SCIENTIFIC COMMITTEE
TWENTY-FIRST REGULAR SESSION

Tonga
13-21 August 2025

CPUE standardization using sdmTMB for skipjack tuna stock assessment

WCPFC-SC21-2025/SA-IP-05
14 July, 2025

Makoto, Nishimoto[1], Yoshinori Aoki[1], Naoto Matsubara[1], Yuichi Tsuda[1], Hidetada Kiyofuji[1], Thomas Teears[2], Nan Yao[2], Paul Hamer[2]

[1] National Research and Development Agency, Japan Fisheries Research and Education Agency (FRA), Yokohama. Japan
[2] Oceanic Fisheries Program, The Pacific Community (SPC), B.P. D5, 98848 Nouméa, New Caledonia

Executive summary

In previous skipjack tuna stock assessments, the VAST model has been used to standardize Japanese pole-and-line (JPPL) CPUE. However, it has not been possible to easily reproduce the analysis results due to reasons such as compatibility issues with dependent packages accompanying updates to the software version. This was highlighted in the previous SC18 assessment (Castillo Jordan et al. 2022). The purpose of this document is to consider the use of sdmTMB as an alternative, reproducible analytical method and presents an updated standardized JPPL CPUE. We constructed and applied a spatiotemporal statistical model incorporating spatial autocorrelation using sdmTMB to JPPL logbook data from 1972 to 2023. We constructed a statistical model at the yearly level and included the month as a fixed effect, because the year-quarter level model did not converge due to the excessive number of estimated parameters. As a result, the model successfully met the convergence criteria for the gradient, and we obtained the annual and quarterly CPUE index by converting to the monthly-quarter level. The dynamics of the CPUE index estimated using sdmTMB model were consistent with the dynamics of past estimates using VAST model and also captured the quarterly trend. This result demonstrates the usefulness of sdmTMB as an alternative method, and the CPUE standardization approach that we proposed using the month-quarter conversion is practical, even in situations with a large number of parameters.

1. Introduction

Current stock assessments of skipjack tuna primarily rely on CPUE indices based on data from pole-and-line fisheries. In previous assessments, the VAST model has been employed to standardize CPUE from the Japanese pole-and-line (JPPL) fishery. However, at SC18, it was pointed out that reproducing past analyses was difficult due to software version updates and compatibility issues with dependent packages. To ensure robust assessments, it is imperative to establish standardized procedures based on reproducible methodologies.

Recently, the R package sdmTMB (Anderson et al., 2022) has been developed, offering flexible modeling with options to incorporate temporal and spatial autocorrelation, identical to VAST. The package has garnered attention in fisheries science, having been applied in international stock assessments for WCPO yellowfin and bigeye tuna, where it yielded results consistent with those from VAST (Day et al., 2023; Magnusson et al., 2023). A notable advantage of sdmTMB lies in its simplicity and transparency, which facilitates reproducibility. Therefore, sdmTMB presents a promising alternative for standardizing JPPL CPUE, particularly in response to reproducibility concerns raised in previous skipjack assessments.

This Information Paper explores the use of sdmTMB as a reproducible alternative and presents updated standardized CPUE indices for JPPL. We developed a spatiotemporal statistical model incorporating spatial autocorrelation and applied it to JPPL logbook data from 1972 to 2023. The model was constructed at the annual level, with month effects included as fixed effects. Estimated monthly parameters were converted into quarterly indices to produce CPUE at the year-quarter level. We compared the resulting indices with those from previous assessments and evaluated the performance of this alternative approach. Finally, we summarized insights gained during the modeling process, highlighting future research needs.

2. Materials and Methods

2.1 Japanese Pole-and-Line Fishery Data

We used JPPL logbook data from 1972 to 2023, consistent with the eight-area spatial structure used in the 2022 skipjack tuna stock assessment (Castillo Jordan et al., 2022) **(Figure 1)**. This dataset was updated with three additional years (2021–2023). Data was not yet available for the full 2024 year. Although the spatial distribution of JPPL fishing grounds has been shrinking in recent years, JPPL data still covers a large area of the western and central Pacific Ocean (**Figure 2**). JPPL vessels are categorized as coastal or offshore based on size, with different fishing strategies employed by each type. Coastal operations typically occur from April to December in waters north of 30°N near Japan, while offshore operations are conducted year-round across broader areas in the WCPO (Teears et al., 2022).

Data screening largely followed previous protocols (Kinoshita et al., 2019), but we additionally excluded days with no skipjack catch but positive albacore catch, treating these as non-target operations. This reflects the operational reality of JPPL vessels, which shift targeting strategies between skipjack and albacore based on fishing efficiency and market conditions.

2.2 CPUE Standardization Using sdmTMB

CPUE standardization was conducted using a delta lognormal GLMM, consistent with prior assessments (Kiyofuji, 2016; Castillo Jordan et al., 2019; Teears et al., 2022). We initially attempted a year-quarter model similar to that used in VAST, but due to the combination of long time series data (over 50 years) and a wide spatial range, memory overload errors occurred due to the high computational load in sdmTMB

To reduce computational load while maintaining model structure comparable to VAST (Teears et al., 2022), we constructed a delta lognormal GLMM at the annual level. Spatial locations at which the effects were estimated (i.e. the knots; n=285) were uniformly distributed across the spatial domain (**Figure 2**) consistent with the previous VAST analysis.

The model used daily catch per logbook entry (kg) as the response variable and included year, month, fishing type (coastal/offshore), vessel GRT, number of poles as fixed effects, and SST as a smooth term with k = 5.. Vessel ID, spatial, and spatiotemporal random effects were included as random effects:

$p_i \sim Year + Month + Class + grt + npoles + s(sst, k = 5) + (1|VesselID) + \omega\,(x_i) + \phi(x_i, t_i),$

$c_i \sim Year + Month + Class + grt + npoles + s(sst, k = 5) + (1|VesselID) + \omega\,(x_i) + \phi(x_i, t_i),$

where p_i is the encounter probability and c_i is the positive CPUE. The spatial random effect $\omega\,(x_i)$ was the spatial random effect at knot x associated with the observer datarecord i and the spatiotemporal random effect $\phi(x_i, t_i)$was modeled as i.i.d. across Year and knot x (see mesh and node structure in **Figure 3).**

Using estimated parameters, we calculated CPUE indices for the southwest Pacific region. We converted month effects from February, May, August, and November into quarter effects to derive year-quarter CPUE indices. A sea surface temperature (SST) is also been used as a spatial filter. In the tested MP, the 18◦C threshold was used to exclude grid cells with environmentally unsuitable habitats from the CPUE prediction (Kiyofuji et al. 2019). The CPUE satandardization were conducted in R version 4.4.2 using the sdmTMB package version 0.6.0.

Additionally, we compared the estimated time-varying coefficient of variation (CV) for regions 1 to 8 in the global model and regional-level models using sdmTMB to refine the models and understand the information content of the data (Appendex 1). This CV is calculated using the log-normal equation characteristic of the SE predicted during CPUE standardization, specifically $sqrt(\exp(SE^2) - 1)$. In the regional level model, we used a year-quarterly level model and created a mesh with an equivalent number of knots in INLA for each of the subsets of regions 1&2, regions 3&4, and regions 7&8. Region 5&6 was excluded from the regional level modeling due to insufficient data. Additionally, to ensure stable calculations for the annual-quarterly level model, we used time series data from the past 30 years. Since the data was recently updated, the results presented in the regional-level model are based on data up to 2022.

3. Results

The annual-level model including month effects as fixed effects successfully converged, with gradients <0.001 and a positive definite Hessian, allowing estimation of standard errors.

We converted monthly effects into quarterly values to obtain CPUE indices at quarterly scales (**Figure 4**). From the plots of the centered mean and absolute scale, the dynamics of the sdmTMB-based CPUE closely resembled those from VAST, capturing quarterly trends (**Figure 5**).

Residual diagnostics for the model showed that residuals were largely normally distributed, indicating that the model fit the data (**Figure 6**). Spatial residuals indicated larger errors at data-sparse outer regions, suggesting potential under- or overestimation in these areas, likely due to recent contractions in fishing grounds. This highlights the need for caution when interpreting estimates in poorly sampled regions.

When comparing the estimated temporal variation CV results for regions 1 to 8 between the global model and regional-level models using sdmTMB, it was found that, overall, modeling regions 1 to 8 together resulted in a smaller CV than modeling each region separately (Appendix 1: **Figure A1**). Additionally, in the tropical regions of areas 5 to 8, there was a slight increase in CV from the 1970s to recent years.

4. Discussion

We explored a CPUE standardization approach using sdmTMB for skipjack stock assessment. The model converged and successfully reproduced trends from previous assessments, demonstrating its effectiveness. Models such as VAST and sdmTMB, which incorporate temporal and spatial autocorrelation, are valuable for generating reliable CPUE indices.

While both VAST and sdmTMB have their strengths and limitations, we selected sdmTMB due to reproducibility concerns in past assessments (Castillo Jordan et al., 2022). VAST offers a high degree of automation for mesh creation and parameter space configuration, which is convenient but can hinder troubleshooting. In contrast, sdmTMB requires manual configuration of mesh and parameter spaces but offers greater transparency and traceability. For accurate and reproducible assessments, such transparency is crucial.

The transparency of sdmTMB allowed us to investigate convergence issues. Although we could not fit a full year-quarter model due to memory constraints, we identified that introducing random effects at the year-quarter level over long time series caused the computational failure. While sdmTMB may handle some complexity, estimating dense spatiotemporal structures remains challenging. In the case of JPPL skipjack, the 35-year model supported year-quarter structures, but failed even after adjusting the mesh size for longer time series. Even when the model was divided into finer regional levels as a regional-level model, it did not work well for time series exceeding 50 years.

As an alternative, we proposed a model that fully utilizes time series data and converts monthly effects fixed at the annual level into quarterly CPUE indices. Our results show that sdmTMB is a practical alternative, and this month-to-quarter conversion approach enables standardized CPUE estimation even in complex models. In our analysis, when comparing global-level and regional-level CVs, we found that global models tend to have smaller CVs overall because they borrow information from neighboring regions in the spatiotemporal autocorrelation term. However, in some tropical regions, CVs were small despite limited data due to recent fisheries ground shrinking, suggesting that CVs may be underestimated, and that data supplementation has both positive and negative aspects. Going forward, we believe that exploring modeling techniques for implementing annual-quarterly models within realistic ranges and techniques for responding to recent spatial reductions in fishing grounds will lead to even better resource assessments.

5. References

Anderson, S. C., Ward, E. J., English, P. A., & Barnett, L. A. K. (2022). sdmTMB: An R package for fast, flexible, and user-friendly generalized linear mixed effects models with spatial and spatiotemporal random fields (p. 2022.03.24.485545). bioRxiv.

Castillo Jordan, C., Tears, T., Hampton, J., Davies, N., Day, J., Magnusson, A., Scott, R., Pilling, G., Scott, F., and Hamer, P. (2022). Stock assessment of skipjack tuna in the western and central Pacific Ocean: 2022. WCPFC-SC18-2022/SA-WP-01, Online, August 10-18, 2022.

Day, J., A. Magnusson, T. Teears, J. Hampton, N. Davies, C. Castillo Jordan, T. Peatman, R. Scott, J. Scutt Phillips, S. McKechnie, F. Scott, N. Yao, R. Natadra, G. Pilling, P. Williams, & P. Hamer. (2023). Stock assessment of bigeye tuna in the western and central Pacific Ocean. Working paper WCPFC-SC19-2023/SA-WP-05, Koror, Palau.

Kinoshita, J., Aoki, Y., Ducharme-Barth, N., and Kiyofuji, H. (2019). Standardized catch per unit effort (CPUE) of skipjack tuna of the Japanese pole-and-line fisheries in the WCPO from1972 to 2018. Technical Report WCPFC-SC15-2019/SA-WP-14, Pohnpei, Federated States of Micronesia.

Kiyofuji, H. and Ochi, D. (2016). Proposal of alternative spatial structure for skipjack stock assessment in the WCPO. Technical Report WCPFC-SC12-2016/SA-IP-09, Bali, Indonesia,3–11 August 2016

Kiyofuji, H., Aoki, Y., Kinoshita, J., Okamoto, S., Masujima, M., Matsumoto, T., Fujioka, K., Ogata, R., Nakao, T., Sugimoto, N., and Kitagawa, T. (2019). Northward migration dynamics of skipjack tuna (katsuwonus pelamis) associated with the lower thermal limit in the western pacific ocean. Progress in Oceanography, 175:55–67.

Magnusson, A., J. Day, T. Teears, J. Hampton, N. Davies, C. Castillo Jordan, T. Peatman, R. Scott, J. Scutt Phillips, S. McKechnie, F. Scott, N. Yao, R. Natadra, G. Pilling, P. Williams, P. Hamer. (2023). Stock assessment of yellowfin tuna in the Western and Central Pacific Ocean. Working paper WCPFC-SC19-2023/SA-WP-04, Koror, Palau.

Teears, T., Aoki, Y., Matsubura, N., Tsuda, Y., Castillo Jordon, C., Hampton, J., Schneiter, E., Scutt Phillips, J., Peatman, T., and Hamer, P. (2022). Background analyses and data inputsfor the 2022 skipjack tuna stock assessment in and central Pacific Ocean. Technical ReportWCPFC-SC18-2022/SA-IP-05.

6. Tables and Figures

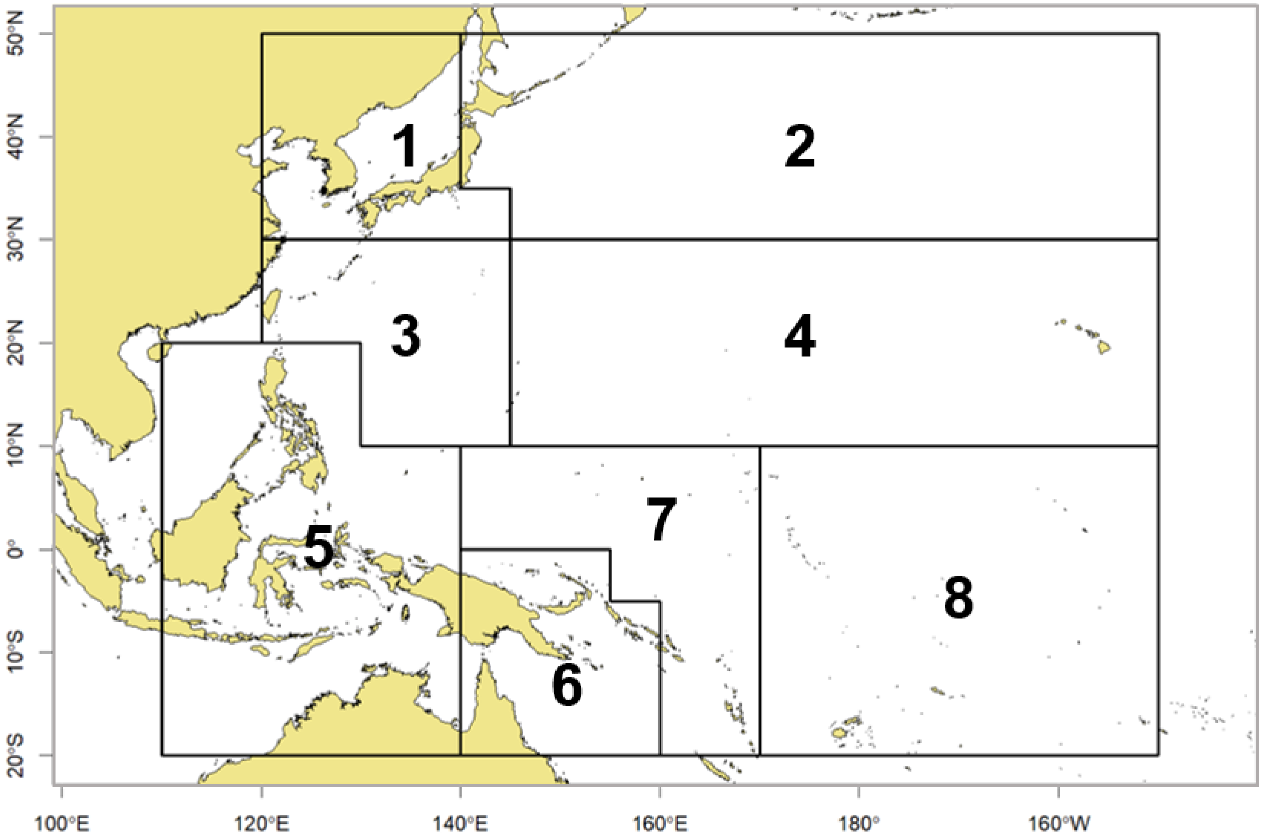


Figure 1. Geographical areas covered by the WCPO skipjack tuna stock assessment and boundaries of the eight model regions.

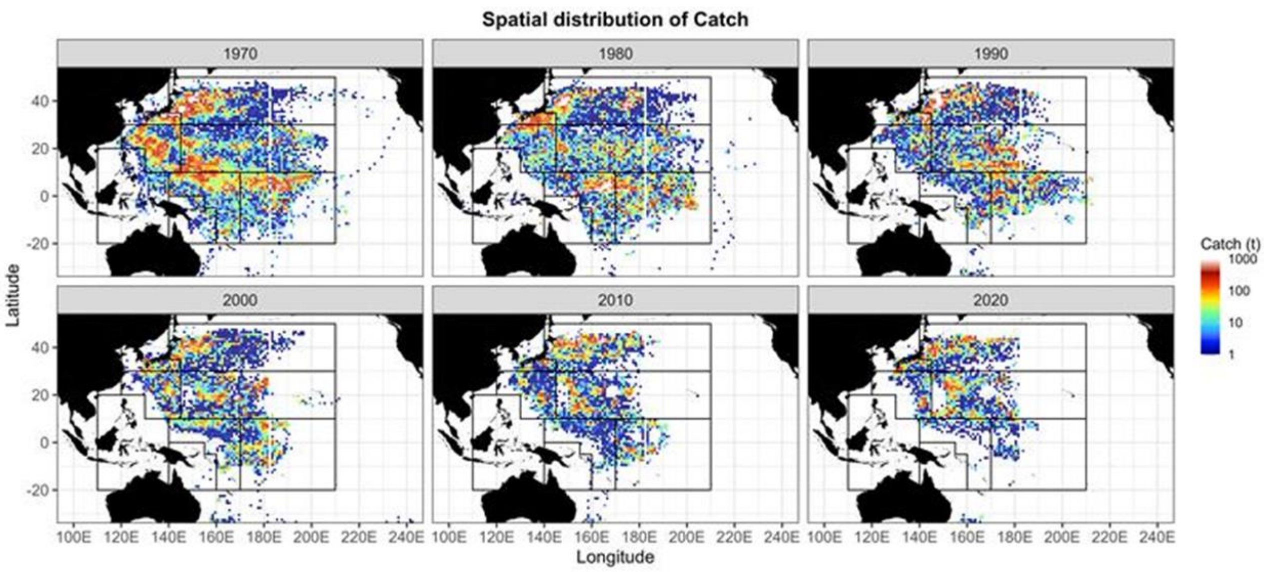


Figure 2. Decadal shifts in spatial distribution of Japanese pole-and-line (JPPL) fishery skipjack catch (metric tons) from 1972 to 2023. Each map shows the catch for each decade.

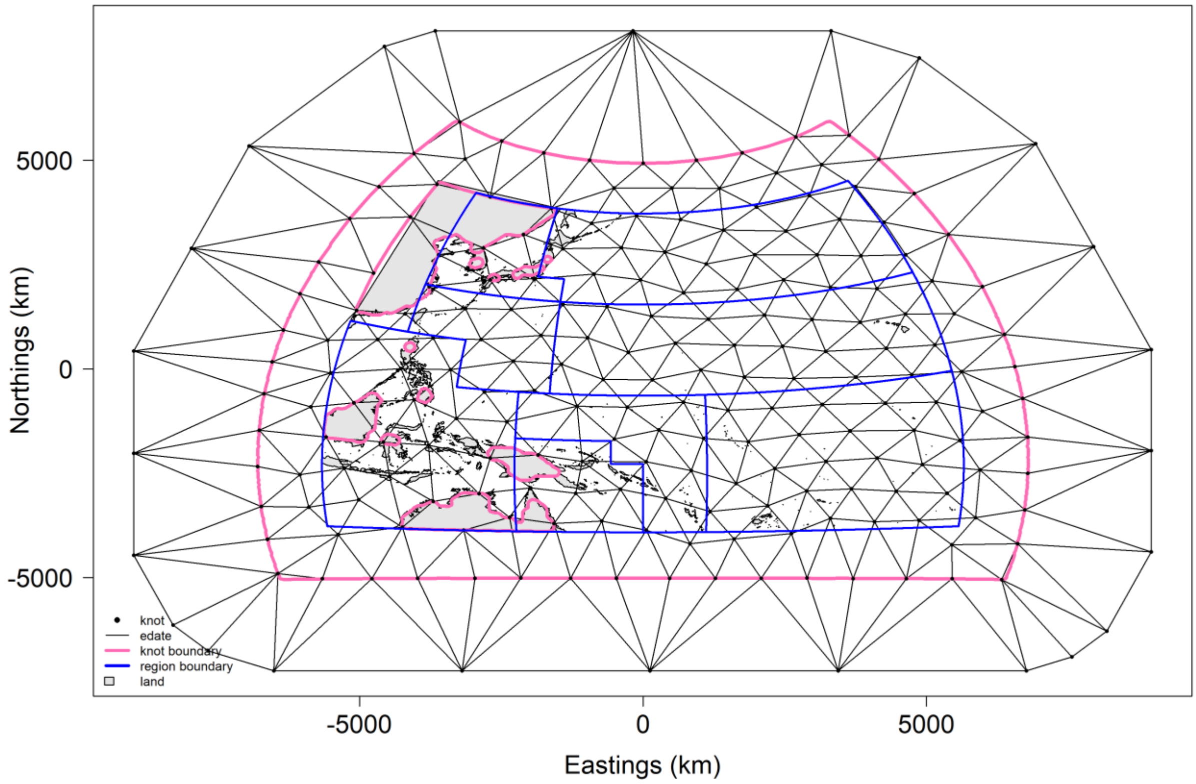


Figure 3. Spatial mesh structure used in the standardization of JPPL CPUE. In this mesh, each knot corresponds to a triangle vertex (i.e., a mesh node). The mesh was constructed using the INLA package (package ver. 24.06.27).

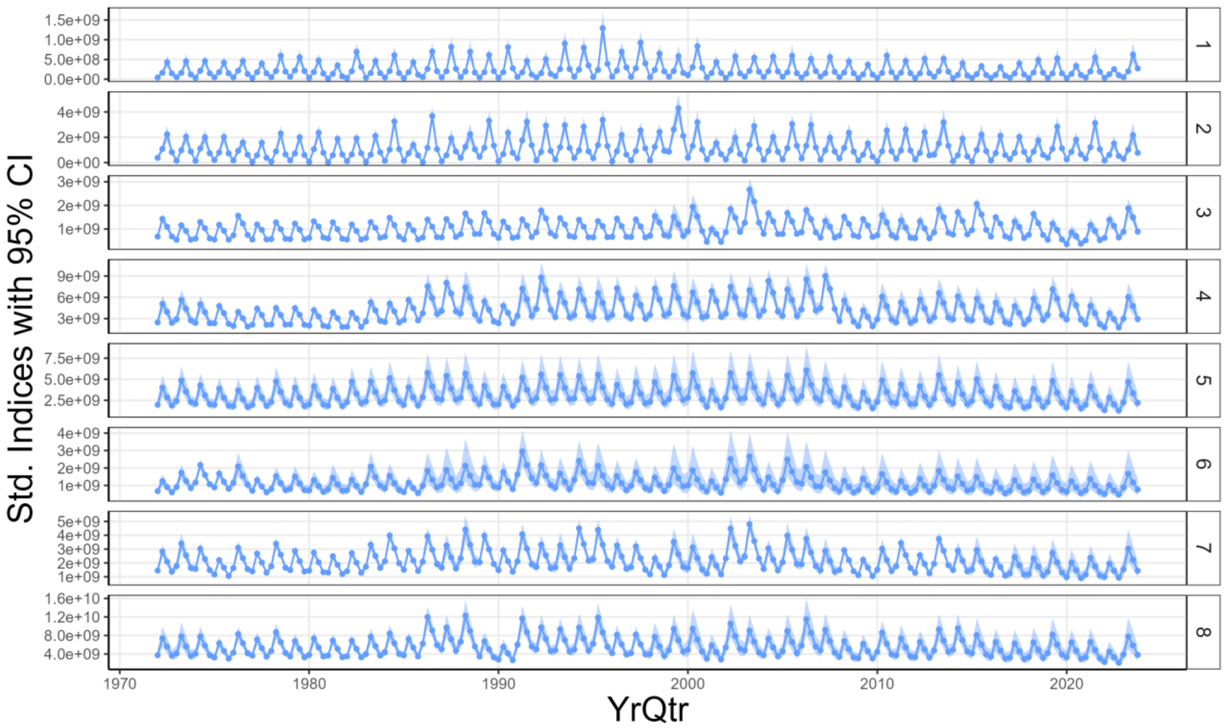
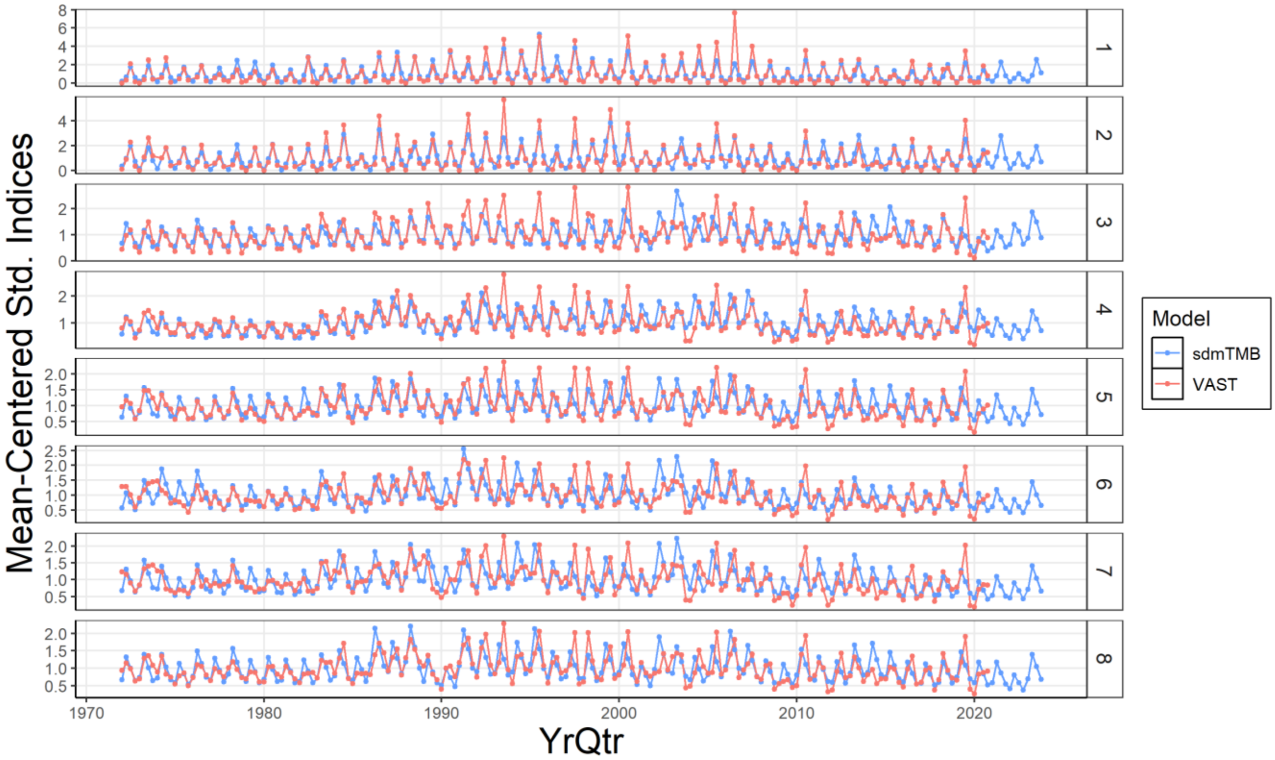


Figure 4. CPUE index estimation results in the sdmTMB models: Area-weighted CPUE standardized abundance indices by year and quarter in regions 1-8. The light blue shaded areas indicate the 95% confidence interval.

(a)



(b)

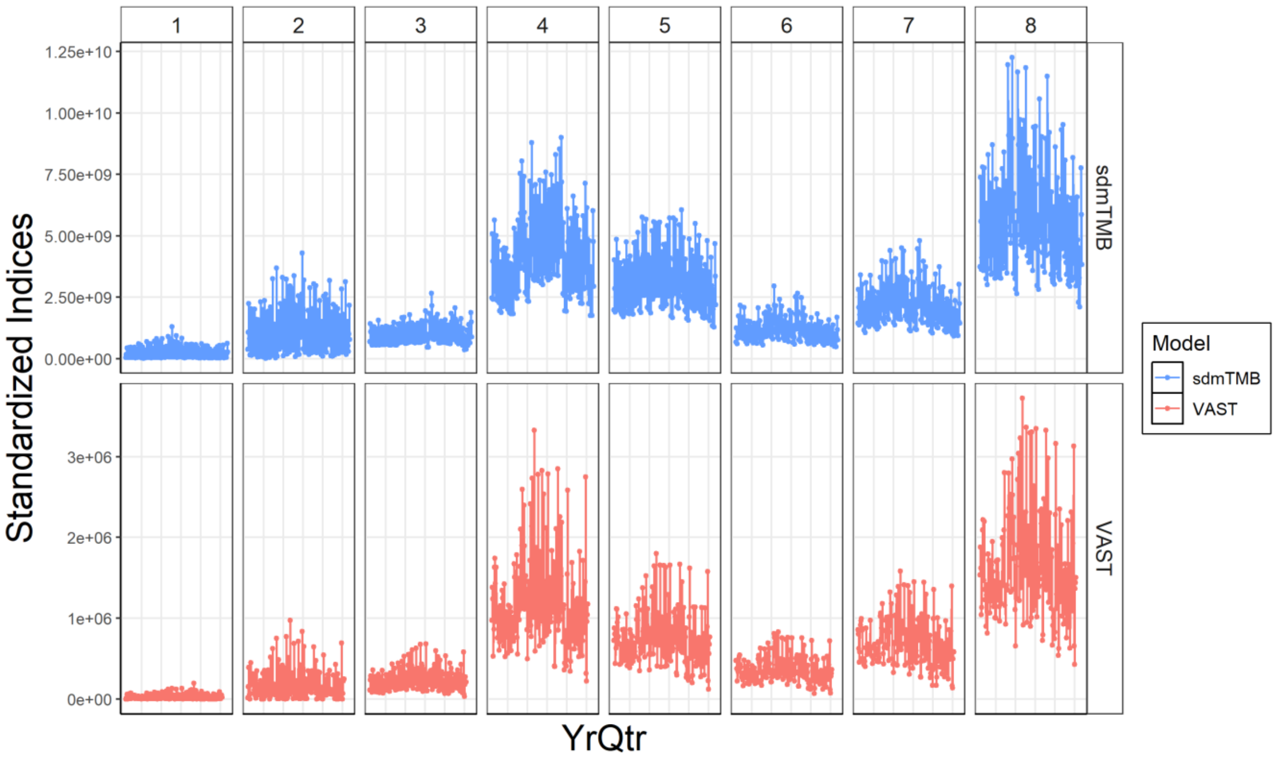
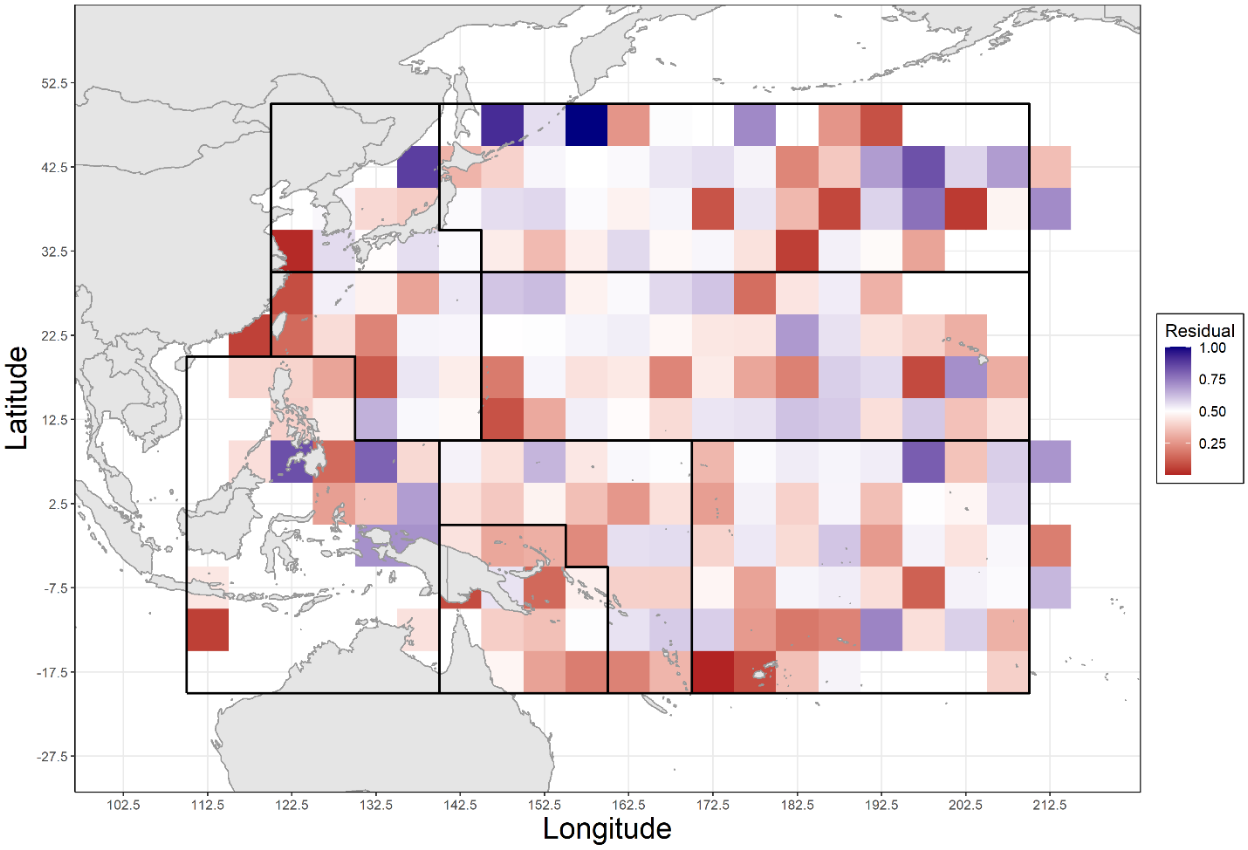


Figure 5. CPUE index estimation results in the sdmTMB and VAST models: (a) mean-centered standardized CPUE index and (b) Area-weighted CPUE standardized abundance indices for regions 1-8. The VAST model output results were based on results from the previous resource assessment in 2022 (annual and quarterly estimates from 1972 to 2020).

(a)



(b)

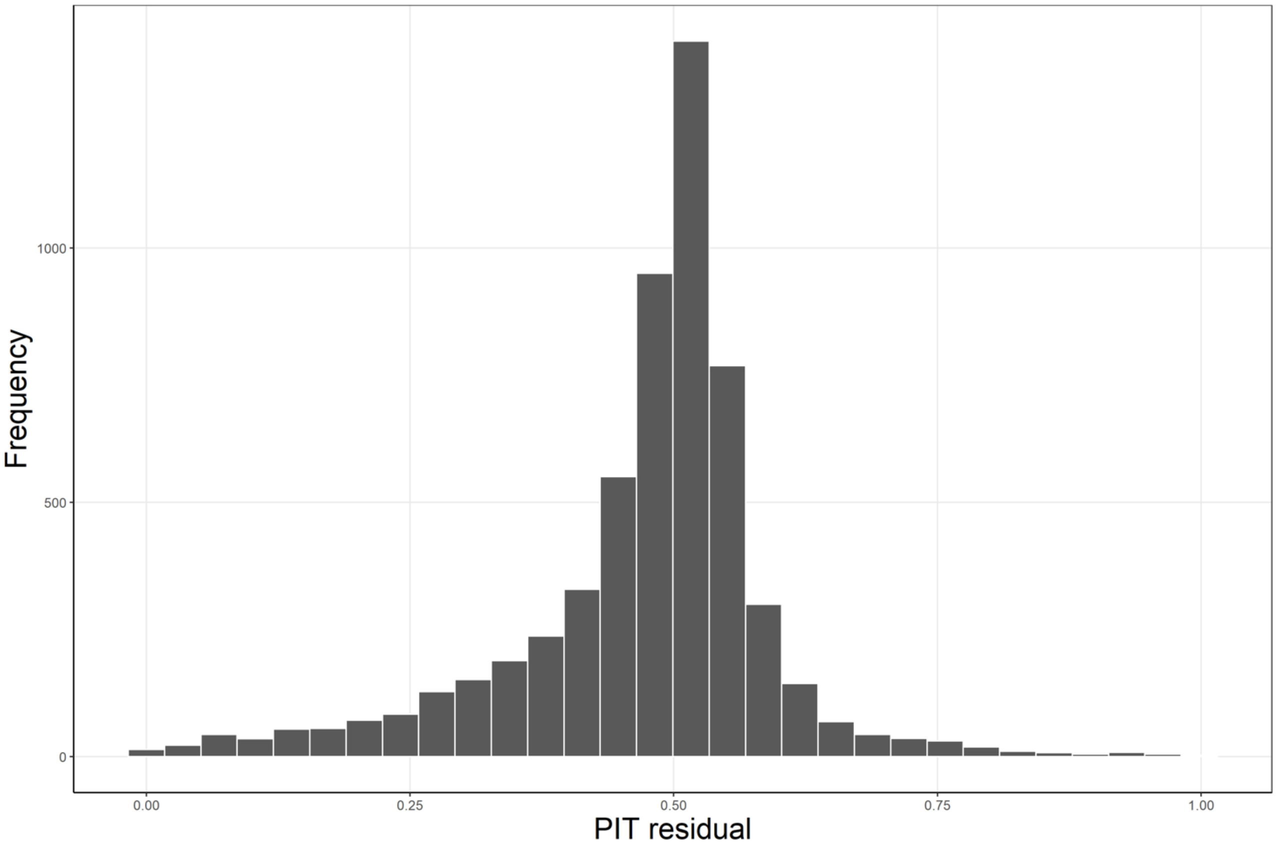


Figure 6. (a) Spatial distribution of probability integral transform (PIT) residuals aggregated over the full time series at the 5×5-degree grid-cell level. (b) Histogram of aggregated PIT residuals.

Appendix 1. Comparison of estimated time-varing CVs for regions 1 to 8 in the global model and regional-level models.

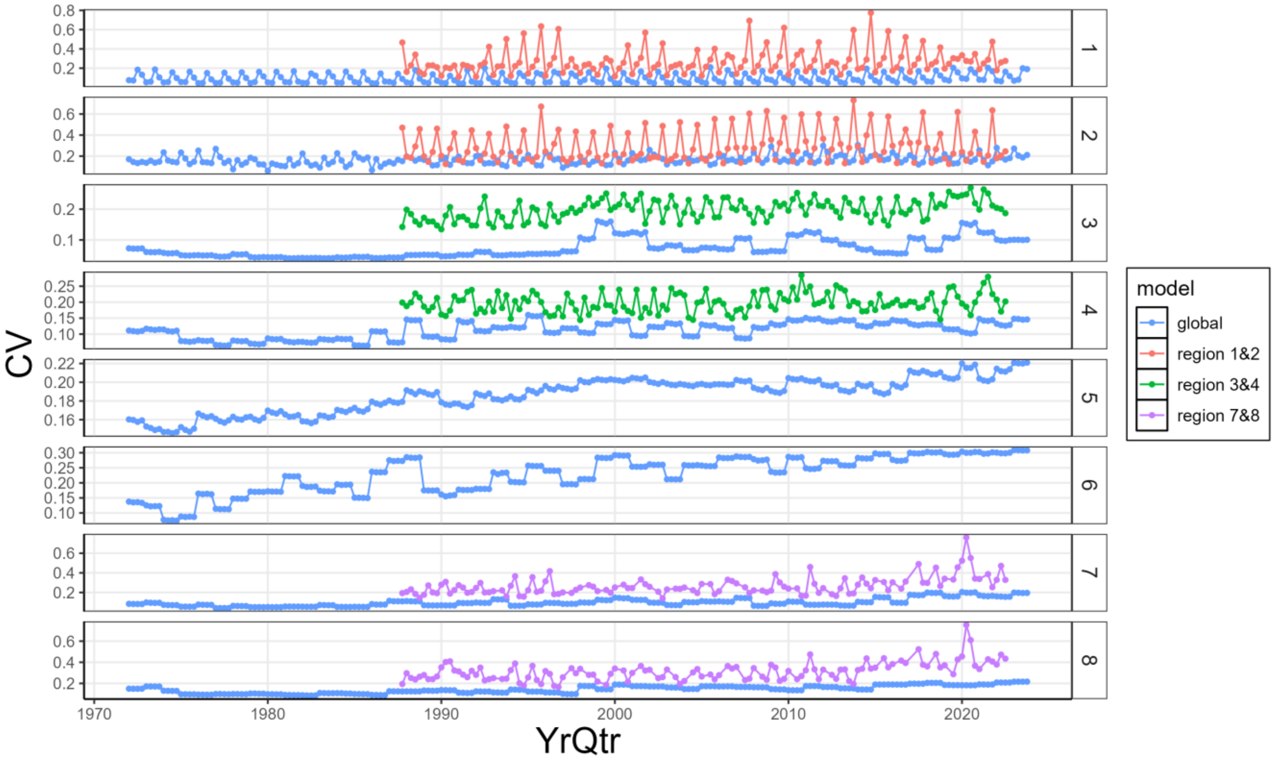


Figure A1. Comparison of estimated time-varing CVs for regions 1 to 8 in the global model and regional-level models using sdmTMB. Blue indicates the global model calculated using the model equation in the text, including all regions 1 to 8. Red, green, and purple represent individual models created using INLA for each subset of Area 1&2, Area 3&4, and Area 7&8, respectively. Note that the regional-level model uses a model that includes YrQtr as a spatiotemporal random effect and uses data from the most recent 30 years to converge stably.