NPFC-2022-TWG CMSA06-WP10

Standardizing monthly egg survey data as an abundance index for spawning stock biomass of chub mackerel in the Northwest Pacific

Shota Nishijima[1], Kaito Kanazawa[1,2], Momoko Ichinokawa[1] and Ryuji Yukami[1], Akihiro Manabe [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

We estimated a relative abundance index of spawning stock biomass (SSB) for the Pacific stock of chub mackerel with monthly egg density data obtained by research surveys. The monthly egg surveys have been conducted from 2005 to 2021 off the Pacific coast of Japan to cover the spawning ground of chub mackerel. We applied the vector-autoregressive spatio-temporal (VAST) model to the survey data to derive the index of egg abundance, which should represent relative SSB. This document provides important references and diagnostics on this standardization according to the "CPUE Standardization Protocol for Chub Mackerel". Since we found no serious problems in the diagnostics and confirmed the convergence of the spatio-temporal model, we suggest the estimated index can be utilized as an SSB abundance index for the forthcoming stock assessment of chub mackerel in the Technical Working Group for the Chub Mackerel Stock Assessment.

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

Spatial variables: To account for the spatial autocorrelation and spatio-temporal interaction of the egg density (Kanamori et al. 2019), we incorporated the spatial and spatio-temporal random effects in the model.

Environmental variables: Although sea surface temperature (SST) is known as possibly affecting spatial distribution of the spawning ground of chub mackerels (Kanamori et al. 2019), a previous document suggested that its influence was weak probably because the temporal resolution of our available SST data (annual mean) does not match the biological time scale of the spawning and egg hatching (Kanamori et al.

2018). Therefore, we did not include SST as the explanatory variable here, which would be a room for improvement of the standardization.

(2). Spatio-temporal distributions of catch, effort, and CPUE.

Survey summary: Conical or cylindrical conical plankton nets with mouth ring diameters of 45 or 60 cm and mesh sizes of 0.33 or 0.335 mm were towed vertically form 150 m depth (if the depth was <150 m, nets were lowered to just above the bottom) (see details of survey method for Takasuka et al. 2008 a,b, Takasuka et al. 2017).

Temporal distribution: The survey is conducted monthly from 2005 to 2021. Although the survey data was available throughout the year, we used the data obtained during January to June so that the main spawning season of chub mackerel was covered. Indeed, the mean egg density was substantially higher during January to June (6.90 /m^2, with 55.7 SD) than during the other months (0.40 /m^2, with 6.61 SD). Number of surveys did not systematically vary among years (Table 1, Fig. 1).

Spatial distribution: The surveys were conducted in the area from 131.5º–149.5º E and 26.5º–42.5º N (Fig. 1).

Figure 1. The spatio-temporal distribution of the monthly egg surveys. The black rectangles indicate that the survey was conducted in that area in at least one month during the year.

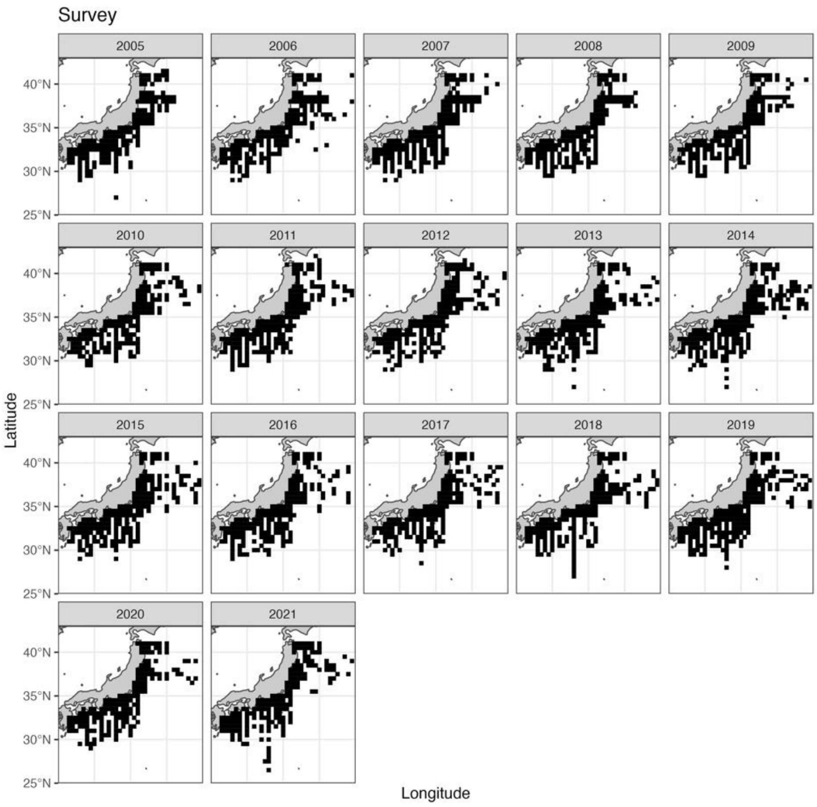
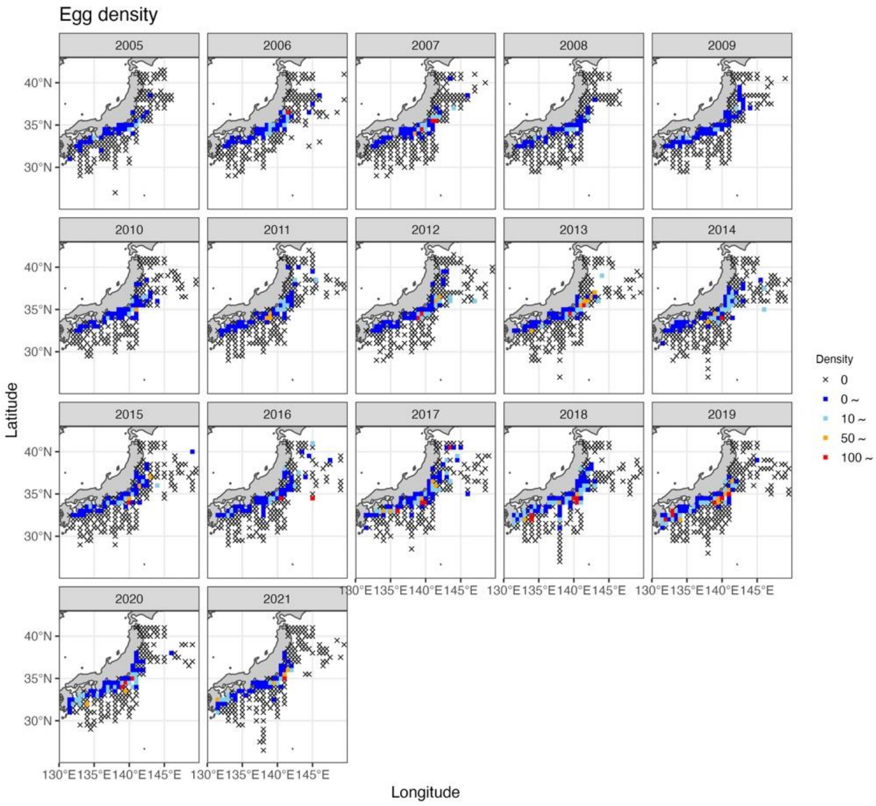


Table 1. The summary of the survey (number of surveys, number of positive catches, and the nominal value of egg density) and the result of standardization (estimated abundance and log SD) for each year.

Year	**Number of surveys (Grids x Months)**	**Number of positive catches**	**Mean catches**	**Estimated abundance**	**Estimated log SD**
2005	555	56	1.243	0.200	0.684
2006	573	81	3.839	0.426	0.609
2007	616	81	7.910	1.015	0.573
2008	583	70	2.302	0.366	0.584
2009	597	98	2.235	0.355	0.541
2010	571	92	2.553	0.422	0.583
2011	579	89	4.095	0.635	0.581
2012	585	88	6.801	0.999	0.602
2013	592	99	7.841	1.036	0.577
2014	624	105	4.886	0.738	0.572
2015	593	89	4.472	0.804	0.597
2016	609	109	3.149	0.776	0.577
2017	560	144	12.700	1.906	0.549
2018	571	162	18.594	2.308	0.502
2019	632	147	15.241	1.850	0.483
2020	542	125	11.589	2.139	0.499
2021	517	105	7.084	1.024	0.520

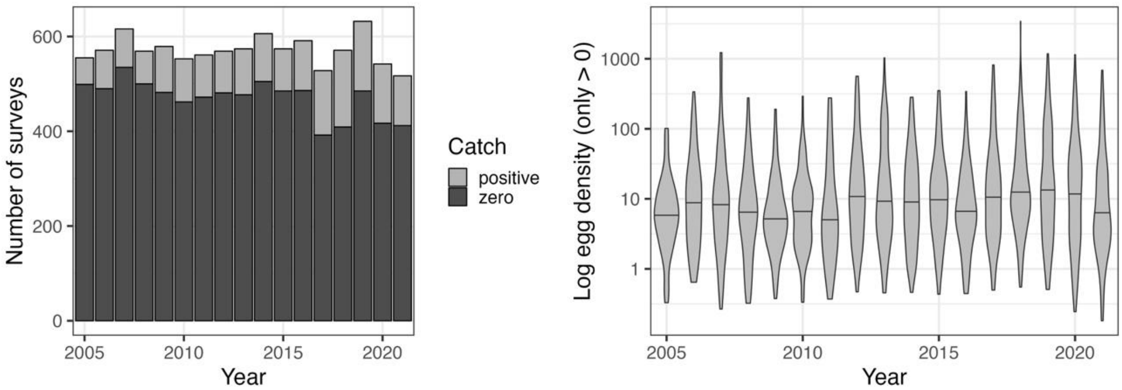
Figure 2. The spatio-temporal distribution of the egg density. For each grid (0.5º latitude * 0.5º longitude), egg density was averaged over the different months of the same year for the illustration purposes.



(3). Plots representing the correlation between the variables

Here, we present the yearly trend of the proportion of positive catches and egg density in Figure 3.

Figure 3. The yearly trend of the number of positive catches (left panel) and the egg density (right panel). Note that, in the right panel, y-axis is log-scale and only positive egg density is shown.



(4). Explanatory variables in the full model

Following variables were included as the explanatory variables.
Fixed effects: Year (categorical)
Random effects: Spatial and spatio-temporal random factors

(5). Model details

We used the vector autoregressive spatio-temporal (VAST) model (Thorson 2019), which accounts for the spatio-temporal changes in survey design and observation rates and can accurately estimate relative local densities at high resolution. The model has been used for various objectives such as standardization of CPUE (e.g., Thorson et al. 2015) and distribution shifts (e.g., Thorson et al. 2016, Kanamori et al. 2019).

The model includes two components, (i) the encounter probability $p_{y,i}$ for year y at location i and (ii) the expected egg density $d_{y,i}$ when spawning egg are encountered. Encounter probability $p_{y,i}$ and positive density $d_{y,i}$ are approximated using Gaussian random fields (a multidimensional generalization of Gaussian process):

$$\operatorname{logit} p_{y,i} = {\beta_y}^{(p)} + {L_\omega}^{(p)}\omega_i + {L_\varepsilon}^{(p)}\varepsilon_{y,i},$$
$$\log d_{y,i} = {\beta_y}^{(d)} + {L_\omega}^{(d)}\omega_i + {L_\varepsilon}^{(d)}\varepsilon_{y,i},$$

where β_ys are the year specific intercepts, L_ωs and L_εs are spatial and spatio-temporal random effects. More detailed information about this model was provided by Thorson (2019).

After estimating the parameters using the *VAST* package in R, the index of abundance in year y at location i (i.e., local egg density), $\hat{d}_{y,i}$, and the index of abundance in year y is (i.e., yearly egg density), $\widehat{D}_y$, were obtained as:

$$\hat{d}_{y,i} = \operatorname{logit}^{-1}\left[{\beta_y}^{(p)} + {L_\omega}^{(p)}\omega_i + {L_\varepsilon}^{(p)}\varepsilon_{y,i}\right] \times \exp\left[{\beta_y}^{(d)} + {L_\omega}^{(d)}\omega_i + {L_\varepsilon}^{(d)}\varepsilon_{y,i}\right],$$
$$\widehat{D}_y = \sum_i a_i\, \hat{d}_{y,i},$$

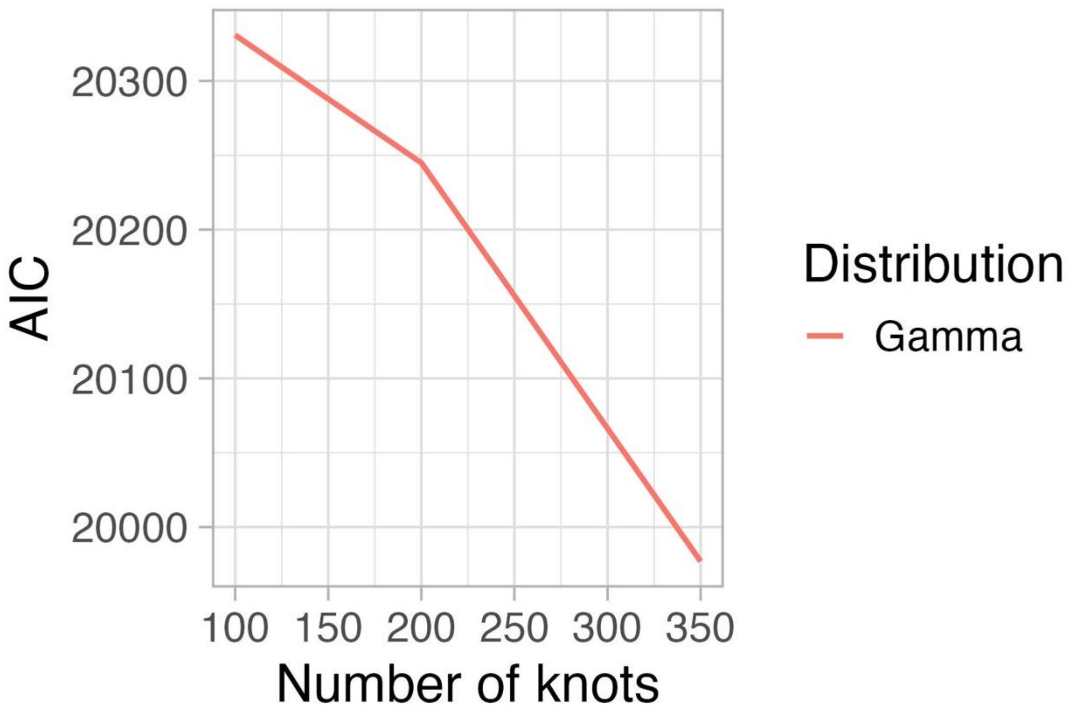
where a_i is area associated with location i. In this document, a_i is fixed as 1 because the area of each location was equal.

The response variables in the positive density were assumed to follow a gamma distribution with 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). Spatial resolution (number of knots) for the spatio-temporal variation was set as 100, 200, or 350 in the approximation of $\varepsilon_{y,i}$. These sets of model settings were submitted to model selection and the one with the lowest AIC was selected at the best model that was adopted in this document.

(6). Best model

Based on AIC, we determined the model with gamma distribution (log link) and with 350 knots as the best model (Fig. 4), after confirming the convergence of the optimization using the *check_fit* function.

Figure 4. AIC values of the different model settings.



(7). Diagnostics of the model and the residuals

There apparently were no systematic biases in the spatio-temporal distribution of standardized residuals that were obtained using the R package 'DHARMa' (Hartig 2022) (Fig. 5).

The parameter estimates were stable as the final gradients of all parameters were nearly zero (< 0.01). The prediction of encounter probability was diagnosed by investigating 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 was 0.93 (Fig. 6), suggesting its good prediction. The Q-Q plots for the standardized residuals indicates that the distribution assumption is met (Fig. 7).

Figure 5. Spatio-temporal distribution of the residuals of the encounter probability $p_{y,i}$ (logit scale) and positive density $d_{y,i}$.

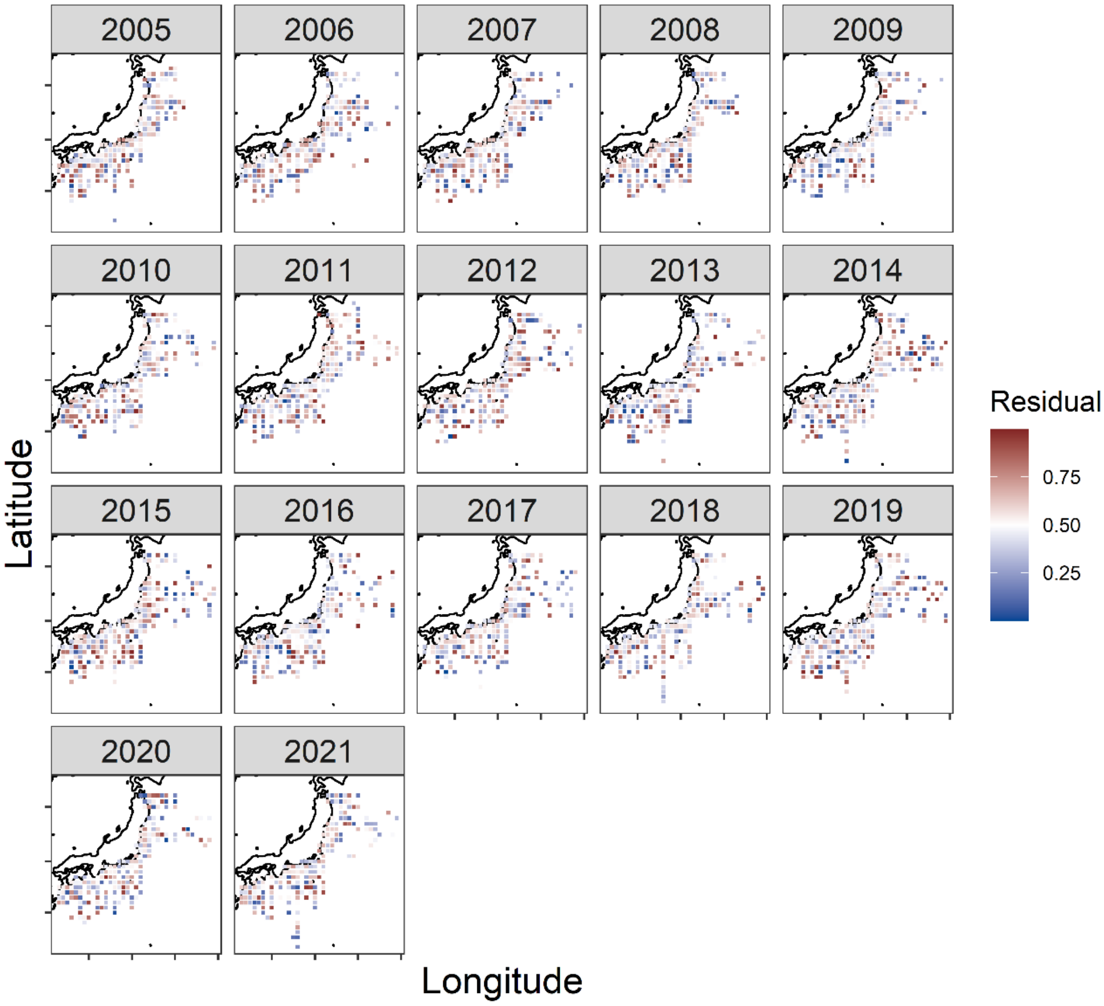


Figure 6. Predicted and observed probability of the encounter probability $p_{y,i}$. The red shared area represents the 95% CI of the prediction.

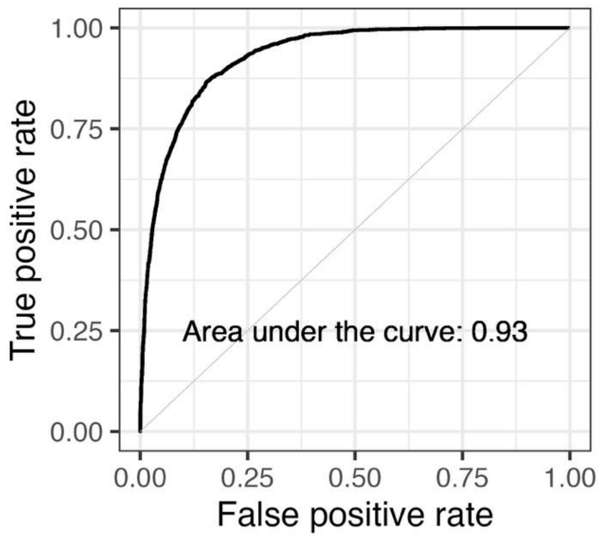
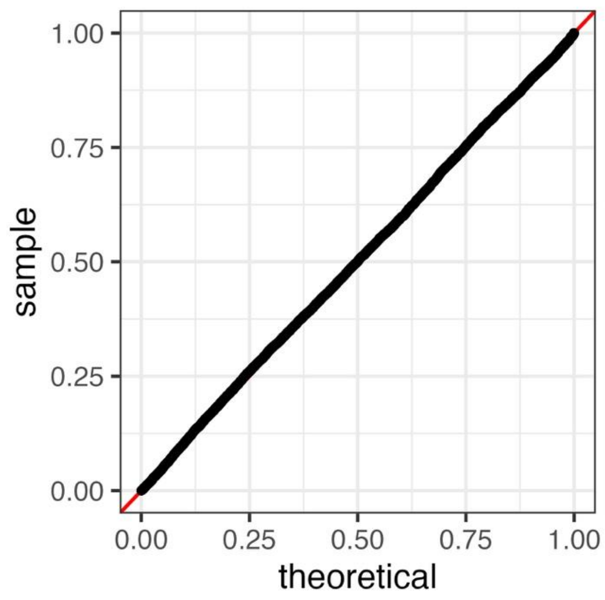


Figure 7. Quantile-quantile plotthat compares the distribution of the observation and prediction of egg density.



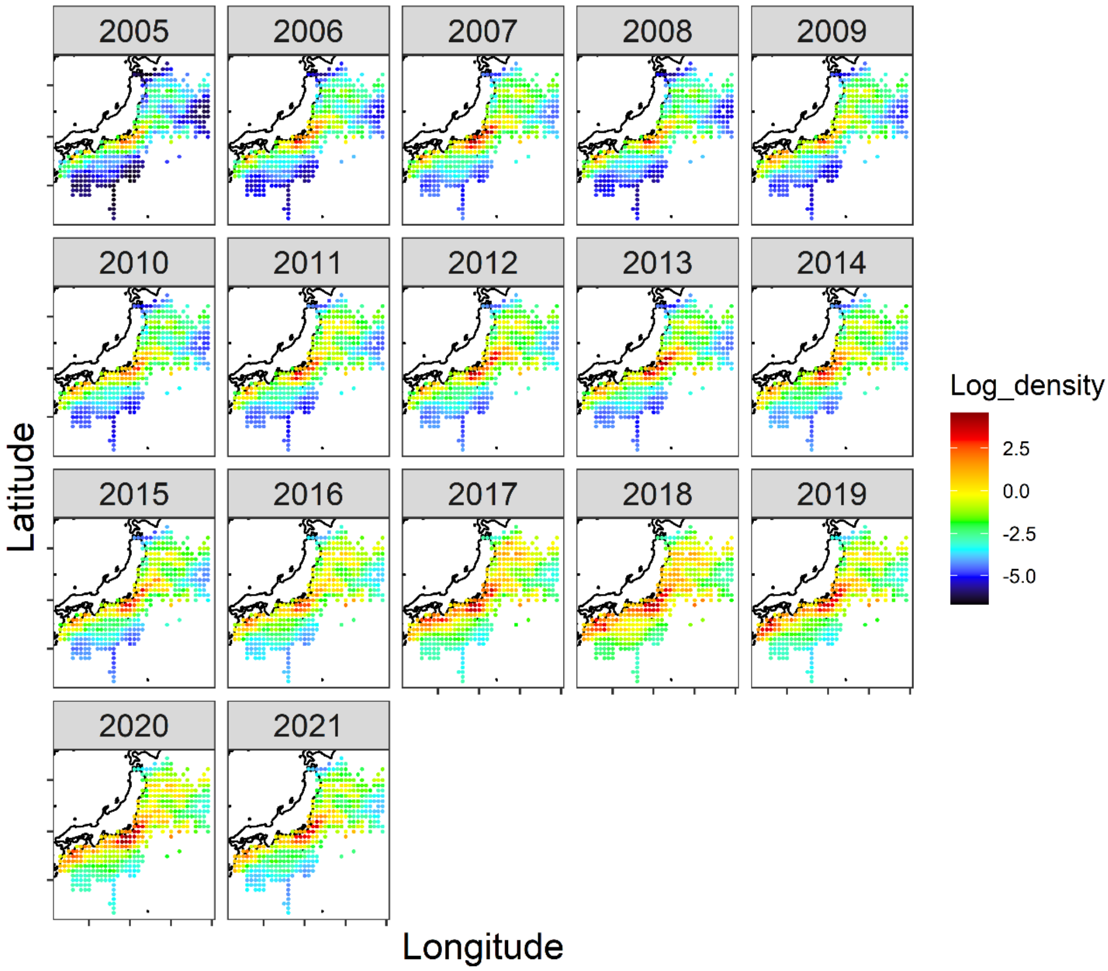
(8). Estimated relationships between the explanatory variables and the response variable

The explanatory variables in the model were only spatial and temporal ones and we did not incorporate other covariables. The spatial and temporal patterns of the response variable (predicted abundance) are shown in Fig. 8 in the next section.

(9). Yearly standardized CPUE and its uncertainty

We present the spatio-temporal distribution of the predicted egg density (Fig.8). The uncertainty of the model (95% CI) is shown in the next section.

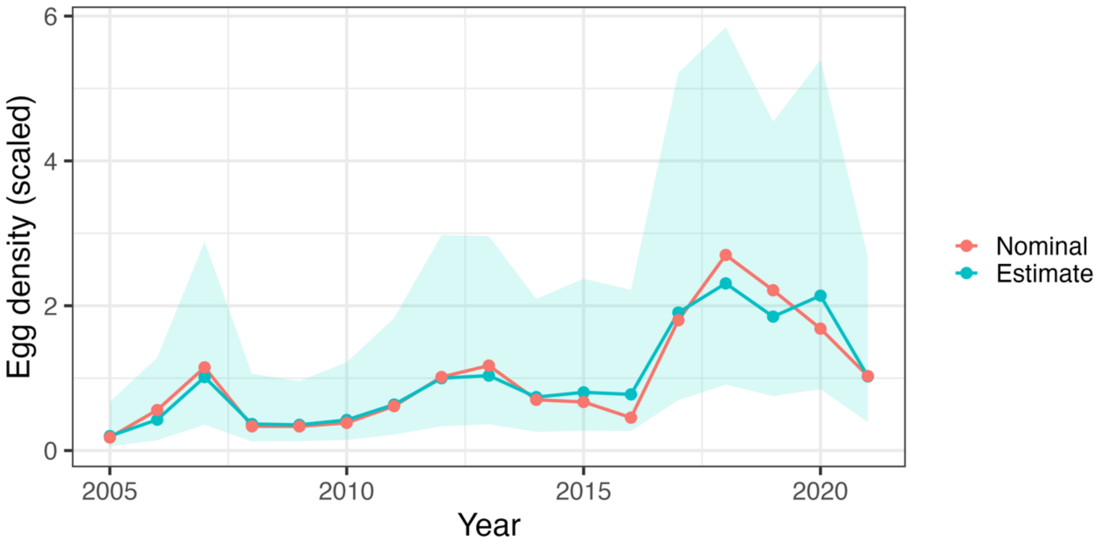
Figure 8. Spatio-temporal distribution of the predicted egg abundance.



(10). Comparison of the nominal and standardized CPUEs

The yearly patterns of index trends were similar between nominal and standardized CPUEs, and the inter-annual variability in standardized CPUE was smaller than nominal CPUE's one (Fig. 9). Both indices indicate that SSB has increased since 2016 but peaked in 2018 and has been decreasing recently.

Figure 9. The yearly patterns of scaled (divided by mean) nominal and standardized SSB indices. Blue area is 95% confidence interval of the standardized index.



Reference lists

Cadigan NG, Myers RA 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

Hartig F (2022)._HARMa: Residual Diagnostics for Hierarchical (Multi-Level / Mixed) Regression Models. R package version 0.4.5, <https://CRAN.R-project.org/package=DHARMa>.

Kanamori, Y., A. Takasuka, S. Nishijima, M. Ichinokawa, and H. Okamura. (2018) Standardizing abundance index for spawning stock biomass of chub mackerel in the Northwestern Pacific. NPFC-2018-TWG CMSA02-WP03.

Kanamori, Y., A. Takasuka, S. Nishijima, and H. Okamura. (2019) Climate change shifts the spawning ground northward and extends the spawning period of chub mackerel in the western North Pacific. Marine Ecology Progress Series 624:155–166.

Takasuka A, Oozeki Y, Kubota H (2008a) Multi-species regime shifts reflected in spawning temperature optima of small pelagic fish in the western North Pacific. Marine Ecology Progress Series 360:211-217

Takasuka A, Kubota H, Oozeki Y (2008b) Spawning overlap of anchovy and sardine in the western North Pacific. Marine Ecology Progress Series 366:231-244

Takasuka A, Tadokoro K, Okazaki Y, Ichikawa T, Sugisaki H, Kuroda H, Oozeki Y (2017) In situ filtering rate variability in egg and larval surveys off the Pacific coast of Japan: Do plankton nets clog or over-filter in the sea? Deep-Sea Res I 120: 132−137

Thorson JT (2019) Guidance for decisions using the Vector Autoregressive Spatio-Temporal (VAST) package in stock, ecosystem, habitat and climate assessments. Fish Res 210:143-161

Thorson JT, Pinsky ML, Ward EJ (2016) Model-based inference for estimating shifts in species distribution, area occupied and center of gravity. Methods Ecol Evol 7:990-1002.

Thorson JT, Shelton AO, Ward EJ, Skaug HJ (2015) Geostatistical delta-generalized linear mixed models improve precision for estimated abundance indices for West Coast groundfishes. ICES J Mar Sci J Cons 72:1297-1310

Thorson JT, Cunningham CJ, Jorgensen E, Havron A., Hulson PJF, Monnahan CC., 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