



Size pattern and relative CPUE of Taiwanese PBF fisheries using delta-generalized linear mixed models (GLMM) and vector-auto-regressive spatiotemporal model (VAST)

Tzu-Lun Yuan¹, Shui-Kai Chang¹, Hung-I Liu², and Chao-Chin Huang¹

¹ National Sun Yat-sen University, Kaohsiung, Taiwan

² Center for Artificial Intelligence & Data Science, Ming Chi University of Technology, New Taipei City, Taiwan

April 2021

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 20 to 27 April 2021, virtual meeting through Teams.

Document should not be cited without author's permission.

Summary

Size of Taiwanese offshore PBF was measured through official port sampling program, with data coverage higher than 95% since 2010. Average fish size in the northern fishing ground was stable at 217-224 cm since 2010, while the size in the southern fishing ground was showing an increasing trend from 212 cm in 2007 to 235 cm in 2012 when the PBF catch was at the historical lowest level, and then a continuous declining trend to 210 cm in 2020 when the catch was recovered to highest record since 2008. The 2020 catch was estimated to be 1149 mt, doubled the average of 2008–2019.

Taking advantage of voyage data recorder (VDR) data, trip data and CDS data, historical offshore PBF catch and effort data was reconstructed. Applying traditional delta-generalized linear mix model (delta-GLMM) (without consideration of spatial effect) and vector-auto-regressive spatiotemporal model (VAST) (considered spatial effect in the model) to the reconstructed data, standardized relative CPUE was estimated for three regions: southern fishing ground, northern fishing and both fishing grounds combined, respectively. The results of the GLMM and the VAST models both exhibit a substantial increase in 2020 for all the three study regions. For the southern region where its series was considered more representative in previous PBFWG meetings, while the trends from VAST show an obvious sharp decline in the beginning of the study period and a sharp increase in the last year (2020), both GLMM and VAST models suggested that the level of 2020 has recovered to the level of 2006–2007 level.

Introduction

PBF is an important seasonal target species for the Taiwanese offshore longline fishery. The catch peaked in 1999 (3,089 mt) and then continuously declined to the lowest level of 213 mt in 2012, less than 10% of the peak catch, and recovered thereafter. Since the catch was relatively small in Taiwan's tuna fisheries, logbook submission was not required until 2010 when specific management regulations (catch documentation scheme, CDS) on PBF fishery were implemented (Chang et al., 2015, 2017). With implementation of the new regulations, detail catch and size information on PBF are available since 2010. In addition, detail GPS-locations are available from voyage data recorder (VDR) system since 2007 and vessel monitoring system (VMS) since 2018 (FA, 2019). With these available data and additional daily market landing data by vessel and landing port (since 2001), the historical catch and effort was reconstructed for estimation of abundance (Chang et al., 2017).

Small scale port sampling program to collect length data of PBF was conducted since 2000. The scale gradually increased and since 2010, over 95% of PBF landed were measured. This study provides historical catch and size information of PBF from Taiwanese offshore longline fishery as well as relative CPUE series standardized by delta-generalized linear mix model (delta-GLMM) (without consideration of spatial effect) which result was adopted for stock assessment purpose in previous years (Chang and Liu, 2016, 2017; Liu and Chang, 2019; Yuan et al., 2019; Chang et al., 2020) and vector-auto-regressive spatiotemporal model (VAST; Thorson and Barnett, 2017) (considered spatial effect in the model) (Chang et al., 2020).

Materials and Methods

Data

The size data of PBF for 2003–2009 was collected by port sampler through port sampling program. The data since 2010 was collected by inspectors of CDS program which come with detail fishing location for separation of fishing grounds (northern and southern fishing grounds).

The catch and effort data (number of fish and fishing days per trip) used in this study is the same as that used in ISC/20/PBFWG-2/11 (Chang et al., 2020) with newly available 2020 data. The whole series data is from 2001 to 2020, but the first two years data (2001 and 2002) have the lowest data representativeness for the whole period in terms of the proportion of the catches in the catch/effort data for analyses to the annual total catch (20~40%). Also, the quality of the trip information to construct the effort data (fishing days) for the two years was concerned. Therefore, based on the discussion on ISC/20/PBFWG-2/11 (Chang et al., 2020) which tested standardizations on both the data of 2001–2019 and of 2003–2019, the study used the data since 2003 only for GLMM analyses.

Detailed spatial information for the spatiotemporal standardization was available only since 2007; data before that year contains only the information that assigns the trip to the northern or southern fishing ground (as defined in Fig. 1). Therefore, only 2007–2020 data is used for spatiotemporal VAST analyses. The fish was caught in a wide range of the northwestern Pacific Ocean, however, most PBF was caught in the waters off eastern Taiwan. To avoid the effect from the sparse data away from eastern Taiwan, the data used for spatiotemporal study was narrowed down to the core-area of eastern Taiwan: 120–126°E, 18–28°N. The core-area data has removed only 8.05% of PBF catch from the original data for the whole period of 2007–2020.

Non-spatial model

The design of traditional non-spatial model is identical to the one used in ISC/19/PBFWG-1/02: 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). The standardized index is the product of these two estimated components. Akaike and Bayesian information criteria was used to determine the most favorable variable composition of standardization models.

Covariates considered in the GLMM included: year (2003–2020), month (May–July), fishing area (northern and southern fishing ground separated by 24.3°N), and vessel size (CT1–CT4). Since the number of explanatory variables considered in the study was small (due to limitation in available information), simpler backward (decreasing variables) and forward (increasing variables) methods were applied when determining the variables to be included in the model ($\alpha=0.05$). All the explanatory variables were included initially in the model and were determined in the final models through backward method. First order interactions of the explanatory variables were

also considered for the model and were determined through forward method. The interaction of year and month was treated as random variable, while the others were treated as fix variables. Three standardization runs were performed: (1) that on the area-combined data (fishing ground effect was treated as a covariate in the model); (2) that on the data from the southern fishing ground; and, (3) that on the data from the northern fishing ground.

The GLMM runs in the previous years were performed in SAS. This year, the program runs were performed in R package considering the package is free and powerful and no license is available for SAS. To see if there is substantial difference in the results, this study performed a simple comparison on the results from the two packages.

Spatiotemporal model

The R package VAST (Thorson and Barnett, 2017; Xu et al., 2019) was applied to the abovementioned data for PBF. 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. In this study, we model the encounter probability (p) for observation i using a logit-linked linear predictor

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

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

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

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 $\delta_i(v_i) \sim \text{Normal}(0,1)$, $i = 1,2$. L_{ω} , L_{ε} and L_{δ} are the scaling coefficients of the random effect distributions.

Both the spatial and spatiotemporal random effects are assumed to be correlated in space. We assume that the spatial random effect is $\omega_i \sim \text{MVN}(0, R_i)$, $i = 1,2$ and the spatiotemporal random effect in year t is $\varepsilon_i(s, t) \sim \text{MVN}(0, R_i)$, $i = 1,2$, where R_1 and R_2 are the correlation matrices approximating the similarity of encounter probability and positive catch rate among observation locations. The correlation between both the spatial and the spatiotemporal residuals at two locations (s and s') is assumed to decline over distance at a rate specified by the Matérn function: $R_i(s, s') = \kappa_i |H(s - s')|$, $i = 1,2$, where κ is the decorrelation scaling parameter, which controls the rate of decrease in spatial correlation with increasing distance, and H is a 2 by 2 transformation matrix describing geometric anisotropy (correlation decrease with increasing distance faster in some directions than in the others). Thus, $\kappa_i |H(s - s')|$ is the standardized distance between location s and s' after accounting for geometric anisotropy (Thorson et al., 2015).

The observed catch rate (c_i) for each observation is C_i/E_i , Where C and E represent observed catch and effort, respectively. 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.

The index of abundance of Pacific bluefin tuna (in year t) is then predicted using an area-weighted approach:

$$I_{std}(t) = \sum_{k=1}^{n_k} (a(k) \times d(k, t))$$

where n_k denotes the number of knots, $a(k)$ is the area associated with knot k , and $d(k, t)$ is the predicted density for knot k and year t :

$$d(k, t) = \text{logit}^{-1}(\beta_1(t_i) + L_{\omega_1} \omega_1(s_i) + L_{\varepsilon_1} \varepsilon_1(s_i, t_i)) \\ \times \exp(\beta_2(t_i) + L_{\omega_2} \omega_2(s_i) + L_{\varepsilon_2} \varepsilon_2(s_i, t_i))$$

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.

The fishing activities analyzed in this study took place in more than 5000 unique $0.1^\circ \times 0.1^\circ$ grid cells. For computational purposes, we use the k-means algorithm to cluster all the grid cells into 50 spatial knots and assume 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

The longline catch had been as high as 3,089 mt in 1999 but continuously declined to the lowest record of 210 mt in 2012 (Fig. 2). Thereafter, the catch slowly bounced back and stayed at the level of 400–550 mt during 2014–2019, with the peak of 552 mt in 2015. After the stable level, the catch jumped up to 1149 mt in 2020. The number of fishing vessels has slightly increased from 2019 level (7% increase), however, the catch in 2020 is doubled than that in 2019 (492 mt) and was the highest record since 2008. Both the catches in the northern and southern fishing grounds have increased from 2019, while the increase is more substantially in the southern ground (500 mt more) than in the northern ground (10 mt more). Similar to 2018 situation, over 80% of the catch in 2020 was made in the southern fishing ground. Comparing the nominal CPUE of the vessels that have fished for PBF in 2019 and 2020, 65% of them shows higher CPUE in 2020 than in 2019 (Fig. 3). With these observations, we assumed that the conspicuous increase of catch was highly likely a result of abundance increases rather than a change of fishing pattern.

Average size of PBF caught by the Taiwanese longline fishery was around 212–220 cm before 2008 (Fig. 4). Thereafter, the average in the northern fishing ground stably stayed at 218–227 cm during 2008–2020; while in the southern fishing ground the average gradually increased since 2008, to 234 cm in 2012, and declined all the way to

211 and 210 cm in 2019 and 2020, respectively, showing a different trend from the northern ground. The substantial increase of average size in the southern ground was considered resulting from the decline of recruitment to the fishing ground; and the decrease since 2013 was a response to more smaller fish recruited to the fishing ground and more large fish removed from the fishing ground (Fig. 5).

Non-spatial model runs

GLMM run fitted the data well (based on the qq-plots and the residual histograms, Fig. 6) for the sub-models on southern, northern and whole fishing grounds. From AIC and BIC analyses, standardization separately by fishing grounds has better performance than the one combined both fishing grounds (Table 1). Thus, area-separate models were considered more preferable because the size composition of the two fishing grounds apparently different. The index of the southern fishing ground was considered relatively better representing the PBF abundance index than the northern one considering its features of better data stability and with much higher proportion of historical catches; and was recommended to be used for the PBF stock assessment by the PBFWG since 2016.

The resulted relative CPUE series from GLMM under R environment were provided in Fig. 7, together with the relative series under SAS environment. The results showed some small differences between the two different environments for southern and northern fishing grounds, even with the same model structures (same set of covariates as in Table 1), while the difference was more obvious for the whole fishing ground (combined both southern and northern grounds) in the beginning period. The reason causing such differences is still unknown.

Spatiotemporal model

This model is to address the research priority identified in the 2018 PBFWG (ISC, 2018, Attachment 4) that to “improve Taiwanese index with focus on spatiotemporal change”. Two major fishing grounds are noted from the distribution of fishing effort in the core area (fishing days, Fig. 8): southeast Taiwan and northeast Taiwan. All the spatiotemporal models have successfully converged, which were confirmed by the fact that the Hessian matrix was positive definite, and the maximum gradient component was smaller than 0.001. Moreover, quantile diagnostics suggested the spatiotemporal model fitted the catch and effort data well (Fig. 9).

Pronounced spatiotemporal variations in density were predicted for the period of 2007–2020 (Fig. 10). Predicted 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 in 2020 was apparently substantial. The pattern in southern ground in 2020 was similar to that in 2007, with an additional high density in eastern Taiwan.

Both encounter probability and positive catch rate are more coherent along directions of slightly southeast-northwest for southern ground, southwest-northeast for northern ground, and south-north for both fishing grounds combined (Fig. 11). The

center of biomass of PBF in the east-west (left) and north-south (right) directions are computed for the model runs (Fig. 12). For southern ground, there was no clear pattern in east-west movement but has a trend of moving northward in recent years. No long-term trend was noted for the northern fishing ground. For whole core area, the center of biomass has been moving eastward, and no clear pattern in south-north movement.

The standardized indices computed based on the spatiotemporal distribution of predicted density showed a clear trend of sharp decline since 2007, reaching the lowest level in 2012, and then slowly and continuously recovered to the current year and sharply increased in 2020 for the southern fishing ground (Fig. 13 left panels). The sharp increase in 2020 was also noted in the northern fishing ground. The effective area occupied computed from the models (Fig. 13 right panels) showed roughly similar trends with the corresponding standardized indices, but the patterns were less clear and the associated uncertainty is much larger.

General trend of the indices

The standardized indices using GLMM and VAST were plotted together in Fig. 14 for comparison. Indices from spatial VAST standardizations generally showed similar trends with those from non-spatial GLMM runs, but that the spatial model results exhibited substantial fluctuations of relative CPUEs in both the beginning and the ending years. However, focusing on the indices for the southern fishing grounds that was adopted for the stock assessment purposes in the previous years, the level of 2020 estimated was the same level of 2006 for GLMM and was 10% higher than the level of 2007 (the beginning year of data period) for VAST, implying that both models suggested similarly the index of 2020 has recovered to the 2006-2007 level. This sharp increase of 2020 index was coincident with the substantial increase of 2019 index of Japanese coastal and offshore longline fishery (Tsukahara et al., 2020) rendering further exploration worthwhile.

References

- Chang, S.-K. and H.-I Liu. 2016. Update of Standardized PBF CPUE Series for Taiwanese Longline Fishery. Pacific Bluefin Tuna Working Group Intersessional Workshop of the ISC, La Jolla, USA, 29 February – 11 March 2016. ISC/16/PBFWG-1/02 (revised).
- Chang, S.-K. and H.-I Liu. 2017. Standardized PBF CPUE series for Taiwanese longline fishery up to 2016. Pacific Bluefin Tuna Working Group Intersessional Workshop of the ISC, Shimizu, Japan, 15 – 20 February 2017. ISC/17/PBFWG-1/02.
- Chang, S.-K., H.-I Liu, H. Fukuda and M.N. Maunder. 2017. Data reconstruction can improve abundance index estimation: An example using Taiwanese longline data for Pacific bluefin tuna. PLoS ONE 12, e0185784.
- Chang, S.-K., H.-I Liu. and Y.-W. Fu. 2015. Estimation of standardized CPUE series on Pacific bluefin tuna for Taiwanese longline fishery under incomplete data. Pacific Bluefin Tuna Working Group Intersessional Workshop of the ISC, Kaohsiung, Taiwan, November 18-25, 2015. ISC/15/PBFWG-2/10.
- Chang, S.-K., T.-L. Yuan, H.-I Liu and H. Xu. 2020. Abundance index of Taiwanese PBF fisheries based on traditional and spatiotemporal delta-generalized linear mixed models.

- Pacific Bluefin Tuna Working Group Intersessional Workshop of the ISC, 2 to 12 March 2020, Webinar meeting. ISC/20/PBFWG-1/03.
- FA. 2019. National Report on Chinese Taipei (Taiwanese Tuna and Tuna-like Fisheries in the North Pacific Ocean in 2018). ISC/19/PLENARY/05.
- ISC. 2018. Report of the Pacific Bluefin Tuna Working Group Intersessional Workshop. ANNEX 9 of the 18th Meeting of the International Scientific Committee for Tuna and Tuna-Like Species in the North Pacific Ocean. ISC/18/ANNEX/09.
- Liu, H.-I and S.-K. Chang. 2019. CPUE standardizations of Taiwanese PBF fisheries with/without geostatistical consideration. Pacific Bluefin Tuna Working Group Intersessional Workshop of the ISC, 18-22 March, 2019, Jeju, Korea. ISC/19/PBFWG-1/02.
- Thorson, J.T. and L.A.K. Barnett. 2017. Comparing estimates of abundance trends and distribution shifts using single- and multispecies models of fishes and biogenic habitat. ICES Journal of Marine Sciences 74 (5), 1311–1321.
- Tsukahara, Y., H. Fukuda and S. Nakatsuka. 2020. Standardized CPUE by spatiotemporal model for Pacific bluefin tuna (*Thunnus Orientalis*) caught by Japanese coastal and offshore longline. ISC/20/PBFWG-1/02.
- Xu, H., C.E. Lennert-Cody, M.N. Maunder and C.V. Minte-Vera. 2019. Spatiotemporal dynamics of the dolphin-associated purse-seine fishery for yellowfin tuna (*Thunnus albacares*) in the eastern Pacific Ocean. Fisheries Research 213:121-131.
- Yuan, T.L., S.K. Chang, H. Xu and H.I Liu. 2019. Standardized index of relative abundance of Taiwanese PBF fisheries based on traditional and spatiotemporal delta-generalized linear mixed models. Pacific Bluefin Tuna Working Group Intersessional Workshop of the ISC, La Jolla, USA, 18 – 23 November 2019. ISC/19/PBFWG-2/11.

Table 1. Best variable combinations of the delta-lognormal mixed models for GLMM, and the Akaike information criterion (AIC) and Bayesian information criterion (BIC). (ZPM: zero-proportion model; PCM: positive-catch model)

Model type	Final model formulation	AIC	BIC
<i>Southern fishing ground</i>			
ZPM:	Year+Month+CT	22190.44	23561.62
PCM:	Year+Month+CT+Year*Month	22372.88	23737.35
<i>Northern fishing ground</i>			
ZPM:	Year+Month+CT+Year*Month	7787.58	7938.82
PCM:	Year+Month+CT+Year*Month	6671.36	6807.47
<i>Combined southern and northern fishing grounds</i>			
ZPM:	Year+CT+Area	36301.90	36482.97
PCM:	Year+Month+CT+Area+Year*Month	30454.08	30644.24

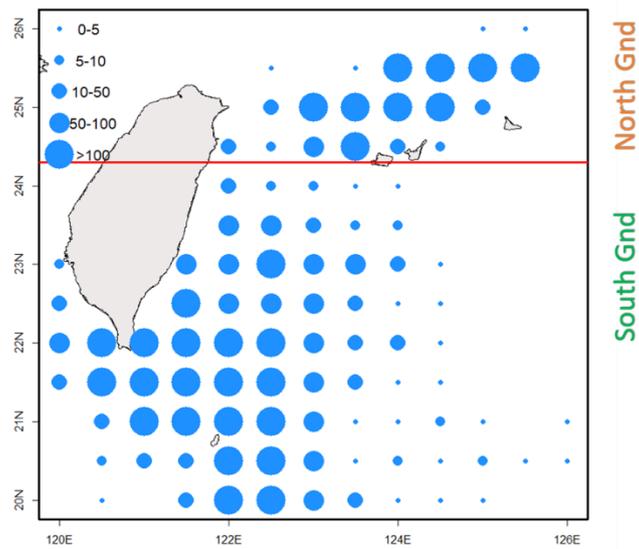


Fig. 1. Average PBF catch distribution off Taiwan for 2010–2020. The line splits the fishing grounds into southern ground and northern ground by 24.3°N.

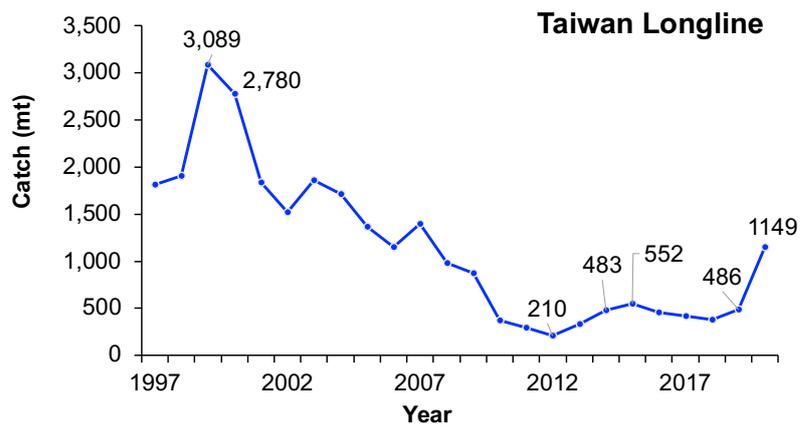


Fig. 2. Annual PBF catches by Taiwanese offshore longline fishery.

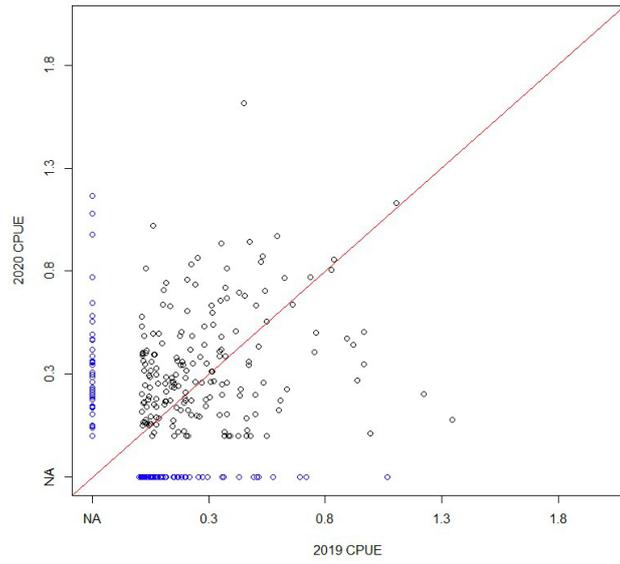


Fig. 3. Nominal CPUE of the vessels that have fished for PBF in 2019 and 2020 (black circles). The blue circles indicate the CPUE of vessels have either fishing in 2019 or 2020 only.

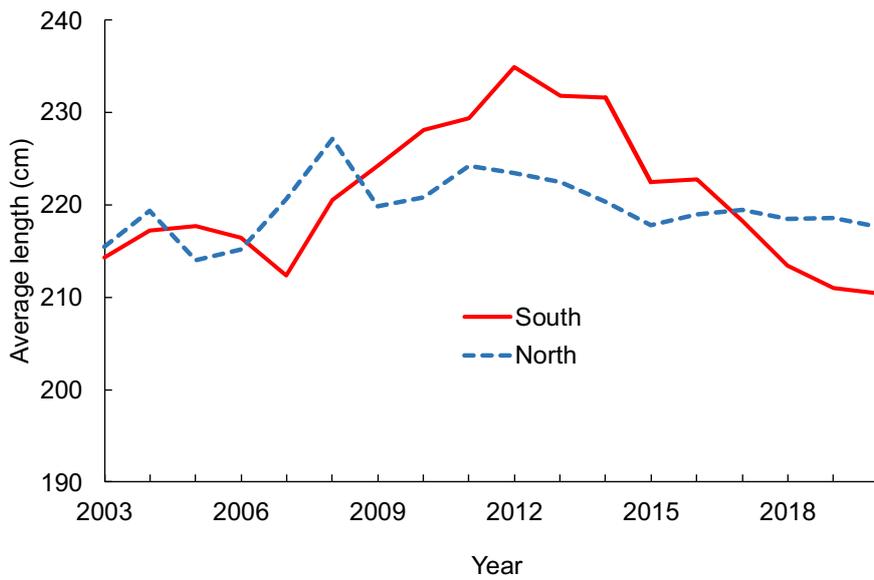


Fig. 4. Annual trend of average length of PBF of Taiwanese longline fishery, by southern and northern fishing grounds.

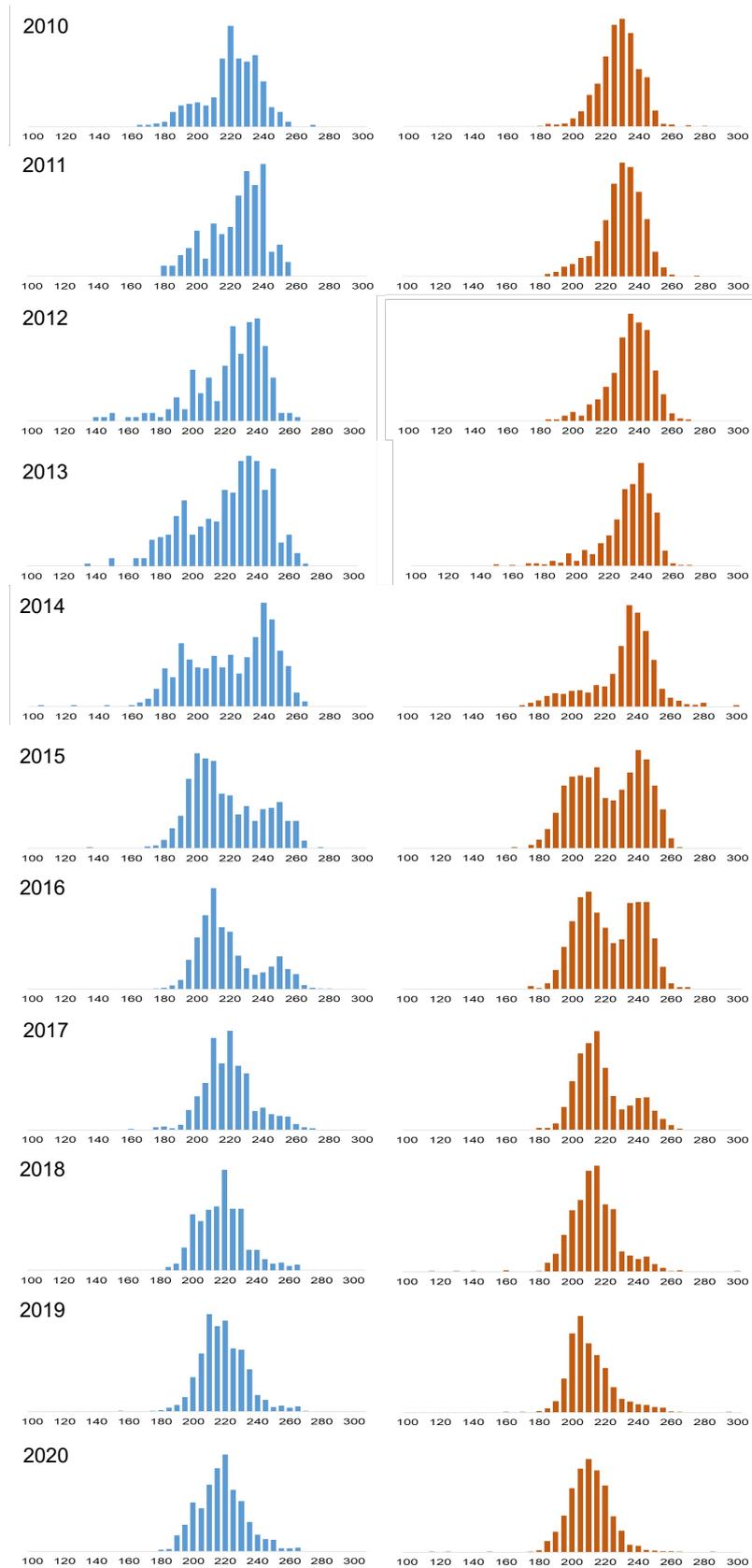


Fig. 5. Length frequencies of Taiwanese PBF during 2010 – 2020 for northern fishing ground (left in blue) and southern fishing ground (right in red).

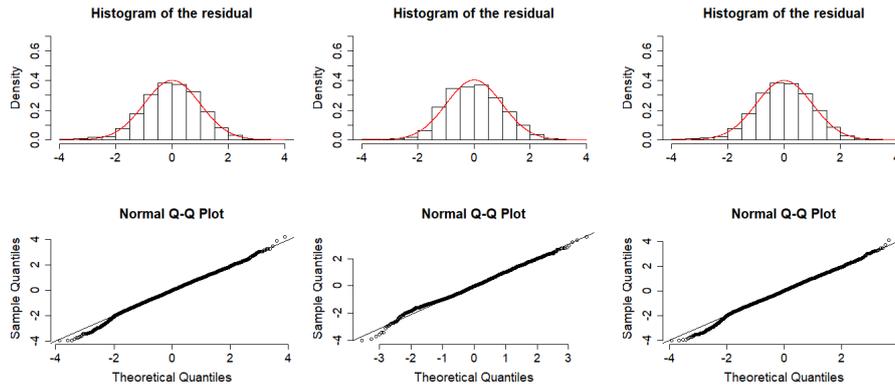


Fig. 6. Diagnostic residual plots (the posterior-predictive residual histogram and qq plot comparing the observed and predicted quantiles of CPUE given encounter) for the traditional delta-GLMM analyses. Panels from left to right are for the southern, northern and whole fishing grounds.

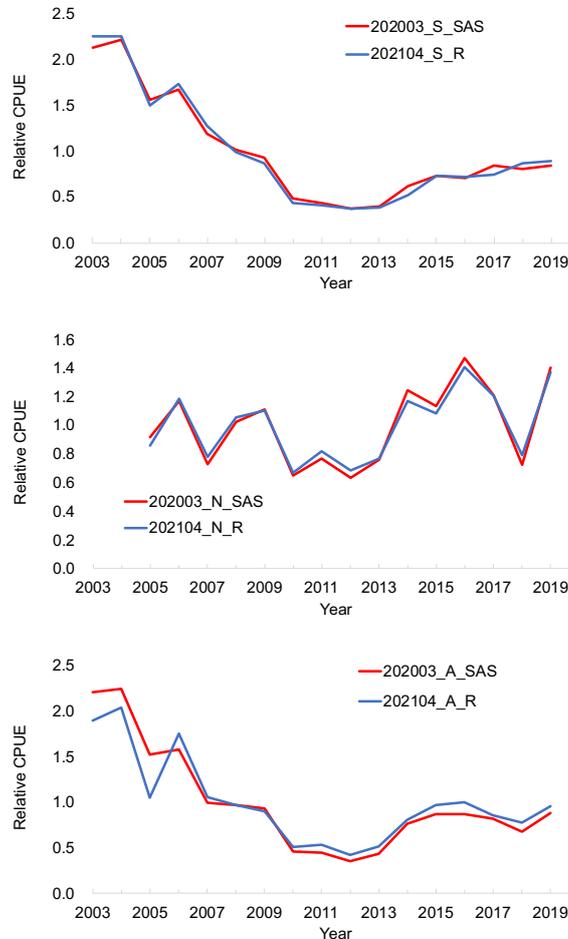
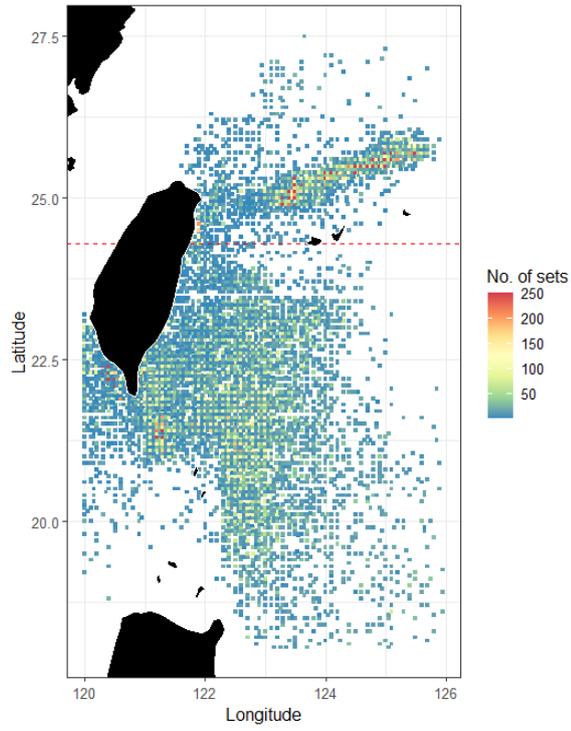


Fig. 7. Relative CPUE series from GLMM under R environment (blue lines) and under SAS environment (red lines), for the southern, northern and whole fishing grounds, respectively.

(A)



(B)

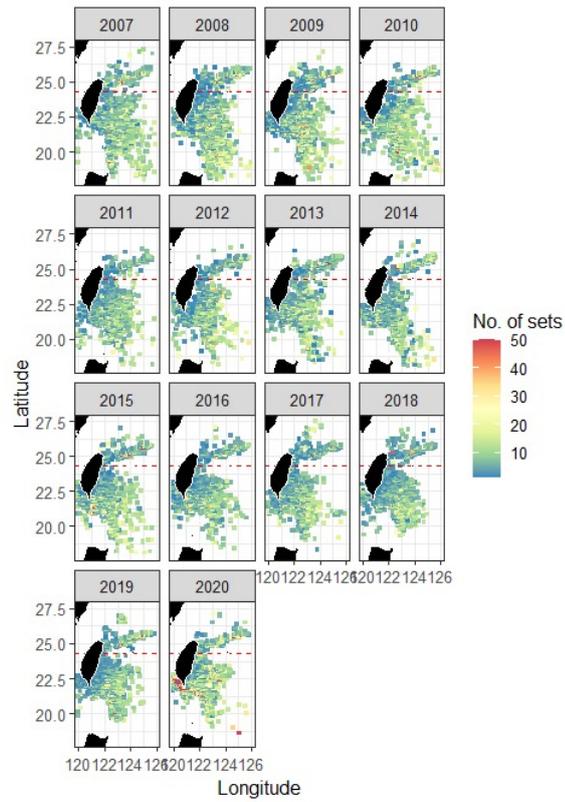


Fig. 8. Geographic distribution of fishing days per $0.1^\circ \times 0.1^\circ$ grid cell during 2007–2020: (A) whole period; (B) by year.

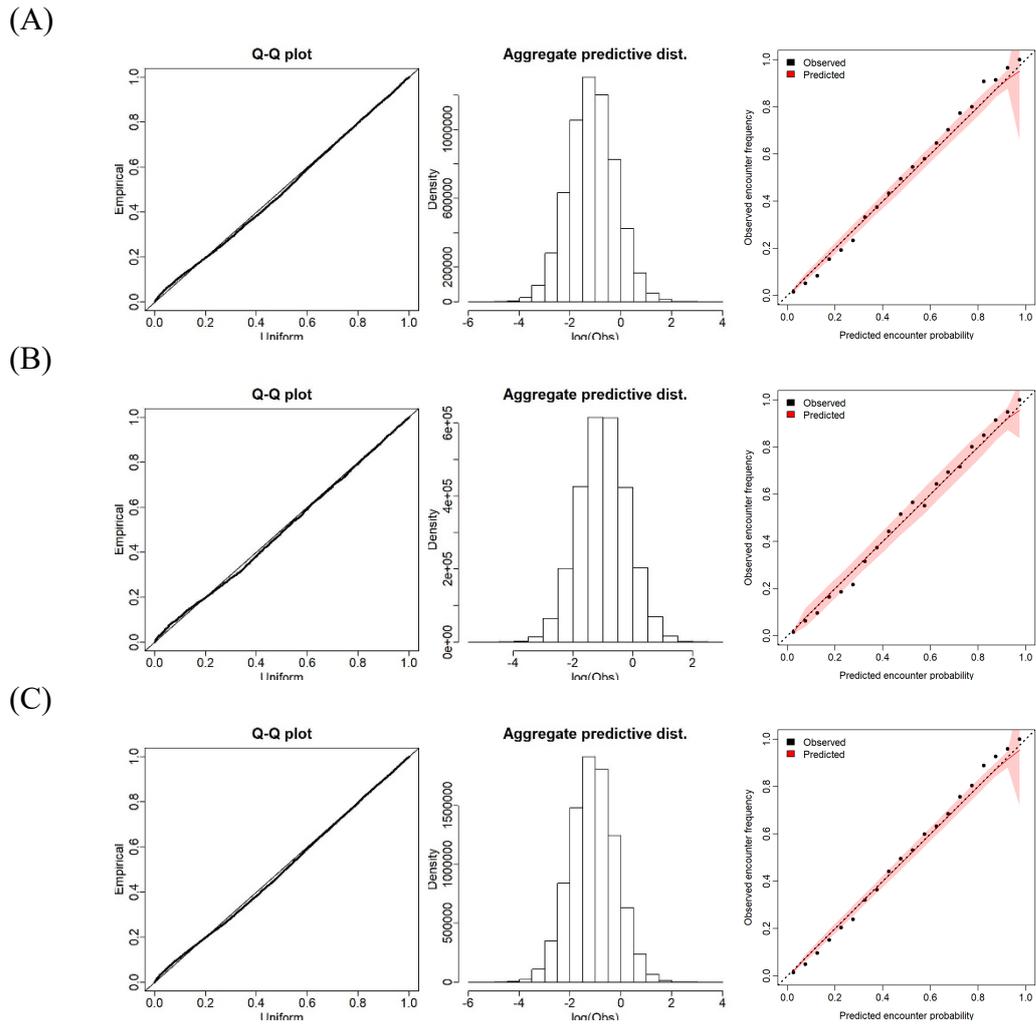


Fig. 9. Diagnostic residual plots for VAST-1 (A for Core-South fishing ground and B for Core-North fishing ground), VAST-2 (C, whole core fishing ground combined). The graphs from left to right: the qq plot comparing the observed and predicted quantiles of CPUE given encounter, the posterior-predictive residual histogram, and encounter probability.

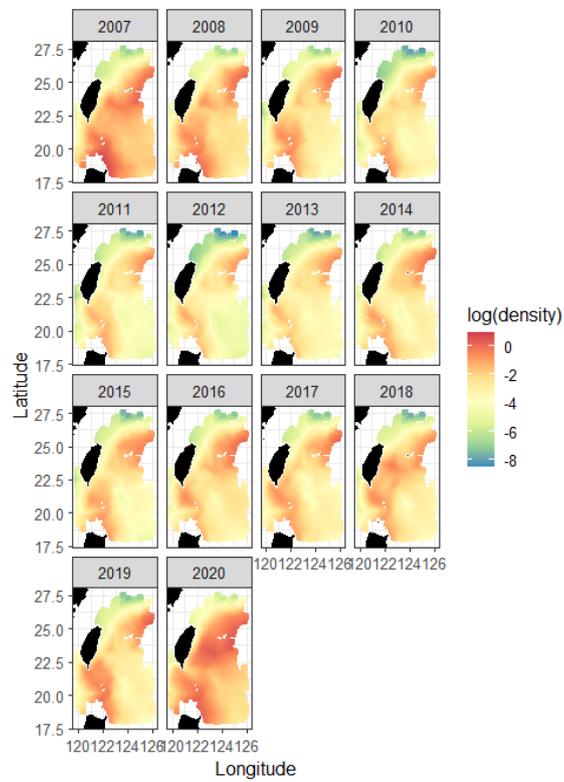
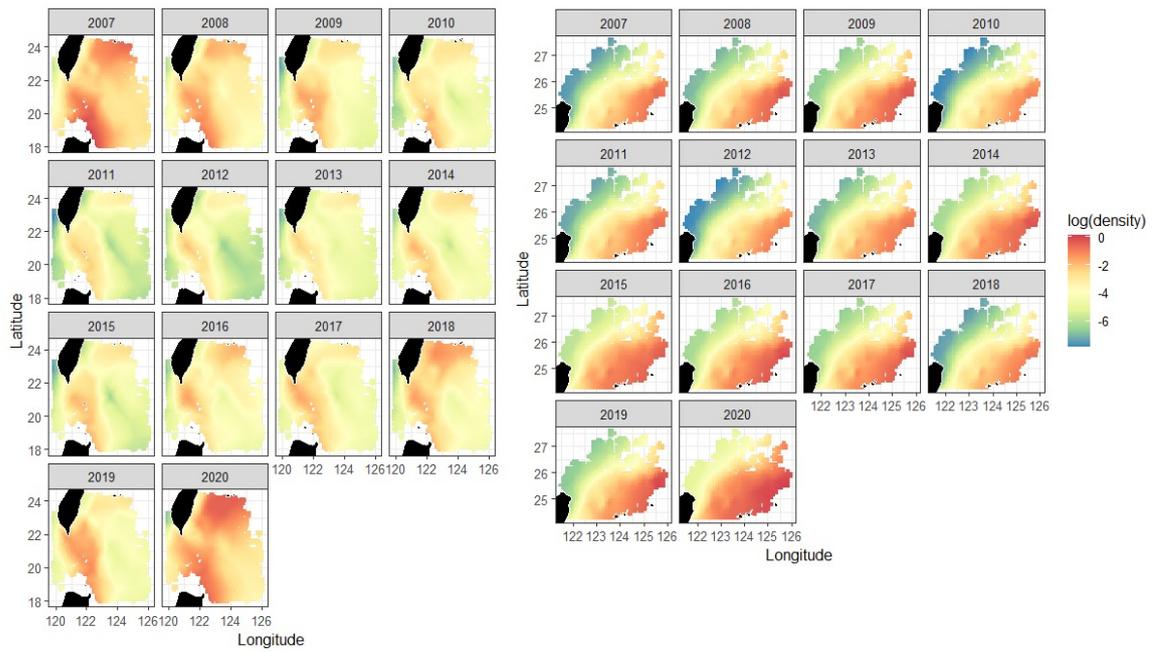


Fig. 10. Spatiotemporal distribution of predicted log density of PBF during 2007–2020 from VAST analyses (upper left – South, upper right – North, bottom: South and North combined).

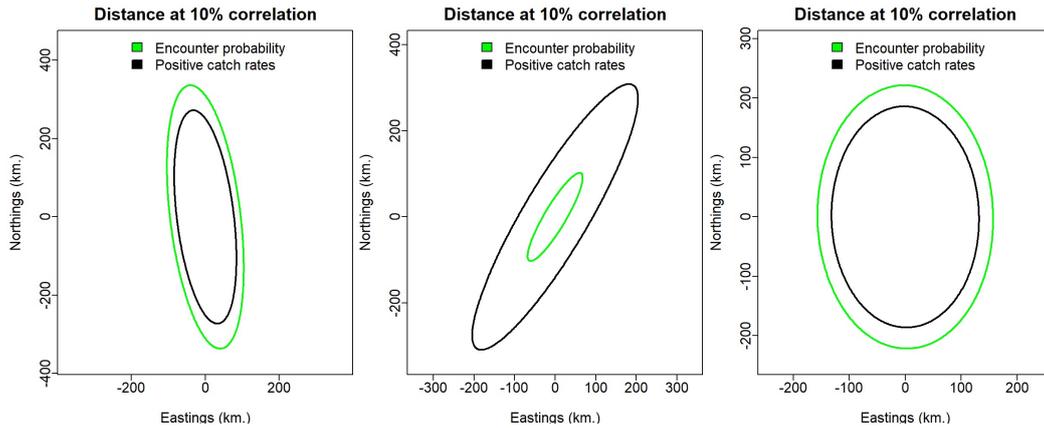


Fig. 11. Distance of 10% correlation of encounter probability and positive catch rate. From left to right: southern region, northern region and both regions combined.

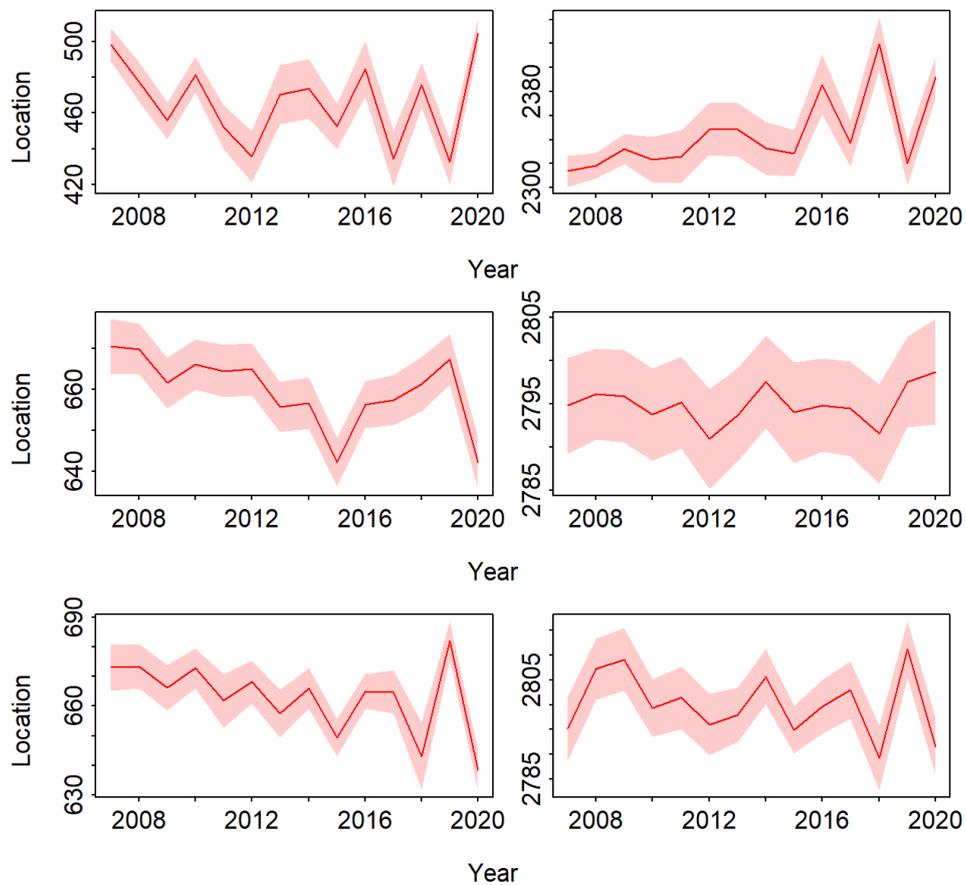


Fig. 12. The center of biomass of PBF in the east-west (left) and north-south (right) direction. From top to bottom: southern region, northern region and both regions combined.

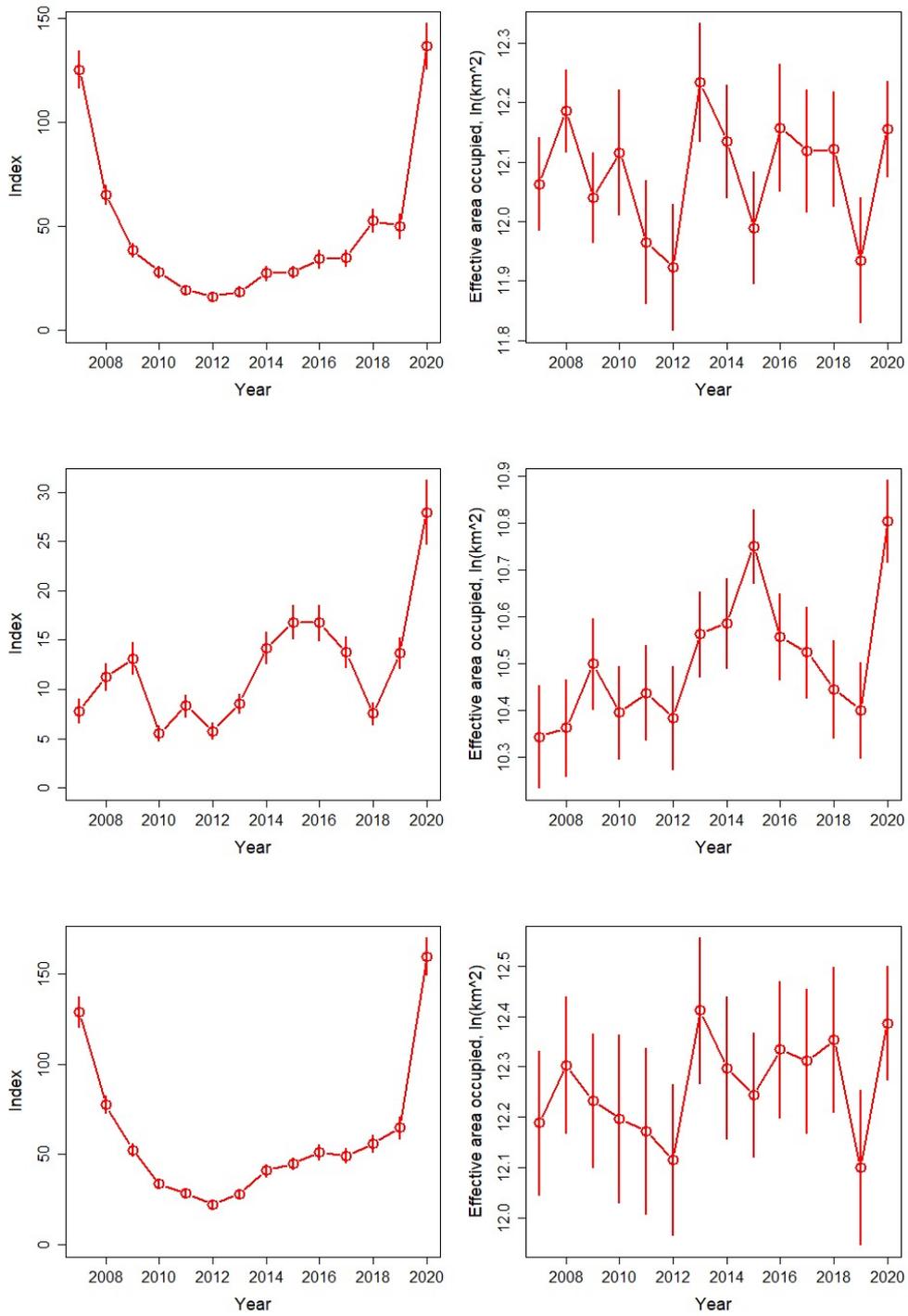


Fig. 13. Standardized index of relative abundance (left) and estimated effective area occupied (right) of PBF. The bars represent the standard errors. From top to bottom: southern region, northern region and both regions combined.

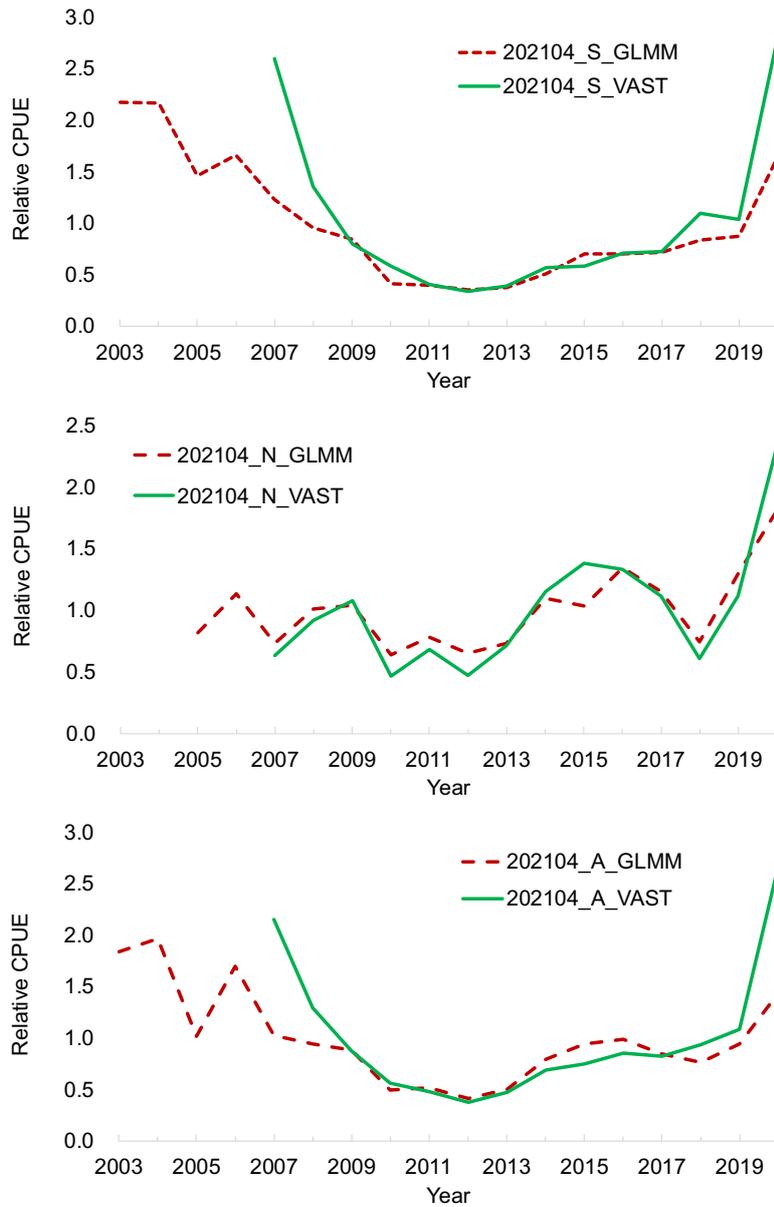


Fig. 14. Relative CPUE series based on traditional delta-GLMM on the whole dataset and on VAST analyses on the core area data. “202104” in the codes indicate the results of April meeting of 2021 (this study). “S”, “N”, and “A” in the codes indicate the results for the southern, northern, whole fishing grounds (southern and northern fishing grounds combined), respectively.