Developing abundance indices for Taiwanese PBF longline fishery using GLMM and VAST, incorporating SST and size data

Tzu-Lun Yuan[1], Shui-Kai Chang[2] and Haikun Xu[3]

[1] Department of Applied Mathematics, Tunghai University, Taichung, Taiwan
[2] National Sun Yat-sen University, Kaohsiung, Taiwan
[3] Inter-American Tropical Tuna Commission, La Jolla, CA, USA

February 2024

Working document submitted to the ISC Pacific Bluefin Tuna Working Group, International Scientific Committee for Tuna and Tuna-Like Species in the North Pacific Ocean (ISC), from 29 February to 9 March 2024, Kaohsiung, Taiwan.

Summary

The total catch of PBF of Taiwanese coastal and offshore fisheries (mainly from longline fishery) had been as high as 3,089 mt in 1999, continuously declined to the lowest record of 214 mt in 2012, and recovered to 1,154–1,497 mt in 2020–2022. The preliminary estimate of the 2023 catch was 2,122 mt, the highest record in recent twenty years.

The average size of PBF was around 212–220 cm before 2008; after that, the average in the North region stably maintained at 218–224 cm during 2008–2022, while in the South region, the average gradually increased to 235 cm in 2012 and declined to about 210 cm during 2020–2022. However, in 2023, the average size of both regions declined to 208–210 cm. A regression tree analysis on the LF data of 2010–2019 suggested that there was no strong spatial and monthly pattern in the mean length of PBF: the substantial increase in average size in the South has resulted from the decline of recruitment, and the decrease since 2013 was a response to smaller fish recruited to the region and more large fish was removed.

Based on the suggestions of the last PBFWG meeting in 2023, five sets of standardizations were conducted in this study using GLMM and VAST, with covariate specifications revised to be similar. The results showed that the patterns of the relative CPUEs from GLMM and VAST differed. However, the trends were similar: all CPUE series suggested a decreasing trend from the beginning of the data series to the lowest level in 2011–2012 and a slow recovery after that until 2000, when a fast recovery occurred.

Introduction

PBF SSB has steadily declined from 1996 to the historically low level in 2010. Affected by this trend, the total catch of Taiwanese offshore fisheries (including longline fishery and some minor coastal fisheries) that harvest mostly large adult fish peaked in 1999 (3,089 mt) and then continuously declined to the lowest level of 214 mt in 2012, less than 10% of the peak catch. With the implementation of management measures to restrict catches of small PBF since 2010, the stock started recovering after that, and the recovery has been more rapid in recent years (ISC 2023).

For monitoring the status of PBF catch from Taiwanese offshore longline fishery, this report provides historical catch and size information as well as relative CPUE series that were standardized by delta-generalized linear mix model (delta-GLMM) from which the result was adopted for stock assessment purposes in previous PBFWG meetings, and by the vector-auto-regressive spatiotemporal model (VAST; Thorson and Barnett, 2017).

Large scale port sampling program to collect length (in cm) data of PBF was implemented since 2010, and over 95% of PBF landed in Taiwan were measured. Because of the high coverage of fish measurements from the catch, length frequency (LF) data from the port sampling program was directly used to represent the catch at size.

For CPUE standardizations, two additional data types were incorporated in the analyses. The first data of spatiotemporal sea surface temperature (SST) was used as a factor in the standardization models to see the relationship of SST variation with the CPUE series. The second data was the size measurement of individual PBF reported by fishermen and verified by the port inspector with spatiotemporal information. The size data was converted to age group data before being incorporated into the standardization models using the von Bertalanffy growth equation (VBGE) of PBF developed by Shiao et al. (2016).

Materials and Methods

Catch and effort data (number of fish and fishing days per trip) of TWN PBF longline fishery for 2003–2023 was reconstructed and compiled using the approach documented in Chang et al. (2017). This study used three sets of data for performing three standardization models: (1) data from 2003–2023 for traditional delta-GLMM analyses, (2) 2007–2023 for spatiotemporal VAST analyses because the spatiotemporal information was available since 2007, and (3) 2010–2023 for spatiotemporal VAST models incorporating size data because the spatiotemporal size data of individual fish was available since 2010. Three model designs were applied for PBF CPUE standardization: traditional delta-GLMM with or without SST effect, VAST with or without SST effect, and VAST with SST effect and size data incorporated. Akaike Information Criterion (AIC) was applied to select the best model.

Traditional delta-GLMM with SST effect

The design of the traditional non-spatial model in this study was identical to the one used in ISC/21/PBFWG-2/02: standardizing the catch and effort data using delta-GLMM, which separately estimates the proportion of positive PBF catches assuming a binomial error distribution (zero-proportion model) and the mean catch rate of positive catches by assuming a lognormal error distribution (positive-catch model). Akaike and Bayesian information criteria were used to determine standardization models' most favorable variable composition. Covariates considered in the model included year (2003–2023), month (May–July), fishing area (the North region and the South region separated by 24.3°N), vessel size (CT1–CT4), and SST. The SST was obtained from the ERDDAP Data Server (*ERDDAP - SST and SST Anomaly, NOAA Global Coral Bleaching Monitoring, 5km, V.3.1, Monthly, 1985-Present, Lon0360 - Data Access Form*), which provides various forecasting products for marine conditions. Three standardization runs were performed: on the area-combined data, on the data from the South region, and on the North region. R package LME4 (version 1.1-31) of R (version 4.2.2) was used for GLMM calculations.

Spatiotemporal VAST with SST effect

The R package VAST (Thorson, 2019)(VAST, version 3.10.0) was applied to the data for spatiotemporal analyses. VAST is a delta-generalized linear mixed model that separately estimates the proportion of positive PBF catches and the mean catch rate of positive catches. This study considered time-invariant spatial variations, time-varying spatiotemporal variations, as well as effects of month and vessel on catchability and SST on density.

We modeled the encounter probability (p) for observation *i* using a logit-linked linear predictor:

$$\text{logit}(p_i) = \beta_1(t_i) + L_{\omega_1}\omega_1(s_i) + L_{\varepsilon_1}\varepsilon_1(s_i,t_i) + L_{\delta_1}\delta_1(v_i) + \sum_{k=1}^{n_k}\varphi_1(k)Q_1(i,k) + \gamma_{1,1}X(s_i,t_i) + \gamma_{1,2}X^2(s_i,t_i)$$

and modeled the positive catch rate (λ) for observation *i* using a log-linked linear predictor:

$$\log(\lambda_i) = \beta_2(t_i) + L_{\omega_2}\omega_2(s_i) + L_2\varepsilon_2(s_i,t_i) + L_{\delta_2}\delta_2(v_i) + \sum_{k=1}^{n_k}\varphi_2(k)Q_2(i,k) + \gamma_{2,1}X(s_i,t_i) + \gamma_{2,2}X^2(s_i,t_i)$$

where $\beta(t_i)$ is the intercept for in year t_i, $\omega(s_i)$ denotes time-invariant spatial variations at location s_i, $\varepsilon(s_i,t_i)$ denotes time-varying spatiotemporal variations at location s_i in year t_i, and $\delta(v_i)$ denotes the effect of vessel v_i on catchability and $\delta_j(v_j) \sim \text{Normal}(0,1), j = 1,2$. L_ω, L_ε and L_δ are the scaling coefficients of the random effect distributions, the $\varphi_j(k)$ is the k^{th} catchability covariates $Q(i,k)$ for observation i (i.e., the impact of month and the vessel factor) and $\gamma_{j,1}$and $\gamma_{j,2}$ represents the impact of the linear and square term of covariate (SST) with value $X(s_i,t_i)$ and $X^2(s_i,t_i)$on density at location s_i in year t_i.

Spatiotemporal VAST incorporating size data

The spatiotemporal model incorporating size data of individual fish can predict density at unsampled locations, times and length classes and provide the relative trends of total abundance various age compositions (Thorson et al., 2017a; Kai et al., 2017). Due to the limitation of computation capability, the size data was first converted to fish ages using the VBGE of Shiao et al. (2016) [$L_t = 251.2 \times (1 - e^{-0.16(t+2.57)})$] and then grouped into seven age groups (age-bins) per 3 ages: 6-8, 9-11, 12-14, 15-17, 18-20, 21-23, and >23 yrs. Each temporal, spatial, and age group has the encounter probability (p) for observation i:

$$\text{logit}(p_i) = \beta_1(t_i) + L_{\omega_1}\omega_1(s_i) + L_{\tau_1}\tau_1(l_i) + L_{\varepsilon_1}\varepsilon_1(s_i, t_i, l_i) + L_{\delta_1}\delta_1(v_i) + \sum_{k=1}^{n_k} \varphi_1(k)Q_1(i,k) + \gamma_{1,1}X(s_i, t_i, l_i) + \gamma_{1,2}X^2(s_i, t_i, l_i)$$

and positive catch rate (λ) for observation i:

$$\log(\lambda_i) = \beta_2(t_i) + L_{\omega_2}\omega_2(s_i) + L_{\tau_2}\tau_2(l_i) + L_{\varepsilon_2}\varepsilon_2(s_i, t_i, l_i) + L_{\delta_2}\delta_2(v_i) + \sum_{k=1}^{n_k} \varphi_2(k)Q_2(i,k) + \gamma_{2,1}X(s_i, t_i, l_i) + \gamma_{2,2}X^2(s_i, t_i, l_i)$$

where $\tau(l_i)$ denotes the impact of age group on the expected catch rate l_i, $\varepsilon(s_i, t_i, l_i)$ denotes interaction item of the spatio-temporal and age variation the age l_i at location s_i in year t_i, and $X(s_i, t_i, l_i)$ (SST) is a habitat covariate that explain variation in density at location s_i in year t_i for the l_i age. The spatiotemporal-at-age variation, $\varepsilon(s, t, l)$, is modeled by combining the Gaussian random field for spatial variation with a first-order autoregressive process (AR1) for temporal and age variation.

Standardized and nominal indices

The observed catch rate (c_i) for each observation is C_i/E_i. The probability function for c_i is

$$\Pr(c_i = c) = \begin{cases} 1 - p_i & \text{if } c = 0 \\ p_i \times \text{Lognormal}(c_i | \log(\lambda_i), \sigma^2) & \text{if } c > 0 \end{cases}$$

where σ^2 is a dispersion parameter.

Calculation of the index of abundance of PBF (in year t) for delta-GLMM models:

$$I_{non}(t) = \text{logit}^{-1}(p_i) \times \exp(\lambda_i).$$

For the spatiotemporal model with the SST effect, the index is predicted using an area-weighted approach:

$$I_{st}(t) = \sum_{s=1}^{n_k} d(s,t)$$

where n_k denotes the number of locations and $d(s,t)$ is the predicted density for location s and year t:

$$d(s,t) = \text{logit}^{-1}\left(\begin{array}{c}\beta_1(t_i) + L_{\omega_1}\omega_1(s_i) + L_{\varepsilon_1}\varepsilon_1(s_i,t_i) + L_{\delta_1}\delta_1(v_i) \\ + \sum_{k=1}^{n_k} \varphi_1(k)Q_1(i,k) + \gamma_{1,1}X(s_i,t_i) + \gamma_{1,2}X^2(s_i,t_i)\end{array}\right) \times \exp\left(\begin{array}{c}\beta_2(t_i) + L_{\omega_2}\omega_2(s_i) + L_2\varepsilon_2(s_i,t_i) + L_{\delta_2}\delta_2(v_i) \\ + \sum_{k=1}^{n_k} \varphi_2(k)Q_2(i,k) + \gamma_{2,1}X(s_i,t_i) + \gamma_{2,2}X^2(s_i,t_i)\end{array}\right)$$

For the spatiotemporal model with additional size data, the index is the sum of the abundance of each temporal, spatial, and age group:

$$I_{sta}(t) = \sum_{k=1}^{n_k}\sum_{l=1}^{n_l} d(k,t,l)$$

where n_k denotes the number of locations, n_l denotes the number of ages, and $d(s,t,l)$ is the predicted density for the age l at location s in year t:

$$d(s,t,l) = \text{logit}^{-1}\left(\begin{array}{c}\beta_1(t_i) + L_{\omega_1}\omega_1(s_i) + L_{\tau_1}\tau_1(l_i) + L_{\varepsilon_1}\varepsilon_1(s_i,t_i,l_i) + L_{\delta_1}\delta_1(v_i) \\ + \sum_{k=1}^{n_k} \varphi_1(k)Q_1(i,k) + \gamma_{1,1}X(s_i,t_i,l_i) + \gamma_{1,2}X^2(s_i,t_i,l_i)\end{array}\right) \times \exp\left(\begin{array}{c}\beta_2(t_i) + L_{\omega_2}\omega_2(s_i) + L_{\tau_2}\tau_2(l_i) + L_{\varepsilon_2}\varepsilon_2(s_i,t_i,l_i) + L_{\delta_2}\delta_2(v_i) \\ + \sum_{k=1}^{n_k} \varphi_2(k)Q_2(i,k) + \gamma_{2,1}X(s_i,t_i,l_i) + \gamma_{2,2}X^2(s_i,t_i,l_i)\end{array}\right)$$

Then, the average spatial distribution of the predicted catch rate for each year is calculated as:

$$I_{sta}(t,l) = \sum_{s=1}^{n_k} d(s,t,l).$$

Essentially, the area-weighted approach computes total abundance as the weighted sum of estimated density across the pre-defined spatial domain of knots, with weights equal to the area associated with each knot. For computational purposes, k-means algorithm was used to cluster all the grid cells into 50 spatial knots and assumed that both the spatial and spatiotemporal random effects for a grid cell are from the closest knot in space.

Results and Discussions

Catch trend and size pattern

PBF catch of Taiwan had been as high as 3,089 mt in 1999. However, it continuously declined to the lowest record of 214 mt (210 mt from longline fishery) in 2012 (Fig. 1). After a gradual increase to the level of 400–550 mt during 2014–2019, PBF catch recovered to 1,154 mt in 2020, and to 1,475 mt and 1,497 mt in 2021 and 2022, similar to the level of 2007 when the stock was relatively healthy (Fig. 1). Preliminary estimate of the 2023 catch reached 2,122 mt, close to the quota of 2023 (2,299 mt, including 1,965 mt of annual quota and 17% carryover quota from 2022).

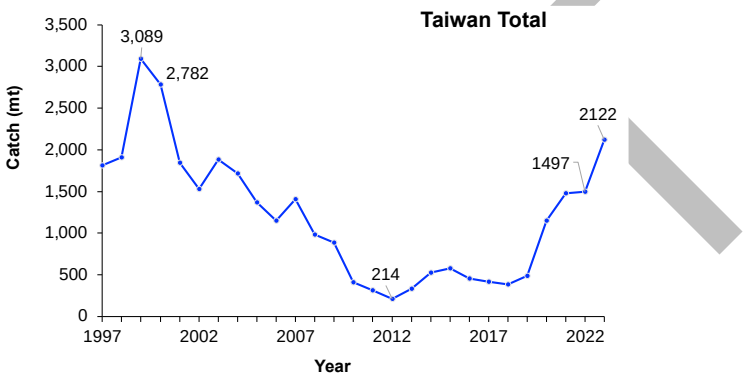

Fig. 1. Annual PBF catches by Taiwanese offshore fishery (mainly from longline fishery with minor catches from other small coastal fisheries).

The average size of PBF caught by the Taiwanese longline fishery was around 212–220 cm before 2008 (Fig. 2). after that, the average in the North region stably maintained at 218–224 cm during 2008–2022, while in the South region, the average size gradually increased to 235 cm in 2012 and declined to 210 cm during 2020–2022, showing a different trend from the North region. However, in 2023, the average size declined to 208–210 cm for both regions. The substantial increase in average size in the South region was considered to result from the decline of recruitment to the fishing ground; and the decrease since 2013 was a response to more small-fish recruited to the fishing ground and more large-fish removed from the fishing ground (Fig. 3).

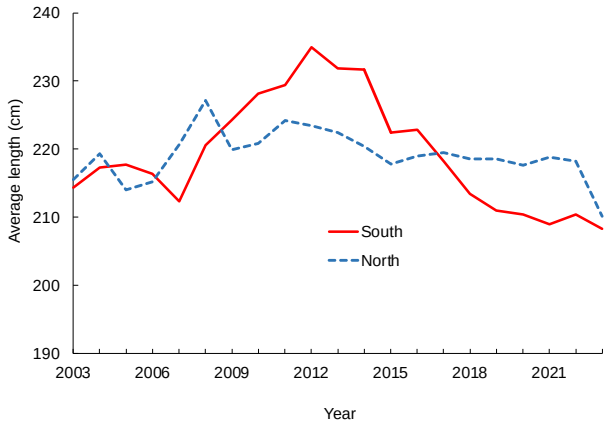

Fig. 2. Annual trend of average length of PBF of Taiwanese longline fishery by region.

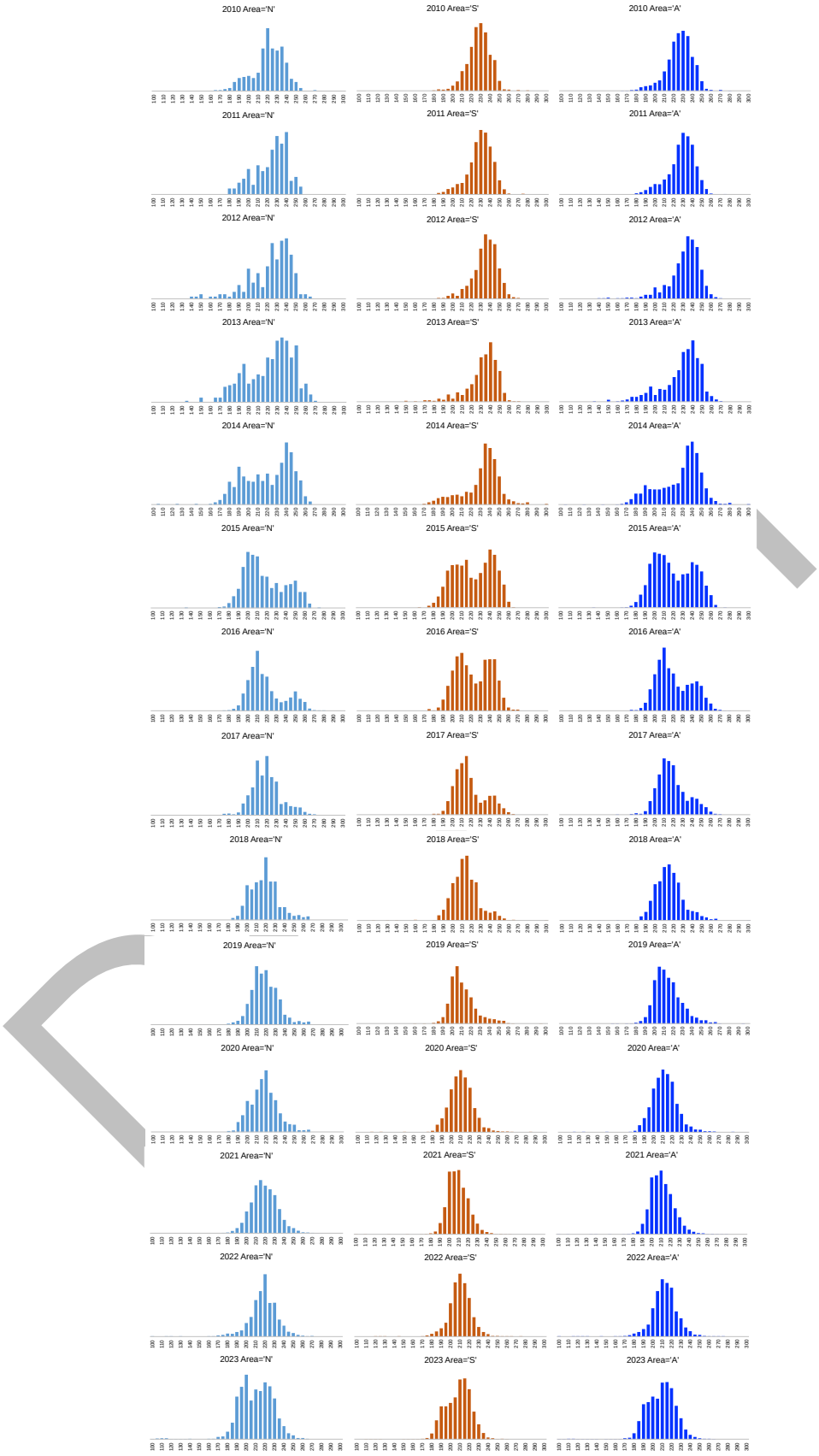

Fig. 3. Length frequencies of Taiwanese PBF during 2010 – 2023 for North region (left in sky blue), South region (middle in red), and All area (left in blue).

A regression tree analysis on the LF data of 2010–2019 suggested that there was no strong spatial and monthly pattern in the mean length of PBF around Taiwan and that the variation of LF was mainly attributed to recruitment fluctuation (Chang et al., 2024). The analysis also suggested a weak latitudinal split (24.5ºN, the definition for separating this study's north and south fishing grounds) and no necessary monthly or longitudinal split for the fishery. With these suggestions, this study performed standardizations on the whole area (the Whole region), with some also on the South region to be consistent with previous years' analyses.

Relative CPUE from delta-GLMM and VAST

Yuan et al. (2023) performed 18 standardizations using delta-GLMM and VAST with different combinations of covariates. Based on discussions on the results and suggestions on the results in PBFWG of November 2023, this study re-ran four sets of selected standardizations that have similar specifications of covariates (Table 1). Year, month (fixed effect), vessel size (4 tonnages, fixed effect), and vessel effect (vessel ID, random effect) were included in both delta-GLMM and VAST models. The GLMM also included the year and month interaction term (as a random effect). Previous tests showed that including the interaction term could improve model performance (with smaller AIC), and the trend was slightly different from that did not include the interaction term. For the VAST, however, the model could not include year and month interaction as a random effect, and the run that included interaction term as the fixed effect could not converge. Therefore, this interaction term was not included.

Table 1. Covariates and the treatments of GLMM and VAST standardizations investigated in this study. The new treatments of covariates different from the previous study (Yuan et al., 2023) were marked in red.

	GLMM	VAST
Year	Yes	Yes
Month	Fixed	**Fixed**
Year*Month	Random	NA
Vessel - size	Fixed (4 tonnages)	**Fixed (4 tonnages)**
Vessel - individual	**Random (vessel ID)**	Random (vessel ID)
Spatial effect	No	Yes
Spatiotemporal	No	Yes
SST	Yes/No	Yes

Five additional sets of relative CPUE series are provided in this study (Table 2): (1) delta-GLMM series (not including SST data) for the South and the Whole regions since 2003; (2) delta-GLMM-sst series (including SST data) for the South and the Whole regions since 2007; (3) VAST-sst series (including SST data) for the South and the Whole regions since 2007; (4) VAST-sst/size series (including SST and size converted age-group data) for the Whole region since 2010; and, (5) VAST-sst/size series by age-group for the Whole region since 2010.

Table 2. Additional five sets of standardization suggested by the 2023 November PBFWG meeting.

Sets	Statistical model	Region	Starting year
1	GLMM	South/Whole	2003
2	GLMM-sst	South/Whole	2007
3	VAST-sst	South/Whole	2007
4	VAST-sst/size	Whole	2010
5	VAST-sst/size by age group	Whole	2010

Distribution of fishing effort in the core area (fishing days, Fig. 4) suggested the fishing ground could be split into two fishing grounds by 24.5ºN, with the South region being the major one. High fishing effort was observed from 2020 onwards in the SW waters off Taiwan. Those efforts targeted PBF and non-tuna species, with high effort but low catch rate.

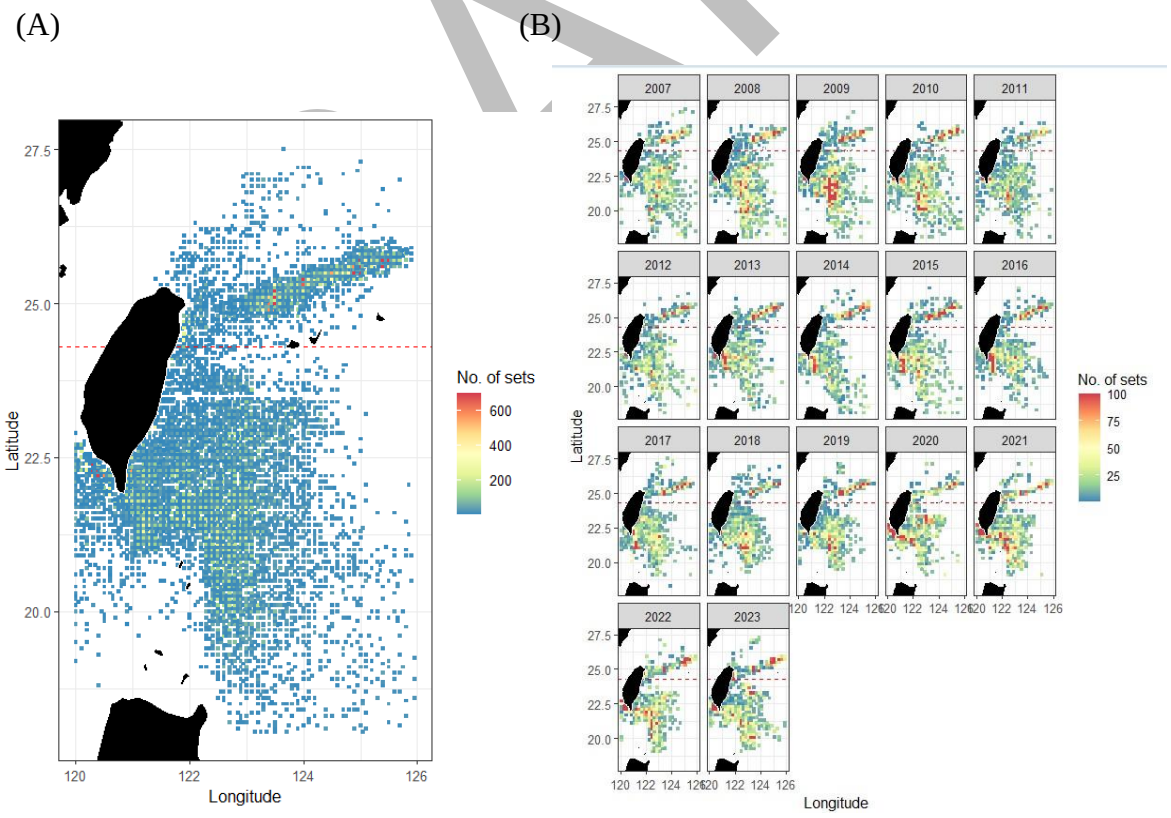

Fig. 4. Geographic distribution of fishing days during 2007–2023: (A) whole period (original position); (B) by year (aggregated to 0.2º square area).

The resulting CPUE series of the 1st set of standardization is shown in Fig. 5. Yuan et al. (2023) demonstrated that, although including SST in the standardization models produced smaller AICs for both GLMM and VAST approaches, the resulting CPUE series were almost identical for each approach. Due to the lack of spatial information in the data of 2003–2006, if SST is not necessary to be considered, the GLMM standardization can have a longer series of data starting from 2003. Fig. 5 shows the resulting new series (GLMM.2024) for the South and the Whole regions, which were very similar to the previous ones (GLMM.2023), except for 2023 in the South region, indicating that inclusion of the random effect of vessel ID generally did not have much effect on the standardization. The reason for the large deviation in 2023 in the South region is still unclear.

Similar comparisons for VAST standardizations are also shown in Fig. 5 (right panels). Obvious deviations between the new (VAST-sst.24) and old (VAST-sst.23) series were noted for the years since 2020. Including month (fixed effect) and vessel size (fixed effect) in the VAST model runs (Table 1) apparently have effects on the standardizations, although the trends are basically similar.

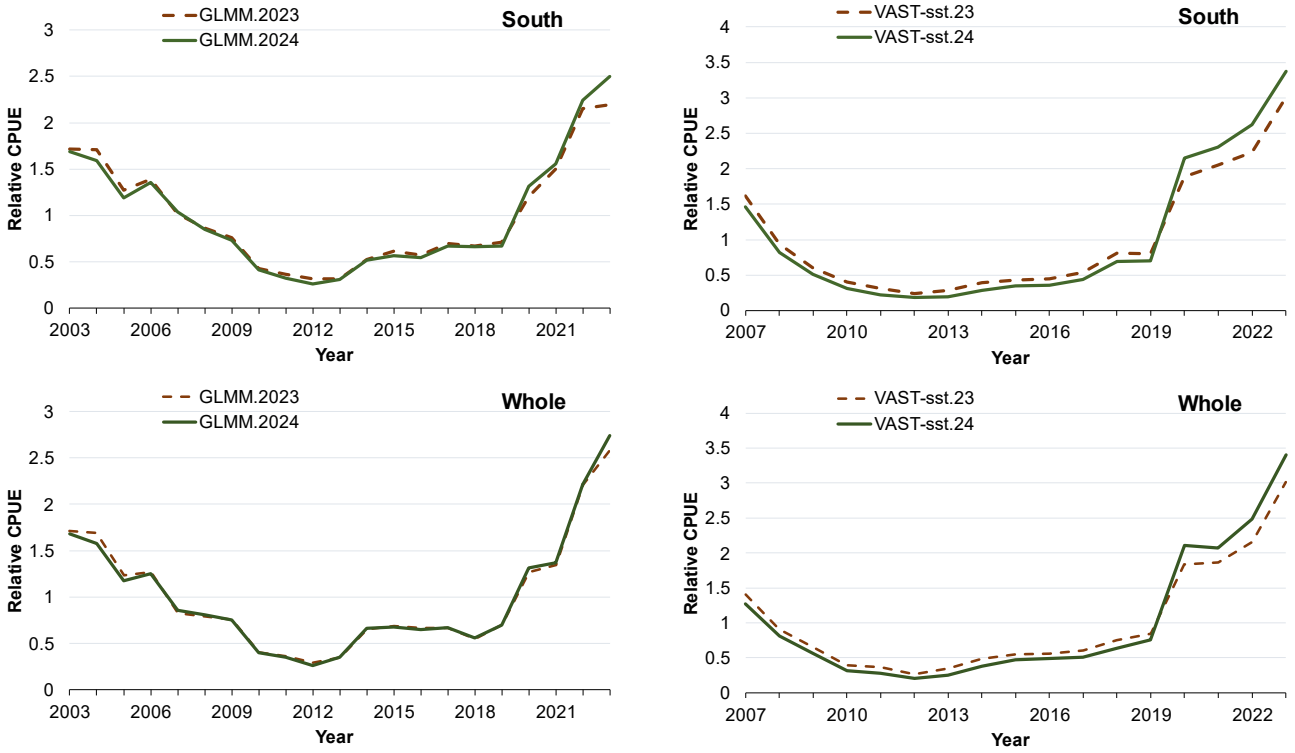

Fig. 5. Comparisons of relative CPUE series from this study (suffixed by "24") and the previous study (Yuan et al., 2023; suffixed by "23") for GLMM (left panels, starting from 2003) and VAST (right panels, starting from 2007) standardizations.

The CPUE series of the 2nd to 4th sets of standardizations (Table 2) with model specifications of Table 1 are plotted in Fig. 6. The series incorporating size data was started in 2010 because the size data with spatiotemporal information was available only in 2010. The figure suggested that incorporating size data in the VAST standardizations did not show a substantial impact; the two CPUE series with (VAST-sst/size.24) and without (VAST-sst.24) incorporation of size data were similar in trends for both the South and the Whole regions. Nevertheless, the two series were different from the results of GLMM standardizations. This difference might be resulted from the spatiotemporal effect that has been addressed in the VAST standardizations, but not in the GLMM.

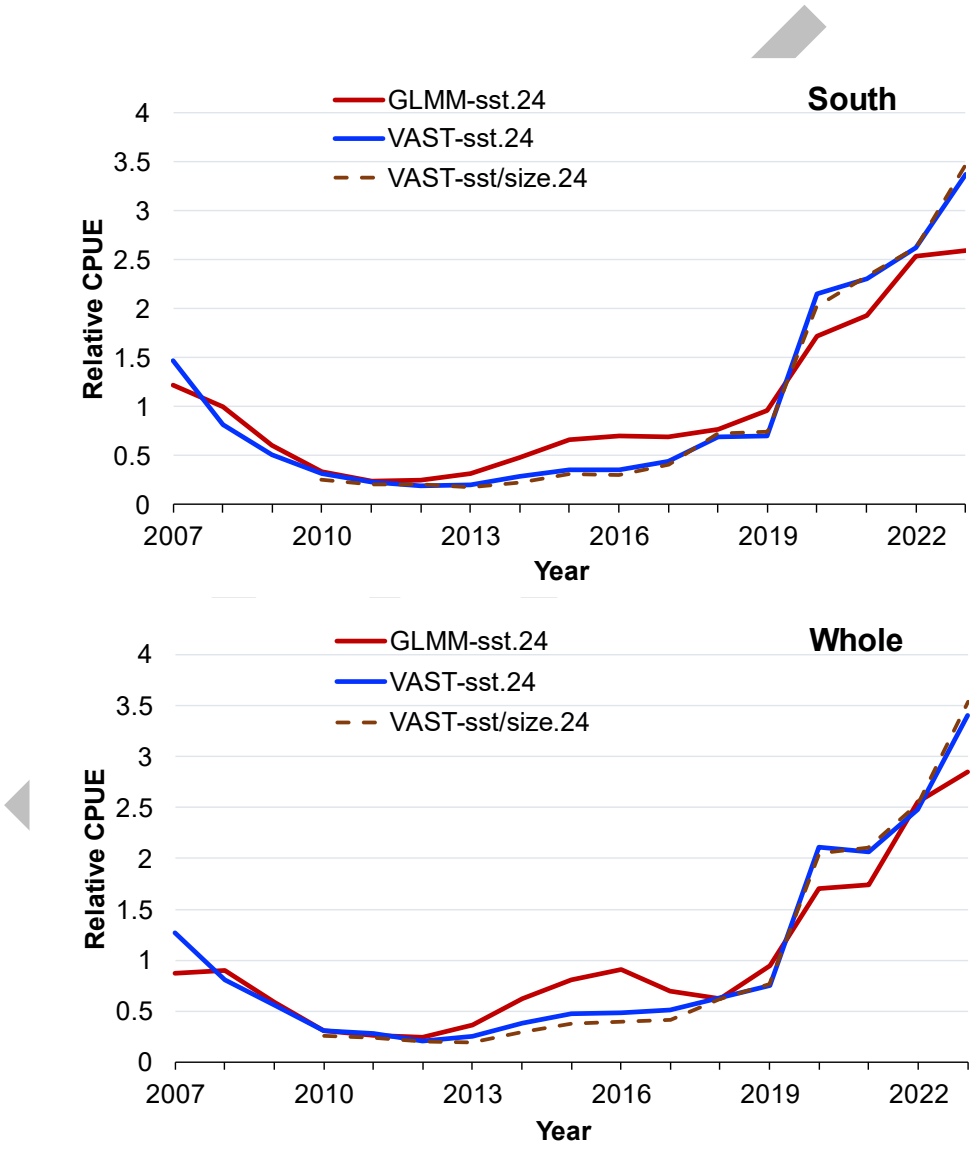

Fig. 6. Relative CPUE series resulted from GLMM and VAST with and without incorporation of size data (age groups) (the 2nd to 4th sets standardizations of Table 2).

All the CPUE series suggested a decreasing trend from the beginning of the data series to the lowest level in 2011–2012 and a recovery after that to the recent year.

Pronounced spatiotemporal variations in density were predicted from 2007 through 2023 (Fig. 10): the densities decreased from the starting year of the study (2007) to the lowest level in 2011 and 2012 and then started to increase gradually toward the end year of the study, while the increase during 2020–2023 apparently was substantial. Compared with 2007 (the beginning year of the study), the density pattern was similar in 2020 and higher in the later years, especially for 2022 and 2023, suggesting a recovery of PBF stock from the low level of 2011–2012.

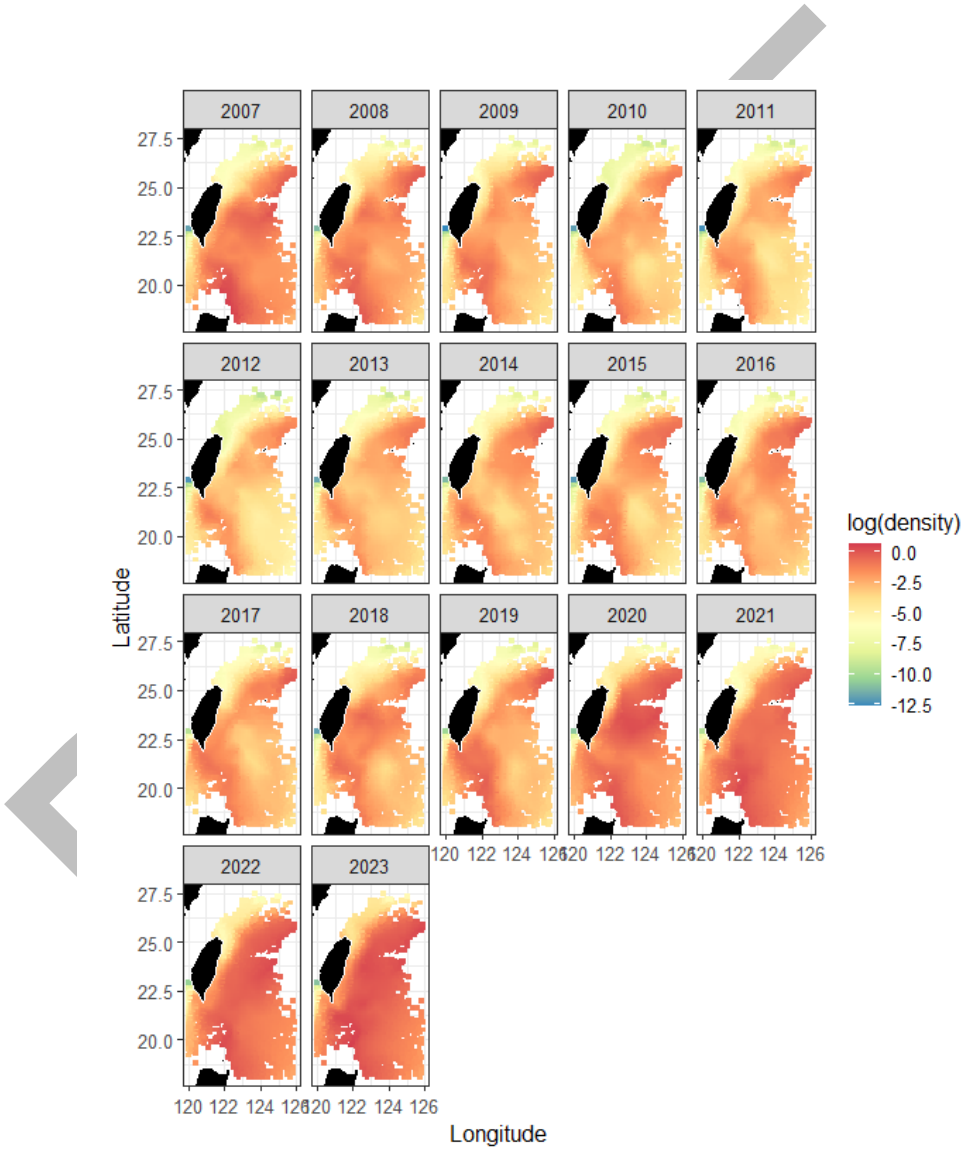

Fig. 7. Spatiotemporal distribution of predicted log density of PBF during 2007–2023 from VAST analyses (upper left – South, upper right – North, bottom: South and North combined).

Relative CPUE from VAST incorporating size data

Relative CPUE and density of PBF of the Whole region for 2010–2023 by the seven age groups were shown in Fig. 6. Age group 9-11 was the most dominant fish, and then the 6-8 age group, for recent years in the fishing ground. During the low stock status in 2010-2012, the density of age groups of 6-8 and 9-11 was very low for both regions, but it started to increase slowly around 2015 and then jumped up since 2020. The year that the density of the age group 12-14 started to increase was about three years later than the younger fish groups, and the density also jumped up in 2020. Contrarily to this, age groups 18-20 and 21-23 have declined since 2015 and have been at a low level since 2019. For the age group > 23 years old, a substantial decrease was observed in 2017.

Age groups older than 15 generally showed obvious declining trends since 2014 or 2015. For age groups 18-20, 21-23, and >23 (large, matured fish), the density showed a concave shape with a peak at around 2014 to 2016 and then showed a continuous decline. The reasons for this observation are unknown.

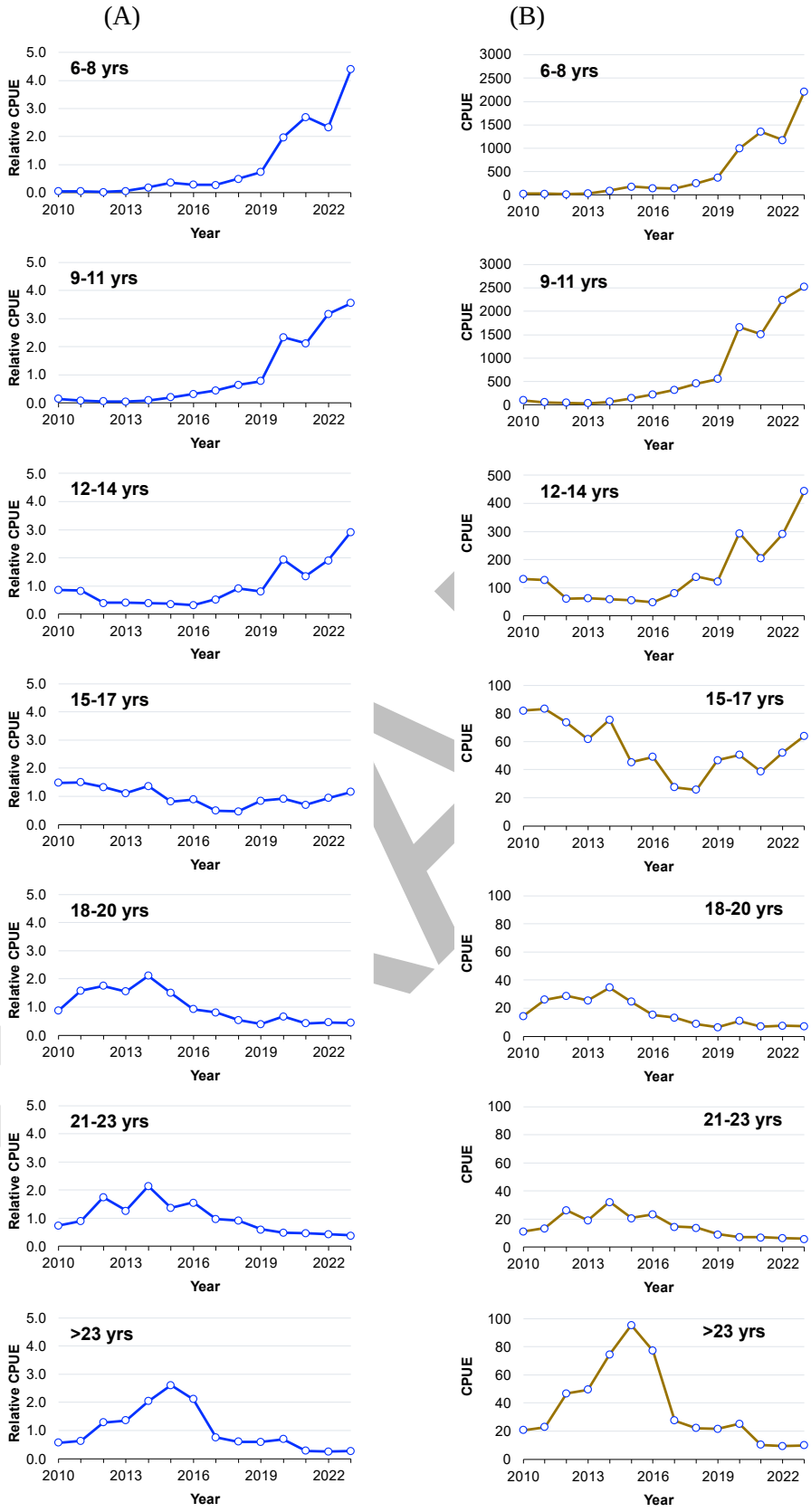

Fig. 6. Relative CPUE (panel A) and density (panel B, scales are different by age group) of PBF for 2010 – 2023, by age group.

References

Shiao, J.-C., Lu, H.-B., Hsu, J., Wang, H.-Y., Chang, S.-K., Huang, M.-Y., Ishihara, T., 2016. Changes in size, age, and sex ratio composition of Pacific bluefin tuna (Thunnus orientalis) on the northwestern Pacific Ocean spawning grounds. ICES J. Mar. Sci. 74, 204–214. https://doi.org/10.1093/icesjms/fsw142

Thorson, J.T., 2019. Guidance for decisions using the Vector Autoregressive Spatio-Temporal (VAST) package in stock, ecosystem, habitat, and climate assessments. Fish. Res. 210, 143–161. https://doi.org/10.1016/j.fishres.2018.10.013

Yuan, T.-L., S.-K. Chang, H. Xu. 2023. PBF abundance indices from Taiwanese offshore longline fisheries using delta-GLMM and VAST, incorporating SST and size data. Pacific Bluefin Tuna Working Group Intersessional Workshop of the ISC, 27 November to 1 December 2023, Webniar. ISC/23/PBFWG-2/04.