Exploring the spatio-temporal dynamics of standardized CPUE for shortfin mako shark (*Isurus oxyrinchus*) caught by the Taiwanese large-scale tuna longline fishery in the Indian Ocean

Hoang Huy Huynh[1,2,3] and Wen-Pei Tsai[2,*]

[1] Institute of Aquatic Science and Technology, National Kaohsiung University of Science and Technology, Kaohsiung, Taiwan
[2] Department of Fisheries Technology and Management, National Kaohsiung University of Science and Technology, Kaohsiung, Taiwan
[3] Research Institute for Aquaculture No.2, Ho Chi Minh 710000, Vietnam

[*] Corresponding author. E-mail: wptsai@nkust.edu.tw (W.-P. Tsai).

(Copyright: Evans Baudin)

ABSTRACT
Understanding spatiotemporal variability is essential in stock assessment and fishery conservation to accurately track changes in the distribution and abundance of fish stocks over time. This study investigates recent trends in the relative abundance of shortfin mako sharks (*Isurus oxyrinchus*) in the Indian Ocean, utilizing catch rate data from the Taiwanese large-scale longline fishery. We standardized the catch per unit effort (CPUE), defined as the number of fish caught per 1,000 hooks, using a vector autoregressive spatiotemporal (VAST) model. The results indicate that the standardized CPUE of shortfin mako sharks has remained stable, with a slight upward trend. While nominal CPUE exhibited significant fluctuations, particularly in 2005 and 2015, the standardized CPUE showed a more consistent increase, especially during 2015 and 2023. This suggests that shortfin mako shark stocks were optimally utilized between 2005 and 2023. The application of a spatiotemporal model, combined with comprehensive data from the Indian Ocean, provided valuable insights into the abundance trends of shortfin mako sharks. Future research should consider integrating environmental factors and extending the observation period to further enhance the analysis.

Keywords: CPUE standardization, Indian Ocean, shortfin mako shark, VAST model.

1. Introduction

Referring to the shortfin mako shark *(Isurus oxyrinchus)*, like other Elasmobranchs, it is highly vulnerable due to certain life history traits. These traits include long life, low fecundity, late maturity, and an extended reproductive cycle of 2–3 years (Mollet et al., 2000; Francis et al., 2005; Semba et al., 2011). Because of these characteristics, the shortfin mako shark is particularly susceptible to overfishing (Dulvy et al., 2021; Pacoureau et al., 2021). In 2019, the International Union for Conservation of Nature (IUCN) classified the shortfin mako as Endangered (EN) (Rigby et al., 2019; Pacoureau et al., 2021). Additionally, the shortfin mako shark is often caught as a bycatch in the Taiwanese large-scale tuna longline fishery (TLTL). According to the Indian Ocean Tuna Commission (IOTC), fisheries based in Taiwan are the second-largest harvesters of shortfin mako sharks in the Indian Ocean, accounting for up to 23% of the total catch (IOTC, 2019).

To address the growing concerns of international and regional fisheries management organizations (RFMOs) regarding elasmobranch conservation, it is crucial to analyze the latest trends in shark populations using tuna fishery logbooks. Catch-per-unit-effort (CPUE) indices are fundamental in fisheries science for assessing resource abundance. However, standardizing catch and effort data is necessary to reduce potential biases and provide a more accurate indicator of fishery resource abundance (Maunder and Punt, 2004; Hoyle et al., 2024). The use of spatiotemporal statistical models has become critical in modern stock assessments. These models enable precise estimates by accounting for spatial and temporal variations (Grüss et al., 2023). The Vector Autoregressive Spatio-Temporal (VAST) model (Thorson, 2019) has been found to be effective in estimating relative abundance indices for highly migratory species, including blue sharks *(Prionace glauca)* and mako sharks in the North Pacific (Kai, 2019).

Therefore, this paper aims to use the VAST model and logbook data to update the standardization estimates of shortfin mako sharks caught by the TLTL from 2005 to 2023. The resulting abundance index for the Indian Ocean shortfin mako shark population will provide valuable insights and help address research gaps in assessing the status of mako sharks in the region.

2. Materials and methods

2.1. Overview description

The research methodology includes two key steps: collecting data and using the VAST model to standardize shark CPUE, as shown in **Figure 1**.

2.2. Fisheries logbook data

The logbook dataset obtained from the Overseas Fisheries Development Council of Taiwan covers the years 2005 to 2023. It provides information on 18 species, including major tunas, billfishes, and sharks. Using this dataset, the study analyzed the catch and discard amounts of shortfin mako sharks (MSK), as well as the number of hooks used and spatial-temporal data: year, month, day, latitude (Lat), and longitude (Lon) for each fishing operation.

The year was divided into four seasons, denoted as quarters (Q1 to Q4), as follows: Q1 represented spring from January to March, Q2 represented summer from April to June, Q3 represented fall from July to September, and Q4 represented winter from October to December. Other variables included in our models are branch lines between floats, also known as hooks per basket (HPB). The gear configuration HPB (HPBC) is divided into four classes: shallow set (HPB < 6), middle (6 ≤ HPB < 10), deep (10 ≤ HPB <16), and ultra-deep (HPB ≥ 16).

Due to the extensive reach of the Taiwanese longline fishery in the Indian Ocean, these statistics are highly valuable for assessing the population status of pelagic sharks. In particular, the catch and discard data for MSK from 2005 to 2023 were used to standardize CPUE of MSK in the Taiwanese large-scale longline fishery in the Indian Ocean.

2.3. Filtering and exploration of data

Incomplete datasets were excluded before standardization if they were missing key information like latitude, longitude, and hooks. CPUE was calculated by dividing the number of MSK captured *(n)* by the number of hooks deployed per 1,000. Nominal CPUE was computed using the formula: $CPUE = \frac{\sum n_i}{\sum e_i} = \frac{\bar{n}}{e} \times 1000$, where n_i represents the catch and discard number of MSK, e_i denotes the corresponding fishing effort (number of hooks in this instance), and i refers to individual observations within the dataset. **Table 1** provides a summary of this data for the period 2005 to 2023.

Figures 2 and 3 show the geographical distribution of observed fishing effort and catch. This includes area stratification (Wu and Tsai, 2022) and nominal CPUE data for four specific regions: (1) Northwest Indian Ocean (North of 10°S, East of 70°E); (2) Northeast Indian Ocean (North of 10°S, 70°E-120°E); (3) Southwest Indian Ocean (South of 10°S, 20°E-60°E); (4) Southeast Indian Ocean (South of 10°S, 60°E-120°E).

2.4. The Vector Autoregressive Spatio-Temporal model (VAST)

In our study, we utilized the VAST model (Thorson, 2019), which is well-known for effectively handling spatiotemporal correlations, addressing changes in catch over time and space, and capturing spatial heterogeneity and autoregressive effects. The model is flexible enough to accommodate individual differences and non-normal data distributions.

By default, the VAST model is structured as a delta-generalized linear mixed model. It divides the catch probability distribution into encounter probability and expected catch rate based on catch occurrence (Thorson, 2019). To improve computational efficiency, VAST employs predefined spatial knots to assess spatial and spatiotemporal correlations. The estimation process in this study utilized the K-means algorithm. This algorithm divides all grid cells into 200 spatial knots that are fixed on a 15'×15' (arcmin) extrapolation grid, as shown in **Figure 4**. This approach was used to construct the map of MSK in this study. It is assumed that spatial and spatiotemporal random effects originate from their nearest spatial knot, following the methodology of Grüss et al. (2019).

The prediction of MSK CPUE is detailed as follows:

We use a logit-linked linear predictor to model the encounter probability *(p)* for observed CPUE (*i*).

$$logit(p_i) = \beta_1(t_i) + L_{\omega 1}\omega_1(s_i) + L_{\varepsilon 1}\varepsilon_1(s_i,t_i) + L_{\delta 1}\delta_1(v_i) + \sum_{k=1}^{n_k}\lambda_1(k)Q(i,k) + \sum_{p=1}^{n_p}\gamma_1(p)X(s_i,t_i,p)$$

Furthermore, we utilize a log-linked linear predictor to model the positive catch rate (λ) for observed CPUE *(i)*:

$$\log(\lambda_i) = \beta_2(t_i) + L_{\omega 2}\omega_2(s_i) + L_{\varepsilon 2}\varepsilon_2(s_i,t_i) + L_{\delta 2}\delta_2(v_i) + \sum_{k=1}^{n_k}\lambda_2(k)Q(i,k) + \sum_{p=1}^{n_p}\gamma_2(p)X(s_i,t_i,p)$$

In these equations, $\beta(t_i)$: intercept in year t_i $\omega(s_i)$: spatial variation at location s_i; L_ω: scaling factor (sd); $\varepsilon(s_i,t_i)$: spatiotemporal variation at location s_i in year t_i; L_ε: scaling factor (sd); $\delta(v_i)$: vessel/targeting effects on catchability, and $\delta(v_i)$~Normal(0,1); L_δ: scaling factor (sd); $Q(i,k)$: catchability covariate (*s*); $\lambda(k)$: associated catchability parameter (*s*); $X(s_i,t_i,p)$ habitat covariate (*s*); $\gamma(p)$: associated habitat parameter (*s*).

The probability of catching data *c* for sample *i*:

$$\Pr\begin{pmatrix} c_i = 0 \\ c_i > 0 \end{pmatrix} = \begin{pmatrix} 1 - p_i \\ p_i \times Lognormal(c_i \,|\, \log(\lambda_i), \sigma_m^2) \end{pmatrix}$$

Annual abundance index *I* was estimated as:

$$I(t) = \sum_{k=1}^{n_k}\left[Area(k) \times Density(k,t)\right] = \sum_{k=1}^{n_k}\left\{Area(k) \times \left[logit(p_i) \times \log(\lambda_i)\right]\right\}$$

2.5. Models selection

The final models for VAST were determined by selecting those with the lowest Akaike Information Criterion (AIC) weight (Akaike, 1973). To ensure that the models accurately represented the observed data patterns, we conducted a rigorous assessment of goodness-of-fit. Once the optimal models were selected, we performed a thorough goodness-of-fit test using the Analysis of Deviance method.

2.6. Computational procedures

The statistical analyses and visualizations for this study were performed using version 4.4.1 of the R language for statistical computing. VAST models were created using the VAST R package version 3.10.1. Graphical outputs were generated using the "ggplot2" package (Wickham et al., 2016).

3. Results and discussion

Exploring the characterized dataset

The dataset from the TLTL in the Indian Ocean between 2005 and 2023 reveals a crucial observation: a significant proportion of recorded sets, specifically 88.6%, reported zero MSK catches **(Figure 5)**. This highlights the potential influence of confounding factors on catch rates and underscores the necessity of standardization to address these factors effectively.

Selection of the best model

Multiple runs were conducted using different sets of covariates to determine the optimal VAST model. All the models demonstrated good convergence, as indicated by a positive definite Hessian matrix and a small maximum gradient **(Table 2)**. The saturated model (M-4), which included the variables Year, Lat, Lon, Quarter, and Area, was identified as the most parsimonious model based on AIC values **(Table 2)**. This model was utilized for further analyses.

Model diagnostics indicate that model M-4 performs well. The Q-Q plot demonstrates a nearly linear pattern, indicating that the residuals reasonably conform to a normal distribution.

Furthermore, the residual vs. predicted plot reveals no apparent patterns, suggesting that the model assumptions are satisfied **(Figure 6)**.

The trend of nominal CPUE and standardized CPUE

The nominal CPUE of MSK in the Indian Ocean showed significant inter-annual fluctuations, particularly in 2005 and 2015 **(Figure 7)**. However, this variability was slightly smoothed in the standardized CPUE series. Overall, the standardized CPUE of MSK caught by the TLTL showed a consistent increase, especially in 2015 and 2023 **(Figure 7)**. This trend suggests that MSK stock in the Indian Ocean may have been optimally utilized during the period from 2005 to 2023. The 95% confidence intervals of the CPUE estimates were substantially larger after 2010 **(Figure 7)**, likely due to the reduction in fishing effort (number of hooks) by longline fisheries **(Table 1)**. While there is an overall upward trend, it is important to note that the relative abundance of MSK still fluctuates rather than showing a smooth, consistent increase from 2005 to 2023. This indicates that while there are positive signs, there is still persistent variability and fluctuations.

Spatial maps of predicted CPUE of MSK

Figure 8 shows the predicted spatial patterns of MSK log density from 2005 to 2023, using the parsimonious VAST model. MSK is found predominantly in the Indian Ocean, with a wide distribution range and some consistency over the years. The annual spatial maps indicate that MSK is more abundant at higher latitudes **(Figure 8)**, while lower abundances are observed in near-equatorial regions, particularly from 2006 to 2012 in the Northeast Indian Ocean. In the Indian Ocean, hotspots of MSK are distributed more irregularly across a wider temporal and spatial context.

The spatial anisotropy for the encounter probability and positive catch-rate model components

Figure 9, based on M-4 (the selected model), shows that the spatial residual of the encounter probability (i.e., 1st linear predictor) changes less than the positive catch rate (i.e., 2nd linear predictor) over a longer distance. In other words, the ellipse representing the encounter probability is larger than the one representing the positive catch rate. The spatial correlation exhibits a higher degree of variation around the equatorial area, where both ellipses stretch along an east-west axis. This suggests that the distribution of mako sharks follows an east-west direction.

Additionally, **Figure 10** presents the coefficient-distribution-influence (CDI) plot of the spatial random effect for the VAST model. The annual influence values of the spatial random effect were slightly above one before 2010 and after 2019, as a greater proportion of data tended to be distributed in the grouped knots (grouped knots 16–20) with coefficients larger than one.

This report highlights the benefits of the spatiotemporal standardization approach (Grüss et al., 2019; Cacciapaglia et al., 2024) and identifies areas for improvement in the future. Currently, the model does not consider environmental factors. To improve results, it is recommended to include longer time series of observer data and environmental factors. For abundance indices, we suggest utilizing the predicted annual CPUEs of mako sharks caught by Taiwanese tuna longline fisheries in the Indian Ocean from 2005 to 2023. This recommendation is based on the extensive time series data, wide geographical coverage, and statistical reliability of the spatiotemporal model. The model uses spatial and temporal correlations through random effects to impute missing data (Thorson, 2019). The spatiotemporal generalized linear mixed effects model (GLMM) developed by Thorson et al. (2015) allows for interaction terms between spatial and temporal effects with high spatial resolution.

References

Akaike, H., 1973. Maximum likelihood identification of Gaussian autoregressive moving average models. Biometrika, 60(2), pp.255-265.

Cacciapaglia, C., Brooks, E.N., Adams, C.F., Legault, C.M., Perretti, C.T. and Hart, D., 2024. Developing workflow and diagnostics for model selection of a vector autoregressive spatiotemporal (VAST) model in comparison to design-based indices. Fisheries Research, 275, p.107009.

Dulvy, N.K., Pacoureau, N., Rigby, C.L., Pollom, R.A., Jabado, R.W., Ebert, D.A., Finucci, B., Pollock, C.M., Cheok, J., Derrick, D.H. and Herman, K.B., 2021. Overfishing drives over one-third of all sharks and rays toward a global extinction crisis. Current Biology, 31(21), pp.4773-4787.

Francis, M.P. and Duffy, C., 2005. Length at maturity in three pelagic sharks (Lamna nasus, Isurus oxyrinchus, and Prionace glauca) from New Zealand.

Grüss, A., McKenzie, J.R., Lindegren, M., Bian, R., Hoyle, S.D. and Devine, J.A., 2023. Supporting a stock assessment with spatio-temporal models fitted to fisheries-dependent data. Fisheries Research, 262, p.106649.

Grüss, A., Walter III, J.F., Babcock, E.A., Forrestal, F.C., Thorson, J.T., Lauretta, M.V. and Schirripa, M.J., 2019. Evaluation of the impacts of different treatments of spatio-temporal variation in catch-per-unit-effort standardization models. Fisheries Research, 213, pp.75-93.

Hoyle, S.D., Campbell, R.A., Ducharme-Barth, N.D., Grüss, A., Moore, B.R., Thorson, J.T., Tremblay-Boyer, L., Winker, H., Zhou, S. and Maunder, M.N., 2024. Catch per unit effort modelling for stock assessment: A summary of good practices. Fisheries Research, 269, p.106860.

IOTC, 2019. OQS Reports. Retrieved 2024/6/10, http://iotc.org/oqs.

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. Fisheries Research, 216, pp.177-195.

Maunder, M.N. and Punt, A.E., 2004. Standardizing catch and effort data: a review of recent approaches. Fisheries research, 70(2-3), pp.141-159.

Mollet, H.F., Cliff, G., Pratt, H.L. and Stevens, J.D., 2000. Reproductive biology of the female shortfin mako, Isurus oxyrinchus Rafinesque, 1810, with comments on the embryonic development of lamnoids. Fishery Bulletin, 98(2), pp.299-318.

Pacoureau, N., Rigby, C.L., Kyne, P.M., Sherley, R.B., Winker, H., Carlson, J.K., Fordham, S.V., Barreto, R., Fernando, D., Francis, M.P. and Jabado, R.W., 2021. Half a century of global decline in oceanic sharks and rays. Nature, 589(7843), pp.567-571.

Rigby, C.L.; Barreto, R.; Carlson, J.; Fernando, D.; Fordham, S.; Francis, M.P.; Jabado, R.W.; Liu, K.-M.; Marshall, A.; Pacoureau, N.; et al. Isurus oxyrinchus. The IUCN Red List of Threatened Species. 2019. Available online: https://www.iucnredlist.org/species/39341/2903170 (accessed on 4 August 2024).

Semba, Y., Aoki, I. and Yokawa, K., 2011. Size at maturity and reproductive traits of shortfin mako, Isurus oxyrinchus, in the western and central North Pacific. Marine and Freshwater Research, 62(1), pp.20-29.

Thorson, J.T., 2019. Guidance for decisions using the Vector Autoregressive Spatio-Temporal (VAST) package in stock, ecosystem, habitat and climate assessments. Fisheries Research, 210, pp.143-161.

Thorson, J.T., Shelton, A.O., Ward, E.J. and Skaug, H.J., 2015. Geostatistical delta-generalized linear mixed models improve precision for estimated abundance indices for West Coast groundfishes. ICES Journal of Marine Science, 72(5), pp.1297-1310.

Wickham, H., Chang, W. and Wickham, M.H., 2016. Package "ggplot2". Create elegant data visualisations using the grammar of graphics. Version, 2(1), pp.1-189.

Wu. X.H., and Tsai, W.P., 2022. Update on the CPUE standardization of the shortfin mako shark caught by the Taiwanese large-scale tuna longline fishery in the Indian Ocean. IOTC–2022–WPEB18-20.

Table 1. Summary logbook information from MSK used in this study summarized.

Year	No. of Hooks	No. of Sets	No. of MSK catches and discards
2005	229,125,876	72,206	11,358
2006	111,539,175	34,699	5,842
2007	141,462,466	44,026	4,181
2008	102,533,017	31,810	6,196
2009	129,191,560	40,105	7,932
2010	97,638,819	29,863	5,533
2011	73,003,298	22,551	5,093
2012	76,970,711	25,285	4,398
2013	75,819,812	23,724	3,807
2014	58,376,963	18,475	2,522
2015	70,899,449	22,537	3,089
2016	101,592,087	31,567	6,508
2017	99,408,067	29,983	7,924
2018	93,070,520	28,034	6,864
2019	97,308,263	28,692	6,282
2020	89,664,075	26,584	6,945
2021	90,950,835	27,262	6,483
2022	98,697,865	30,459	5,937
2023	99,205,141	29,942	7,017
Total	1,936,457,999	597,804	113,911

Table 2. Summary of model structure and outputs among different models. All models include fixed effects. Model selection information, including deviance, and AIC values, for the VAST model of the MSK caught by TLTL in the Indian Ocean during 2005-2023. The ΔAIC indicates the decrease in AIC compared to the best-fitting model.

Model	Model structure covariates	Number of parameters	Deviance	AIC	ΔAIC	Maximun gradient
M-1	Year + Lat + Lon	47	396724.3	376542.7	8.8	< 0.0001
M-2	Year + Lat + Lon + Quarter	49	396724.1	376546.5	12.6	< 0.0001
M-3	Year + Lat + Lon + HPBC	49	396726.5	376545.7	11.8	< 0.0001
M-4	Year + Lat + Lon + Quarter + Area	51	396721.4	376533.9	0	< 0.0001
M-5	Year + Lat + Lon + Quarter + Area + HPBC	53	396723.9	376536.7	2.8	< 0.0001

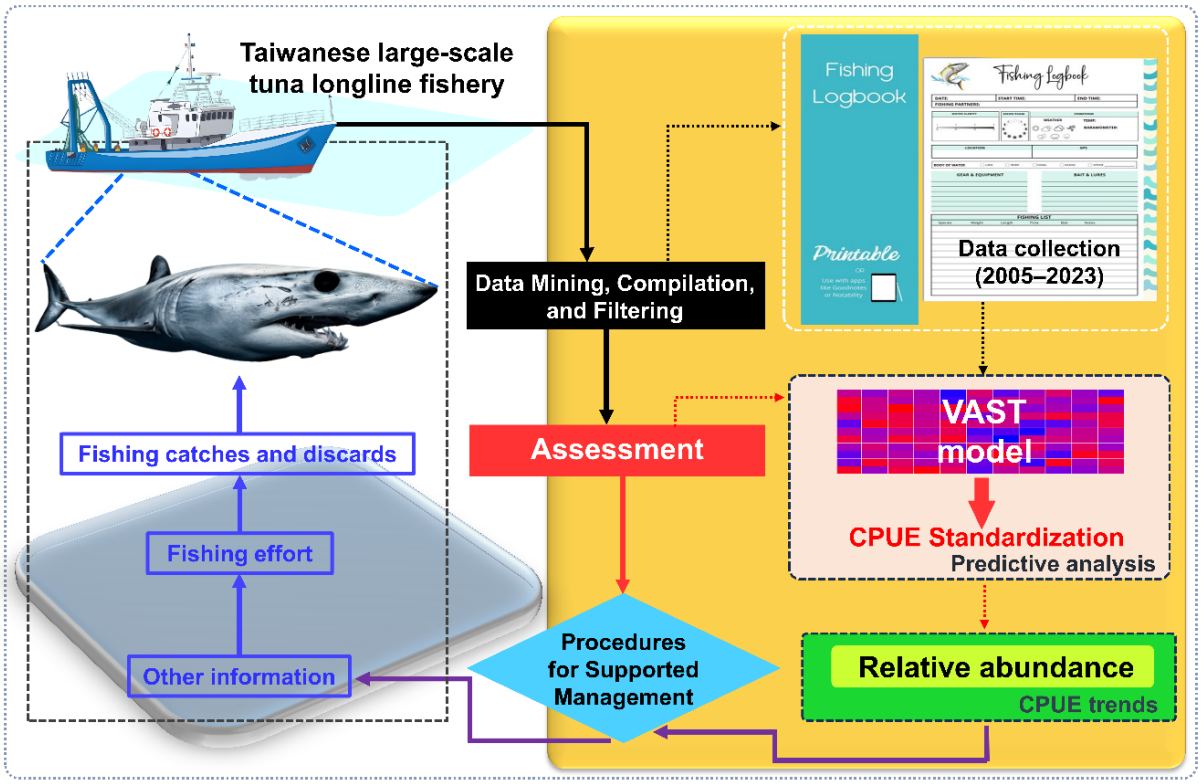

Figure 1. A framework illustrating the key steps in standardized CPUE, with photo credit to Evans Baudin for MSK.

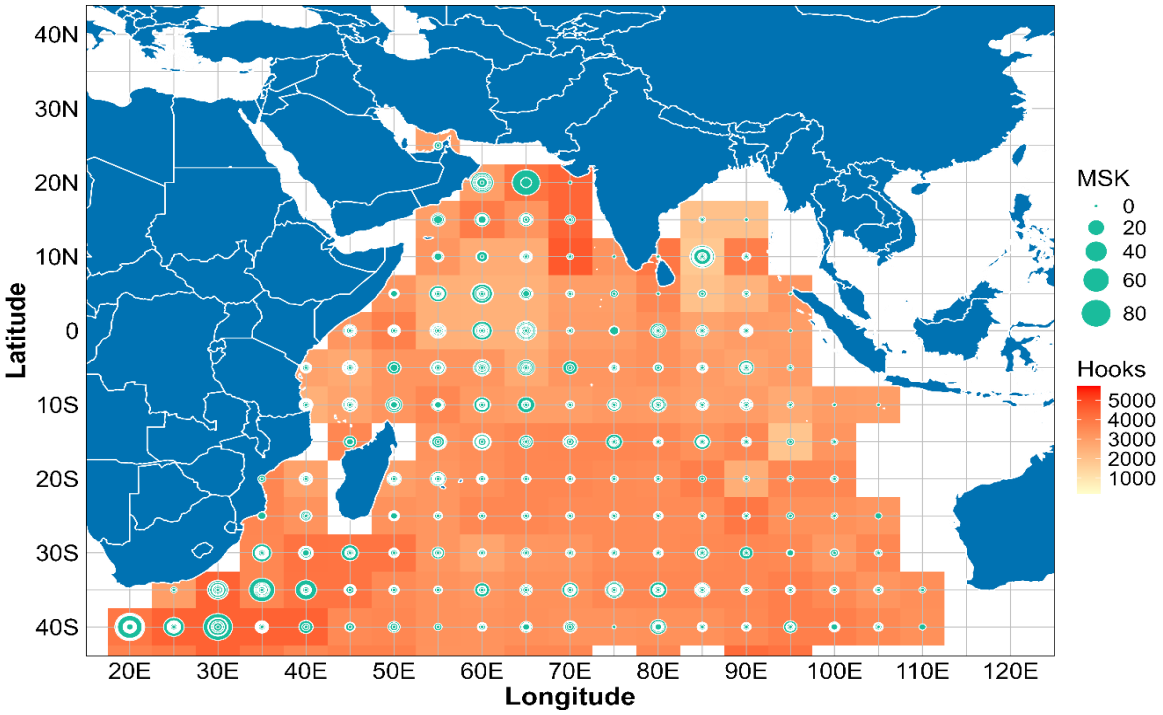

Figure 2. The geographical distribution of logbook fishing effort (measured by the number of hooks), and logbook fishing catch and discard (measured by the number of individuals) are presented for the Taiwanese large-scale longline fishery operating in the Indian Ocean from 2005 to 2023.

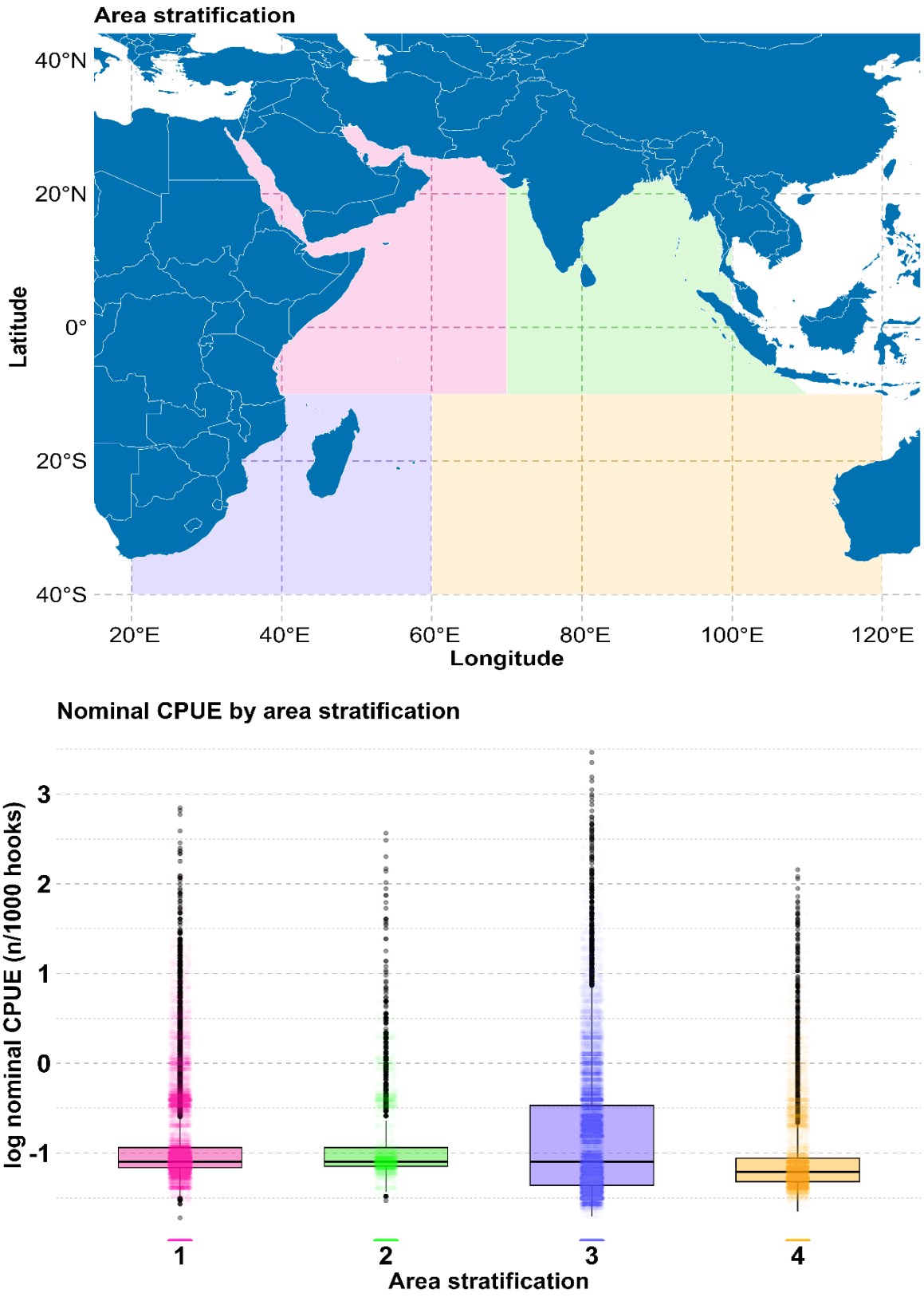

Figure 3. Area stratification and log nominal CPUE data for each area stratum, are presented for the Taiwanese large-scale longline fishery operating in the Indian Ocean from 2005 to 2023. The colors used to indicate each area stratification correspond to log nominal CPUE data; for example, the color yellow represents the first area.

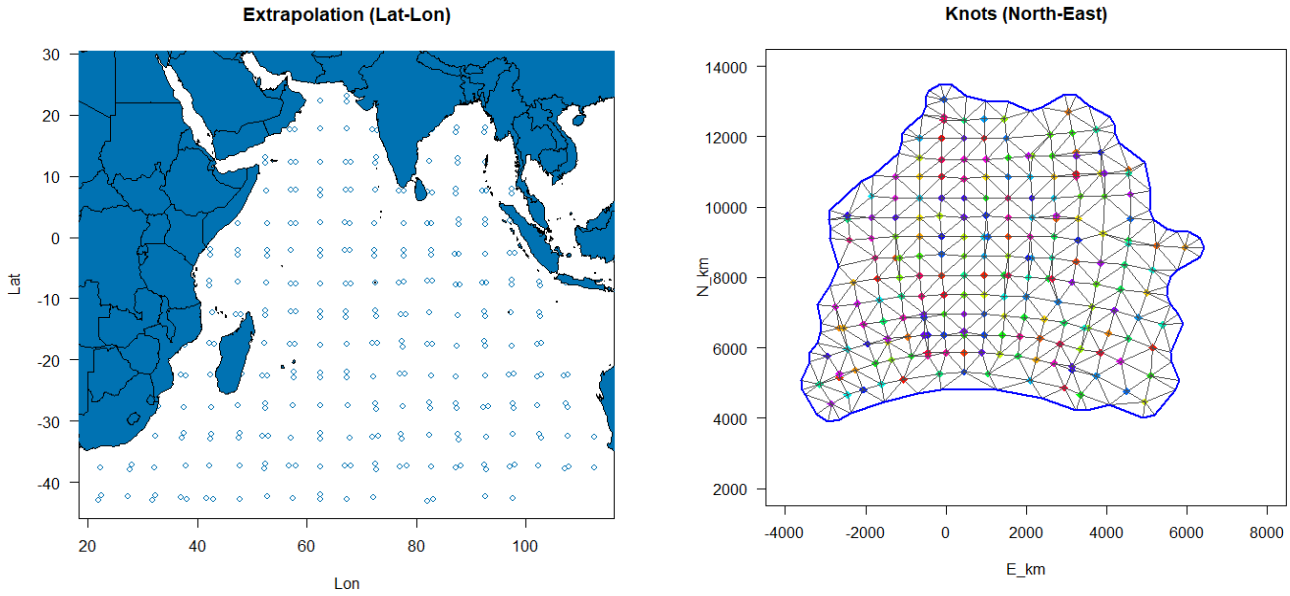

Figure 4. (Left) Map of extrapolation grid cell barycenters (15'×15') and (Right) distribution of the 200 core knots within VAST, with northing (N_km) indicating northward distance and easting (E_km) representing eastward distance, specifically at sea.

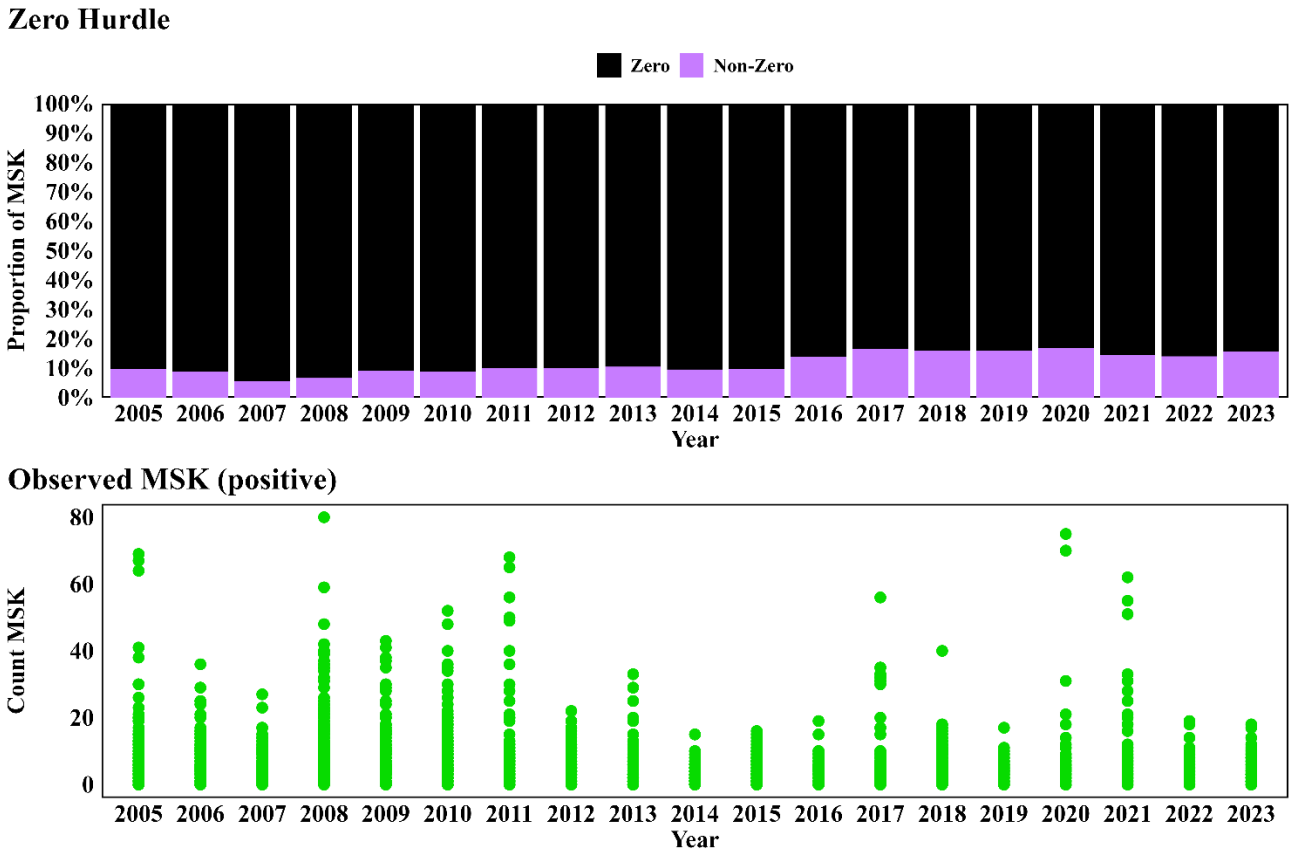

Figure 5. Distribution of MSK caught numbers (zero and positive values) in the Taiwanese large-scale longline fishery in the Indian Ocean.

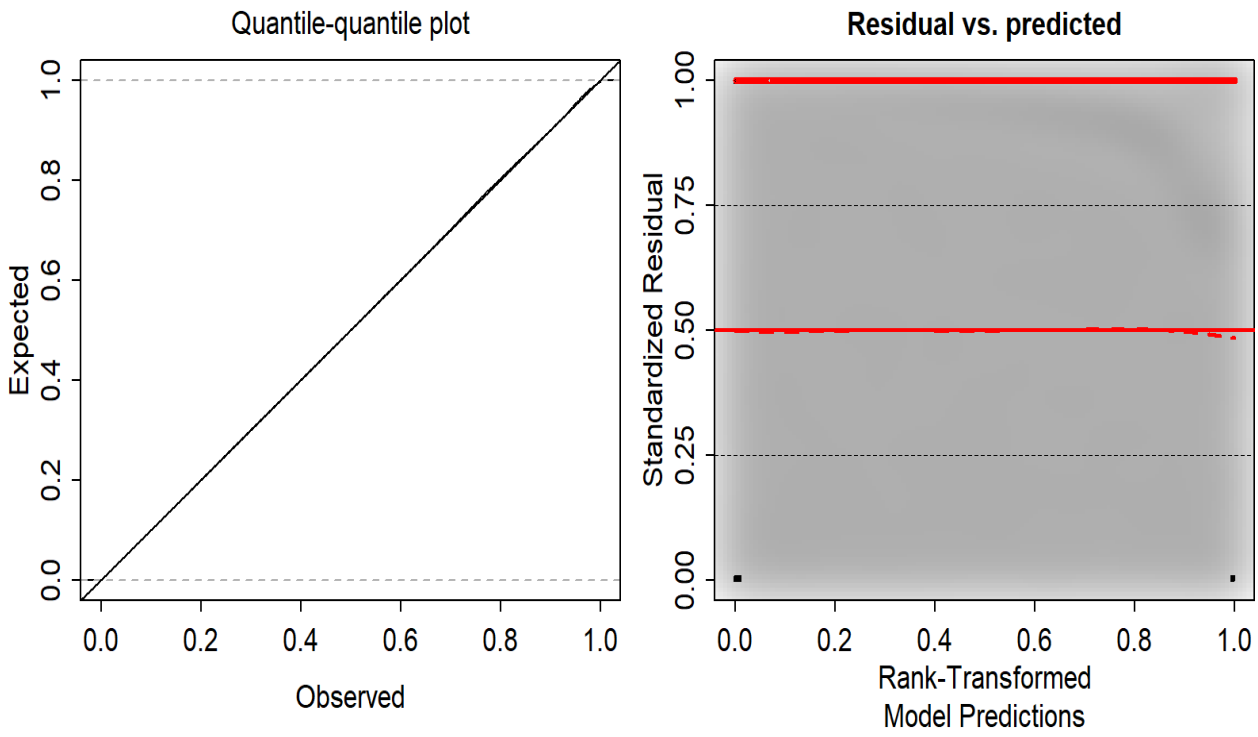

Figure 6. Diagnostic results for VAST models applied to longline MSK dataset.

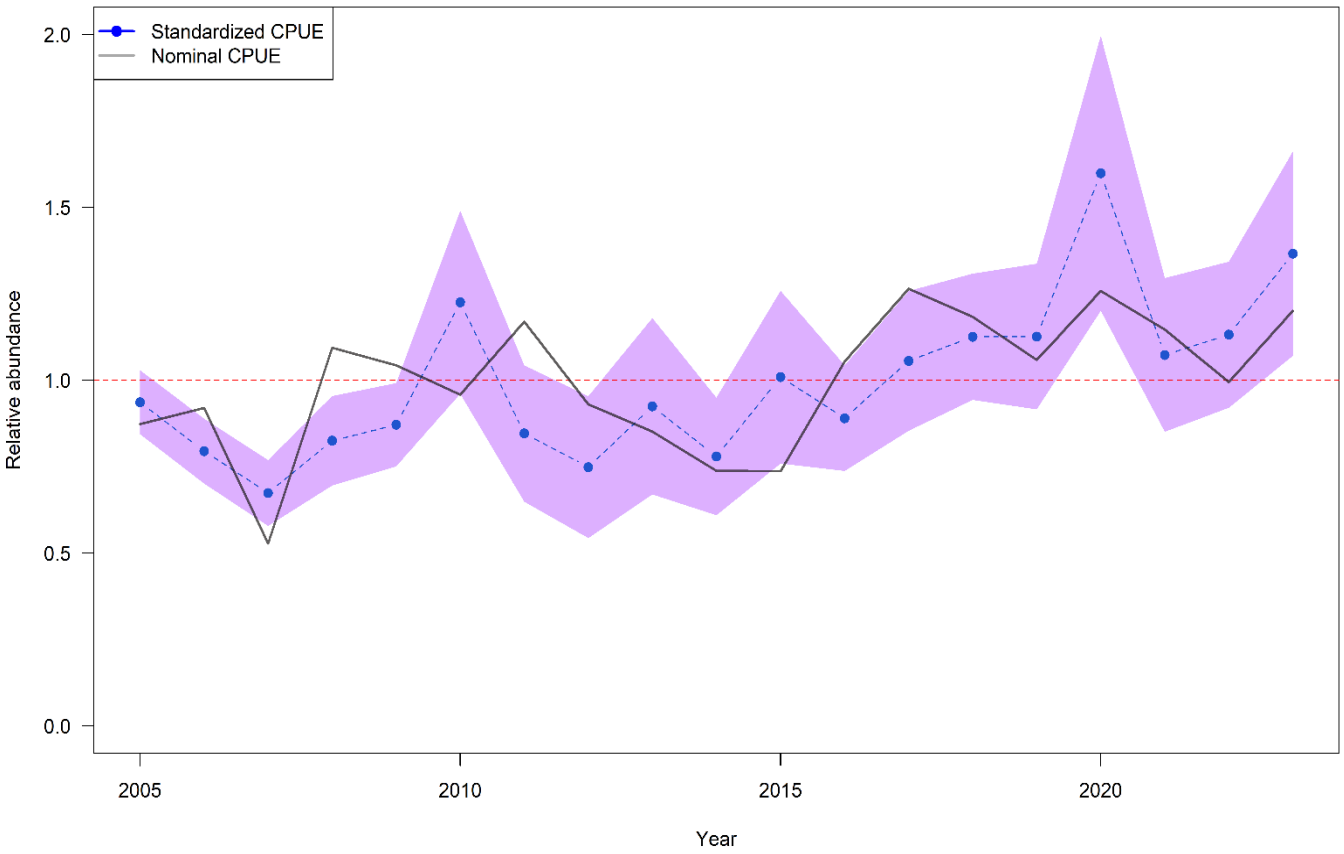

Figure 7. Annual relative nominal CPUE and standardized CPUE compared to their average. The shaded area represents the 95% confidence intervals, and the horizontal dotted red line indicates the mean of relative values (=1.0).

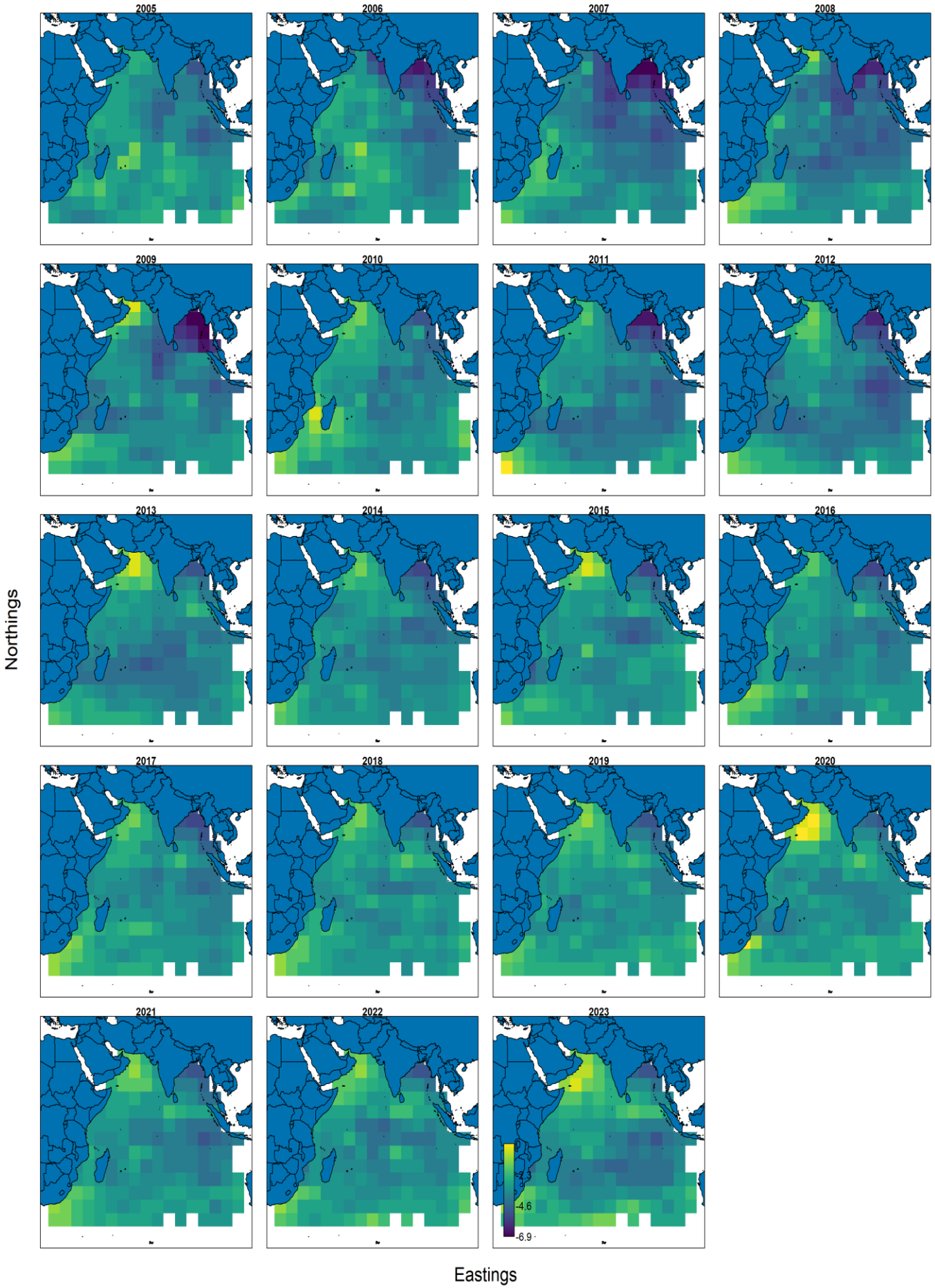

Figure 8. The log density distribution of MSK is based on the parsimonious VAST model from 2005 to 2023.

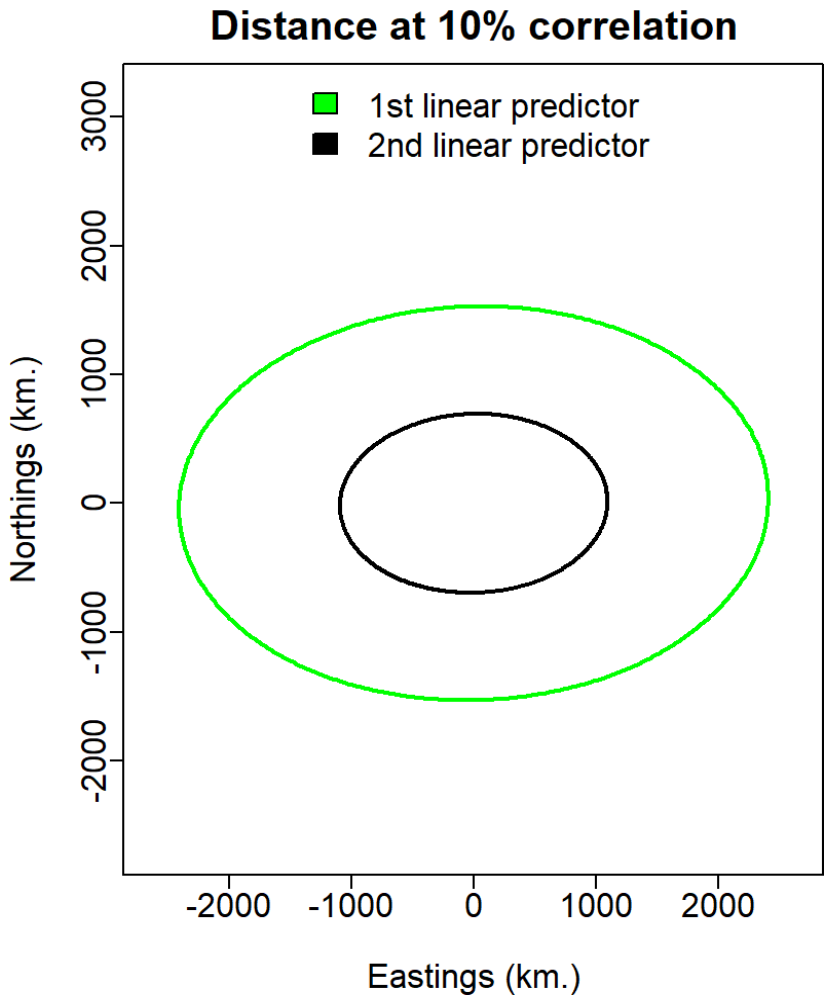

Figure 9. Illustrates the estimation of spatial anisotropy using the delta-Gamma spatiotemporal model (M-4) developed in this study. The ellipses represent the estimates of geometric anisotropy for MSK, with the green line indicating the spatial variation in encounter probability and the black line indicating the spatial variation in positive catch rate. The distance from a point located at (0, 0) where the correlation drops to 10% is represented by the line.

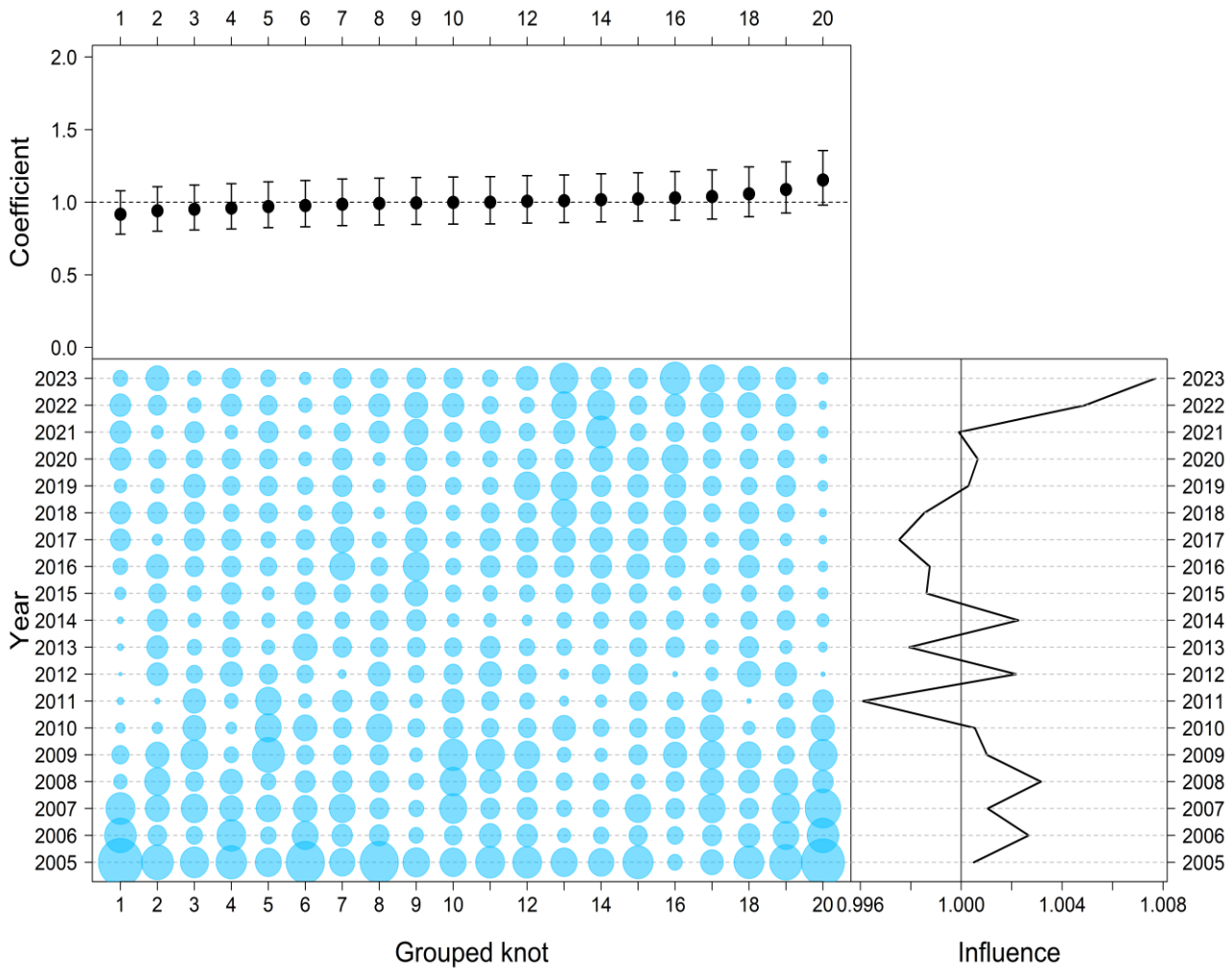

Figure 10. The CDI plot of the spatial random effect for the VAST model applied to MSK in the Indian Ocean from 2005 to 2023. The top section of the plot shows the normalized coefficients with their standard errors. In the bottom left section, bubbles represent the annual distribution of observed CPUEs for each group of five knots, with larger bubbles indicating more data records. The bottom right section illustrates the annual influence value for the spatial random effect.