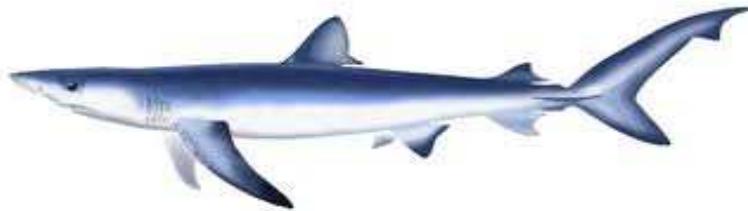


**Spatio-temporal model for CPUE standardization:
Application to blue shark caught by Japanese
offshore and distant water shallow-set
longliner in the western North Pacific¹**

Mikihiko Kai²

² Fisheries Resources Institute, Japan Fishery Research and Education Agency
5-7-1 Orido, Shimizu-ku, Shizuoka 424-8633, JAPAN
Email: kaim@affrc.go.jp



¹ Working document submitted to the ISC Shark Working Group Workshop, 9-19 November 2021, Web-meeting. **Document not to be cited without author's permission.**

Abstract

This working paper provides a standardized CPUE of blue shark caught by Japanese offshore and distant-water shallow-set longline fishery from 1994 to 2020 in the western North Pacific Ocean. Since the catch data of sharks caught by commercial tuna longline fishery is usually underreported due to discard of sharks, the author filtered the logbook data using the similar filtering methods applied in the previous analysis. 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 spatially and seasonally changes in the fishing location due to the target changes between blue shark and swordfish. The estimated annual changes in the CPUE of blue shark revealed an upward trend from 1994 to 2005, and then downward trend until 2008. Thereafter the CPUE gradually increased until 2015 and then slightly decreased in recent years. The estimated CPUE trends from the spatio-temporal model with a large amount of data collected in the most abundant waters in the North Pacific Ocean is a very useful information about the abundance of North Pacific blue shark.

Introduction

In the previous stock assessment in 2017, the ISC shark working group (WG) used standardized CPUE of blue shark (*Prionace glauca*) caught by Japanese offshore and distant water shallow-set longline fishery in the western North Pacific Ocean from 1994 to 2015 (Kai and Shiozaki, 2016) as an abundance indices of base case model (ISC, 2017). The CPUE was chosen as the best available indicators of stock abundance for the late time (i.e., 1994-2015) due to their broad spatial-temporal coverage in the main distribution area (i.e., temperate water), statistical soundness of the standardization process, size and sex composition, and larger catch relative to other fisheries (ISC, 2017). However, the WG acknowledged insufficient analysis for an issue of the CPUE standardization on the targeting/fishing strategy shifts from blue shark to swordfish and vice versa and requested in depth research efforts on this issue in future work.

The VAST (Vector Autoregressive Spatio-Temporal) software package for R (Thorson, 2019), which enables us to analyze fishery data using the spatio-temporal generalized linear mixed model (GLMM) (Thorson et al., 2015), has recently attracted attention as a new approach and is now commonly used globally to predict spatial changes in species distribution and temporal variations in a population range and density. The basic

model structure of VAST adopts a delta-GLMM which can consider spatio-temporal correlations among categories such as a species. Actually, the spatio-temporal model was applied to the estimation of abundance from multi-species fishery data accounting for spatio-temporal variation and fisher targeting (Thorson et al. 2017). This fact implies that the multispecies spatio-temporal model has a high potential to improve the CPUE standardization of the blue shark caught by Japanese shallow-set longline fishery in the western North Pacific. However, the multiple reports from the Japanese skipper's notes revealed that the Japanese shallow-set longline fleets only changed their operational area by season without changing the gear configurations (e.g. hooks between floats and length of the blanch line) even if they changed their target species. The main reason why the previous authors (Hiraoka et al., 2016; Kai and Shiozaki, 2016) directly considered the target effects in the CPUE standardization is that the generalized linear model (GLM) commonly requires to provide a spatial area with a low resolution (e.g., five areas) in the analysis. Therefore, they directly included the target effect in the model as it was impossible to directly explain the spatio-temporal changes in the operational area using the GLM. If a single-species spatio-temporal model can account for the seasonal and interannual variation in the spatial changes of the density with higher spatial resolutions, it is reasonable to use the spatio-temporal model for the CPUE standardization without directly considering the target effect in the analysis. Recently, Thorson et al (2020) developed such a model.

The objective of this working paper is to estimate the standardized CPUE of blue shark caught by Japanese offshore and distant-water shallow-set longline fishery for 1994-2020 using the spatio-temporal GLMM in consideration with seasonal and interannual changes in the density.

Materials and Methods

Data sources

Set-by-set logbook data from Japanese offshore and distant water longline fisheries in the western North Pacific Ocean (20-45° N, 130° E -160° W) were used to estimate the standardized CPUE for 1994-2020. 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 float: 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 was not included in this analysis because the fleets are not commercial fishery.

Data filtering

Using the similar method (Kai and Shiozaki, 2016), 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 baskets (HPB; 3~5) to select a shallow-set fishery because the shallow-set fishery targets the 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 ≥ 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 in relation to the presence of Japan’s largest shark market at “Kesenuma”, Miyagi prefecture. Another reason is explained by the relatively higher abundances in the region, and thus 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 are quite high.

CPUE standardization with spatio-temporal model

The spatio temporal model is consisted of two components of encounter probability and positive catch in a delta model. The first predictor was fixed at a constant value because of high positive catches (> 98%). Second predictor was modeled using a negative binomial (NB) model to account for the count data with over-dispersion (variance/mean =176.5):

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

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

where c is observed catch, $\text{NegBin}(a, b)$ 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 $d_0(t)$ are 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. Temporal 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 ($n_s = 100$) was specified in a balance between computational speed and spatial resolution.

Model selection and diagnostics

To select the best model, the explanatory variable was sequentially added to the year-season random effect model. The best model was selected using the AIC (Akaike 1973) and BIC (Schwarz 1978). For the best model, the goodness of fits was examined using the Pearson residuals and 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 Ocean from 1,830,668 to 104,278.

Selection of the best model

All models were reasonably converged with the positive definite of hessian matrix and a small value of maximum gradient (**Table 1**). The model (M_3) including spatial (station) and spatio-temporal variances (season and station, and year-season and station) as random effects were identified by AIC and BIC as the most parsimonious model (**Table 1**). The estimated CPUE changed substantially if random effect components were sequentially added to the simplest model (M_1) which only includes an interaction term of year-season and station as a random effect (**Fig. 1**). Diagnostic plots of goodness-of-fit for the best model didn't show serious deviations from normality and model misspecification (**Fig. 2**). These results

suggested that the fitting of the best model to the data was good. Lists of all parameters and estimates of the best models are shown in **Table 2**.

Temporal trends in CPUE

The estimated annual changes in the CPUE of blue shark revealed an upward trend from 1994 to 2005, and then downward trend until 2008. Thereafter the CPUE gradually increased until 2015 and then slightly decreased in recent years (**Fig. 3**). The 95% confidence intervals of the CPUEs were larger in 1990s and after 2015 (**Fig. 3**). The estimated seasonal changes in the CPUE of blue shark indicated the highest CPUE in Q2 and followed by that in Q3, Q4 and Q1 (**Fig. 4**). The year-season spatial maps of predicted CPUE clearly showed the higher CPUEs of blue sharks at the higher latitudes (30-40° N, 130° E -160° W) in Q2 and Q3 (**Figs. 5-7**). Meanwhile, the lower CPUEs of blue sharks were observed in the sub-tropical areas throughout the whole seasons.

Discussions

This document paper estimated a historical trend in abundance indices of blue shark caught by Japanese shallow-set longline fishery in the western North Pacific Ocean from 1994 to 2020 using a spatio-temporal GLMM in considerations with seasonal and interannual variations of the 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 previous assessment (Kai and Shiozaki, 2016), 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 spatio-temporal variations with high spatial resolution had a large impact on the seasonal trends in the estimated CPUE (i.e., the highest CPUE in Q2, see **Fig 4**) and that resulted in the substantial changes in the annual CPUEs (**Fig. 3**).

The annual trends of the selected model (M-3) suggested that the abundance indices of blue shark significantly increased in 1990s and remained stable with higher CPUEs until 2005 (**Fig. 2**) due to the reduction of high fishing pressure of a drift net fishery prior to 1993 (Fujinami et al., 2021). Thereafter the abundance indices sharply decreased and reached a historical lowest level in 2008 as the increase of fishing effort around 2000 (**Fig. 6**). In consideration with the time lags between the increase of the fishing effort and the reduction of the CPUEs, the capture sizes of blue sharks could be sub-adults and adults (ages 4-6). The abundance indices gradually increased from 2008 to 2015 as the decrease of the fishing effort

(**Fig. 3**) and then those in recent years slightly decreased. Since the fishing effort in 2020 was almost the same as that in 2019 (**Fig. 6**) when the COVID-19 was widespread in Japan, the reasons for the sharp decline in 2020 is unrelated to the Pandemic. It is therefore necessary to monitor carefully the future catch rates as well as the operational patterns of this fishery.

A comparison of standardized CPUE between present and previous studies indicated a similar historical trend but different magnitude of CPUEs were observed in some years (**Fig. 3**). The correlations of season-year/other seasons of the same year might have effects on the prediction of the standardized CPUEs. In this study, the author used small number of knots (only 100) in this study due to the constraint of the time/machine power. In future work it might be better increasing the number of knots to improve the fitting of the model to the fine spatial scale data.

The author recommends using the predicted annual CPUEs of blue shark caught by Japanese shallow-set longline fishery in the western North Pacific from 1994 to 2020 as a representative of abundance indices for North Pacific blue shark because a wide coverage of the main distributional areas of blue shark (20-45 N° and 130E°-160W°), wide coverage of the size classes as well as both sexes for blue sharks (around 50~250 cm in PCL), and statistical soundness of the spatio-temporal model.

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., 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., Kai, M. 2021. Blue shark catches in the Japanese large-mesh drift-net fishery in the North Pacific Ocean from 1974 to 1993. ISC/21/SHARKWG-1/08.
- Hiraoka, Y., Kanaiwa, M., Ohshimo, S., Takahashi, N., Kai, M., Yokawa, K. (2016) Relative abundance trend of the blue shark *Prionace glauca* based on Japanese distant-water and offshore longliner activity in the North Pacific. *Fish. Sci.* **82**, 687–699.
- ISC (2017) Stock assessment and future projections of blue shark in the North Pacific Ocean through 2015. ISC 17 Plenary report and document.
- Kai, M., Shiozaki, K. (2016) Update of Japanese abundance indices for blue shark caught by Japanese offshore and distant water shallow-set longliner in the North Pacific. ISC/16/SHARKWG-1/10.
- Schwarz, G. E. (1978). Estimating the dimension of a model. *Annals of Statistics*, **6**, 461-464.
- 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.
- Thorson, J.T., Fonner, R., Haltuch, M.A., Ono, K., and Winker, H. (2017) Accounting for spatiotemporal variation and fisher targeting when estimating abundance from multispecies fishery data. *Can. J. Fish Aquat. Sci.* **74**: 1794–1807.
- Thorson, J.T., Shelton, A.O., Ward, E.J., and Skaug, H. (2015) Geostatistical delta-generalized linear mixed models improve precision for estimated abundance indices for West Coast groundfishes. *ICES J. Mar. Sci.* **72**: 1297–1310.

Tables

Table 1. Summary of model structure and outputs among different models. All models include fixed effects. “ Δ ” denotes a difference between the value of criteria and the minimum value for AIC and BIC.

Model	Catch rate predictors of random effect	Number of parameters	Deviance	Δ AIC	Δ BIC	Maximum gradient
M-1	Year-season and station	9	1063256	553	506	< 0.001
M-2	Year-season and station + Station	10	1062884	183	145	< 0.001
M-3	Year-season and station + Station + Season and station	14	1062693	0	0	< 0.001
M-4	Year-season and station + Station + Year and station	37	1067505	194	415	< 0.001
M-5	Year-season and station + Station + Year and station + Season and station	41	1062640	1	259	< 0.01

Table 2. 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	0.0045
2	Northings anisotropy	h_1	Fixed	1.54
3	Anisotropic correlation	h_2	Fixed	1.02
4	Parameter governing pointwise variance (Spatial random effect)	η_ν	Fixed	0.56
5	Parameter governing pointwise variance (Spatio-temporal (season) random effect)	η_ω	Fixed	0.32
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	2.20
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	0.77
10	Residual variation 2 of negative binomial model	σ_2	Fixed	0.44
11	Intercept for first predictor	β_1	Fixed	4.71
12	Intercept for second predictor	β_2	Fixed	-4.25
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 3. 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. CPUEs are predicted using the best fitting model and scaled by the average CPUE.

Year	Predicted CPUE	Nominal CPUE	CV	Number of hooks (millions)
1994	0.84	0.42	0.15	19.9
1995	0.90	0.46	0.15	18.5
1996	0.85	0.53	0.14	17.5
1997	1.04	0.77	0.13	16.7
1998	1.03	0.75	0.13	17.6
1999	1.09	0.92	0.12	17.5
2000	1.06	1.03	0.12	20.2
2001	1.22	1.20	0.10	20.5
2002	1.03	1.13	0.11	17.7
2003	1.08	1.32	0.09	15.8
2004	1.03	1.20	0.10	15.6
2005	1.26	1.49	0.10	13.9
2006	1.06	1.34	0.10	13.3
2007	0.84	0.97	0.10	16.4
2008	0.73	0.96	0.12	13.5
2009	0.97	1.19	0.11	12.1
2010	1.04	1.10	0.13	10.9
2011	0.86	0.77	0.13	6.3
2012	0.88	1.02	0.14	8.0
2013	0.92	0.73	0.15	8.6
2014	1.04	0.82	0.15	8.5
2015	1.17	1.12	0.15	7.3
2016	1.14	1.25	0.15	7.8
2017	1.06	1.43	0.15	7.4
2018	1.04	1.25	0.15	7.6
2019	1.01	1.07	0.15	7.2
2020	0.81	0.73	0.17	7.4

Figures

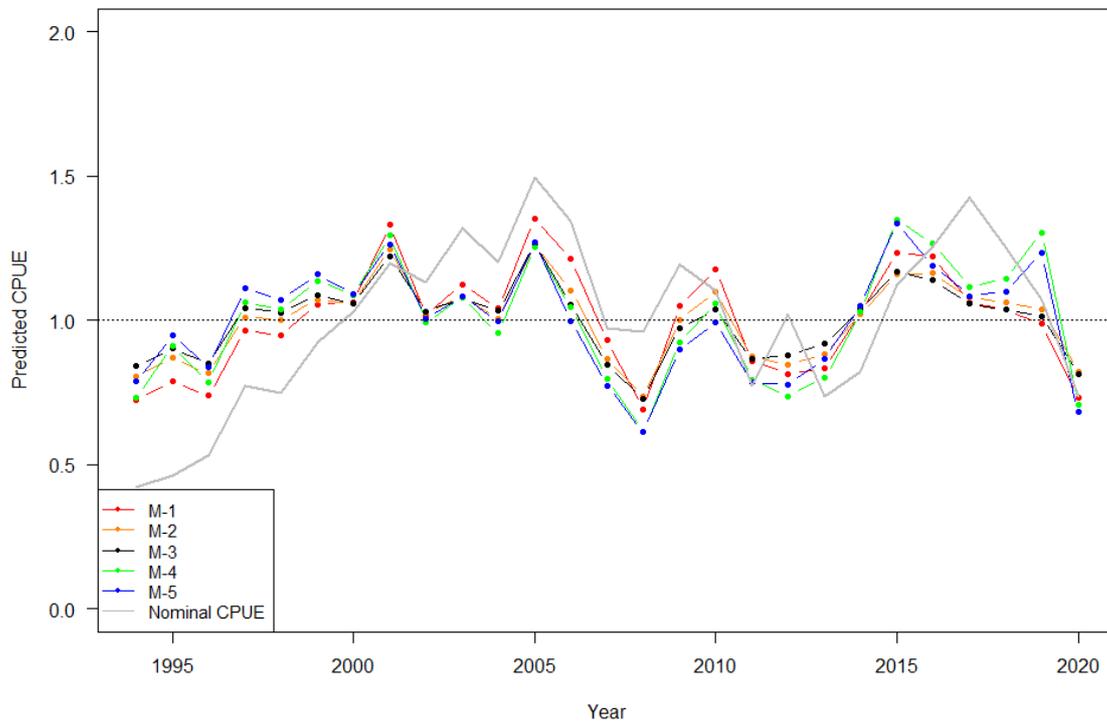


Fig. 1 Comparisons of annual predicted CPUE relative to its average among different model structures. For the details of the models, see table 1. The horizontal dotted line denotes mean of relative values (1.0).

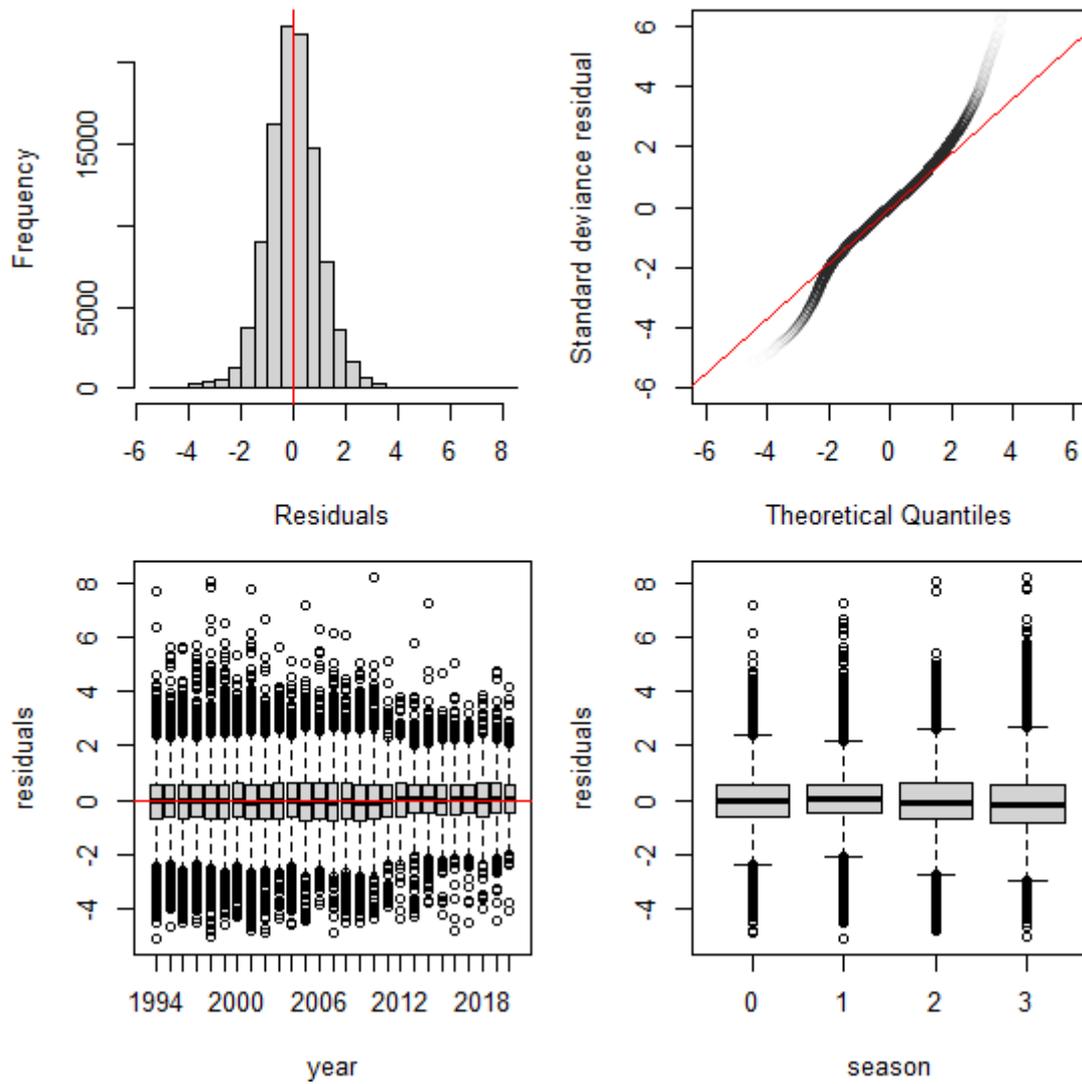


Fig. 2 Diagnostic plots of goodness-of-fit for the most parsimonious model (M-3).

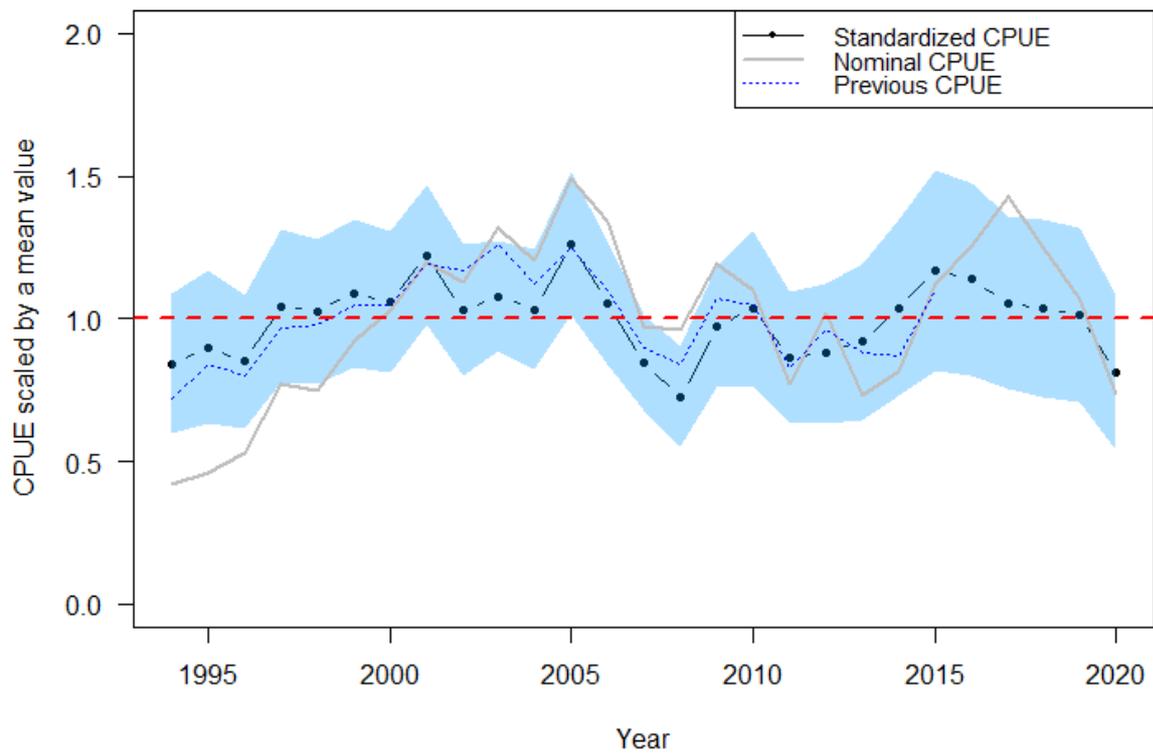


Fig. 3 Annual predicted CPUE relative to its average of the best model (M-3). Gray solid line denotes nominal CPUE relative to its average, shadow denotes 95% confidence intervals, blue dotted line denotes standardized CPUE used in the previous assessment and horizontal dotted line denotes mean of relative values (1.0).

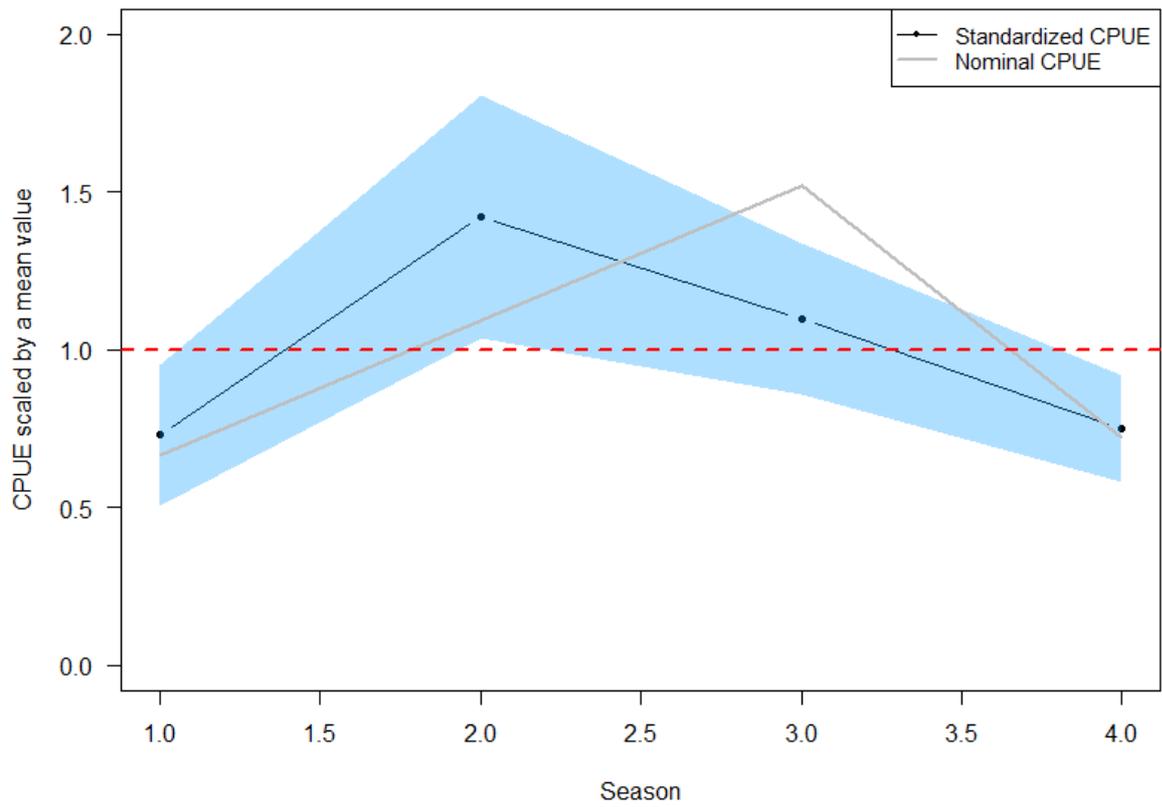


Fig. 4 Seasonal predicted CPUE relative to its average. Gray solid line denotes nominal CPUE relative to its average, shadow denotes 95% confidence intervals, and horizontal dotted line denotes mean of relative values (1.0).

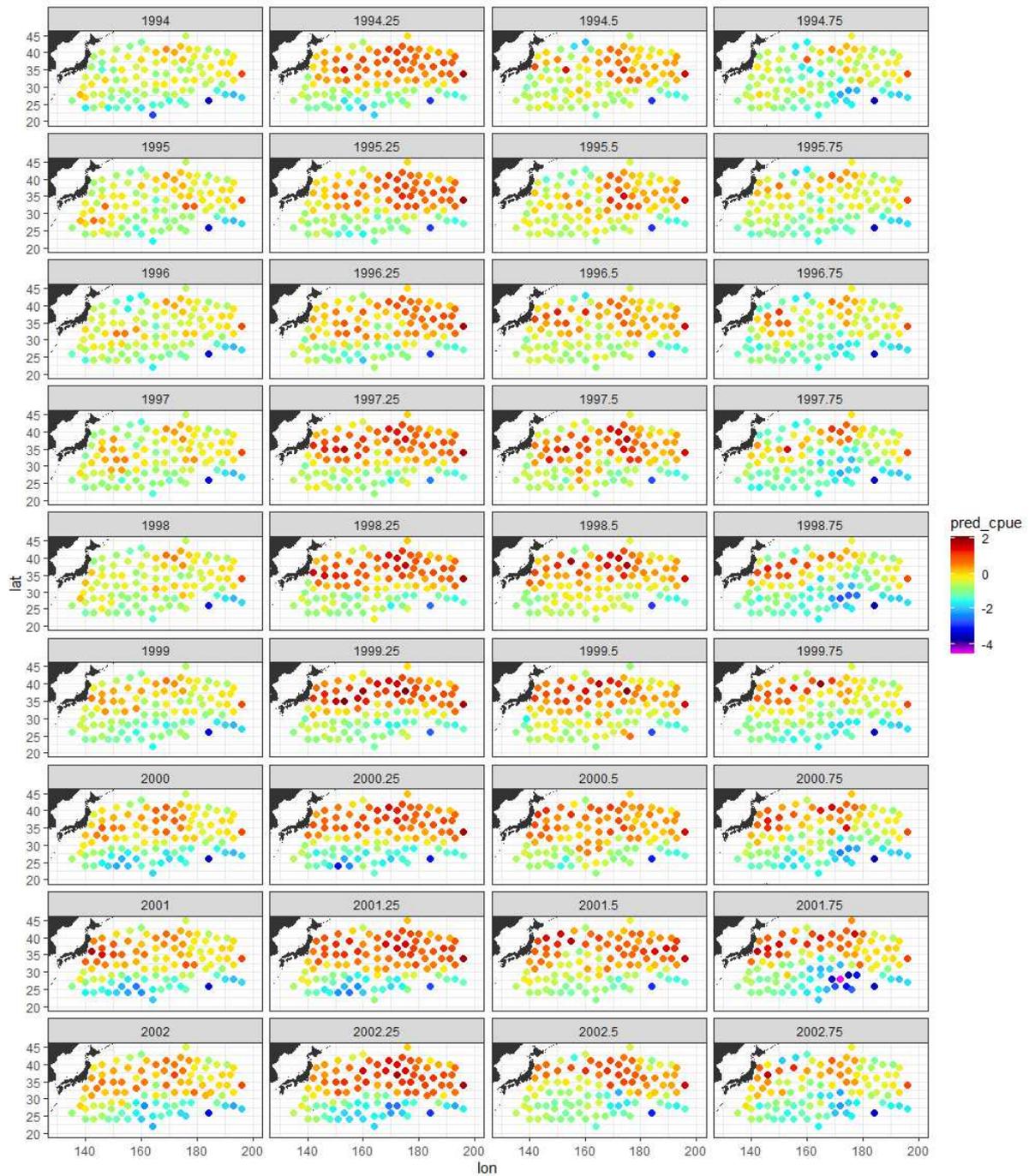


Fig. 5 Year- and season- specific spatial distribution of log-scaled predicted CPUE for blue shark from 1994 to 2002. Each point denotes the location of knot.

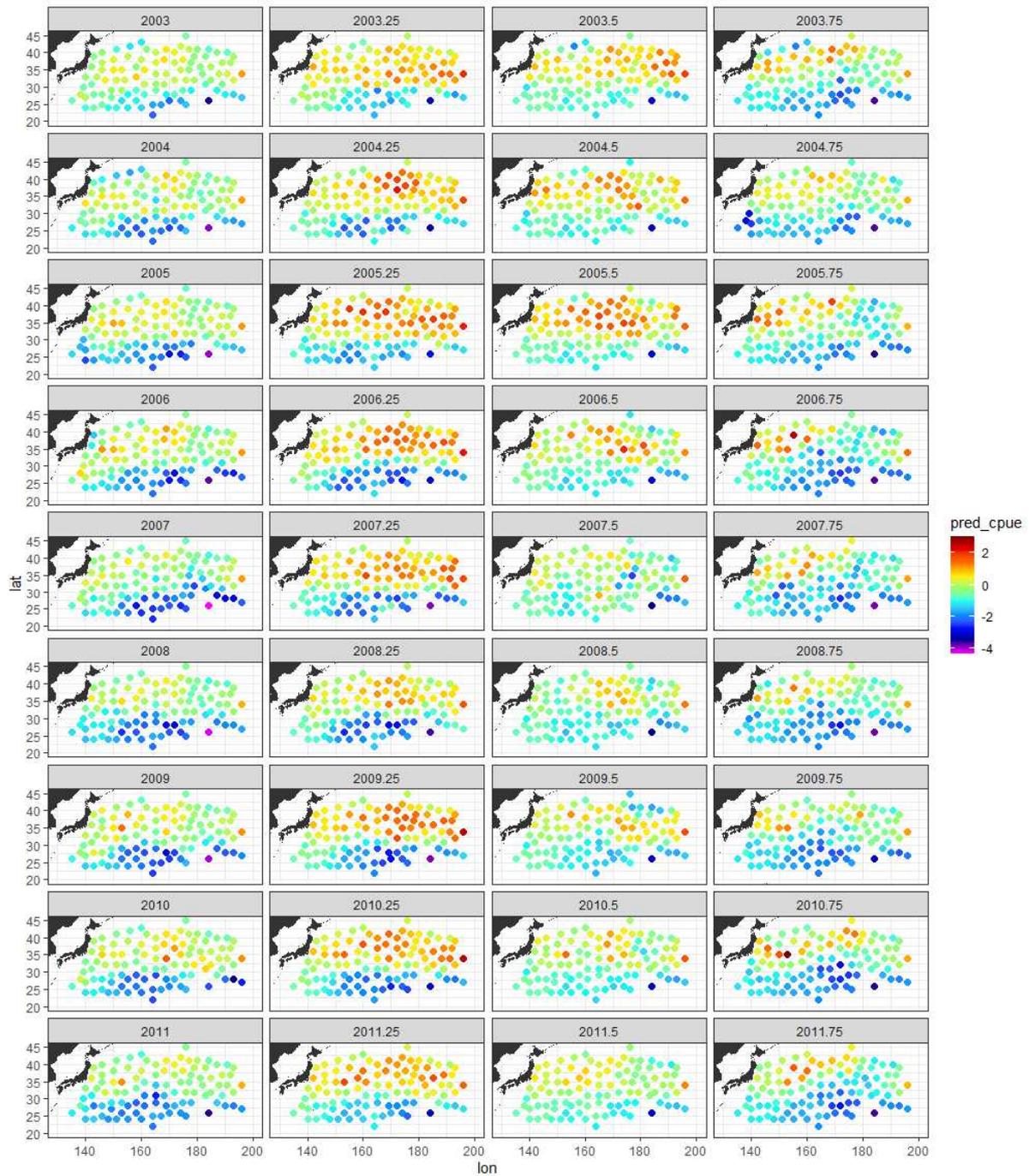


Fig. 6 Year- and season- specific spatial distribution of log-scaled predicted CPUE for blue shark from 2003 to 2011. Each point denotes the location of knot.

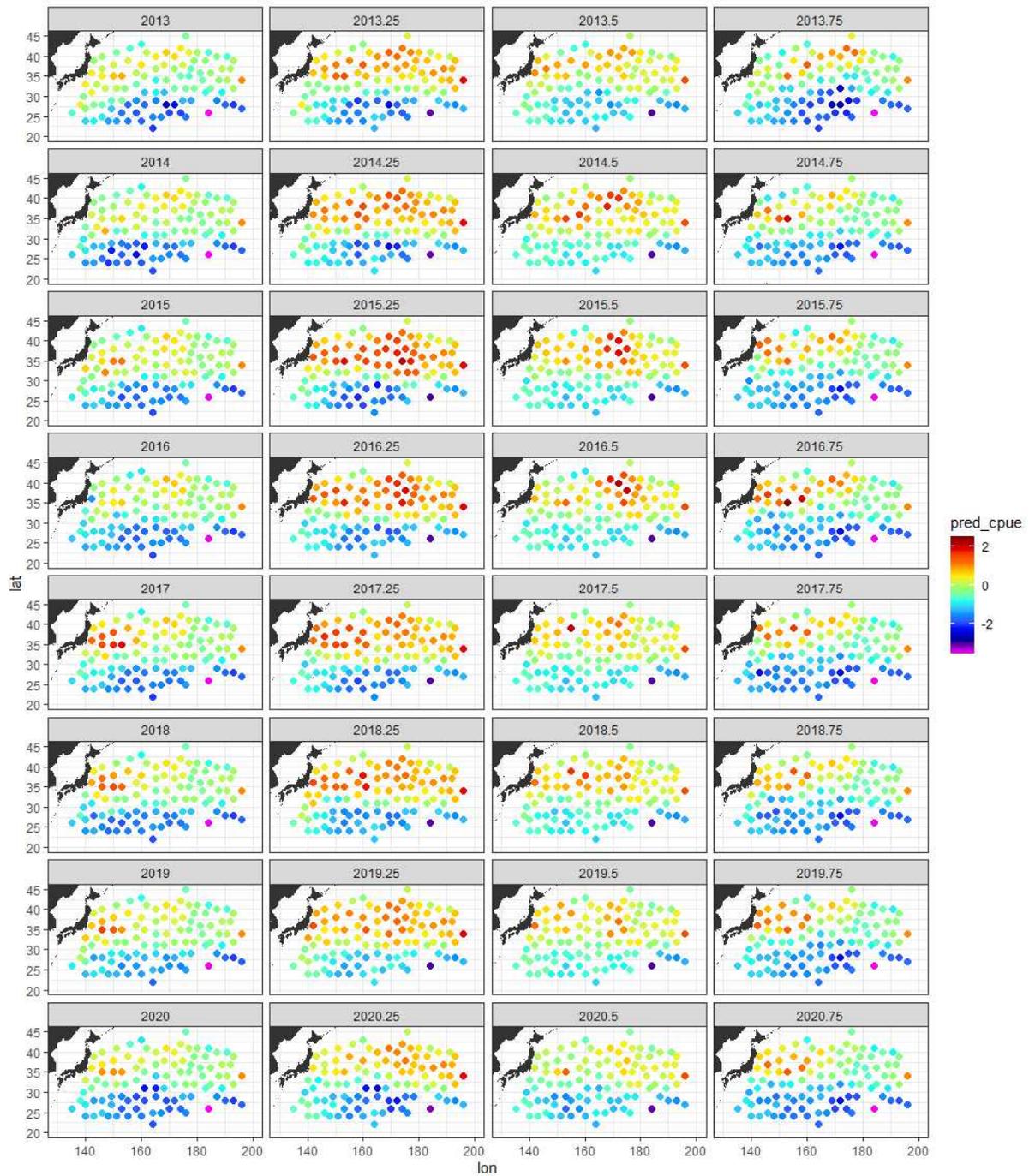


Fig. 7 Year- and season- specific spatial distribution of log-scaled predicted CPUE for blue shark from 2012 to 2020. Each point denotes the location of knot.

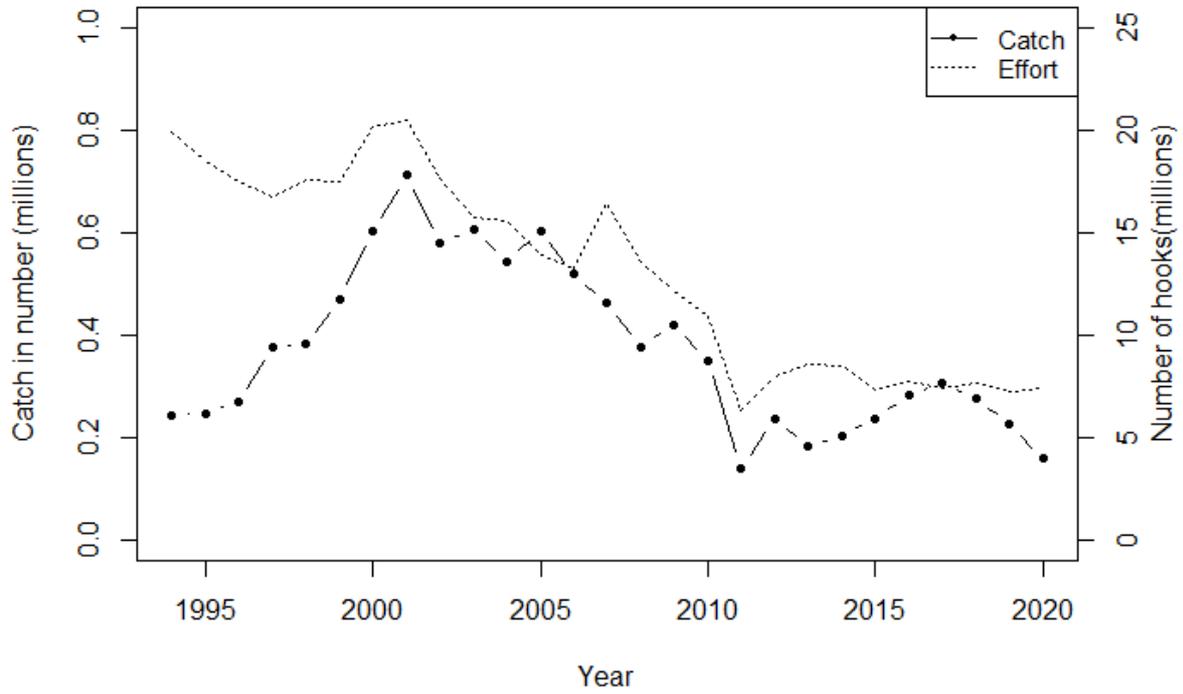


Fig. 8 Annual change in catch number (millions) of blue sharks and fishing effort (number of hooks in millions).

Appendix

Temporal (year-season) changes in the distribution shift, range expansion, and predicted CPUEs

The temporal changes in the location in Eastings and Northings indicated periodic fluctuations (**Fig. A1**). The centroid of the population's distribution shifted from southwest to northeast in Q1 and Q2 and vice-verse in Q3 and Q4. These results are synchronized with the movements of the Japanese shallow-set longliner who seasonally changes their operational areas in accordance with the target shift from swordfish to blue shark and vice versa (Hiraoka et al., 2016; Kai and Shiozaki, 2016). For the eastings, the centroid of the population's distribution clearly showed a substantial change in the locations toward the western water after tsunami attack on March 11th, 2011 (**Fig. A1**) because Japanese shallow-set longliner changed their operational area from far-seas to the coastal and offshore areas off Japan in this period due to the effect of the disaster. Meanwhile, the temporal changes in the location in Northings (**Fig. A1**) indicated that the centroid of the population's distribution gradually shifted from south to north in 1990s. These results suggested that the Japanese shallow-set longliner changed their main operational area from south to north in accordance with the target shift to blue sharks due to high demands of Asian market of shark's meats and fins in 1990s and 2000s.

The temporal changes in the effective area occupied showed periodic fluctuations (**Fig. A2**). Overall, the range expansion was remarkable in Q1, while the range contraction was remarkable in Q4. Notable, the large range contractions were observed in Q4 in 2006, 2010, and 2016. However, the reasons for the sharp decline were unclear.

The temporal changes in the predicted CPUE indicated periodic fluctuations (**Fig. A3**). The highest CPUE was Q2. This result is reasonable because the Japanese shallow-set longliner changes their target species from swordfish to blue shark from Q1 to Q2.

Appendix figures

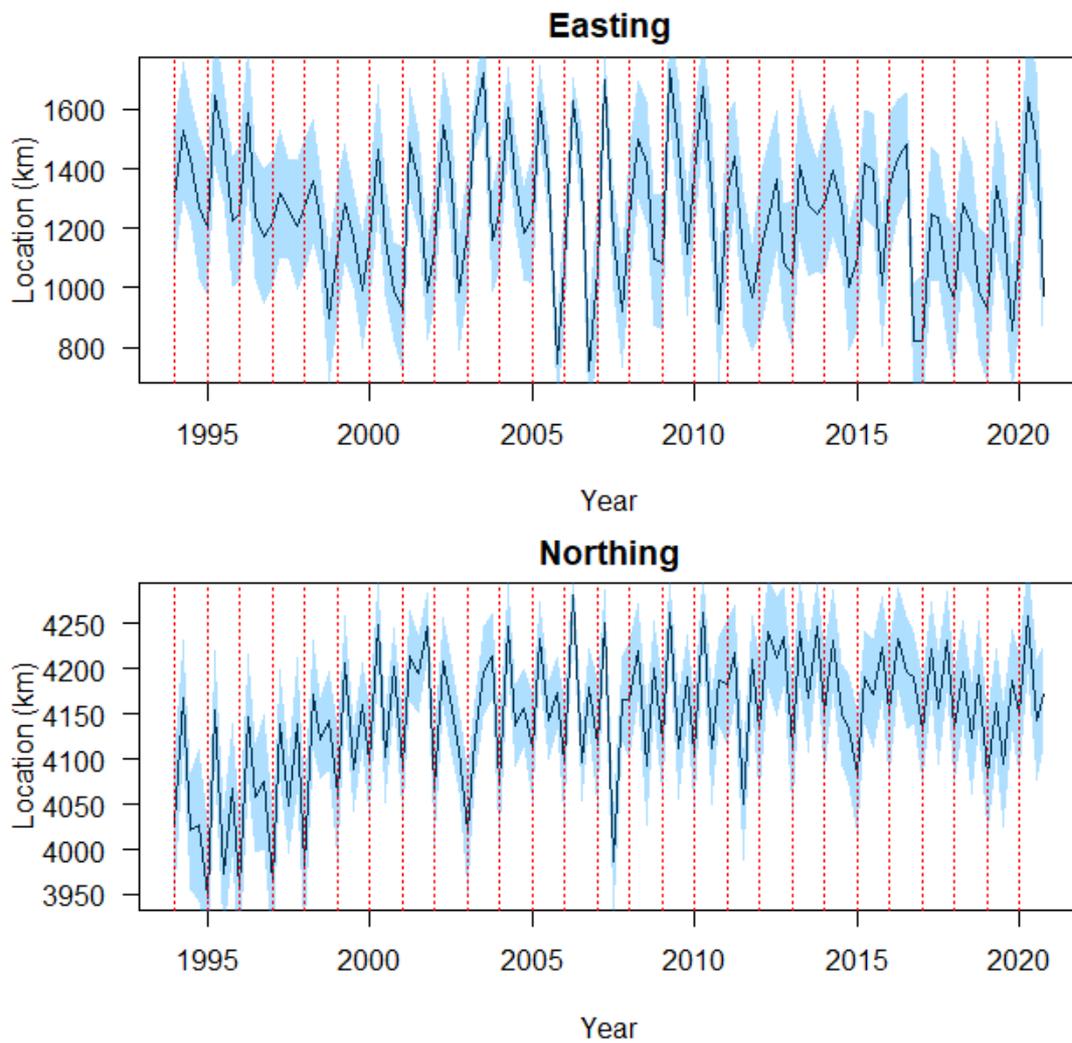


Fig. A1 Year-season changes in the centroid of the population's distribution (location in Eastings and Northings of each knot) for 1994-2020 with 95% confidence intervals (light blue shades). Upper panel denotes the movement of East-West and lower panel denotes the movement of North-South.

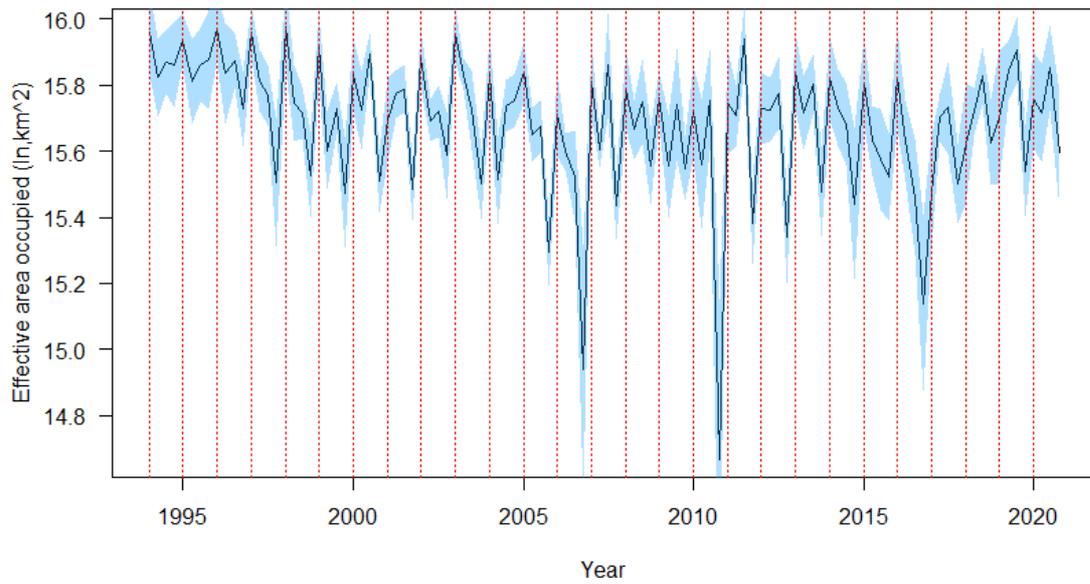


Fig. A2 Year-season changes in the effective area occupied for 1994-2020 with 95% confidence intervals (light blue shades).

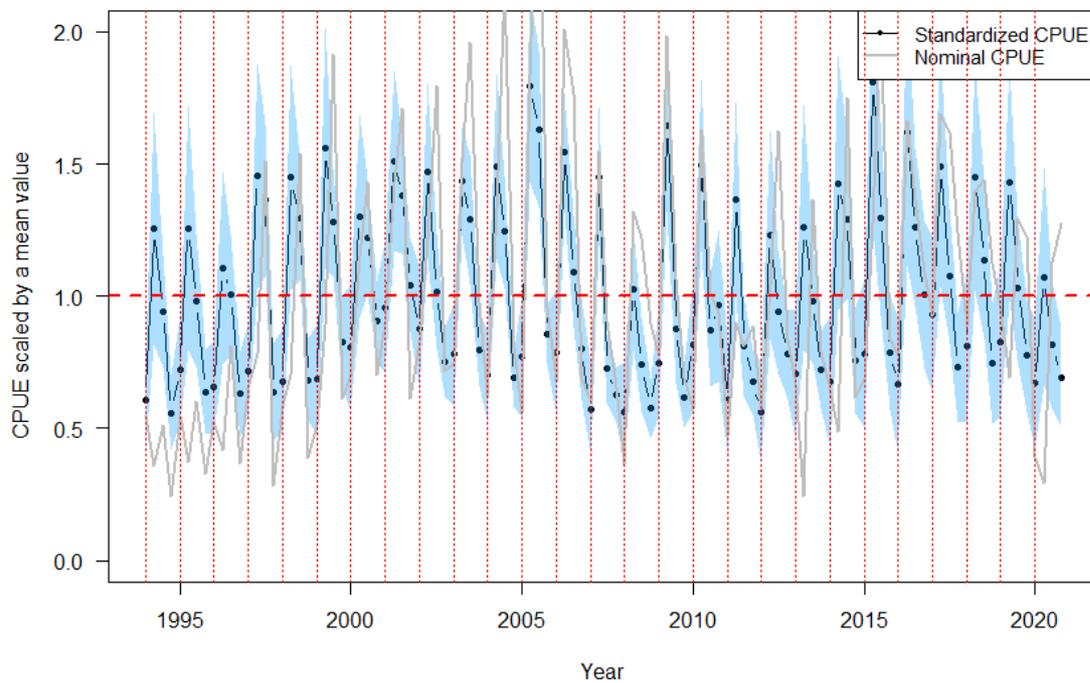


Fig. A3 Year-season changes in the predicted CPUE relative to its average. Gray solid line denotes nominal CPUE relative to its average, shadow denotes 95% confidence intervals, and horizontal dotted line denotes mean of relative values (1.0).