North Pacific Fisheries Commission

NPFC-2022-TWG CMSA06-WP09

Standardizing CPUE of Japanese commercial dip-net fishery targeting spawners of chub mackerel in the Northwest Pacific

Shota Nishijima[1], Yuga Kisara[2], Momoko Ichinokawa[1], Akihiro Manabe[1], and Ryuji Yukami[1]

1. *Fisheries Resources Institute, Japan Fisheries Research and Education Agency*
2. *Fishery Innovation Programs, Division of Regional Innovation and Management, Graduate School of Arts and Science, Iwate University*

Summary

In this document, we provide the summary of the CPUE standardization of Japanese commercial dip-net fishery for chub mackerel following the “CPUE Standardization Protocol for Chub Mackerel”. The year trend of the spawning stock biomass (SSB) was derived from standardized CPUE, by applying the catch-and-effort data of the dip-net fishery targeting spawners of chub mackerel to the delta GLM. Since we found no serious problems in the standardization, we recommend this SSB index to be utilized in the Technical Working Group for the Chub Mackerel Stock Assessment.

(1). Literature review to identify the candidate explanatory variables

Spawning chub mackerel was caught around the Izu Island chain, the main spawning ground of this stock (Fig.1) by the dip-net fishery. Although the catch amount is much smaller than that of the other common fishery such as the purse seine net fishery (Matsuda et al. 1994), the fact that dip-net fishery directly targets the spawning chub mackerels during the spawning season makes the CPUE of the dip-net fishery suitable as an abundance index of SSB for the stock in the Japanese domestic stock assessment.

In the previous document, we reported the standardized CPUE values from 2003 to 2020 (Shinohara et al. 2021). Following this, we conducted CPUE standardization by removing the effects of environmental and spatial variables and updated the result. Since the dip-net CPUE of chub mackerel is known to be affected by water temperature (Nishijima et al. 2021), we used the sea surface temperature (SST) as an explanatory variable. The in-situ SST was recorded in each set. Furthermore, to account for the possible spatial effects on the CPUE, we added a spatial explanatory variable. We used the area category instead of the exact locations of fishing (Figs 1-3) because a large proportion of data (367 out of 2256) lacked the information of the coordinates (longitude and latitude) but included the categorical name of the area of each catch.

(2). Plot of the spatio-temporal distributions of catch, effort, and CPUE.

The data of dip-net fishing from 2003 to 2021 were used (Table 1). We exclusively focused on the data from January to July, the spawning season of chub mackerel, and removed the data obtained during the other months (174 out of 2256). The dip-net fisheries were conducted in the area approximately from 138º–140.5º E and 32.5º–35º N (Fig. 1).

Table 1. The summary of the fishery (number of fisheries, number of positive catches, and the mean nominal CPUE) and the result of standardization (standardized CPUE and the confidence interval).

year	**Number of samples**	**Number of positive catches**	**Mean nominal CPUE**	**Standardized CPUE**	**Lower 95% CI**	**Upper 95% CI**
2003	113	45	5.49	4.22	2.34	6.93
2004	161	74	4.46	5.85	3.98	8.76
2005	116	49	3.29	2.89	1.61	4.83
2006	69	30	25.46	13.26	7.72	21.83
2007	176	176	86.56	48.06	34.35	68.42
2008	81	81	45.53	19.69	13.88	27.72
2009	82	82	56.51	26.38	18.09	39.04
2010	98	94	54.51	24.53	16.84	35.47
2011	70	60	116.21	46.09	30.55	68.79
2012	65	60	120.54	47.13	31.03	69.80
2013	13	13	131.91	70.83	33.75	131.38
2014	117	115	110.94	50.16	34.51	72.47
2015	83	77	120.32	47.34	31.01	70.54
2016	126	120	172.48	80.31	56.74	113.45
2017	123	107	81.48	42.15	28.92	61.18
2018	113	106	142.86	71.10	49.46	101.80
2019	120	110	142.44	51.87	37.07	72.18
2020	177	165	167.34	54.35	39.85	74.46
2021	179	146	115.21	47.34	33.30	65.75

Figure 1. Map of the area category (area 1 to 7). Each point represents the center of the fishing location of each category of the area, and error bars represent the dispersion (1 SD) of the fishing locations within the same category.

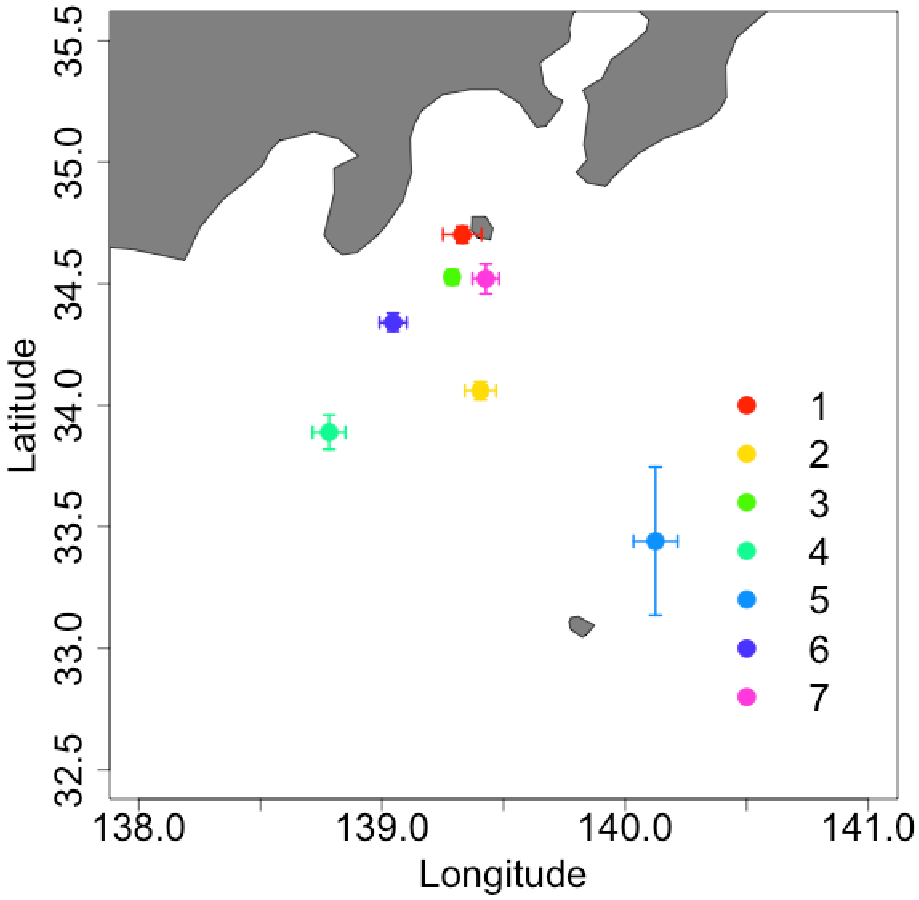


Figure 2. Spatio-temporal trends of the dip-net fishing efforts (man-hour).

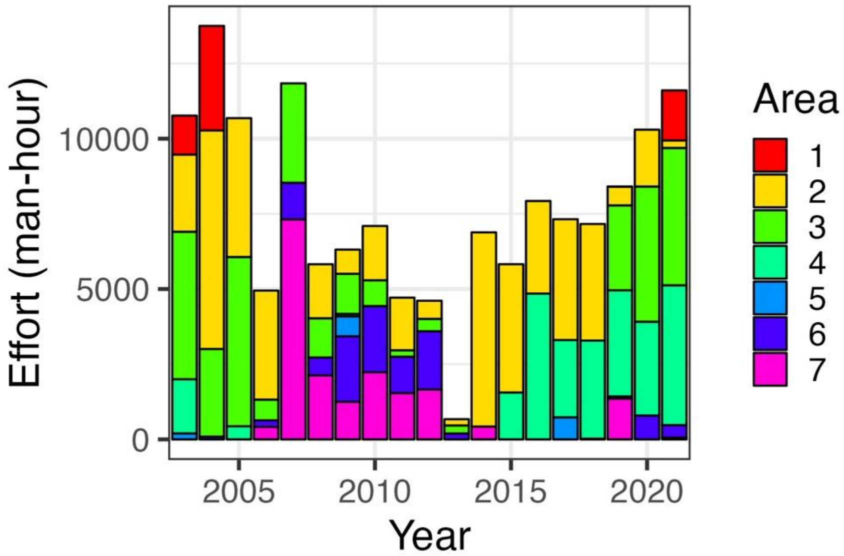
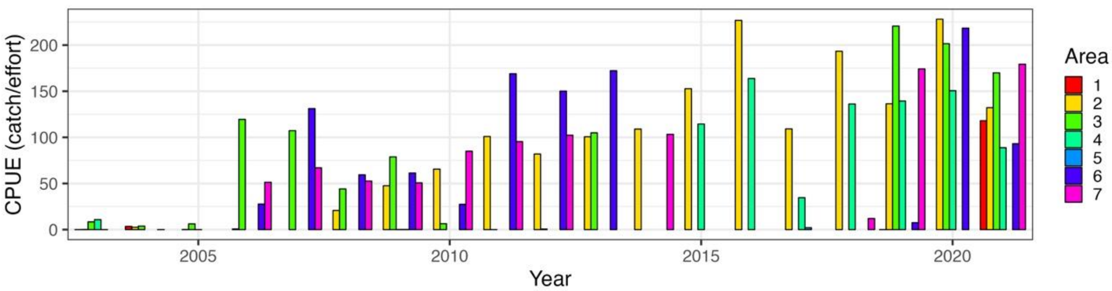


Figure 3. Spatio-temporal trends of the dip-net fishing CPUEs (catch biomass per man-hour).



(3). **Plots representing the correlation between the variables**

We present (i) the yearly trend of SST (Fig. 4), (ii) the spatial difference in SST (Fig. 5), (iii) the yearly trend of the CPUE (Fig. 6), (iv) the spatial distribution of the CPUE (Fig. 7), and (v) the relationship between SST and the CPUE (Fig. 8).

Figure 4. The yearly trend of sea surface temperature (SST).

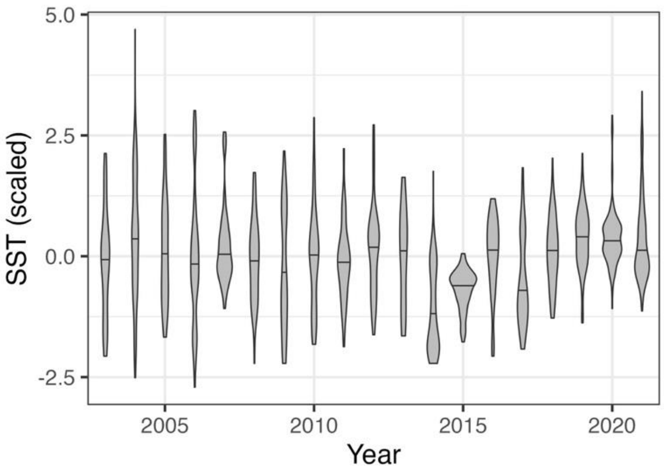


Figure 5. The spatial difference in sea surface temperature (SST).

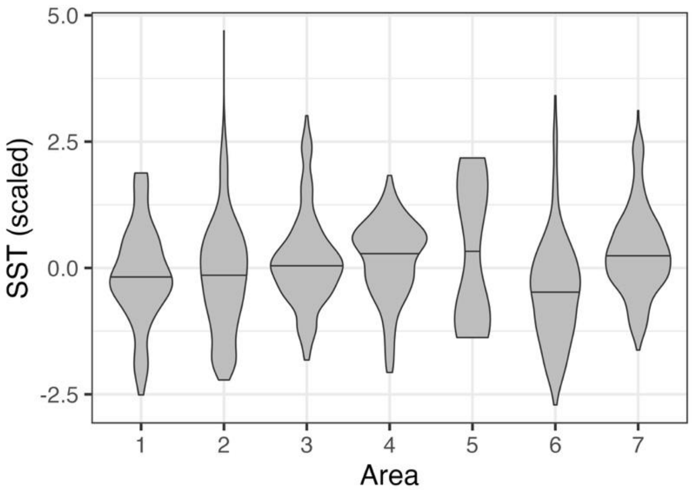


Figure 6. The yearly trend of the number of positive CPUE (left panel) and the average positive CPUE (right panel). The y-axis of the right panel is log-scaled.

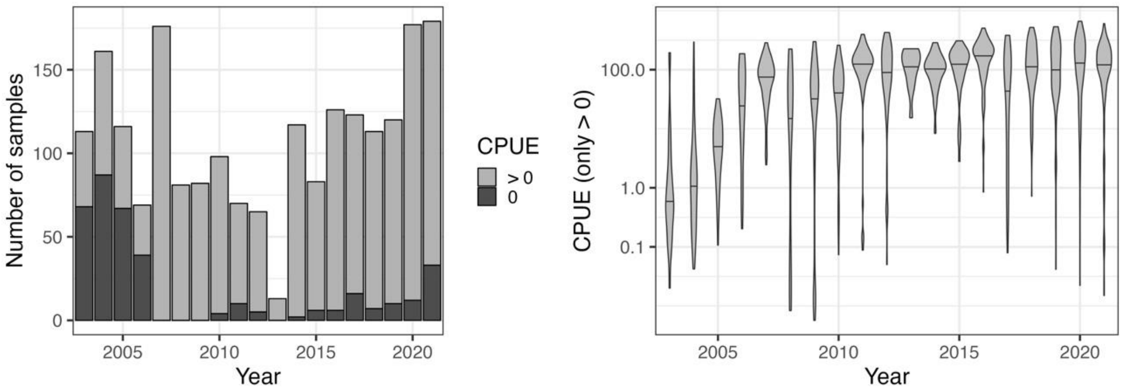


Figure 7. The spatial distribution of the number of positive CPUE (left panel) and the average positive CPUE (right panel). The y-axis of the right panel is log-scaled.

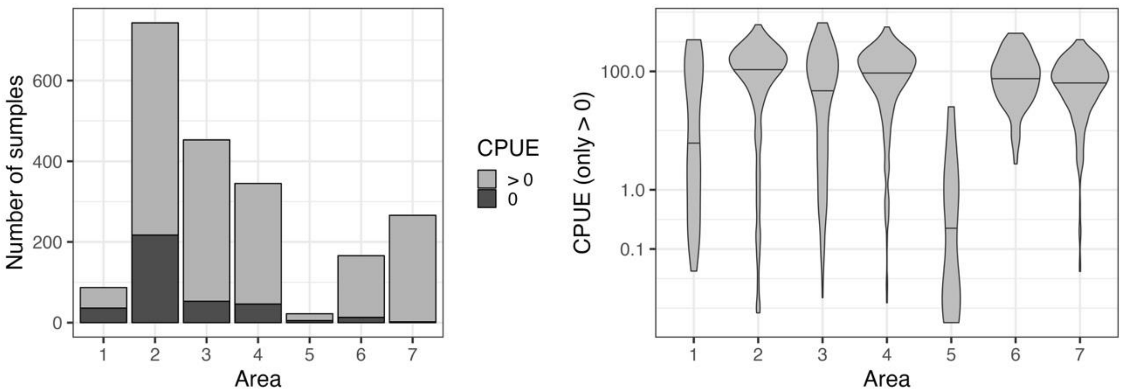
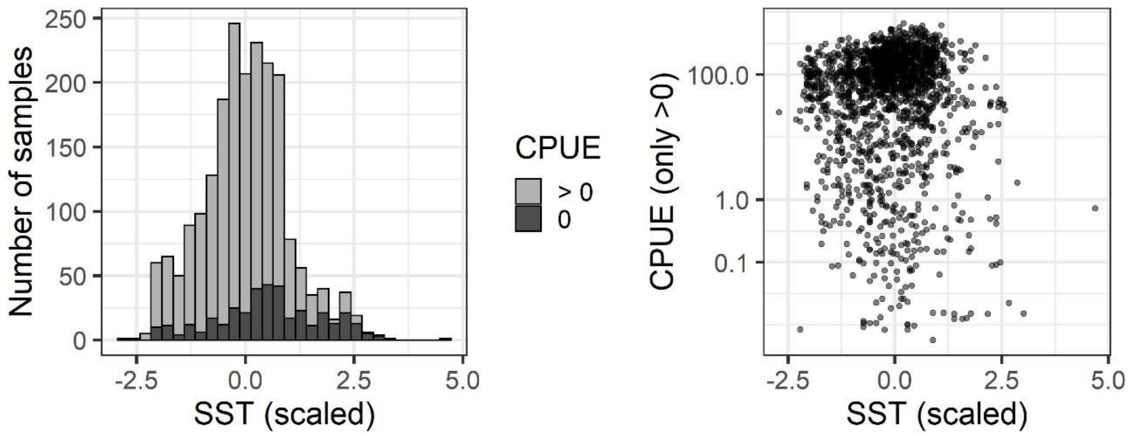


Figure 8. Relationship between sea surface temperature (SST) and the proportion of positive CPUE (left panel) or the positive CPUE values (right panel, log-scale).



(4). **Explanatory variables in the full model**

We incorporated the following as the fixed effects: (i) year (categorical), (ii) month (categorical), (iii) area (categorical), (iv) ship (categorical), (v) sea surface temperature (SST) (continuous), and (vi) SST^2.

(5). **Model details**

We used delta GLM for the standardization of the dip-net fishery CPUE. Delta GLM is the two-step generalized linear model where the probability of occurrence and the density (or CPUE) when occurred were modelled separately. We modelled the probability of occurrence with binomial distribution (logit link) and the CPUE when occurred with gamma distribution (log link). The gamma distribution was used because gamma models generally obtained less biased and more robust estimates than lognormal models and, therefore, it is suggested to use a gamma distribution for index standardization (Cadigan and Myers 2001; Thorson et al., 2021). The distribution of the CPUE modelling was selected based on AICc.

(6). **Best model**

We performed the brute-force model selection approach and determined the best model based on AICc (Table 2). The best model with the lowest AICc was used for standardization.

Table 2. Model selection for the standardization of summer recruitment CPUE. The selected explanatory variables in each model are indicated as “+” notation for categorical variables or coefficient values for continuous variables.

Occurrence model (binomial)										Positive CPUE model (gamma)									
Explanatory variables					Year	df	logLik	AICc	ΔAICc	Explanatory variables						df	logLik	AIC	ΔAICc
Area	Month	Ship	SST	SST^2						Area	Month	Ship	SST	SST^2	Year				
+	+		-0.466	-0.207	+	28	-479.941	1016.837	0.000	+	+	+	0.108	-0.076	+	40	-9184.278	18450.521	0.000
+	+		-0.469	-0.203	+	29	-479.067	1017.158	0.321	+	+	+	0.133		+	39	-9187.509	18454.886	4.364
+	+		-0.446		+	28	-483.566	1024.087	7.250	+	+	+			+	38	-9191.413	18460.600	10.079
+	+		-0.443		+	27	-484.629	1024.147	7.310	+	+		0.152	-0.063	+	34	-9200.577	18470.575	20.054
+	+	+	-0.450	-0.204	+	34	-477.412	1024.229	7.392	+	+		0.170		+	33	-9202.823	18472.984	22.463
+	+	+	-0.432		+	33	-481.917	1031.157	14.320	+	+		0.140	-0.060	+	33	-9206.216	18479.770	29.249
+	+				+	26	-491.252	1035.329	18.492	+	+		0.157		+	32	-9208.261	18481.782	31.261
+	+				+	27	-490.244	1035.377	18.540	+	+				+	32	-9209.462	18484.183	33.662
+	+	+			+	32	-488.072	1041.390	24.552	+	+				+	31	-9213.976	18491.133	40.612
+			-0.920	-0.920	+	24	-513.606	1073.860	57.023	+		+	0.024	-0.134	+	35	-9233.543	18538.591	88.070

(7). Diagnostics of the model and the residuals

The best binomial (Fig.9) and gamma (Fig.10) models were diagnosed by checking the distribution of the residuals along important variables (here, year and area).

Although there were no systematic trends in the residuals of the binomial models along years, the residuals appear to be more clustered after 2007 (Fig. 9a) probably because the probability of positive CPUE is high (Fig. 6). The residuals of the binomial models were not biased by area (Fig. 9b). In addition, the best binomial model was diagnosed by the area under the ROC (receiver operating characteristic) curve (AUC), which quantifies the performance of the classification model and ranges from 0 to 1 where 0.5 suggests the random prediction and 1 suggests 100% correct prediction. Generally, 0.8 to 0.9 AUC value is considered as a good prediction ability. The AUC of the binomial model was 0.911 (Fig. 9c), suggesting its good prediction.

The residuals of the best gamma model were not apparently biased by area (Fig. 10c). As for the temporal trends, however, we observed lower residuals during 2003 to 2005 (Fig. 10a), probably reflecting that most of the positive catches during that time period were observed in Area 3 where relatively higher CPUE is expected (Fig. 12). Nevertheless, the QQ-plot shows that residuals did not strongly deviate from the expected gamma distribution.

Figure 9. Diagnostics of the best binomial model. (a) Temporal and (b) spatial trends of the deviance residuals, and (c) the receiver operating characteristic (ROC) curves with the area under the curve (AUC) value.

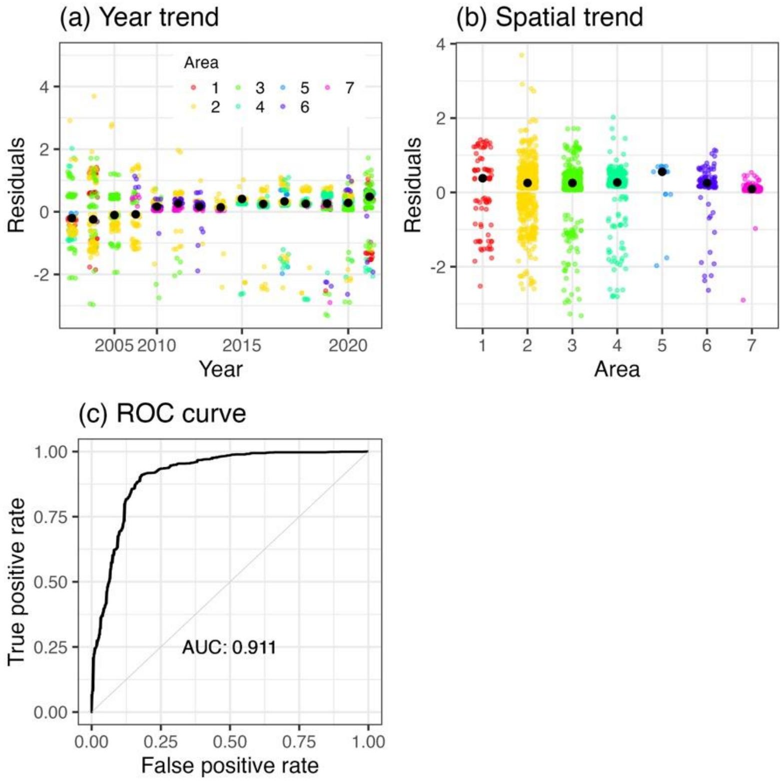
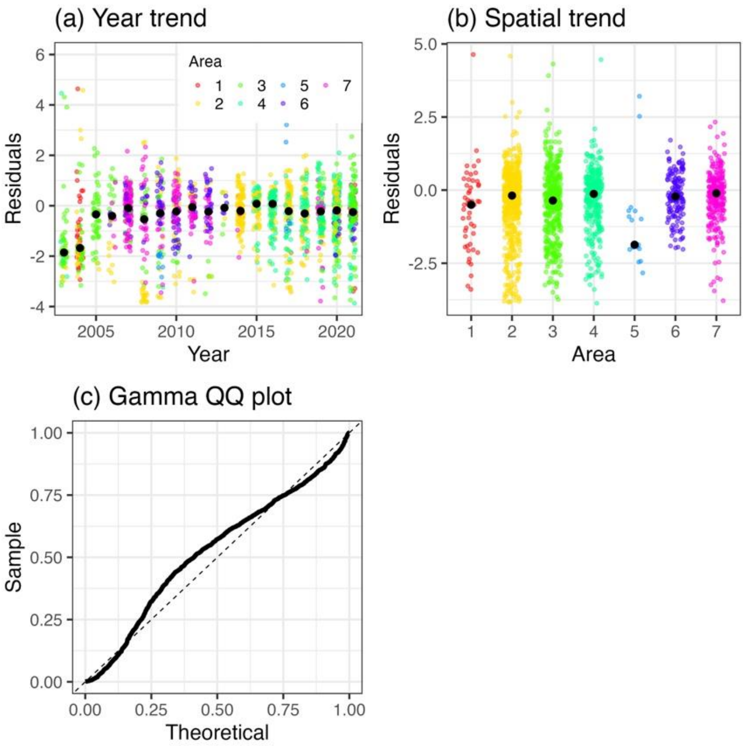


Figure 10. Diagnostics of the best gamma model. (a) Temporal and (b) spatial trends of the deviance residuals, and (c) the QQ-plot.



(8). Estimated relationships between the explanatory variables and the response variable

In the best binomial and gamma models, sea surface temperature and its squares (SST and SST^2) were retained as an explanatory variable (Table 2). A peak mode was found 16.4°C in the binomial model (Fig. 11a) and 20.1°C in the gamma model (Fig. 11b). The probability of occurrence and the positive CPUE values differed among the areas (Fig. 12a, b). Both occurrence probability and positive CPUE were high from February to April (Fig. 12c, d). The identity of ship affected positive CPUE, but not occurrence probability (Fig. 12e,f).

Figure 11. Relationship between sea surface temperature (SST) with the probability of positive catch (a) or with the positive CPUE values (b), estimated from the best binomial or gamma model, respectively. The lines show predicted values obtained by averaging effects of other explanatory variables.

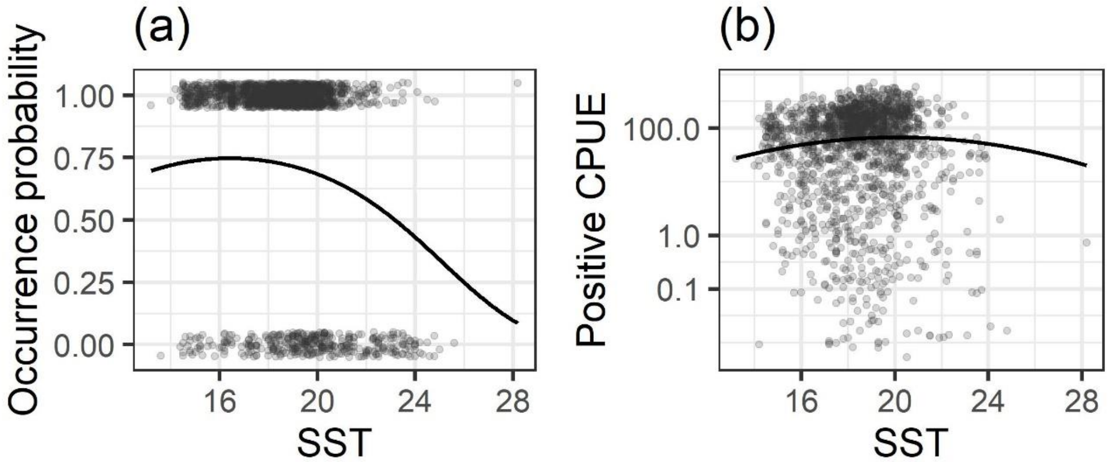
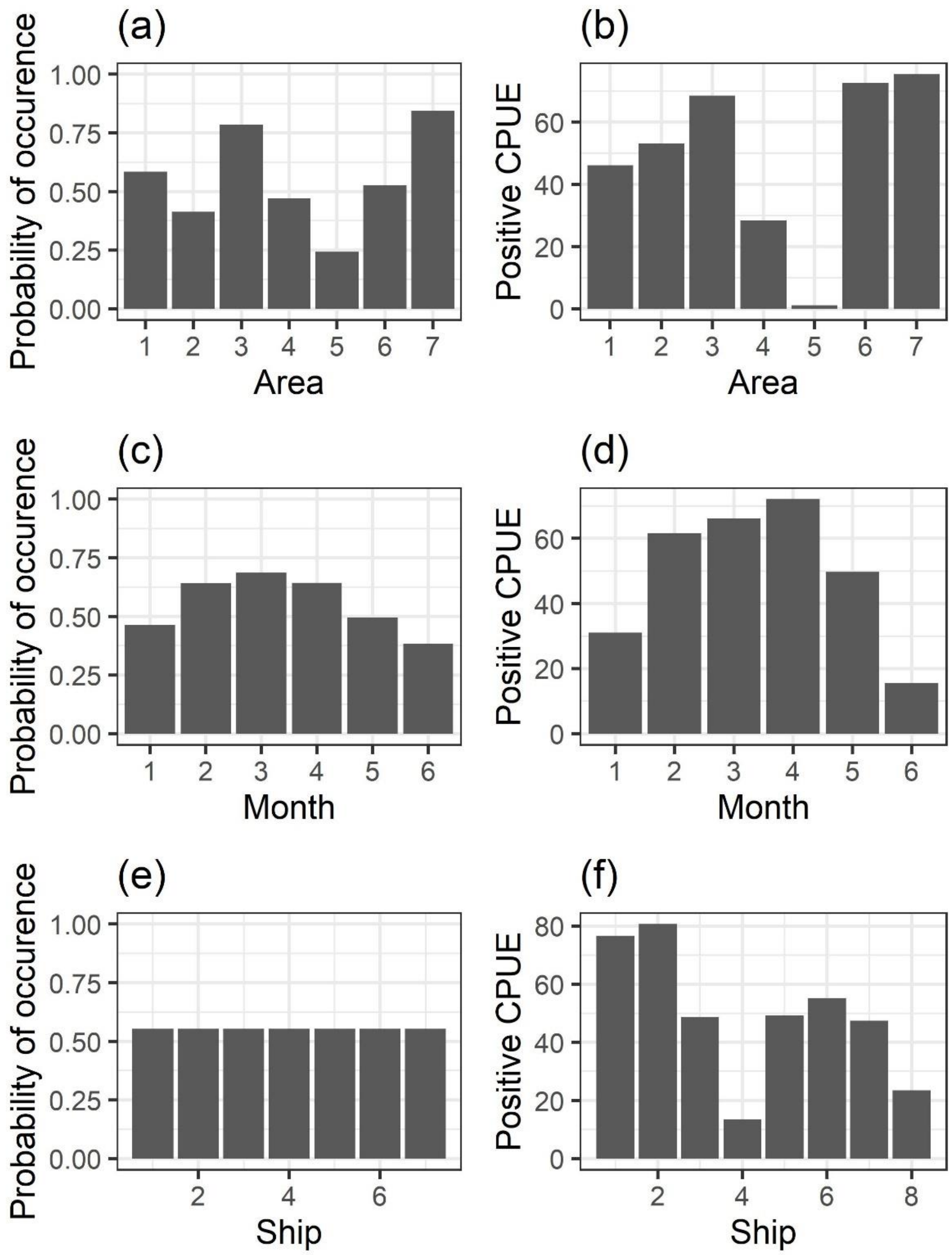


Figure 12. Difference in the probability of occurrence (left) and positive CPUE (right) among areas (upper), months (middle), and ships (lower). The bars show predicted values obtained by averaging effects of other explanatory variables.



(9). **Yearly standardized CPUE and its uncertainty**

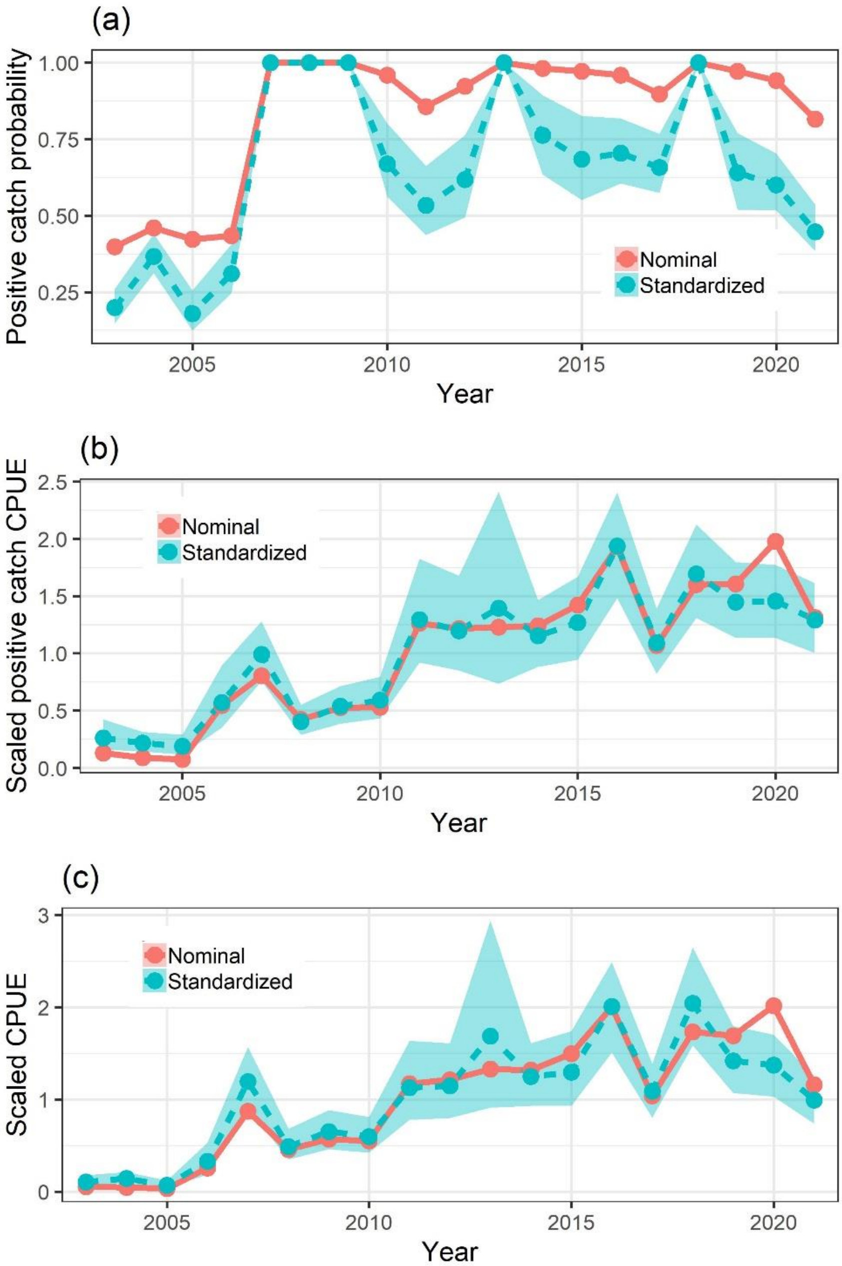
To derive the standardized CPUE values, we calculated predicted CPUE values per each category (for the continuous variables, we divided their range at small regular intervals) of selected variables (e.g., Area = 1, 2, 3…, Year = 2002, 2003, 2004…, SST = 10.0, 11.0, 12.0…), and calculated the arithmetic mean of the yearly predicted values. This averaging for extracting the year trend was necessary due to the nonlinearity of the logit link function in the delta-gamma model.

Confidence intervals were evaluated by sampling posterior distributions of parameter estimates with 1000 replicates using the R package 'arm' (Gelman and Su 2021). The standardized CPUE values and confidence intervals are shown in the next section.

(10). **Comparison of the nominal and standardized CPUEs**

The overall yearly trend was similar between nominal and standardized CPUEs (Fig.13b, c), except for recent two years (2019, 2020, Fig. 13b, c) when the standardized CPUE was lower than nominal value. In 2019 and 2020, a large proportion of the fishery was conducted in area 3 (Fig. 2). SST was relatively high in these two years (19.5°C in 2019 and 19.4°C in 2020, Fig. 5) and this led to high positive CPUE because the peak mode of quadratic function was 20.1 °C (Fig.11b). Therefore, standardized CPUE values in these two years were lower than nominal values by removing the sampling bias of SST. In 2020, furthermore, higher CPUE was observed in Area 2 than predicted (Figure 3), which caused higher nominal CPUE than standardized CPUE.

Figure 13. The yearly patterns of nominal and standardized values of (a) the probability of positive CPUE, (b) positive CPUE, and (c) average CPUE after scaling (divided by means). Blue shaded areas are 95% confidence intervals of standardized CPUE.



Reference lists

Cadigan, N. G., Myers, R. A. 2001. A comparison of gamma and lognormal maximum likelihood estimators in a sequential population analysis. Canadian Journal of Fisheries and Aquatic Sciences, 58, 560–567. https://doi.org/10.1139/cjfas-58-3-560

Gelman, A, Su, Y. 2021. arm: Data Analysis Using Regression and Multilevel/Hierarchical Models. R package version 1.12-2, <https://CRAN.R-project.org/package=arm>.

Matsuda, H., I. Mitani, K. Asano. 1994 Impact factors of purse seine net and dip net fisheries on a chub mackerel population. Researches on population ecology. 36, 201–207.

Shinohara N., Nishijima, S., M. Ichinokawa. 2021. Standardizing CPUE of Japanese commercial dip-net fishery targeting spawners of chub mackerel in the Northwest Pacific. NPFC-2021-TWG CMSA04-WP03.

Thorson, J. T., Cunningham, C. J., Jorgensen, E., Havron, A., Hulson, P. J. F., Monnahan, C. C., von Szalay, P. 2021. The surprising sensitivity of index scale to delta-model assumptions: Recommendations for model-based index standardization. Fisheries Research, 233, 105745. https://doi.org/10.1016/J.FISHRES.2020.105745