# CPUE standardization analyses for southern bluefin tuna based on the Taiwanese longline fishery data from 2002 to 2025

Ching-Ping Lu[1], Sheng-Ping Wang[1], Shu-Ting Chang[2], Yun-Ying Hsiao[3]

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

## ABSTRACT

The CPUE standardization analyses were conducted using statistical data from the Taiwanese longline fleet covering the period 2002–2025. The fishing operations of large-scale Taiwanese longliners were primarily distributed in the Indian Ocean south of 20°S. For the CPUE standardization analyses, the Indian Ocean was divided into two fishing areas using 60°E as the boundary: the central-eastern area (Area E), which represents the primary fishing ground, and the western area (Area W), which serves as a secondary fishing ground. Following the CPUE standardization procedure for both areas, cluster analysis was first applied to identify fishing-targeting strategies and to develop data filters for subsequent analyses. Subsequently, a simple delta-lognormal model without interaction terms was adopted to minimize potential confounding effects associated with interactions in the CPUE standardization process. The standardized CPUE results indicated an increasing trend both in Area E and Area W in 2025 after incorporating the updated data. And the retrospective analyses showed that the overall patterns of standardized CPUE trends in both fishing areas remained consistent patterns with those observed in previous analyses.

## 1. INTRODUCTION

Since the 1990s, SBT [Southern bluefin tuna (*Thunnus maccoyii;* SBT)] fishery was developed as the seasonal targeting style with some Taiwanese fishing vessels equipment deep-frozen freezers in the Indian Ocean. And before the 1990s, SBT was the bycatch species of Taiwanese tuna longline fishery while targeting the albacore in the Indian Ocean. Taiwanese fishing statistics system was reformed for improving the quality of the related SBT fishing information in 2002. And the reporting rate of SBT catch data had substantially improved since then (Anon, 2014). It is helpful for further understanding of the SBT fishery abundance indicators with the improved Taiwanese

SBT statistics fishing information. Therefore, this study aims to explore the variation of temporal and spatial patterns of SBT using the major species of catch and effort data of Taiwanese longline fishery operated in the waters of the south of 20°S of the Indian Ocean. Here, we presented the results of the CPUE standardization of Taiwanese SBT fishery using the updated statistics fishing information from 2002 to 2025.

## 2. MATERIALS AND METHODS

### 2.1. Catch and Effort data

The operational catch and effort data of large scale Taiwanese longline fisheries from 2002 to 2025 provided by the Overseas Fisheries Development Council (OFDC) of Taiwan were applied for the CPUE standardization in this study. The resolution of the dataset was compiled by 5×5 degree fishing location grids from Taiwanese longline fleets fishing information.

According to 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: primary fishing ground) and western area (Area W: secondary fishing ground) by the boundary at the 60°E (Fig. 1). Also, it is suitable for the CPUE standardization analyses of the SBT caught by large scale Taiwanese longline fishery. Here, all of the analyses were conducted with this area stratification with the boundary at the 60°E in the Indian Ocean.

### 2.2. Cluster analysis

Following the methodology of the previous study (Wang et al. 2015) and the suggestions by the experts of CCSBT ESC meetings in 2015 and 2016, at first, we conducted 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 two areas (the Area E and Area W) defined in this study. And the major species composition of the catches in both areas were applied for processing the cluster analyses. There were six major species were used including 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 the CCSBT ESC meeting in 2016, clustering analyses based on operational set-by-set data may contain a substantial amount of noise because most southern bluefin tuna (SBT) caught by Taiwanese vessels were bycatch,

with only a subset of vessels targeting SBT during specific fishing operations within the SBT fishing season. Therefore, the ESC recommended conducting the cluster analyses using aggregated data rather than operational set-by-set data to reduce noise and improve the reliability of the clustering results. Based on 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. However, 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). Then, we decided to conduct 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

To avoid the impact of large amount of SBT zero catch ("0") occurred in the fishing data sets, we applied the delta-lognormal models for the CPUE standardization of SBT caught by Taiwanese longline fishery. The main effects of year (Y), month (M), 5x5 grid (G), cluster (C) and number of hooks between floats (NHBF) were suggested by the previous ESC and all included in both of lognormal and delta models. The effect of 5x5 grid was represented as the effect of latitude and longitude. Additionally, the effects of cluster and NHBF were also included because various catch compositions can be observed in a cluster (Wang et al., 2017). The interactions between main effects were not considered in the models to avoid the confounding resulted from interactions. The models were conducted as below:

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

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,
$\mu$ 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,
$\varepsilon$ is the error term, $\varepsilon \sim N(0, \sigma^2)$.

The effects of year, month, cluster and 5x5 grid were treated as categorical variables. The effect of NHBF was treated as three styles 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 calculated 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 analyses of Area E and Area W

Cluster analyses were performed with the major species composition of the catches of Taiwanese large-scale longliners in the Indian Ocean for both two areas. The annually

composition of the catches of six major species from 2002 to 2025 were presented 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 in the Indian Ocean. As the main fishing area, there were four clusters identified in Area E (Figure 4). And according to the catch proportion by species of each cluster, multiple species were observed in the Cluster 1, the majority of components in the Cluster 1 were contributed by the ALB, BET, SBT and OTH operations, and the rest of operations with less proportion of YFT and SWO were also composed in this mixing Cluster. Then for the Cluster 2, the ALB operation was the majority of this Cluster and the catch proportions were almost over 0.8 each year, compared to the Cluster 1, it also included relatively fewer operations for BET, SBT and OTH. Similar pattern of the catch proportions was also observed in Cluster 3, the ALB operation took the majority with catch proportions approach 0.6 annually, the OTH operation showed the higher catch proportions than the rest of fewer operations for BET and SBT. 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 2 (Fig. 7). For SBT operation Cluster (Cluster 4), several fishing characteristics were observed and described by the main effects as bellowed and shown in Fig. 8 and Fig. 9: (1) the data mainly consisted of the data in the early 2000s; (2) the majority of fishing operations were concentrated from June to September; (3) the NHBF concentrated around 9~12 hooks; and (4) the operations also concentrated in the waters between 30°S and 35°S. 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. And the spatial distribution of higher SBT catch proportion both in Cluster 1 and Cluster 4 may identified the concentrated area in the waters between 65°E and 100°E.

For the Area W, the cluster analysis was performed and two clusters were selected in this secondary fishing Area (Fig. 10). According to the result, the ALB operations was found as the primary parts of composition in Cluster 1, the OTH operations was the secondary parts and contained the rest of operations such as BET, YFT, SBT, and SWO. The OTH operations took the majority parts in Cluster 2 which mostly composed by oilfish (Figs. 11 and 12). Primary parts of SBT catches were found in Cluster 1 in the early 2000s. (Fig. 13). For the fishing characteristics with various factors were described: (1) For the factor of year, Cluster 1 broadly consisted of the data period between 2006 and 2023, while the majority data of Cluster 2 were concentrated between 2009 and 2019; (2) NHBF of Cluster 1 was distributed wider than Cluster 2; And the NHBF of Cluster 2 was concentrated at the deep category, some operations performed

with ultra-deep category; (3) Fishing areas by longitude and latitude were different between two Clusters (Fig. 14). The SBT catch proportion of two cluster were illustrated by the spatial distribution and there were no obvious differences in the Area W. (Fig. 15).

### 3.2 CPUE standardization of Areas E and Areas W

For both of Areas E and W, the final results of model selection were decided with the lowest value of AIC. The results of ANOVA analysis for the lognormal models are shown in Table 1 for both areas. And all of the effects were statistically significant for both areas. About 25.0% and 33.6% 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 34.3% and 21.8% 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 different patterns in the period of 2002 to 2025 (Fig. 17). First, in the Area E, the standardized CPUE series gradually showed the increasing trend after 2004, and reached the first high point in 2007, and also revealed variation in the period of 2004 to 2011. Then substantially increased and reached the highest peak in 2012. After that the trend gradually decreased until 2015. And after 2016, it remained higher and relatively stable increasing pattern and reached the higher values in 2022 and 2025. This year, with the updated data of 2025, it presented the increased status continually. Moreover, for Area W, as the secondary fishing area, the standardized CPUE series 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 presented the highest in 2021 then back to lower value as the past. And with the updated data of 2025, it revealed increased to the similar level of 2020. The pattern of CPUE trends in two areas were relatively both increasing with the updated information of 2025.

### 3.3 Retrospect analysis of Areas E and Areas W

The retrospect analyses were conducted to understand the influence of incorporating updated 2025 data into the CPUE standardization for both Area E and Area W. The analysis was performed by removing the data from 2025 to 2012. The results indicated that the influence of adding the updated data on the CPUE standardization was negligible for both two areas. Although the index values of standardized CPUE series varied 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.

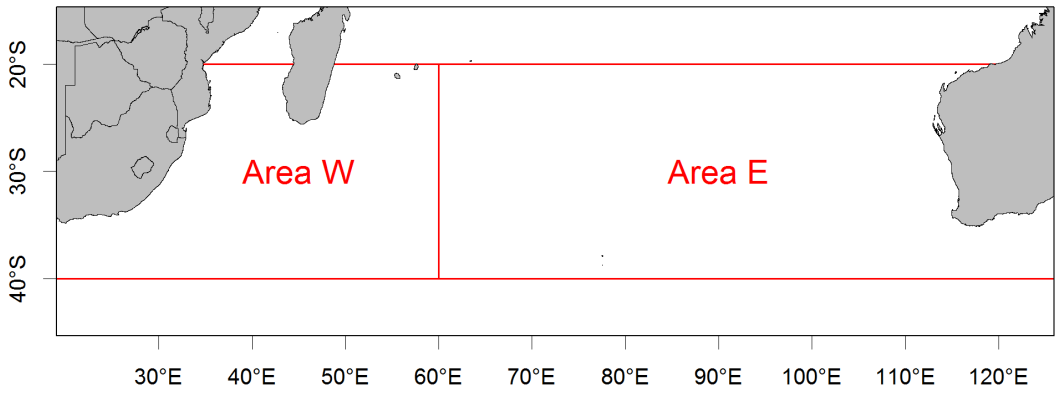



Fig. 1. The map of fishing area stratification for southern bluefin tuna of Taiwanese large scale longline fishery in the Indian Ocean.

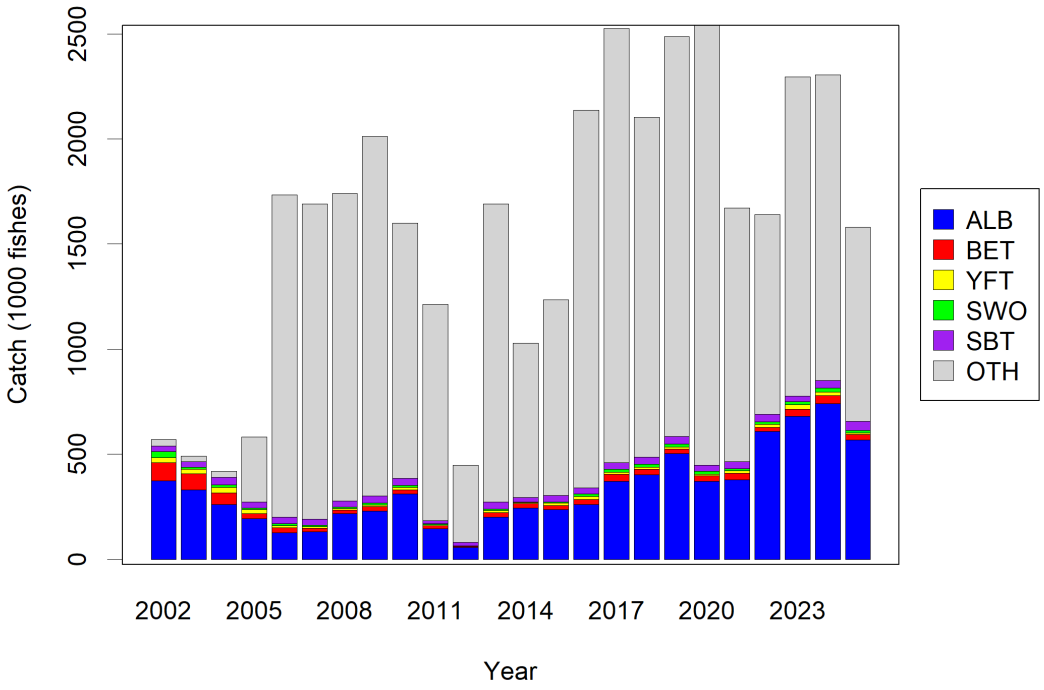



Fig. 2. Annual catch composition of the major species caught by Taiwanese large-scale longline fleets operated in the waters of south of 20°S in the Indian Ocean from 2002 to 2025.

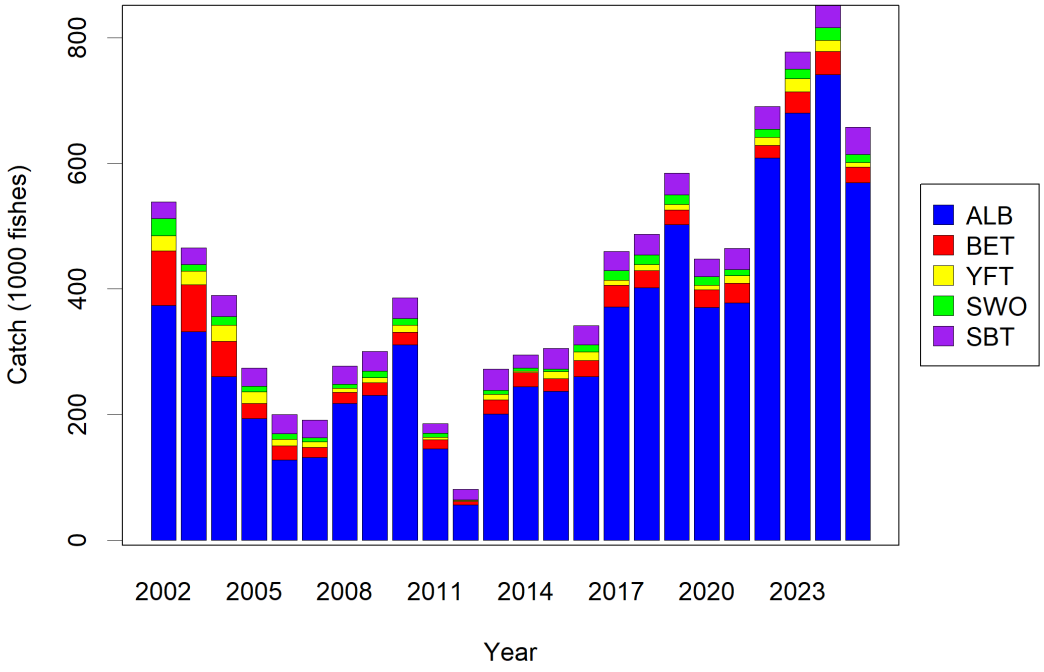



Fig. 3. Annual catch composition of the major species caught by Taiwanese large-scale longline fleets operated in the waters of south of 20°S in the Indian Ocean from 2002 to 2025. The catches of OTH are excluded.

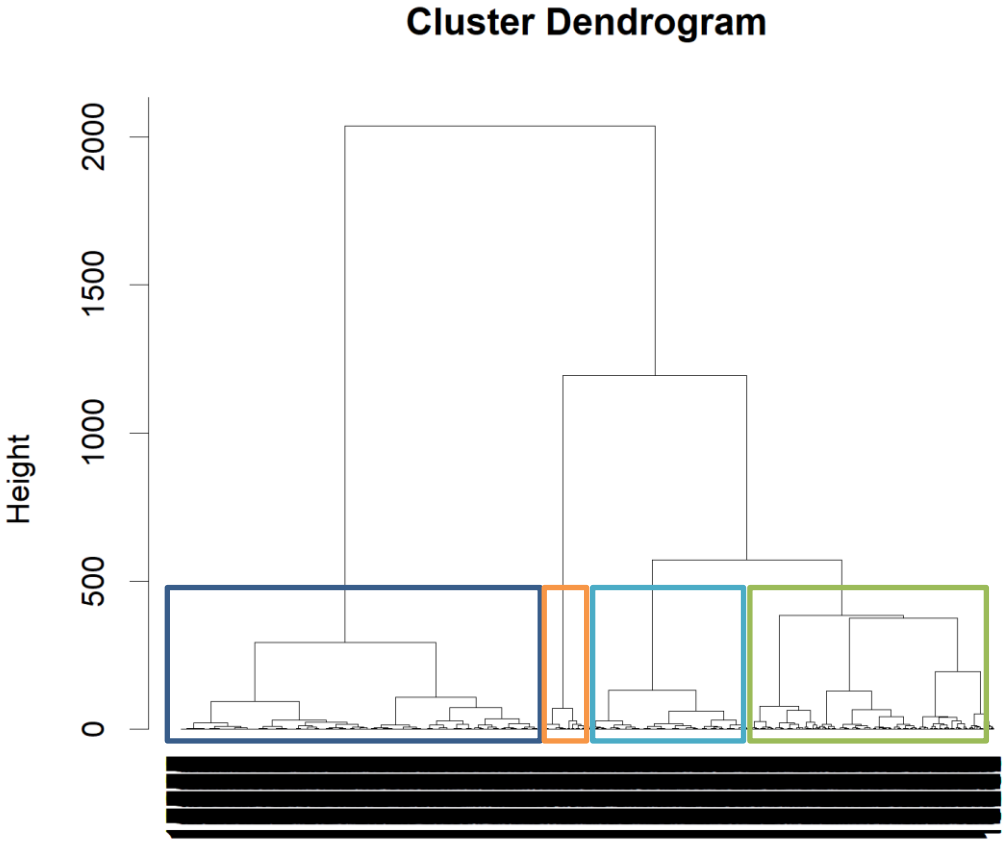



Fig. 4. The tree of cluster analysis using the data of Taiwanese large scale longline fishery in Southern Bluefin Tuna (SBT) Area E of the Indian Ocean.

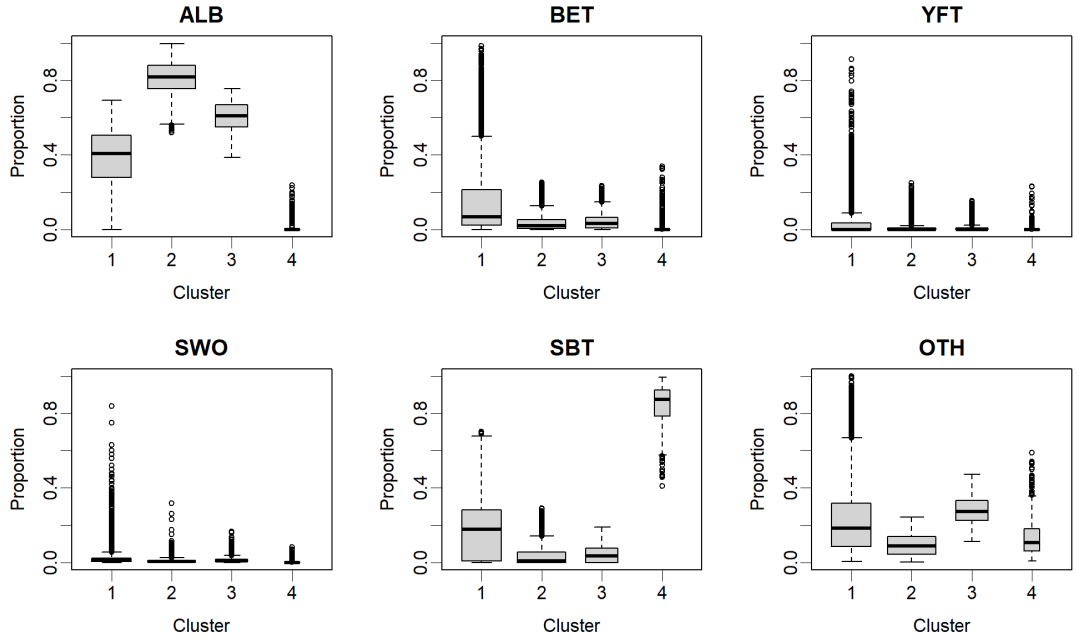



Fig. 5. Catch proportion by species for each cluster of Taiwanese large scale longline fishery in SBT Area E of the Indian Ocean.

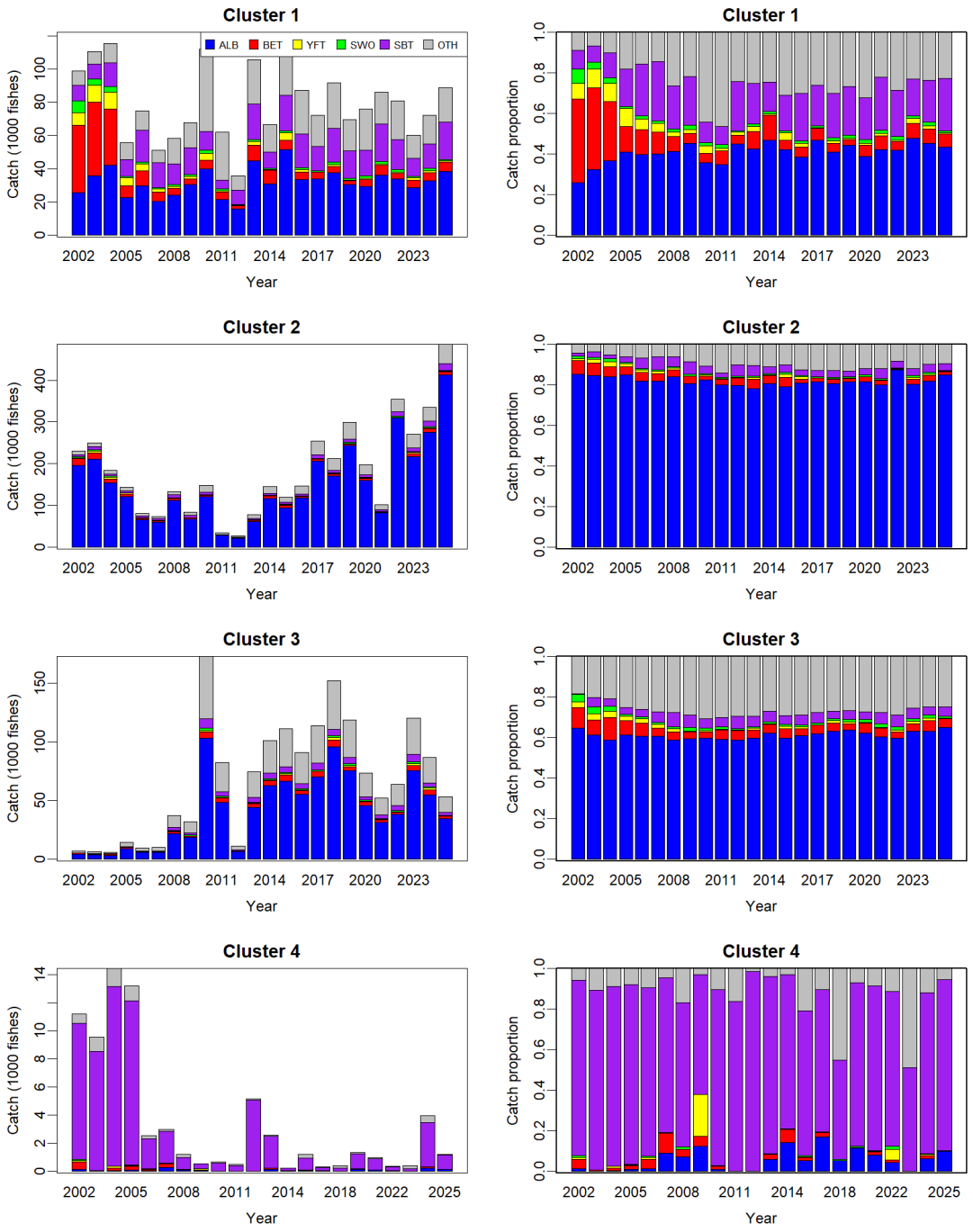



Fig. 6. Annual catch and catch proportion by species for each cluster of Taiwanese large scale longline fishery in SBT Area E of the Indian Ocean.

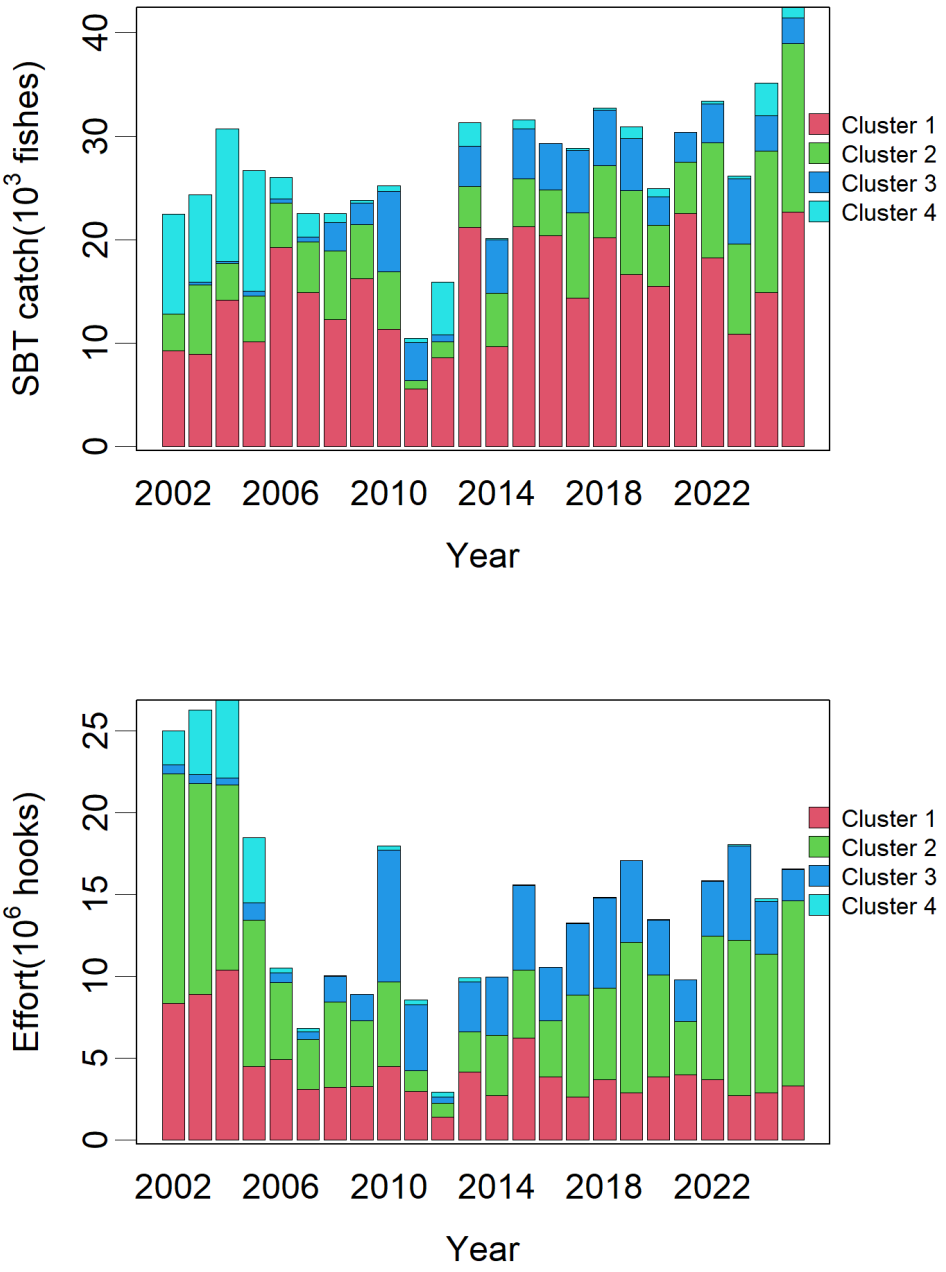



Fig. 7. Annual Southern Bluefin Tuna catches and efforts for each cluster of Taiwanese large scale longline fishery in Area E of the Indian Ocean.

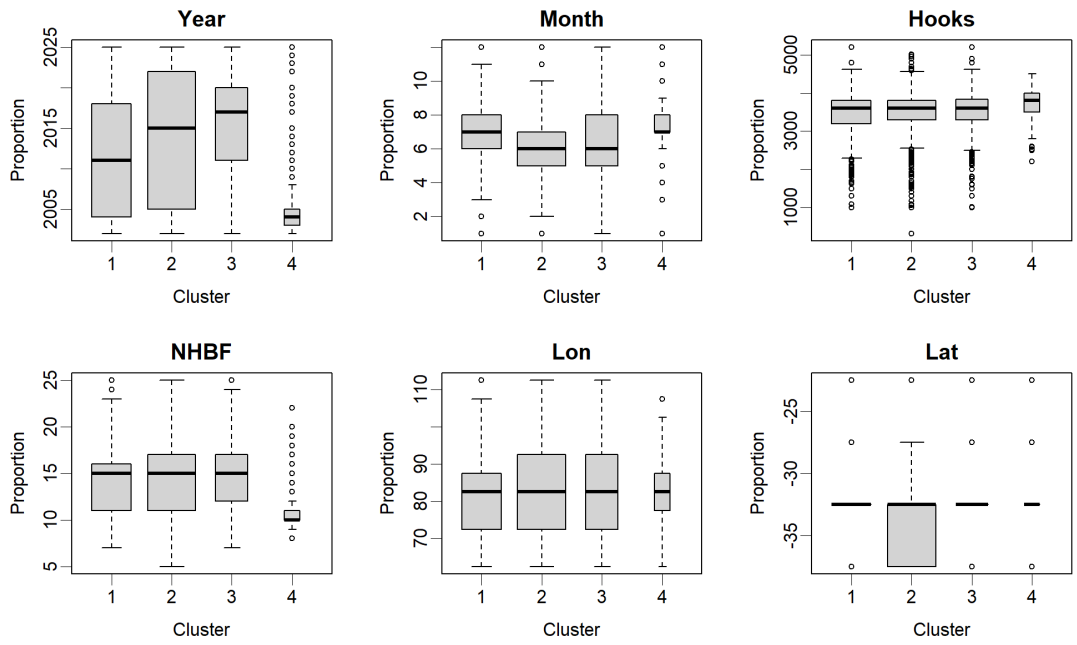



Fig. 8. Data composition by multiple factors for each cluster of Taiwanese large scale longline fishery in SBT Area E of the Indian Ocean.

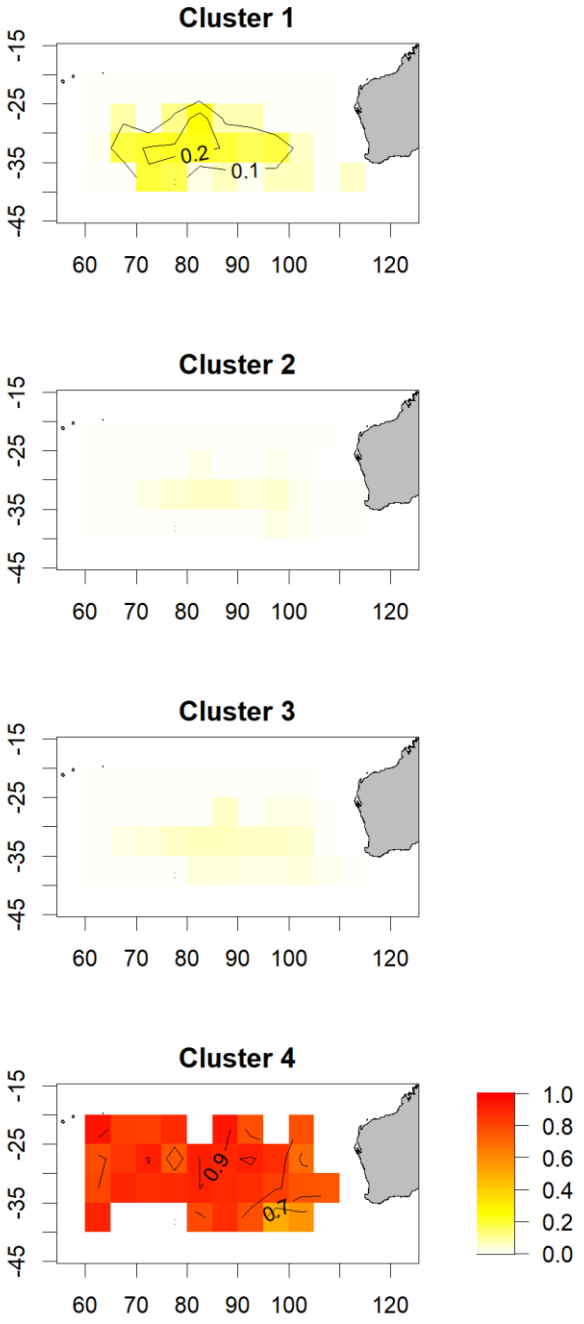



Fig. 9. Southern Bluefin Tuna catch distribution for each cluster of Taiwanese large scale longline fishery in Area E of the Indian Ocean. Red color represents high catch proportion and yellow color presents low catch proportion.

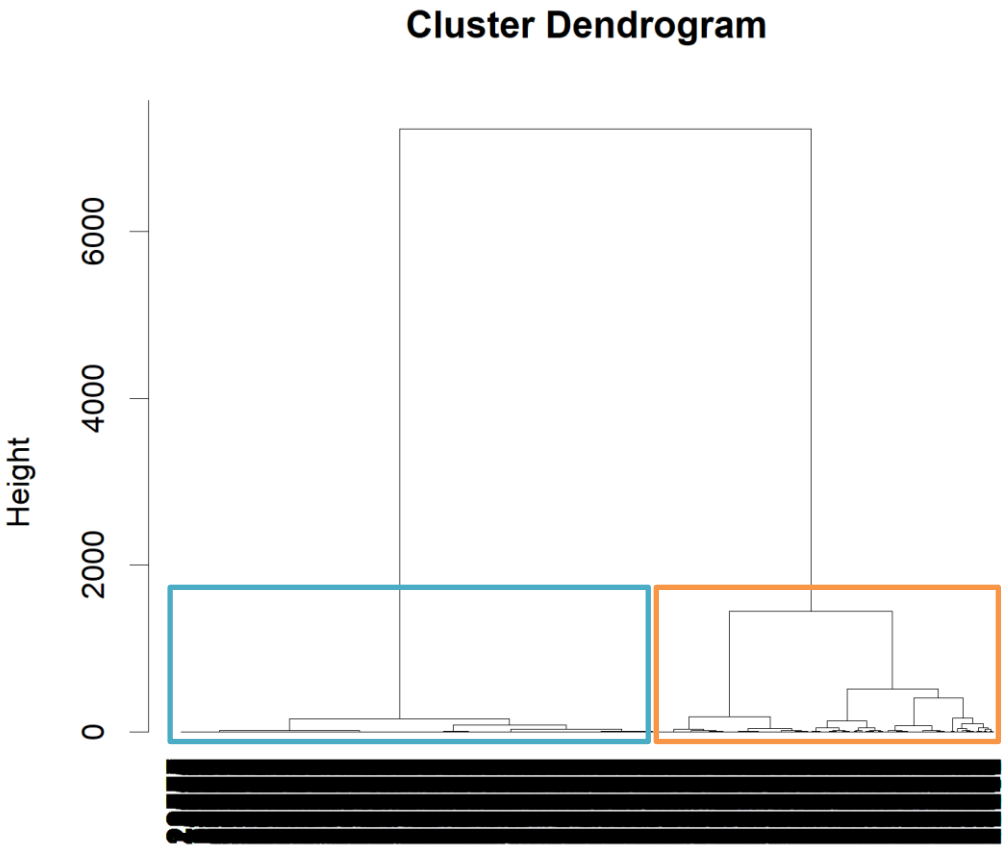



Fig. 10. The tree of cluster analysis using the data of Taiwanese large scale longline fishery in SBT Area W of the Indian Ocean.

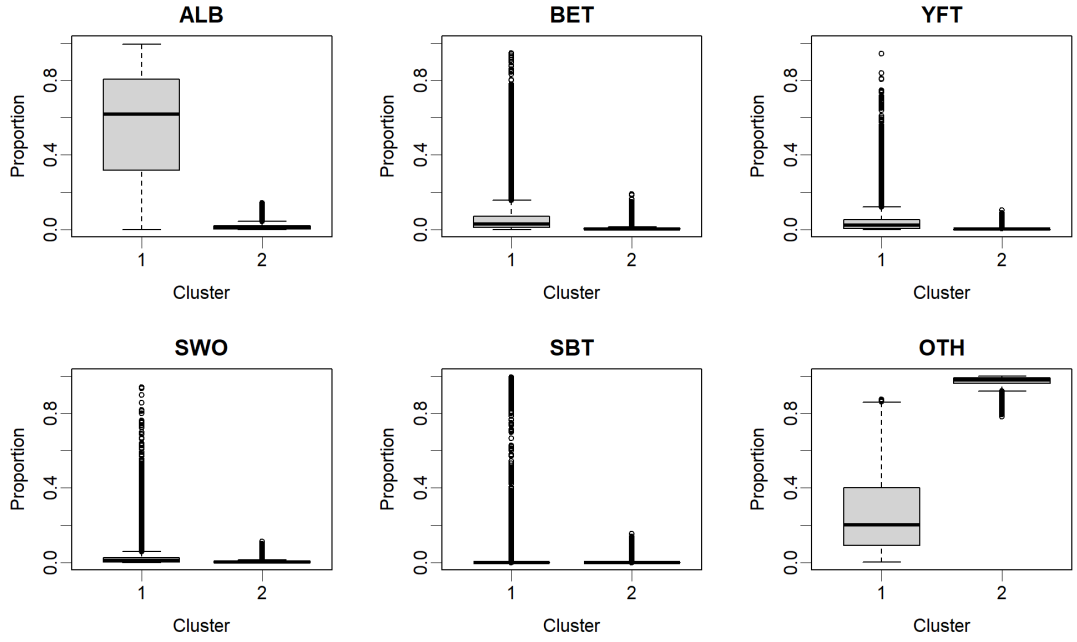



Fig. 11. Catch proportion by species for each cluster of Taiwanese large scale longline fishery in SBT Area W of the Indian Ocean.

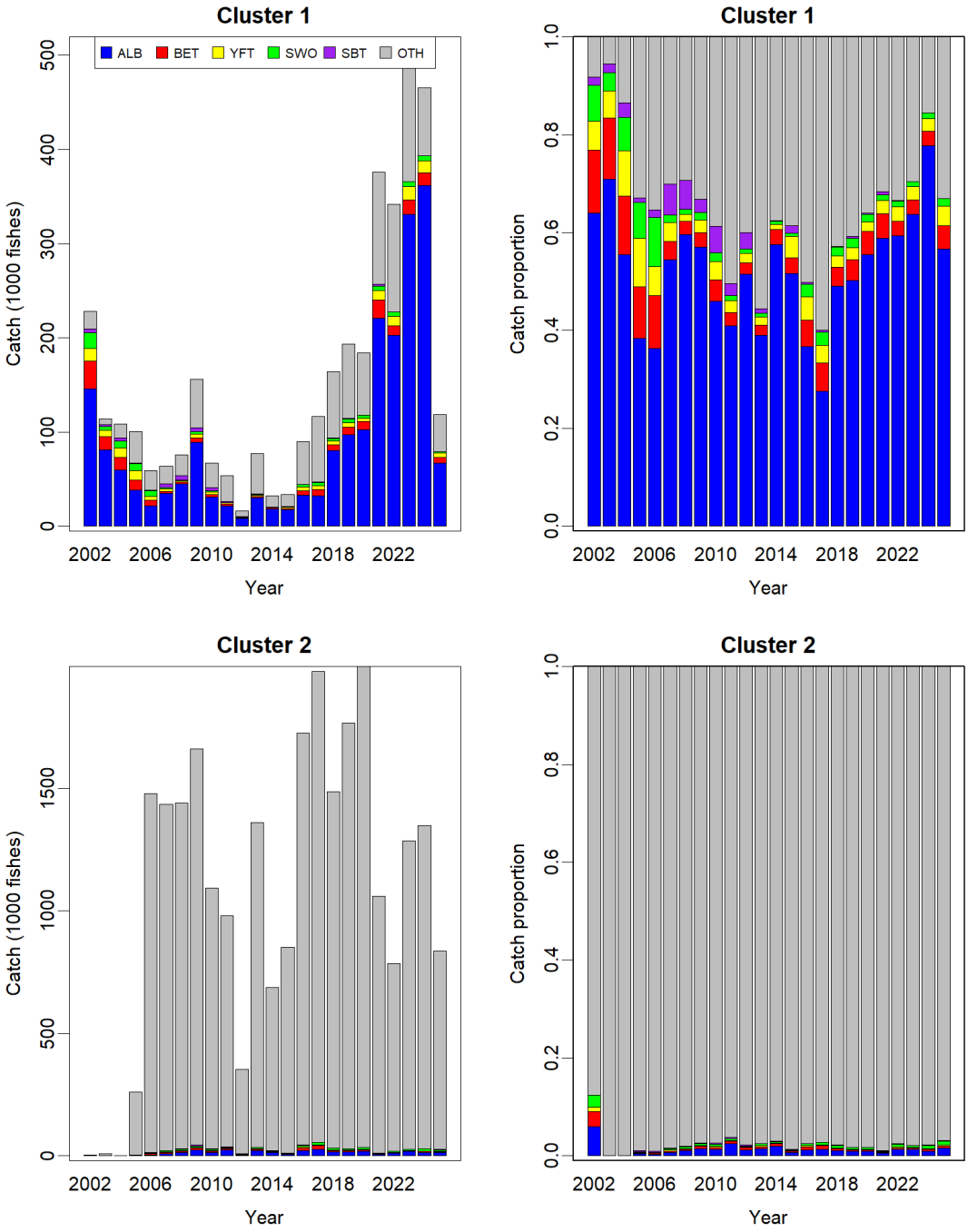



Fig. 12. Annual catch and catch proportion by species for each cluster of Taiwanese large scale longline fishery in SBT Area W of the Indian Ocean.

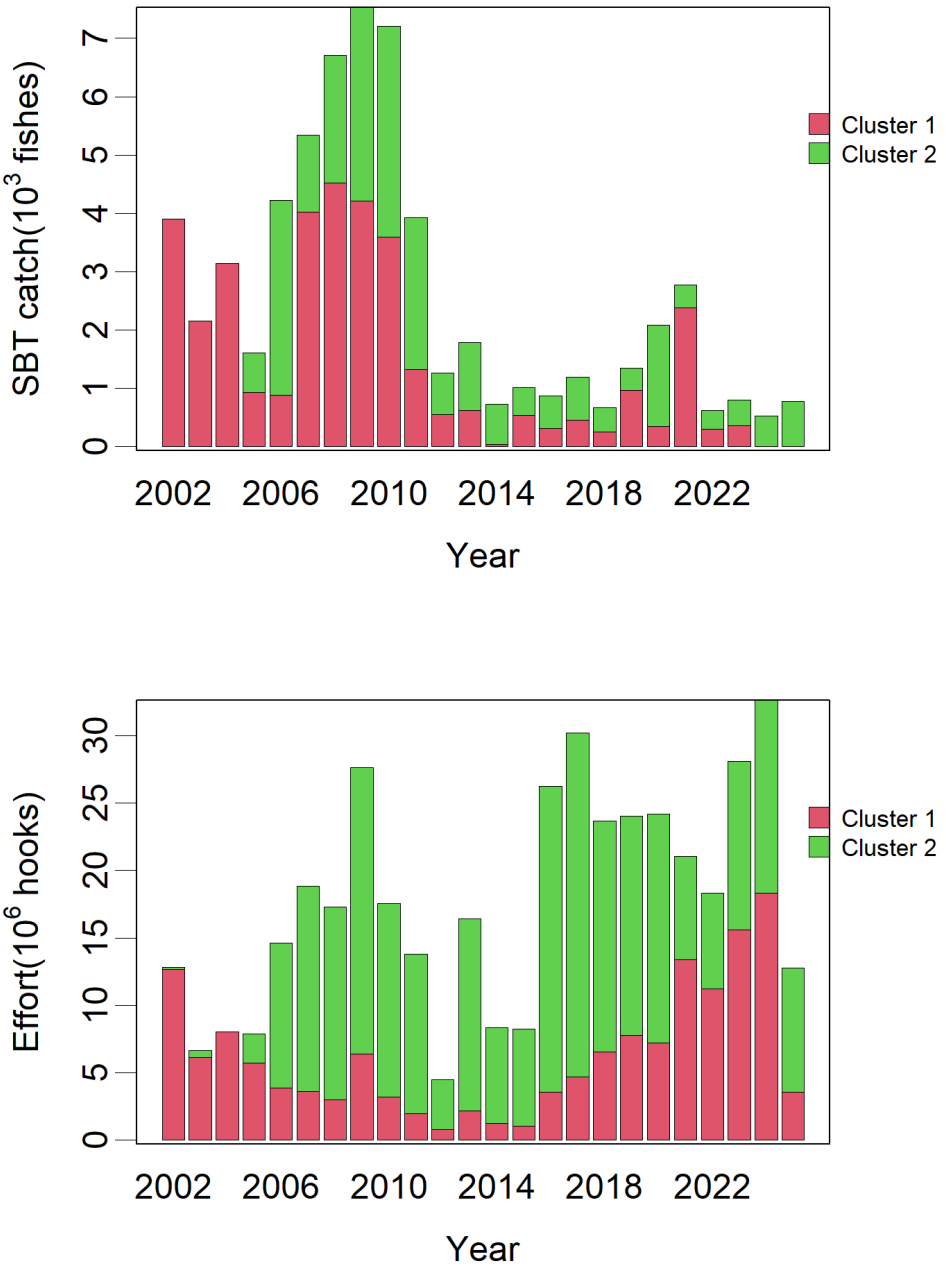



Fig. 13. Annual Southern Bluefin Tuna catches and efforts for each cluster of Taiwanese large scale longline fishery in Area W of the Indian Ocean.

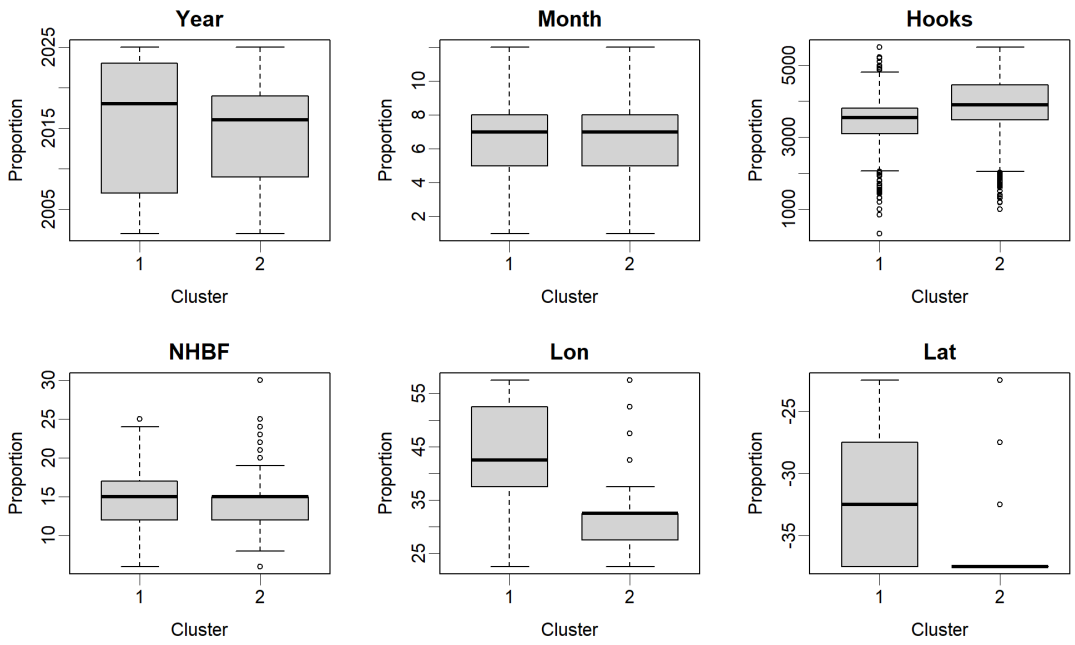



Fig. 14. Data composition by multiple factors for each cluster of Taiwanese large scale longline fishery in Southern Bluefin Tuna Area W of the Indian Ocean.

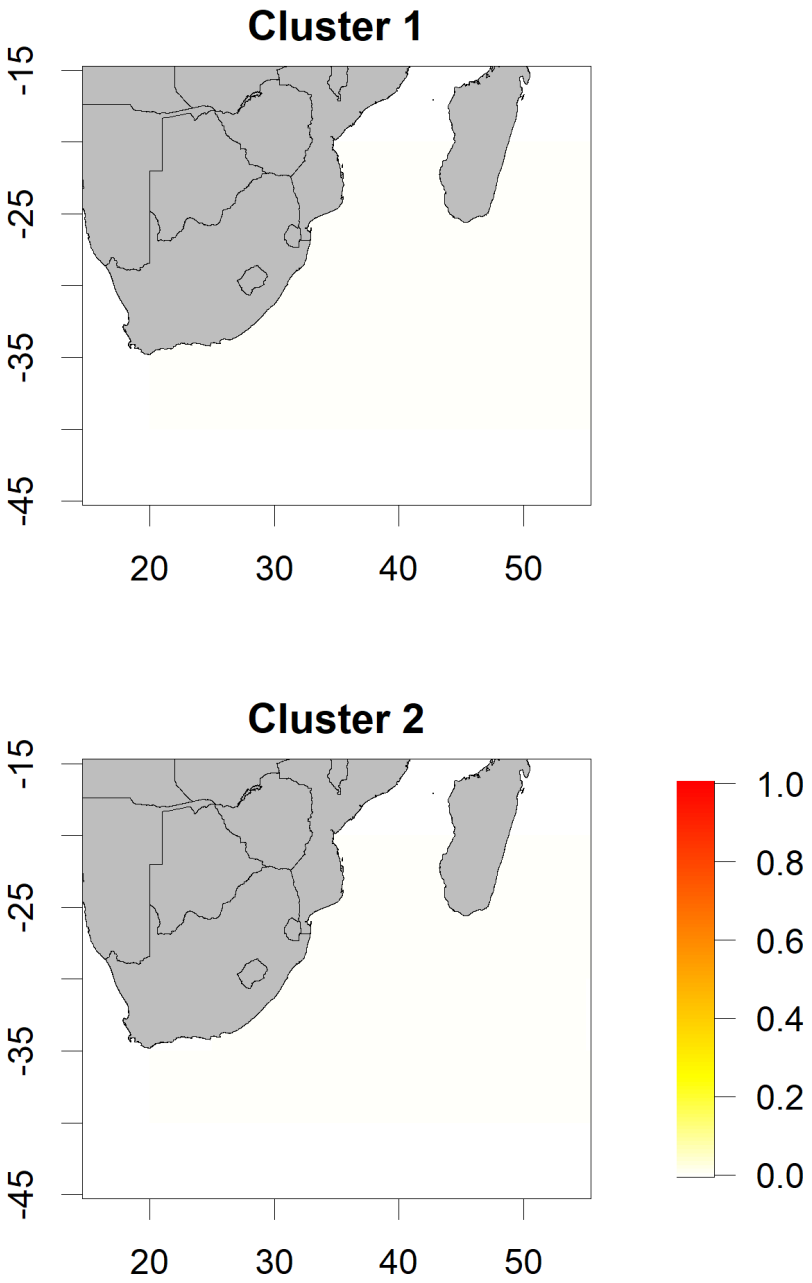
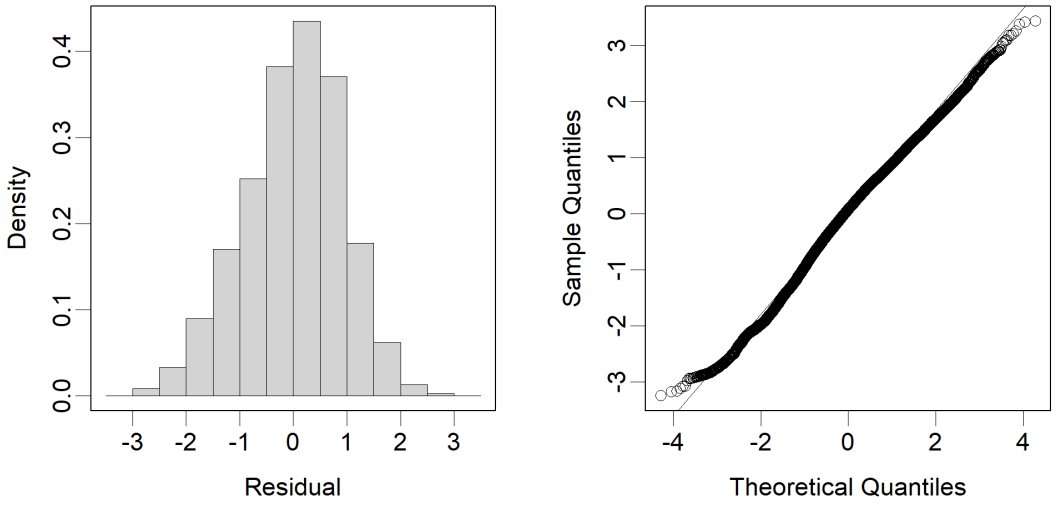



Fig. 15. Southern Bluefin Tuna catch distribution for each cluster of Taiwanese large scale longline fishery in Area W of the Indian Ocean. Red color represents high catch proportion and yellow color presents low catch proportion.

Area E




Area W

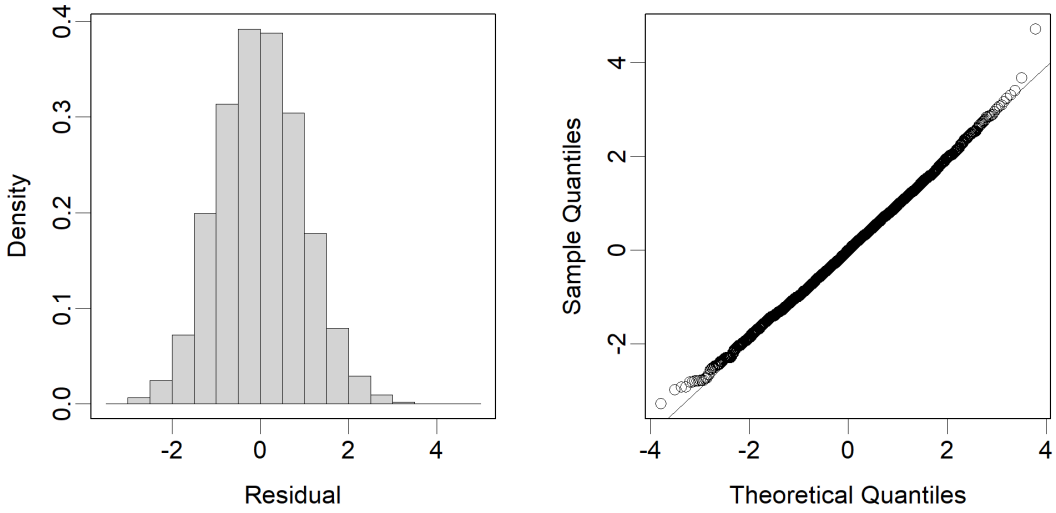
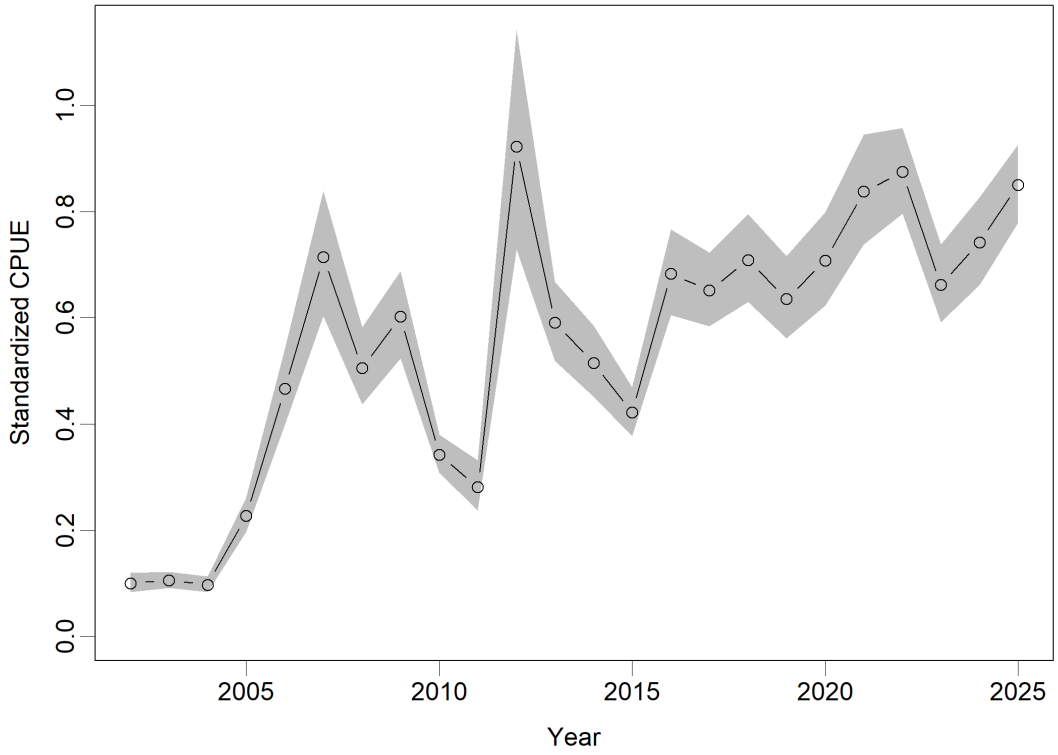



Fig. 16. The frequency distributions and Quantile-Quantile Plots for standardized residuals obtained from lognormal models for Area E and Area W.

Area E




Area W

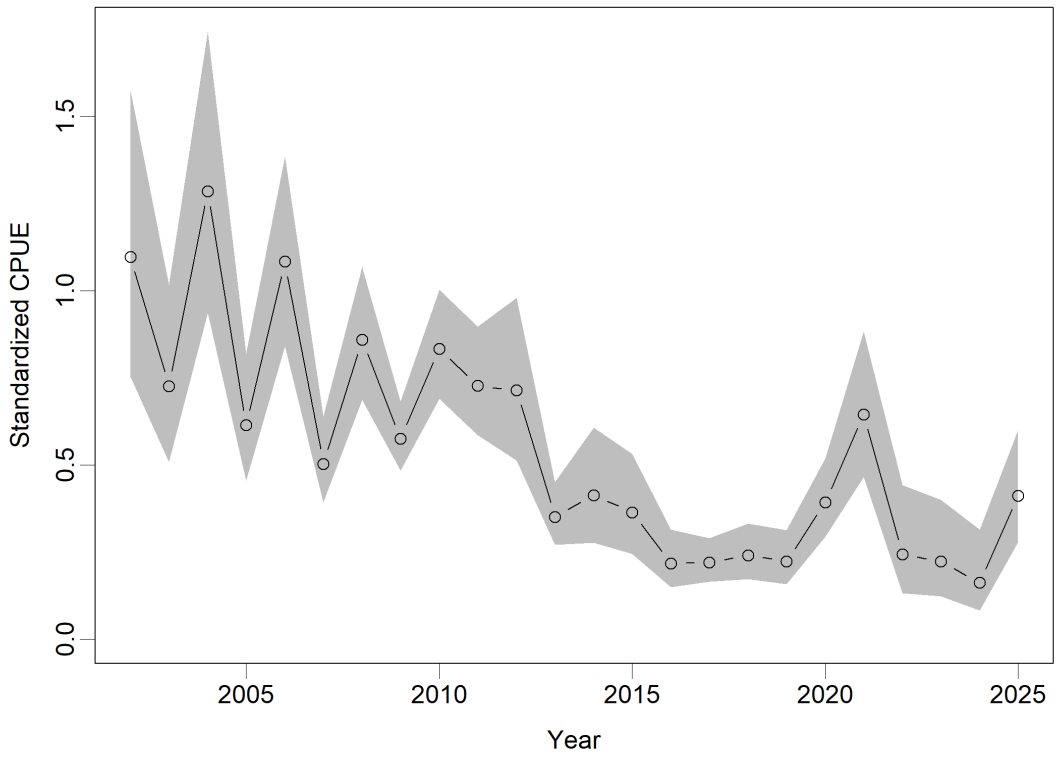
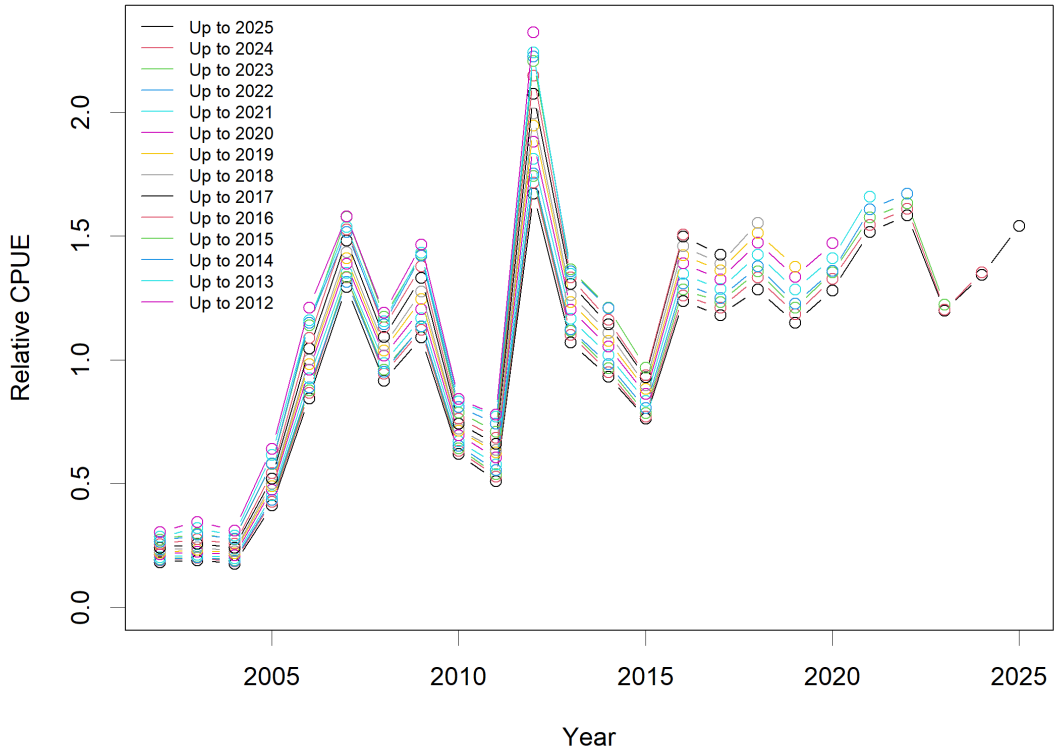



Fig. 17. Area-specific standardized CPUE of SBT caught by Taiwanese longline fishery. Shaded areas illustrate the 95% confidence intervals.

Area E




Area W

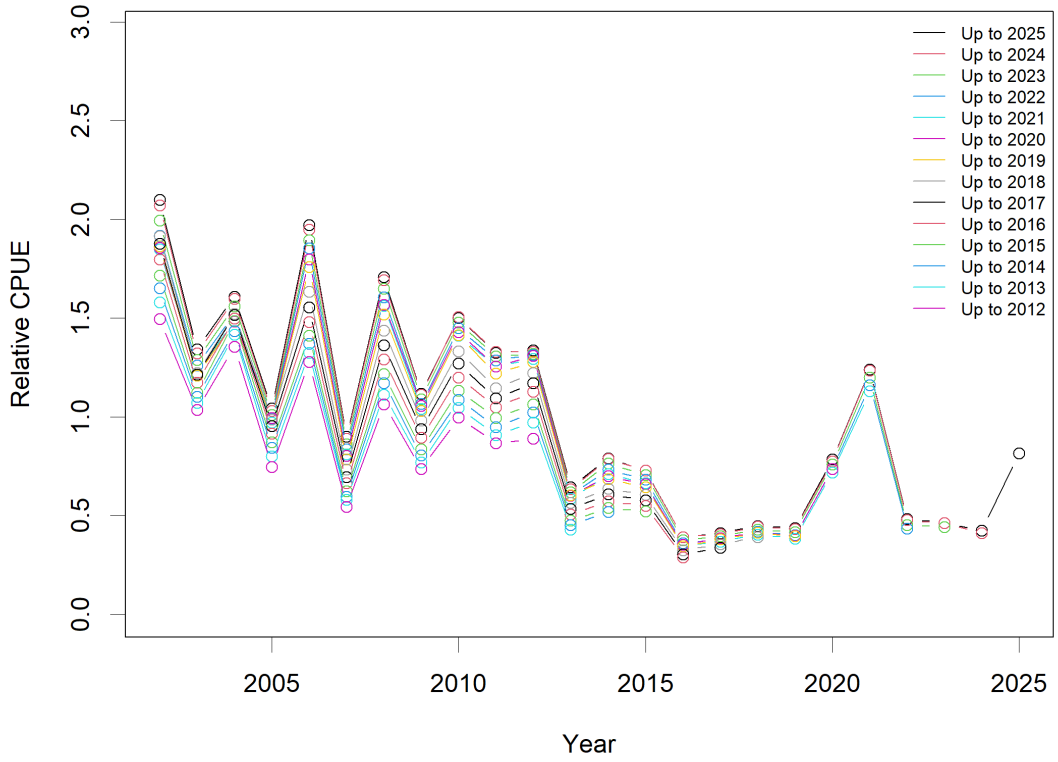



Fig. 18. The retrospect results of CPUE standardization based on including the updated data from different years.

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 | 2305 | 23 | 117.478 | < 2.2e-16 *** |
| M | 710 | 11 | 75.652 | < 2.2e-16 *** |
| G | 737 | 37 | 23.347 | < 2.2e-16 *** |
| C | 4839 | 3 | 1890.536 | < 2.2e-16 *** |
| NHBF | 48 | 2 | 28.343 | 4.983e-05 *** |
| Residuals | 45591 | 53435 | | |

Significant level: '***' 0.001 '**' 0.01 '*' 0.05

Area W

| Source of variance | SS | Df | F | Pr(>F) |
|---|---|---|---|---|
| Y | 710.5 | 23 | 33.3438 | < 2.2e-16 *** |
| M | 409.0 | 11 | 40.1406 | < 2.2e-16 *** |
| G | 145.6 | 22 | 7.1423 | < 2.2e-16 *** |
| C | 103.3 | 1 | 111.5178 | < 2.2e-16 *** |
| NHBF | 43.7 | 2 | 23.6078 | 6.074e-13 *** |
| Residuals | 6166.0 | 6656 | | |

Significant level: '***' 0.001 '**' 0.01 '*' 0.05

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 | 7343.1 | 23 | < 2.2e-16 *** |
| M | 5663.2 | 11 | < 2.2e-16 *** |
| G | 10991.1 | 41 | < 2.2e-16 *** |
| C | 7235.8 | 3 | < 2.2e-16 *** |
| NHBF | 96.7 | 2 | < 2.2e-16 *** |

Significant level: '***' 0.001 '**' 0.01 '*' 0.05

Area W

| Source of variance | LR Chisq | Df | Pr(>Chisq) |
|---|---|---|---|
| Y | 1798.4 | 23 | < 2.2e-16 *** |
| M | 4056.7 | 11 | < 2.2e-16 *** |
| G | 4404.4 | 27 | < 2.2e-16 *** |
| C | 244.2 | 1 | < 2.2e-16 *** |
| NHBF | 22.0 | 2 | 1.694e-05 *** |

Significant level: '***' 0.001 '**' 0.01 '*' 0.05