CPUE standardization for southern bluefin tuna caught by Taiwanese longline fishery

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

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

ABSTRACT

This study attempts to conduct the CPUE standardization for southern bluefin tuna caught by Taiwanese longline fishery using a general linear model. In this study, we attempted to conduct vessel selection based on the cluster analysis and the results indicated that the cluster of fishing operation was the most explanatory main effect. Standardized CPUEs generally reveal quite different trends for different area. Generally, standardized CPUEs substantially decreased for all areas in recent years. CPUEs for fishes with age of 3-5 years were much higher than other age groups. However, CPUEs also substantially decreased in recent years for most age groups, except for age 10+ group.

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 2002-2013.

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 2002-2014 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

In Taiwanese National Report and previous Taiwanese SBT CPUE analyses, the trend of SBT CPUE series were generally calculated based on the data from active longline vessels authorized to seasonally target SBT. In this study, we attempted to conduct vessel selection based on the cluster analysis. 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, latitude, operation cluster, and their interactions. The GLM was conducted as below:

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

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,

Lat	is the latitude 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, while the effect of longitude and latitude were treated as continuous 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:

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

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 7, 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 3, 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 higher before 2005. It should be noted that the data in 2014 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 70-85% and the efforts can reduced to less than 20% when the criterion of the 1st quartile (core25) was used. This indicated that exclusive vessels spent large amount of efforts on not catching SBT. The SBT catches were reduced to 50% but the efforts were not further decreased when the criterion of the 2nd quartile (core50) was used. Only 30% SBT catches were remained when the criterion of the 3rd quartile (core75) was used. 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 latitude effect was excluded from the GLM because it was not statistically significant. In addition, the interaction between month and area effects was excluded because SBT catch did not occur in every months and areas and too many missing values were occurred. Therefore, a final GLM was selected as

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

The ANVOA table for the final GLM is shown in Table 6 and all remaining main effects and interactions were statistically significant. The model can explain 59% of CPUE variance. The explanatory abilities of main effects of fishing operation cluster and longitude and some interactions between these two main effects were much more than other main effects and interactions.

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 latitude effect was excluded from the GLM because it was not statistically significant. The interaction between month and area between month and age effects were excluded because of missing values. Therefore, a final GLM was selected as

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

The ANVOA table for the final GLM is shown in Table 7 and all remaining main effects and interactions were statistically significant. The model can explain 53% of CPUE variance. The fishing operation cluster effect was the most explanatory main effect. The explanatory abilities of interactions between longitude and age effects and between cluster and age effects were also much more than other main effects and interactions.

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. 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. Generally, the standardized CPUEs revealed increasing trends before about 2010 but decreased in recent years for

Areas 2, 8 and 9. The standardized CPUEs in Area 14 reveals increasing trend before 2007 but continuously decreased thereafter. In recent years, however, standardized CPUEs substantially decreased for all areas.

Based on age-specific standardized CPUEs, 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. CPUEs of age 10+ group revealed a relatively stable trend.

Acknowledgments

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., Lin, S.L., 2013. CPUE analysis for southern bluefin tuna caught by Taiwanese longline fleet. CCSBT-ESC/1309/37.

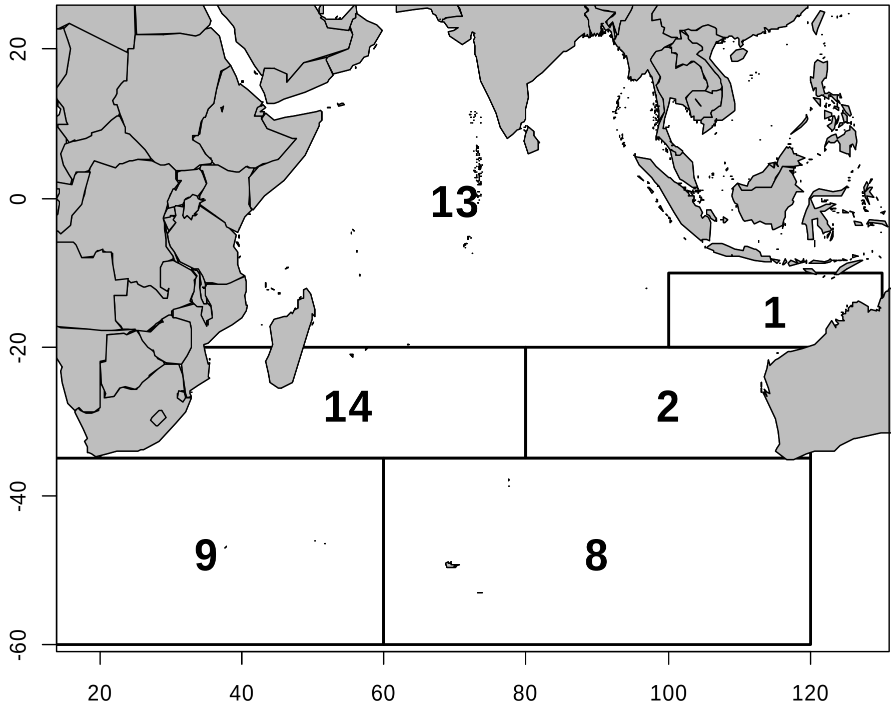


Fig. 1. The definition of CCSBT statistical areas.

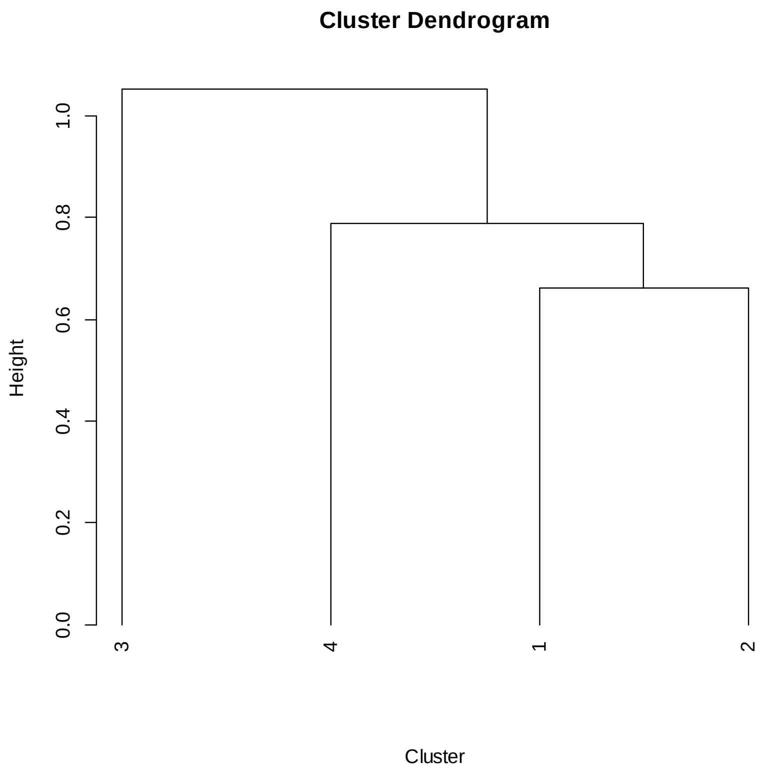


Fig. 2. The dendrogram of hierarchical cluster analysis for classifying the data sets of Taiwanese southern bluefin tuna longline fishery in the Indian Ocean.

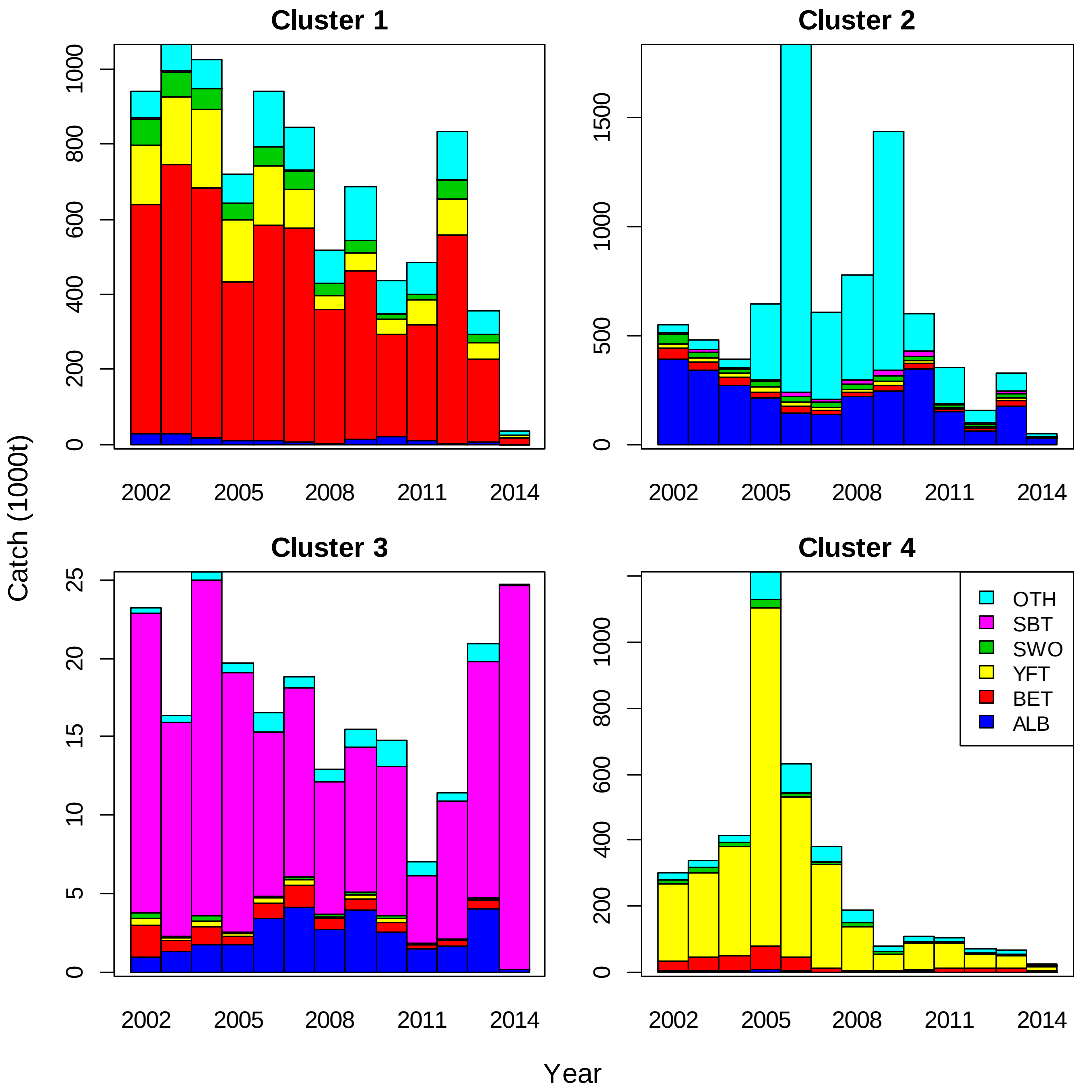


Fig. 3. Annual catches by species of Taiwanese southern bluefin tuna longline fishery in the Indian Ocean for nine clusters.

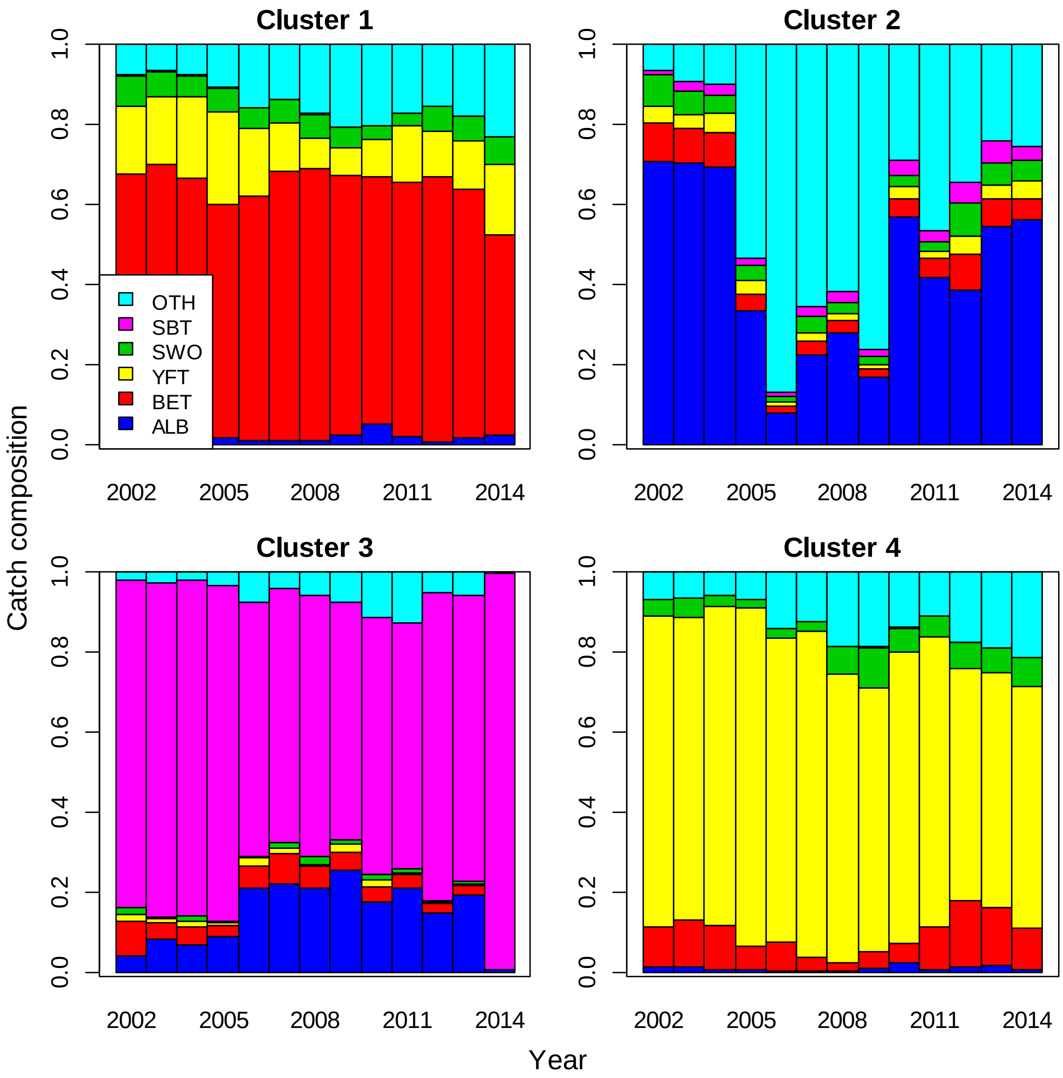


Fig. 4. Annual catch compositions of Taiwanese southern bluefin tuna longline fishery in the Indian Ocean for nine clusters.

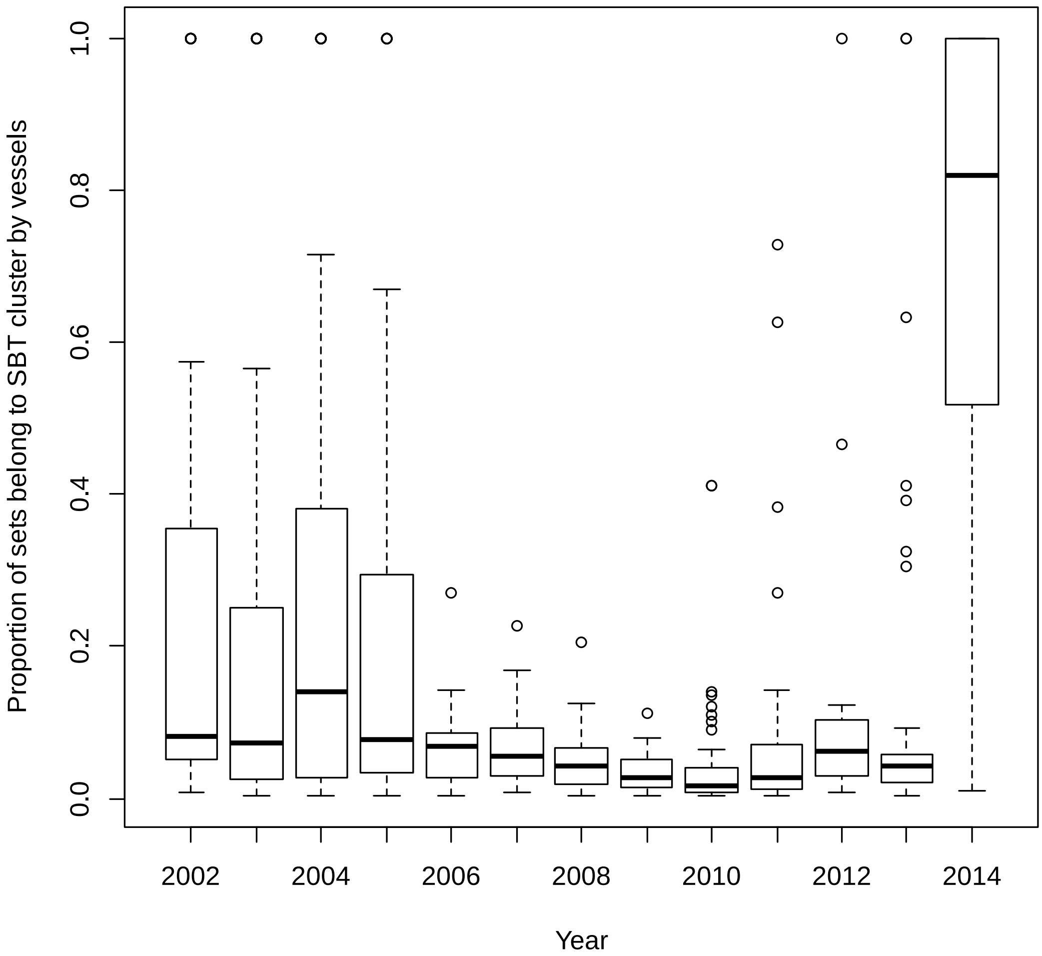
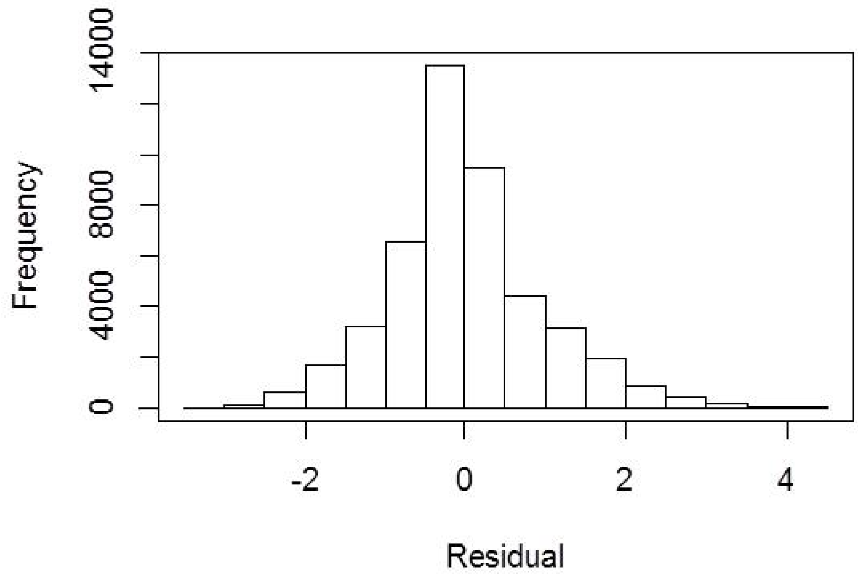


Fig. 5. The box-plot (quartile) of proportion of fishing sets belong to southern bluefin tuna cluster for Taiwanese southern bluefin tuna longline fishery in the Indian Ocean. The data in 2014 are preliminary.



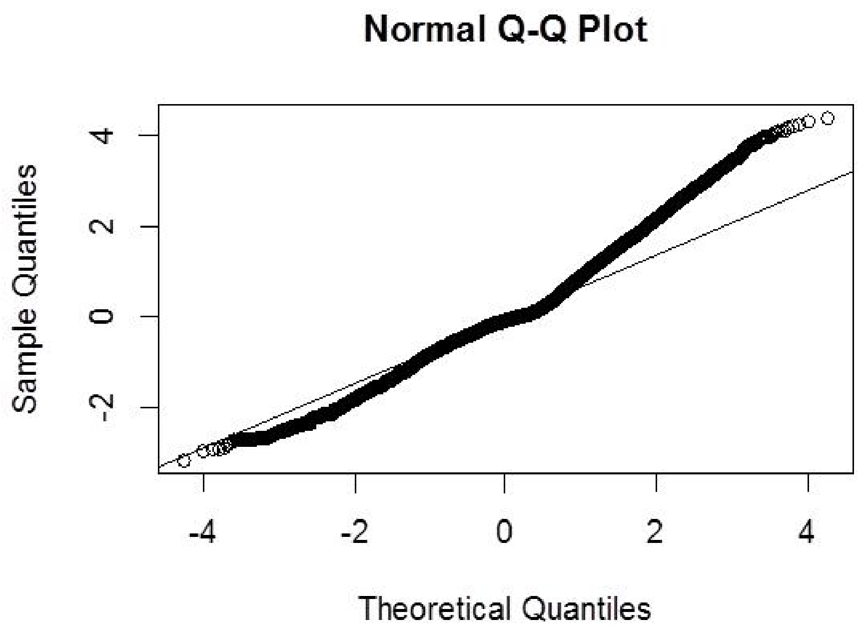


Fig. 6. The frequency distribution and Quantile-Quantile Plot for standardized residuals obtained from area-specific GLM analysis.

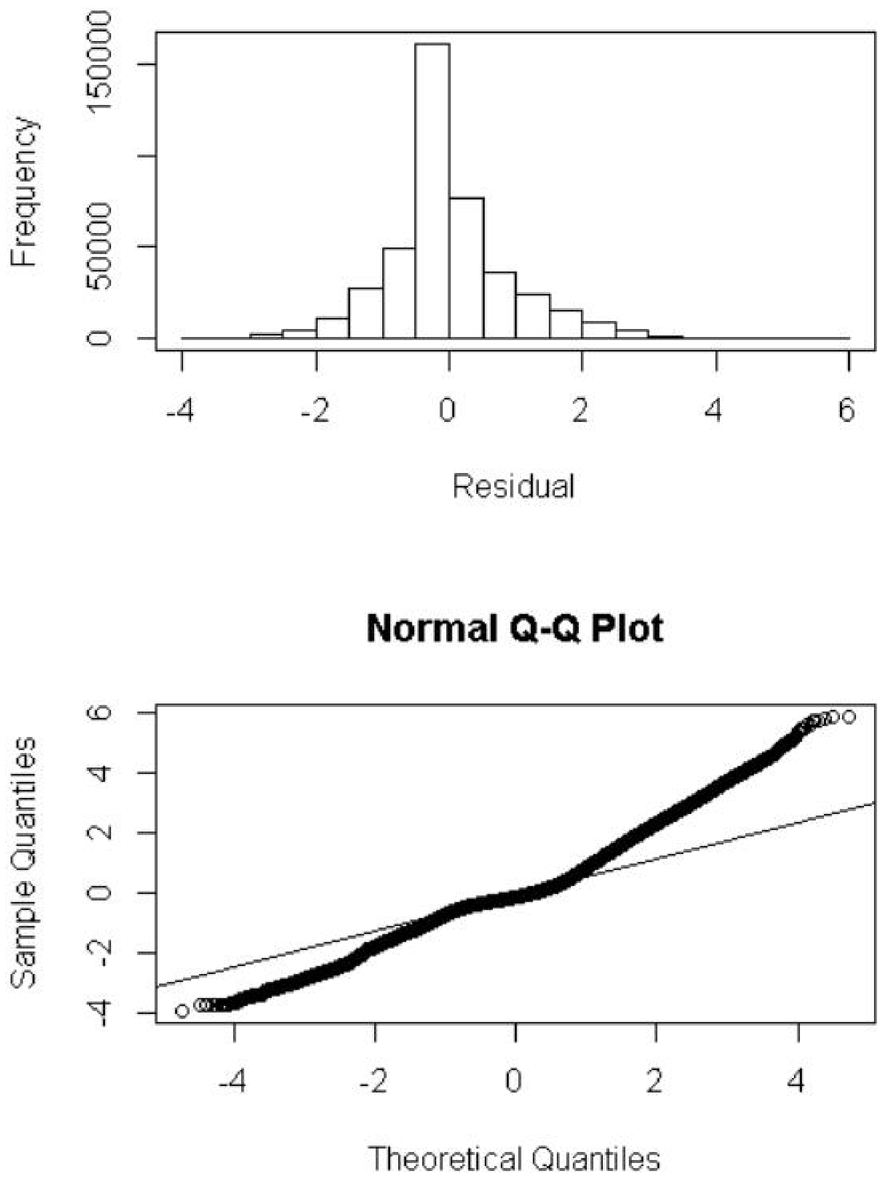


Fig. 7. The frequency distribution and Quantile-Quantile Plot for standardized residuals obtained from age-specific GLM analysis.

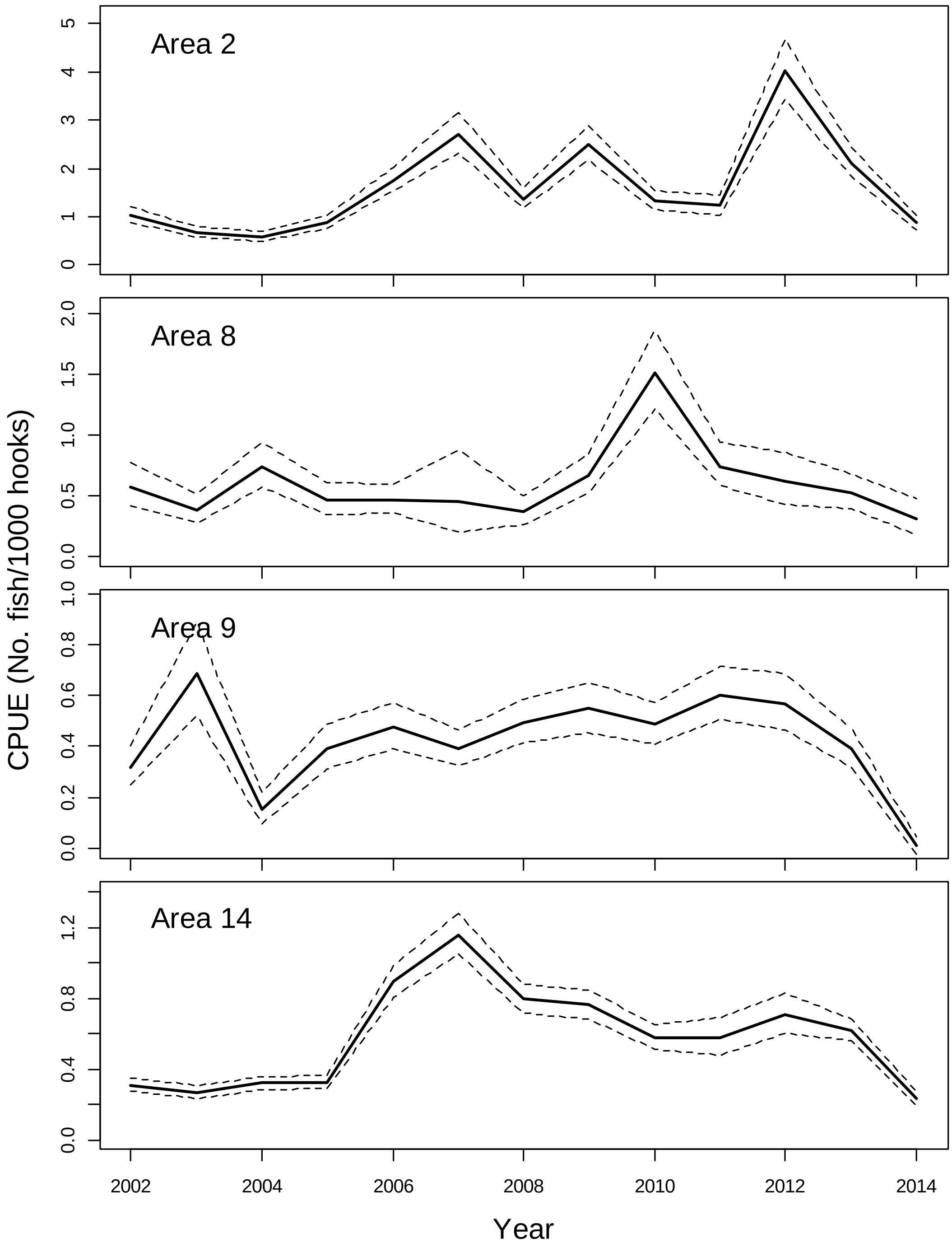


Fig. 8. Area-specific standardized CPUE of southern bluefin tuna caught by Taiwanese longline fishery. Dashed lines represents the 95% confidence intervals.

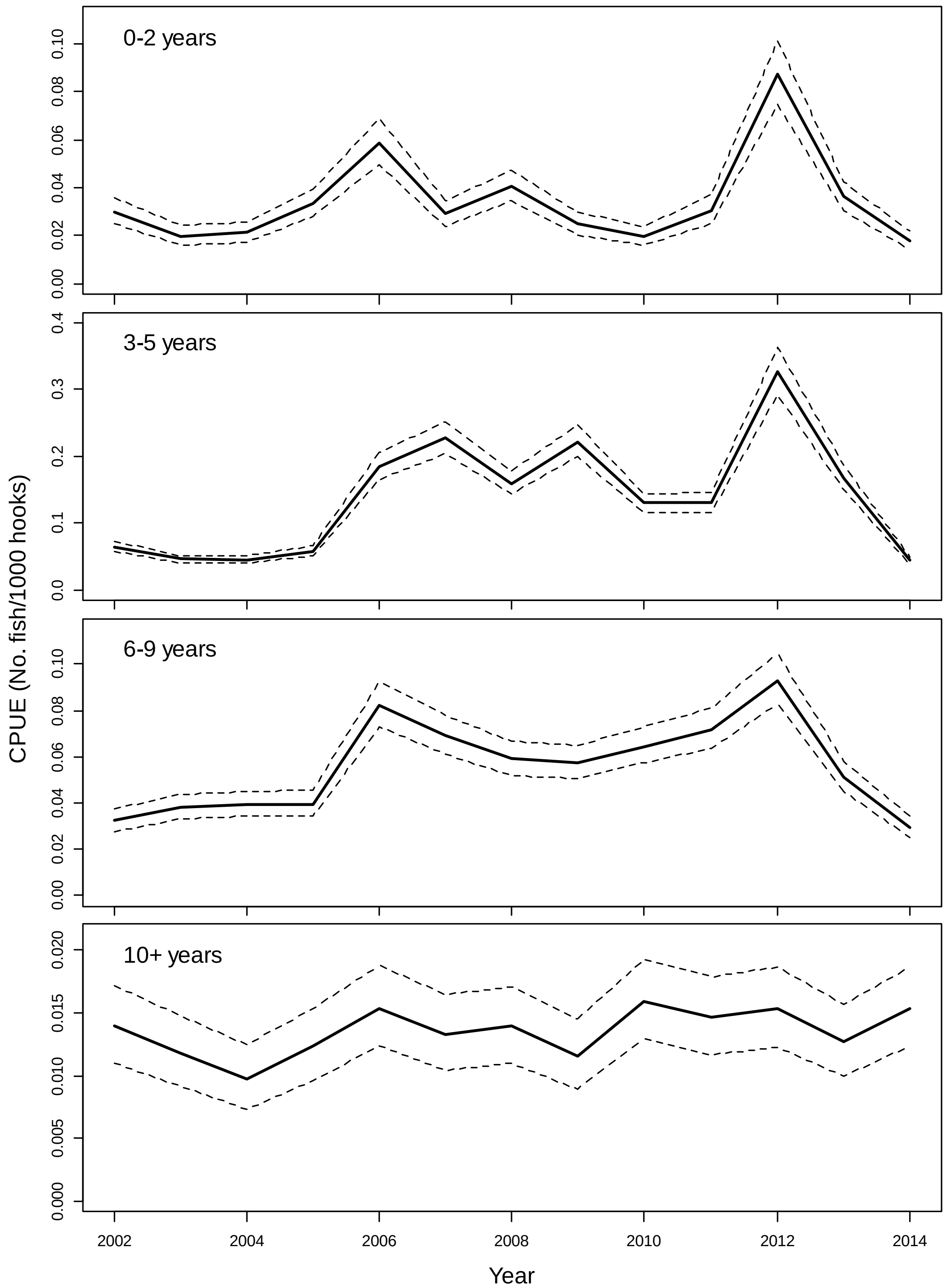


Fig. 9. Age-specific tandardized CPUE of southern bluefin tuna caught by Taiwanese longline fishery. Dashed lines represents the 95% confidence intervals.

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	0.009	0.461	0.339	0.000	0.050
2	0.023	0.171	0.095	0.001	0.478
3	0.731	0.063	0.027	0.039	0.019
4	0.044	0.416	0.080	0.004	0.112
5	0.004	0.841	0.036	0.000	0.027
6	0.058	0.059	0.046	0.011	0.037
7	0.047	0.022	0.006	0.766	0.006
8	0.001	0.020	0.868	0.000	0.019
9	0.012	0.653	0.111	0.001	0.060
10	0.017	0.168	0.518	0.001	0.072

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	0.016	0.616	0.130	0.001	0.060
2	0.389	0.078	0.044	0.024	0.094
3	0.047	0.022	0.006	0.766	0.006
4	0.009	0.096	0.687	0.001	0.046

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	248	39	26	13
2003	261	55	37	19
2004	262	62	42	21
2005	278	36	24	12
2006	269	19	13	7
2007	194	19	13	7
2008	180	25	17	9
2009	163	24	16	8
2010	158	39	26	13
2011	127	21	14	7
2012	128	10	7	4
2013	133	31	21	11
2014	72	32	22	20

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	All data	Core25	Core50	Core75
2002	100	80.3	48.7	20.7
2003	100	76.4	49.4	23.9
2004	100	70.5	47.0	23.5
2005	100	78.3	56.1	29.8
2006	100	72.9	54.7	29.1
2007	100	84.3	60.4	34.5
2008	100	74.3	50.4	26.0
2009	100	77.3	49.1	25.9
2010	100	67.8	45.7	22.7
2011	100	63.9	43.0	17.7
2012	100	79.3	64.6	33.5
2013	100	86.4	61.8	35.3
2014	100	76.1	49.1	44.0

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	All data	Core25	Core50	Core75
2002	100	14.2	6.9	1.6
2003	100	19.3	10.0	3.7
2004	100	18.8	9.2	3.5
2005	100	13.3	7.8	3.3
2006	100	9.0	5.9	3.0
2007	100	11.3	7.9	4.3
2008	100	16.5	11.2	5.8
2009	100	16.5	11.0	5.5
2010	100	31.8	20.8	9.7
2011	100	17.9	10.6	3.4
2012	100	8.9	6.1	2.6
2013	100	29.4	19.3	7.4
2014	100	42.5	21.5	18.8

Table 6. ANOVA table for area-specific model.

	SS	Df	F	Pr(>F)
Y	2581	12	243.353	< 2.2e-16 ***
M	338	11	34.803	< 2.2e-16 ***
A	528	3	199.126	< 2.2e-16 ***
Lon	181	1	204.61	< 2.2e-16 ***
C	677	3	255.386	< 2.2e-16 ***
Y*A	1481	36	46.529	< 2.2e-16 ***
M*Lon	1581	11	162.616	< 2.2e-16 ***
M*C	802	33	27.486	< 2.2e-16 ***
A*Lon	711	3	267.963	< 2.2e-16 ***
A*C	97	9	12.167	< 2.2e-16 ***
Lon*C	291	3	109.776	< 2.2e-16 ***
Residuals	40516	45839		

Table 7. ANOVA table for age-specific model.

	SS	Df	F	Pr(>F)
Y	1291	12	123.0586	< 2.2e-16 ***
M	234	11	24.371	< 2.2e-16 ***
A	373	3	142.4078	< 2.2e-16 ***
Lon	6	1	7.0183	0.008068 **
C	2753	3	1049.6512	< 2.2e-16 ***
AG	927	3	353.5273	< 2.2e-16 ***
Y:AG	13410	36	426.1503	< 2.2e-16 ***
M:Lon	867	11	90.1908	< 2.2e-16 ***
M:C	2799	33	97.034	< 2.2e-16 ***
A:Lon	613	3	233.9491	< 2.2e-16 ***
A:C	450	9	57.2453	< 2.2e-16 ***
A:AG	7111	9	903.9531	< 2.2e-16 ***
Lon:C	1764	3	672.519	< 2.2e-16 ***
Lon:AG	3700	3	1411.0862	< 2.2e-16 ***
C:AG	22319	9	2837.1027	< 2.2e-16 ***
Residuals	371120	424569		