Update on CPUE Standardization of Blue Marlin (*Makaira nigricans*) from Indonesian Tuna Longline Fleets 2006-2023

Bram Setyadji[1,5], **Matthew Spencer**[1], **Laurence Kell**[2], **Serena Wright**[3], **Scott Ferson**[4]

[1]School of Environmental Sciences, University of Liverpool, UK
[2]Imperial College, London, UK
[3]Centre for Fisheries and Aquaculture Sciences (Cefas), UK
[4]School of Engineering, University of Liverpool, UK
[5]National Research and Innovation Agency (BRIN), Indonesia

Abstract

Blue marlin (*Makaira nigricans*) is commonly caught as bycatch by Indonesian tuna longline fleets targeting albacore, yellowfin, and bigeye tunas, contributing approximately 600 tons annually. Indices of relative abundance, derived from commercial catch data, are essential inputs for stock assessment models that inform fisheries management and decision-making processes. In this study, a delta-lognormal generalized linear model (GLM) was used to standardize catch per unit effort (CPUE) and estimate relative abundance indices for blue marlin, based on data collected by the Indonesian scientific observer program from August 2005 to September 2023. Most observed vessels operated out of Benoa Port, Bali.

The results indicated that year, quarter, latitude, and longitude were statistically significant predictors and were retained in the lognormal component of the model, while moon phase and fishing cluster were excluded. Notably, fishing cluster (representing targeting practices) and longitude had no significant effect on the probability of blue marlin catch occurrence. In contrast, year, quarter, longitude, and moon phase appeared to influence catch rates. Overall, the standardized CPUE trend remained relatively stable during the first five years, then approximately doubled in 2012 before returning to a stable level in the subsequent years. However, high uncertainties seemed as lingering issue, which inevitable due to low coverage of scientific observer data.

Keywords: By-catch, catch and effort, longline, stock abundance

1 Introduction

Blue marlin *Makaira nigricans* (Lacépède, 1082) is an apex predator, highly migratory species and considered as a non-target species from Indonesian industrial and small-scale tuna fishery (Sulistyaningsih, Barata, and Siregar 2011; Widodo, Prisantoso, and Suprapto 2016; Nugraha and Setyadji 2013). They are solitary species, prefers the warm offshore surface waters above 24°C and known to have high commercial value in the tropical and subtropical Indian and Ocean Pacific (Nakamura 1985). However, due to its characteristics, blue marlin is threatened by over-exploitation (Collette et al. 2011).

In Indian Ocean, blue marlin was largely caught by longline (68%), followed by gillnets (15%), with remaining catches recorded under coastal longline, troll and handlines (IOTC Secretariat 2024). Contribution of blue marlin from Indonesian fleet between 2018-2022 was around 8% (~600 tons) of total catch in Indian Ocean, ranked third after Taiwan, Srilanka, and India (IOTC Secretariat 2024). Results of latest stock assessment undertaken in 2022, as calculated based on the JABBA model (Just Another Bayesian Biomass Assessment) indicated that, blue marlin stock of the Indian Ocean is overfished and subject to overfishing (IOTC-WPB22 2024), with 100% change of violating the MSY-based reference points in next 10 years if the catch level at the time of the assessment is maintained. However, there were some uncertainties in the robustness of the data available (nominal catch) and the CPUE series, especially in the north eastern Indian Ocean which may hampers the assessment.

Through this paper we attempt to bridge the research's gap in term blue marlin abundance in the north eastern Indian Ocean. Hopefully, the results will be useful for assessing the status of the stock of blue marlin, which is an important fishery resource in the Indian Ocean.

2 Materials and Methods

2.1 Data Collection

This research analyzed the data gathered by the Indonesian scientific observers on commercial tuna longline vessels, which are mainly situated in Benoa Fishing Port, Bali. The observation program started in 2005 through an Australia-Indonesia collaboration (Project FIS/2002/074 of Australian Centre for International Agricultural Research). Continued by the Research Institute for Tuna Fisheries (RITF) from 2012-2021 and took over by Directorate General of Capture Fisheries (DGCF) from 2022 onward.

A total of 2,337 set-by-set data span in detail 1x1 degree latitude and longitude grid from January 2006 to September 2023 were obtained from Indonesia scientific observer, which covers commercial tuna longline vessels mostly based in Port of Benoa, Bali. Fishing trips usually last from three weeks to three months. Main fishing grounds cover from west to southern part of Indonesian waters, stretched from 75°E to 35°S (Figure 1). It also informed concerning the number of fish caught by species, total number of hooks, number of hooks between floats (HBF), start time of the set, start time of haul, soak time, and geographic position where the longlines were deployed into the water.

2.2 Cluster Analysis

Cluster analysis was performed based on species composition as proposed by (He, Bigelow, and Boggs 1997). Further, for each set, the catch composition was calculated and expressed as proportions relative to the total of the four tuna species (e.g. albacore, bigeye tuna, southern bluefin tuna, and yellowfin tuna) and five billfish species (i.e., black marlin, blue marlin, striped marlin, sailfish, swordfish). Clustering a large dataset could be a major stumbling block. Sadiyah, Dowling, and Prisantoso (2011) suggested to perform two step clustering methods, by using non-hierarchical k-means and followed by agglomerative hierarchical clustering. Thus, for this purpose the analyses were performed using NbClust package (Charrad et al. 2014), which was intended to perform k-means and hierarchical clustering with different distance measures and aggregation methods at one go.

The hierarchical cluster analysis with Ward minimum variance method ("ward.D2") followed the criterion by Murtagh and Legendre (2014) was applied, which requires the dissimilarities to be squared before cluster updating. It then processed to the squared Euclidean distances across 21 indices in order to select the optimal number of clusters based on majority rule. The result then passed to CLARA (clustering large applications) under cluster package (Kaufman and Rousseeuw 1990).

2.3 Data Filtering

The major issue for modelling the abundance for billfishes from Indonesian tuna longline fishery was the high proportion number of zero-catch-per-set (Setyadji, Andrade, and Proctor 2018). It was acknowledged that predominance of zero catches could be driving the model outputs as the CPUE trends do not appear to be biologically plausible (IOTC-WPB16 2018). Originally the mean annual proportion of zero catches from the data was very high, up to 90%. In attempt to reduce it, several ways were conducted as follows:

1. Exclude 2005 data from analysis, since it was the beginning of the scientific observer program, therefore it might contain species misidentification;

2. Exclude sets which doesn't contain blue marlin for the whole trip.

As a result of the application of the procedures above, total number of sets used in the analysis was 2,337 and zero catch ratio were slightly reversed to ~85%. Moreover, the filtering process also intended to find spatial consistency across years of observation.

2.4 CPUE standardization

A delta-lognormal GLM was applied to standardize the CPUE. As the approach of Wang (2018) with some modifications, the models were simply conducted with the main effects considered in this analysis were as follow

a. **Year**, set as categorical variable (2006-2020);
b. **Quarter**, set as categorical variable (1-4);
c. **Cluster**, set as categorical variable (1-3);
d. **Moon**, referred to the eight shapes of the directly sunlit portion of the moon that we can see from Earth. The moon phase was calculated using lunar package (Lazaridis 2014);
e. **Lat/Lon**, defined as georeferenced information in 5x5 degree and presented in absolute value to avoid negative mark. Incorporated as a continuous variable in the GLM analysis.

The interactions between main effects were not incorporated into the models to avoid overfitting. The lognormal and delta models were conducted as follows:

Lognormal model for CPUE of positive catch:

$$log\ (CPUE) = \mu + Year + Quarter + Cluster + Moon + Lat + Lon + \varepsilon^{lognormal}$$

Binomial model for presence and absence of catch:

$$PA = \mu + Year + Quarter + Cluster + Moon + Lat + Lon + \varepsilon^{del}$$

We used a forward approach to select the explanatory variables and the order they were included in the full model. The first step was to fit simple models with one variable at a time. The variable included in the model with lowest residual deviance was selected first. As second step the model with the selected variable then received other variables one at a time, and the model with lowest residual deviance was again selected. This procedure continued until residual deviance did not decrease as new variables were added to the previous selected model. Finally, all main effects and first order interactions were considered and a backward procedure based on Akaike Information Criterion (AIC) (Akaike 1974).

The area-specific standardized CPUE trends were estimated based on the exponentiation of the adjust means (least square means) of the year effects (Maunder and Punt 2004; Butterworth 1996). The standardized relative abundance index was calculated by the product of the standardized CPUE of positive catches and the standardized probability of positive catches:

$$index = e^{\log(CPUE)} \left(\frac{e^{\tilde{P}}}{1 + e^{\tilde{P}}} \right)$$

Where:

$CPUE$: is the adjust means (least square means) of the year effect of the gamma model;

$\tilde{P}$: is the adjust means (least square means) of the year effect of the delta model.

Maps and the statistical analyses were carried out using R software version 4.2.0 R Core Team (2018), particularly the package *emmeans* (Lenth 2018), and *MASS* (Venables and Ripley 2002).

3 Results

3.1 Cluster Result

Based on majority rules (Figure 2), the optimal number of clusters was three. Cluster 1 was consisted of mixed ALB, YFT and BET, whereas cluster 2 was dominantly filled with BET, and the biggest proportion in cluster 3 was ALB (Figure 3).

3.2 Descriptive Statistic

Onboard observers recorded catch and operational data following Indonesian tuna longline commercial vessels from 2006-2023. The filtered dataset contained 80 trips, 2337 sets, and around 3.2 million hooks observed, respectively (Table 1). The distribution of sets mainly gathered in area of eastern Indian Ocean with most of the positive catches occurred in the area south of Indonesian waters, between 0º-20º S and 75º-125º E (Figure 4).

3.3 CPUE Data Characteristics

BUM nominal CPUE series is presented in Figure 5. In general, the catches of BUM during the last decade were fluctuating, with tendency of rising since 2007, and slowing down after reached its peak in 2012. The series bounced back after reached its lowest in 2017 and relatively constant afterward with some noises. The lowest CPUE recorded was in 2018 (0.03$\pm$0.16), as the highest was in 2012 (0.28$\pm$0.63). In addition, the proportion of zero catch for BUM was quite high. As opposed to nominal CPUE, the trend was varying annually between a maximum of 0.96% in 2017 and a minimum of 0.77% in 2013 with average proportion 86% (Figure 6).

3.4 CPUE Standardization

Year, quarter, latitude and longitude were kept and statistically significant in the lognormal model (Table 2), whereas only cluster was excluded from the binomial model (Table 3). The positive catch of blue marlin (BUM) was allegedly influenced by spatial (latitude and longitude) and temporal factor (year and quarter), an indication of known seasonality and association with fishing ground. In addition, targeting effect was neither play part on possibility of catching or positive catches.

Overall, the standardized CPUE trend remained relatively stable during the first five years, then approximately doubled in 2012 before returning to a stable level in the subsequent years. However, high uncertainties seemed as lingering issue, which is inevitable due to low coverage of scientific observer data.

Acknowledgement

The Authors would like to thank to all scientific observers of Research Institute for Tuna Fisheries (RITF) for their contribution in collecting data throughout the years. We also would like to extend our gratitude to various organization, namely, Commonwealth Scientific and Industrial Research Organization (CSIRO), the Australian Center for International Agricultural Research (ACIAR) and the Research Institute for Capture Fisheries (RCCF) through research collaboration in the project FIS/2002/074: Capacity Development to Monitor, Analyze and Report on Indonesian Tuna Fisheries. Also, Fisheries Improvement Project (FIP) for their funding support during Covid-19 pandemic.

References

Akaike, Hirotugu. 1974. "A New Look at the Statistical Model Identification." *IEEE Transactions on Automatic Control* 19 (6): 716–23. https://doi.org/10.1109/TAC.1974.1100705.

Butterworth, D. S. 1996. "A Possible Alternative Approach for Generalised Linear Model Analysis of Tuna CPUE Data." *Collective Volume of Scientific Papers* 45: 123–24.

Charrad, Malika, Nadia Ghazzali, Véronique Boiteau, and Azam Niknafs. 2014. "Nbclust: An r Package for Determining the Relevant Number of Clusters in a Data Set." *Journal of Statistical Software* 61 (6): 1–36.

Collette, B. B., K. E. Carpenter, B. A. Polidoro, M. J. Juan-Jordá, A. Boustany, David J. Die, C. Elfes, W. Fox, J. Graves, and L. R. Harrison. 2011. "High Value and Long Life, Double Jeopardy for Tunas and Billfishes." *Science (New York, N.Y.)* 333 (6040): 291–92. https://doi.org/10.1126/science.1208730.

He, Xi, Keith A. Bigelow, and Christofer H. Boggs. 1997. "Cluster Analysis of Longline Sets and Fishing Strategies Within the Hawaii-based Fishery." *Fisheries Research* 31 (1-2): 147–58.

IOTC Secretariat. 2024. "Review of the Statistical Data Available for Indian Ocean Blue Marlin (1950-2022)." IOTC-2024-WPB22-INF06-BUM. Indian Ocean Tuna Commission (IOTC).

IOTC-WPB16. 2018. "Report of the 16th Session of the IOTC Working Party on Billfish." Working {{Party Report}} IOTC– 2018– WPB16–R[E]. Cape Town, South Africa: Indian Ocean tuna Commission (IOTC).

IOTC-WPB22. 2024. "Report of the 22nd Session of the IOTC Working Party on Billfish." Beau Vallon, Seychelles. https://iotc.org/sites/default/files/documents/2024/10/IOTC-2024-WPB22-RE.pdf.

Kaufman, Leonard, and Peter J. Rousseeuw. 1990. "Finding Groups in Data: An Introduction to Cluster Analysis." In *Probability and Mathematical Statistics. Applied Probability and Statistics*. Wiley Series.

Lazaridis, Emmanuel. 2014. "Lunar: Lunar Phase & Distance, Seasons and Other Environmental Factors."

Lenth, Russell. 2018. "Emmeans: Estimated Marginal Means, Aka Least-Squares Means."

Maunder, Mark N., and André E. Punt. 2004. "Standardizing Catch and Effort Data: A Review of Recent Approaches." *Fisheries Research* 70 (December): 141–59. https://doi.org/10.1016/j.fishres.2004.08.002.

Murtagh, Fionn, and Pierre Legendre. 2014. "Ward's Hierarchical Agglomerative Clustering Method: Which Algorithms Implement Ward's Criterion?" *Journal of Classification* 31 (3): 274–95. https://doi.org/10.1007/s00357-014-9161-z.

Nakamura, Izumi. 1985. "Billfishes of the World. An Annotated and Illustrated Catalogue of Marlins, Sailfishes, Spearfishes and Swordfishes Known to Date." *FAO Species Catalogue; FAO Fisheries Synopsis* 5 (125): 65.

Nishida, Tom, and Sheng-Ping Wang. 2006. "Standardization of Swordfish (*Xiphias Gladius*) CPUE of the Japanese Tuna Longline Fisheries in the Indian Ocean." *Paper Presented at IOTC 5th WPB Meeting, March 27-31, 2006, Colombo, Sri Lanka*, 10.

Nugraha, Budi, and Bram Setyadji. 2013. "Kebijakan Pengelolaan Hasil Tangkapan Sampingan Tuna Longline Di Samudera Hindia." *Jurnal Kebijakan Perikanan Indonesia* 5 (2): 67–71. https://doi.org/http://dx.doi.org/10.15578/jkpi.5.2.2013.67-71.

R Core Team. 2018. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing.

Sadiyah, Lilis, Natalie Dowling, and Budi Iskandar Prisantoso. 2011. "Changes in Fishing Pattern from Surface to Deep Longline Fishing by the Indonesian Vessels Operating in the Indian Ocean." *Indonesian Fisheries Research Journal* 17 (2): 87–99. https://doi.org/10.15578/ifrj.17.2.2011.87-99.

Setyadji, Bram, Humber Agnelli Andrade, and Craig Hutton Proctor. 2018. "Standardization of Catch Per Unit Effort with High Proportion of Zero Catches: An Application to Black Marlin *Istiompax Indica* (Cuvier, 1832) Caught by the Indonesian Tuna Longline Fleet in the Eastern Indian Ocean." *Turkish Journal of Fisheries and Aquatic Sciences* 19 (2): 119–29. https://doi.org/10.4194/1303-2712-v19_2_04.

Sulistyaningsih, Ririk Kartika, Abram Barata, and Kiroan Siregar. 2011. "Perikanan Pancing Ulur Tuna Di Kedonganan, Bali." *Jurnal Penelitian Perikanan Indonesia* 17 (3): 185–91. https://doi.org/http://dx.doi.org/10.15578/jppi.17.3.2011.185-191.

Venables, W. N., and B. D. Ripley. 2002. *Modern Applied Statistics with S*. 4th ed. New York: Springer.

Wang, Sheng-Ping. 2018. "CPUE Standardization of Striped Marlin (*Tetrapturus Audax*) Caught by Taiwanese Large Scale Longline Fishery in the Indian Ocean." *Paper Presented on 16th Working Party on Billfish, Cape Town, South Africa, 4-8 September 2018, IOTC– 2018– WPB16– 18*, 31.

Widodo, Agustinus Anung, Budi Iskandar Prisantoso, and Suprapto Suprapto. 2016. "Perikanan Pancing Ulur Di Samudera Hindia: Hasil Tangkapan Ikan Berparuh Yang Di Daratkan Di Sendang Biru, Malang, Jawa Timur." *Jurnal Penelitian Perikanan Indonesia* 18 (3): 167–73. https://doi.org/http://dx.doi.org/10.15578/jppi.18.3.2012.167-173.

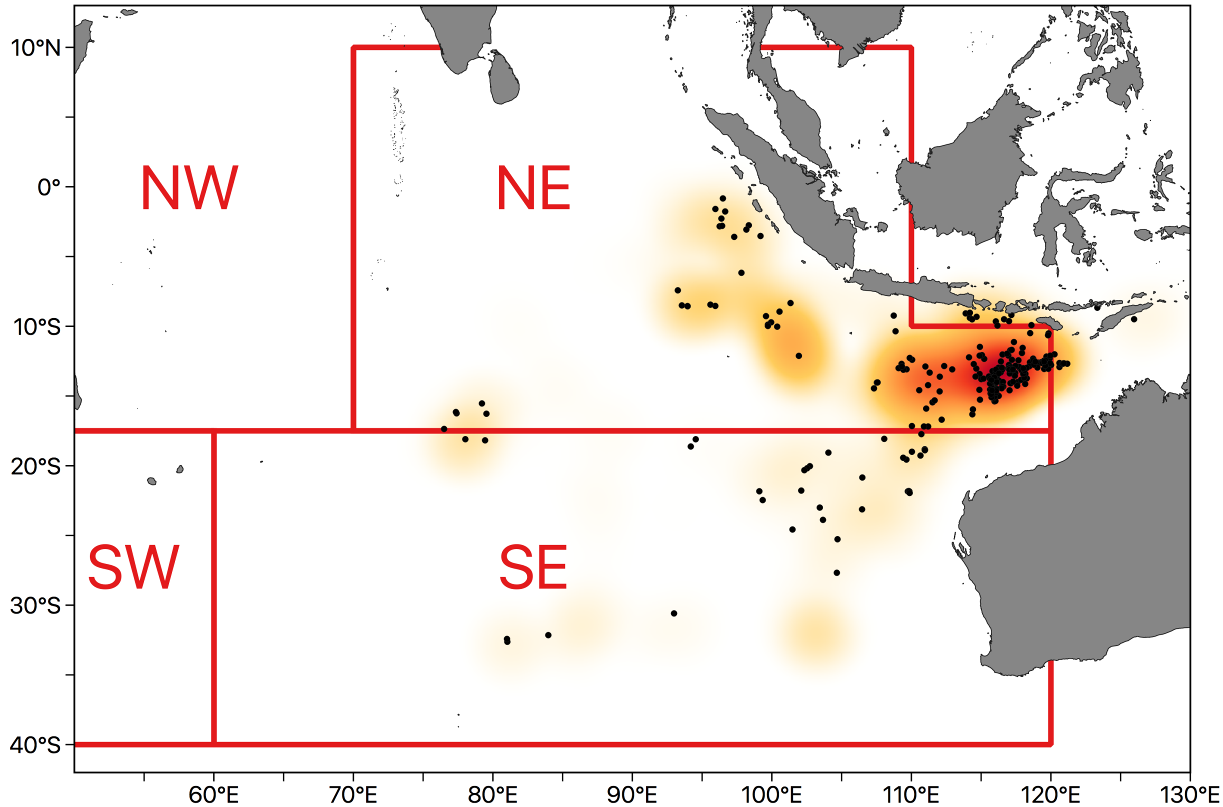

Figure 1: Area stratification used in the analysis (Wang 2018) based on the aggregation of the relative sizes from nine IOTC statistics areas for swordfish in the Indian Ocean (Nishida and Wang 2006)

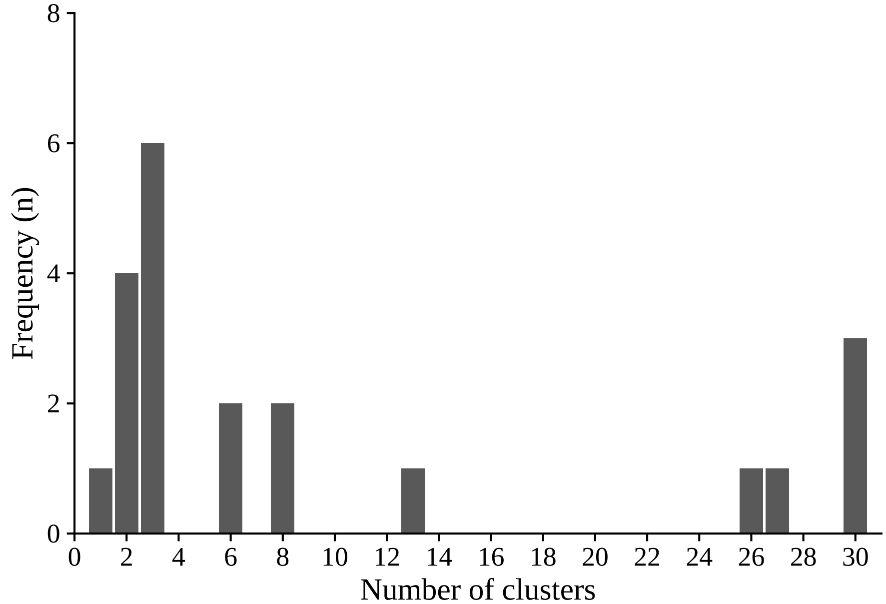

Figure 2: Selection of optimum number of clusters, based on the majority rules.

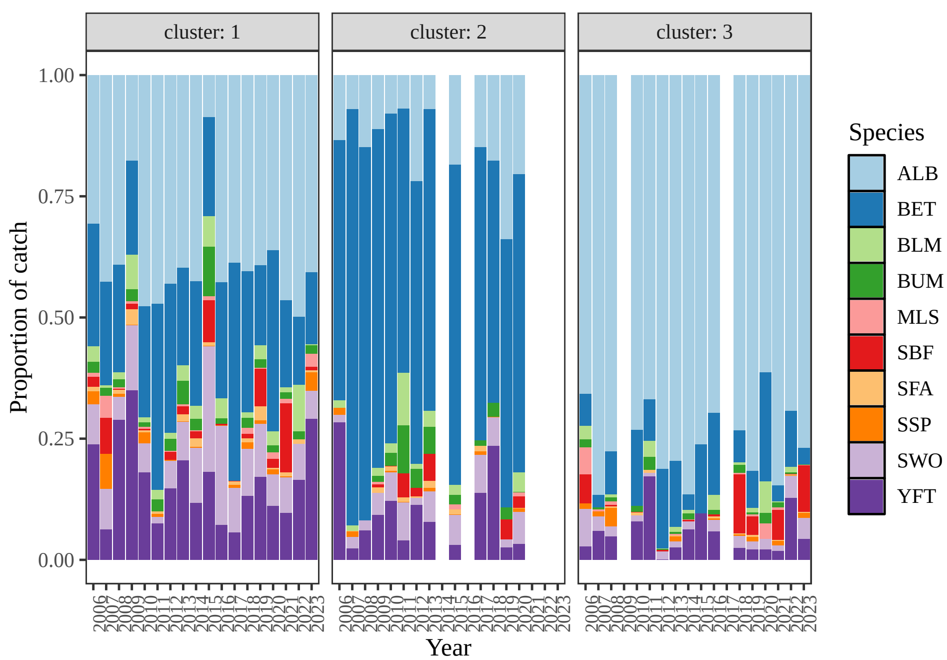

Figure 3: Catch proportions of BUM caught by Indonesian longline fleets operated in the north-eastern Indian Ocean.

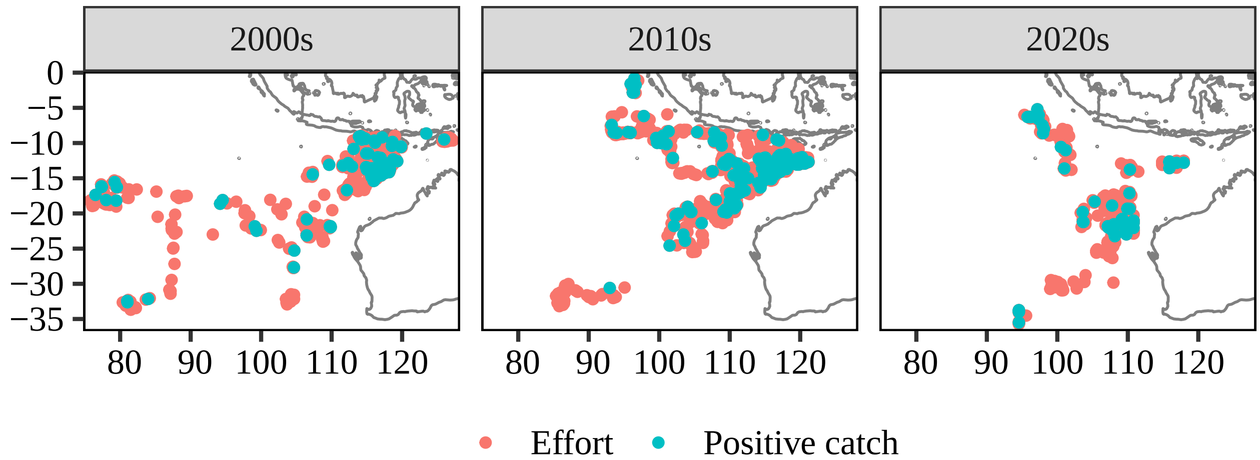

Figure 4: Effort and positive catch distribution of BUM by Indonesian longline fleets operated in the north-eastern Indian Ocean.

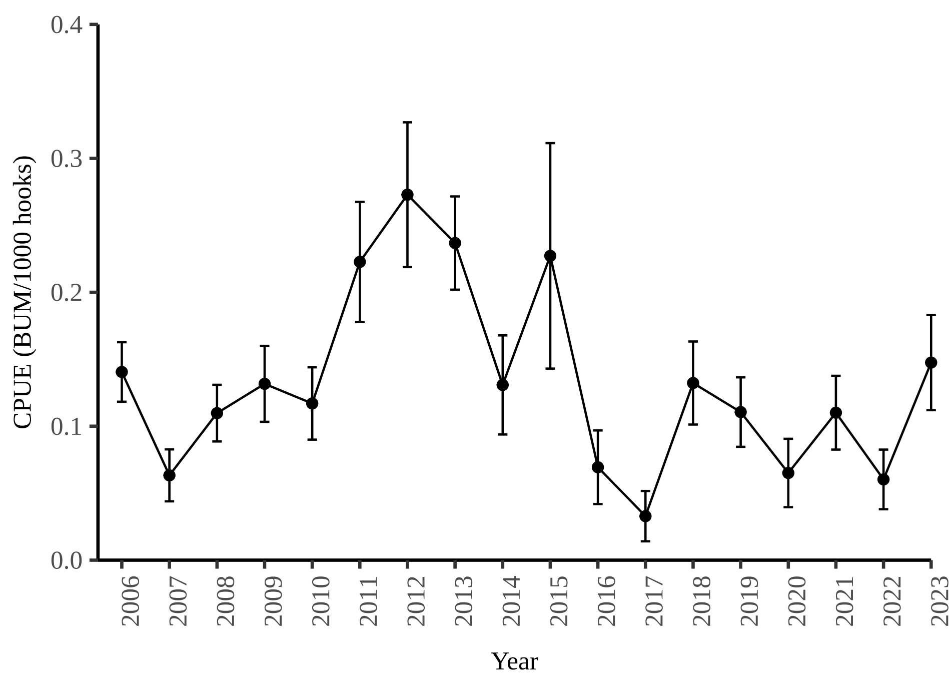

Figure 5: Nominal CPUE series (N/1000 hooks) for BUM from 2006 to 2023. The error bars refer to the standard errors.

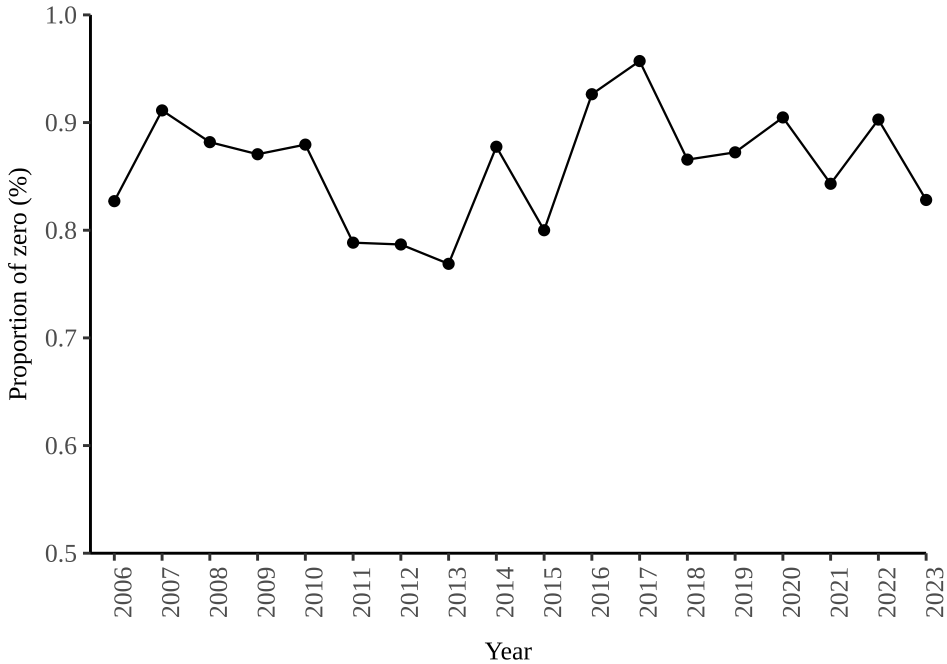

Figure 6: Proportion of zero-catch-per-set from 2006 to 2023. The error bars refer to the standard errors.

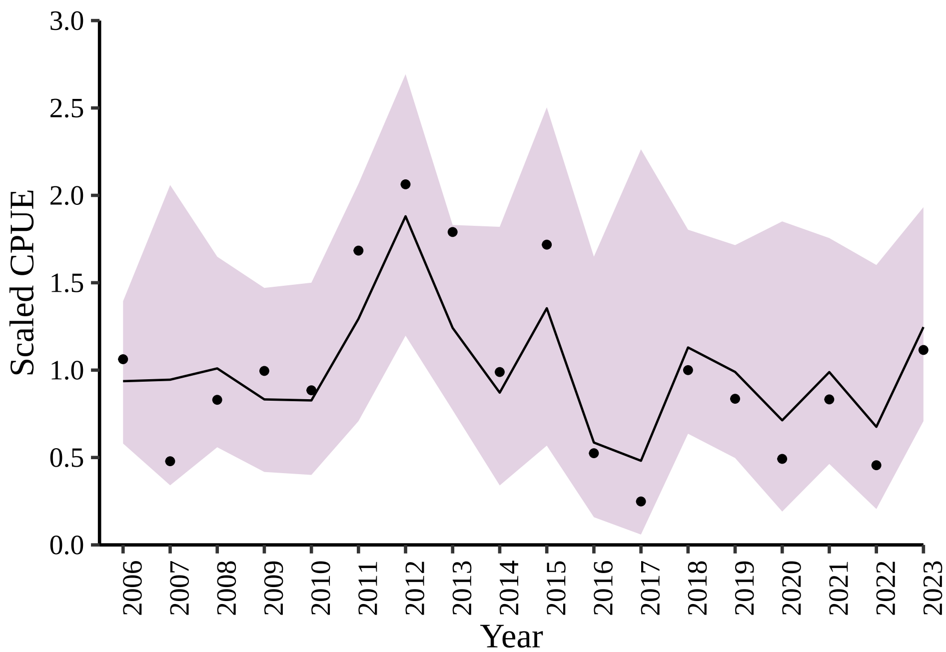

Figure 7: Standardized catch-per-unit-effort (CPUE) calculated using delta-lognormal model. Values were scaled by dividing them by their means.

Table 1: Summary of observed effort from Indonesian tuna longline fishery during 2006–2023. Results are pooled and also presented by year of observation

Year	Trips	Sets	Total Hooks	Mean Hooks	se	Mean HBF	se
2006	8	237	350081	1477.14	13.11	11.35	0.22
2007	4	124	211434	1705.11	27.95	13.44	0.23
2008	8	220	278357	1265.26	30.84	10.24	0.29
2009	5	170	202241	1189.65	16.59	11.64	0.38
2010	6	166	221274	1332.98	35.51	13.61	0.40
2011	3	105	110384	1051.28	16.97	12.00	0.00
2012	5	136	206237	1516.45	55.60	13.26	0.13
2013	6	173	190262	1099.78	16.77	11.61	0.11
2014	4	98	110616	1128.73	22.14	14.29	0.24
2015	2	51	60911	1194.33	27.85	11.84	0.61
2016	2	95	118118	1243.35	12.54	11.42	0.40
2017	2	70	86048	1229.26	25.37	15.64	0.06
2018	5	186	246086	1323.04	14.45	14.90	0.19
2019	7	141	190106	1348.27	15.95	11.74	0.35
2020	2	63	86845	1378.49	18.20	13.48	0.11
2021	3	102	166554	1632.88	22.41	11.61	0.25
2022	3	72	134424	1867.00	33.93	13.61	0.49
2023	5	128	257005	2007.85	35.92	11.52	0.21

Table 2: The deviance table for selected lognormal model.

	Df	Deviance	Resid. Df	Resid. Dev	F	Pr(>F)
NULL	NA	NA	336	51.7166	NA	NA
Year	17	8.0558	319	43.6607	3.9763	0.0000
Quarter	3	1.5320	316	42.1287	4.2851	0.0055
Lat2	1	2.6726	315	39.4561	22.4262	0.0000
Lon2	1	2.0358	314	37.4203	17.0828	0.0000

Table 3: The deviance table for selected binomial model.

	Df	Deviance	Resid. Df	Resid. Dev	Pr(>Chi)
NULL	NA	NA	2334	1927.489	NA
Year	17	42.7036	2317	1884.785	0.0005
Moon	7	22.0107	2310	1862.774	0.0025
Quarter	3	17.3516	2307	1845.423	0.0006
Lat2	1	6.7264	2306	1838.696	0.0095
Lon2	1	3.0491	2305	1835.647	0.0808