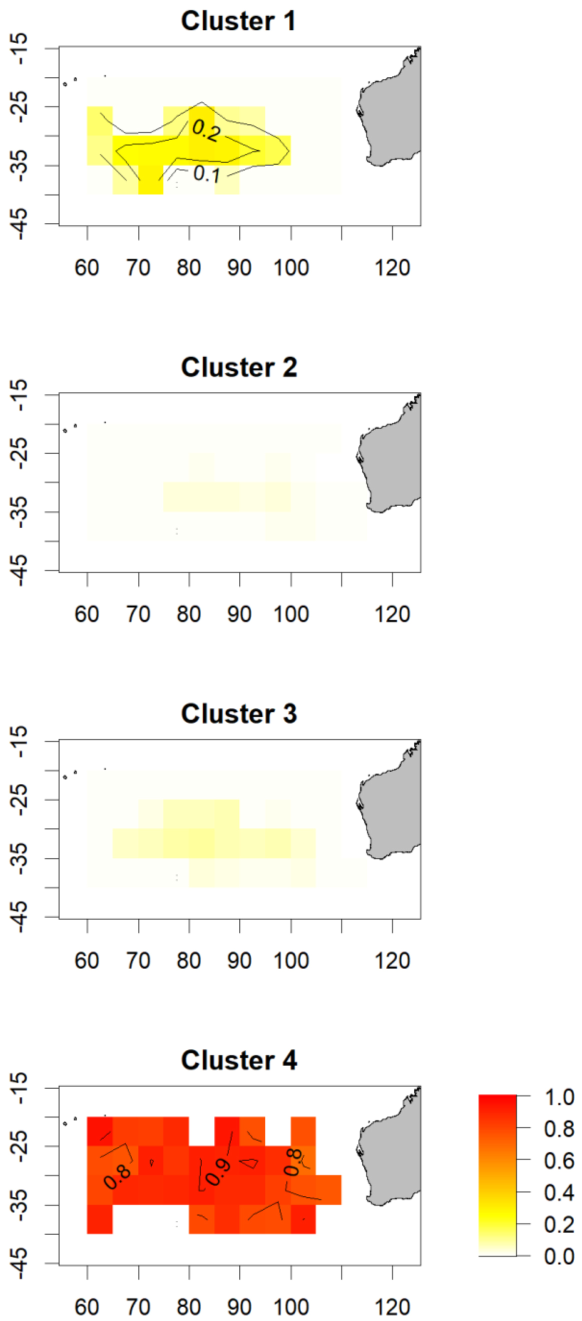


Fig. 9. Southern Bluefin Tuna catch distribution for each cluster of Taiwanese large scale longline fishery in Area E of the Indian Ocean. Red color represents high catch proportion and yellow color presents low catch proportion.

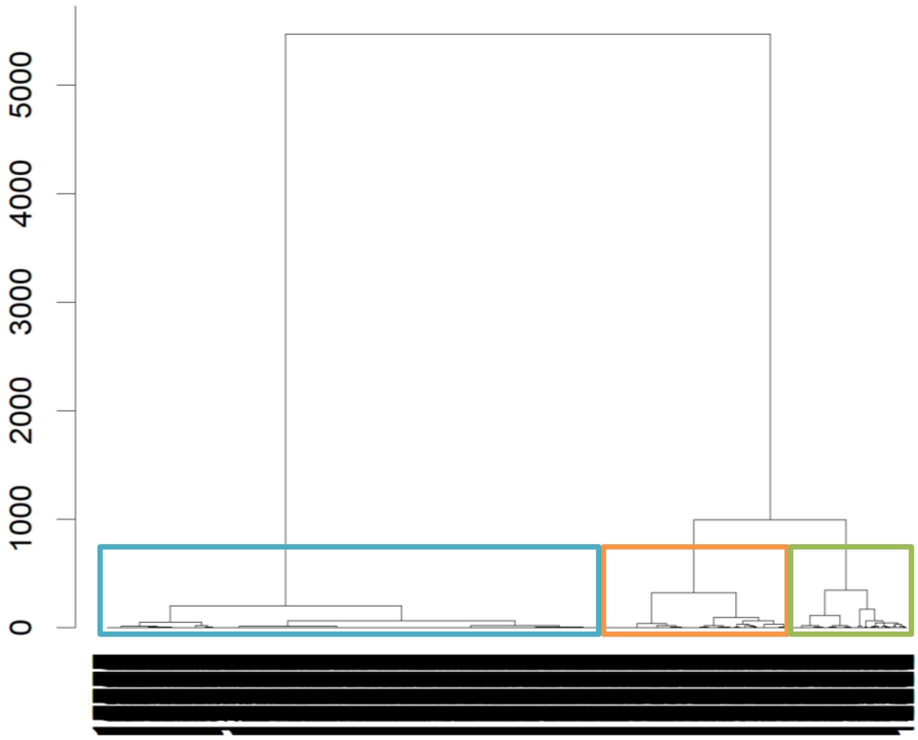


Fig. 10. The tree of cluster analysis using the data of Taiwanese large scale longline fishery in SBT Area W of the Indian Ocean.

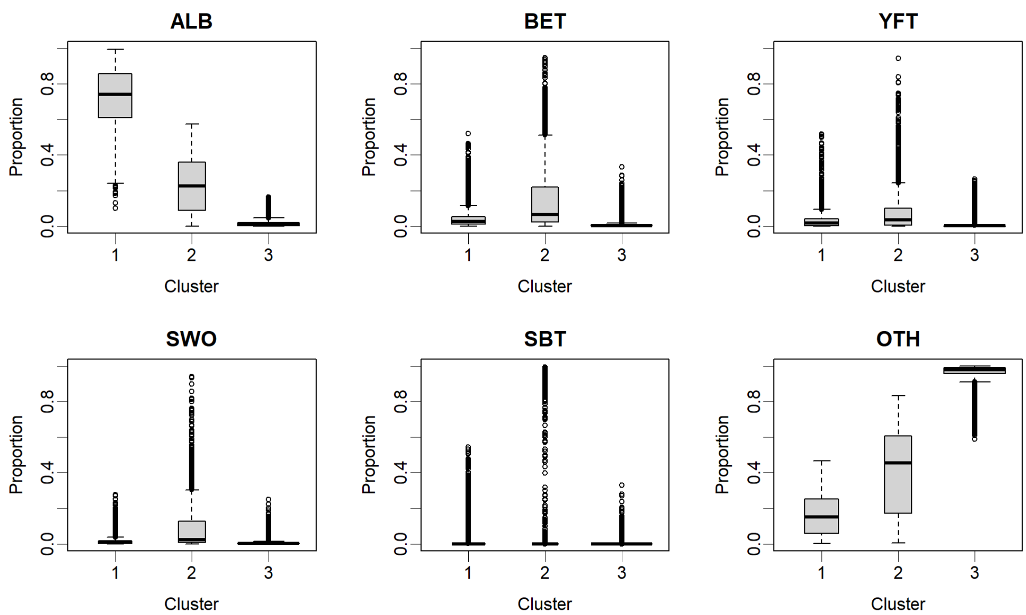


Fig. 11. Catch proportion by species for each cluster of Taiwanese large scale longline fishery in SBT Area W of the Indian Ocean.

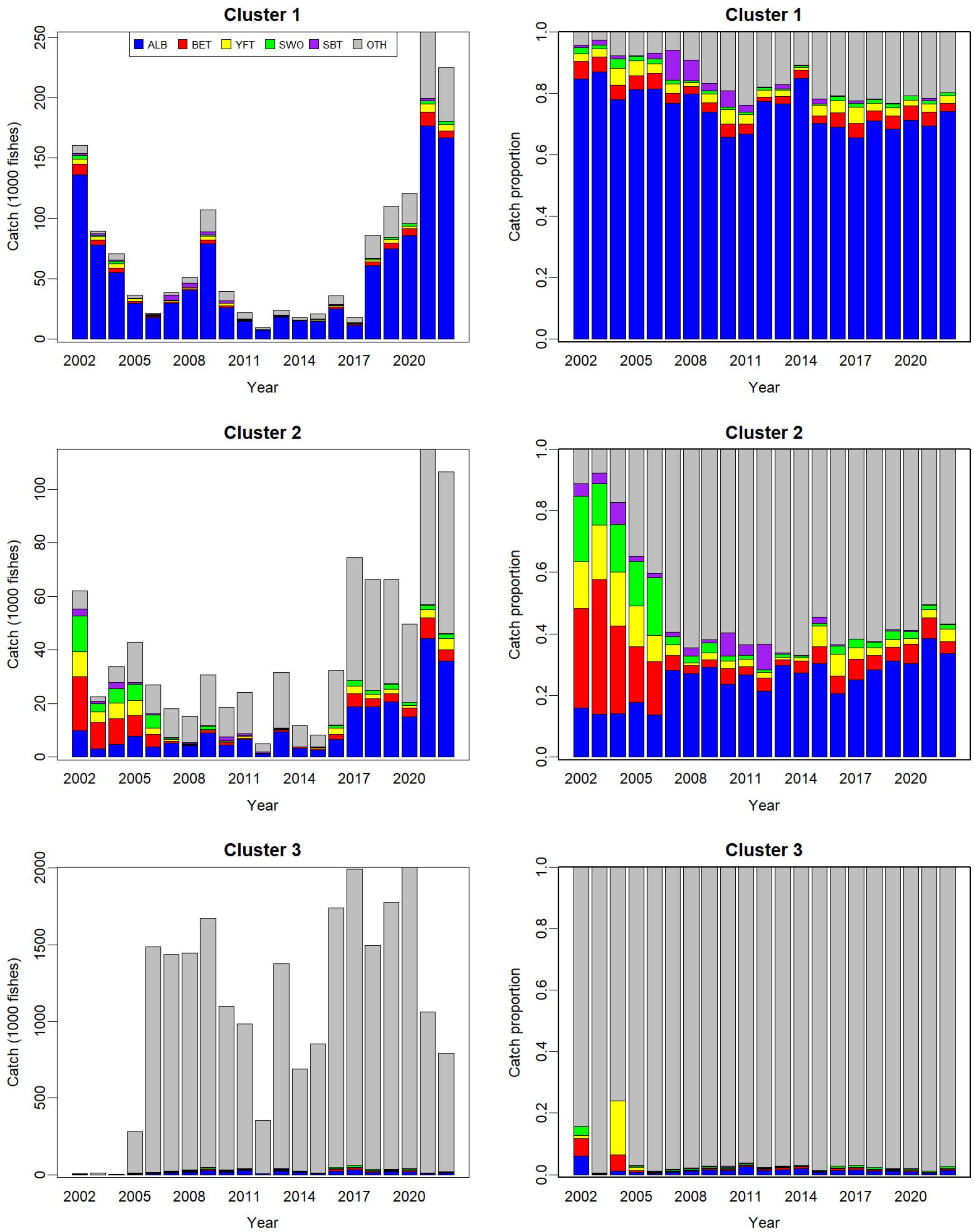
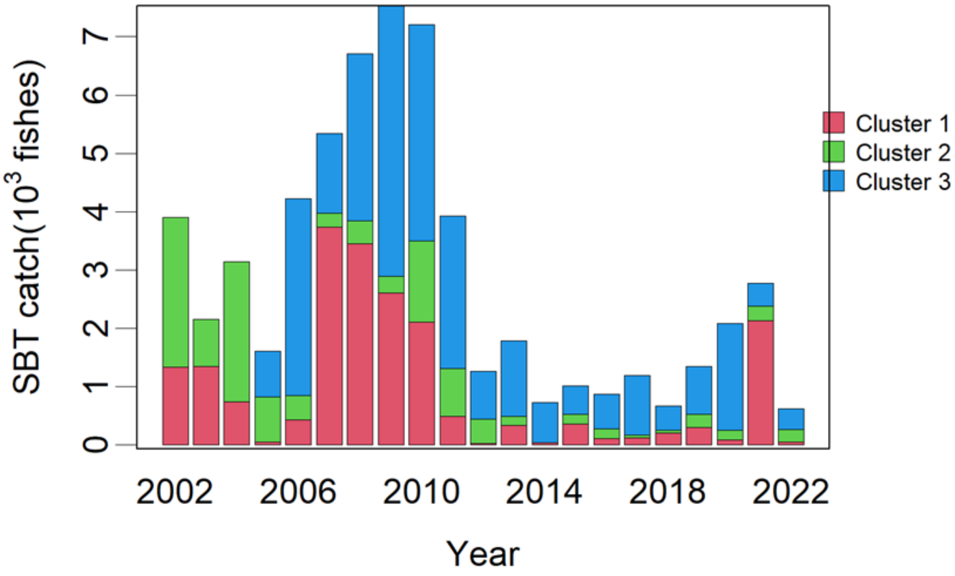


Fig. 12. Annual catch and catch proportion by species for each cluster of Taiwanese large scale longline fishery in SBT Area W of the Indian Ocean.



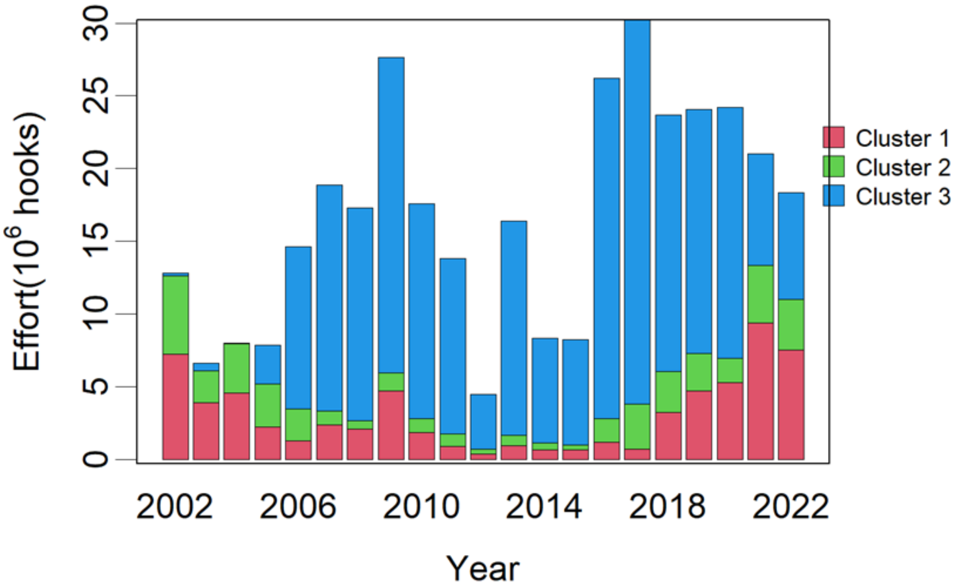


Fig. 13. Annual Southern Bluefin Tuna catches and efforts for each cluster of Taiwanese large scale longline fishery in Area W of the Indian Ocean.

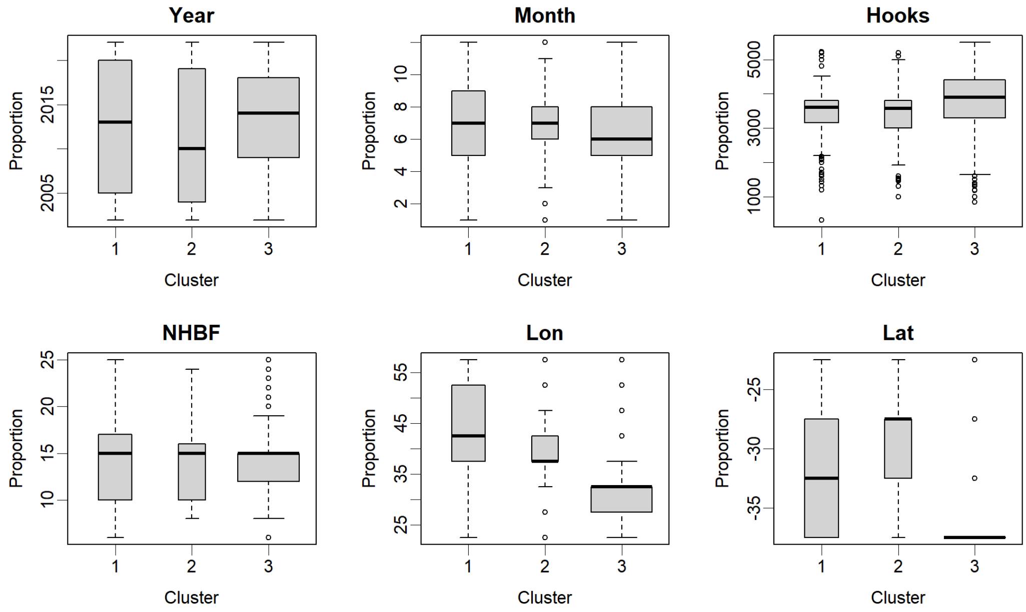


Fig. 14. Data composition by multiple factors for each cluster of Taiwanese large scale longline fishery in Southern Bluefin Tuna Area W of the Indian Ocean.

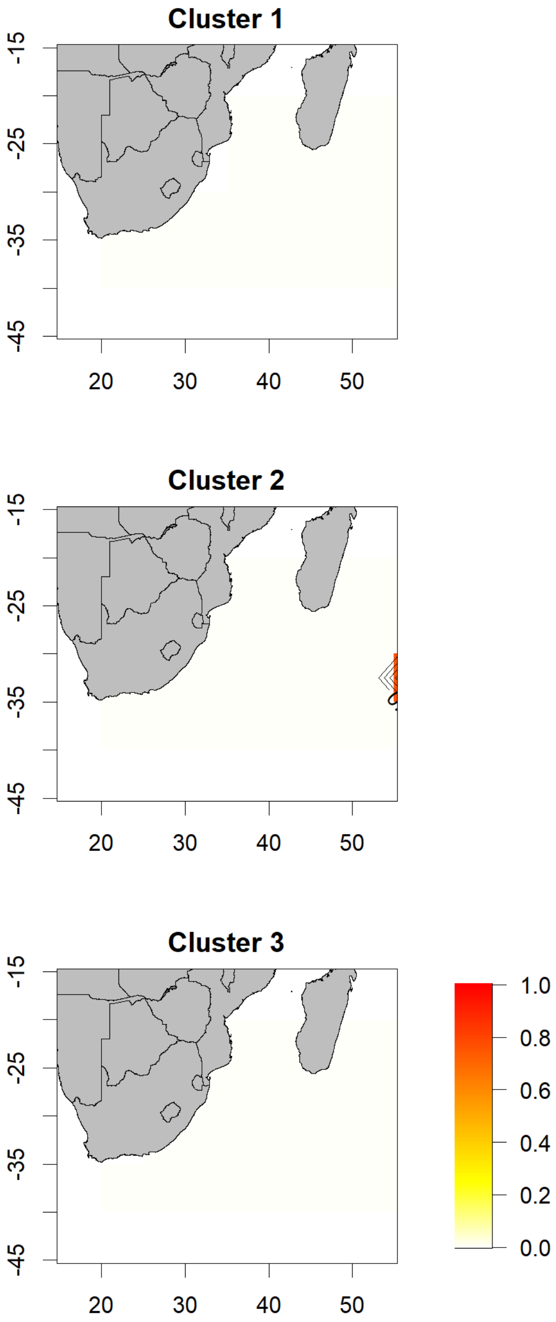
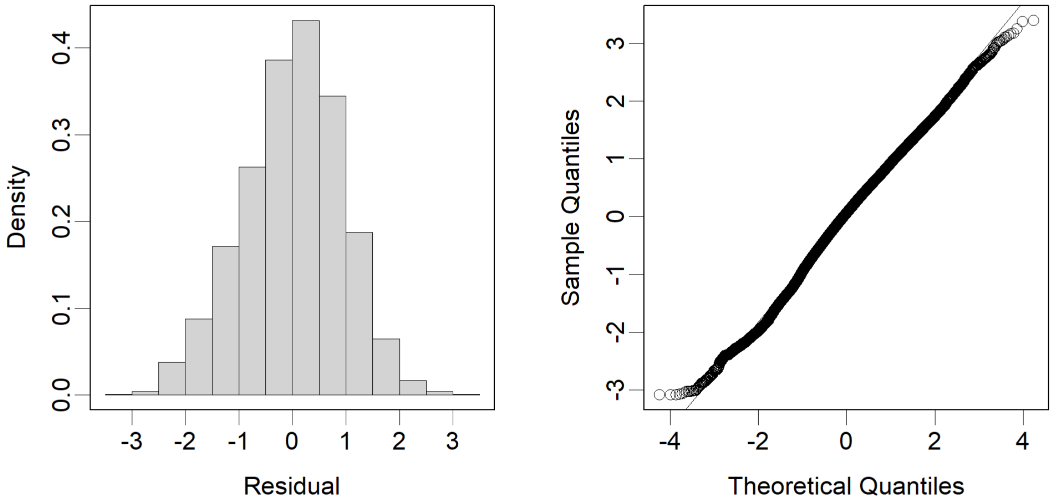


Fig. 15. Southern Bluefin Tuna catch distribution for each cluster of Taiwanese large scale longline fishery in Area W of the Indian Ocean. Red color represents high catch proportion and yellow color presents low catch proportion.

Area E



Area W

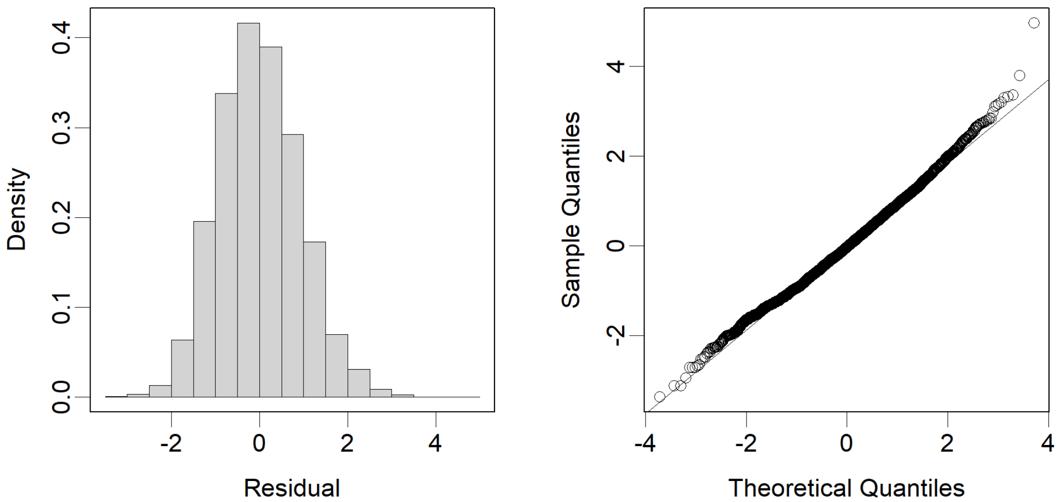
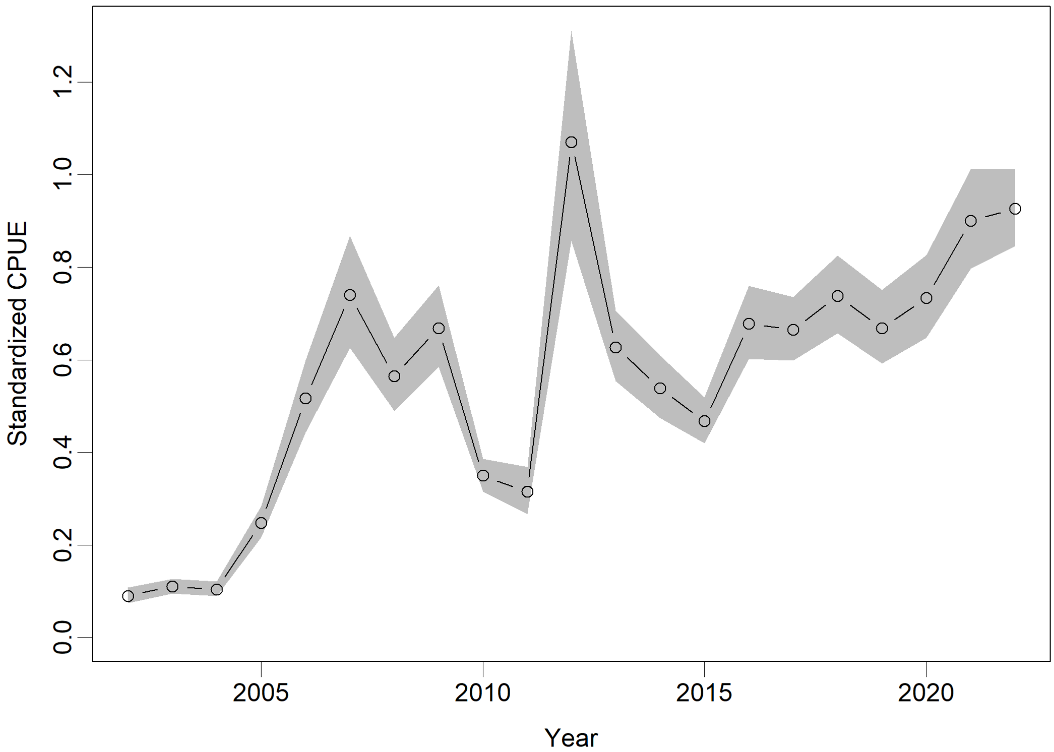


Fig. 16. The frequency distributions and Quantile-Quantile Plots for standardized residuals obtained from lognormal models for Area E and Area W.

Area E



Area W

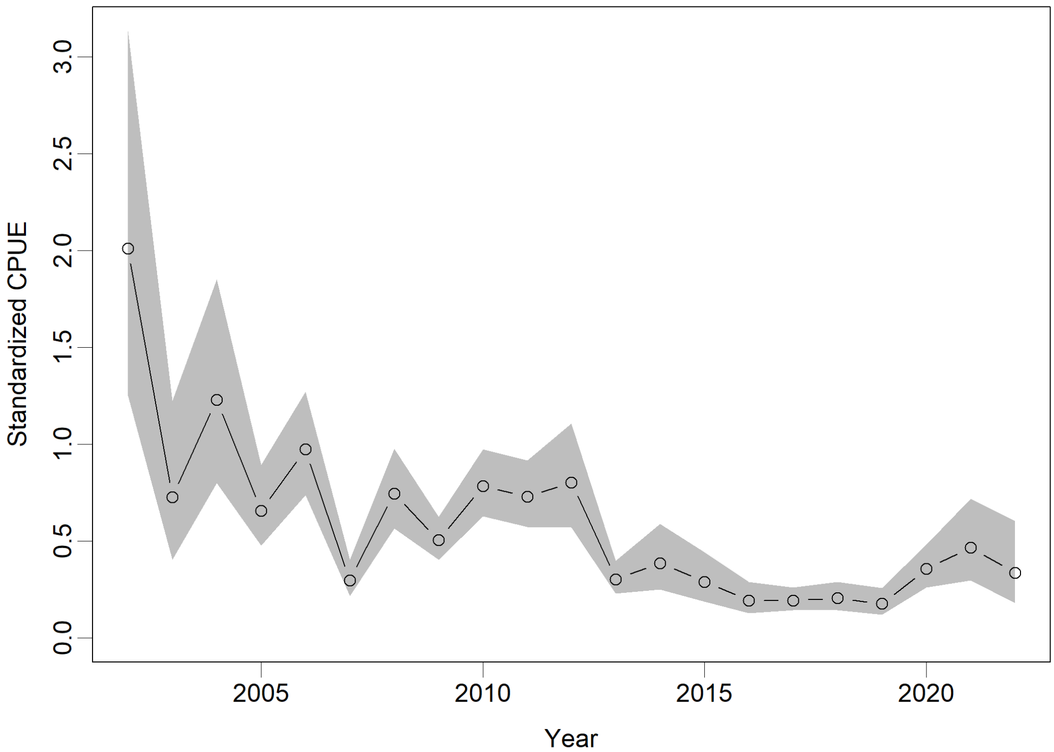
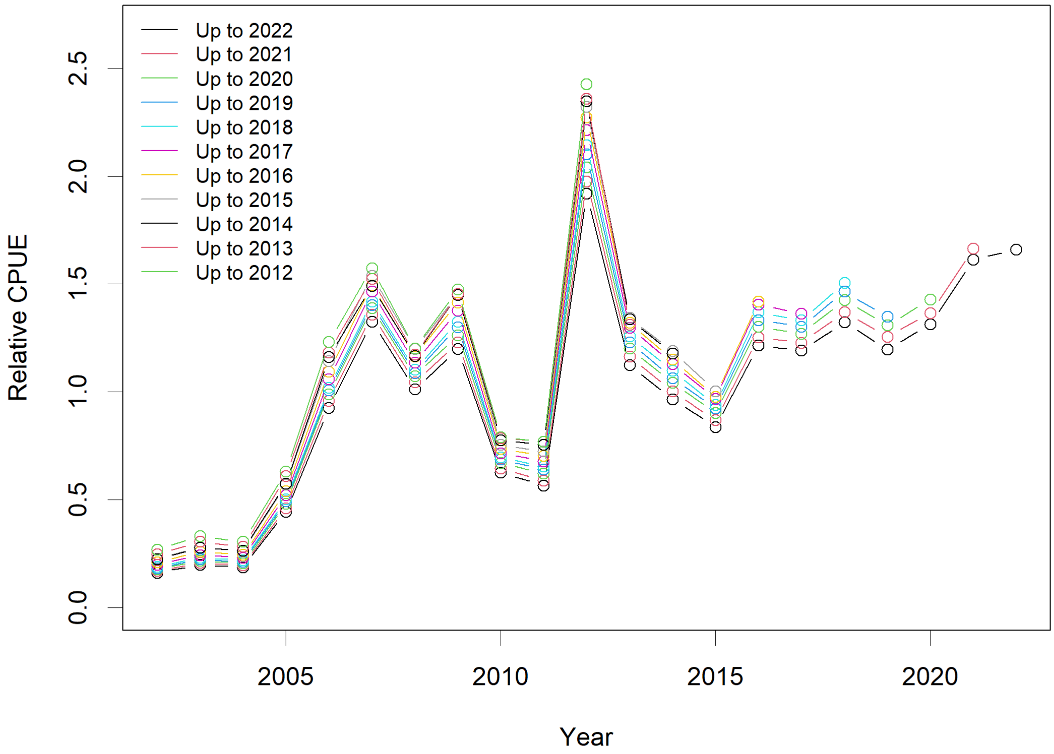


Fig. 17. Area-specific standardized CPUE of SBT caught by Taiwanese longline fishery. Shaded areas illustrate the 95% confidence intervals.

Area E



Area W

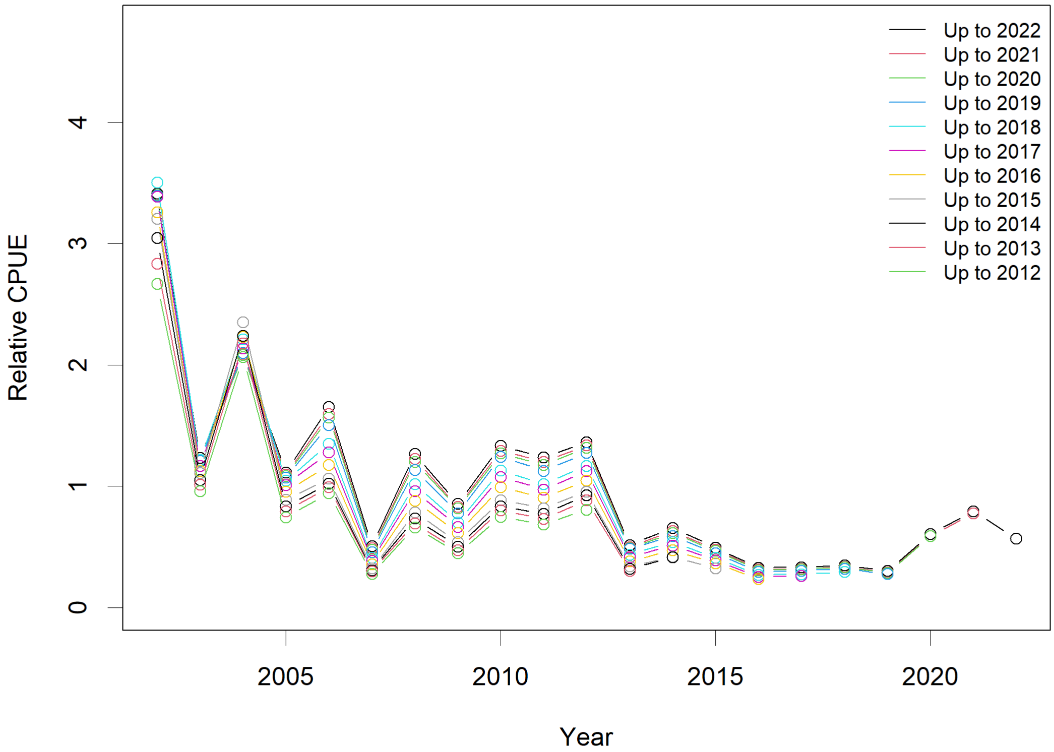


Fig. 18. The results of CPUE standardization based on including the updated data from different years.

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	2376	20	131.017	< 2.2e-16 ***
M	484	9	61.998	< 2.2e-16 ***
G	765	37	23.838	< 2.2e-16 ***
C	3550	3	1364.889	< 2.2e-16 ***
NHBF	51	2	29.131	2.275e-13 ***
Residuals	39037	45026		

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

Area W

Source of variance	SS	Df	F	Pr(>F)
Y	643.5	20	37.1035	< 2.2e-16 ***
M	427.7	10	49.3197	< 2.2e-16 ***
G	119.9	22	6.2845	< 2.2e-16 ***
C	12.9	1	14.8271	0.0001193 ***
NHBF	33.3	2	19.1778	5.047e-09 ***
Residuals	4328.2	4991		

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	7184.8	20	< 2.2e-16 ***
M	4978.6	11	< 2.2e-16 ***
G	7895.4	41	< 2.2e-16 ***
C	7183.5	3	< 2.2e-16 ***
NHBF	155.6	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	814.4	20	< 2.2e-16 ***
M	3172.7	11	< 2.2e-16 ***
G	2334.3	27	< 2.2e-16 ***
C	37.2	1	1.076e-09 ***
NHBF	11.8	2	0.002702 **

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