CPUE standardization of blue marlin (Makaira nigricans) caught by the Taiwanese large-scale longline fishery in the Indian Ocean using GLM and sdmTMB

Wen-Qi Xu, Chih-Yu Lin, Sheng-Ping Wang*

Department of Environmental Biology and Fisheries Science, National Taiwan Ocean University, Taiwan.
* Corresponding author: wsp@mail.ntou.edu.tw

Paper submitted to the 23th meeting of the Working Party on Billfish, Indian Ocean Tuna Commission, 15-18 September, 2025.

ABSTRACT

This study presents the standardization of catch-per-unit-effort (CPUE) indices for blue marlin (Makaira nigricans) caught by the Taiwanese large-scale longline fishery in the Indian Ocean from 2005 to 2023. Daily logbook data were analyzed for the northwestern (NW) and northeastern (NE) regions, where blue marlin catches are most prevalent. Species targeting clusters were identified using hierarchical cluster analysis, and two CPUE standardization approaches were applied: delta-GLM with various distributions for positive catches, and the spatio-temporal mixed-effects model sdmTMB. Throughout the study period, a high proportion of zero-catch observations was recorded in both areas, highlighting the challenges of standardizing CPUE for bycatch species in this fishery. Both methods produced similar general trends in standardized CPUE, with a pronounced peak in the NW region during 2010–2015 and a subsequent decline, while the NE region showed more stable indices over time. The sdmTMB model provided smoother CPUE trajectories and better accounted for spatial and temporal heterogeneity.

1. INTRODUCTION

Blue marlin (*Makaira nigricans*) is an ecologically and economically significant billfish species inhabiting tropical and subtropical waters of the Indian Ocean. Although blue marlin is generally regarded as a non-target species by both industrial and artisanal fisheries, it remains an important component of regional catches due to its high market value. According to the most recent IOTC assessments, longline

fisheries accounted for approximately 43.8% of the total blue marlin catch in the Indian Ocean between 2019 and 2023, followed by line fisheries (27.4%) and gillnets (23%). The remainder of the catch was attributed to other minor gears, contributing 5.8% of the total catch during this period. The principal fleets included those flagged to Sri Lanka (32%), Taiwan (22%), and India (21%), with the remaining 26 fleets collectively responsible for 34.8% of the recent total catch (IOTC, 2024).

Over the past several decades, blue marlin catches in the Indian Ocean have exhibited marked variability. Drifting longline catches were relatively stable at 3,000 to 4,000 metric tons until the late 1970s, after which they increased steadily, surpassing 8,000 metric tons since the early 1990s. More recently, the average annual catch from 2019 to 2023 was estimated at 7,049 metric tons, which remains below the most recent estimate of maximum sustainable yield (MSY) at 8,740 metric tons (IOTC, 2024). However, the 2022 stock assessment indicated that the Indian Ocean blue marlin stock is overfished and subject to overfishing (IOTC, 2022).

The standardization of catch-per-unit-effort (CPUE) indices is essential for robust stock assessments and effective resource management. Historically, CPUE standardization for the Taiwanese large-scale longline fishery in the Indian Ocean primarily employed generalized linear models (GLMs) (Lin et al., 2022). However, recent methodological advances have enabled the integration of spatial and spatiotemporal variability, which are increasingly recognized as important for accurately modeling relative abundance. In response to these developments, this study aims to standardize CPUE for blue marlin caught by the Taiwanese large-scale longline fishery in the Indian Ocean using both GLM and the spatial dynamic modeling framework sdmTMB. The resulting abundance indices are intended to support future stock assessments and inform management decisions regarding blue marlin in the Indian Ocean.

2. MATERIALS AND METHODS

2.1. Catch and Effort data

This study utilized daily operational catch and effort logbook data for the Taiwanese longline fishery, provided by the Oversea Fisheries Development Council of Taiwan (OFDC). The data were spatially stratified by 1° × 1° longitude and latitude grids and covered the period from 1979 to 2023. For spatial analysis, the four-area stratification originally developed for swordfish by Wang and Nishida (2011) was adopted (Fig. 1).

Consistent with recommendations from previous IOTC meetings (IOTC, 2021), Taiwanese data prior to 2005 were excluded from analyses of fishing operation

targeting and from CPUE standardization for billfish, due to concerns about data quality. These quality issues could have affected not only the accuracy of reported catches of major tropical tunas, but also introduced uncertainties in the catch and effort data for other species, including billfish. Therefore, CPUE standardizations in this study were based on data collected from 2005 to 2023, as recommended by previous IOTC scientific discussions.

2.2. Cluster analysis

The cluster analysis procedures followed the methodology described by Wang et al. (2021). A direct hierarchical clustering approach was used, employing an agglomerative algorithm for efficient computation and memory usage, as implemented in the "fastcluster" package in R (Müllner, 2013; Müllner, 2021). Specifically, Ward's minimum variance method ("ward.D" in the "hclust.vector" function) was applied to squared Euclidean distances calculated from the species composition data.

The optimal number of clusters was determined using the elbow method, which evaluates the change in deviance between and within clusters as the number of clusters increases. The number of clusters was selected when the improvement in the sum of within-cluster variation was less than 10%.

2.3. CPUE Standardization

2.3.1 delta GLM approach

In the Taiwanese large-scale longline fishery, blue marlin is not a primary target species, resulting in a high proportion of zero-catch observations in the operational data. In the context of CPUE standardization, it has been common practice to either exclude zero-catch records or substitute them with a constant. However, recent literature recommends the use of delta methods to address the predominance of zeros. The delta approach consists of two components: a binomial model for the probability of a nonzero catch, and a separate model for the positive catch rates. This methodology is widely recognized as appropriate for datasets characterized by a high frequency of zero catches (Pennington, 1983; Lo et al., 1992; Pennington, 1996; Hinton and Maunder, 2004; Maunder and Punt, 2004; Andrade, 2008; Lauretta et al., 2016; Langley, 2019).

For the standardization of blue marlin CPUE in the Indian Ocean, the positive catch rates were modeled using generalized linear models (GLMs) with alternative skewed error distributions. Model covariates included year, quarter, vessel identification, spatial grid (5° longitude × 5° latitude), and the targeting effect as defined by clusters identified through hierarchical cluster analysis. Significant

interactions among the main effects were also considered in the modeling process. CPUE standardizations were performed separately for each area (Fig. 1).

For positive CPUE observations, the following candidate distributions and link functions were evaluated:

- Lognormal model: The natural logarithm of CPUE was used as the response variable, with an identity link function:
  $\log(CPUE) = \mu + Y + VesselID + Q + LonLat5 + Cluster + \text{interaction} + \varepsilon$,
  $\varepsilon \sim N(0, \sigma^2)$

- Gamma and inverse Gaussian models: CPUE was modeled as the response variable with a log link function:
  $E[CPUE] = \exp(\mu + Y + VesselID + Q + LonLat5 + Cluster + \text{interaction})$

  or equivalently,

  $\log(E[CPUE]) = \mu + Y + VesselID + Q + LonLat5 + Cluster + \text{interaction}$

For the binomial component of the delta model (presence or absence of catches):

$PA = \mu + Y + VesselID + Q + LonLat5 + Cluster + \text{interaction} + \text{interaction} + \varepsilon$,
$\varepsilon \sim$ Binomial distribution

where:

CPUE is defined as the number of fish caught per 1,000 hooks.
μ is the model intercept.
Y is the year effect,
VesselID denotes the vessel effect,
Q is the quarter (seasonal) effect,
LonLat5 represents the effect of the 5° longitude × 5° latitude spatial grid,
Cluster indicates the targeting effect derived from cluster analysis,
interaction refers to significant interaction terms among main effects.

Stepwise regression was performed for model selection based on the Akaike information criterion (AIC), employing both forward selection and backward elimination to determine the optimal set of covariates for each model. The coefficient of determination (R^2) was also calculated to further assess model fit. The AIC values produced by the glm() and glm.nb() functions in R, which are based on the full likelihood including constant terms, permit comparison between models with different

error distributions.

Due to differences in link functions among the candidate models, AIC values from the lognormal model (identity link) are not directly comparable to those from models with log link functions. Therefore, a Jacobian correction was applied to the AIC of the lognormal model to ensure comparability:

$$AIC^J = AIC + \sum \log(CPUE)$$

where:

AIC is the original value from the lognormal model,

AIC^J is the Jacobian-corrected AIC,

$CPUE$ represents the observed CPUE values.

The standardized CPUE series were calculated as the product of the estimated least square means of the year effect from the positive CPUE and delta GLM models. Specifically, the annual index was computed as the product of the standardized CPUE for positive catches and the predicted probability of a nonzero catch:

$$DL^{index} = e^{\log(CPUE)} \times \left(\frac{e^{PA}}{1 + e^{PA}} \right)$$

where: DL^{index} is the standardized CPUE index.

2.3.2 Spatio-temporal modelling approach

The sdmTMB R package (Anderson et al., 2024) was used to fit generalized linear mixed-effects models (GLMMs) that incorporate both spatial and spatio-temporal random fields. This method requires the construction of a stochastic partial differential equation (SPDE) mesh, which enables estimation of a Gaussian random field (GRF) with a Matérn covariance structure (Lindgren et al., 2011). The mesh is composed of triangular elements, and the SPDE is estimated at mesh vertices using integrated nested Laplace approximation (INLA; Rue et al., 2009).

Model structures were implemented in sdmTMB through the Template Model Builder (TMB; Kristensen et al., 2016) framework, with SPDE meshes generated via INLA. This approach enables efficient spatial and spatio-temporal modeling and is well suited for deriving standardized CPUE indices from fishery-dependent data.

A delta-lognormal model structure was applied, consisting of two components: a binomial model for encounter probability and a lognormal model for positive catch rates. Fixed effects included year, quarter, and cluster (as an indicator of targeting behavior) to account for systematic variation in catchability. Spatial random effects and

spatio-temporal interaction terms were included in both components and modeled as Gaussian Markov random fields (GMRFs) defined by the SPDE mesh.

The encounter probability p_i was modeled as:

$$\text{logit}(p_i) = \beta_1(t_i) + \omega_1(s_i) + \varepsilon_1(s_i, t_i) + \delta_1(v_i)$$

The positive catch rate λ_i was modeled as:

$$\log(\lambda_i) = \beta_2(t_i) + \omega_2(s_i) + \varepsilon_2(s_i, t_i) + \delta_2(v_i)$$

where:

$\beta(t_i)$: the intercept for each time t as a fixed effect,

$\omega(s_i)$: spatial random field at location s_i,

$\varepsilon(s_i, t_i)$: spatio-temporal random field at location s_i and time t_i,

$\delta(v_i)$: vessel effect on catchability.

This modeling framework enables flexible estimation of spatial and temporal patterns in relative abundance, providing robust standardized CPUE indices for subsequent stock assessment and management applications.

3. RESULTS AND DISCUSSION

3.1. Fishing trends

Figs. 2 and 3 illustrate the temporal trends in catch (number of fish) and the proportional distribution of blue marlin catches by area, based on logbook data from the Taiwanese large-scale longline fishery operating in the Indian Ocean. The results indicate that blue marlin were predominantly caught in the tropical and subtropical regions of the Indian Ocean, with the majority of the catch consistently originating from the northwestern area (Area NW).

Annual catches varied substantially during the study period, with a notable peak observed around 2014 (Fig. 2). Although catches have generally declined in recent years, the northwestern area continued to contribute the largest proportion of the total catch each year. The northeastern area (Area NE) represented the second most important fishing ground, while catches from the southwestern (Area SW) and southeastern (Area SE) regions remained comparatively low throughout the study period.

The proportional distribution of catch by area (Fig. 3) further demonstrates the dominance of Area NW, which accounted for over 70% of the total annual catch in most years. The relative contribution of Area NE fluctuated between 10% and 40%, while the combined proportion of Areas SW and SE rarely exceeded 5%. These patterns reflect the spatial heterogeneity in blue marlin abundance and fishing effort, with the majority of Taiwanese longline operations concentrated in the northern Indian Ocean.

Figs. 4 and 5 present the temporal changes in fishing effort (number of hooks) and the proportional distribution of effort by area for the Taiwanese large-scale longline fishery in the Indian Ocean. While the total fishing effort was highest in the northwestern area (Area NW) throughout the study period, considerable effort was also allocated to the northeastern (Area NE), southwestern (Area SW), and southeastern (Area SE) regions, especially in recent years. The proportion of effort devoted to the SW and SE areas increased over time, reflecting a gradual spatial expansion of fishing activities.

Despite the widespread distribution of fishing effort, the spatial pattern of blue marlin catches remained highly concentrated. As shown in Figs. 1 and 2, the vast majority of blue marlin were consistently caught in the northern Indian Ocean, particularly in Area NW, with Area NE contributing a smaller but still notable proportion. In contrast, blue marlin catches from the SW and SE regions were minimal across all years, regardless of effort allocation.

This pronounced spatial disparity between fishing effort and catch indicates that blue marlin are predominantly distributed in the northern Indian Ocean. Therefore, subsequent CPUE standardization analyses in this study were restricted to data from the NW and NE areas, where both catch and effort were substantial and the data were considered sufficiently informative for robust modeling.

3.2. Cluster analysis

Cluster analysis was conducted separately for each sub-area using species composition data. The optimal number of clusters was determined by the elbow method, which evaluates the reduction in within-cluster sum of squares as the number of clusters increases. As illustrated in Fig. 6 both the relative sum of squares and the improvement in sum of squares indicated that four clusters represented an appropriate solution for both the northwestern (NW) and northeastern (NE) areas. The criterion for selection was based on the point where the incremental improvement in within-cluster variation fell below 10%, which is consistent with established practices for cluster determination.

The multivariate dispersion plots further visualize the partitioning of fishing sets

into four distinct clusters in each area, as shown in Fig. 7. These plots display the distribution of fishing sets along the first two principal component axes, with each cluster represented by a unique color and confidence ellipse. The resulting clusters reflect differences in species composition and likely correspond to different targeting strategies or fishing grounds utilized by the Taiwanese longline fleet in the Indian Ocean. The identification of four distinct clusters in both NW and NE areas provides a basis for incorporating cluster membership as a proxy for targeting effect in the subsequent CPUE standardization models.

The species compositions exhibited clear differences among clusters (Fig. 8). In the northwestern (NW) area, Cluster 1 was marked by high catches in the early years, mainly of yellowfin tuna and bigeye tuna, along with notable amounts of swordfish and other species. Cluster 2 maintained relatively stable catches over time, with yellowfin tuna and bigeye tuna as the dominant species, supplemented by albacore, swordfish, and sharks. Cluster 3 was unique, with most of the catch classified as "other species," and only minor contributions from yellowfin tuna and bigeye tuna, suggesting different targeting or higher bycatch. Cluster 4 was dominated by yellowfin tuna, which consistently accounted for the largest share of the catch, while other species made up only a small proportion.

In the northeastern (NE) area, each cluster also showed distinct patterns. Cluster 1 had relatively high catches in recent years with a balanced mix of yellowfin tuna, bigeye tuna, albacore, swordfish, and other species. Cluster 2 was characterized by higher catches in the early years, gradually declining and dominated by yellowfin tuna with some bigeye tuna, albacore, and swordfish. Cluster 3 showed a predominance of yellowfin tuna in both total catch and proportion, especially before 2012. Cluster 4 had higher catches before 2010, after which catch levels declined; in this cluster, yellowfin tuna and bigeye tuna together were the main species, but the proportion of yellowfin tuna varied over time, with other species also contributing in some years.

Fig. 9 shows the annual trends in blue marlin catch and fishing effort for the northwestern (NW) and northeastern (NE) areas, grouped by cluster. In both areas, blue marlin catches remained low relative to total fishing effort, consistent with their bycatch status in the fishery. The contribution of each cluster to total catch and effort varied across years, reflecting shifts in fishing strategy or targeting patterns. In the NW area, Cluster 2 generally contributed the most to both catch and effort, while other clusters, such as Cluster 4, showed notable increases during certain years. In the NE area, catch and effort were more evenly distributed among clusters in earlier years, with Clusters 3 and 4 becoming more prominent before 2010. After 2012, both catch and effort decreased across all clusters in both regions.

Despite relatively higher catches of blue marlin in the northern areas, the

proportion of zero-catch observations consistently exceeded 60 percent across all years and clusters (Fig. 10). From 2005 to 2023, both the NW and NE areas exhibited persistently high rates of zero-catch records for blue marlin, regardless of cluster membership. While the proportion of zero-catch declined somewhat between 2011 and 2016, it increased again in recent years and remained substantial throughout the time series. This pattern underscores the challenge of modeling blue marlin as a bycatch species, emphasizing the need for statistical approaches specifically designed for zero-inflated data. Incorporating these considerations into CPUE standardization is essential for producing robust indices for stock assessment and management.

3.3. CPUE standardization

3.3.1 delta GLM approach

CPUE standardizations were separately conducted for only northern areas (NW and NE, Fig. 1) since the catches and CPUE of blue marlin in the southern areas were much lower than those in the northern areas (Figs. 2-10).

Model selection results based on the Akaike Information Criterion (AIC) indicated that the inverse Gaussian error distribution provided the best fit for both the NW and NE areas, as these models yielded the lowest AIC values among the candidate distributions (Table 1). Although the R^2 values for the inverse Gaussian models were slightly lower than those for the gamma models, the inverse Gaussian models demonstrated superior overall performance according to AIC.

Although the inverse Gaussian models were selected based on the lowest AIC values, the residual diagnostic plots (Fig. 11) indicate that model residuals still exhibit some patterns and deviations from ideal model assumptions. While the inverse Gaussian model shows relatively less pronounced trends or heteroscedasticity than the lognormal and gamma models, none of the candidate models achieved fully satisfactory residual diagnostics. These results highlight the inherent challenges of modeling highly zero-inflated bycatch data such as blue marlin. Despite these limitations, the delta-inverse Gaussian model was retained as the most appropriate option among those tested, providing a reasonable basis for standardizing CPUE in both areas for further stock assessment.

The ANOVA results for each area are summarized in Table 2. In both NW and NE regions, year, vessel ID, spatial grid (Lonlat5), cluster, and their interactions were significant factors influencing CPUE. However, the quarter effect was not significant in either the positive catch or delta models. This suggests that after accounting for the main factors such as year, vessel, and spatial effects, seasonal variation no longer plays a major role in explaining CPUE variability. The results indicate that spatial and operational characteristics of the fishery are more important than seasonal changes in

determining blue marlin catch rates.

The standardized CPUE series with 95% confidence intervals from the delta-inverse Gaussian model are presented in Figure 12. In the NW area, CPUE increased steadily from 2005 and peaked in 2015, then declined sharply and remained at lower levels in recent years. In the NE area, CPUE gradually increased until about 2019, followed by a decreasing trend in the most recent years. These results reflect substantial interannual variation in both areas.

3.3.2 Spatio-temporal modelling approach

Spatial prediction maps for the NW region from 2005 to 2023 show that higher blue marlin densities mainly occurred between 10°S and 10°N, especially in the central and western Indian Ocean. These high-density areas were most evident during 2011 and 2015, corresponding to the years with higher standardized CPUE. After 2016, predicted densities generally declined and the distribution became less concentrated.

For the NE region, predicted densities were much lower throughout the study period. Only in the earlier years were there localized high-density patches near the equatorial zone. In recent years, these high-density areas have become rare and scattered, with most of the region showing very low predicted densities.

Overall, blue marlin in the NW region exhibited more pronounced temporal and spatial variation, while the NE region consistently had lower and more dispersed densities across years.

The standardized CPUE time series estimated by the sdmTMB spatial-temporal models for both NW and NE areas are shown in the Fig. 14. In the NW region, CPUE exhibited substantial year-to-year fluctuations, with a clear increase from 2005 that peaked around 2011 and again in 2015, followed by a continuous decline in subsequent years. The confidence intervals were noticeably wider during years with higher CPUE variability, especially during the peak periods, reflecting greater uncertainty in those annual estimates. These pronounced changes may relate to shifts in fishing effort, environmental conditions, or spatial distribution of blue marlin.

In contrast, the NE area demonstrated a more stable trend over time. The CPUE gradually increased from 2005, reached a relatively modest peak around 2015, and then slightly declined and stabilized in recent years. The confidence intervals in the NE area were relatively narrower and more consistent, suggesting lower interannual variability and more stable fishing conditions compared to the NW.

Overall, these results highlight that spatial and temporal variability is critical in CPUE standardization, and that the two areas may experience different population dynamics or fishery impacts, which supports the use of area-specific management and

assessment strategies for blue marlin in the Indian Ocean.

3.4 Comparison of Standardized CPUE Indices

The standardized CPUE time series estimated by the delta-GLM (with inverse Gaussian distribution for positive CPUE) and the sdmTMB spatio-temporal model showed generally similar trends in both the NW and NE regions (Fig. 15). However, several important differences were observed in the details of the trajectories and in the annual fluctuations.

In the NW region, both methods identified pronounced fluctuations and a clear peak in CPUE between 2010 and 2015, followed by a sharp decline in recent years. The delta-GLM results exhibited higher peaks and larger year-to-year variability, while the sdmTMB estimates were more smoothed and less influenced by extreme values. This difference suggests that the sdmTMB model better accounts for spatial and temporal heterogeneity and is less sensitive to anomalously high CPUE values that may result from localized or sporadic fishing events.

In the NE region, both methods produced more stable CPUE trends compared to the NW. The trajectories generally increased from 2005 and then showed a modest decline in recent years. Again, the delta-GLM results exhibited slightly larger annual variability, whereas the sdmTMB estimates were more consistent across years.

The observed differences between the two approaches may be attributed to several key factors, such as spatial and temporal random effects, handling of zero-inflated and extreme data, and model robustness. The sdmTMB model incorporates spatial and spatio-temporal random fields, which allow it to explicitly account for spatial structure and non-independence in the data. This improves its ability to address uneven sampling and heterogeneous fishing effort distributions, especially when vessel distributions change over time. While both methods are designed to handle zero-inflated data, the sdmTMB model's mixed-effects framework provides more flexibility in capturing the complex patterns typical of bycatch species such as blue marlin, especially in the presence of high zero-catch rates and outliers. The delta-GLM is more sensitive to extreme CPUE values, which can amplify short-term fluctuations in the standardized index. In contrast, sdmTMB distributes this variance across random effects, resulting in smoother, and arguably more reliable, CPUE trajectories.

Based on the current results, sdmTMB is clearly more appropriate for standardizing CPUE in this context. The method provides more robust and interpretable trends, effectively addresses spatial and temporal heterogeneity, and is less affected by extreme or anomalous catch events. Given the advantages in both theoretical justification and empirical performance, we recommend using the

sdmTMB-based standardized CPUE indices as the primary basis for future stock assessment and management advice. The delta-GLM approach may be retained for comparison purposes if required by management bodies, but is no longer necessary as a routine part of the analysis. In light of these findings and the demonstrated performance of sdmTMB, we recommend that future standardization of CPUE for blue marlin in the Indian Ocean be based primarily on the sdmTMB framework, with delta-GLM results reserved only as supplementary references if needed.

REFERENCE

Anderson, S.C., Ward, E.J., English, P.A., Barnett, L.A.K., Thorson, J.T., 2024. sdmTMB: an R package for fast, flexible, and user-friendly generalized linear mixed effects models with spatial and spatiotemporal random fields. bioRxiv preprint. https://doi.org/10.1101/2022.03.24.485545.

Andrade, H.A., 2008. Using delta-gamma generalized linear models to standardize catch rates of yellowfin tuna caught by Brazilian bait-boats. ICCAT SCRS/2008/166.

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

IOTC, 2021. Report of the 19th Session of the IOTC Working Party on Billfish. 13–16 September 2021, Microsoft Teams Online. IOTC–2021–WPB19–R[E].

IOTC, 2024. Report of the 27th Session of the IOTC Scientific Committee. 2–6 December 2024, South Africa. IOTC–2024–SC27–R[E].

Kristensen, K., Nielsen, A., Berg, C.W., Skaug, H., Bell, B.M., 2016. TMB: Automatic Differentiation and Laplace Approximation. J. Stat. Softw. 70(5), 1-21. https://doi.org/10.18637/jss.v070.i05.

Langley, A.D., 2019. An investigation of the performance of CPUE modelling approaches – a simulation study. New Zealand Fisheries Assessment Report 2019/57.

Lauretta, M.V., Walter, J.F., Christman, M.C., 2016. Some considerations for CPUE standardization; variance estimation and distributional considerations. ICCAT Collect. Vol. Sci. Pap. ICCAT, 72(9): 2304-2312.

Lin, C.Y., Wang, S.P.*, Xu, W.Q. (2022). CPUE standardization of blue marlin (*Makaira nigricans*) caught by Taiwanese large-scale longline fishery in the Indian Ocean. The 20th Session of the IOTC Working Party on Billfish, Indian Ocean Tuna Commission, 12–15 September 2022, Online. IOTC-2022-WPB20-09.

Lindgren, F., Rue, H., Lindström, J., 2011. An Explicit Link between Gaussian Fields

and Gaussian Markov Random Fields: The Stochastic Partial Differential Equation Approach. J. R. Stat. Soc. Ser. B-Stat. Methodol. 73(4), 423-498. https://doi.org/10.1111/j.1467-9868.2011.00777.x

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.

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

Müllner, D., 2013. fastcluster: Fast Hierarchical, Agglomerative Clustering Routines for R and Python. Journal of Statistical Software, 53(9): 1-18.

Müllner, D., 2021. The fastcluster package: User's manual, Version 1.2.3. https://cran.r-project.org/web/packages/fastcluster/vignettes/fastcluster.pdf

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.

Rue, H., Martino, S., Chopin, N., 2009. Approximate Bayesian Inference for Latent Gaussian Models by Using Integrated Nested Laplace Approximations. J. R. Stat. Soc. B 71(2), 319-392. https://doi.org/10.1111/j.1467-9868.2008.00700.x

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., Xu, W.Q., Lin, C.Y., Kitakado, T., 2021. Analysis on fishing strategy for target species for Taiwanese large-scale longline fishery in the Indian Ocean. OTC–2021–WPB19–11.

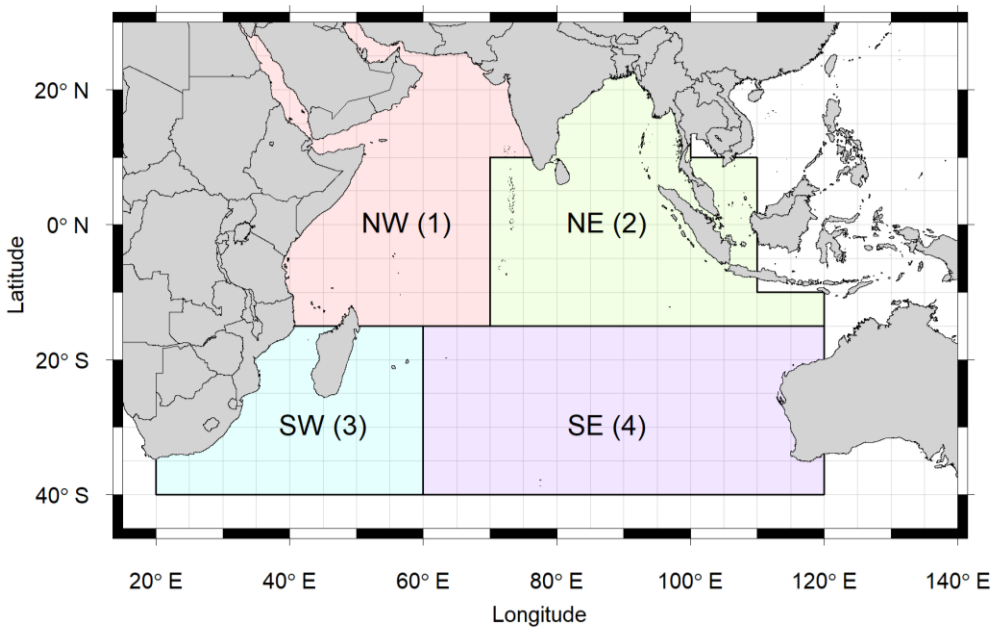


Fig. 1. Area stratification for swordfish in the Indian Ocean.

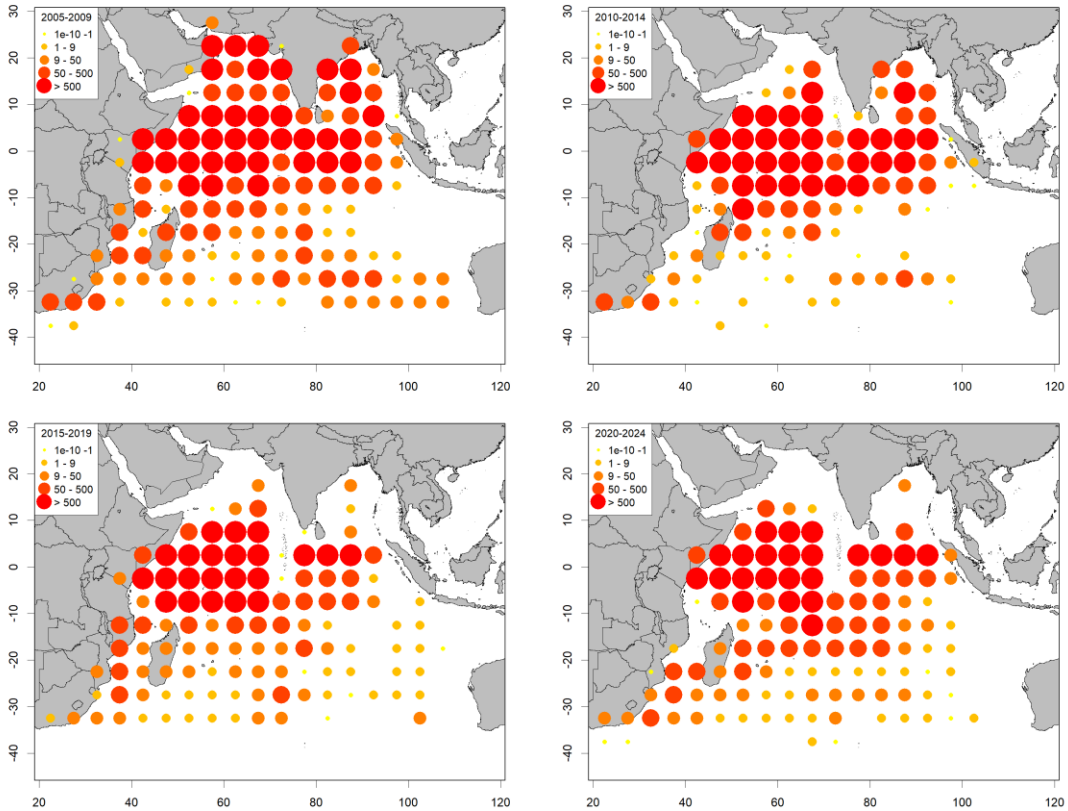


Fig. 2. Blue marlin catch distribution of Taiwanese large-scale longline fishery in the Indian Ocean.

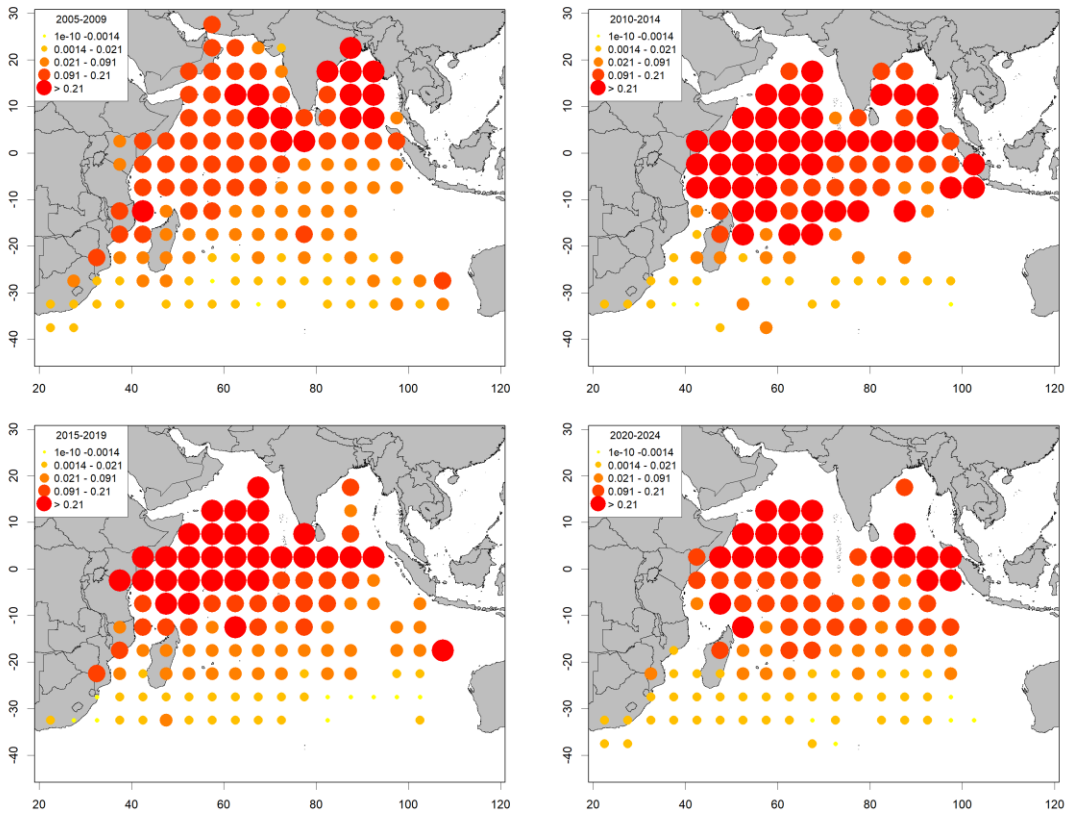


Fig. 3. Blue marlin CPUE distribution of Taiwanese large-scale longline fishery in the Indian Ocean.

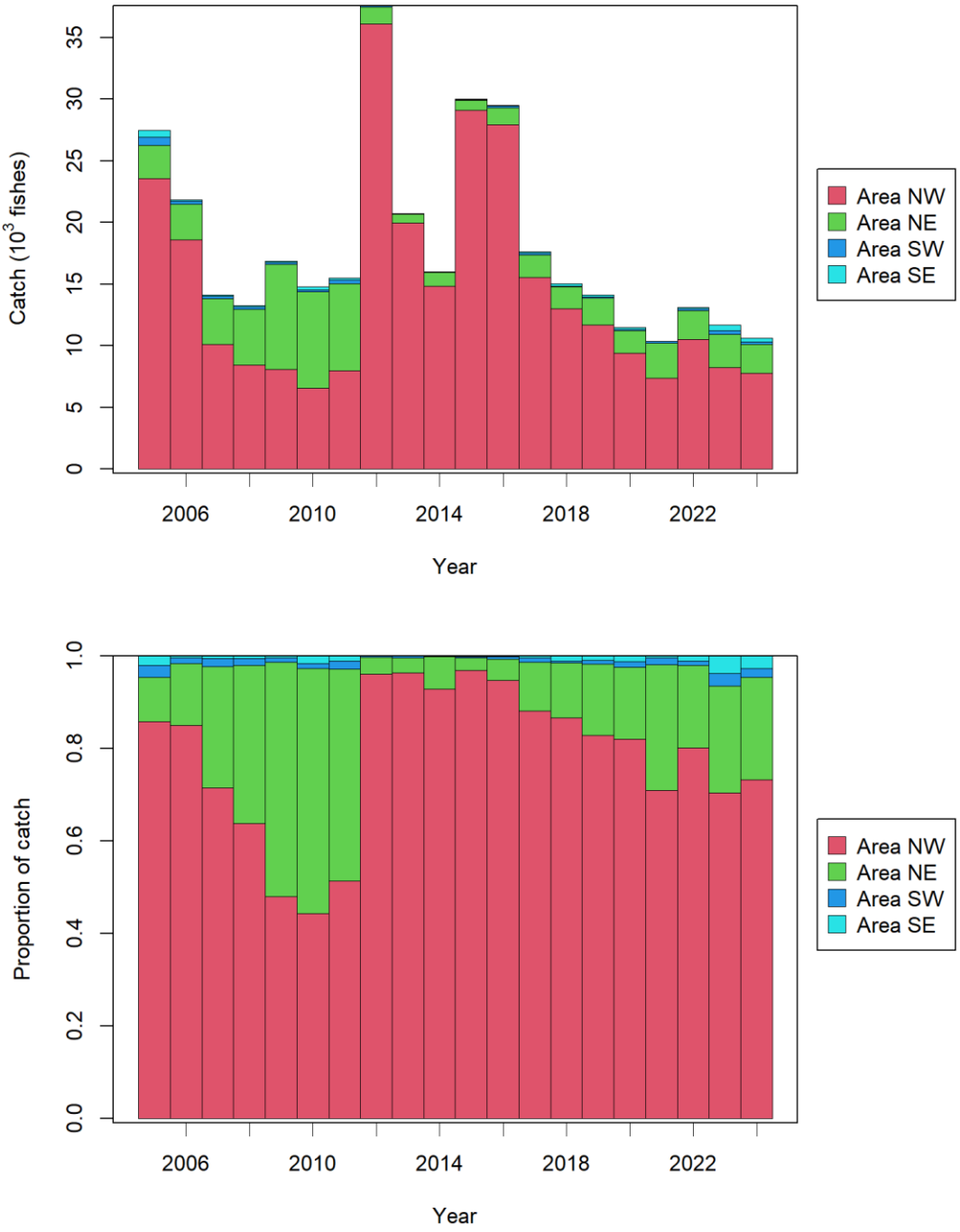


Fig. 4. Annual blue marlin catches of Taiwanese large-scale longline fishery in the Indian Ocean.

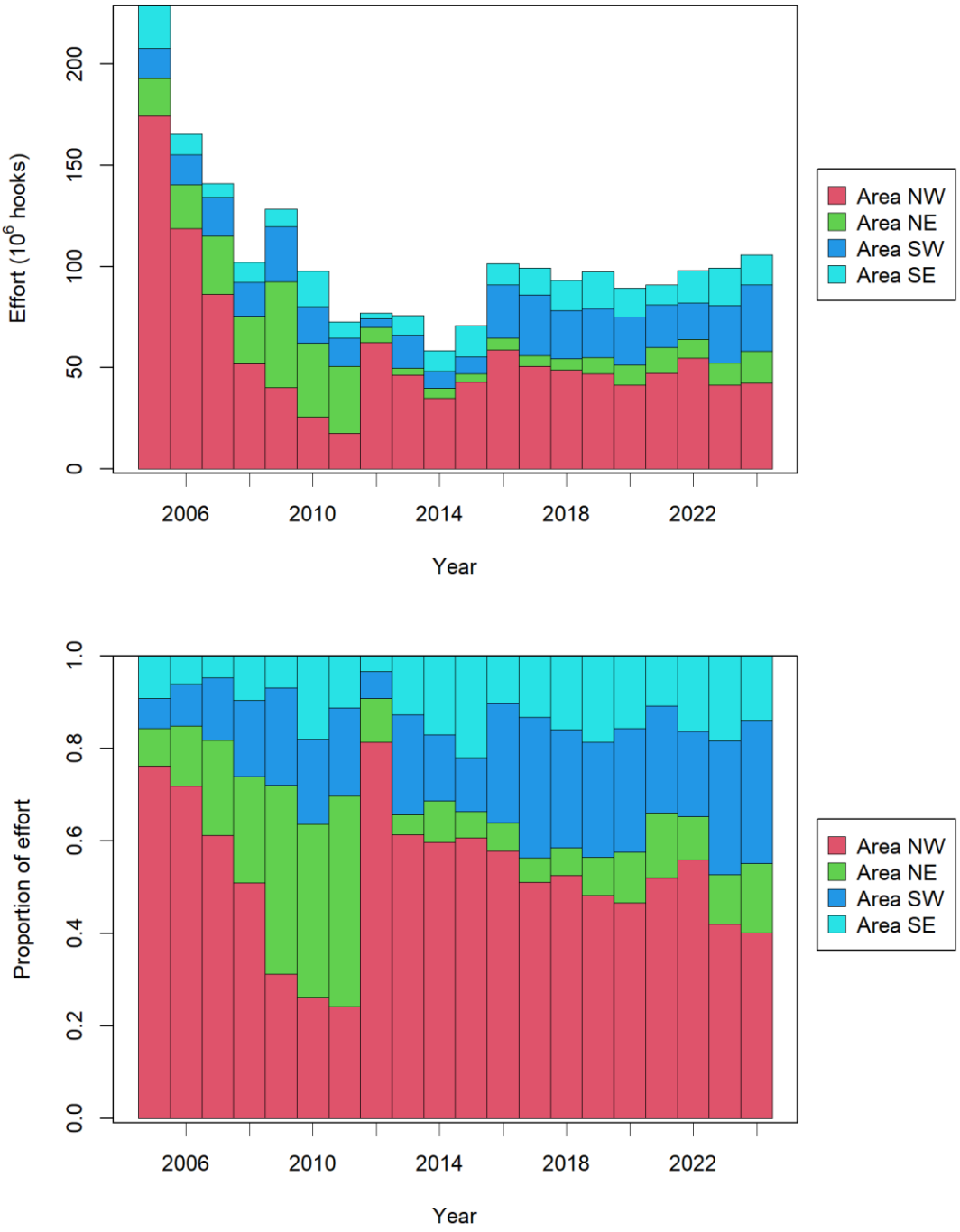
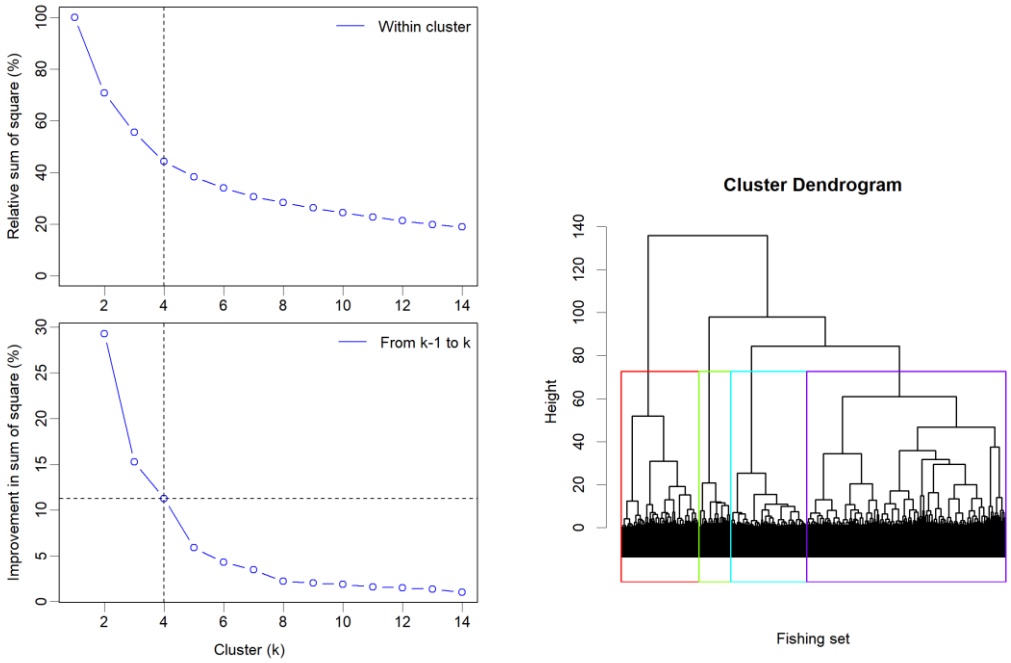


Fig. 5. Annual efforts (number of hooks) of Taiwanese large-scale longline fishery in the Indian Ocean.

NW



NE

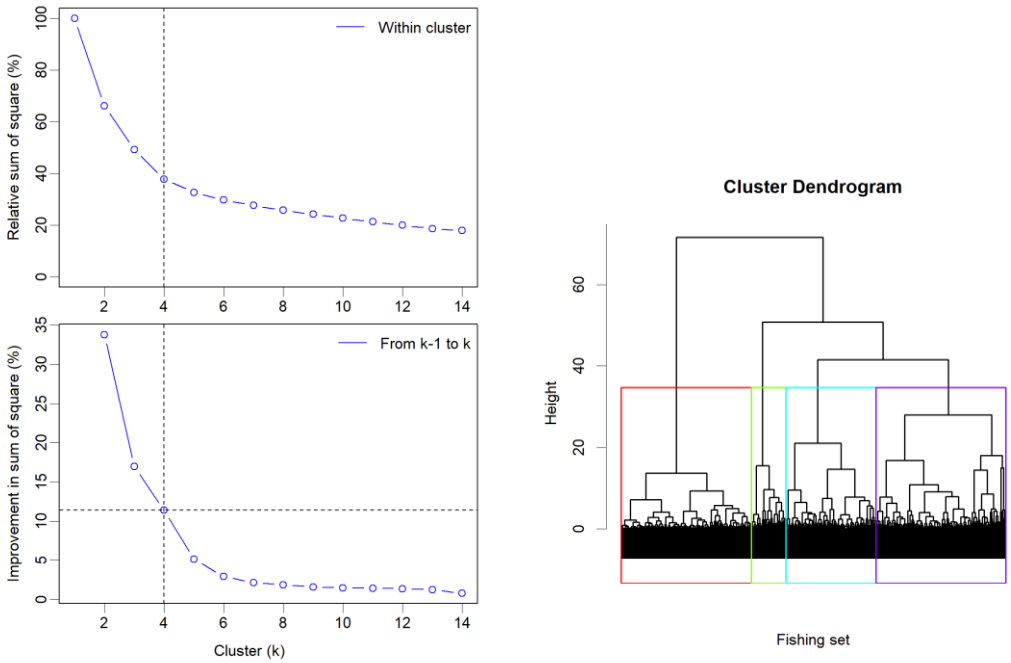
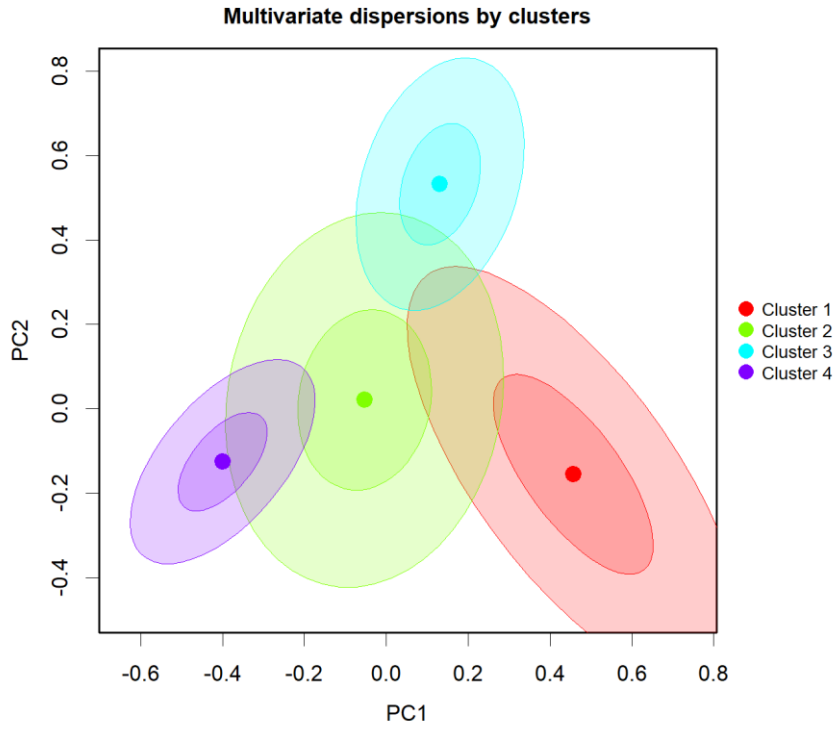


Fig. 6. Sum of squares within clusters for the data of Taiwanese large-scale longline fishery in billfish area of the Indian Ocean.

NW



NE

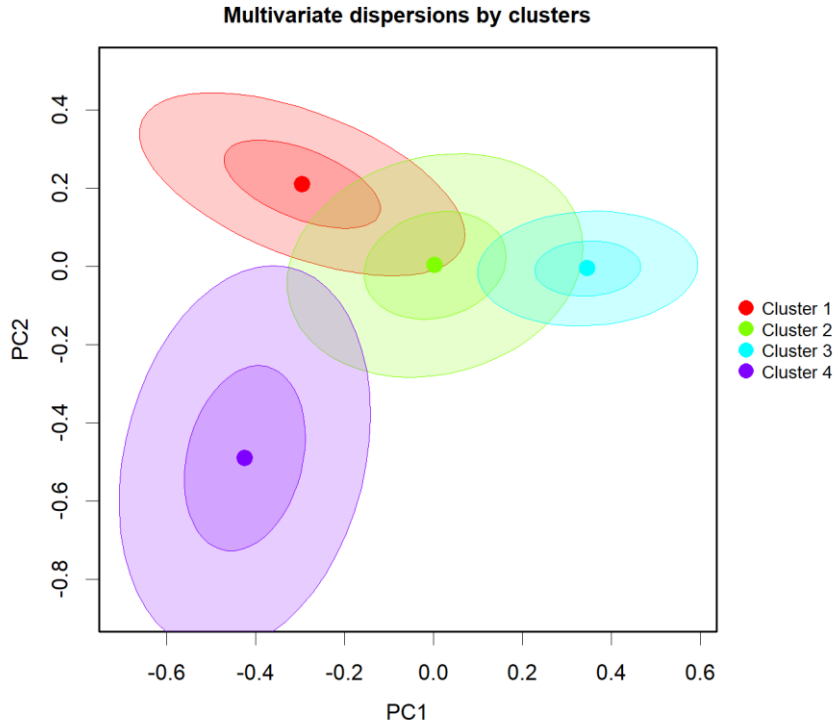


Fig. 7. Multivariate dispersions of the centroids by clusters derived from PCA for the data of Taiwanese large-scale longline fishery in billfish area of the Indian Ocean.

NW

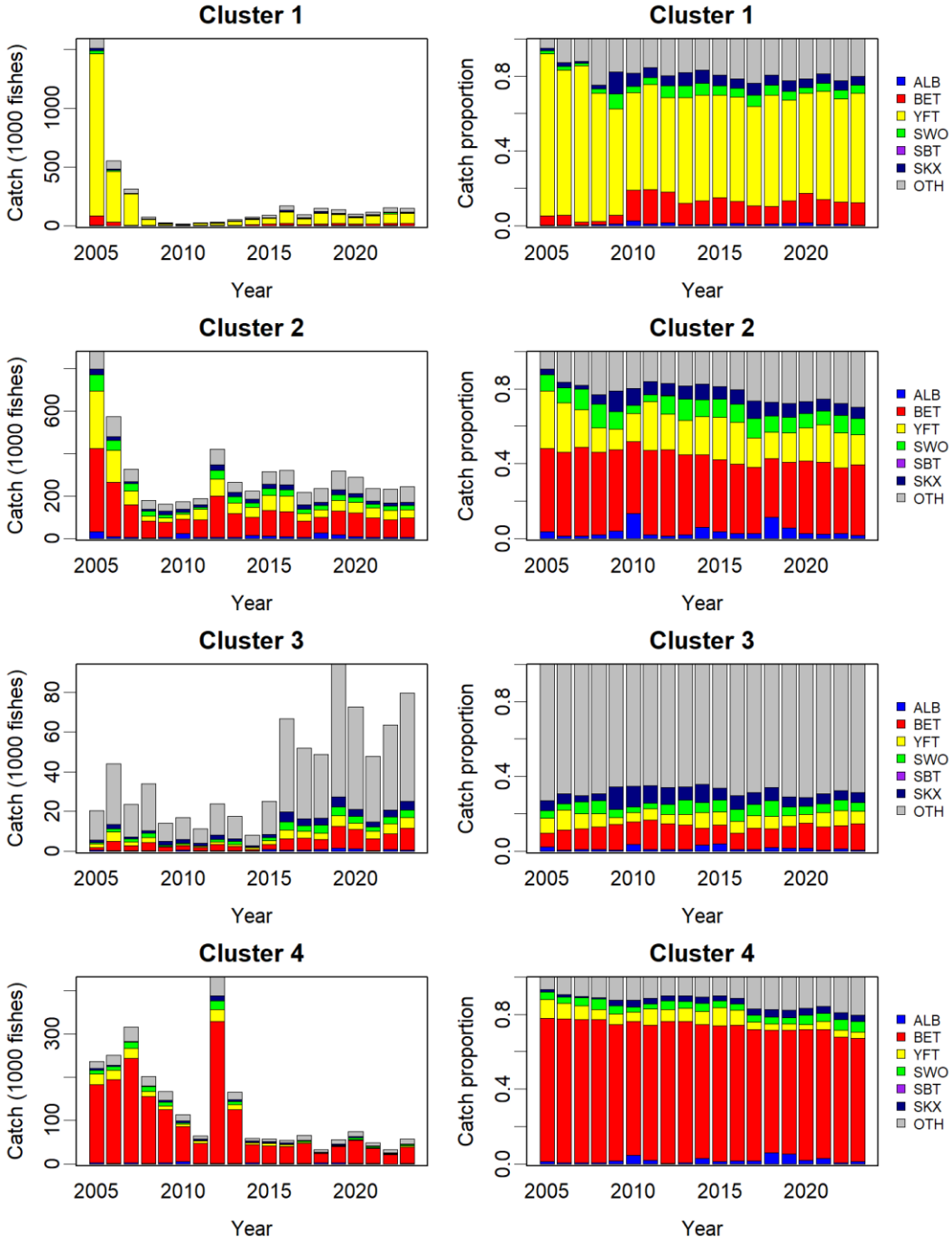


Fig. 8. Annual catches and compositions by species for each cluster of Taiwanese large-scale longline fishery in billfish area of the Indian Ocean.

NE

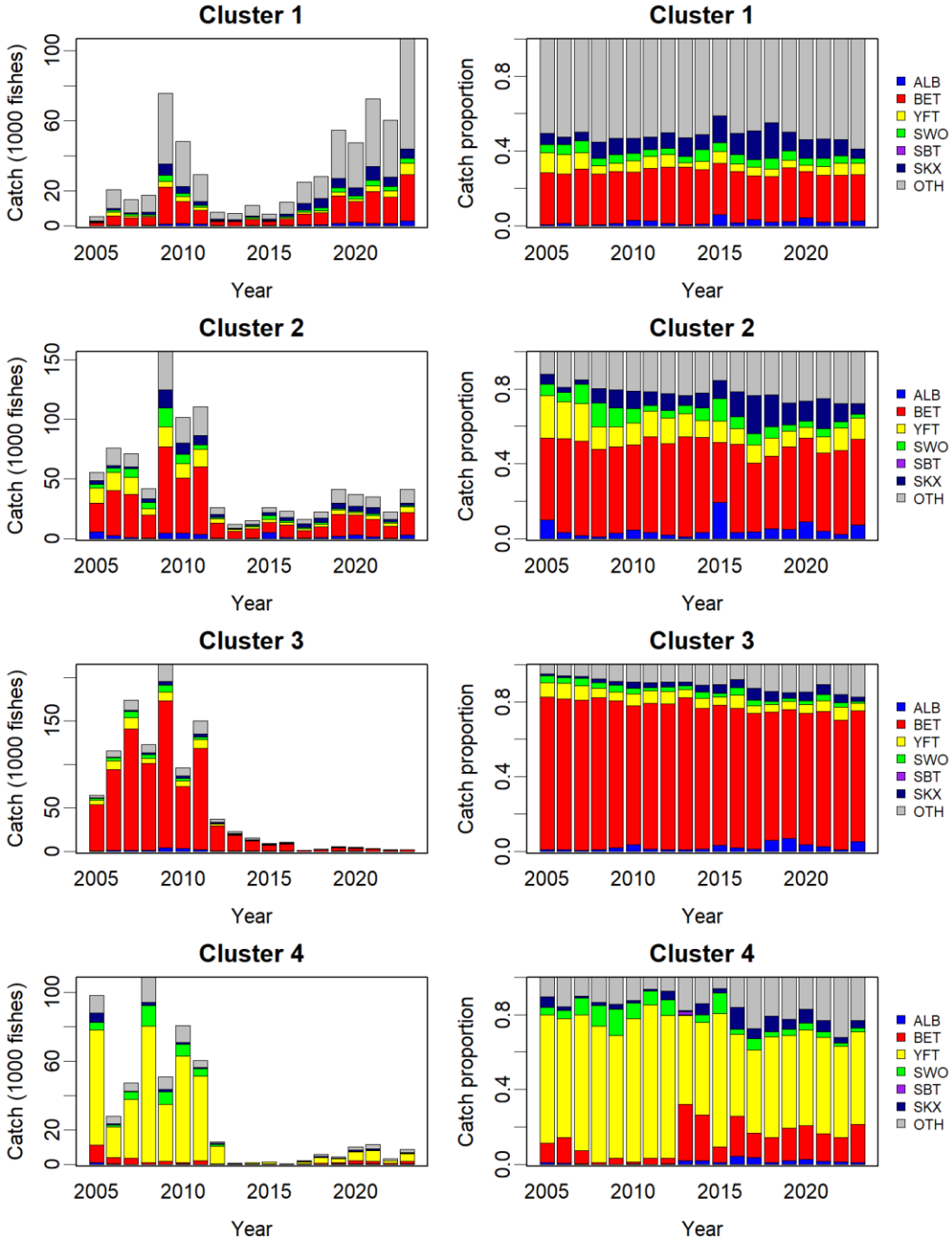


Fig. 8. (Continued).

NW

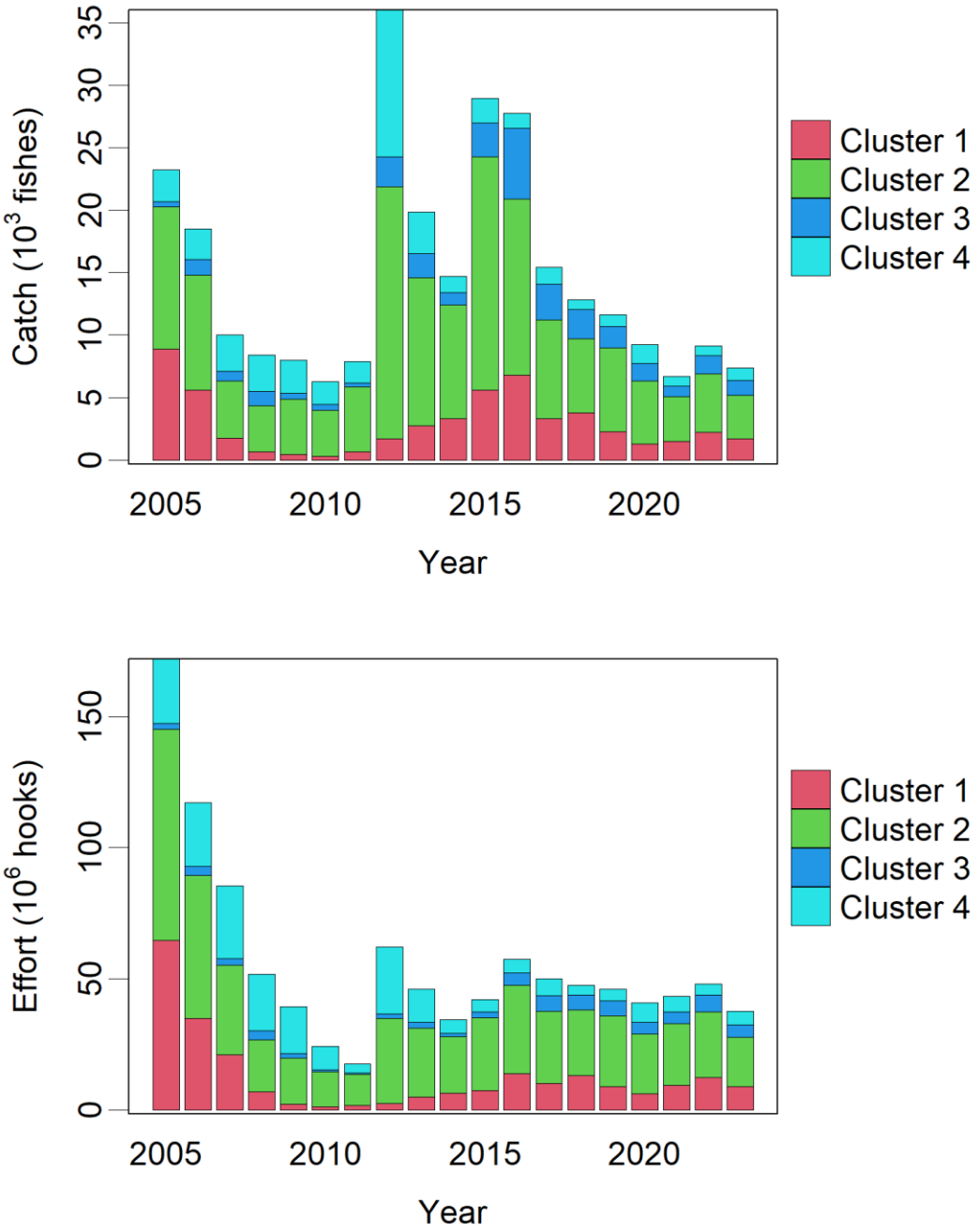


Fig. 9. Annual blue marlin catches and efforts for each cluster of Taiwanese large-scale longline fishery in billfish area of the Indian Ocean.

NE

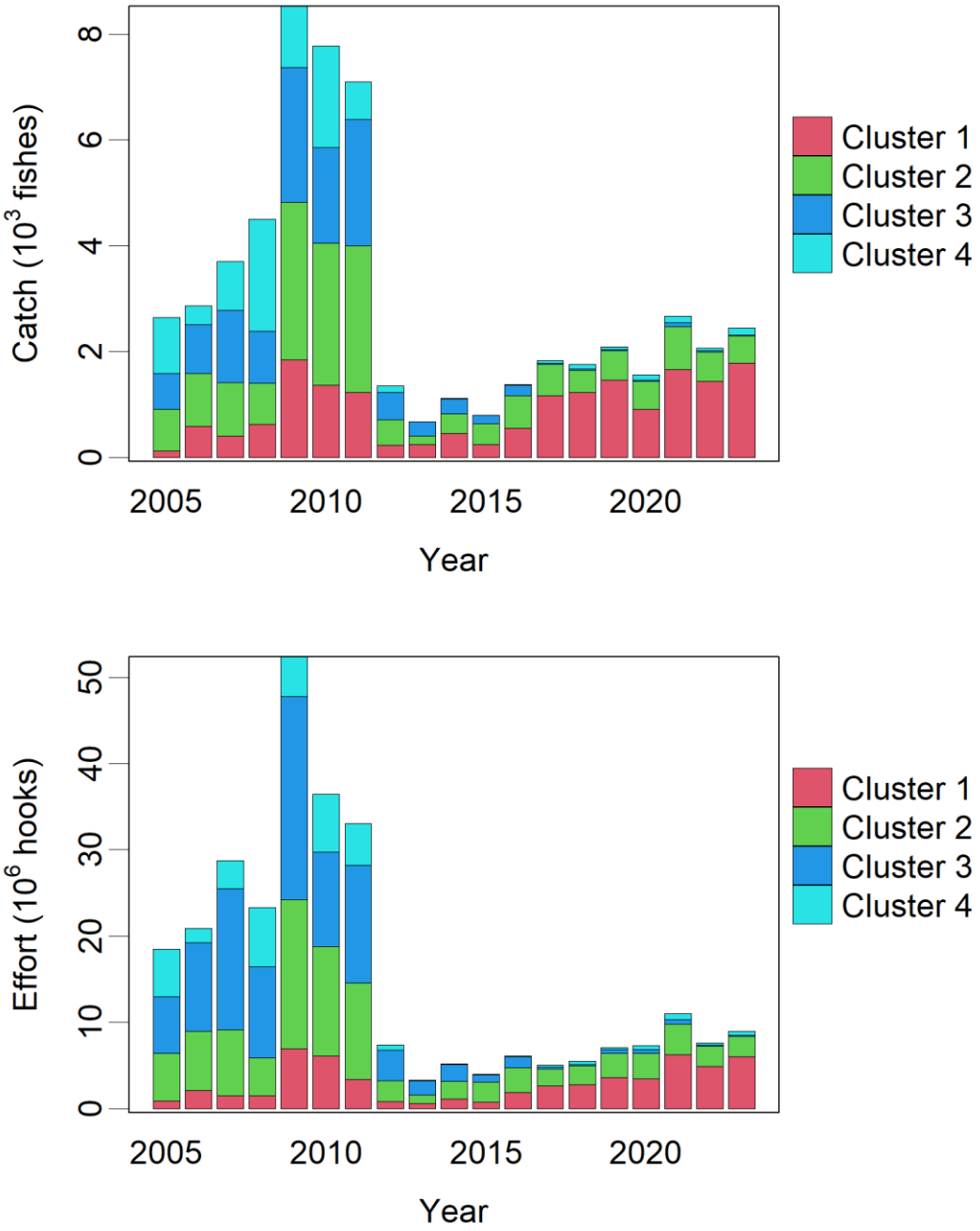
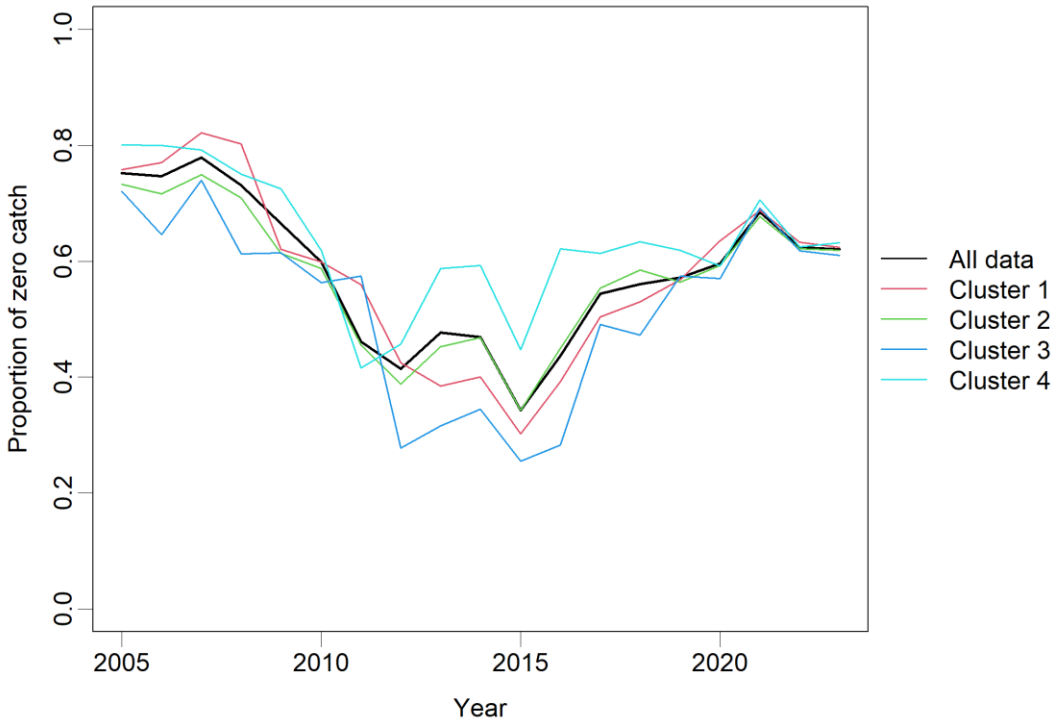


Fig. 9. (Continued).

NW



NE

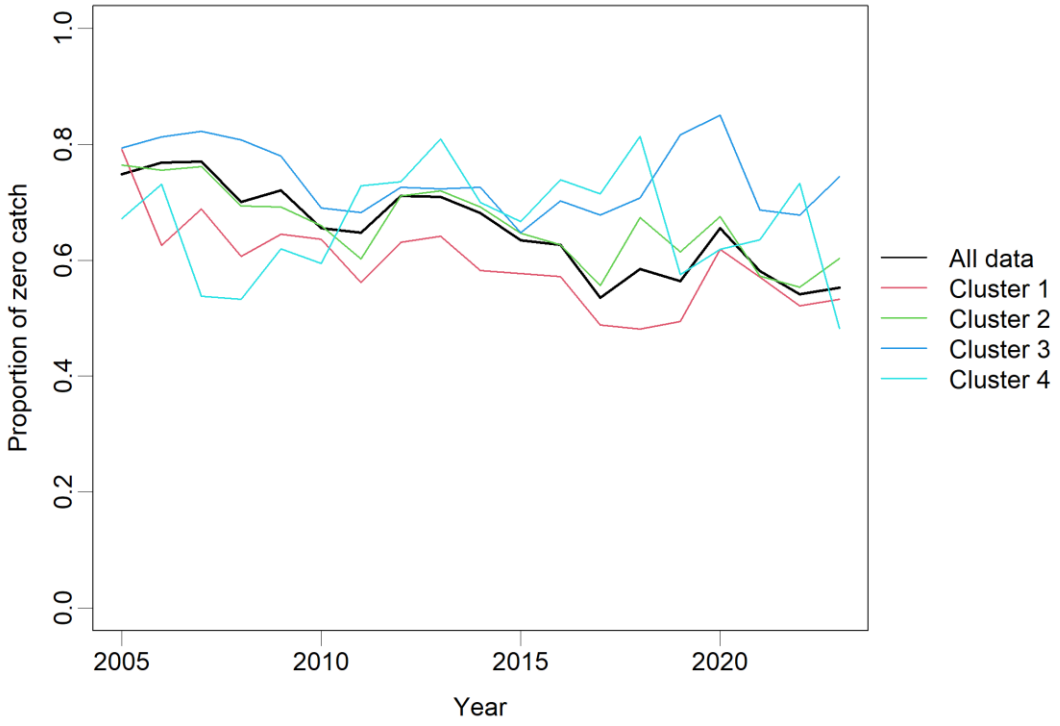
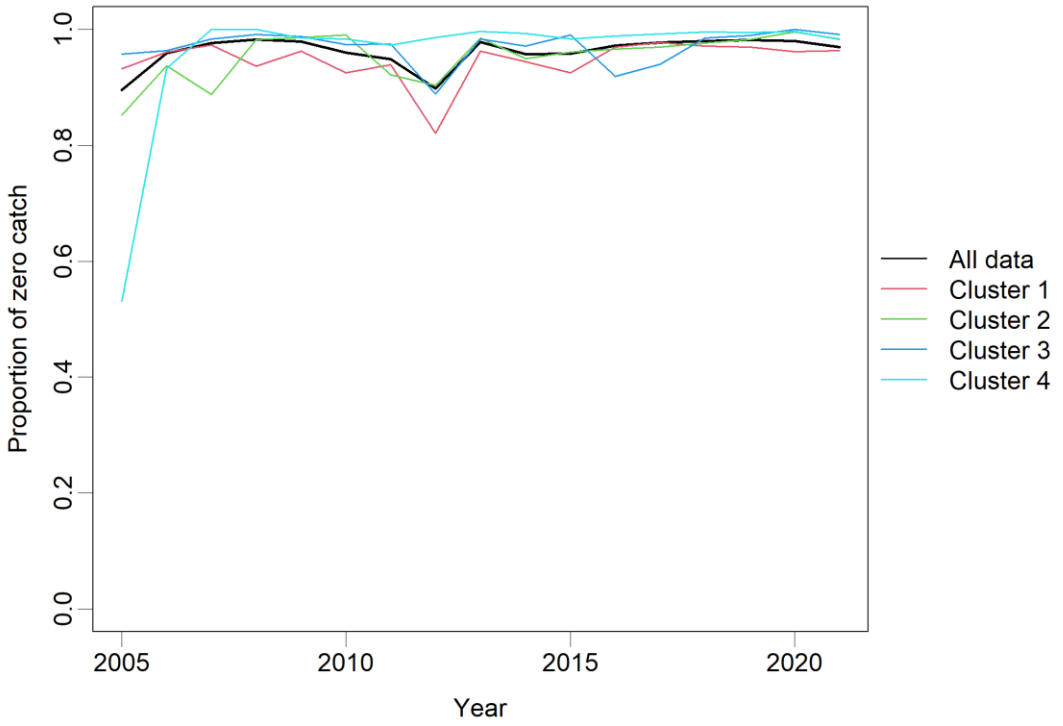


Fig. 10. Annual zero proportion of blue marlin catches for each cluster of Taiwanese large-scale longline fishery in billfish area of the Indian Ocean.

SW



SE

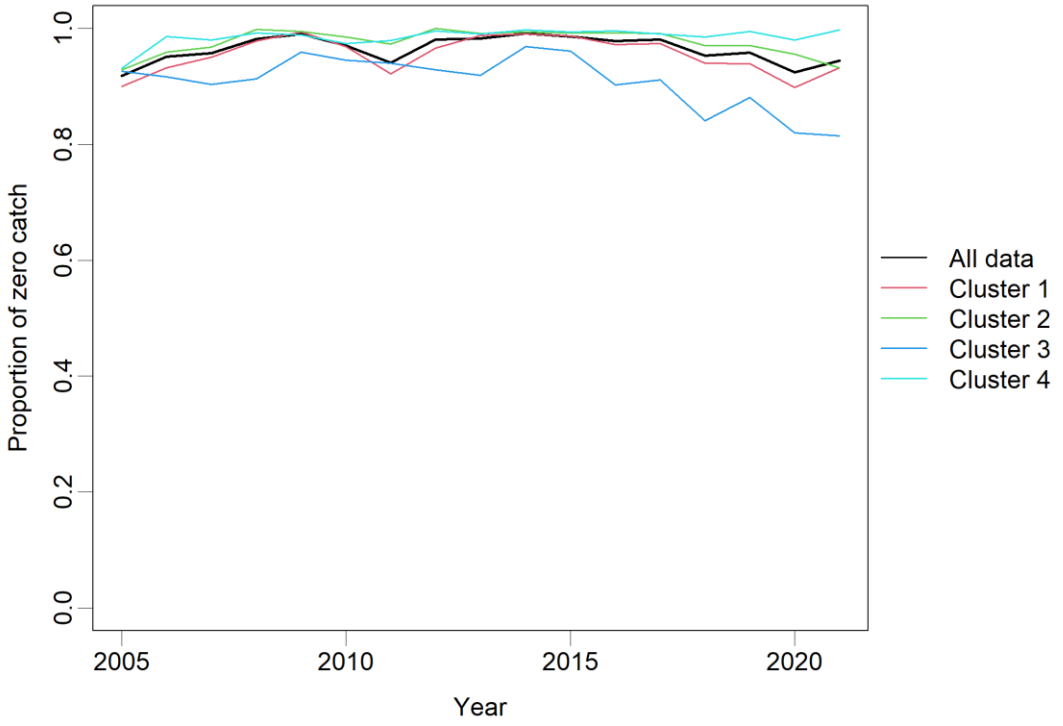


Fig. 10. (Continued).

NW

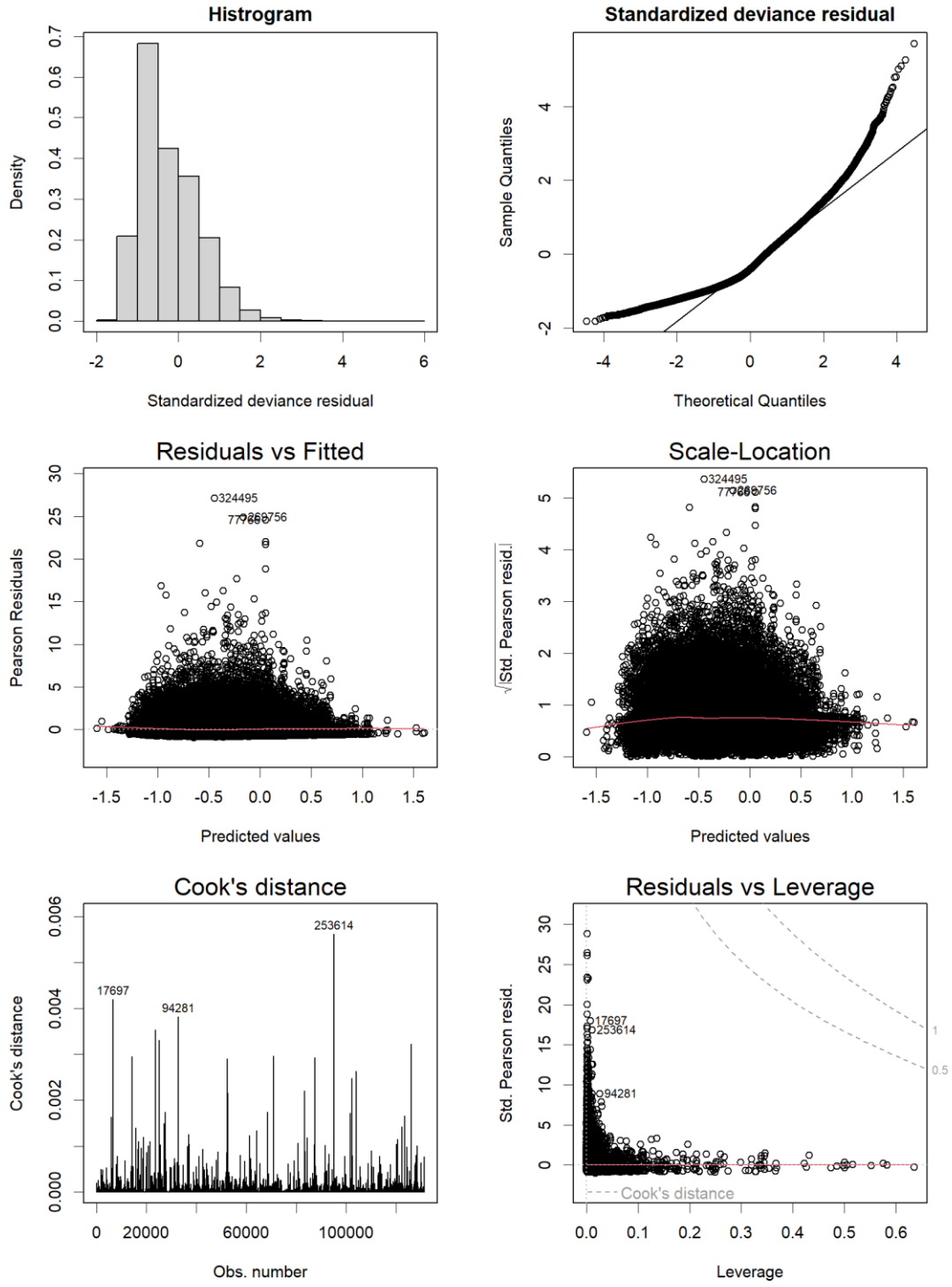


Fig. 11. Diagnostic plots for GLMs with inverse gaussian error distribution assumption for blue marlin caught by Taiwanese large-scale longline fishery in the Indian Ocean from 2005 to 2021.

NE

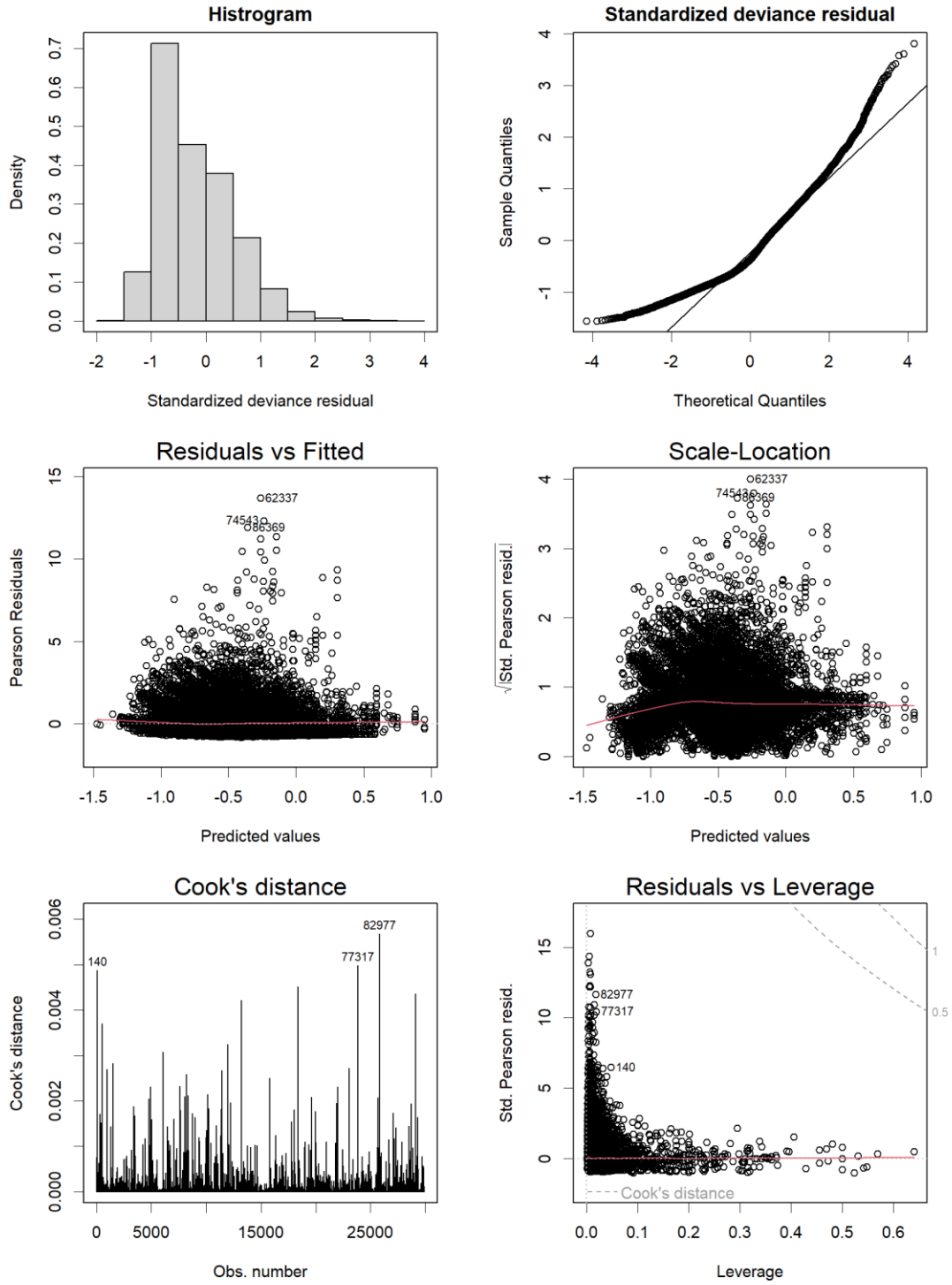


Fig. 11. (Continued).

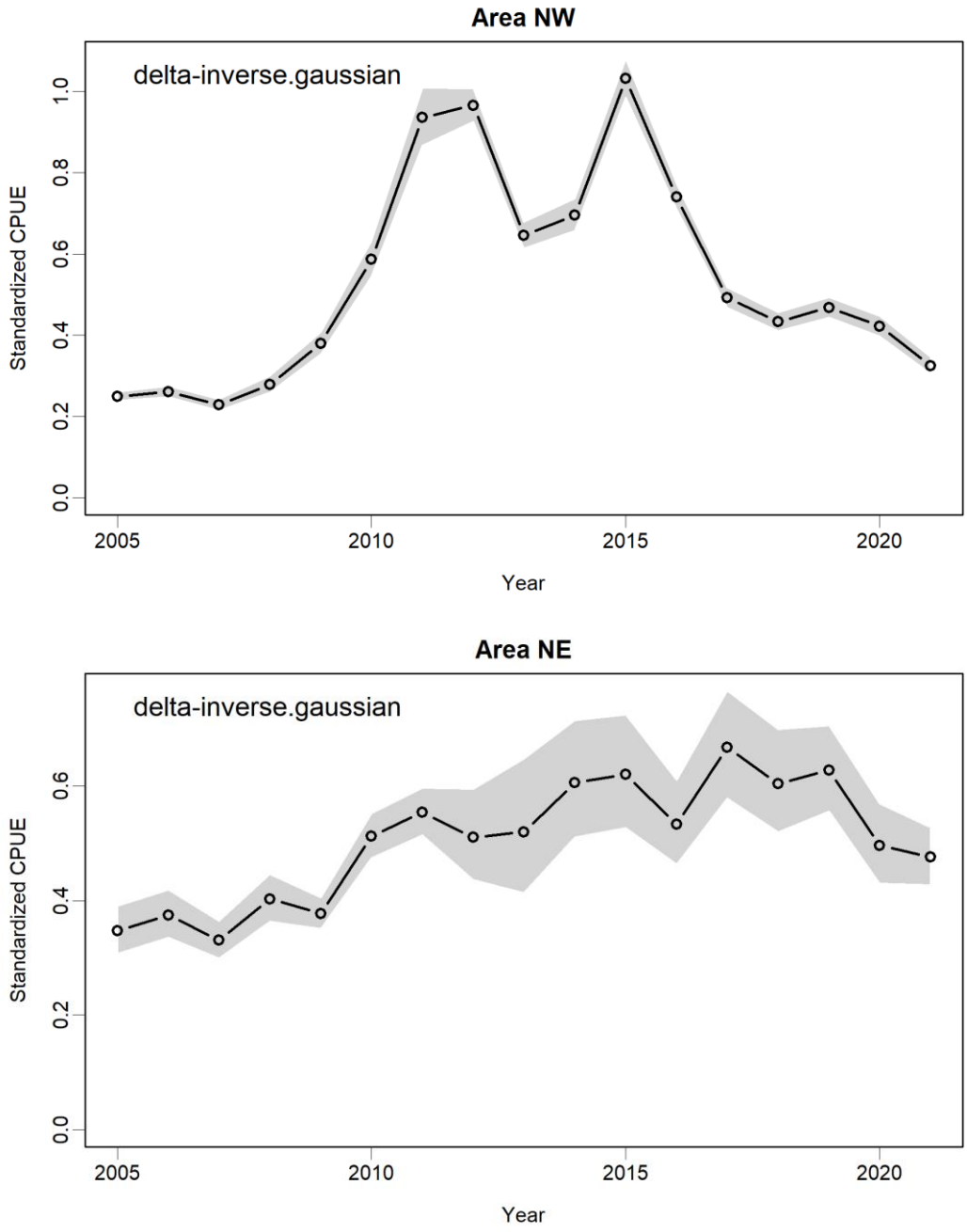


Fig. 12. Standardized CPUE series with 95% confidence intervals based on selected model for blue marlin caught by Taiwanese large-scale longline fishery in the Indian Ocean from 2005 to 2021.

NW

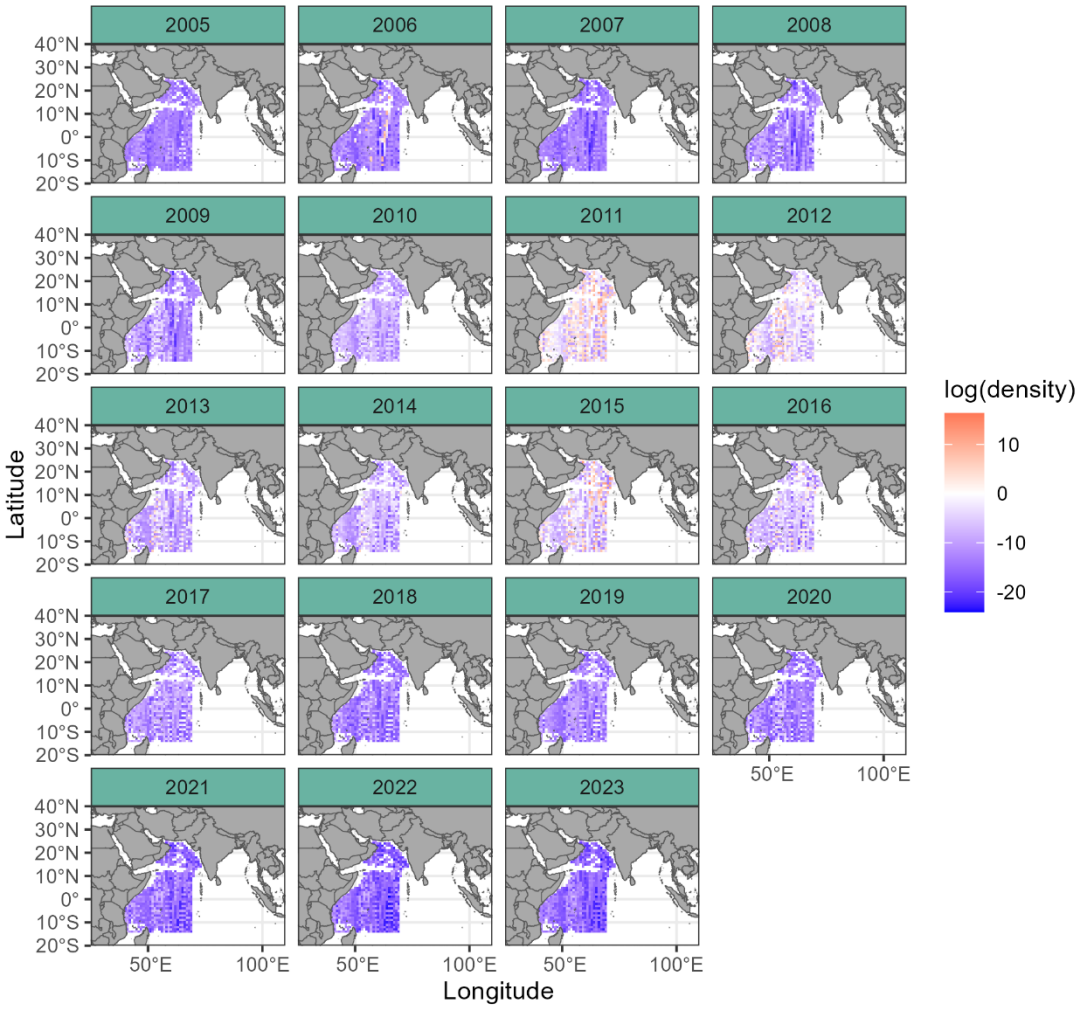


Fig. 13. Spatio-temporal distribution of predicted log density from 2005 to 2023.

NE

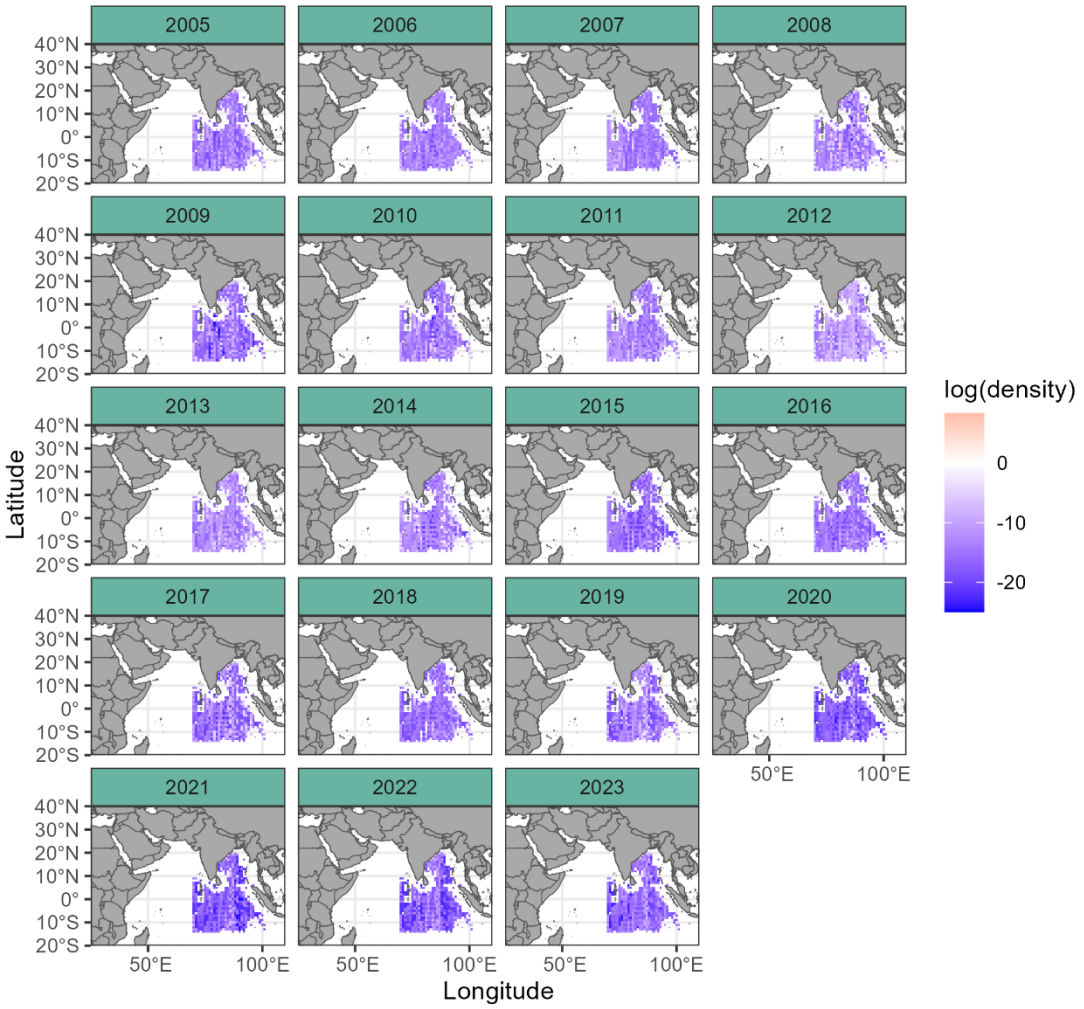


Fig. 13. (Continued).

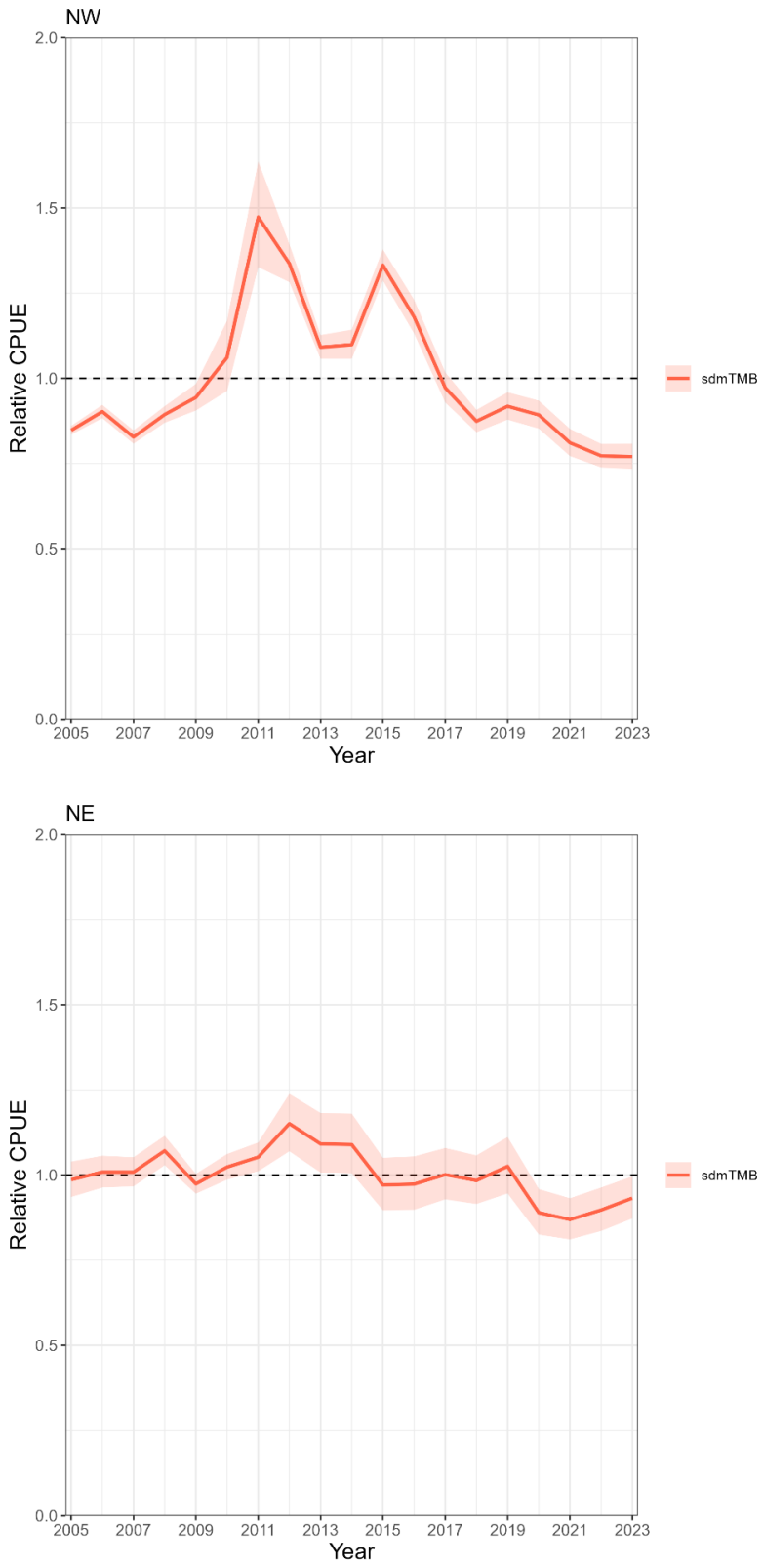


Fig. 14. Standardized CPUE series based on sdmTMB for blue marlin caught by Taiwanese large-scale longline fishery in the Indian Ocean from 2005 to 2023.

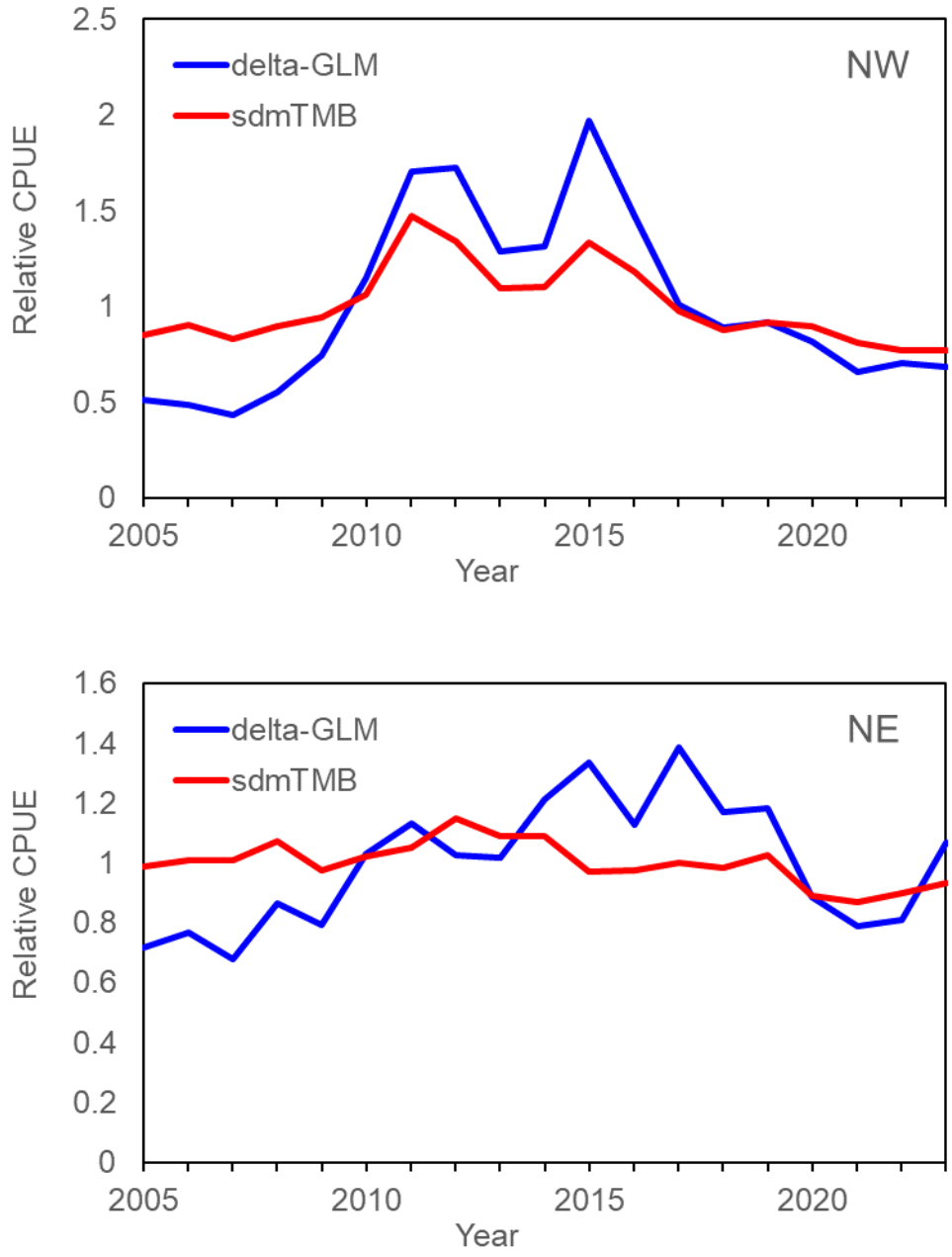


Fig. 15. Standardized CPUE indices for blue marlin in the NW and NE Indian Ocean from 2005 to 2023, estimated using delta-GLM and sdmTMB models.

Table 1. Diagnostic statistics for standardized CPUE series based on various models for positive catches of blue marlin caught by Taiwanese large-scale longline fishery in the Indian Ocean from 2005 to 2021.

Area	Model	R^2	AIC
NW	lognormal	0.192	73,531
NW	gamma	0.271	99,874
NW	inverse.gaussian	0.257	60,876
NE	lognormal	0.186	5,060
NE	gamma	0.252	10,509
NE	inverse.gaussian	0.244	2,758

Table 2. ANOVA table for selected standardized CPUE series based on selected GLMs for blue marlin caught by Taiwanese large-scale longline fishery in the Indian Ocean from 2005 to 2021.

Area NW

Positive catch model:

	SumSq	Df	Fvalues	Pr(>F)	
Year	5,084	18	319.3	< 2.2e-16	***
VesselID	7,722	377	23.2	< 2.2e-16	***
Quarter	0.7	3	0.233	0.852	
Lonlat5	758	41	20.9	< 2.2e-16	***
Cluster	1,570	3	591.7	< 2.2e-16	***
Quarter:Lonlat5	607	109	6.3	< 2.2e-16	***
Residuals	115,710	130,801			

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

Delta model

	LRChisq	Df	Pr(>Chisq)	
Year	15,694	18	<2.2e-16	***
VesselID	12,244	387	<2.2e-16	***
Quarter	0.4	3	0.547	
Lonlat5	1,709	41	<2.2e-16	***
Cluster	1,175	3	<2.2e-16	***
Quarter:Lonlat5	1,574	116	<2.2e-16	***

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

Table 2. (Continued).

Area NE

Positive catch model:

	SumSq	Df	Fvalues	Pr(>F)	
Year	173	18	13.036	< 2.2e-16	***
VesselID	2,406	288	11.304	< 2.2e-16	***
Quarter	0.4	3	0.173	0.915-	
Lonlat5	240	34	9.531	< 2.2e-16	***
Cluster	348	3	156.935	< 2.2e-16	***
Quarter:Lonlat5	181	87	2.820	< 2.2e-16	***
Residuals	21,776	29464	NA		

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

Delta model

	LRChisq	Df	Pr(>Chisq)	
Year	841	18	<2.2e-16	***
VesselID	4,050	313	<2.2e-16	***
Quarter	3	3	0.216	
Lonlat5	629	35	<2.2e-16	***
Cluster	788	3	<2.2e-16	***
Quarter:Lonlat5	544	93	<2.2e-16	***

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