INTER-AMERICAN TROPICAL TUNA COMMISSION

SCIENTIFIC ADVISORY COMMITTEE

NINTH MEETING

La Jolla, California (USA)

14-18 May 2018

DOCUMENT SAC-09 INF F

CPUE STANDARDIZATION FOR BIGEYE TUNA AND YELLOWFIN TUNA CAUGHT BY TAIWANESE LONGLINE IN THE EASTERN PACIFIC OCEAN

Sheng-Ping Wang[1], Mark N. Maunder[2], Cleridy E. Lennert-Cody[2], Alexandre Aires-da-Silva[2]

1. Department of Environmental Biology and Fisheries Science, National Taiwan Ocean University, Taiwan
2. Inter-American Tropical Tuna Commission

SUMMARY

In this study, cluster analysis methods were applied to identify the fishing operations defining targeting for each set. The cluster identifier and the levels of number of hooks between floats were incorporated as predictors into the delta-lognormal models for conducting the CPUE standardizations for bigeye tuna and yellowfin tuna. The results indicated that cluster identifier had the most explanatory effect for both of species. In addition, the models that included cluster identifier as a proxy for a targeting effect provided better performance than the model that included NHBF as targeting effect. The standardized CPUE series from 2000 to 2016 were also provided in this study.

INTRODUCTION

Based on the discussions at the Eighth Meeting of the Scientific Advisory Committee of the Inter-American Tropical Tuna Commission, this study analyzed the operational catch and effort data for Taiwanese longline fishery in the eastern Pacific Ocean using cluster analysis methods and also conducted the CPUE standardizations for bigeye tuna and yellowfin tuna.

MATERIALS AND METHODS

Catch and Effort data

The analyses are conducted based on the daily operational-level catch and effort data (logbook) with 5x5 degree longitude and latitude resolution for the Taiwanese longline fishery operating in the eastern Pacific Ocean during 1980-2016, which were provided by Oversea Fisheries Development Council of Taiwan (OFDC). It should be noted that the data for 2016 are preliminary.

Data on number of hooks between floats (NHBF) were available since 1994 and the collection of NHBF data were more complete since 1995. Therefore, the data of NHBF may not be applicable to conduct the long-term CPUE standardization for fishes caught by Taiwanese longline fishery.

Categorization of fishing targeting using cluster analysis

Cluster analysis methods can group sets into categories by automatically identifying sets with similar species-composition. This approach has been applied to tuna and billfish species using data from vessels in multiple fleets (e.g. Bigelow and Hoyle, 2012; He et al., 1997; Wang and Nishida, 2014; Wang, 2015; Wang et al., 2017).

Cluster analyses was conducted based on species composition of the catches. Six main species groups were used in this study, including albacore (ALB), bigeye tuna (BET), yellowfin tuna (YFT), swordfish (SWO), sharks (SKX) and others (OTH, mainly skipjack) (Fig. 1). He et al. (1997) suggested a cluster analysis with two steps to classify the data because the large number of sets precluded the use of direct hierarchical cluster analysis methods. First, a non-hierarchical cluster analysis (K-means method) was used to group all sets into 15 distinct clusters for taking the mixture of fishing targeting into account (i.e. 15 ways of 2 species can be chosen from a group of 6 species, C_2^6). Secondarily, a hierarchical cluster analysis with Ward minimum variance method was applied to the squared Euclidean distances calculated from the 15 clusters. Non-hierarchical and hierarchical cluster analyses were conducted using R functions kmeans and hclust (The R Foundation for Statistical Computing Platform, 2015). He et al. (1997) indicated that the choice for the number of clusters to produce was largely subjective. At least two clusters were expected and more than two clusters were produced to allow other possible categories to emerge.

CPUE Standardization

The fitted models were simply conducted with the main effects of year, quarter, longitude, latitude and the operations related to the fishing targeting (clusters or NHBF). However, more complex models (e.g., with interaction terms) could be considered in the future. The data on NHBF were treated as a categorical variable with three levels (regular: <=9 hooks; deep: 10-14 hooks; ultra deep: >=15 hooks) (Wang and Nishida, 2011).

The delta-lognormal general linear models (Pennington, 1983; Lo et. al., 1992; Pennington, 1996) is applied to conduct the CPUE standardizations.

Lognormal model for CPUE of positive catch:

$$\log(CPUE) = \mu + YQ + Lon + Lat + T$$

Delta(logistic regression) model for presence and absence of catch:

$$logit(p) = \mu + YQ + Lon + Lat + T$$

where *CPUE* is the mean positive catch of the species of interest (catch in number/1,000 hooks),
p is the probability of catch,
μ is the intercept,
YQ is the effect of year-quarter,
Lon is the effect of longitude,
Lat is the effect of latitude,
T is the effect of fishing targeting (clusters or NHBF),

The final models were selected using a stepwise search ("both" direction, i.e. "backward" and "forward") and the Akaike information criterion (AIC).

The standardized relative abundance index was calculated by the product of the standardized CPUE of positive catches and the standardized probability of positive catches:

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

where $\widetilde{CPUE}$ is the adjusted mean (from the fitted model above) of the year-quarter effect of the lognormal model,

$\tilde{P}$ is the adjusted mean (on the scale of the link function, from the fitted model above) of the year-quarter effect of the delta model.

RESULTS

Cluster analysis

He et al. (1997) suggested considering the number of clusters until the smallest cluster contained less than 10% of the total number of observations. In case of this study, selecting 4 hierarchical clusters derived from 15 k-means groups achieved the criterion of He et al. (1997) (Fig. 2).

The results of the cluster analysis indicated that Cluster 1 represented the fishing operations targeting BET, the operations of Cluster 2 targeted ALB, Cluster 3 represented the operations with multiple species, and operations of Cluster 4 mainly consisted of BET and YFT (Fig. 3). Cluster 2 concentrated in shallow sets with wide range of NHBF, while NHBF were quite similar for Cluster 1, 3 and 4, but the catch compositions were obviously different for these clusters (Figs. 3 and 4). Fig. 5 shows the trends of effort and BET and YFT catches by clusters. Most effort was deployed for ALB (Cluster 2) before 2000 and changed to target on BET and other species (Clusters 1 and 3) thereafter. BET catches were mainly made by the operations of Cluster 1, but YFT catches can be from the operations of various clusters, except for Cluster 2.

The spatial distributions of the effort (hooks) and the catches of BET ad YFT are shown in Figs 6-8. Most effort distributed in the waters of the south of 10°S before 2000 (Cluster 2), shifted to the waters between of 10°S and 10°N during 2000-2010 (Clusters 1, 3 and 4), and moved northward and southward to the temperate waters thereafter (Cluster 2). Consequentially, large amounts of BET and YFT catches made by Clusters 1, 3 and 4 were distributed in the waters between of 10°S and 10°N during 2000-2010.

CPUE standardization

Because very few BET and YFT catches were made by Cluster 2 and before 2000 for all clusters, the data of Cluster 2 and data before 2000 in Clusters 1,3 and 4 were excluded for the CPUE standardization.

For models with cluster and NHBF as targeting effects, the lowest value of AIC were obtained based on the models with all effects included for BET and YFT. The

ANVOA tables for the lognormal (positive catch) and delta (probability of positive catch) models are shown in Tables 1-4. For lognormal models, Cluster had the highest explanatory power for both of BET and YFT, while explanatory power of NHBF was much lower than for other variables. For delta models, Cluster was still the most important factor for BET, but its explanatory power was obviously less than the effects of year-quarter and latitude for YFT, while NHBF was the least factor for both of BET and YFT. The distributions of standardized residuals and the Quantile-Quantile Plots indicated that the distributions of residuals approximately followed the assumption of the normal distribution for lognormal models (Figs. 9 and 10).

The statistics of R^2, AIC and BIC and listed in Table 5. For both lognormal and delta models, the values of R^2 obtained from the models with Cluster as targeting effect were higher than those from the models with NHBF as the targeting effect. The values of AIC and BIC from the models with Cluster as targeting effect were also slightly lower than those from the models with NHBF as targeting effect. The results indicated that the models with Cluster as the targeting effect are more appropriate for both BET and YFT.

Figs 11 and 12 show the standardized CPUE series for BET and YFT. The standardized CPUE series obtained from models with Cluster and NHBF as targeting effects revealed a somewhat different pattern for BET, while similar trends were observed from the standardized CPUE for YFT series obtained from different models, except for some values before 2005.

REFERENCE

Bigelow, K.A., Hoyle, S.D., 2012. Standardized CPUE for South Pacific albacore. WCPFC-SC8-2012/SA-IP-14.

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.

Lo, N.C.H., Jacobson, L.D., Squire, J.L., 1992. Indices of relative abundance from fish spotter data based on delta-lognormal models. Can. J. Fish. Aquat. Sci., 49: 2515-2526.

Pennington, M., 1983. Efficient estimation of abundance, for fish and plankton surveys. Biometrics, 39: 281-286.

Pennington, M., 1996. Estimating the mean and variance from highly skewed marine data. Can. J. Fish. Aquat. Sci., 94: 498-505.

Wang, S. P., Maunder, M.N., Aires-da-Silva, A., 2009a. Implications of model and data assumptions: An illustration including data for the Taiwanese longline fishery into the eastern Pacific Ocean bigeye tuna (*Thunnus obesus*) stock assessment. Fisheries Research, 97: 118-126.

Wang, S. P., Maunder, M.N., Aires-da-Silva, A., Bayliff, W. H., 2009b. Evaluating fishery impacts: Application to bigeye tuna (*Thunnus obesus*) in the eastern Pacific Ocean. Fisheries Research, 99: 106-111.

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.

Wang, S.P., 2015. CPUE standardization of striped marlin (*Kajikia audax*) caught by Taiwanese longline fishery in the Indian Ocean using targeting effect derived from cluster and principle component analyses. IOTC–2015–WPB13–31.

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., Nishida, T., 2014. CPUE standardization with targeting analysis for swordfish (*Xiphias gladius*) caught by Taiwanese longline fishery in the Indian Ocean. OTC–2014–WPB12–22 rev1.

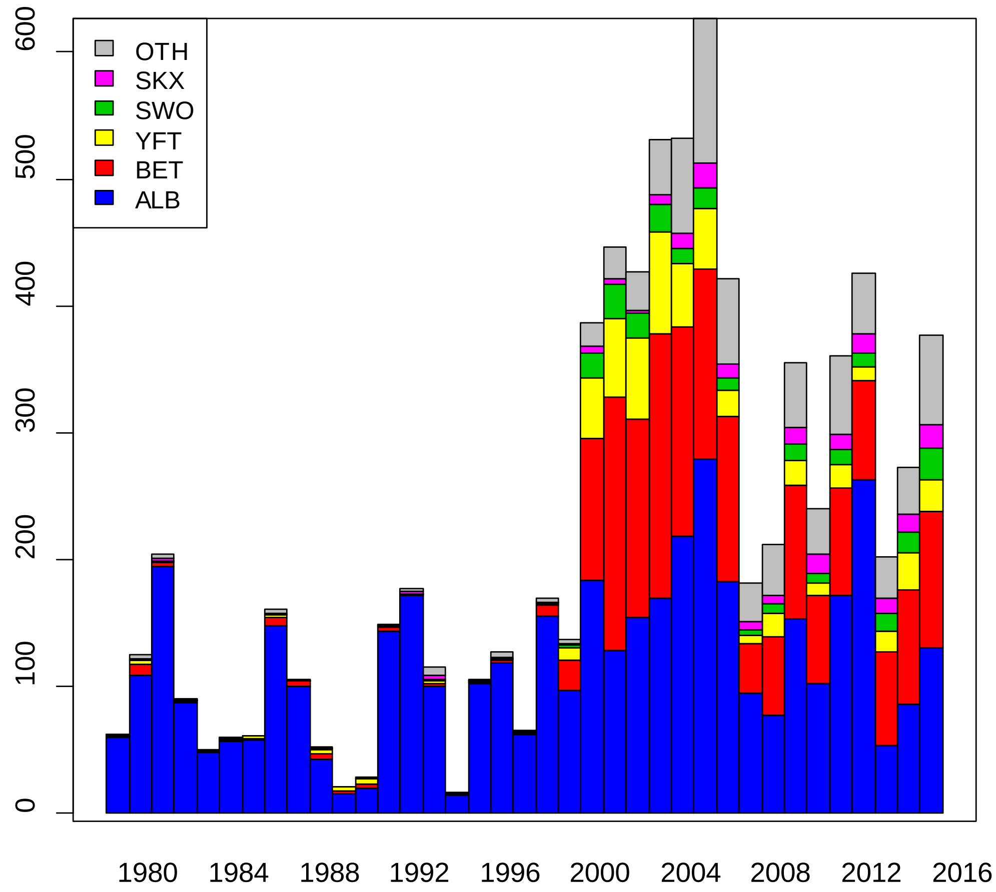


Fig. 1. Annual catch composition of Taiwanese longline vessels operated in the eastern Pacific Ocean.

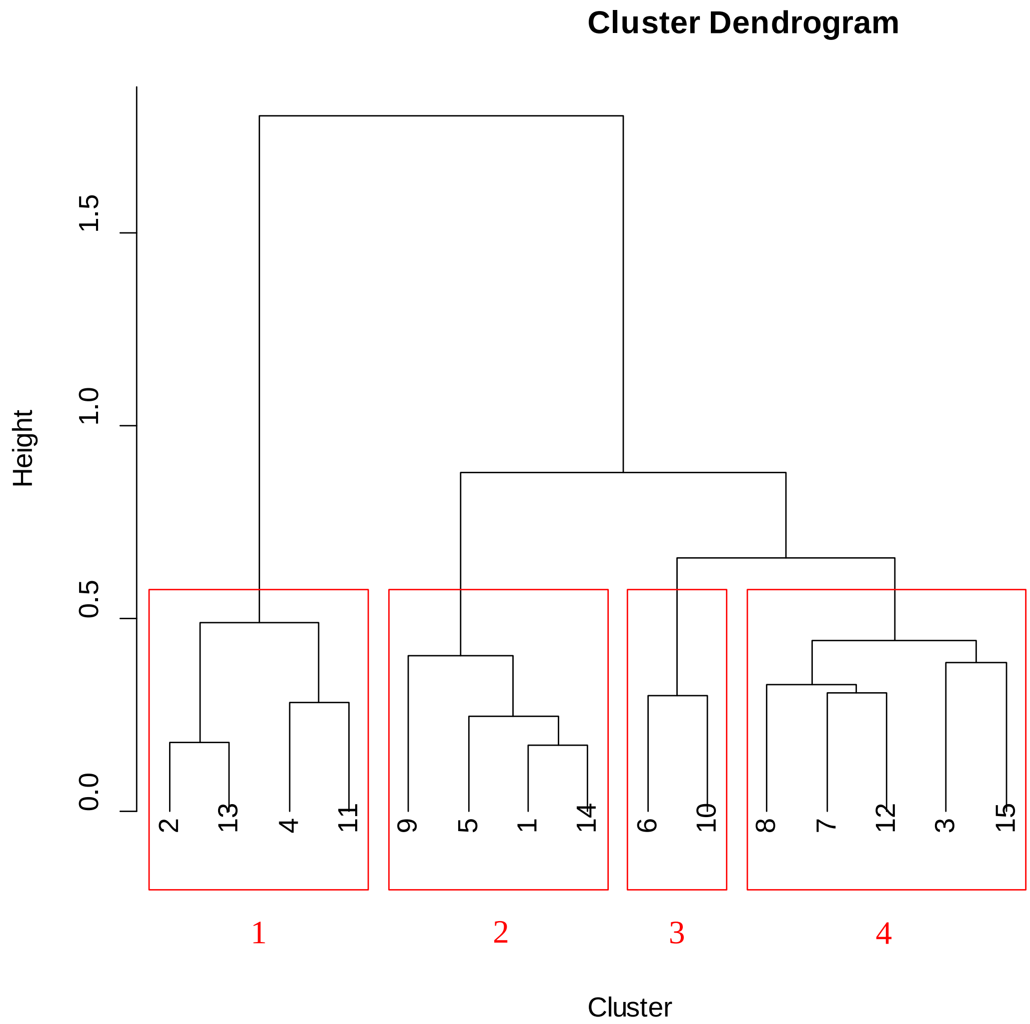


Fig. 2. Hierarchical clusters (labels in red) derived from k-means groups (labels in black) based on the daily operational catch composition data of Taiwanese longline vessels operated in the eastern Pacific Ocean.

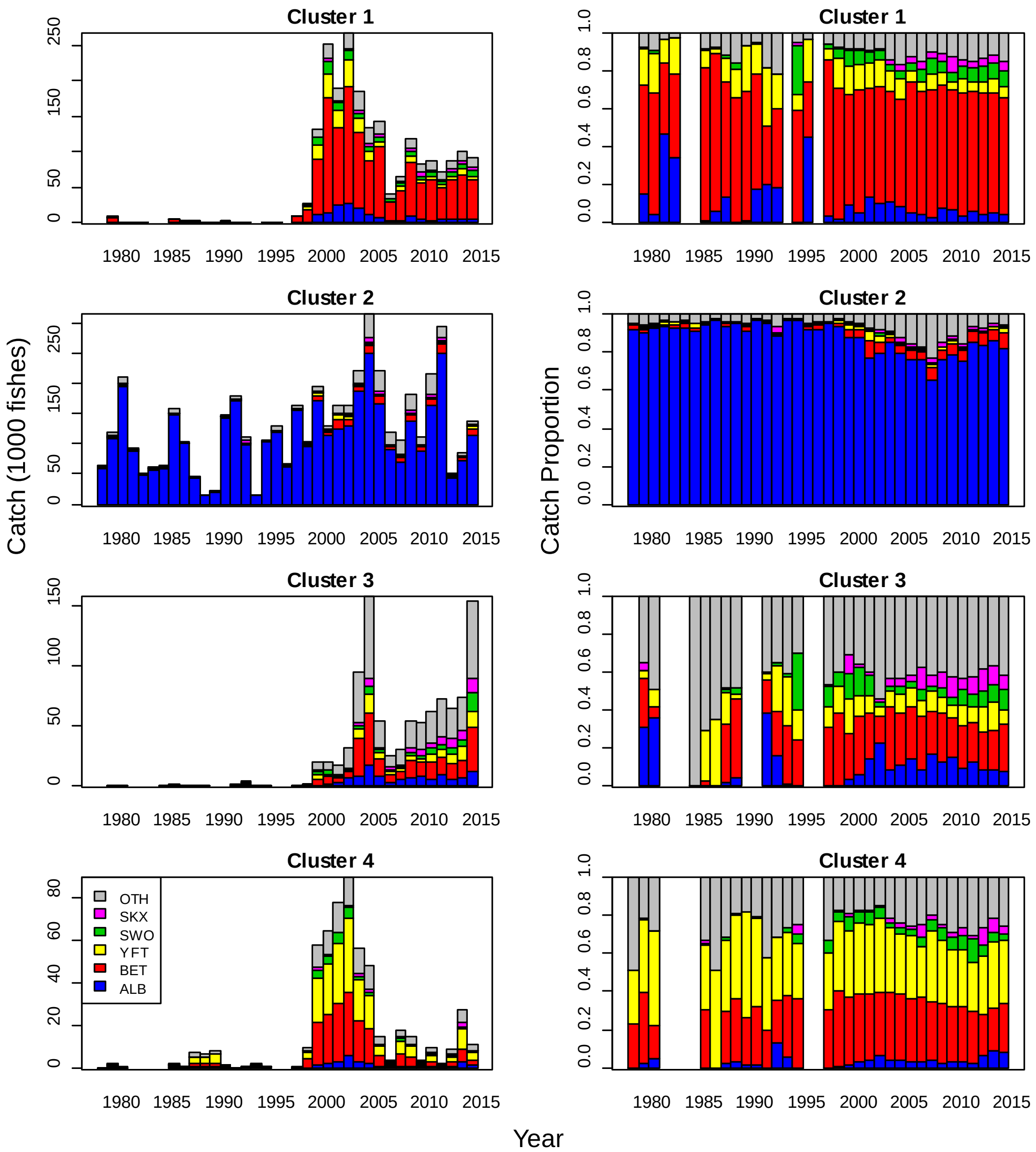


Fig. 3. Annual catch compositions (left panel) and proportions (right panel) by clusters based on the daily operational data of Taiwanese longline vessels operated in the eastern Pacific Ocean.

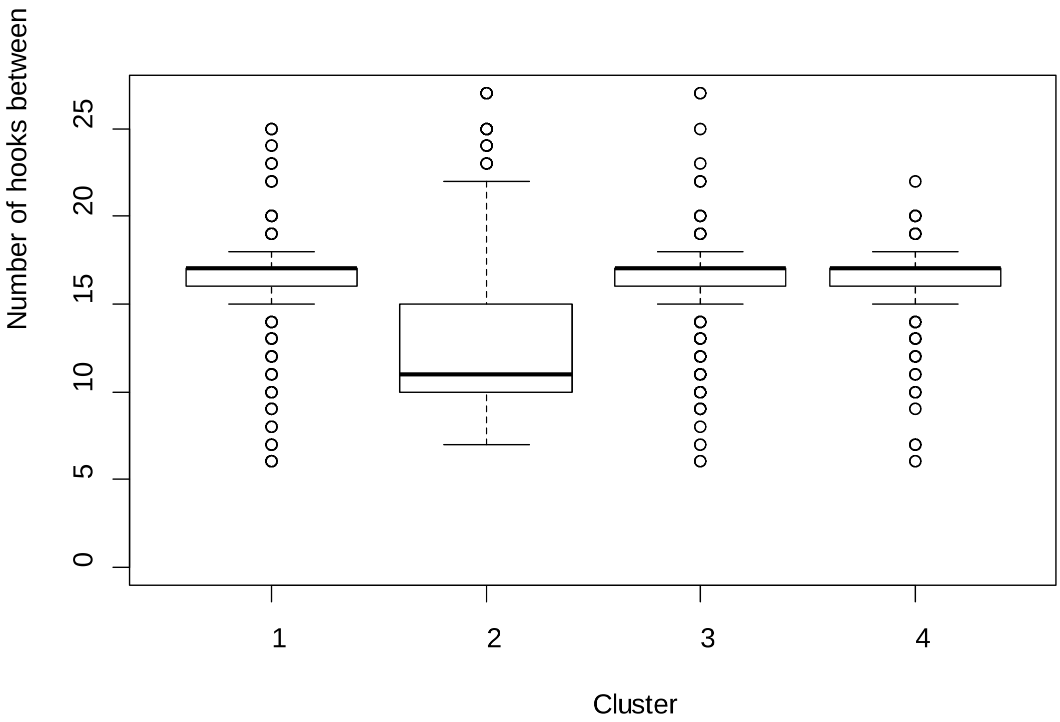


Fig. 4. The number of hooks between float grouped by four clusters based on the data of Taiwanese longline vessels operated in the eastern Pacific Ocean.

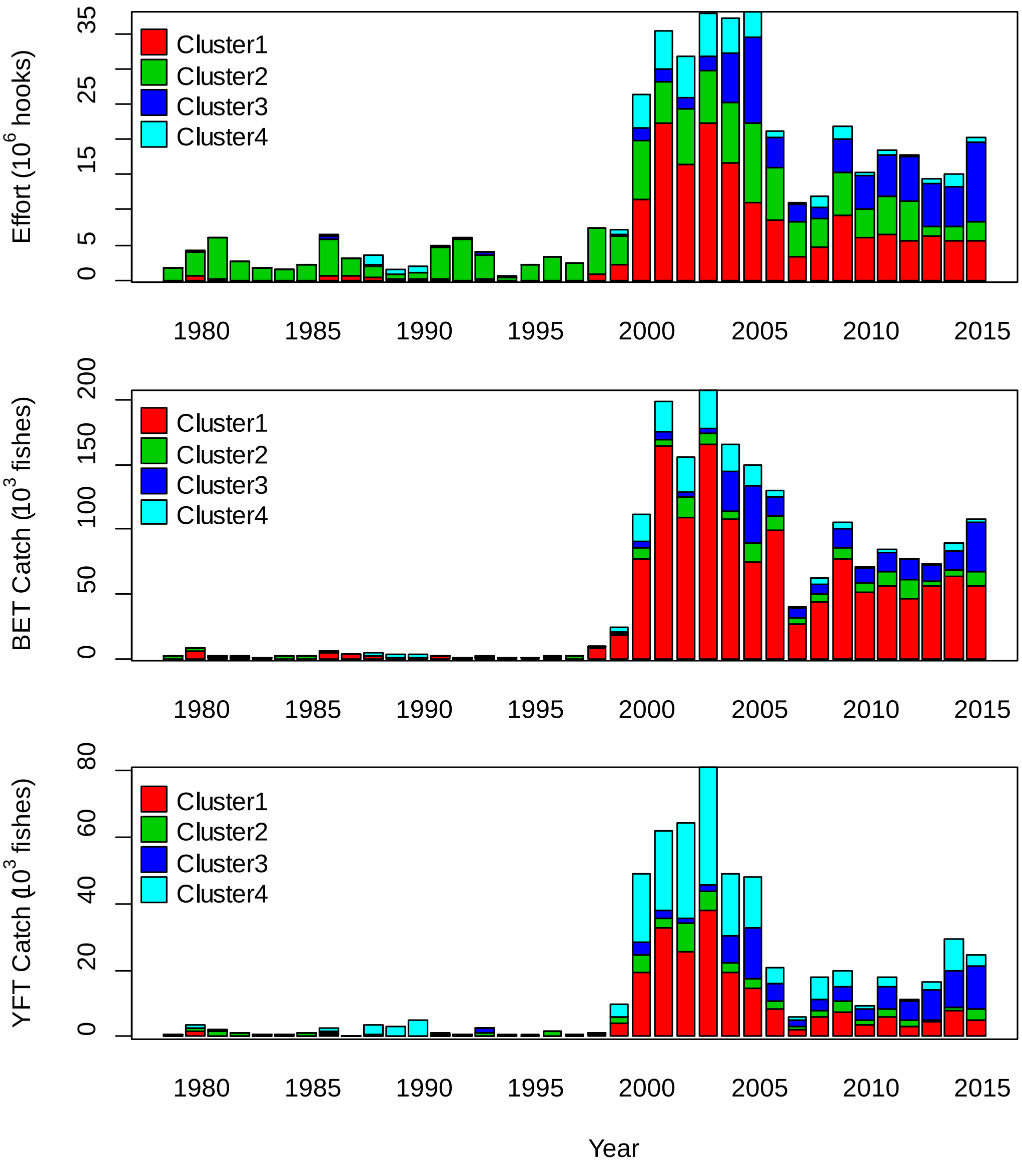


Fig. 5. Annual effort and catches of BET and YFT, by cluster, based on the daily operational data of Taiwanese longline vessels operated in the eastern Pacific Ocean.

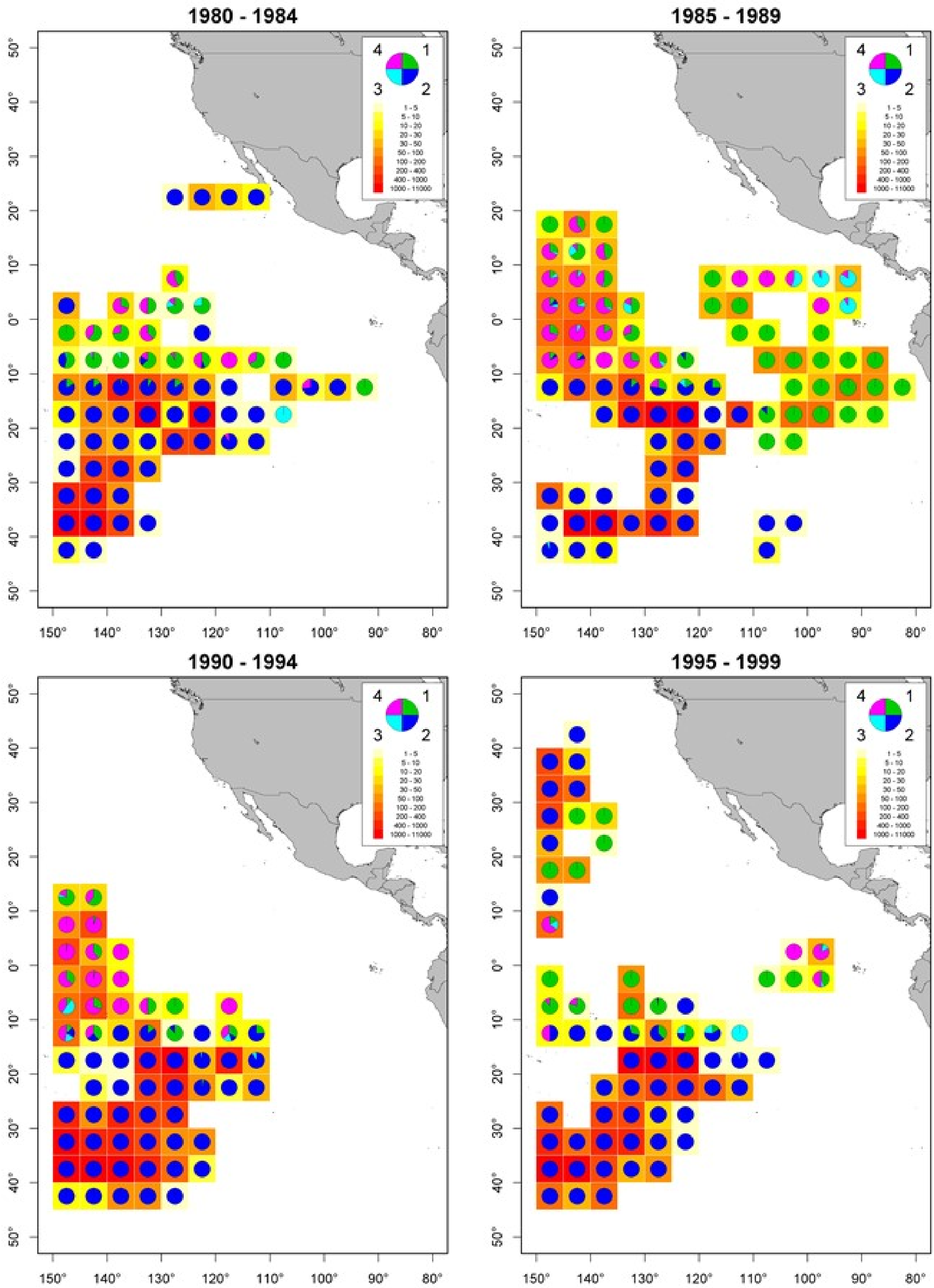


Fig. 6. The distribution of effort (hooks) of Taiwanese longline vessels operated in the eastern Pacific Ocean. Pie plots show the proportions of effort contributed by clusters and squares with heat colors show the levels of effort.

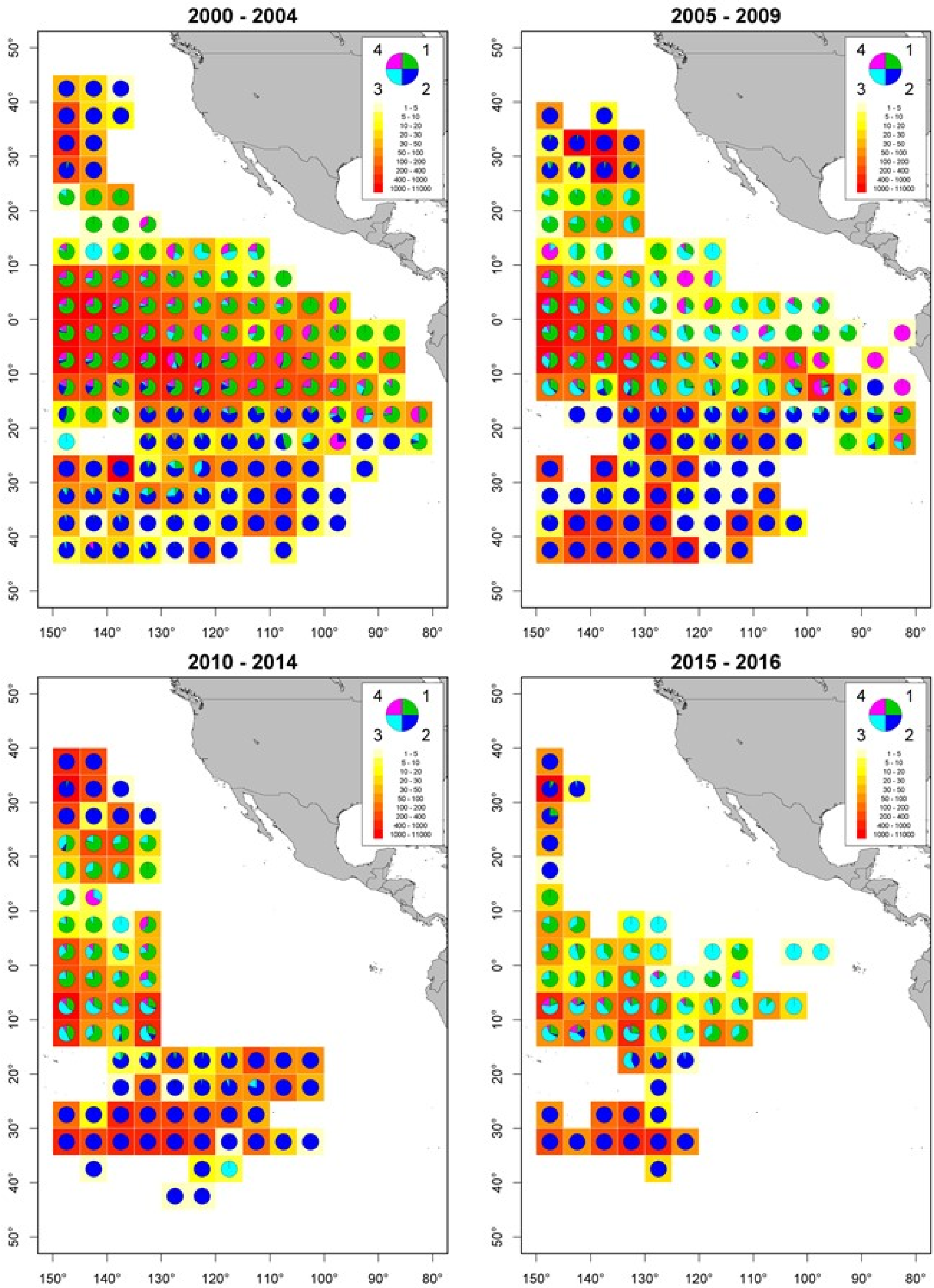


Fig. 6. (Continued).

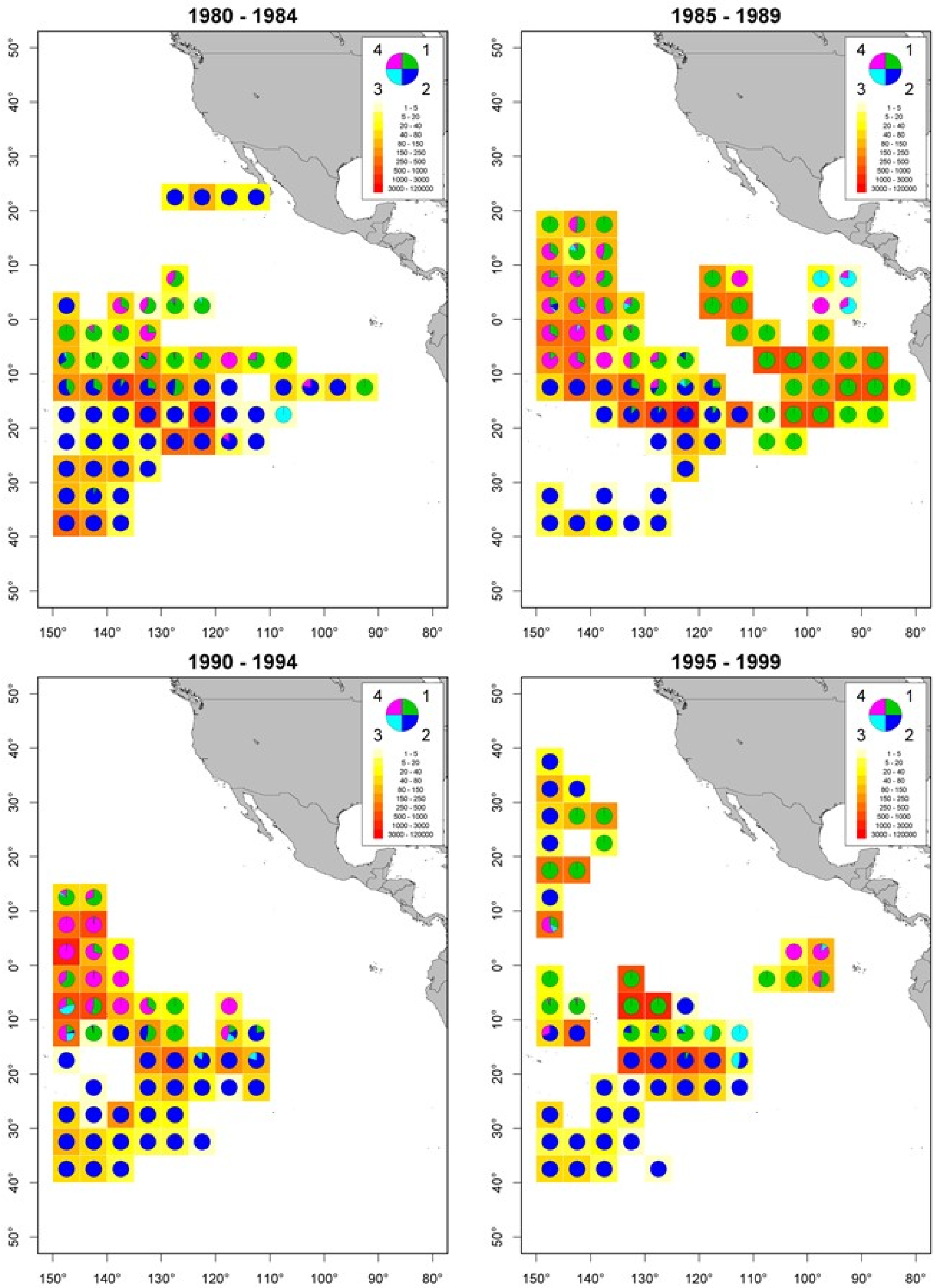


Fig. 7. The distribution of BET catches of Taiwanese longline vessels operated in the eastern Pacific Ocean. Pie plots show the proportions of BET catches contributed by clusters and squares with heat colors show the levels of BET catches.

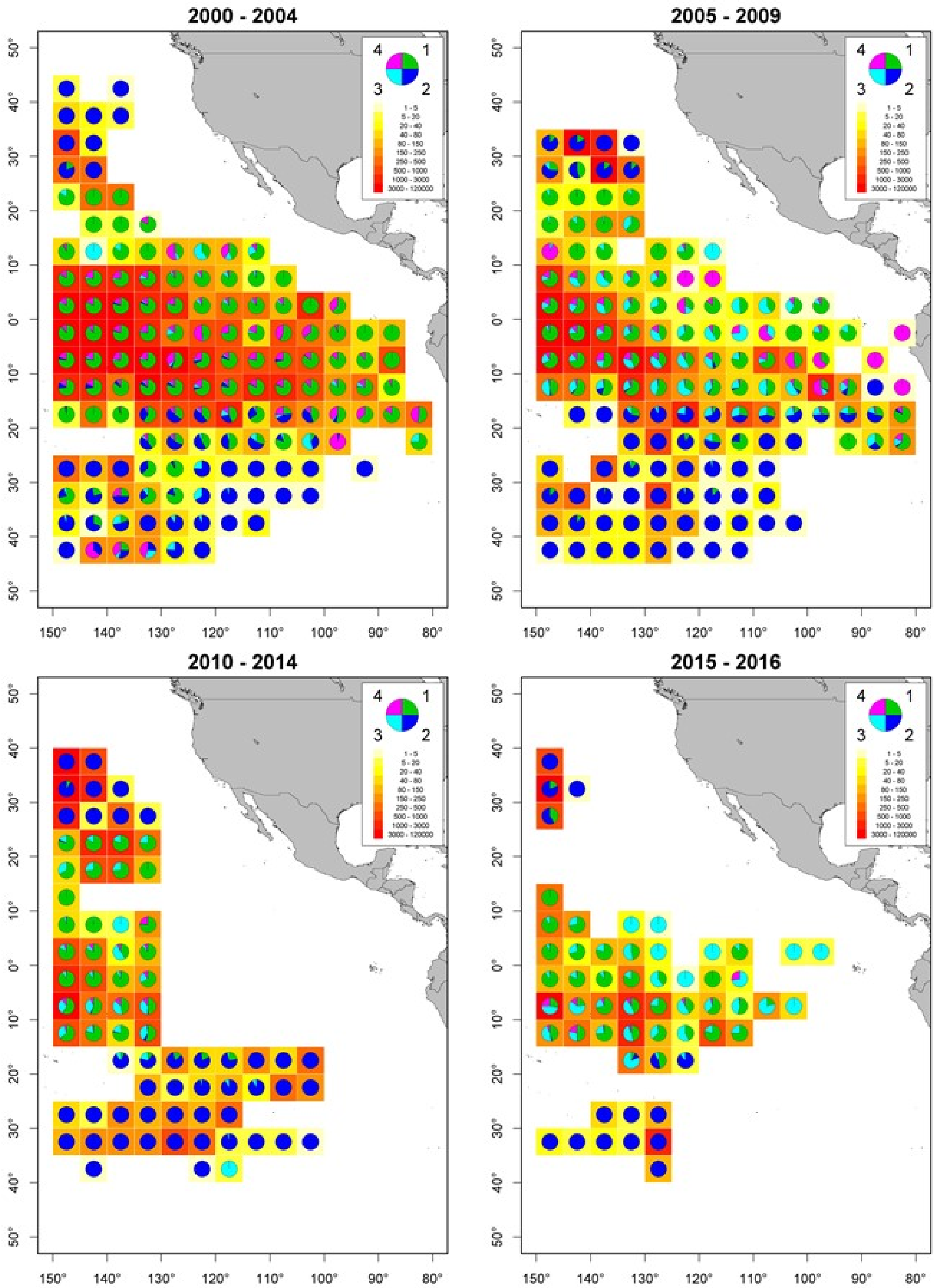


Fig. 7. (Continued).

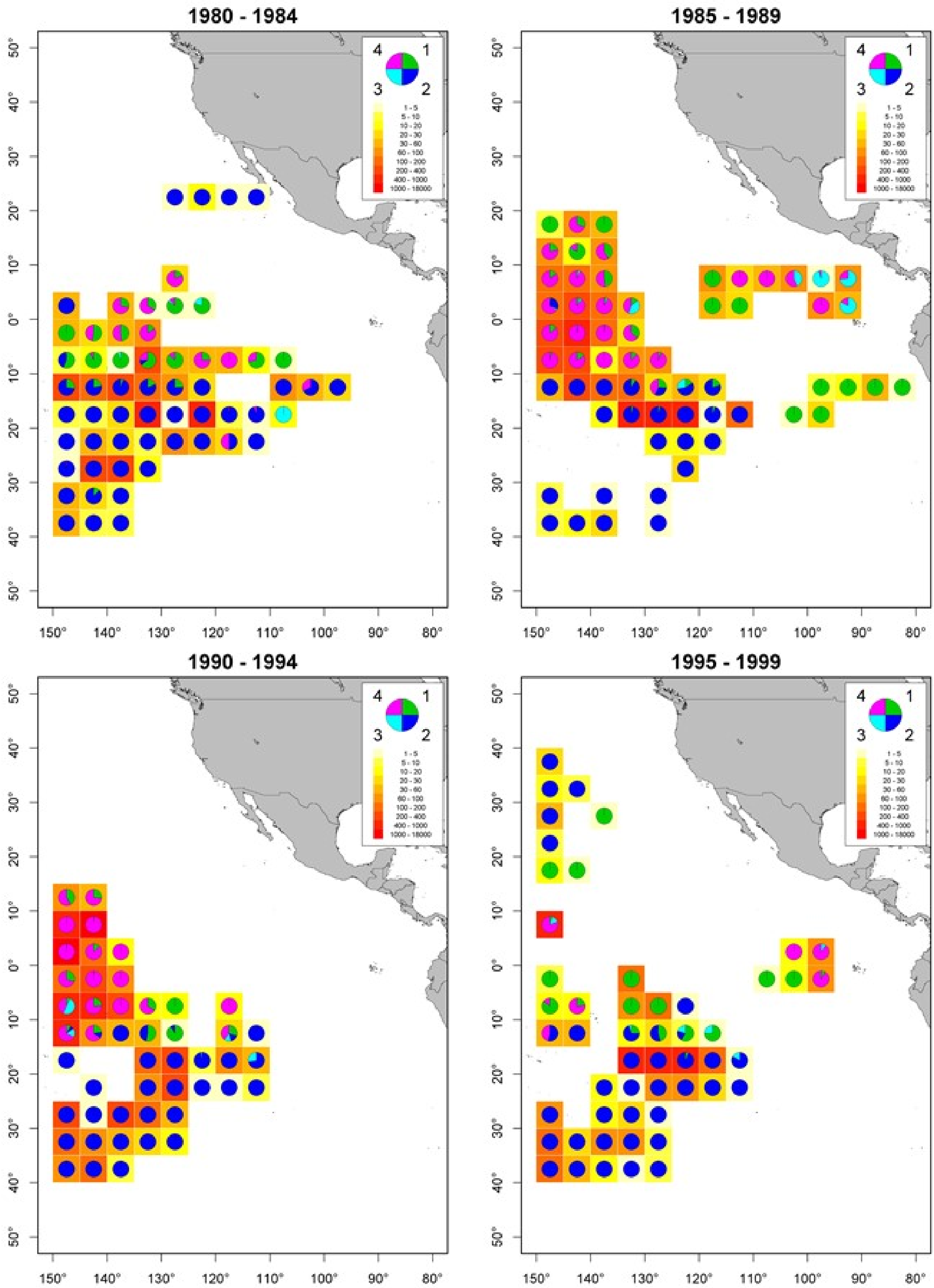


Fig. 8. The distribution of YFT catches of Taiwanese longline vessels operated in the eastern Pacific Ocean. Pie plots show the proportions of YFT catches contributed by clusters and squares with heat colors show the levels of YFT catches.

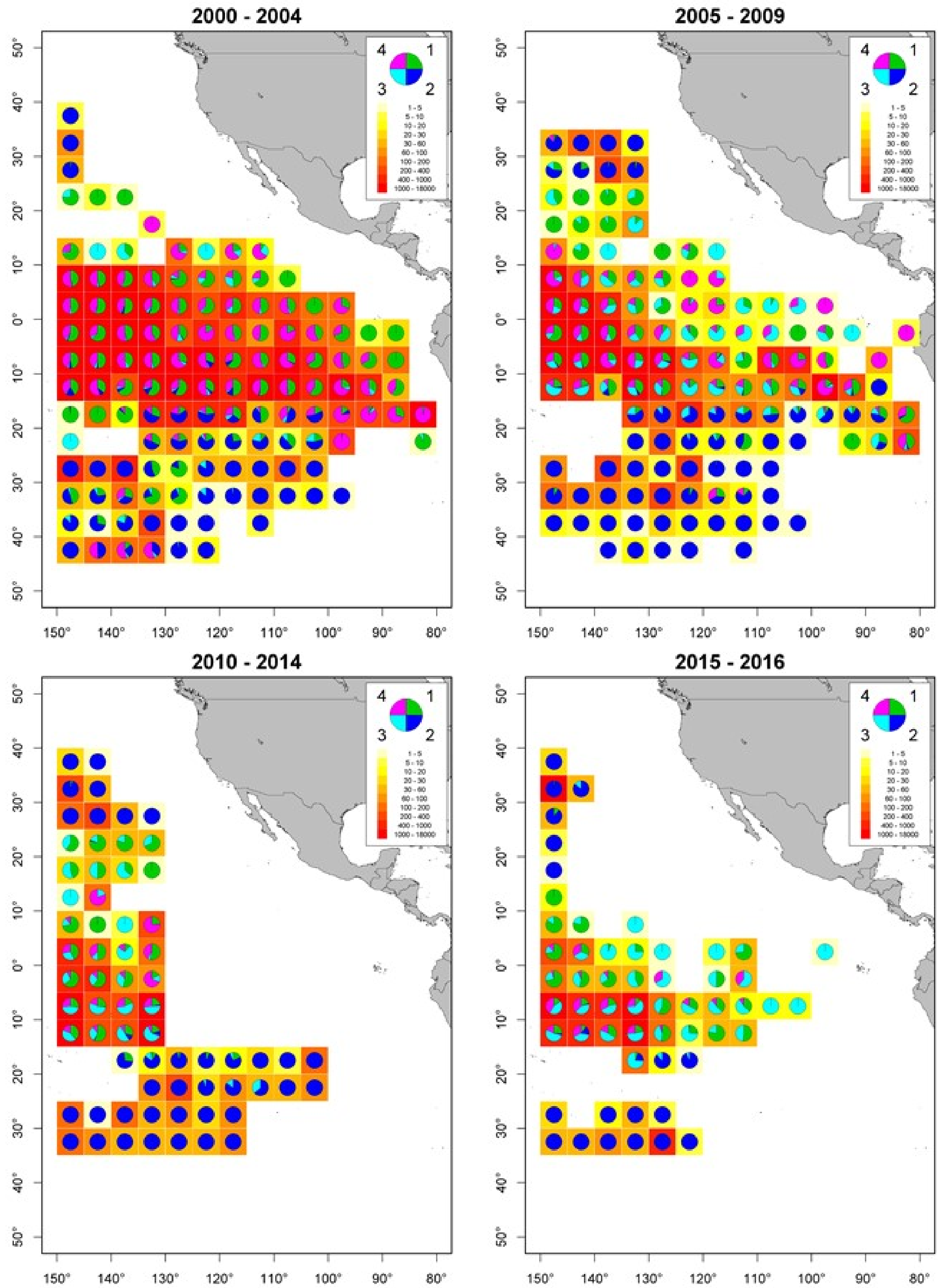
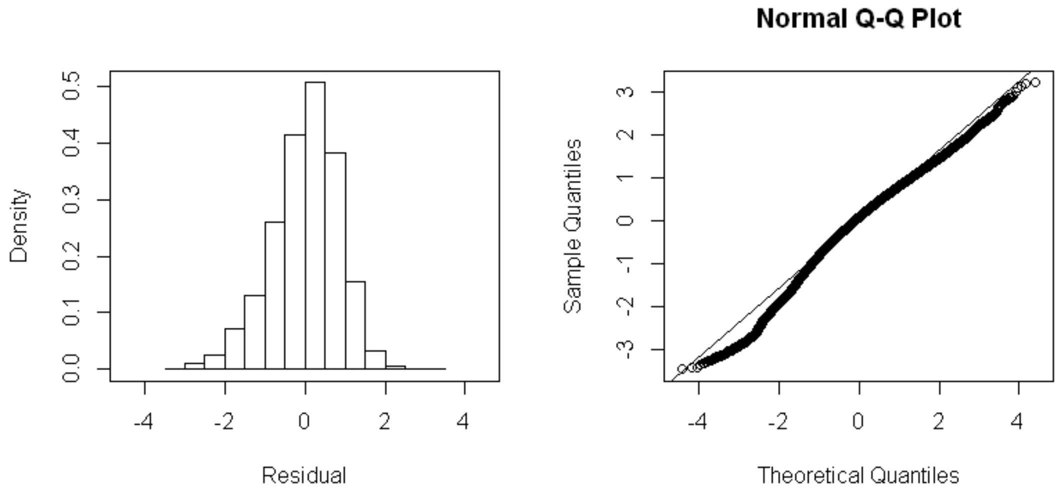


Fig. 8. (Continued).

BET



YFT

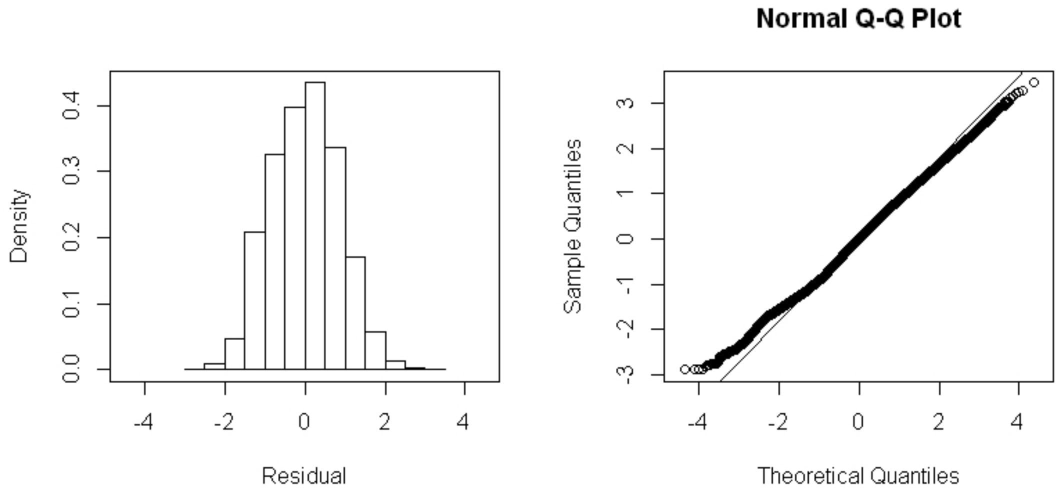
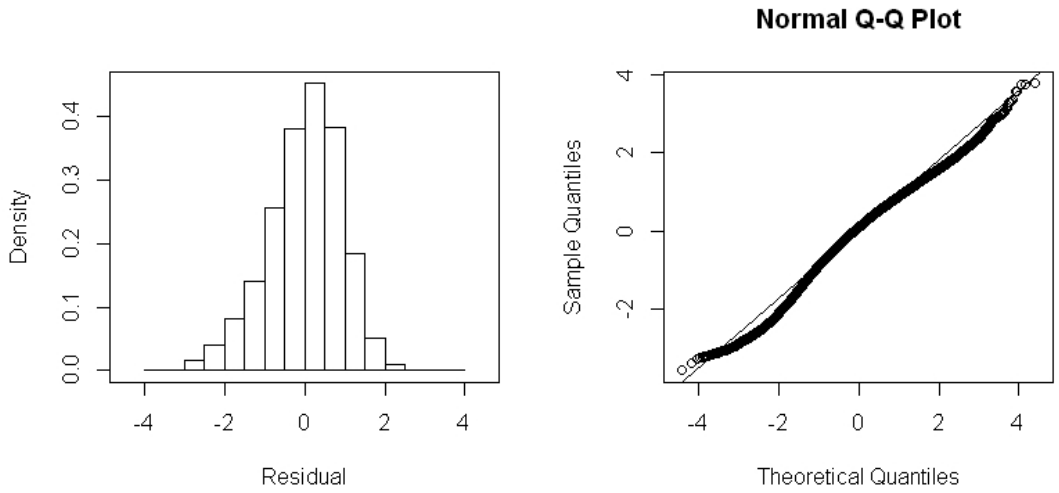


Fig. 9. The frequency distributions and Quantile-Quantile Plots for standardized residuals obtained from lognormal models with Cluster as targeting effect for BET and YFT caught by Taiwanese longline vessels operated in the eastern Pacific Ocean.

BET



YFT

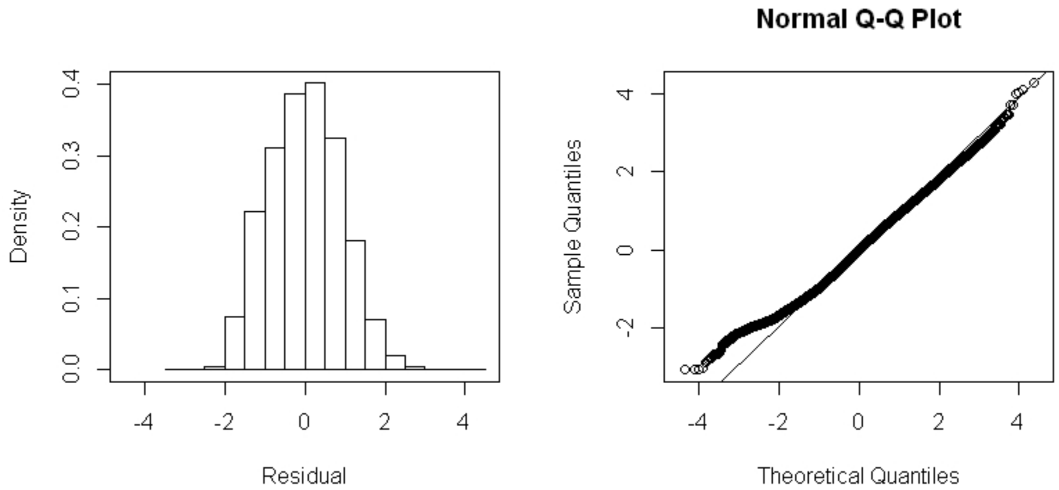


Fig. 10. The frequency distributions and Quantile-Quantile Plots for standardized residuals obtained from lognormal models with NHBF as targeting effect for BET and YFT caught by Taiwanese longline vessels operated in the eastern Pacific Ocean.

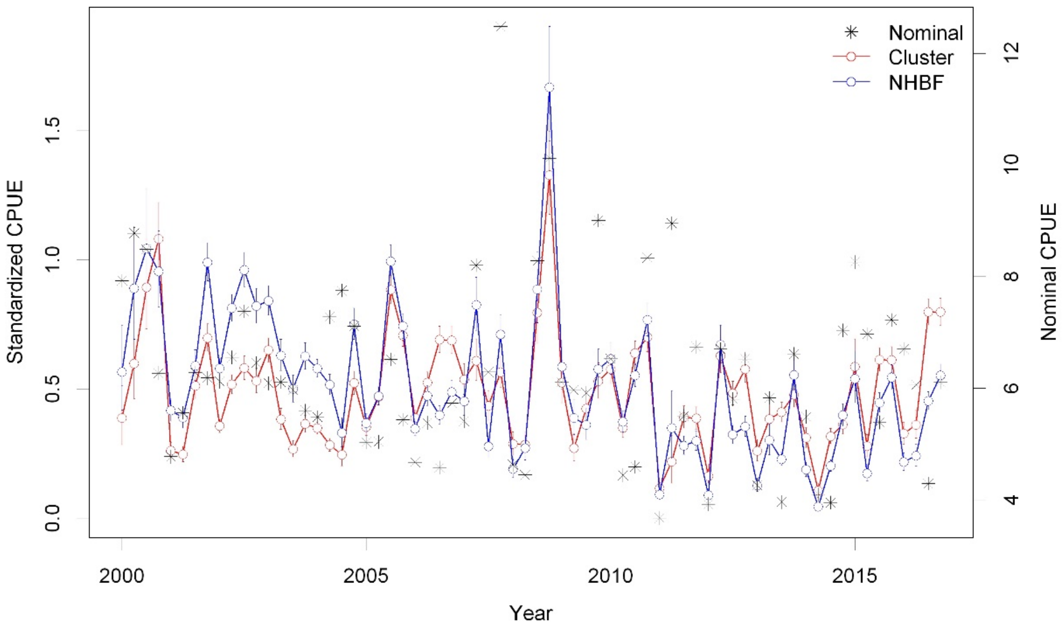


Fig. 11. Standardized CPUE series with 95% confidence intervals obtained from the models with Cluster and NHBF as targeting effects for BET caught by Taiwanese longline vessels operated in the eastern Pacific Ocean.

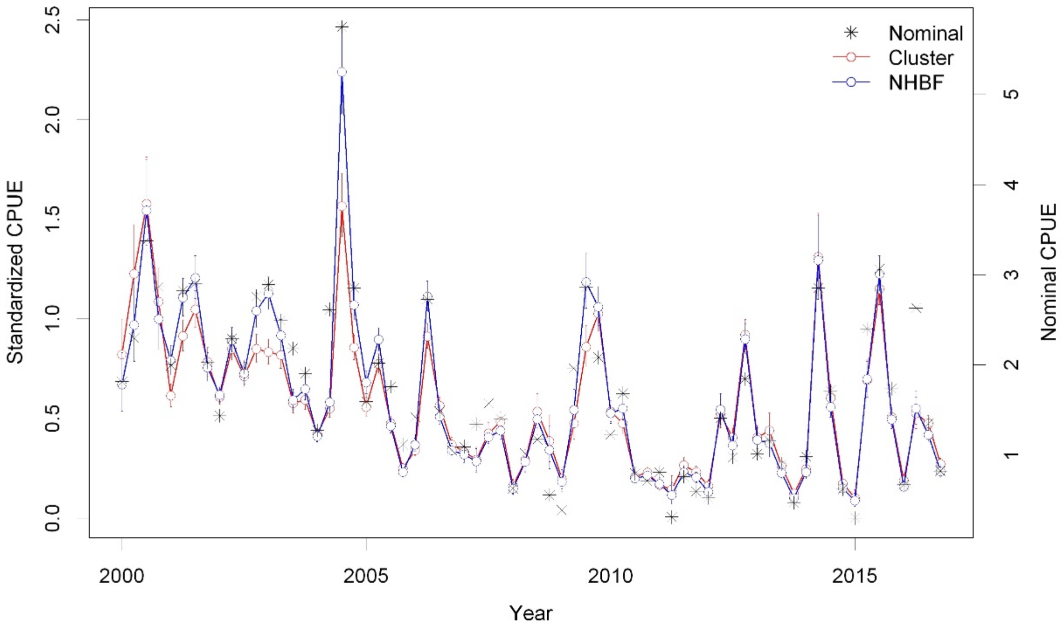


Fig. 12. Standardized CPUE series with 95% confidence intervals obtained from the models with Cluster and NHBF as targeting effects for YFT caught by Taiwanese longline vessels operated in the eastern Pacific Ocean.

Table 1. ANOVA table for the lognormal and delta models with Cluster as targeting effect for BET caught by Taiwanese longline vessels operated in the eastern Pacific Ocean.

Lognormal model

	SS	Df	F	Pr(>F)
YQ	3434	67	72.9	< 2.2e-16 ***
Lon	589	13	64.4	< 2.2e-16 ***
Lat	706	15	66.9	< 2.2e-16 ***
T (Cluster)	13443	2	9556.5	< 2.2e-16 ***
Residuals	67476	95936		

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

Delta model

	LR Chisq	Df	Pr(>Chisq)
YQ	510	67	< 2.2e-16 ***
Lon	165	13	< 2.2e-16 ***
Lat	411	15	< 2.2e-16 ***
T (Cluster)	1109	2	< 2.2e-16 ***

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

Table 2. ANOVA table for the lognormal and delta models with NHBF as targeting effect for BET caught by Taiwanese longline vessels operated in the eastern Pacific Ocean.

Lognormal model

	SS	Df	F	Pr(>F)
YQ	6675	67	118.3	< 2.2e-16 ***
Lon	1134	13	103.6	< 2.2e-16 ***
Lat	2179	15	172.4	< 2.2e-16 ***
T (NHBF)	114	3	45.2	< 2.2e-16 ***
Residuals	80805	95935		

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

Delta model

	LR Chisq	Df	Pr(>Chisq)
YQ	600.4	67	< 2.2e-16 ***
Lon	238.9	13	< 2.2e-16 ***
Lat	367.6	15	< 2.2e-16 ***
T (NHBF)	210.8	3	< 2.2e-16 ***

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

Table 3. ANOVA table for the lognormal and delta models with Cluster as targeting effect for YFT caught by Taiwanese longline vessels operated in the eastern Pacific Ocean.

Lognormal model

	SS	Df	F	Pr(>F)
YQ	5435	67	113.3	< 2.2e-16 ***
Lon	284	13	30.5	< 2.2e-16 ***
Lat	1072	15	99.8	< 2.2e-16 ***
T (Cluster)	7097	2	4957.1	< 2.2e-16 ***
Residuals	53172	74277		

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

Delta model

	LR Chisq	Df	Pr(>Chisq)
YQ	4724.8	67	< 2.2e-16 ***
Lon	763.5	13	< 2.2e-16 ***
Lat	2706.5	15	< 2.2e-16 ***
T (Cluster)	2608.1	2	< 2.2e-16 ***

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

Table 4. ANOVA table for the lognormal and delta models with NHBF as targeting effect for YFT caught by Taiwanese longline vessels operated in the eastern Pacific Ocean.

Lognormal model

	SS	Df	F	Pr(>F)
YQ	8485	67	156.2	< 2.2e-16 ***
Lon	408	13	38.8	< 2.2e-16 ***
Lat	1525	15	125.4	< 2.2e-16 ***
T (NHBF)	56	3	22.9	7.66E-15 ***
Residuals	60214	74276		

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

Delta model

	LR Chisq	Df	Pr(>Chisq)
YQ	5698.6	67	< 2.2e-16 ***
Lon	750.6	13	< 2.2e-16 ***
Lat	3246.3	15	< 2.2e-16 ***
T (NHBF)	87.0	3	< 2.2e-16 ***

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

Table 5. The model selection statistics obtained from lognormal and delta models with Cluster and NHBF as targeting effects for BET and YFT caught by Taiwanese longline vessels operated in the eastern Pacific Ocean.

	Lognormal model			Delta model		
	R^2	AIC	BIC	R^2	AIC	BIC
BET						
Cluster	0.2645	238,838	239,776	0.1461	16,323	17,253
NHBF	0.1192	256,152	257,099	0.0986	17,223	18,162
Diff. (%)	55%	-7%	-7%	33%	-6%	-5%
YFT						
Cluster	0.2550	186,306	187,219	0.1125	96,146	97,077
NHBF	0.1563	195,557	196,479	0.0892	98,670	99,609
Diff. (%)	39%	-5%	-5%	21%	-3%	-3%