SC/24/ALBWG-01/04

Additional Japanese longline logbook data analysis on CPUE of adult North Pacific albacore tuna.[1]

Makoto Nishimoto, Hirotaka Ijima, Naoto Matsubara, and Yuichi Tsuda
Highly Migratory Resources Division, Fisheries Stock Assessment Center, Fisheries Resources Institute (FRI), Japan Fisheries Research and Education Agency. 2-12-4 Fukuura, Kanazawa-ku, Yokohama, Kanagawa, 236-8648, JAPAN
Nishimoto_makoto44@fra.go.jp

This working paper was submitted to the ISC Albacore Working Group Intercessional Workshop, 11-18 March 2024, held at the National Oceanic and Atmospheric Administration Southwest Fisheries Science Center, Victoria, Canada

Abstract

We report the Japanese longline CPUE standardization results conducted during the North Pacific albacore tuna stock assessment. At the request of the ISC Albacore Working Group (ALBWG), we addressed two analyses using the R software package R-INLA, which includes, 1) using data for all quarters in area 2 with the addition of data from the most recent year and the second quarter results, and 2) to show abundance trend with explicit uncertainty in estimation. The model using data from all quarters converged well, and the resampling program for the posterior distribution of the standardized CPUE was modified to work in realistic time by reducing the number of iterations. On the other hand, since the number of iterations is currently kept to the minimum necessary, we plan to address this by modifying the program in the future to improve the accuracy of the uncertainty assessment.

Introduction

The Japanese longline CPUE in area 2 has been modeled as an indicator of the abundance of adult female albacore (Matsubayashi et al. 2023, Ijima and Tsuda 2023). Matsubayashi et al. (2023) presented their results to the ISC Albacore Working Group (ALBWG) with the estimation results of Japan's standardized longline CPUE using data from area 2, quarter 2, using the Stochastic Partial Differential Equations (SPDE) approach, which can appropriately account for spatial autocorrelation for data. The ALBWG discussed the results and agreed to use the proposed spatio-temporal model in the next stock assessment. However, the ALBWG suggested using data from all time periods to reduce estimation uncertainty and increase information content, rather than only using data from quarter 2, for a more accurate stock assessment. We followed this requirement and performed an analysis during the stock assessment meeting that addressed all quarters and was able to show results consistent with previous resource assessments (Ijima and Tsuda 2023). As for the uncertainty in parameter estimation, however, it was not possible to show a population transition that explicitly included it. The reason for this was thought to be the possibility that the code program was stuck due to heavy computational load. Thus, we examined the problems in detail and established a policy to resolve them. This working paper reports the results of these additional analyses.

Material and methods

Japanese longline logbook data

We utilized logbook data from Japanese longline fishery operations. The logbook data was obtained from area 2, where believed to have the highest density of mature fish. The data was aggregated by year, month, number of hooks between floats (HBF), vessel name, and latitude and longitude at a resolution of 1 degree. These data were aggregated at this spatial resolution to reduce data volume. 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).

Generation of mesh for analysis

In order to model the data in INLA, one of the software packages that allow estimation with the SPDE approach, it is necessary to generate a mesh to calculate the spatial autocorrelation between data points. First, we converted each coordinate from latitude and longitude to meters so that the distance between data points would be accurately reflected in the analysis. Next, we set the max.edge value to 500, which determines the maximum allowable length of a triangle in the mesh, and the cutoff value to 170, which determines the minimum allowable distance between points (Fig. 2).

Statistical model

In the analysis, we used the same spatio-temporal statistical model as Ijima and Tsuda (2023). The categorical year variable was treated as a fixed effect, and the vessel name and the HBF as random effects. The number of albacore catches was assumed to follow a Zero-inflated negative binominal distribution, and 1,000 hooks were used as an offset term. The spatiotemporal statistical model is as follows:

$alb \sim intercept + year + f(fleet, model = iid) + f(hbf, model = iid) + f(vessel\ ID, model = iid) + f(w, model = AR1) + offset(hooks/1000)$.

Here, $f(w, model = AR1)$ is the spatiotemporal random effect consisting of the product of discrete time and space, where spatial processes and temporal autocorrelation are aggregated. We used the Auto-regression 1 (AR1) process that time step as a quarter to analyze time-series data of Area 2. We assumed a zero-inflated negative binomial distribution for the error distribution of the objective variable to deal with the many zero data in the catch record.

Standardized CPUE

We calculated the least-square means for standardized CPUE using all combinations of year or year-quarter and location explanatory variables; the R-INLA predictor function outputs the response variable for each explanatory variable, but only the mean and percentile of the posterior distribution, so to compute the variance of the least squares means, the posterior distribution, a resampling based on the posterior distribution must be performed again. In this analysis, 100 resampling were performed to compute the mean and variance of the quarterly standardized CPUE, with adjustments to fit within a realistic time frame.

Result and discussion

Estimation of Standardized CPUE

Standardized CPUE was calculated using data from all quarters, and the model converged to yield parameter estimates. Plots of the randomized quantile residuals suggested that the residuals generally followed a normal distribution, but deviations from the normal distribution were observed at locations with large residuals (Figure 3). In such areas, the existence of systematic errors or model misspecification is considered a possibility, and future work will be needed to determine the causes, such as the selectivity bias of the data and the selection of appropriate explanatory variables. Regarding the shape of the posterior distribution of the parameters, the results showed approximate normality and no major problems were observed (Figure 4). The results of the estimation with the zero-inflated negative binomial distribution provided the standardized CPUE that took into account the effects of latent spatial fields and temporal autocorrelation, and we extracted it at the quarter 2 period of interest to show time-series changes. The annual changes in standardized CPUE were generally consistent with the trends in nominal CPUE, but the values tended to be lower overall. This is similar to the results from the previous study (Ijima and Tsuda 2023). In addition, the standardized CPUE in 2021 was slightly lower than in the previous study, but this can be attributed to an increase in the amount of information due to additional updates of data for the two most recent years, 2022 and 2023. Further analysis along with data updates can be expected to improve the estimation accuracy.

Improvement of the resampling program and clarification of uncertainty

The previous study set the 5,000 resamples in the standardized CPUE calculation, but this program could not finish (Ijima and Tsuda 2023). We changed to a 100-resampling setup, the minimum required to avoid the computational burden of CPUE resampling, and were able to obtain resampled estimates of the standardized CPUE for all parameter space-time combinations. This made it possible to show time-series changes in standardized CPUE in any given quarter, as shown here (Figure 6). However, since the estimation accuracy increases with the number of resamples, it will be necessary to improve the program in the future by adopting a prior distribution with higher computational efficiency or by resampling after model selection. In addition, even after accounting for uncertainty, its 95th percentile quartile estimation range is narrow, and the CPUE for 2020 still shows a steep decline. This issue also needs to be investigated.

References

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 Tsuda, Y. (2023) Additional Japanese longline logbook data analysis for adult albacore tuna CPUE). ISC/23/ALBWG/08

Matsubayashi, J., Ijima, H., Matsubara, N., Aoki, Y., and Tsuda, Y. (2023) CPUE standardization for North Pacific albacore caught by Japanese longline fishery from 1996 to 2021 in Area 2 and Quarter 2. ISC/22/ALBWG-02/08

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.

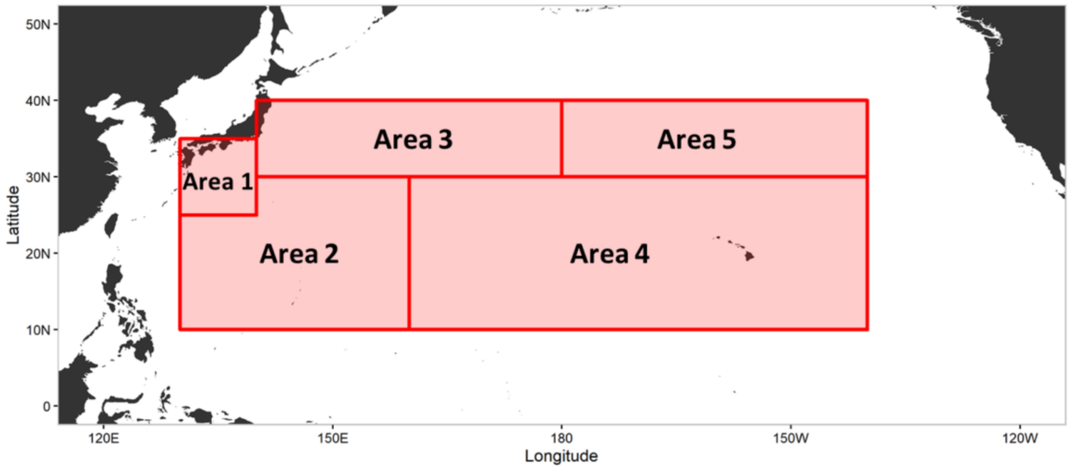

Figure 1. Area definition of Japanese longline fishery for albacore.

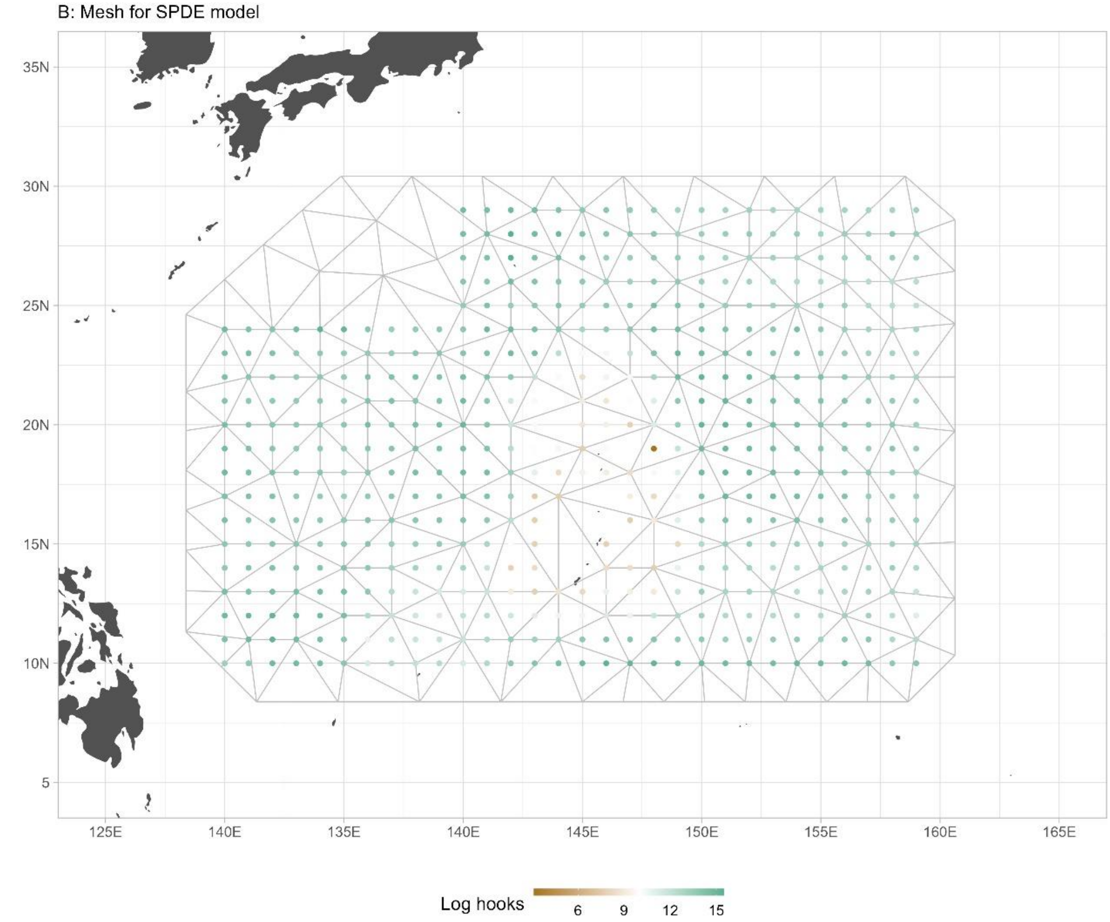

Figure 2. Mesh for SPDE model and Triangulation of Japanese longline data. The color of each point indicates the log-scaled number of hooks (annual mean values).

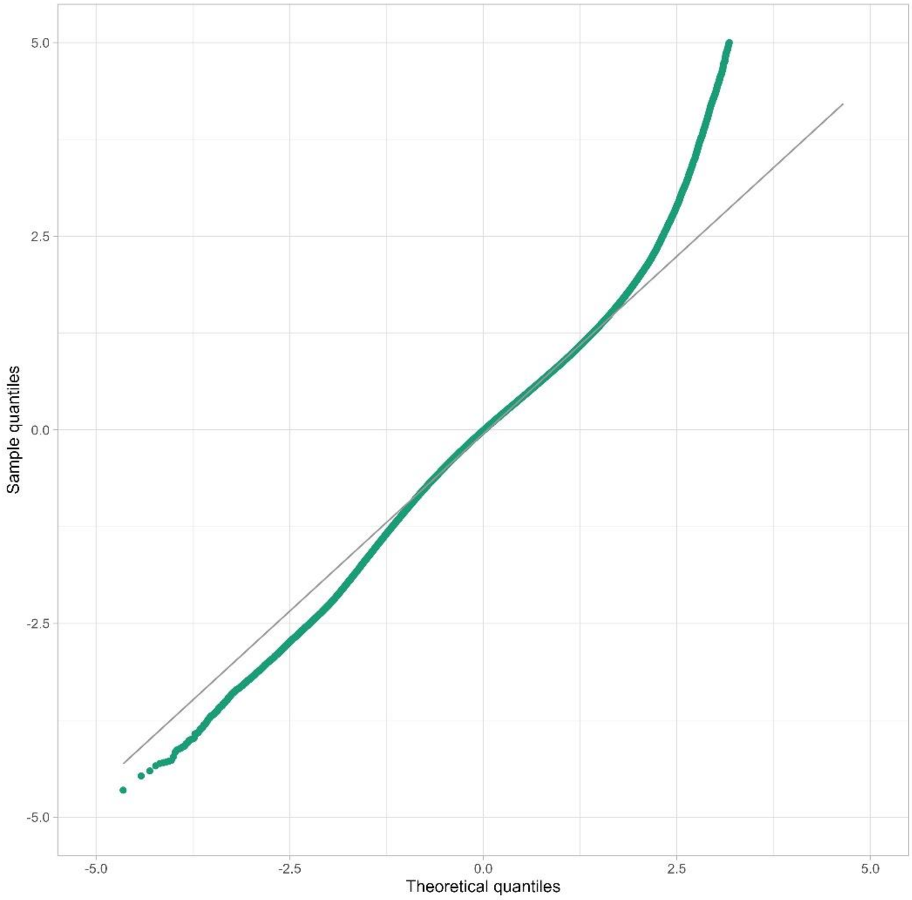

Figure 3. Normal Q-Q Plot of randomized quantile residuals.

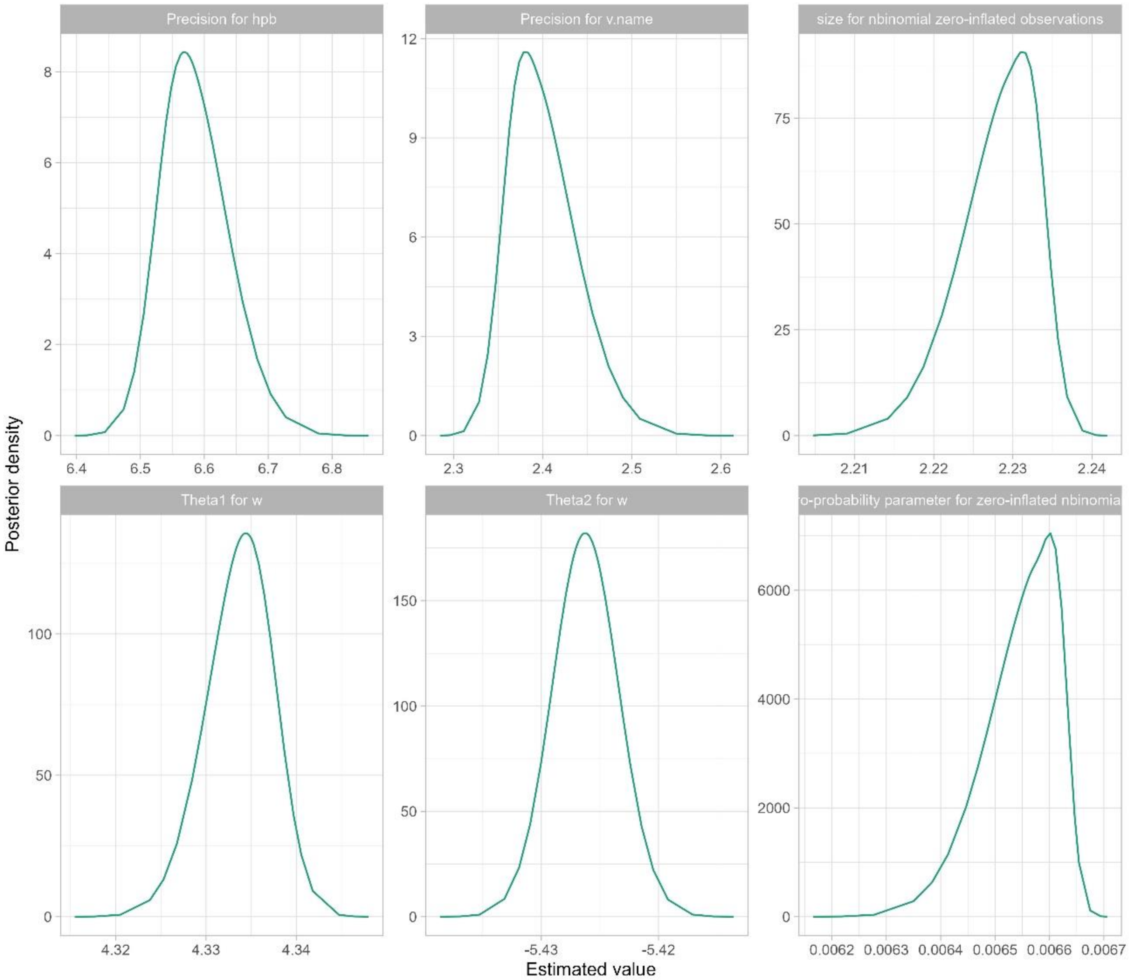

Figure 4. Posterior distributions of each parameter.

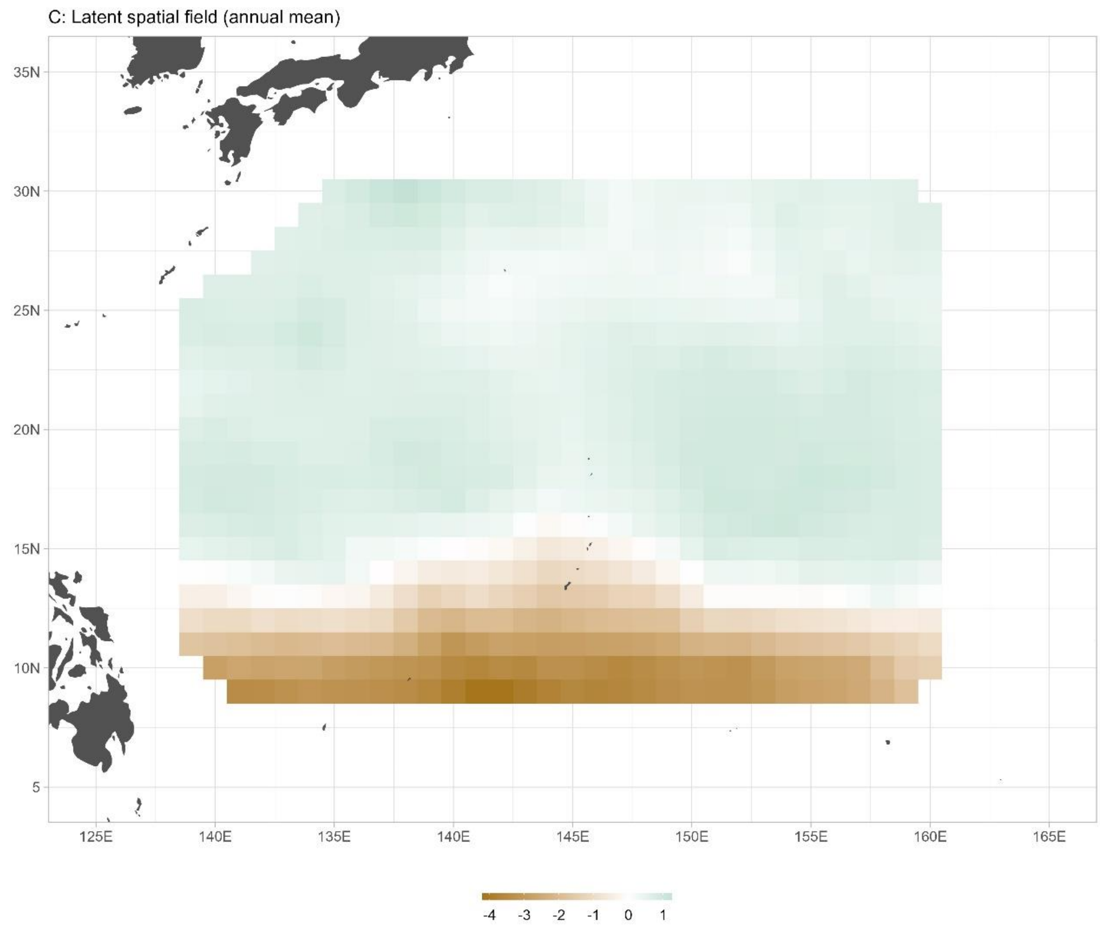

Figure 5. Interpolated spatial random filed (annual mean values).

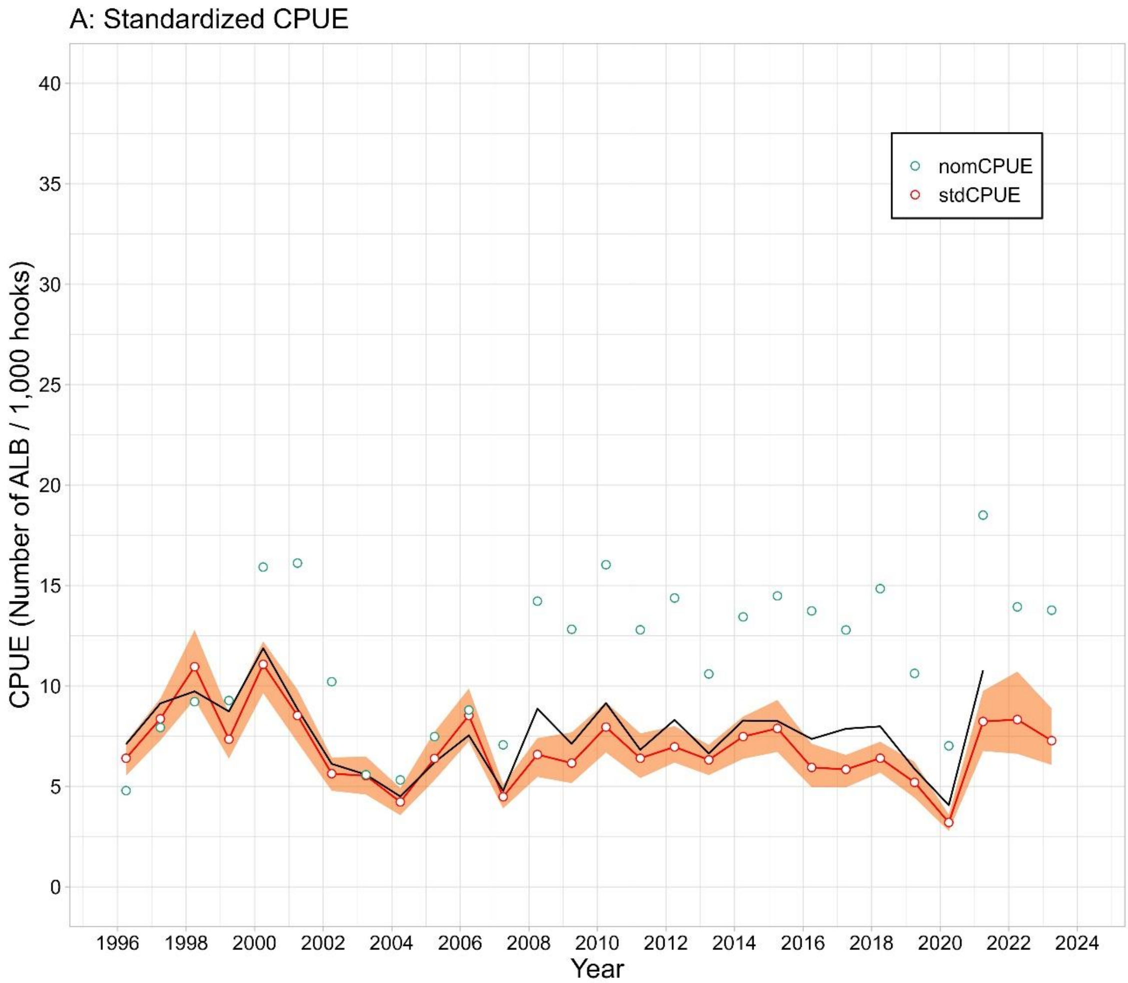

Figure 7. Annual trends in nominal CPUE (green points) and standardized CPUE estimated by using SPDE model (red line) in quarter 2 (Q2). The red ranges indicate the 95% quantile intervals of the estimated standardized CPUE. In this document, as in Matsubayashi et al. (2023), we are interested in time-series changes in Q2, so we show the interannual changes in Q2. The black line in the figure is the estimate from the previous evaluation, depicted for comparison with the update after the addition of the current data.