ISC/23/ALBWG-01/03

CPUE standardization for North Pacific albacore caught by Japanese longline fishery from 1996 to 2021 in Area 2 and Quarter 2

Jun Matsubayashi, Hirotaka Ijima, Naoto Matsubara, Yoshinori Aoki and Yuichi Tsuda

Highly Migratory Resources Division, Fisheries Resources Institute,
Japan Fisheries Research and Education Agency (FRA)

2-12-4, Fukuura, Kanazawa-ku, Yokohama,
Kanagawa 236-8648, JAPAN

Email: Matsubayashi_jun86@fra.go.jp

This working paper was submitted to the ISC Albacore Working Group Intercessional Workshop, 6-12 December 2022, held at the Fisheries Resources Institute, Japan Fisheries Research and Education Agency, Kanagawa, Japan and again at the This working paper was submitted to the ISC Albacore Working Group Stock Assessment Workshop, 20 – 27 March 2023 held at the NOAA/SWFSC, CA, USA.

Summary

This investigation undertook standardization of the CPUE for North Pacific Albacore in Area 2 during Quarter 2, utilizing operational data from the Japanese longline fishery and a geostatistical model. The CPUE of albacore for Area 2 during Quarter 2 can potentially serve as an indicator of the abundance of large, female individuals. This study encompassed spatiotemporal effects by utilizing the Stochastic Partial Differential Equations (SPDE) methodology to develop spatiotemporal models. The objective was to model the catch of Albacore with a zero-inflated negative binomial error distribution, incorporating the year effect, location effect, hooks per basket, fleet type, and vessel name. The model converged efficiently and did not appear to have significant issues, as evidenced by randomized quantile residual plot. The standardized CPUE estimates were relatively consistent with the nominal CPUE until 2007, but decreased compared to the nominal CPUE after 2008.

Introduction

When assessing the North Pacific albacore (*Thunnus alalunga*) stocks, standardized catch per unit effort (CPUE) indices derived from longline fishery catch data in Quarter 1 of Area 2 have been used as a primary input parameter for stock synthesis models (ISC 2020). This is because previous studies have found that Area 2 consistently yields larger adult albacore, based on catch-at-length data, compared to the other four main albacore fishing areas in the North Pacific, regardless of the season (Ijima et al. 2017; Ochi et al. 2016) (refer to Figure 1). It is noteworthy that larger female albacore are often caught in Quarter 2 of Area 2. As such, the CPUE of Quarter 2 of Area 2 may serve as a useful indicator in stock assessment, as it is believed that larger females make the greatest contribution to juvenile recruitment. Therefore, this study focused on standardizing the CPUE of Quarter 2 of Area 2 using a geostatistical model.

The standardized CPUE estimates were previously obtained using generalized linear mixed models (GLMMs), which incorporated temporal and spatial variables as random effects (Ochi et al. 2017; Fujioka et al. 2019). However, this method did not account for the possibility of increased correlations between catch data when collected in closer proximity in both time and space. To address this limitation, geostatistical models utilizing the Stochastic Partial Differential Equations (SPDE) approach, which account for spatial autocorrelation, have been recently utilized in fishery resource management (Ijima and Koike 2021). Thus, in our analysis, we utilized a geostatistical model INLA (Rue et al. 2009) to present the annual trends of the standardized CPUE as potential input data for stock assessment models.

Data and Methods

Longline logbook data

The longline operation dataset includes records of the number of albacores caught, the date and quarter of the catch, the type of fleet location (classified as Distant, Offshore, or Coastal, referred to as "fleet" hereon), the number of hooks per basket (hpb), the total number of hooks, and the vessel ID from 1976 to 2021. Latitude and longitude were recorded in in units of one degree, and data from the same year, month, vessel, hpb, and latitude and longitude were summed to reduce the volume of data. To focus specifically on albacore fishing, we initially excluded data from the Distant fleet and data with fewer than 10 hpb. Previous research has found that Area 2 (see Fig. 1) consistently has larger, adult albacores based on catch-at-length data, regardless of season, among the five main albacore fishing areas in the North Pacific (Ijima et al. 2017, Ochi et al. 2016). Additionally, the method of collecting logbook data for Japanese longline vessels changed in 1994, and data using the new collection method are believed to be stable and reliable since approximately 1996 (Ijima et al. 2017, Ochi et al. 2017, Fujioka et al. 2019, ISC 2019).

Generation of INLA mesh

First, we transformed the coordinates of each data point from latitude and longitude to meters to ensure that the distance between data points was accurately reflected in the analysis. To model the data with INLA, it is necessary to generate a mesh that creates an artificial set of neighborhoods within the study area and calculates the spatial autocorrelation between data points. In creating the mesh, we must specify the values of max.edge and cutoff; max.edge determines the maximum allowable length of triangles in the mesh, while the cutoff value determines the minimum allowable distance between points. Lowering the value of max.edge increases the resolution of the mesh, but also increases computation time. Similarly, smaller cutoff values result in higher-resolution meshes, but also increase computation time. In this study, we set the max.edge value to 500 and the cutoff value to 170 (Fig. 2).

Generation of SPDE models

Geostatistical model treats the effect of location through the SPDE approach, and thus we refer to this model as *SPDE* model. The structure of *SPDE* model with all explanatory variables (full model) is as follows:

$$CPUE_{alb} \sim intercept + year + f(fleet, model = iid) + f(hpb, model = iid) + f(vessel\ ID, model = iid) + f(w, model = AR1) + offset(hooks/1000)$$

where w is the spatial random effect calculated based on the SPDE approach and $AR1$ represents autoregressive model. Thus, the model estimates multiple spatial random fields that are autoregressive by year. As the spatio-temporal model required a significant amount of computation time, we made the assumption of a zero-inflated negative binomial distribution for the error distribution of the objective variable. Furthermore, we constructed models that included all explanatory variables.

Results and Discussion

Validation of the best model

The plot of the Matérn correlation function, which defines the correlation between locations, versus distance, suggested that we have strong spatial correlation up to approximately 300km, and the distance over which the correlation between points decreases to 10% was approximately 880 km (Fig. 3). Therefore, the max.edge value we set for triangularization was within the range of distances where correlations between sites were sufficiently high.

A plot of randomized quantile residuals suggested that the residuals generally followed a normal distribution (Fig. 4), but for points where the residuals were high, there may be some systematic error as they do not clearly follow a normal distribution. To investigate the causes of these systematic errors, we examined the relationships between randomized quantile residuals and year, location, hpb, and fleet, but did not find any clear differences along these variables. The only residual plot by fishing vessel ID showed residuals that clearly deviated from the normal distribution for certain vessels. Even for the same fishing vessel, the residuals sometimes followed a normal distribution and sometimes did not, and it is difficult to identify the cause of these systematic errors. In future stock assessments, more accurate standardized CPUE could be achieved by introducing a process to extract and exclude such problematic data. We did not encounter any other problems with the latent random field or the posterior distribution of parameters in the best model (Fig. 5, 6).

Estimation of standardized CPUE

The trends in annual changes in standardized CPUE estimated based on geostatistical model were generally consistent with the trends in nominal CPUE until

2007, although standardized CPUE tended to be lower than nominal CPUE after 2008 (Fig. 7).

Reference

Fujioka, K., Ochi, D., Ijima, H., Kiyofuji, H. (2019) Update standardized CPUE for North Pacific albacore caught by the Japanese longline data from 1976 to 2018. ISC/19/ALBWG-02/01.

Ijima, H., Ochi, D. and Kiyofuji, H. (2017) Estimation for Japanese catch at length data of North Pacific albacore tuna (*Thunnus alalunga*). ISC/17/ALBWG/04.

Ijima, H. and Koike, H. 2021. CPUE standardization for Striped Marlin (*Kajikia audax*) using spatio-temporal model using INLA. ISC/21/BILLWG-03/01.

ISC (2020) Stock assessment of albacore tuna in the north Pacific Ocean in 2020. Report of the albacore working group, 15–20 July 2020, online.

Ochi, D., Ijima, H., Kinoshita, J. and Kiyofuji, H. (2016) New fisheries definition from Japanese longline North Pacific albacore size data. ISC/16/ALBWG-02/03.

Ochi, D., Ijima, H. and Kiyofuji, H. (2017) Abundance indices of albacore caught by Japanese longline vessels in the North Pacific during 1976-2015. ISC/17/ALBWG/01.

Rue H., Martino S., Chopin N. (2009) Approximate Bayesian inference for latent Gaussian models by using integrated nested Laplace approximations. Journal of the Royal Statistical Society: Series B, 71, 319–392.

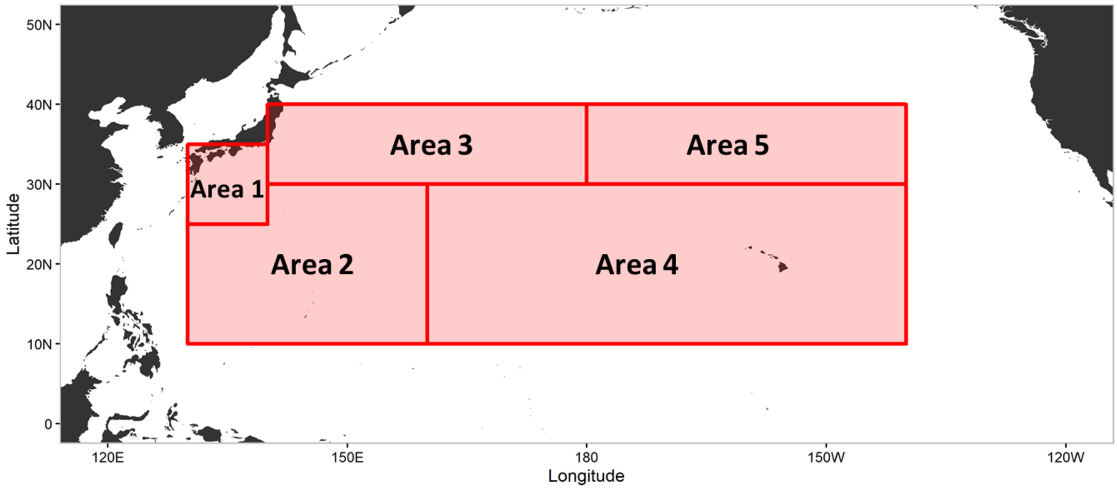

Figure 1. Area definition of Japanese longline fishery for albacore.

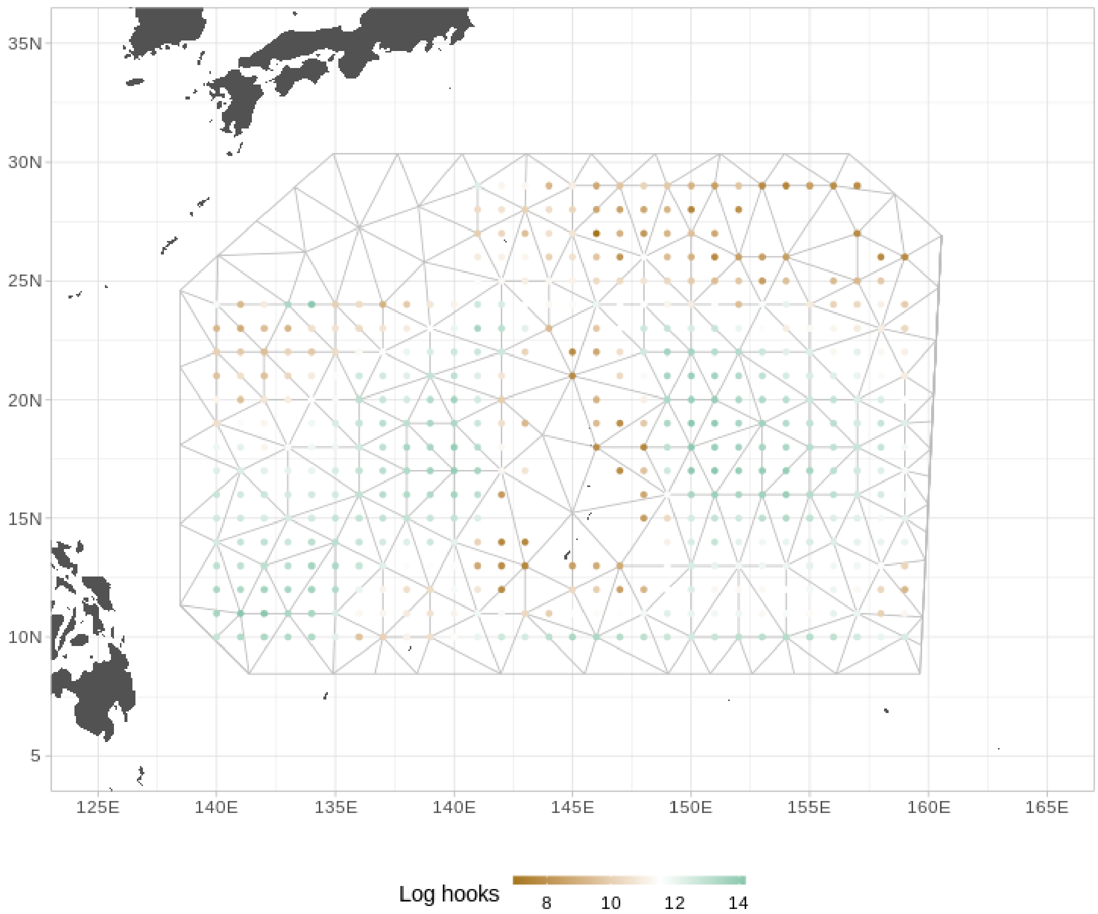

Figure 2. Triangularization of Japanese longline data and points are data locations. The color of each point indicates the number of hooks.

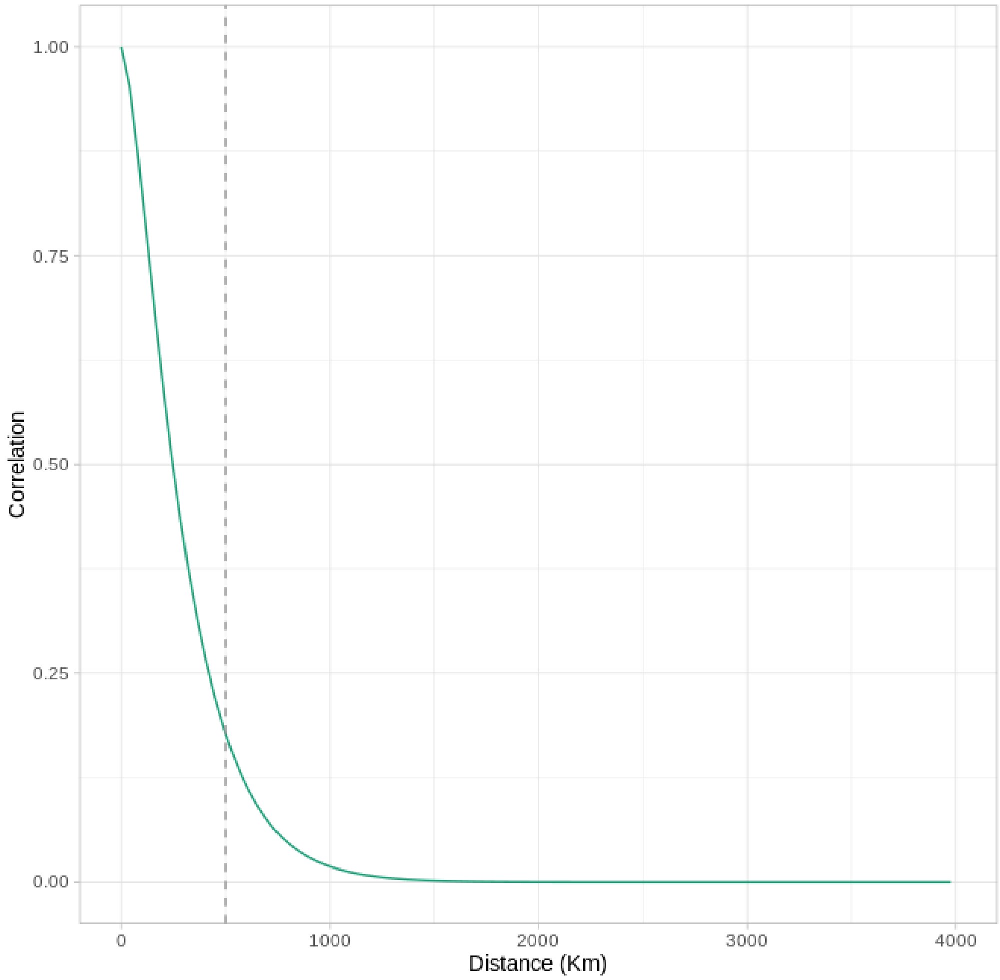

Figure 3. Matern correlation function versus distance. The dashed line indicates the max.edge value we set for triangularization.

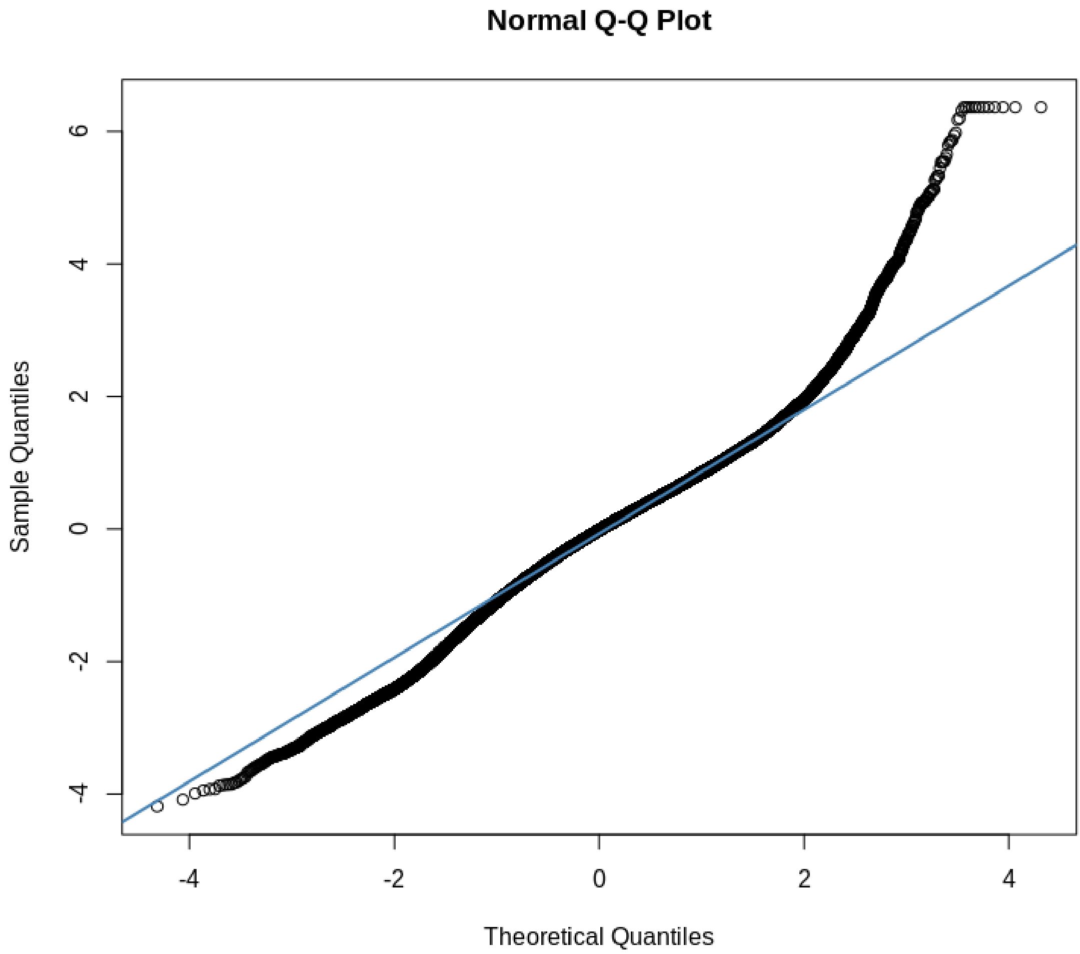

Figure 4. Normal Q-Q Plot of randomized quantile residuals.

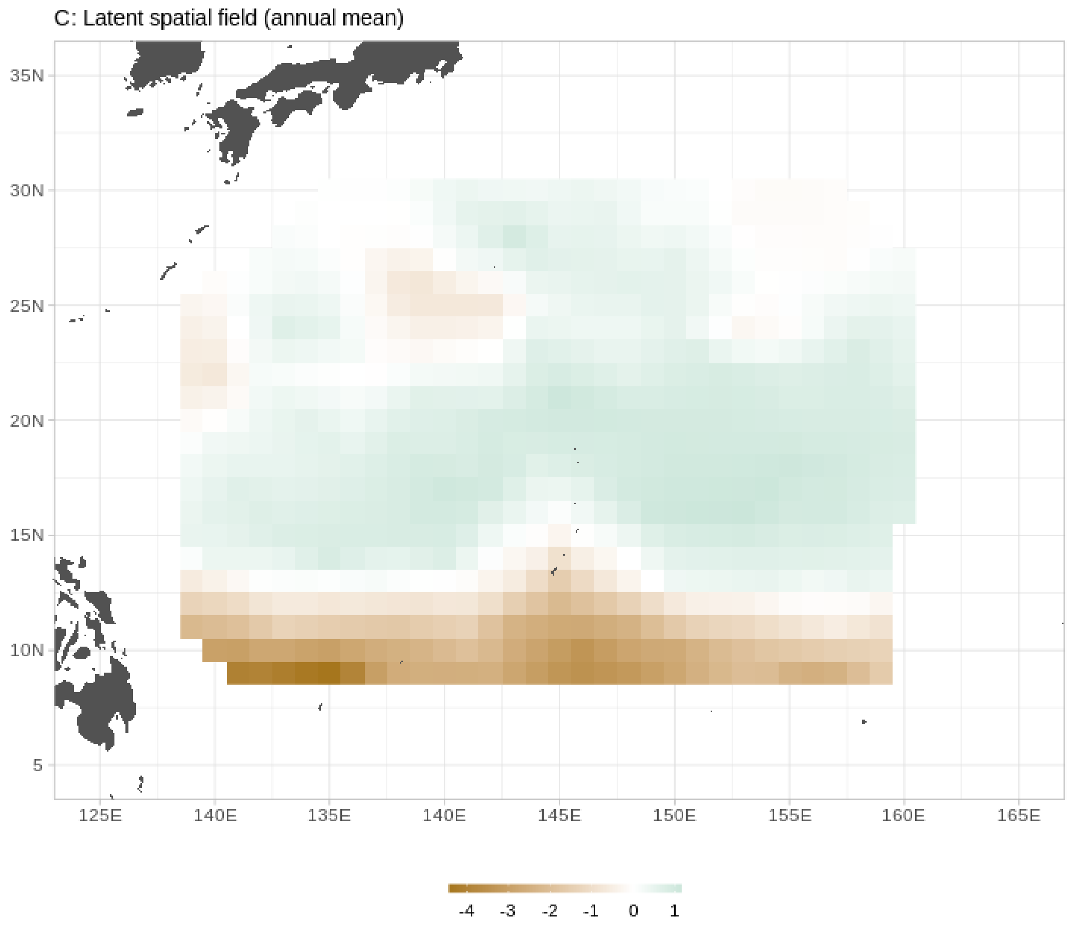

Figure 5. Interpolated spatial random field (annual mean values are shown).

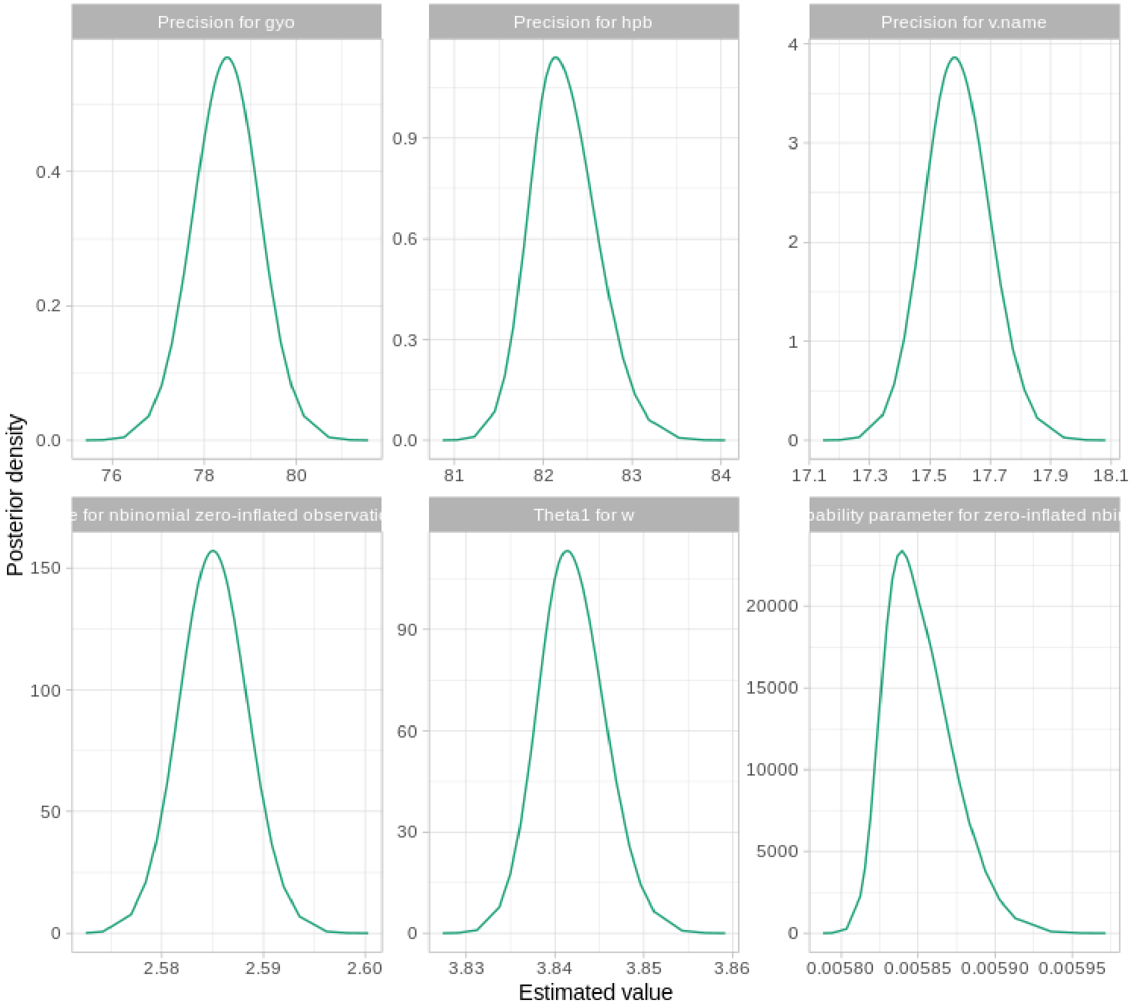

Figure 6. Posterior distributions of each parameter.

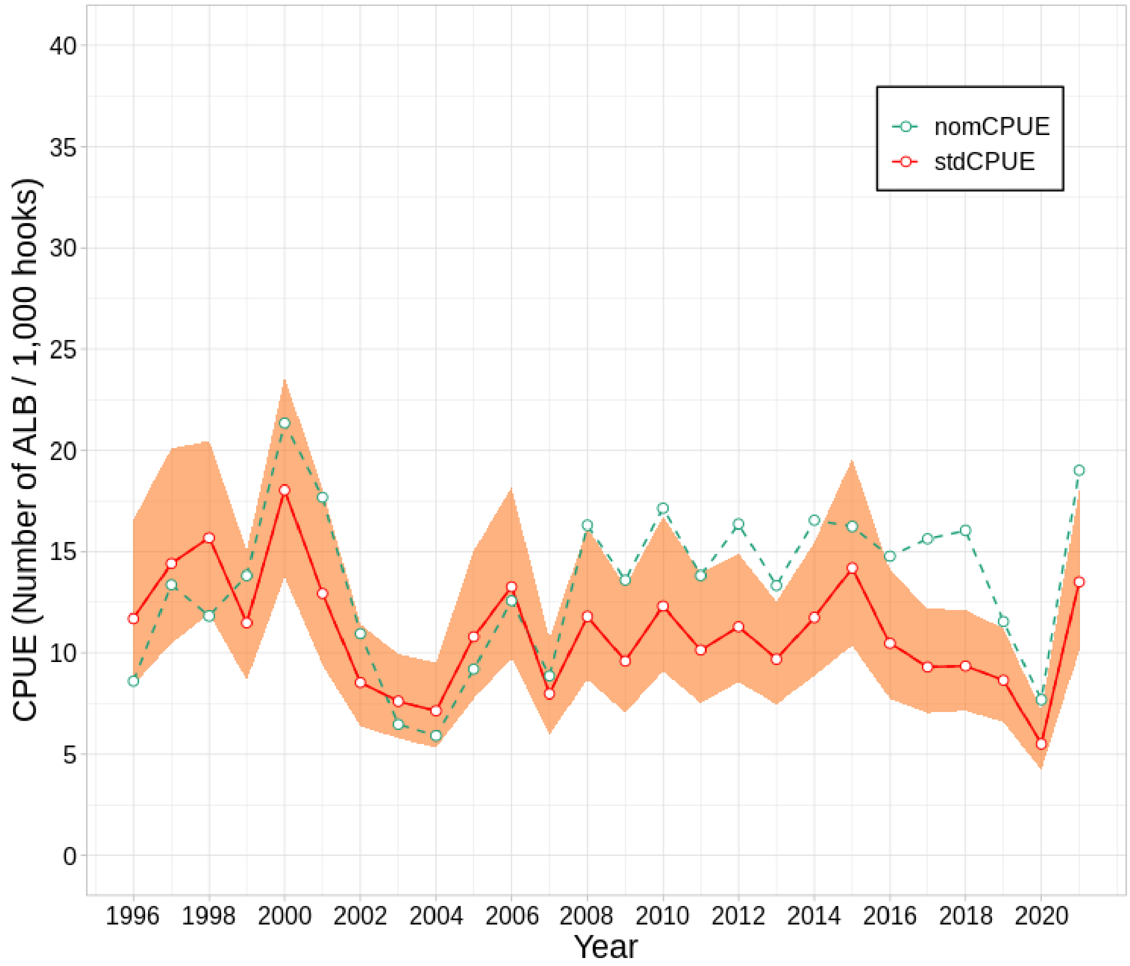

Figure 7. Plot of annual trends in nominal and standardized CPUE estimated by using SPDE model in this study. The red ranges indicate the 5% and 95% quantile intervals of the estimated standardized CPUE.