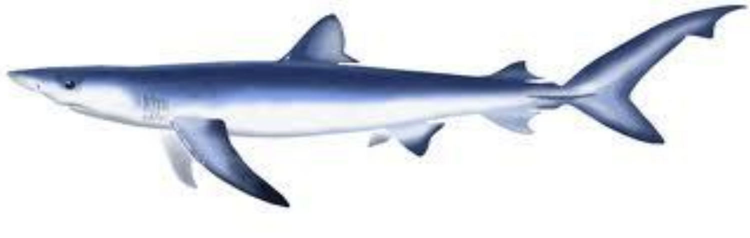
ISC/25/SHARKWG-1/03

Spatio-temporal model for CPUE standardization: Application to blue shark caught by longline of Japanese research and training vessels in the western and central 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 research and training vessels (JRTVs) longline fishery in the western and central North Pacific up to 2023, using the same methodology applied in 2021. A statistical filtering method was employed to remove unreliable set-by-set data collected by JRTVs after the 2000s. The nominal CPUE of the JRTVs was then standardized using a spatio-temporal generalized linear mixed model (GLMM) to provide annual changes in the abundance indices in the North Pacific. The predicted abundance indices of blue shark revealed a downward trend until 2008, followed by an upward trend thereafter, which is similar to trend observed in 2021. The CPUE trends predicted from the fishery-independent data widely collected in the North Pacific provide very useful information about the abundance in this region.

Introduction

For the stock assessment in 2022, Japan provided a newly developed standardized CPUE of blue shark (*Prionace glauca*) caught by Japanese research and training vessels (JRTVs) longline fishery from 1994 to 2020 (Kai, 2021). A statistical filtering method was employed to remove unreliable set-by-set data collected by JRTVs after the 2000s (Kai, 2019). The nominal CPUE was then standardized using a spatio-temporal GLMM (Thorson, 2019). The predicted abundance indices of blue shark revealed a downward trend until 2008, followed by an upward trend, with a stable trend in recent years (Kai, 2021).

In the previous stock assessment in 2022, a composite-CPUE was derived from a Dynamic Factor Analysis (DFA) by combining three indices: the JRTVs longline index, the Hawaii deep-set longline index and the Taiwanese large-scale longline index, which were derived from observer data or research and training vessel data. These indices showed similar annual trends in CPUE (ISC, 2022). The combined index represents fisheries that primarily target tunas through deep-setting behavior across a broad range of the central Pacific Ocean and typically select larger individuals compared to the Japanese shallow-set index.

The JRTVs data were collected from research vessels belonging to, or chartered by, national/prefectural fisheries research institutes, and vocational training vessels attached to fisheries high schools throughout Japan. The data are treated as one dataset because both types of vessels are not commercial fisheries, their operation overlap, and their gear configuration is similar (i.e., deep-set fishery). Since the JRTVs data are fishery-independent, it is expected that the data would be accurately reported and have no issue of target shifts. However, Clarke et al. (2011) raised an issue regarding the lower reporting ratio of sharks after the 2000s. Kai (2019) mentioned that the main reason for the under-reporting between 2001 and 2013 is possibly due to a reduced recording of discarded sharks resulting from a revision of the input form in 2000 (an increase in input items). In 2013, the Japan Fishery Agency instructed accurate recording of the number of all sharks caught, including discards, which led to an increased reporting rate of sharks after 2013.

The objective of this working paper is to update standardized CPUE of blue shark caught by JRTVs longline fishery in the western and central North Pacific Ocean up to 2023 and to provide the annual relative abundance index for the indicator analysis of blue shark in 2025. First, temporal changes in the reporting rate are analyzed, and unreliable set-by-set data are removed using a statistical filtering method. Then, the nominal CPUE is standardized using a spatio-temporal GLMM for the filtered data.

Materials and Methods

The author used a similar statistical filtering method and spatio-temporal model as those used in the paper by Kai (2019, 2021).

Data sources

Set-by-set longline logbook data collected from JRTVs in the western and central North Pacific Ocean from 1992 to 2023 were used. The data include information on shark species, operation time, catch numbers, number of hooks, number of branch line between floats (HBF), location of sets by latitude-longitude resolution of 1° ×1°, and trip identity. As the JRTVs mostly use deep sets (i.e., 6-16 HBF), two types of deep sets (shallower and deeper deep sets; HBF < 11 and 10 < HBF) were used in this analysis. The four seasons (quarters (Q) 1 to 4) of the year were defined as follows: Q1: JAN-MAR; Q2: APR-JUN; Q3: JUL-SEP; Q4: OCT-DEC.

Data filtering

Preliminary filtering was conducted to remove incomplete and insufficient data that had little or no information about HBF and locations (latitude and longitude), a number of hooks that were less than 800, HBF that were less than 6 (i.e., shallow sets), and operations conducted in waters other than the North Pacific Ocean. Additionally, follow-up filtering was conducted to remove unreliable set-by-set data caused by under-reporting of actual shark catches. The author used a statistical filtering method based on the information on shark presence in the catch (Hoyle et al., 2017; Kai, 2019) and applied it to JRTV data from 2001 to 2013 to address a clear decline in annual reporting rates during this period (**Fig. 1a**).

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 (> 94%). Second predictor was modeled using a negative binomial (NB) model to account for datasets with over-dispersion (variance/mean =14.56):

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

$$\log\,(d) = d_0(t) + \gamma(s) + \theta(s,t) + \epsilon(v) + \textstyle\sum_{j=1}^{n_j} \beta_j x_j, \qquad (1)$$

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 the 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 t), $\gamma(s)$ represents spatial variation (s), $\theta(s,t)$ represents spatio-temporal variation (station s and year t), $\epsilon(v)$ represents random variation in catchability for the vth vessel, and β_j represents the impact of covariate j with value x_j on catchability. The shallow and deep sets and three-month quarters (i.e. $n_j = 2$, $x_j = hbf \text{ and } q$) are used as covariates (changing the catchability) corresponding to Eq. (1).

The VAST (Vector Autoregressive Spatio-Temporal; version VAST_v13_0_0) software package for R (Thorson, 2019) was used to standardize the nominal CPUE. To be consistent with the period of late time series for Japanese longline fishery, the author used the JRTVs data after 1993 in the CPUE standardization. The annual abundance index I was estimated as:

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

where n_s is total number of knots (i.e., sampling location in this study) and f is fishing effort (number of hooks) at location s.

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 fits 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

Summary of data filtering

The preliminary filtering reduced the number records for this analysis from 40,019 sets to 35,841 sets. The follow-up filtering further reduced the number of records from 35,841 sets representing 1,499 trips to 32,058 sets representing 1,333 trips. The follow-up filtering appeared to be reasonable because the reduction in catch rates between 2001 and 2013 disappeared (**Fig. 1**). The differences in annual changes in number of catches, number of hooks, and nominal CPUE between the data with and without follow-up filtering are shown in **Fig. 2**.

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 (< 0.00001). The model incorporates spatial and spatio-temporal variances, as well as variation over vessels as random effects. A list of all parameters and estimates of the model is provided in **Table 1**. The predicted annual CPUE indicated a declining trend from 1994 to 2008, followed by a gradual increase until 2023 (**Fig. 3**). Uncertainty (CV) in the CPUE estimates was significantly higher in 1999 and 2020, due to lower fishing efforts (number of hooks) compared to other years (**Table 2**).

Model diagnostics

Diagnostic plots of goodness-of-fit for the model did not show any serious deviations from normality or indications of model misspecification (**Fig. 4**). These results suggest that the model fits the data well.

Discussions

This document paper updates annual abundance indices of blue sharks in the western and central North Pacific Ocean up to 2023 (**Fig. 3**). The author used the same spatio-temporal GLMM model structure as in 2021, utilizing fishery independent JRTVs data. Unreliable set-by-set data with low reporting rates of sharks in 2000s were removed using a statistical filtering method.

The predicted abundance indices of blue shark revealed a downward trend until 2008, followed by an upward trend thereafter, which is similar to trend observed in 2021. The hotspots of blue sharks appeared mainly in the temperate waters above 30 °N, as shown in past studies (Hiraoka et al., 2016; Kai et al., 2017), and in the subtropical areas off the southwest coast of Hawaii (**Fig. 5**). In these regions, adult blue sharks are widely present in low-latitude waters (Nakano and Stevens, 2008), and Hawaii-based pelagic longline vessels frequently operate (ISC, 2017).

The CPUE trends predicted from the fishery-independent data widely collected in the North Pacific provide 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., 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., Smyth, G.K. 1996. Randomized quantile residuals. J. Comput. Graph. Stat. 5, 236−244.

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.

Hoyle, S.D., Semba, Y., Kai, M., Okamoto, H. 2017. Development of Southern Hemisphere porbeagle shark stock abundance indicators using Japanese commercial and survey data. New Zealand Fisheries Assessment Report 2017/07.

ISC 2022. Stock Assessment and Future Projections of Blue Shark in the North Pacific Ocean Through 2020. SA-Report Annex12: Kona, Hawai'i, U.S.A. July 12-18, 2022

Kai, M. 2021. Spatio-temporal model for CPUE standardization: Application to blue shark caught by longline of Japanese research and training vessels in the western and central North Pacific. ISC/21/SHARKWG-2.

Kai, M. 2019. Spatio-temporal changes in catch rates of pelagic sharks caught by Japanese research and training vessels in the western and central North Pacific. Fish. Res. 216, 177−195.

Kai, M., Thorson, J.T., Piner, K.R., Maunder, M.N. 2017. Predicting the spatio-temporal distributions of pelagic sharks in the western and central North Pacific. Fish. Oceanogr. 2, 569−582

Lindén, A., Māntyniemi, A. 2011. Using the negative binomial distribution to model overdispersion in ecological count data. Ecology 92, 1414−1421.

Nakano, H., Stevens, J.D. 2008. The biology and ecology of the blue shark, *Prionace glauca*, in: Camhi, M.D., Pikitch, E.K., Babcock, E.A. (Eds.), Sharks of the open ocean: biology, fisheries and conservation. Blackwell Publishing Ltd., Oxford, pp. 140–151.

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.

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	0.0023
2	Variation over vessel	σ_{ϵ}	Fixed	1.37
3	Northings anisotropy	h_1	Fixed	1.44
4	Anisotropic correlation	h_2	Fixed	0.91
5	Parameter governing pointwise variance (Spatial random effect)	η_{γ}	Fixed	1.56
6	Parameter governing pointwise variance (Spatio-temporal (year) random effect)	η_{θ}	Fixed	0.51
7	Parameter governing autocorrelation (Spatio-temporal: year random effect)	ρ_{θ}	Fixed	1.38
8	Residual variation 1 of negative binomial model	σ_1	Fixed	0.14
9	Residual variation 2 of negative binomial model	σ_2	Fixed	0.20
10	Coefficient of hooks between floats	β_1	Fixed	-0.551
11	Coefficient of three month quarters	β_2	Fixed	0.117
12-37	Intercept for year	d_0	Fixed	Not shown
38	Vessel effect	ϵ	Random	Not shown
39	Spatial residuals	γ	Random	Not shown
40	Spatio-temporal (year) residuals	θ	Random	Not shown

Table 2. Summary of annual CPUE predicted by spatio-temporal model in this study and previous study in 2021 (Kai, 2021) 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 (Kai, 2021)	Nominal CPUE	CV	Number of hooks (millions)
1994	1.47	1.48	1.25	0.11	4.83
1995	1.38	1.44	1.01	0.12	4.63
1996	1.31	1.39	1.22	0.11	4.52
1997	1.37	1.44	1.45	0.10	4.25
1998	1.39	1.39	1.23	0.12	2.76
1999	1.53	1.44	1.55	0.22	0.86
2000	1.19	1.24	0.94	0.13	2.73
2001	1.12	1.17	1.09	0.10	2.93
2002	1.03	1.09	1.03	0.10	3.03
2003	1.04	1.05	1.31	0.11	2.75
2004	0.95	0.96	0.98	0.11	3.09
2005	0.76	0.78	0.82	0.12	2.32
2006	0.69	0.72	0.70	0.13	2.31
2007	0.67	0.64	0.71	0.15	1.51
2008	0.37	0.41	0.51	0.14	1.45
2009	0.61	0.58	0.69	0.14	0.69
2010	0.77	0.79	0.96	0.15	0.75
2011	0.63	0.66	0.83	0.16	0.83
2012	0.59	0.59	0.83	0.16	0.85
2013	0.75	0.79	0.83	0.16	1.15
2014	1.05	1.04	0.89	0.17	1.47
2015	0.80	0.83	0.82	0.16	1.24
2016	0.97	1.09	1.09	0.14	1.19
2017	0.92	1.06	1.00	0.12	1.19
2018	0.91	0.98	0.91	0.14	1.13
2019	0.93	0.98	1.20	0.16	0.91
2020	0.98	0.97	0.80	0.17	0.52
2021	1.36		0.97	0.17	0.52
2022	1.20		0.97	0.18	0.57
2023	1.23		1.43	0.17	0.44

Figures

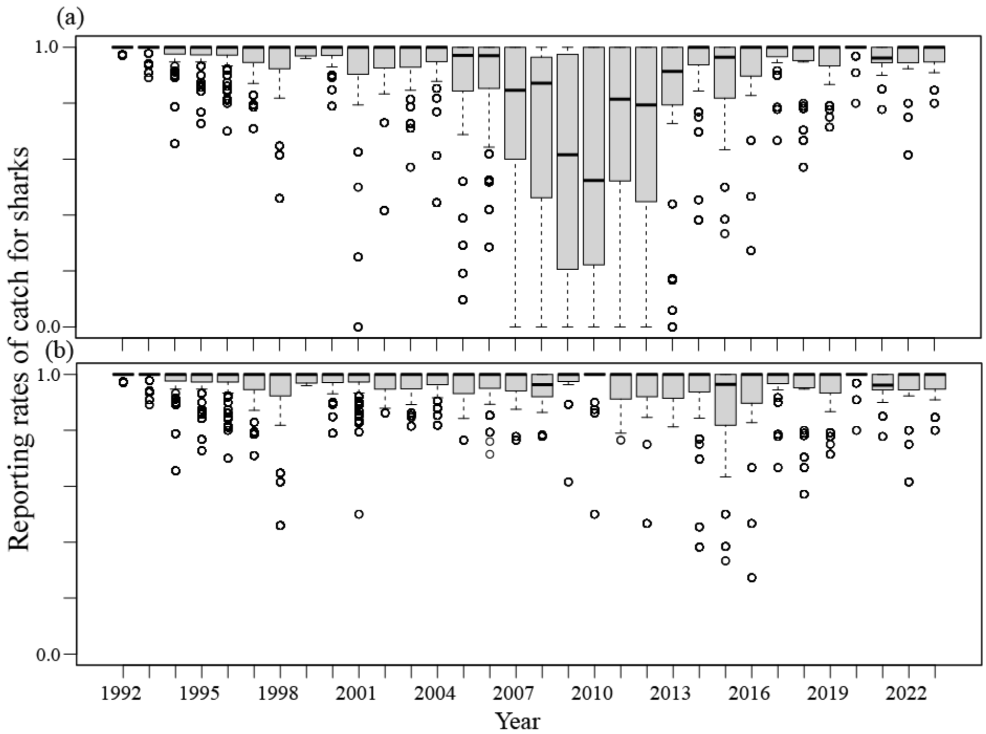


Fig. 1 Annual reporting rates of catch for sharks (a) before filtering and (b) after filtering.

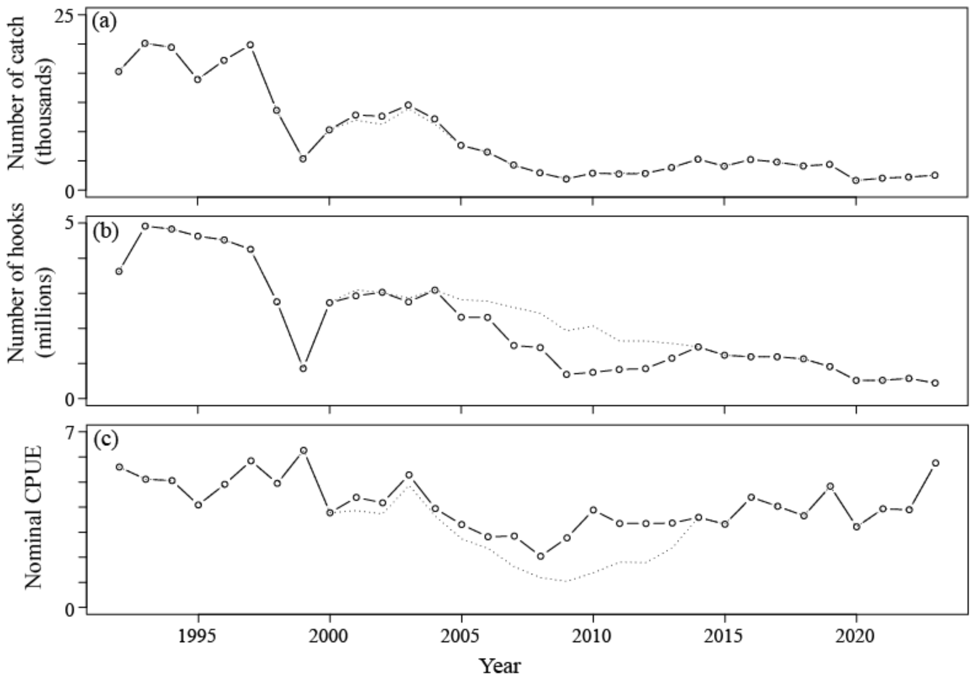


Fig. 2 Annual catch in numbers (thousands) (a), number of hooks (millions) (b), and nominal CPUE (per 1000 hooks) (c) for blue shark before filtering (broken line) and after filtering (solid line with open circle).

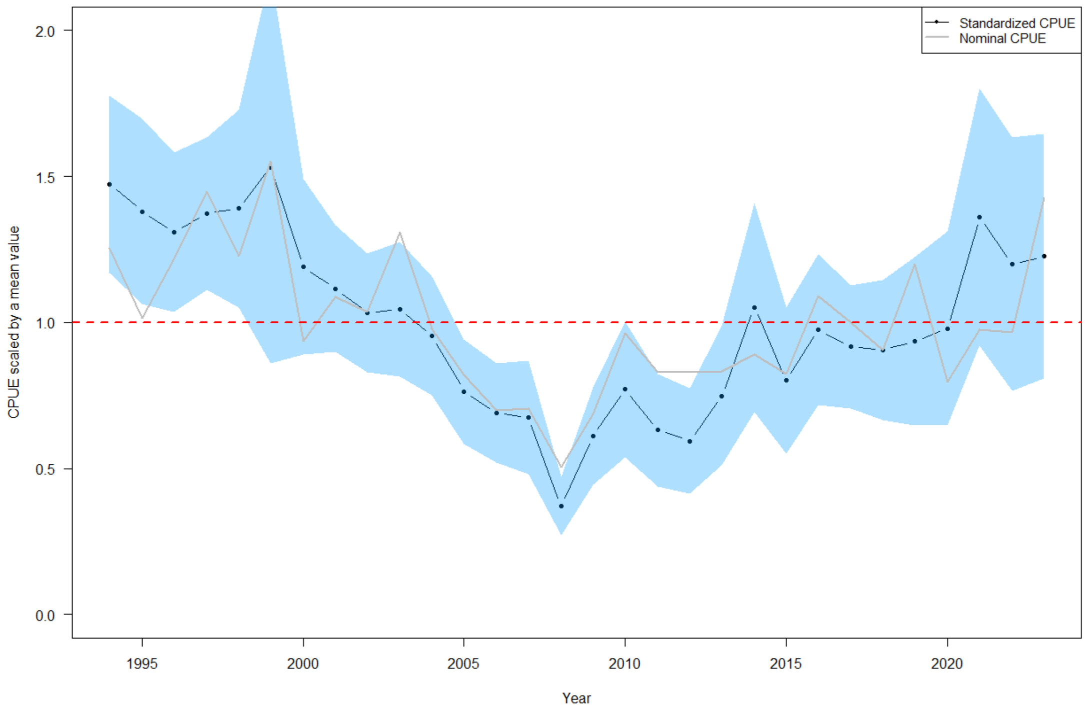


Fig. 3 Annual predicted CPUE relative to its average. Gray solid line denotes nominal CPUE relative to its average, shadow denotes 95% confidence intervals, and horizontal dotted line denotes mean of relative values (1.0).

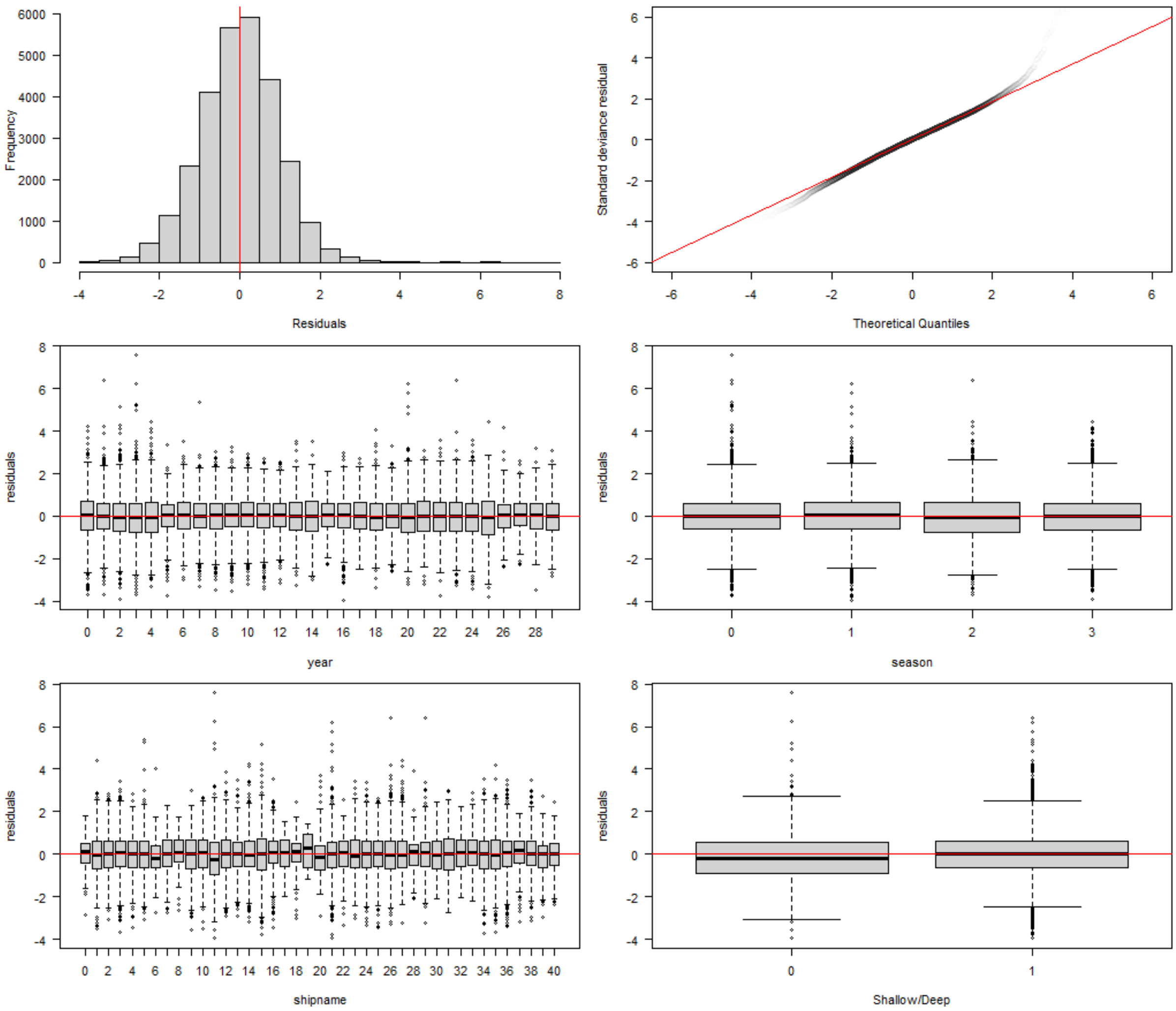


Fig. 4 Diagnostic plots of goodness-of-fit for the most parsimonious model.

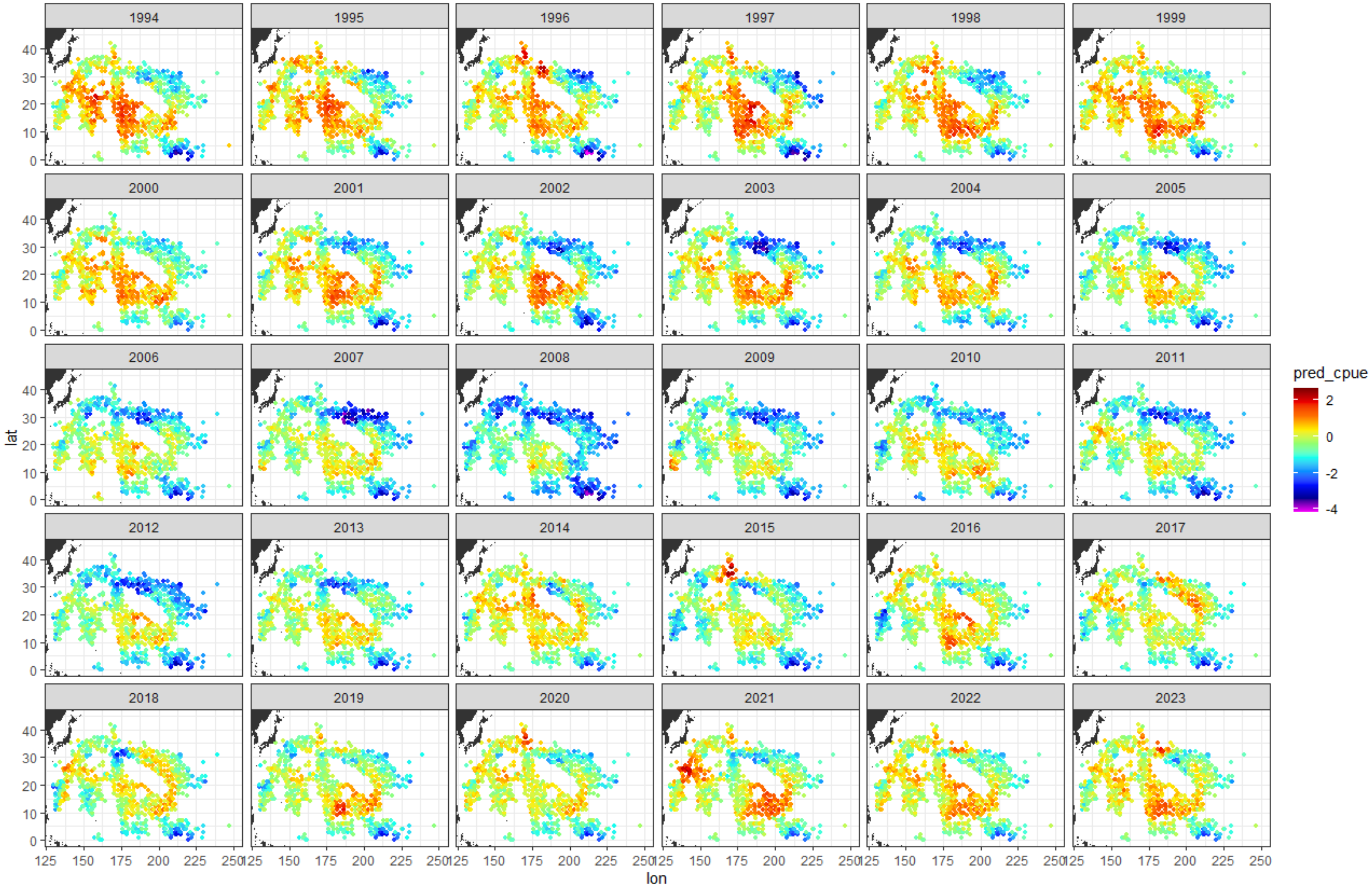


Fig. 5 Annual spatial distribution of log-scaled predicted CPUE for blue shark.