CPUE standardization for southern bluefin tuna caught by Taiwanese longline fishery

Sheng-Ping Wang[1], I-Lu Lai[2], Shu-Ting Chang[3], Shiu-Ling Lin[2]

[1] National Taiwan Ocean University, Taiwan
[2] Fisheries Agency, Council of Agriculture, Executive Yuan, Taiwan
[3] Overseas Fisheries Development Council, Taiwan

ABSTRACT

In this study, cluster analysis was conducted to identify the characteristic of fishing operating for each data set based on catch compositions of Taiwanese longline fishery. The results of cluster analysis were used to be a criterion for extracting data from SBT targeting vessels and also to be a targeting effect in the model. CPUE standardization was conducted using a general linear model. Although standardized CPUE series reveal different trends for different area, they roughly decreased for all areas in recent years and substantially increased for all areas in 2015. CPUE series for fishes with age of 3-5 years were much higher than other age groups, while obvious decline trends were observed in recent years for most age groups. Similarly, CPUE series substantially increased for all age groups in 2015.

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 1990s. Since Taiwanese SBT statistics system was reformed in 2002, the reporting rate of SBT catch has substantially improved since then (Anon, 2014). This study attempted to conduct the CPUE standardization for SBT caught by Taiwanese longline fishery for year of 2002-2015.

2. MATERIALS AND METHODS

2.1. Catch and Effort data

In this study, monthly 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 2002-2015 were provided by Overseas Fisheries Development Council of Taiwan (OFDC).

2.2. Definition of fishing areas

Although two fishing ground (Area1, in the area of 20°S-40°S and east of 50°E; Area 2, in the area of 20°S-45°S and 20°E-50°E) should be appropriate to Taiwanese SBT longline fishery (Anon, 2013), the ESC17 indicated that "current area stratification may be appropriate for the Taiwanese data, but that if the spatial strata were the CCSBT statistical areas then comparisons could be made with the other longline CPUE indices (CCSBT, 2012)." Therefore, the CCSBT statistical areas were adopted for the analysis of Taiwanese CPUE standardization (Fig. 1).

2.3. Vessel selection

Based on the approach of Wang et al. (2015) that was discussed in 2015 CCSBT ESC20, the cluster analysis (He et al., 1997) was adopted to conduct vessel selection and the data of selected vessels were used to conduct CPUE standardization. Cluster analysis was conducted based on species composition of the catches. Five species groups were used in this study, including albacore (ALB), bigeye tuna (BET), yellowfing tuna (YFT), swordfish (SWO) and southern bluefin tuna (SBT). He et al. (1997) suggested a cluster analysis with two steps to classify the data sets because the large number of data sets precluded direct hierarchical cluster analysis. First, a non-hierarchical cluster analysis (K-means method) was used to group all data sets into 10 clusters for taking the mixture of fishing operations into account ($C_2^5 = 10$ ways in which 2 species can be chosen from 5 species groups). Second, a hierarchical cluster analysis with Ward minimum variance method was applied to the squared Euclidean distances calculated from 10 non-hierarchical clusters. Non-hierarchical and hierarchical cluster analyses were conducted using R functions kmeans and hclust (The R Foundation for Statistical Computing Platform, 2014). 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.

Then we calculated the proportion of fishing sets which was designated as SBT cluster for each vessel and for each year. The quartile of SBT cluster proportion for each year was used to be the criteria for vessel selection. If a vessel had no fishing set designated as SBT cluster in a year, the data of this vessel were excluded from the calculation of quartile.

2.4. CPUE standardization

The general linear model (GLM) was applied to standardize the CPUE of SBT caught by Taiwanese longline fishery. The effects included in the models were year, month, fishing area, longitude, operation cluster, and their interactions. The latitude effect was excluded from the GLM because too many missing values were occurred in latitude strata. The GLM was conducted as below:

$$\ln(CPUE + c) = \mu + Y + M + A + Lon + C + \text{interactions} + \varepsilon$$

where	$CPUE$	is the nominal CPUE of SBT (catch in number/1,000 hooks),
	c	is the constant value (i.e. 10% of the average nominal CPUE),
	μ	is the intercept,
	Y	is the year effect,
	M	is the month effect,
	A	is the fishing area effect,

Lon	is the longitude effect,
C	is the fishing operation cluster effect,
ε	is the error term, $\varepsilon \sim N(0, \sigma^2)$.

The effects of year, month, area and operation cluster were treated as categorical variables. As suggested in 2015 CCSBT ESC, the effect of longitude was changed from continuous variable to be categorical variables. Regarding the effect of interaction related to year effect, an interaction between year and area was only included in this study for the further estimates of the area-specific CPUE standardization because interactions with the year effect would lead to problems for the year effect as an index of abundance (Hinton and Maunder, 2004; Maunder and Punt, 2004).

The area-specific standardized CPUE trends were estimated based on the exponentiations of the adjust means (least square means) of the interaction between year and area effects (i.e. $Y \times A$) (Butterworth, 1996; Maunder and Punt, 2004).

The age-specific CPUE standardization was also conducted based on GLM. Ages were grouped into 0-2, 3-5, 6-9 and 10+ years. The GLM was conducted as below:

$$\ln(CPUE + c) = \mu + Y + M + A + Lon + C + Age + \text{interactions} + \varepsilon$$

where Age is the age effect.

Because the age-specific catches data did not occur in every areas and years, we did not attempt to estimate age- and area-specific standardized CPUE. The age-specific standardized CPUE trends were estimated based on the exponentiations of the adjust means of the interaction between year and age effects (i.e. $Y \times Age$).

The model selection is based on the Akaike information criterion (AIC) and the estimations of the models were performed using R with glm() and lsmeans() functions.

3. RESULTS AND DISCUSSIONS

3.1. Cluster analysis for fishing operation

Based on the results of non-hierarchical cluster analysis (K-means method), the catch proportion of SBT is highest in cluster 8, which is 77% (Table 1). Then hierarchical cluster analysis was conducted based on the catch proportion by species obtained from K-means clusters. In this study, the number of clusters was decreased until a cluster still contained 77% SBT catch proportion. Finally, four clusters were selected (Fig. 2) and Table 2 shows the average proportions of catches by species. Figs. 3 and 4 show the annual catches and catch compositions, and SBT catch obviously occurred in the operation sets designated as cluster 4, which was also designated as SBT cluster.

3.2. Vessel selection

Fig. 5 shows the box-plot (quartile) of SBT cluster proportion for each year. The proportions of fishing sets belonged to SBT cluster were generally higher before 2005 but peak was also observed in 2014. It should be noted that the data in 2015 are preliminary and incomplete.

Tables 3-5 show the numbers of vessels, proportion of SBT catches and proportion of efforts of vessels selected based on the criteria of 1st (core25), 2nd (core50) and 3rd (core75) quartiles. The results indicated that the numbers of vessels and efforts can be substantially reduced for three selected criteria. The SBT catches can maintain at about 55-96% of total catches and most annual efforts can be reduced to less than about 50% of total efforts when the criterion of the 1st quartile (core25) was used. This indicated that exclusive vessels spent large amount of efforts on not catching SBT. Although the efforts were further decreased when the criterion of the 2nd quartile (core50) or the 3rd quartile (core75) were used, the catches were also substantially reduced. Therefore, the 1st quartile (core25) was finally adopted as the criterion for vessel selection.

3.3. Summary of GLM statistics

3.3.1. Area-specific model

The final model was selected based on the lowest value of AIC for models with various combinations of effects and interactions. In addition, considering the missing values were occurred for strata of interactions, the final GLM was selected as

$$\ln(CPUE + c) = \mu + Y + M + A + Lon + C + Y \times A + M \times C + A \times C$$

The ANVOA table for the final GLM is shown in Table 6 and all main effects and interactions were statistically significant. The model can explain 55% of CPUE variance. The distribution of standardized residuals obviously concentrates around 0 and the Quantile-Quantile Plot also indicates that the distribution of residuals fits to the assumption of normal distribution (Fig. 6).

3.3.2. Age-specific model

Similar to area-specific model, the final GLM for age-specific CPUE was selected as

$$\ln(CPUE + c) = \mu + Y + M + A + Lon + C + AG + Y \times AG + M \times Lon + A \times C + A \times AG + C \times AG + \varepsilon$$

The ANVOA table for the final GLM is shown in Table 7 and all main effects and interactions were statistically significant. The model can explain 52% of CPUE variance. However, the distribution of standardized residuals obviously concentrates around 0 and the Quantile-Quantile Plot also indicates that the distribution of residuals may be slightly unfit to the normal distribution (Fig. 7).

3.4. Trend of standardized CPUE

Fig. 8 shows the area-specific standardized CPUE trends estimated based on incorporating the definition of CCSBT statistical areas (Fig. 1). Standardized CPUEs generally reveal quite different trends in different areas. For Areas 2 and 8, the standardized CPUEs roughly revealed increasing trends before about 2010 and decreased in recent years. For Area 2, however, a peak was observed for CPUE in 2012. The standardized CPUEs in Area 14 increased before 2007 and gradually decreased thereafter. The trend of standardized CPUE in Area 9 was relatively stable. In 2015, standardized CPUEs substantially increased for all areas.

Based on age-specific standardized CPUEs (Fig. 9), CPUEs for fishes with age of 3-5 years were much higher than other age groups. Except for age 10+ group, however, CPUEs revealed similar trends among age groups, which increased before 2006, fluctuated during 2007-2012, and substantially decreased in recent years. In 2015, CPUE series substantially increased for all age groups.

REFERENCE

Anon., 2013. Review of Taiwan SBT Fishery of 2011/2012. CCSBT-ESC/1309/SBT Fisheries-Taiwan.

Anon., 2014 Preparation of Taiwan's Southern bluefin tuna catch and effort data submission for 2014. CCSBT-ESC/1409/45.

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.

Hinton, M., Maunder, M.N., 2004. Methods for standardizing CPUE and how to select among them. ICCAT Col. Vol. Sci. Pap., 56: 169-177.

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

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.

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

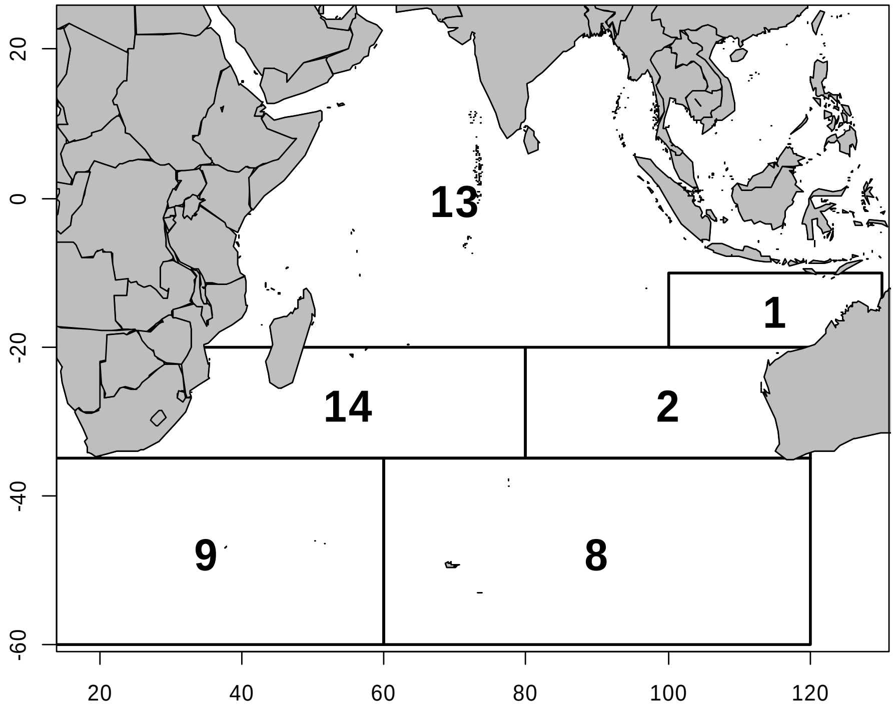

Fig. 1. The definition of CCSBT statistical areas.

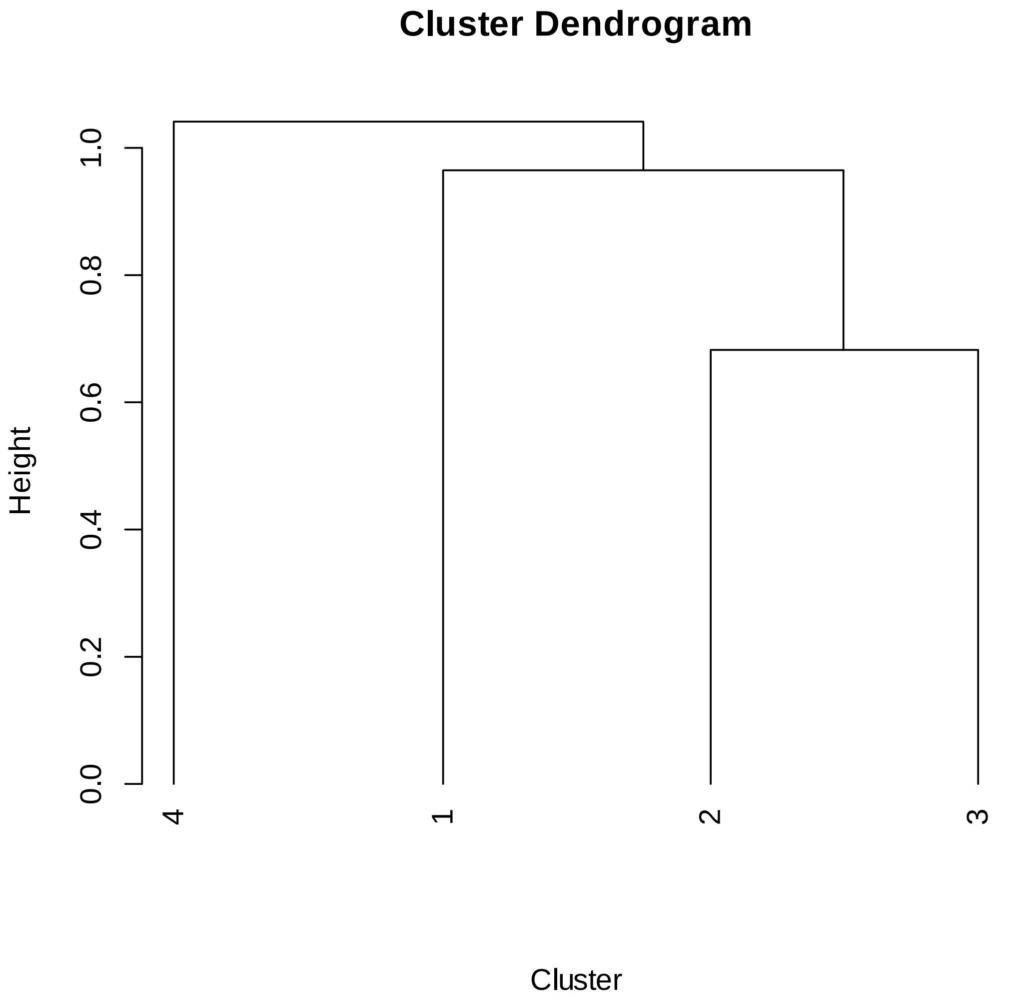

Fig. 2. The dendrogram of hierarchical cluster analysis for classifying the data sets of Taiwanese southern bluefin tuna longline fishery in the Indian Ocean.

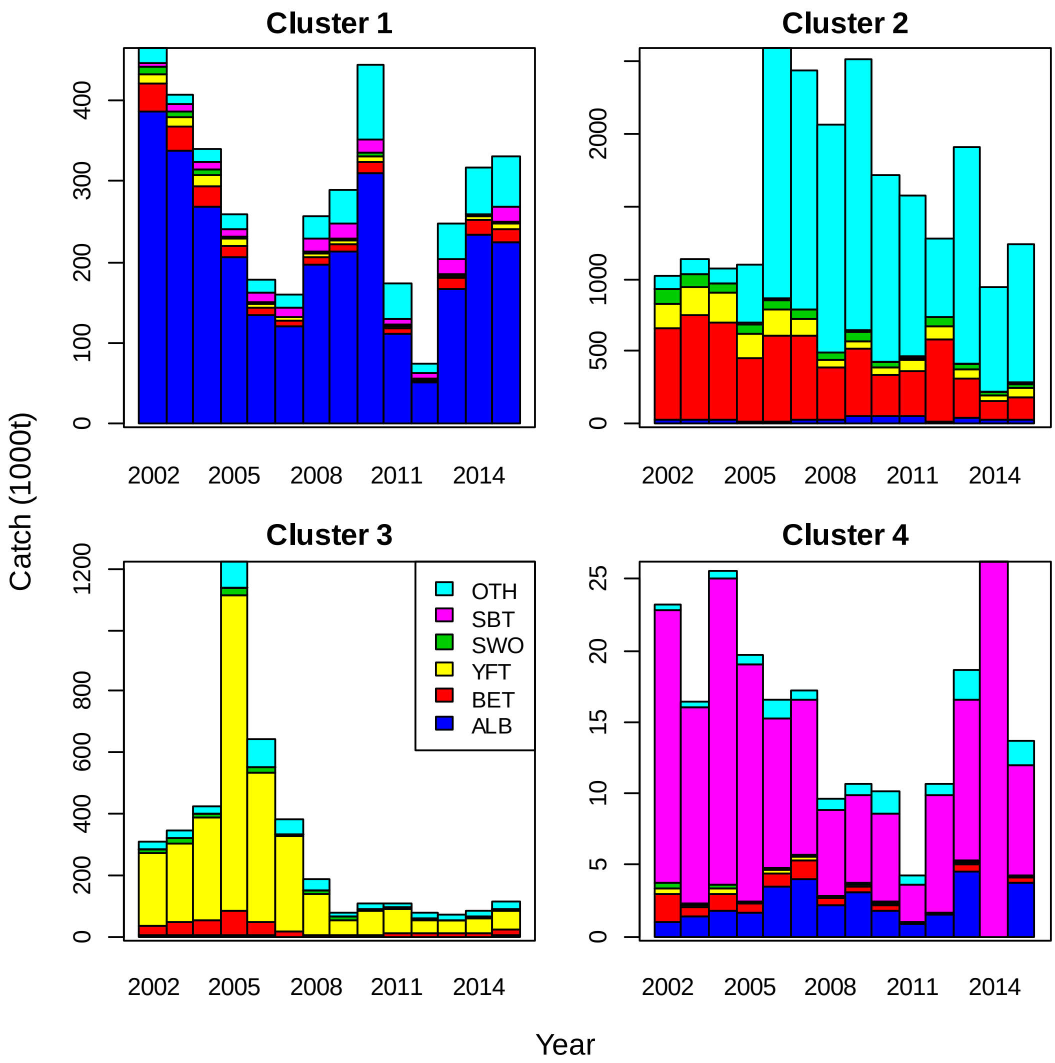

Fig. 3. Annual catches by species of Taiwanese southern bluefin tuna longline fishery in the Indian Ocean for nine clusters.

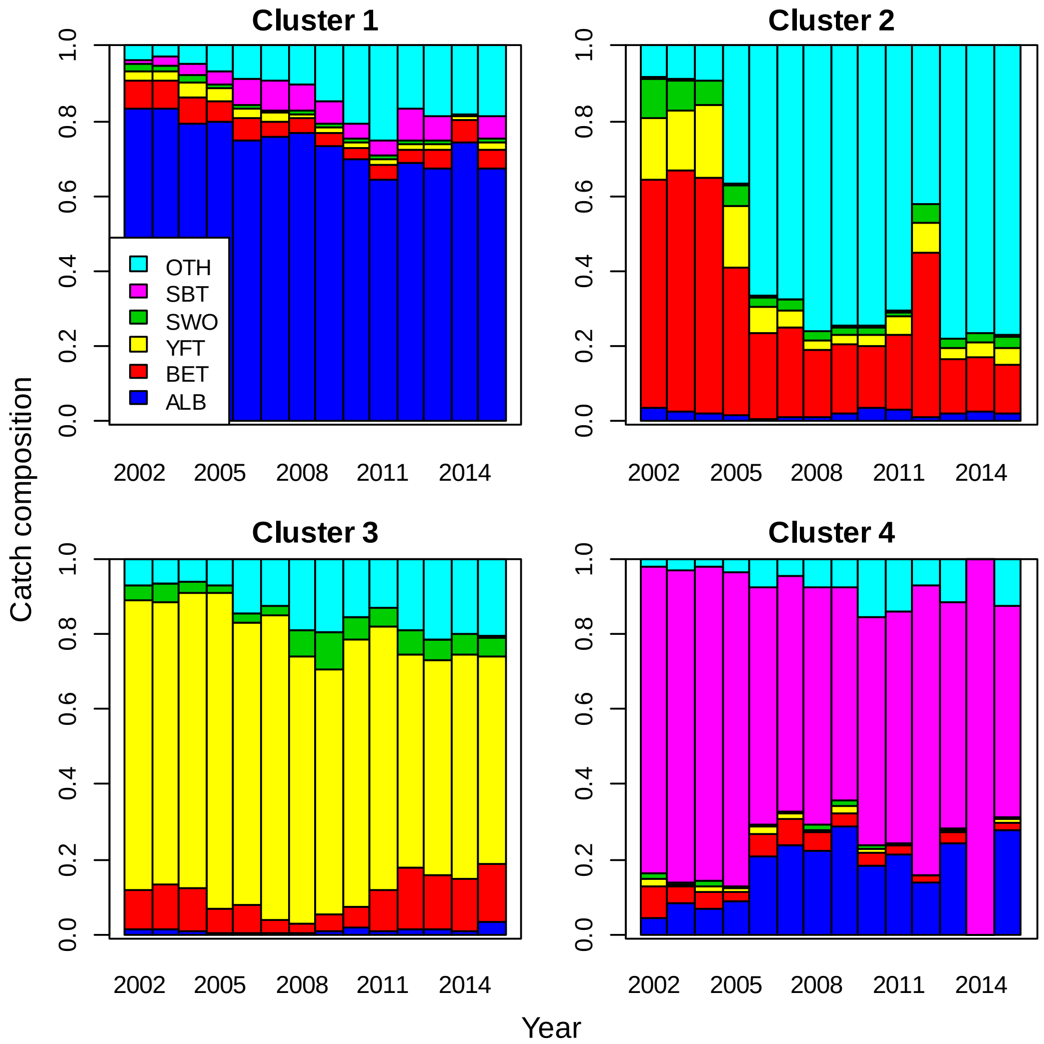

Fig. 4. Annual catch compositions of Taiwanese southern bluefin tuna longline fishery in the Indian Ocean for nine clusters.

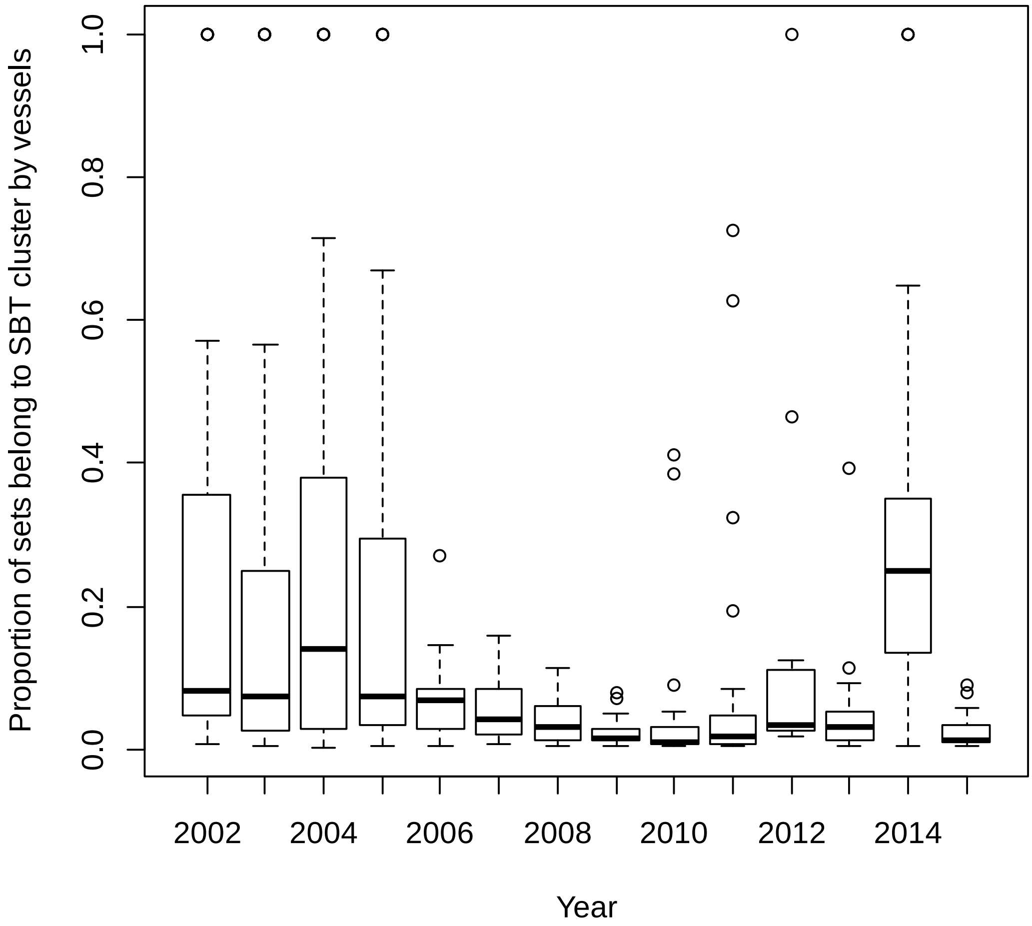

Fig. 5. The box-plot (quartile) of proportion of fishing sets belong to southern bluefin tuna cluster for Taiwanese southern bluefin tuna longline fishery in the Indian Ocean. The data in 2015 are preliminary.

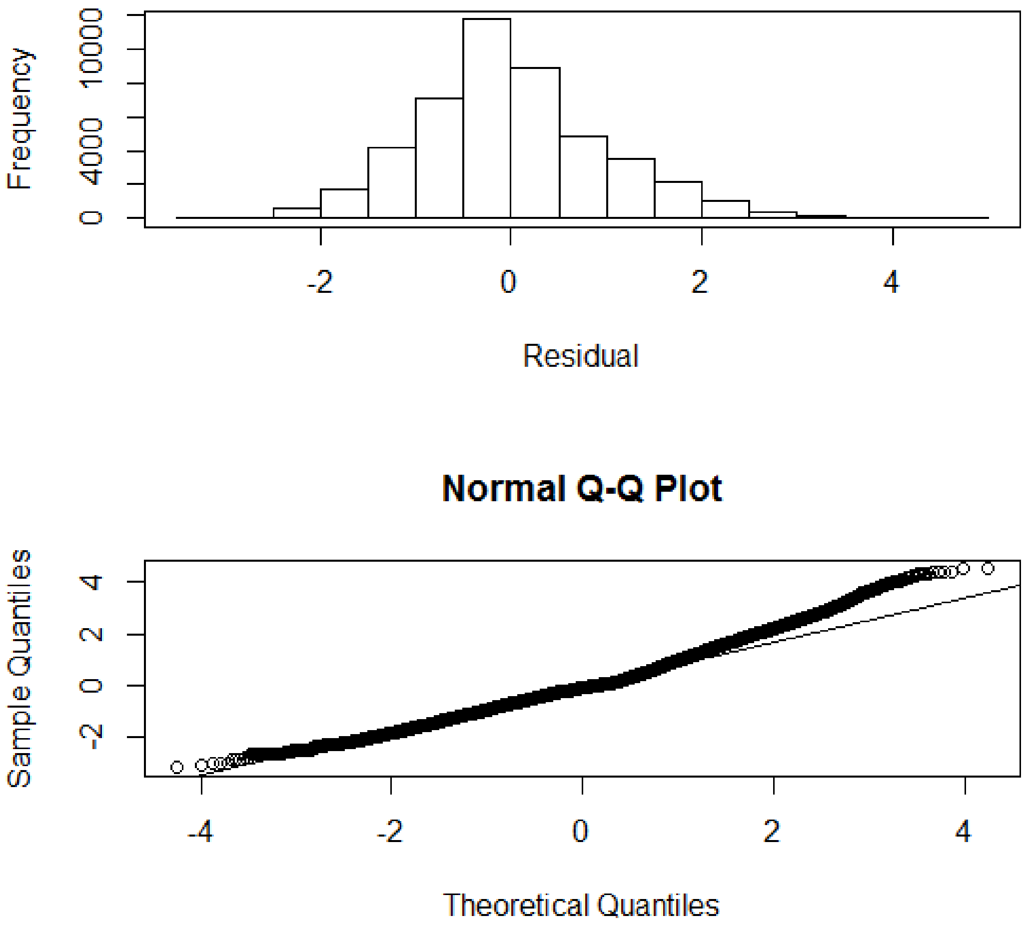

Fig. 6. The frequency distribution and Quantile-Quantile Plot for standardized residuals obtained from area-specific GLM analysis.

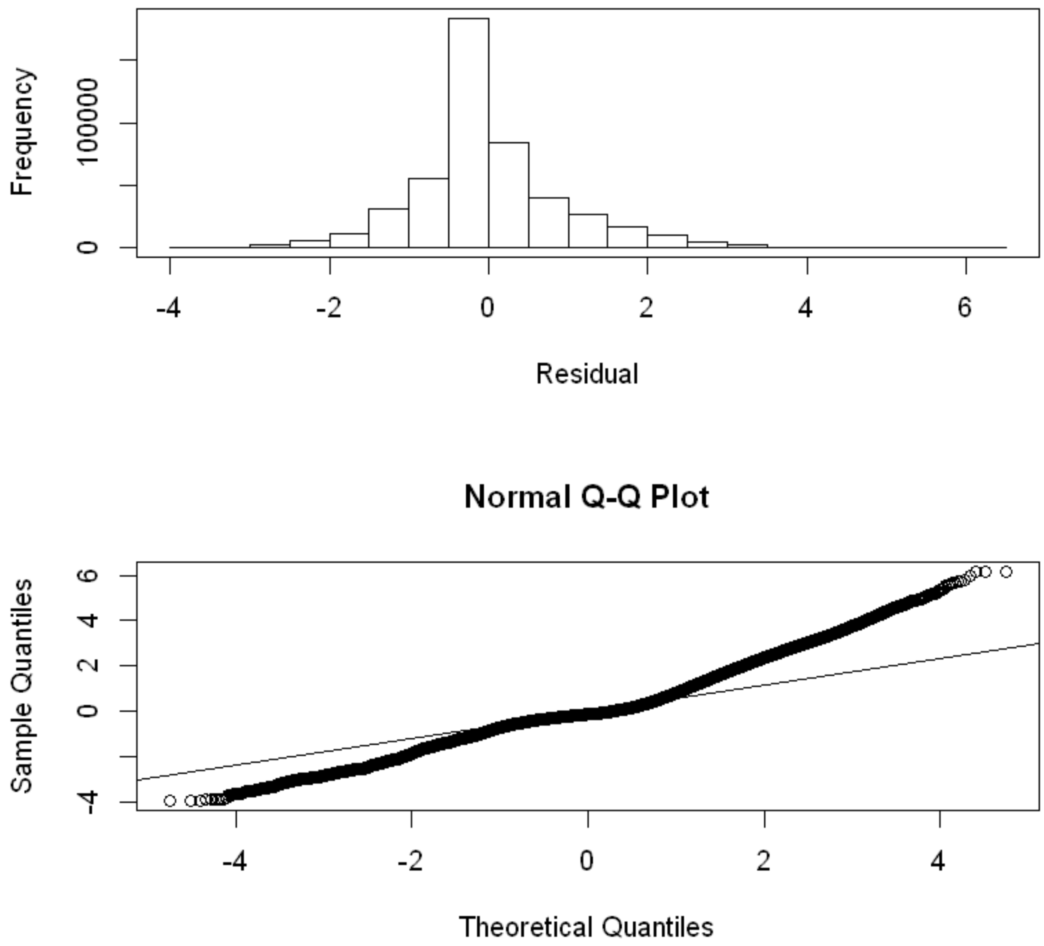

Fig. 7. The frequency distribution and Quantile-Quantile Plot for standardized residuals obtained from age-specific GLM analysis.

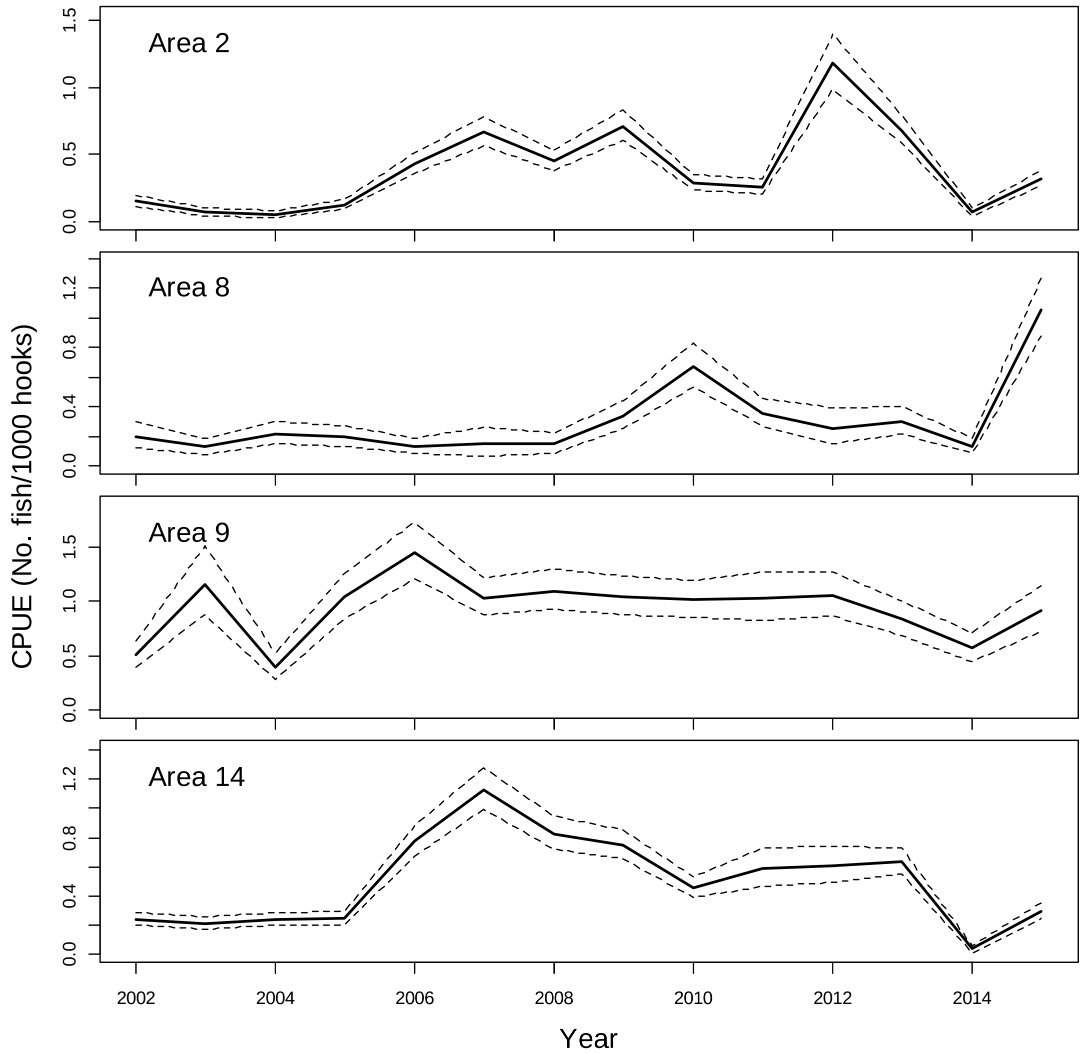

Fig. 8. Area-specific standardized CPUE of southern bluefin tuna caught by Taiwanese longline fishery. Dashed lines represents the 95% confidence intervals.

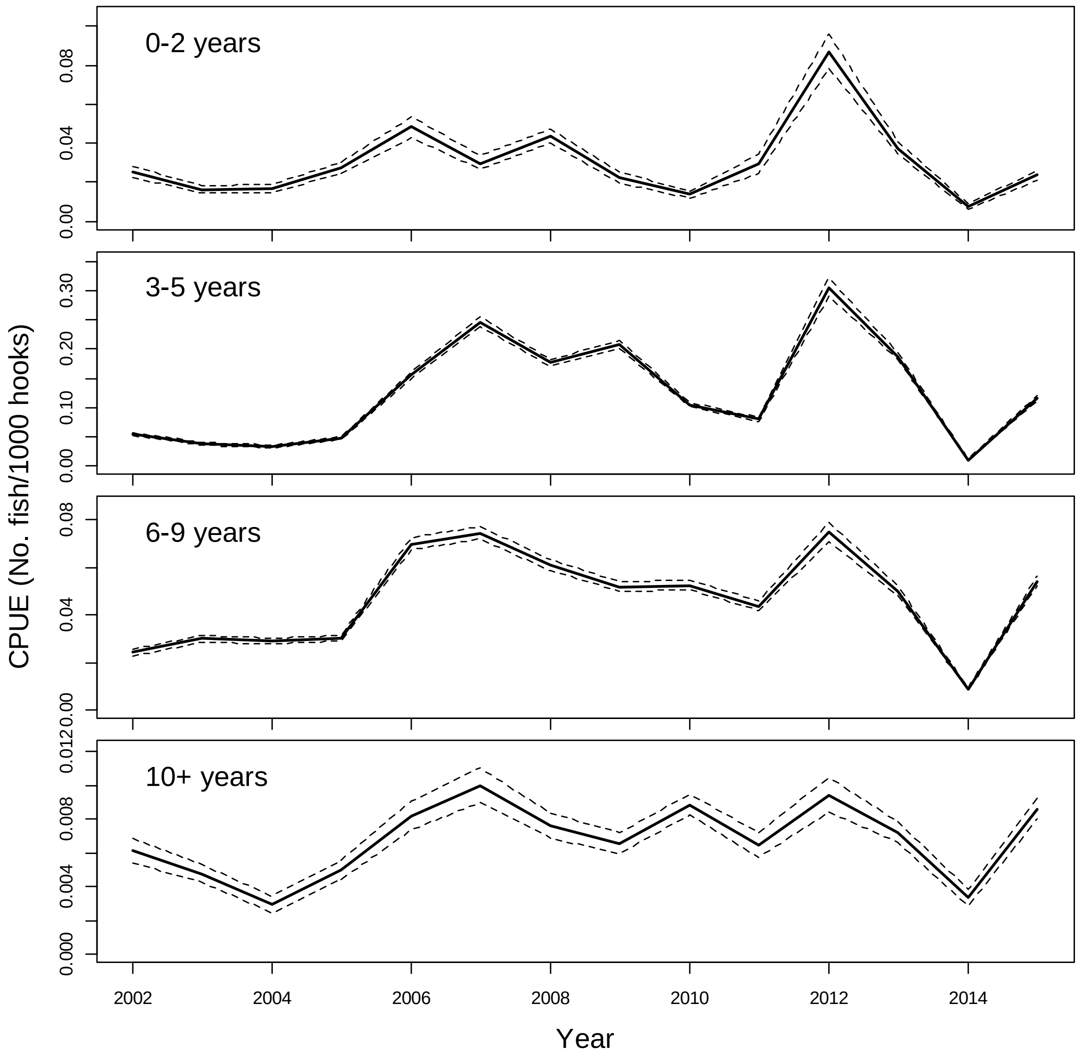

Fig. 9. Age-specific tandardized CPUE of southern bluefin tuna caught by Taiwanese longline fishery. Dashed lines represents the 95% confidence intervals.

Table 1. The average proportions of catches by species for Taiwanese southern bluefin tuna longline fishery in the Indian Ocean based on 10 non-hierarchical (K-means) clusters.

Cluster	ALB	BET	YFT	SBT	SWO
1	0.726	0.061	0.025	0.041	0.014
2	0.009	0.466	0.335	0.000	0.049
3	0.040	0.392	0.083	0.004	0.098
4	0.017	0.173	0.504	0.001	0.070
5	0.036	0.032	0.026	0.008	0.020
6	0.012	0.644	0.107	0.001	0.064
7	0.012	0.193	0.103	0.001	0.459
8	0.046	0.020	0.006	0.769	0.004
9	0.004	0.837	0.038	0.000	0.027
10	0.001	0.021	0.862	0.000	0.021

Table 2. The average proportions of catches by species for Taiwanese southern bluefin tuna longline fishery in the Indian Ocean based on 4 hierarchical clusters.

Cluster	ALB	BET	YFT	SBT	SWO
1	0.726	0.061	0.025	0.041	0.014
2	0.018	0.492	0.111	0.002	0.073
3	0.009	0.102	0.671	0.001	0.047
4	0.046	0.020	0.006	0.769	0.004

Table 3. The number of vessels for Taiwanese southern bluefin tuna longline fishery in the Indian Ocean based on various data selection criteria.

Year	All data	Core25	Core50	Core75
2002	126	39	26	13
2003	117	55	37	19
2004	111	62	42	21
2005	78	36	24	12
2006	57	19	13	8
2007	45	18	12	6
2008	52	21	14	7
2009	71	20	14	7
2010	77	33	23	11
2011	56	17	12	6
2012	16	8	6	3
2013	53	26	18	9
2014	47	33	22	11
2015	58	25	17	9

Table 4. The proportion of southern bluefin tuna catches of selected vessels to all data sets for Taiwanese southern bluefin tuna longline fishery in the Indian Ocean based on various data selection criteria.

Year	Core25	Core50	Core75
2002	80.6	48.5	18.3
2003	76.5	50.1	24.2
2004	70.6	47.0	23.5
2005	78.3	54.4	29.8
2006	74.8	54.7	33.8
2007	88.3	61.3	36.3
2008	70.1	49.4	23.4
2009	69.2	57.5	32.2
2010	56.9	40.0	24.2
2011	54.4	37.5	14.3
2012	68.9	60.8	29.7
2013	84.0	63.6	32.6
2014	95.7	67.9	31.0
2015	66.0	51.1	29.2

Table 5. The proportion of efforts of selected vessels to all data sets for Taiwanese southern bluefin tuna longline fishery in the Indian Ocean based on various data selection criteria.

Year	Core25	Core50	Core75
2002	31.1	16.6	4.3
2003	48.1	26.7	10.7
2004	42.2	21.3	11.3
2005	47.9	25.5	12.4
2006	37.9	24.0	15.2
2007	46.4	29.7	14.6
2008	43.0	29.5	13.7
2009	30.8	19.4	9.9
2010	42.6	29.7	13.7
2011	25.0	15.6	5.2
2012	53.8	42.5	10.0
2013	40.6	29.5	12.7
2014	69.2	44.0	12.9
2015	43.8	28.5	13.7

Table 6. ANOVA table for area-specific model.

Variable	SS	Df	F	Pr(>F)
Y	4733	13	379.817	< 2.2e-16 ***
M	3033	11	287.636	< 2.2e-16 ***
A	181	3	62.807	< 2.2e-16 ***
Lon	2021	18	117.135	< 2.2e-16 ***
C	501	3	174.148	< 2.2e-16 ***
Y*A	2140	39	57.236	< 2.2e-16 ***
M*C	1103	33	34.877	< 2.2e-16 ***
A*C	454	9	52.672	< 2.2e-16 ***
Residuals	44125	46034		

Table 7. ANOVA table for age-specific model.

	SS	Df	F	Pr(>F)
Y	2206	13	191.43	< 2.20E-16 ***
M	7898	11	809.98	< 2.20E-16 ***
A	407	3	153	< 2.20E-16 ***
Lon	4277	17	283.82	< 2.20E-16 ***
C	3238	3	1217.65	< 2.20E-16 ***
AG	5024	3	1888.95	< 2.20E-16 ***
Y:AG	26587	39	769.02	< 2.20E-16 ***
A:C	1595	9	199.96	< 2.20E-16 ***
A:AG	13768	9	1725.65	< 2.20E-16 ***
C:AG	35535	9	4453.95	< 2.20E-16 ***
Residuals	423892	478170		