CPUE standardization for southern bluefin tuna caught by Taiwanese longline fishery for 2002-2017

Sheng-Ping Wang[1], Ching-Ping Lu[1], Shu-Ting Chang[2], An-Chiang Huang[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

In this study, the CPUE standardization analyses were conducted based on the data of Taiwanese longline fleets operated in the waters of the south of 20°S of the Indian Ocean during 2002-2017. An approach of cluster analysis was adopted to explore the targeting of fishing operations and to produce the data filter for selecting the data for CPUE standardizations. The targeting of fishing operation was identified from the cluster analyses performed based on the weekly-aggregated data instead of set-by-set data. For CPUE standardizations, a simple delta-lognormal models without interactions were adopted to avoid the confounding from interactions.

1. INTRODUCTION

Southern bluefin tuna (SBT) (*Thunnus maccoyii*) was by-catch of Taiwanese tuna longline fishery targeting albacore in the past, but after the fishing vessels equipped with deep-frozen freezers, some fishing vessels operating in the Indian Ocean started targeting SBT seasonally since the 1990s. Since Taiwanese SBT statistics system was reformed in 2002, the reporting rate of SBT catch has substantially improved since then (Anon, 2014). In this study, we attempted to explore the temporal and spatial patterns of catch and effort data of Taiwanese longline fishery operated in the waters of the south of 20°S of the Indian Ocean and also conduct the CPUE standardization for SBT caught by Taiwanese longline fishery for the year of 2002-2017.

2. MATERIALS AND METHODS

2.1. Catch and Effort data

In this study, operational catch and effort data with 5x5 degree fishing location grids of Taiwanese active longline vessels authorized to seasonally target SBT operating in the Indian Ocean in the period of 2002-2017 were provided by Overseas Fisheries Development Council of Taiwan (OFDC).

Based on the previous studies (e.g. Wang et al.,), the SBT fishing ground can be divided into the central-eastern area (Area E) and western area (Area W) (Fig. 1). In this study, all of the analyses were conducted based on this area stratification.

2.2. Cluster analysis

Based on the approach of Wang et al. (2015) and suggestions from CCSBT ESC meetings in 2015 and 2016, 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), southern bluefin tuna (SBT) and other species (OTH, most of the catches consisted of oilfishes) (Figs. 1 and 2). However, the 2016 CCSBT ESC considered that clustering operational set-by-set data might include large amount noise because most of SBT was caught by Taiwanese vessels as bycatches and only part of vessels targeted SBT for some fishing operations during the fishing seasons. In addition, ESC suggested that the cluster analysis could be conducted based on the aggregated data rather than the operational data sets. Therefore, the cluster analysis was performed based monthly- and weekly-aggregated data and then merged the clusters with operational data sets to identify the SBT fishing operations. However, the proportion of SBT catches substantially decreased for data sets and it was more difficult to identify the cluster that contained SBT fishing operations when conducting the cluster analysis based on monthly-aggregated data (Wang et al., 2017). Therefore, the cluster analysis was performed based on weekly-aggregated data.

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 (SBT sets and other tuna sets) 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 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 a large amount of zero SBT catch occurred in the fishing sets, the delta-lognormal models were applied to standardize the CPUE of SBT caught by Taiwanese longline fishery. As suggested in 2016 ESC, main effects of year, month, 5x5 grid and number of hooks between float were 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. In addition, the effects of cluster and number of hooks between float (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 float,
	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 (regular: <=9 hooks; deep: 10-14 hooks; ultra deep: >=15 hooks) (Wang and Nishida, 2011).

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

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

Three clusters were selected for the Area E (Fig. 4). Cluster 1 mainly belonged to the ALB operations but also contained the operations for BET, SBT and OTH; Cluster 2 was the ALB operation; Cluster was the SBT operation (Figs. 5 and 6). Although the highest SBT catch proportion was occurred in Cluster 3, most of the SBT catches were contained in Cluster 1 (Fig. 7). For SBT Cluster (Cluster 3), the data mainly consisted of the data in the early 2000s; fishing mainly operated during June and September; NHBF concentrated at around 10 hooks; and the operations also concentrated in the waters between 30°S and 35°S (Fig. 8). Based on the spatial distribution of SBT catch proportion, the SBT catch proportion of Cluster 3 was obviously higher than those of other two Clusters (Fig. 9).

For Area W, we also selected three clusters (Fig. 10). Cluster 1 was ALB operations; Cluster 2 mainly belonged to the ALB operations but also contained the operations for BET, YFT, SWO and OTH; Cluster 3 consisted of operations for OTH (mainly for oilfish) (Figs. 11 and 12). Most SBT catches were contained in Cluster 2 and Cluster 3, and Cluster 1 contained very few SBT catches (Fig. 13). For Clusters 2 and 3, Cluster 2 mainly consisted of the data before 2010, while the data of Cluster 3 were mainly after 2010; NHBF of Cluster 3 was more than that of Cluster 2; fishing areas (longitude and latitude) were also different for Clusters 2 and 3 (Fig. 14). For the spatial distribution of SBT catch proportion, the SBT catch proportion of Cluster 2 was higher than those of other two Clusters (Fig. 15).

3.2 CPUE standardization

For Area W, the data of Cluster 1 were excluded when conducting the CPUE standardizations because very few SBT catches were contained in the Cluster. Clusters 2 and 3 still contained 97.7% SBT catches and 88.5% hooks.

For both of Areas E and W, the models with the lowest value of AIC were selected as the final models. The ANOVA tables for the lognormal models are shown in Table 1. All of the effects were statistically significant for both areas. About 19% 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 37% and 27% of CPUE variances were explained by the models for Area E and Area W, respectively.

Fig. 17 shows the area-specific standardized CPUE trends. Standardized CPUE series generally revealed quite different trends for two areas. For Area E, the standardized CPUEs gradually increased before 2007, revealed decreasing trend from 2007 to 2011, substantially increased in 2012 and then gradually decreased until 2015, and increased again in recent two years. For Area W, the standardized CPUE series generally revealed a decreasing trend with a fluctuation since 2002.

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 Area E, while including updated data changed the standardized CPUE series for Area W although the trends were similar (Fig. 18).

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.

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

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., Nishida, T., 2011. CPUE standardization of swordfish (*Xiphias gladius*) caught by Taiwanese longline fishery in the Indian Ocean. IOTC-2011-WPB09-12.

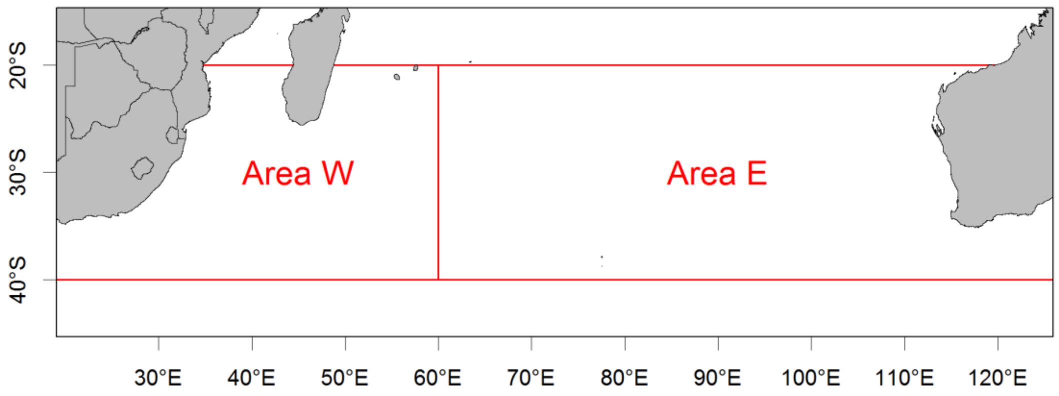


Fig. 1. Area stratification for southern bluefin tuna of Taiwanese large scale longline fishery in the Indian Ocean.

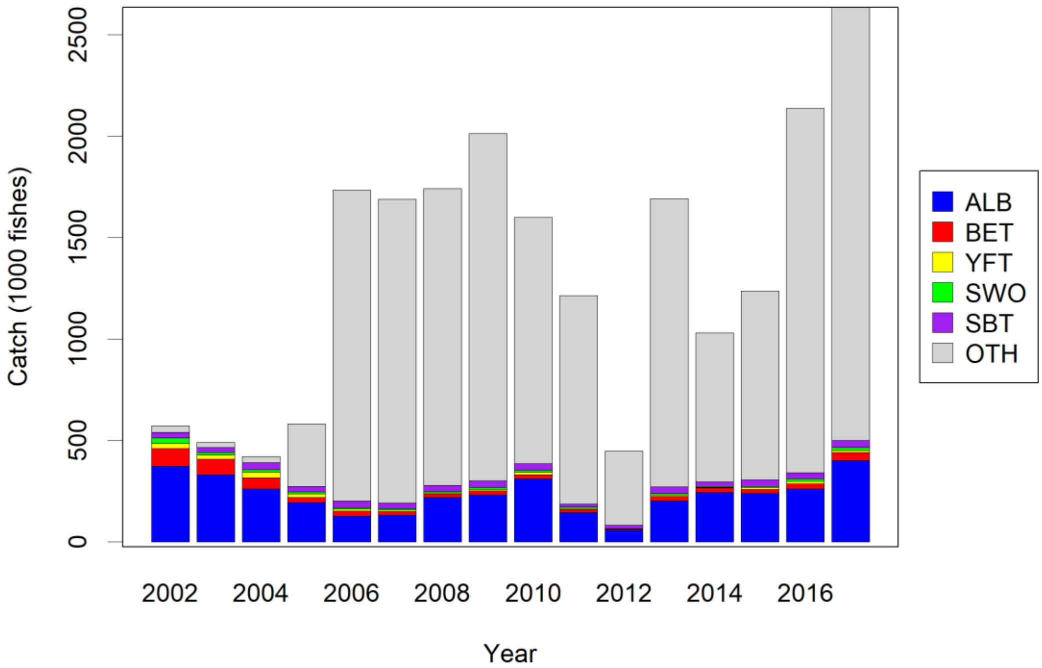


Fig. 2. Annual catch composition of Taiwanese longline fleets operated in the waters of south of 20°S.

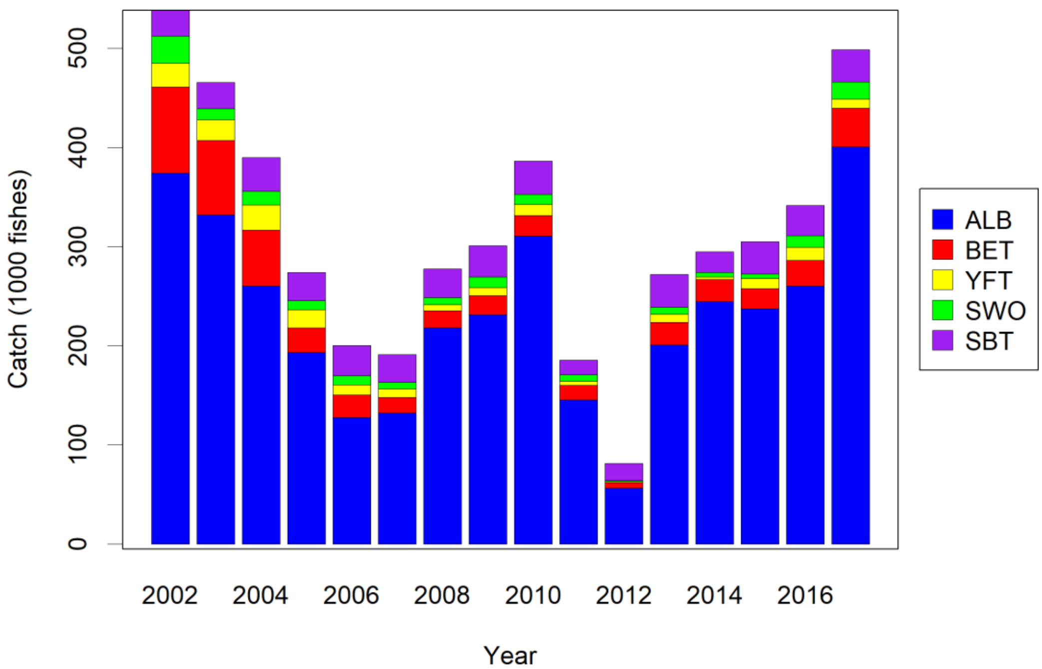


Fig. 3. Annual catch composition of Taiwanese longline fleets operated in the waters of south of 20°S. The catches of OTH are excluded.

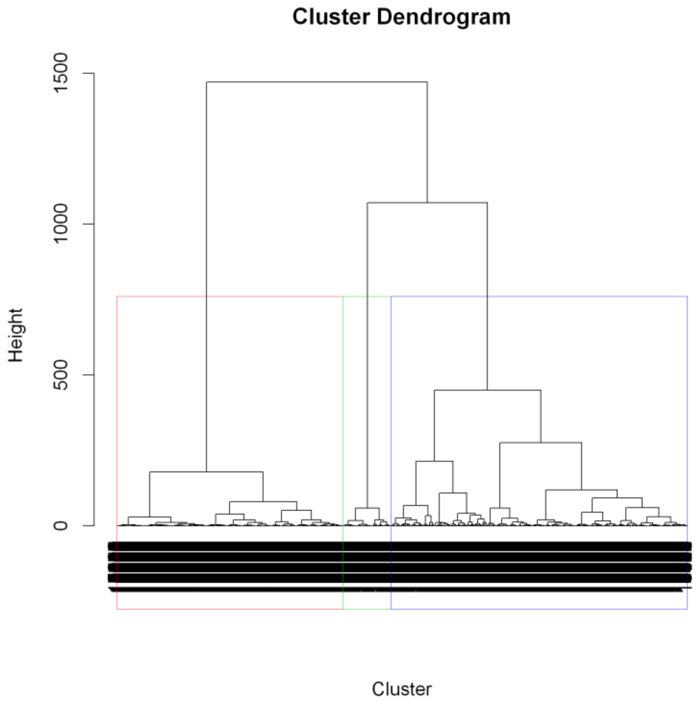


Fig. 4. Cluster tree for the data of Taiwanese large scale longline fishery in Southern Bluefin Tuna Area E of the Indian Ocean.

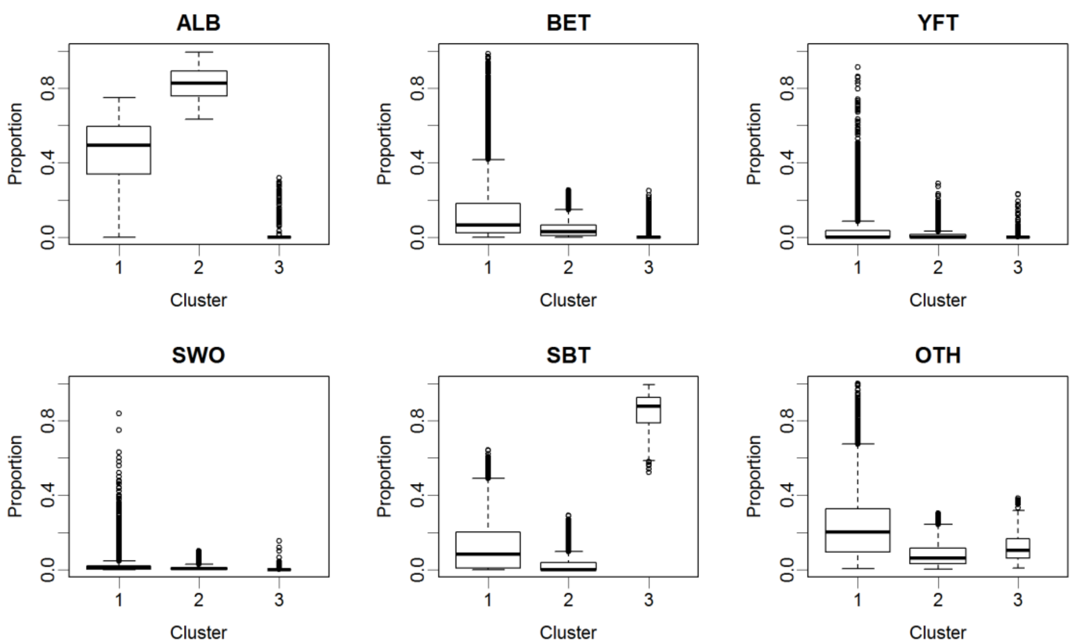


Fig. 5. Catch proportion by species for each cluster of Taiwanese large scale longline fishery in Southern Bluefin Tuna Area E of the Indian Ocean.

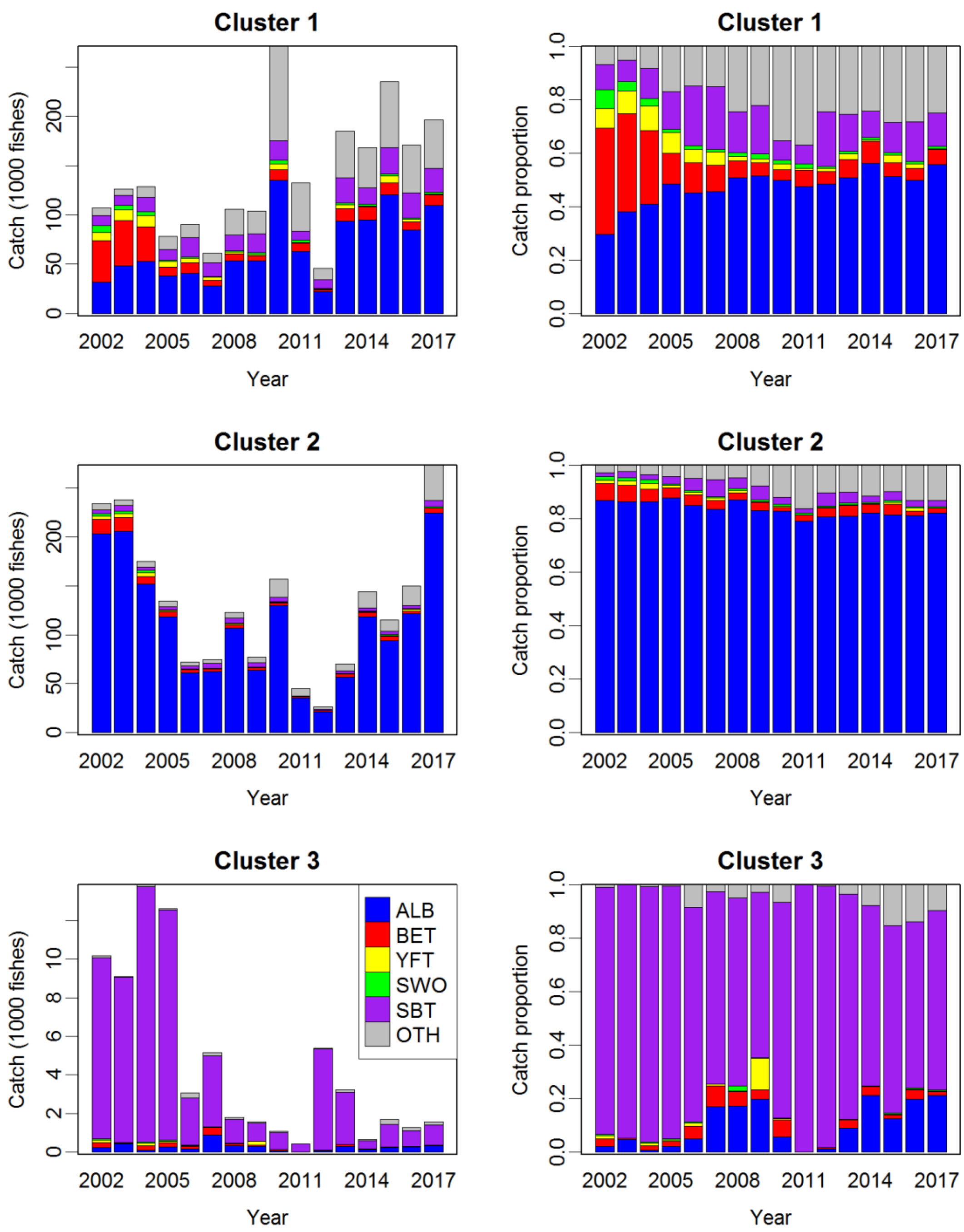
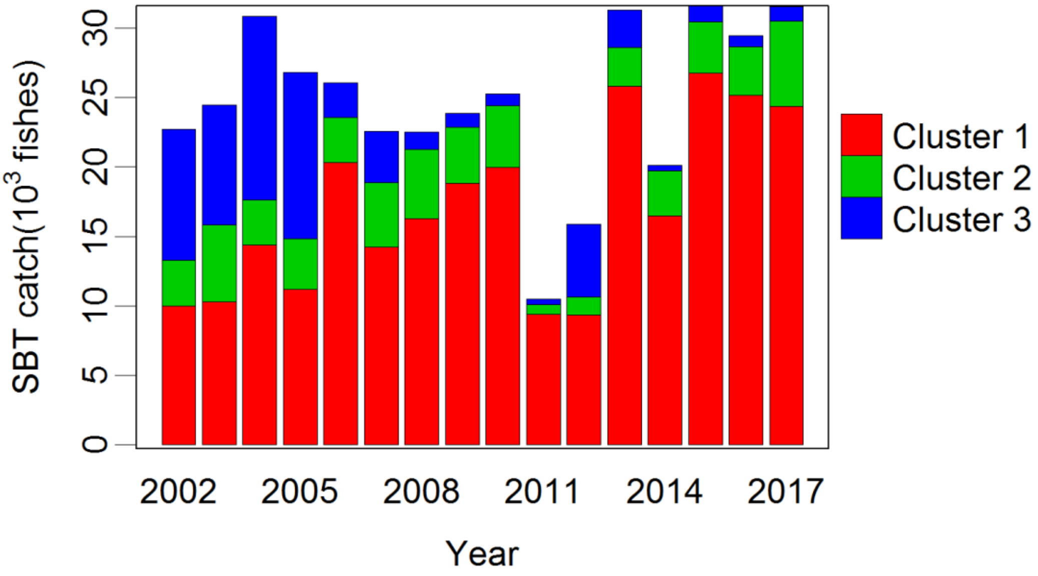


Fig. 6. Annual and catch proportion by species for each cluster of Taiwanese large scale longline fishery in Southern Bluefin Tuna Area E of the Indian Ocean.



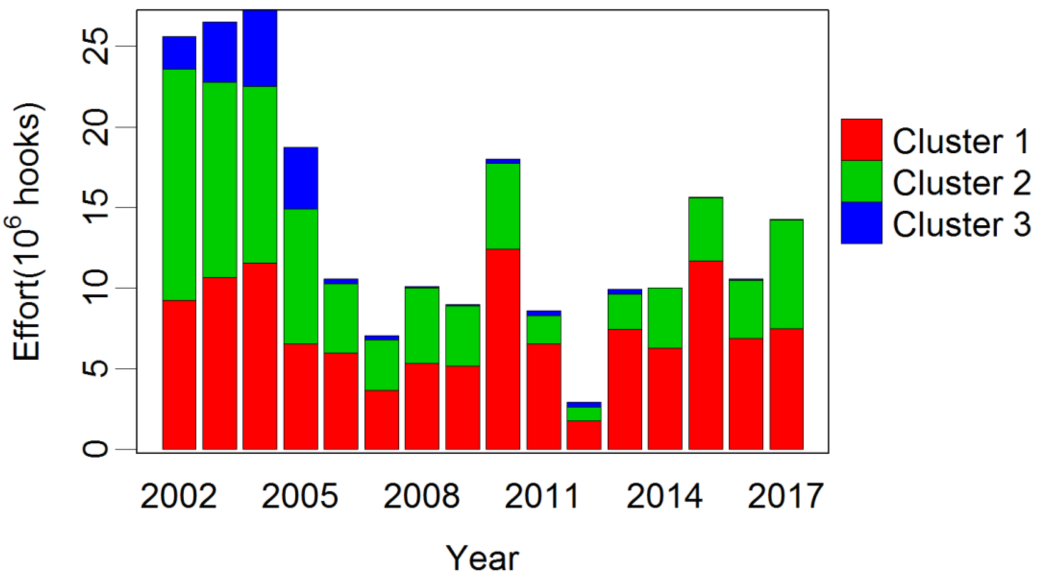


Fig. 7. Annual Southern Bluefin Tuna catches and efforts for each cluster of Taiwanese large scale longline fishery in Southern Bluefin Tuna Area E of the Indian Ocean.

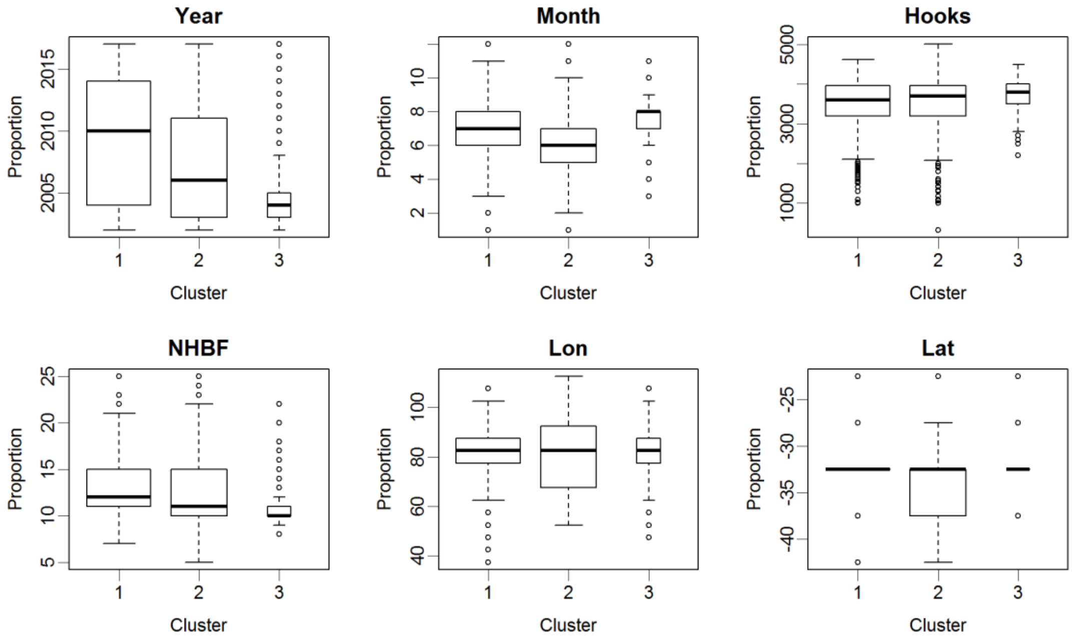


Fig. 8. Data composition by factors for each cluster of Taiwanese large scale longline fishery in Southern Bluefin Tuna Area E of the Indian Ocean.

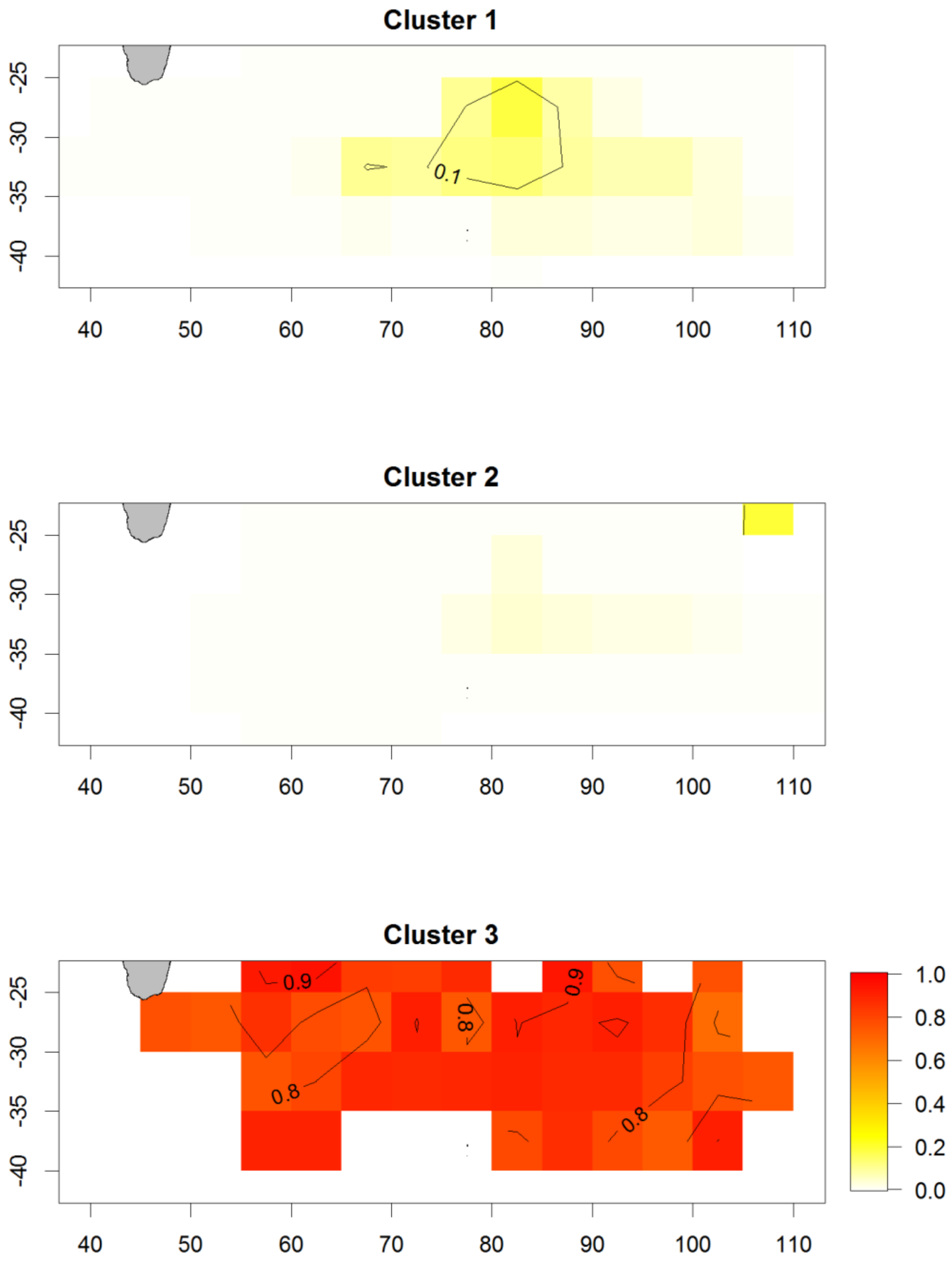


Fig. 9. Southern Bluefin Tuna catch distribution for each cluster of Taiwanese large scale longline fishery in Southern Bluefin Tuna Area E of the Indian Ocean. Yellow is low catch and red is high catch.

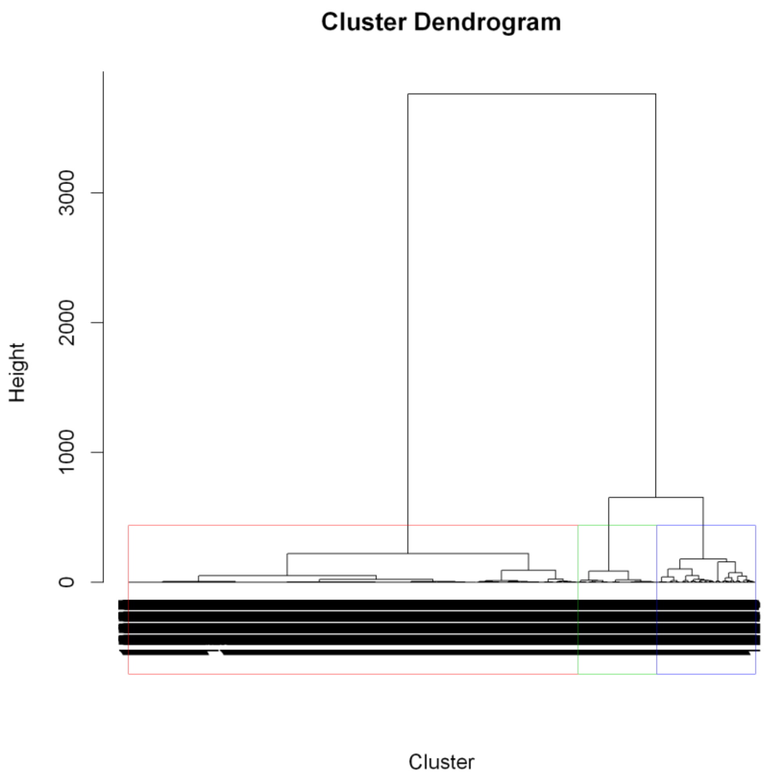


Fig. 10. Cluster tree for the data of Taiwanese large scale longline fishery in Area W of the Indian Ocean.

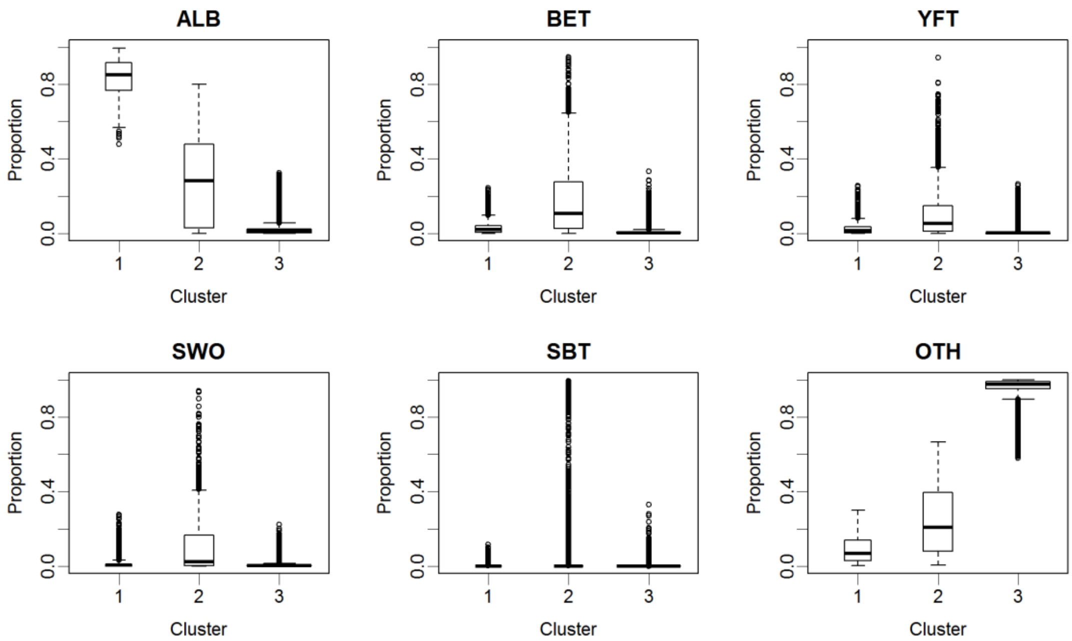


Fig. 11. Catch proportion by species for each cluster of Taiwanese large scale longline fishery in Area W of the Indian Ocean.

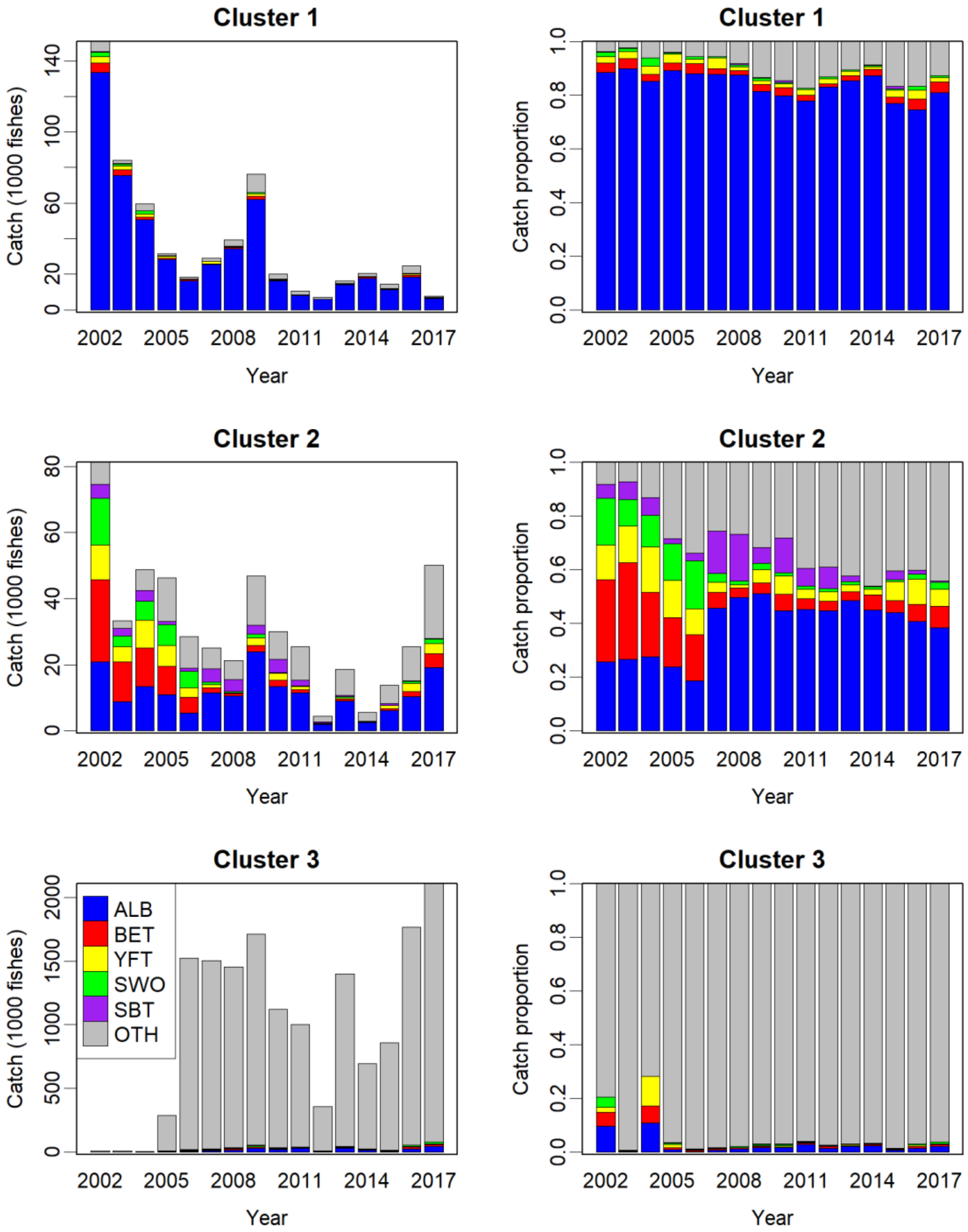
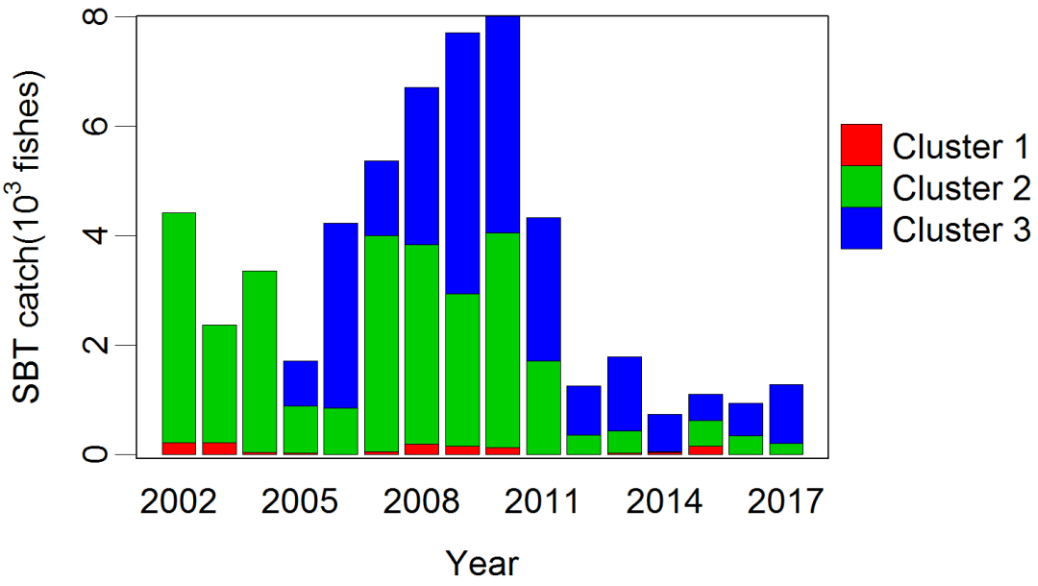


Fig. 12. Annual and catch proportion by species for each cluster of Taiwanese large scale longline fishery in Area W of the Indian Ocean.



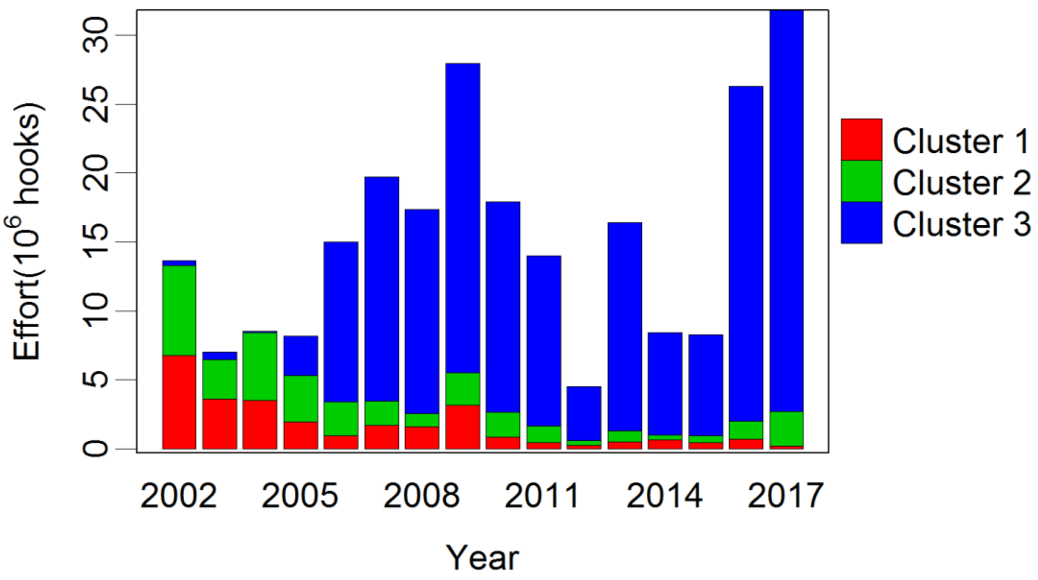


Fig. 13. Annual Southern Bluefin Tuna catches and efforts for each cluster of Taiwanese large scale longline fishery in Area W of the Indian Ocean.

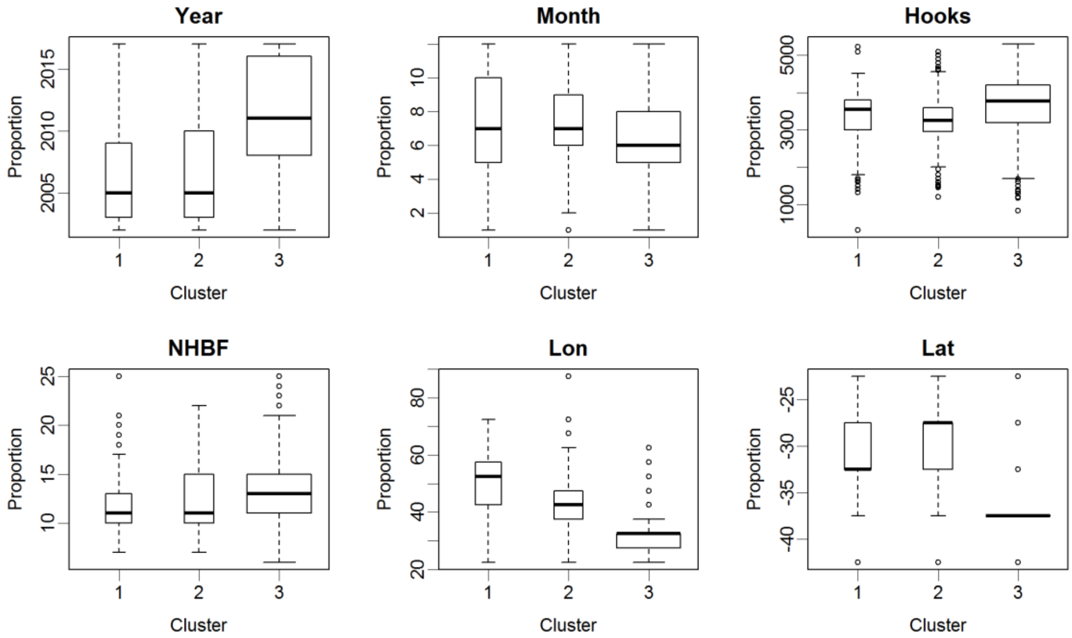


Fig. 14. Data composition by factors for each cluster of Taiwanese large scale longline fishery in Area W of the Indian Ocean.

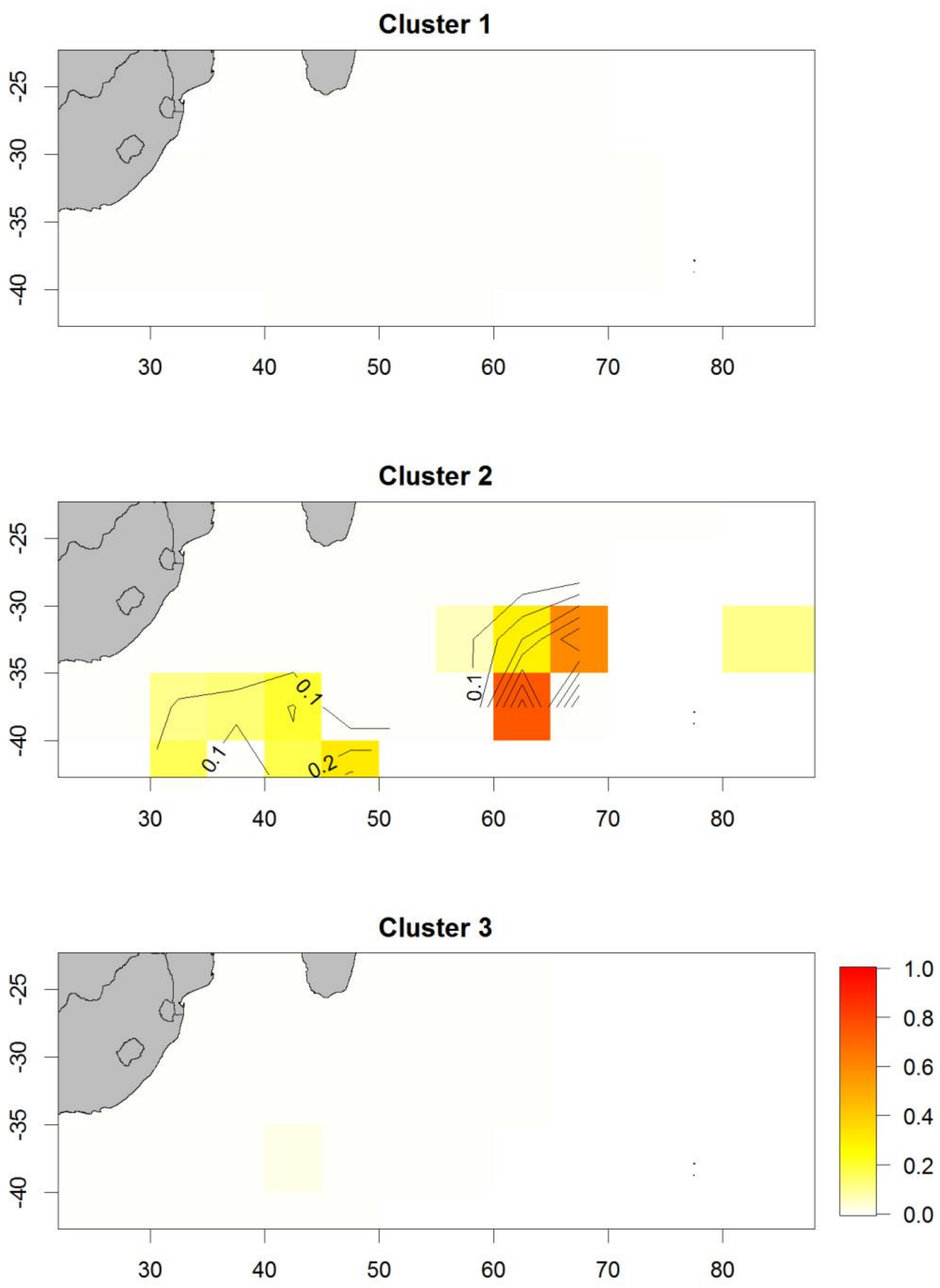
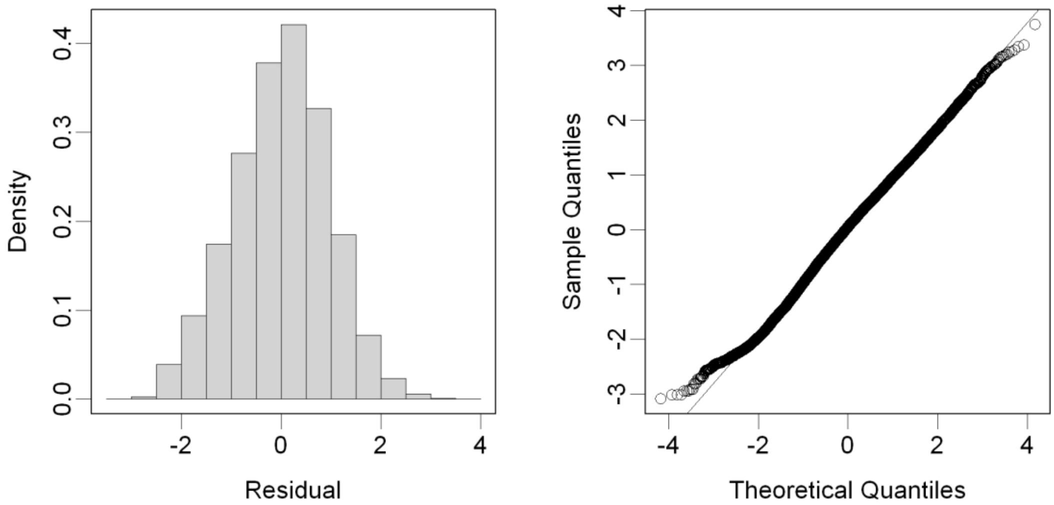


Fig. 15. Southern Bluefin Tuna catch distribution for each cluster of Taiwanese large scale longline fishery in Area W of the Indian Ocean. Yellow is low catch and red is high catch.

Area E



Area W

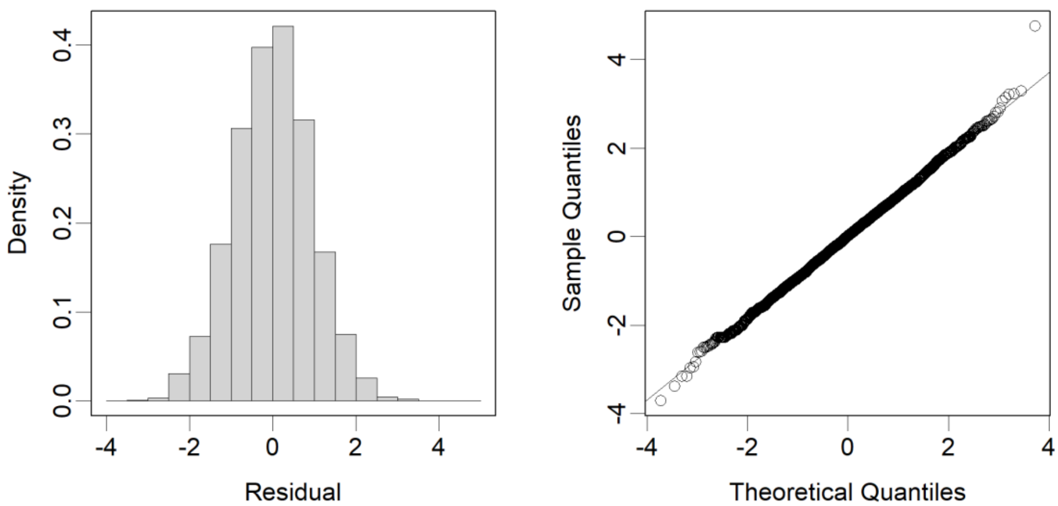
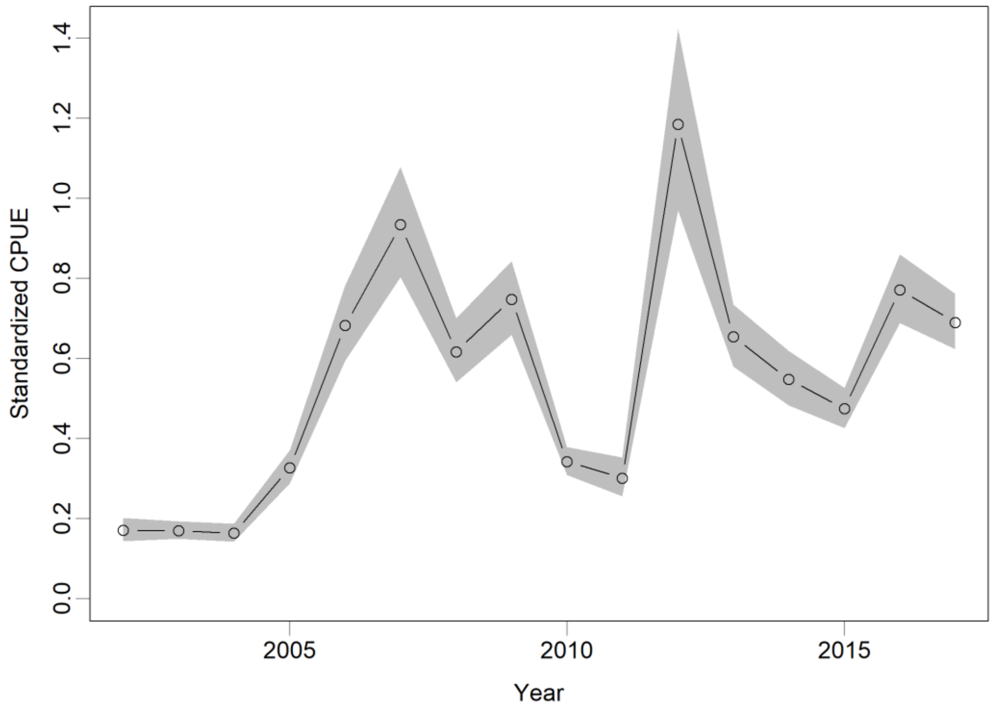


Fig. 16. The frequency distributions and Quantile-Quantile Plots for standardized residuals obtained from lognormal models for Area E and Area W.

Area E



Area W

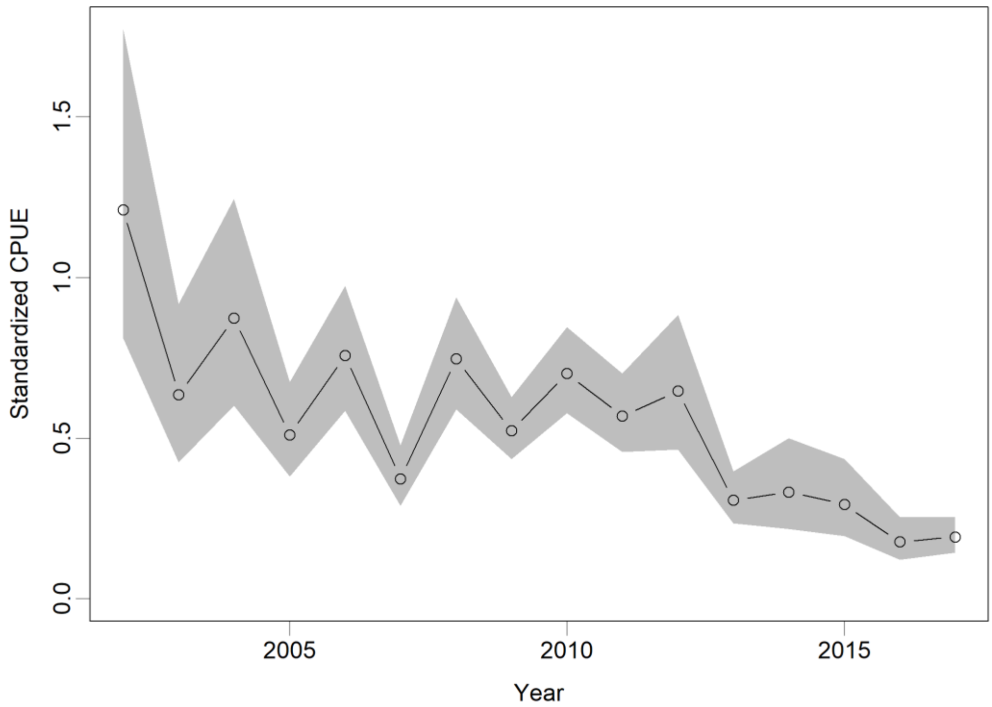
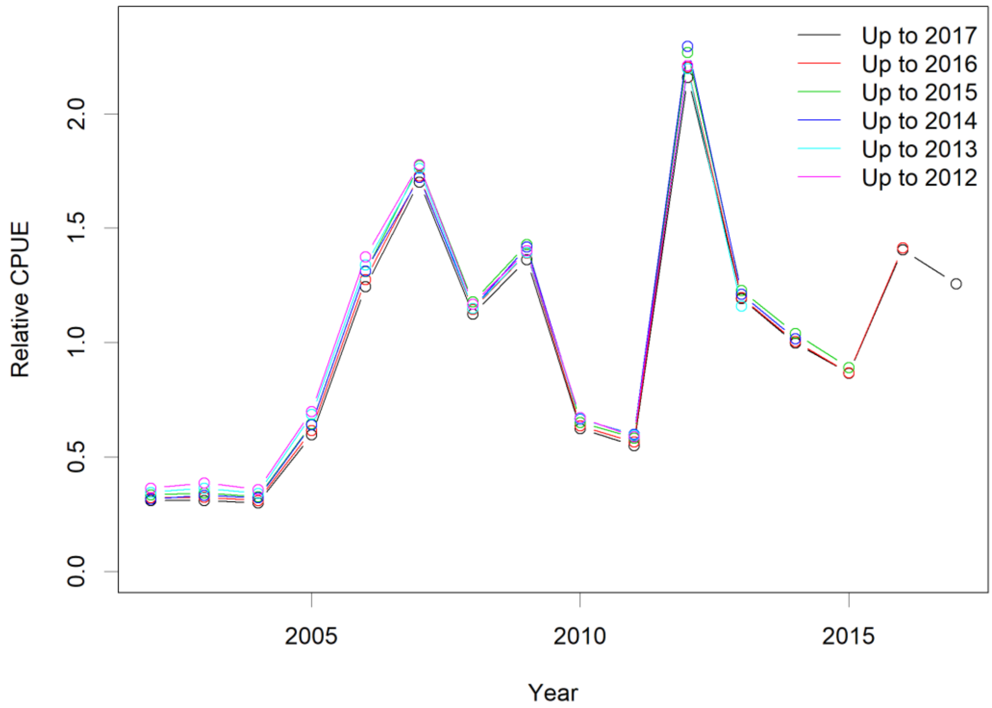


Fig. 14. Area-specific standardized CPUE of southern bluefin tuna caught by Taiwanese longline fishery. Shaded areas show the 95% confidence intervals.

Area E



Area W

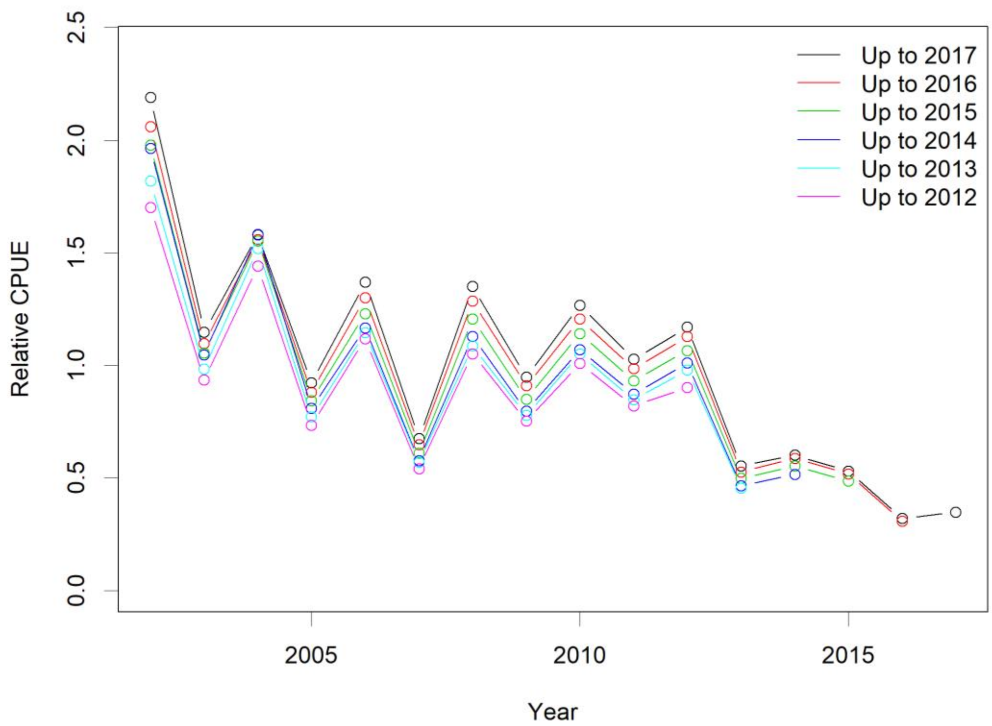


Fig. 18. The results of CPUE standardization based on including the updated data from different years.

Table 1. ANOVA tables for the lognormal models for Area E and Area W.

Area E

Source of variance	SS	Df	F	Pr(>F)
Y	1980.1	15	145.301	< 2.2e-16 ***
M	488.2	9	59.704	< 2.2e-16 ***
CT	230.3	3	84.508	< 2.2e-16 ***
G	766.9	45	18.757	< 2.2e-16 ***
C	871.5	2	479.649	< 2.2e-16 ***
NHBF	32.5	2	17.866	1.76E-08 ***
Residuals	31285.2	34435		

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

Area W

Source of variance	SS	Df	F	Pr(>F)
Y	545.1	15	41.5251	< 2.2e-16 ***
M	288.6	10	32.9745	< 2.2e-16 ***
CT	80.2	2	45.8214	< 2.2e-16 ***
G	94.3	35	3.0793	3.18E-09 ***
C	49.5	1	56.5288	6.48E-14 ***
NHBF	35	2	20.0037	2.22E-09 ***
Residuals	4592.1	5247		

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

Table 2. ANOVA tables for the delta models for Area E and Area W.

Area E

Source of variance	LR Chisq	Df	Pr(>Chisq)
Y	5338.7	15	< 2.2e-16 ***
M	4292.2	11	< 2.2e-16 ***
CT	155.8	3	< 2.2e-16 ***
G	3914.3	58	< 2.2e-16 ***
C	6123	2	< 2.2e-16 ***
NHBF	115.5	2	< 2.2e-16 ***

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Area W

Source of variance	LR Chisq	Df	Pr(>Chisq)
Y	527.4	15	< 2.2e-16 ***
M	2977	11	< 2.2e-16 ***
CT	8.8	2	0.0124086 *
G	3476.7	43	< 2.2e-16 ***
C	447.3	1	< 2.2e-16 ***
NHBF	16.2	2	0.0003091 ***

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1