ISC/25/SHARKWG-1/02

Spatio-temporal model for CPUE standardization: Application of blue shark caught by Japanese offshore and distant water shallow-set longliner in the western North Pacific up to 2023[1]

Mikihiko Kai[2]

[2] Fisheries Resources Institute, Japan Fishery Research and Education Agency
2-12-4 Fukuura, Kanazawa-ku, Yokohama 236-8648, JAPAN
Email: kai_mikihiko61@fra.go.jp

[1] Working document submitted to the ISC Shark Working Group Workshop, 27 November-3 February 2025, Hybrid-meeting. **Document not to be cited without author's permission.**

Abstract

This working paper updates the standardized CPUE of blue shark caught by Japanese offshore and distant-water shallow-set longline fishery in the western North Pacific up to 2023. Since the catch data of sharks caught by commercial tuna longline fishery is usually underreported due to the discard of sharks, the author filtered the logbook data using similar filtering methods applied in 2021. The nominal CPUE of filtered shallow-set data was then standardized using the spatio-temporal generalized linear mixed model (GLMM) to provide the annual changes in the abundance of blue sharks in the northwestern Pacific. The author focused on seasonal and interannual variations of the density in the model to account for spatial and seasonal changes in the fishing location due to the target changes between blue shark and swordfish. The predicted annual changes in the CPUE of blue shark revealed a slight upward trend until 2005, followed by a downward trend after 2005, reaching its lowest level in 2008. After that, it showed an increasing trend again until 2015, but then started to decrease again, and it has been fluctuating in recent years. The abundance indices predicted from the spatio-temporal model, with a large amount of data collected in the most abundant waters in the western North Pacific, are very useful information about the abundance in this region.

Introduction

For the stock assessment in 2022, Japan provided a standardized CPUE of blue shark (*Prionace glauca*) caught by Japanese offshore and distant-water shallow-set longline fishery in the western North Pacific from 1994 to 2020 (Kai, 2021). The CPUE was chosen as one of the available indicators of stock abundance for the late period (i.e., 1994-2020) in the stock assessment due to its broad spatial-temporal coverage in the main distribution area (i.e., temperate waters), the statistical soundness of the standardization process, size and sex composition, and larger catch relative to other fisheries (ISC, 2022). The predicted annual changes in the CPUE of blue shark revealed an increasing trend in the 1990s, a stable trend from 1998 to 2005, a downward trend until 2008, and an upward trend thereafter, except for 2020.

The Japanese shallow-set longline fleets changed their operational area by season without changing the gear configurations (e.g., hooks between floats and length of the branch line) even if they changed their target species from swordfish (*Xiphias gladius*) to blue shark, and vice versa. To account for the spatio-temporal changes in the operational area, the author used the VAST (Vector Autoregressive Spatio-Temporal) software package for R (Thorson, 2019).

The objective of this working paper is to update the standardized CPUE of blue shark caught by Japanese offshore and distant-water shallow-set longline fishery up to 2023 and to provide the annual relative abundance index for the indicator analysis of blue shark in 2025. The nominal CPUE is standardized using the spatio-temporal GLMM (VAST), taking into account seasonal and interannual changes in density.

Materials and Methods

The author used a similar statistical filtering method and spatio-temporal model as those used in the paper by Kai (2021).

Data sources

Set-by-set logbook data from Japanese offshore and distant-water longline fisheries in the western North Pacific (20-45° N, 130° E -160° W) were used to estimate the standardized CPUE for 1994-2023. The set-by-set logbook data included information on catch number, amount of effort (number of hooks), number of branch lines between floats (hooks between floats: HBF) as a proxy for gear configuration, location (longitude and latitude) of set by resolution of 1 × 1 degree square, vessel identity (vessel name/call sign), fishery type (offshore/distant water), and the prefecture in Japan where the longline boats were registered. The offshore "Kinkai" fleet was defined by tonnage of vessels between 20 and 120 MT, while the distant-water "Enyo" fleet consisted of vessels larger than 120 MT. Japanese research and training vessel ("JRTV") data for offshore and distant-water longline fishery were not included in this analysis because these fleets are not commercial fisheries.

Data filtering

The logbook data was filtered to remove the set-by-set data including discard and under-reporting catch. First, the set-by-set data was selected by the number of hooks per basket (HPB; 3-5) to identify a shallow-set fishery, as the shallow-set fishery targets blue sharks or swordfish in the northwestern Pacific, while a deep-set fishery usually targets tunas (HPB; 6-21). Second, the set-by-set data was selected by a reporting rate of shark's positive catch by vessel (RR; number of sets with shark recorded/total number of sets $\geq$ 0.946) because Clarke et al. (2011) mentioned that one of the potential reasons for high reporting rates (i.e., 0.946) for sharks in the northwestern Pacific could be a commercial interest in those catches due to the presence of Japan's largest shark market at Kesen-numa, Miyagi prefecture. Another reason is the relatively higher abundances in the region, leading to higher catch rates of sharks (ISC, 2017). Third, the set-by-set data was selected by the registered prefecture ("Tohoku and Hokkaido," including Hokkaido, Aomori, Iwate, Miyagi, Fukushima, and Toyama) of vessels because the fleets in these prefectures frequently target blue sharks, and the RR of the vessels registered in these prefectures is quite high.

CPUE standardization with spatio-temporal model

The spatio-temporal model consists of two components: encounter probability and positive catch in a delta model. The first predictor was fixed at a constant value due to the high positive catches (> 98.7%). The second predictor was modeled using a negative binomial (NB) model to account for the count data with over-dispersion (variance/mean = 178.5):

$$c \sim NegBin\ (c^*, c^*(1+\sigma_1) + {c^*}^2\sigma_2),$$

$$\log\ (d) = d_0(t) + \gamma(s) + \omega(s,q) + \delta(s,y) + \theta(s,t), \quad (1)$$

where c is observed catch, $NegBin\ (a, b)$ is a negative binomial distribution with mean a and variance b (Lindén and Mäntyniemi, 2011), c^* is an expected catch and a function of density d and fishing effort f (number of hooks = 1), σ_1 and σ_2 are residual variations, $d_0(t)$ represents temporal variation (the intercept for each year-season t), $\gamma(s)$ represents spatial variation (s), $\omega(s,q)$ represents spatio-temporal variation (station s and season q), $\delta(s,y)$ represents spatio-temporal variation (station s and year y), and $\theta(s,t)$ represents spatio-temporal variation (station s and year-season t). The intercept $d_0(t)$ is decomposed into season and year main effects and an autocorrelated interaction of season and year to specify the interpolation for season-year combinations without any data, using information from adjacent season-years, other years of the same season, or other seasons of the same year (see Thorson et al., 2020).

The VAST (v13_0_0) was used to standardize the nominal CPUE. Annual abundance index I was estimated as:

$$I(t) = \sum_{s=1}^{n_s} f(s) \times c^*(s,t) / \{\sum_{t=1}^{n_t} \sum_{s=1}^{n_s} f(s) \times c^*(s,t)\}, \quad (2)$$

where n_s is total number of knots and f is fishing effort (number of hooks) at location s. The number of knots (n_s = 250) was specified to balance computational speed and spatial resolution.

Model selection and diagnostics

In the previous analysis in 2021, to select the best model, the explanatory variables were sequentially removed from the full model in Eq (1). The best model was then selected using the AIC (Akaike, 1973). Since the predicted CPUE index is used for the indicator analysis of blue shark in 2025, the same model structure selected in 2021 was used. For the model, the goodness of fit was examined using Pearson residuals and a QQ-plot. The residuals were computed using a randomized quantile (Dunn and Smyth, 1996) to produce continuous normal residuals.

Results

Data filtering reduced the number of datasets collected in the North Pacific from 1,921,027 to 106,516.

Selection of the best model and annual trends in CPUE

The model, which uses the same structure selected in 2021, showed reasonable convergence with a positive definite Hessian matrix and a small maximum gradient (< 0.001). The model incorporates spatial (station) and spatio-temporal variances (station and year; station and season; and station and year-season) as random effects. A list of all parameters and estimates of the model is provided in Table 1. The predicted annual changes in the CPUE of blue shark revealed a slight upward trend until 2005, followed by a downward trend after 2005, reaching its lowest level in 2008. After that, it showed an increasing trend again until 2015, but then started to decrease again, and it has been fluctuating in recent years. The 95% confidence intervals in the CPUE estimates were substantially larger after 2015 because the fishing efforts (number of hooks) were smaller than those of any other years. The seasonal changes in the predicted CPUE of blue shark indicated the highest CPUE in Q2, followed by Q3, Q4, and Q1. The spatial maps of predicted CPUE showed that the hotspots appeared in the temperate water (30-40 ˚N and 150-180 ˚E) in summer (April-June) and autumn (July-September) throughout the years. These results suggest that many blue sharks are dominant in this area from summer to autumn, as fishermen operate in this area targeting blue sharks during this period.

Model diagnostics

Diagnostic plots of goodness-of-fit for the best model didn’t show a serious deviation from normality or model misspecification (**Fig. 3**). These results suggested that the fitting of the best model to the data was good.

Discussions

This document paper updates the annual abundance indices of blue shark caught by Japanese shallow-set longline fishery in the western North Pacific Ocean up to 2023, using a spatio-temporal GLMM, taking into account seasonal and interannual variations in density. The main advantage of the spatio-temporal model is an imputation for the missing data using spatial and temporal correlations through random effects (Thorson, 2019). Unlike the design based GLM used in the past assessment, the spatio-temporal GLMM developed by Thorson et al. (2020) enabled us to include interaction terms between spatial and temporal effects (season, year and season-year effects) with high spatial resolutions. The consideration of spatial variation with higher resolution had a large impact on the seasonal trends in the predicted CPUE (i.e., the CPUE in Q2 was the highest in Fig 2). The remarkable changes in the predicted CPUE by season and adding of interaction terms between year-season and stations resulted in the substantial changes in the predicted trends of annual CPUE (Fig. 1; Kai, 2021). These results suggested that abundance indices of blue shark significantly increased in 1990s due to the reduction of high fishing pressure of a drift net fishery prior to 1993 (Fujinami et al., 2021a and b). Thereafter the abundance indices remained at higher levels until 2005 and sharply decreased and reached a historical lowest level in 2008 as the increase of fishing pressure from the longline fishery in 2000s. After 2008, the abundance indices gradually increased as the decrease of the fishing effort from the longline fishery. The abundance indices in recent years have been fluctuating. Therefore, the results of this study indicate that there are no signs of a significant decline in the blue shark population in this region in recent years. The abundance indices predicted from the spatio-temporal model, with a large amount of data collected in the most abundant waters in the western North Pacific, are very useful information about the abundance in this region.

References

Akaike, H. (1973) Information theory and an extension of the maximum likelihood principle. *In* Petrov, B.N., Csaki, F. (Eds.) Second International Symposium on Information Theory, Budapest, Akademiai Kiado, pp 267-281.

Clarke, S.C., Yokawa, K., Matsunaga, H., and Nakano, H. (2011) Analysis of North Pacific shark data from Japanese commercial longline and research/training vessel records. WCPFC-SC7-2011/EB-WP-02.

Dunn, K. P., and Smyth, G. K., 1996. Randomized quantile residuals. J. Comput. Graph. Stat. 5, 236-244.

Fujinami, Y., Kanaiwa, M., and Kai, M. (2021a) Estimation of annual catch for blue sharks caught by Japanese high seas squid driftnet fishery in the North Pacific Ocean from 1981 to 1992. ISC/21/SHARKWG-2/07.

Fujinami, Y., Kanaiwa, M., and Kai, M. (2021b) Blue shark catches in the Japanese large-mesh driftnet fishery in the North Pacific Ocean from 1973 to 1993. ISC/21/SHARKWG-2/08.

ISC (2017) Stock assessment and future projections of blue shark in the North Pacific Ocean through 2015. ISC 17 Plenary report and document.

ISC (2022). Report of the 22nd meeting of the International Scientific Committee for Tuna and Tuna-like Species in the North Pacific Ocean. Plenary session, July 12-18, Kona, Hawaii, United States of America.

Kai, M. (2021) Spatio-temporal model for CPUE standardization: Application to blue shark caught by Japanese offshore and distant water shallow-set longliner in the western North Pacific. ISC/21/SHARKWG-2/01.

Lindgren, F., Rue, H., and Lindstr€om, J. (2011) An explicit link between Gaussian fields and Gaussian Markov random fields: The SPDE approach. J. R. Stat. Soc., Ser. B, 73: 423–498.

Thorson, J.T. (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, J.T., Adams, C.F., Brooks, E.N., Eisner, L.B., Kimmel, D.G., Legault, C.M., Rogers, L.A., and Yasumiishi, E.M. (2020) Seasonal and interannual variation in spatio-temporal models for index standardization and phenology studies. ICES J. Mar. Sci. 77: 1879−1892.

Tables

Table 1. List of all parameters and estimates of the selected model.

No	Parameter name	Symbol	Type	Estimates
1	Distance of correlation (Spatial random effect)	κ	Fixed	0.0049
2	Northings anisotropy	h_1	Fixed	1.64
3	Anisotropic correlation	h_2	Fixed	0.98
4	Parameter governing pointwise variance (Spatial random effect)	η_γ	Fixed	0.57
5	Parameter governing pointwise variance (Spatio-temporal (season) random effect)	η_ω	Fixed	0.44
6	Parameter governing pointwise variance (Spatio-temporal (year) random effect)	η_δ	Fixed	No estimation
7	Parameter governing pointwise variance (Spatio-temporal (year-season) random effect)	η_θ	Fixed	1.26
8	Parameter governing autocorrelation (Spatio-temporal: year-season random effect)	ρ_θ	Fixed	1.34
9	Residual variation 1 of negative binomial model	σ_1	Fixed	0.37
10	Residual variation 2 of negative binomial model	σ_2	Fixed	0.44
11	Intercept for first predictor	β_1	Fixed	4.68
12	Intercept for second predictor	β_2	Fixed	-4.43
13	Spatial residuals	γ	Random	Not shown
14	Spatio-temporal (season) residuals	ω	Random	Not shown
15	Spatio-temporal (year) residuals	δ	Random	No estimation
16	Spatio-temporal (year-season) residuals	θ	Random	Not shown

Table 2. Summary of annual CPUE predicted by spatio-temporal model along with corresponding estimates of the coefficient of variation (CV), annual nominal CPUE, and number of hooks in millions. Values are predicted using the best fitting model and scaled by average CPUE.

Year	Predicted CPUE	Predicted CPUE in 2021	Nominal CPUE	CV	Number of hooks (millions)
1994	0.81	0.84	0.41	0.15	19.4
1995	0.98	0.90	0.45	0.17	18.2
1996	0.86	0.85	0.52	0.15	17.5
1997	1.14	1.04	0.74	0.15	16.5
1998	1.11	1.03	0.72	0.14	17.1
1999	1.14	1.09	0.88	0.14	17.4
2000	1.10	1.06	0.99	0.13	20.1
2001	1.32	1.22	1.14	0.11	20.1
2002	0.99	1.03	1.10	0.12	17.7
2003	1.14	1.08	1.28	0.10	15.9
2004	1.05	1.03	1.17	0.10	15.5
2005	1.27	1.26	1.46	0.11	13.6
2006	1.01	1.06	1.32	0.10	13.2
2007	0.83	0.84	0.95	0.10	15.6
2008	0.62	0.73	0.93	0.11	13.5
2009	0.93	0.97	1.15	0.11	12.2
2010	0.91	1.04	1.06	0.12	11.3
2011	0.78	0.86	0.74	0.14	6.2
2012	0.75	0.88	0.99	0.14	7.5
2013	0.81	0.92	0.71	0.16	8.0
2014	1.10	1.04	0.79	0.17	7.9
2015	1.36	1.17	1.07	0.18	6.7
2016	1.14	1.14	1.18	0.18	7.2
2017	1.02	1.06	1.36	0.17	6.9
2018	0.94	1.04	1.19	0.19	7.1
2019	1.16	1.01	1.00	0.19	6.8
2020	0.68	0.81	0.71	0.18	7.1
2021	1.01		1.11	0.16	5.0
2022	1.15		1.52	0.19	3.8
2023	0.86		1.35	0.18	5.1

Figures

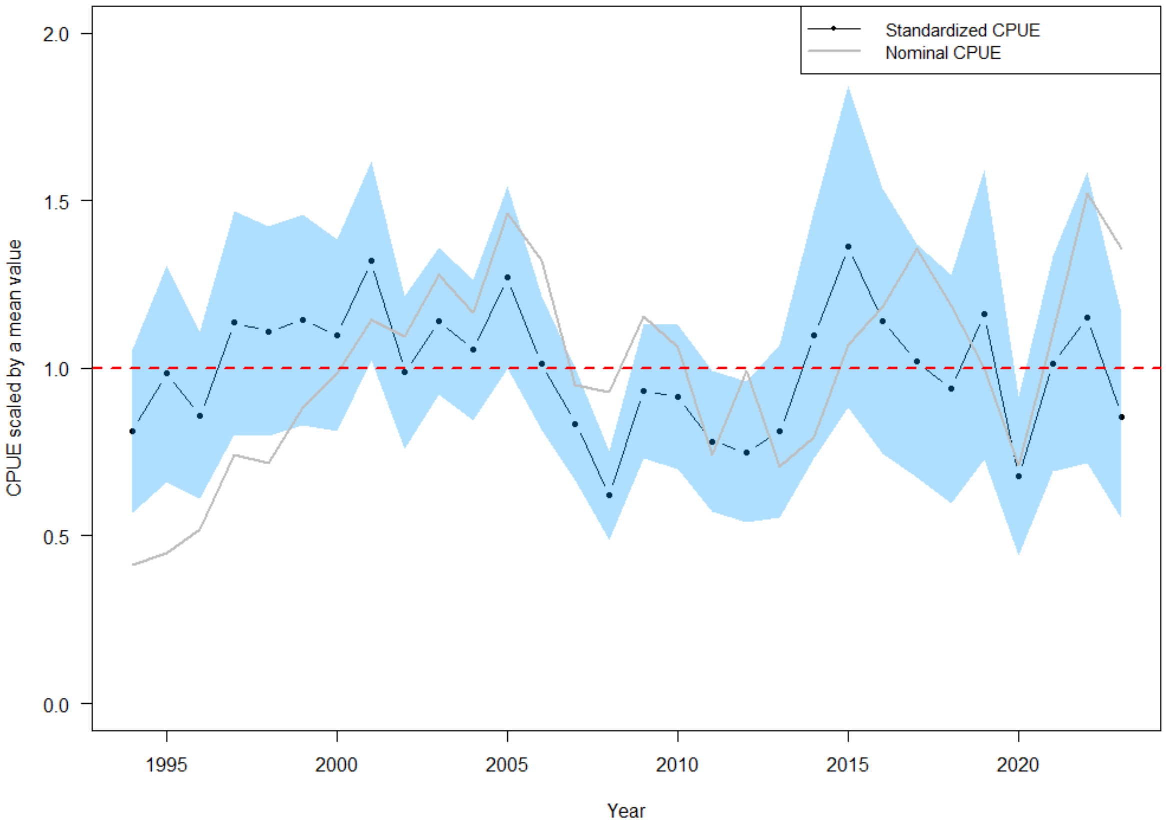

Fig. 1 Annual predicted CPUE relative to its average. Gray solid line denotes nominal CPUE relative to its average, shadow denotes 95% confidence intervals, and horizontal dotted line denotes mean of relative values (1.0).

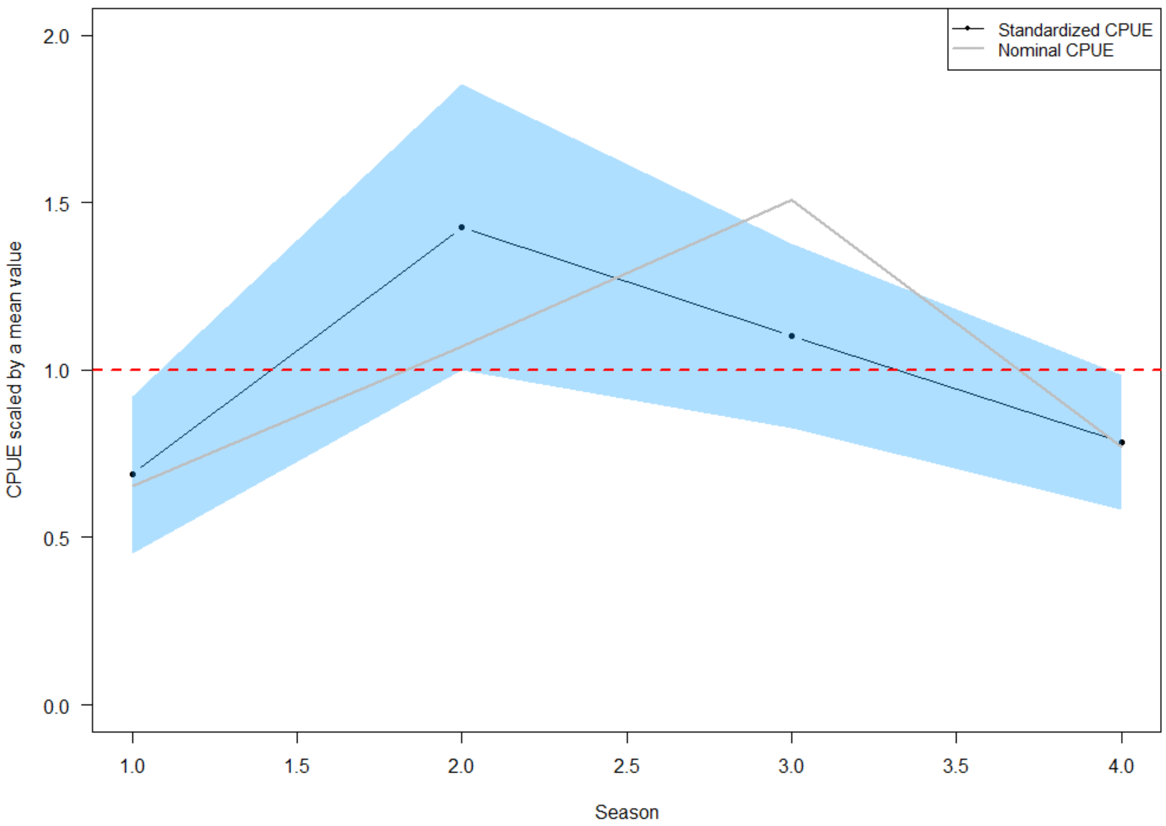

Fig. 2 Seasonal predicted CPUE relative to its average. Gray solid line denotes nominal CPUE relative to its average, shadow denotes 95% confidence intervals, and horizontal dotted line denotes mean of relative values (1.0).

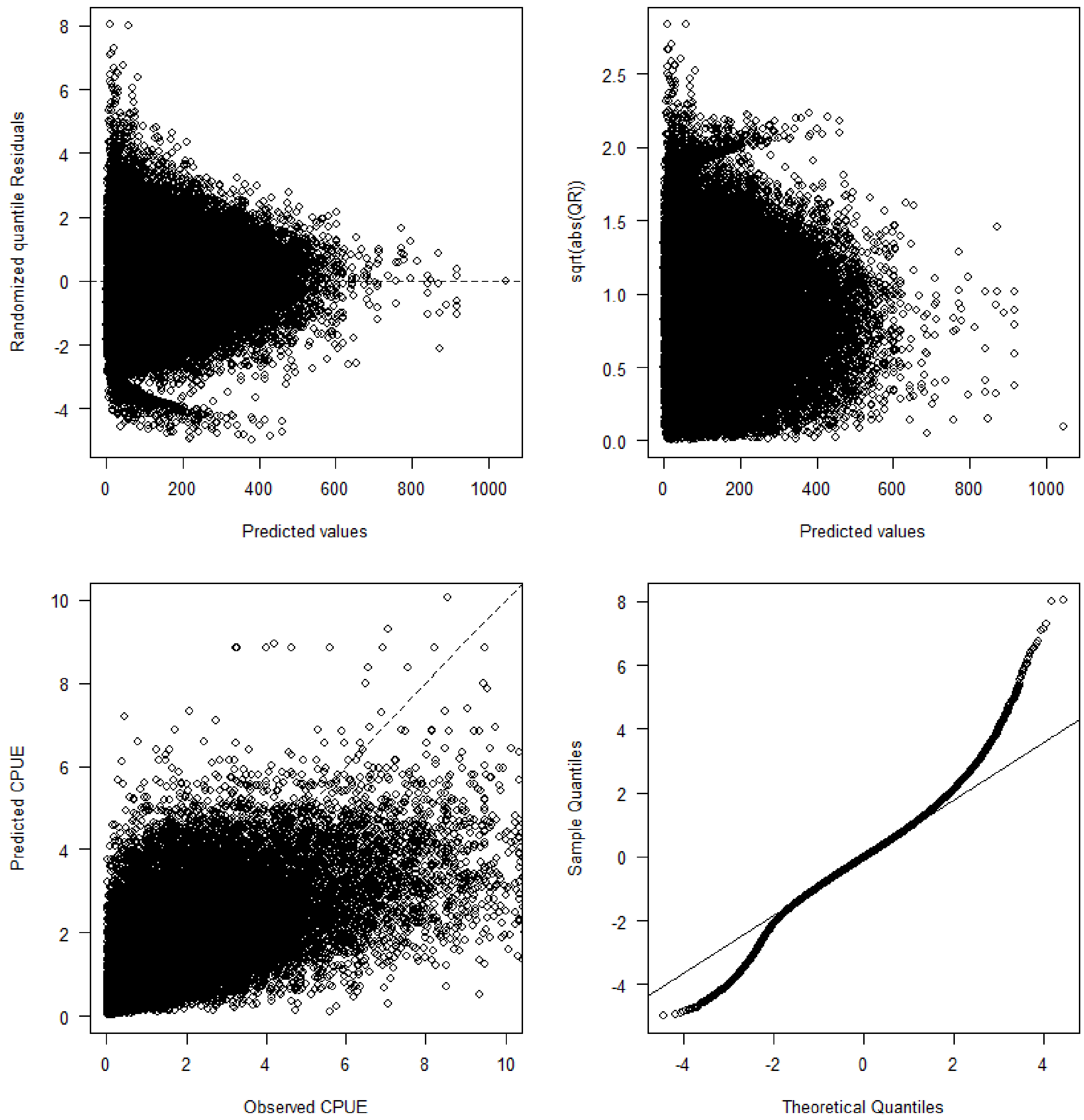

Fig. 3 Diagnostic plots of goodness-of-fit for the most parsimonious model. (a) Standardized residuals versus the fitted values can assess whether model misspecification is occurring; (b) the observed versus the predicted values can assess qualitatively whether the explanatory variables are indeed able to reduce variance in the data; (c) the square root of the absolute values of the standardized residuals versus the fitted values can assess whether variance changes as a function of the predicted value; and (d) quantile-quantile (QQ) plots can assess normality.

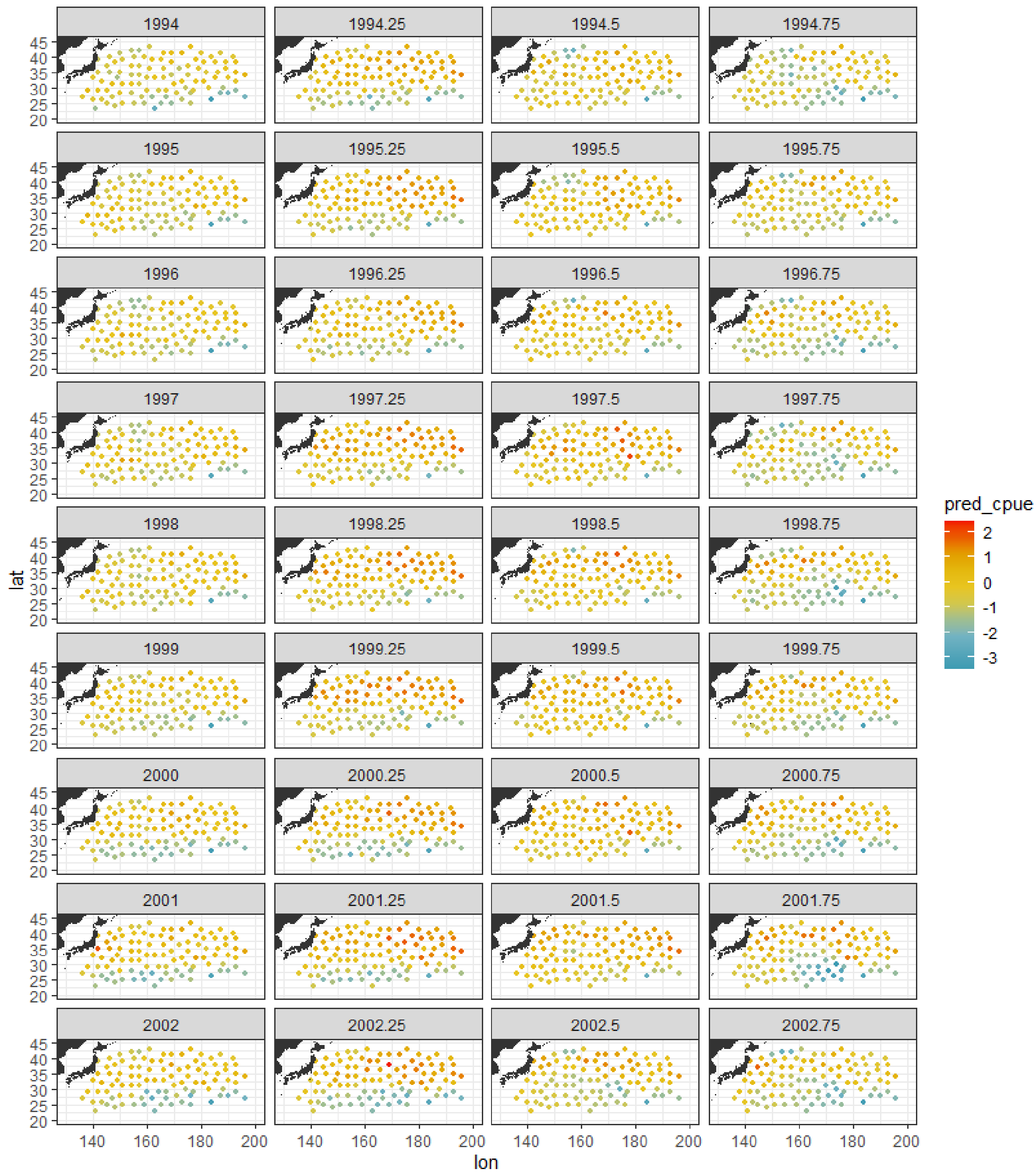

Fig. 4 Year- and season- specific spatial distribution of log-scaled predicted CPUE for blue shark from 1994 to 2002. The number of knots is 250.

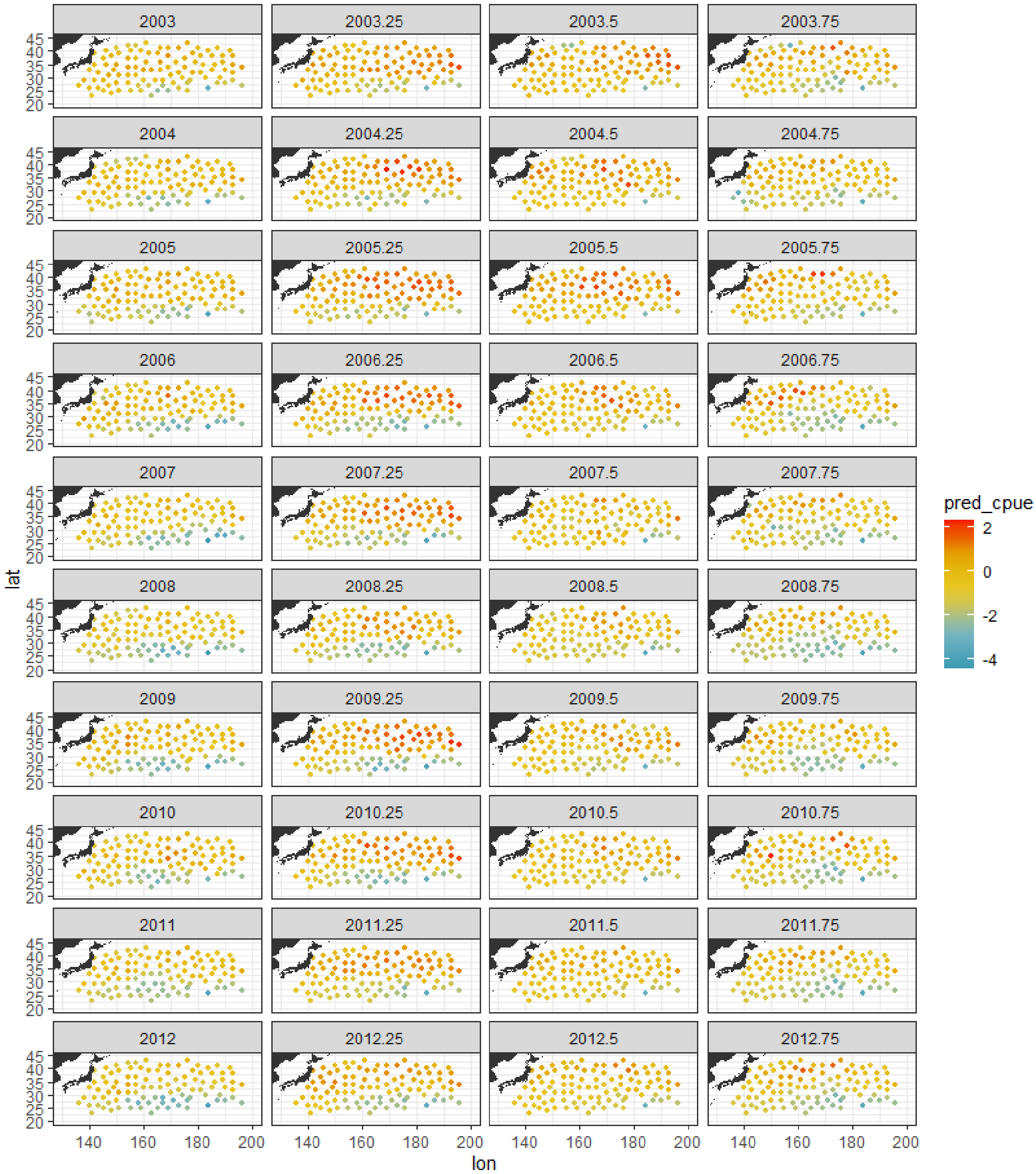

Fig. 5 Year- and season- specific spatial distribution of log-scaled predicted CPUE for blue shark from 2003 to 2011. The number of knots is 250.

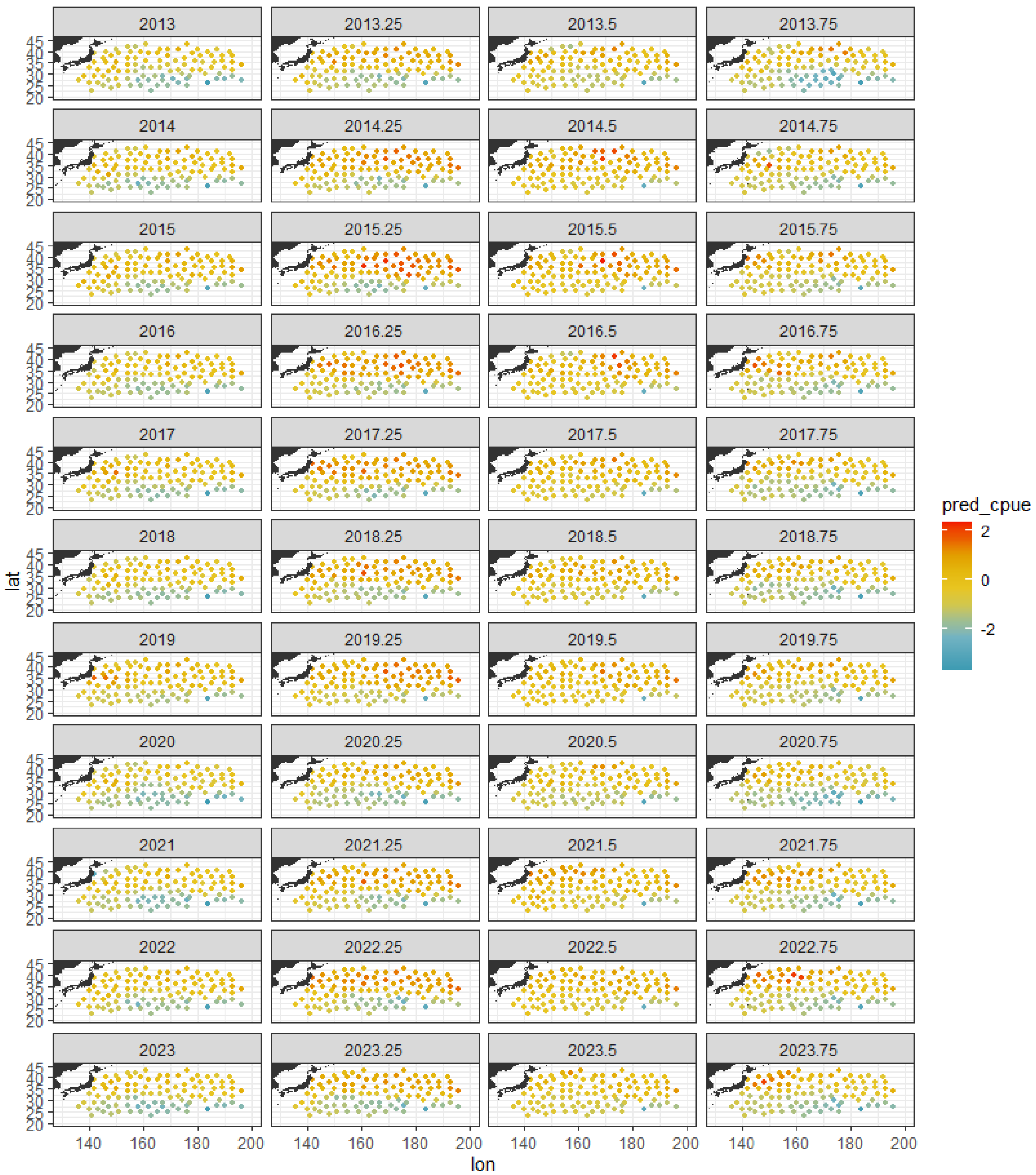

Fig. 6 Year- and season- specific spatial distribution of log-scaled predicted CPUE for blue shark from 2013 to 2023. The number of knots is 250.