Update on CPUE Standardization of Blue Marlin (*Makaira nigricans*) from Indonesian Tuna Longline Fleets 2006-2021

Bram Setyadji[a,*], Denham Parker[b], Sheng-Ping Wang[c]

[a] *National Research and Innovation Agency, Jl. Raya Jakarta - Bogor Km. 46. Cibinong, Bogor, 16911,*

[b] *Forestry, Fisheries and the Environment, Environment House, Cnr. Steve Biko (previously Beatrix Street) and Soutpansberg Road, 473 Steve Biko, Arcadia, Pretoria, 83,*

[c] *National Taiwan Ocean University, Yu-Syce Building, No. 2, Beining Rd, Zhongzheng District, Keelung, 202,*

Abstract

Black marlin (*Makaira indica*) is commonly caught as frozen by-catch from Indonesian tuna longline fleets targeting albacore, yellowfin and bigeye tuna and its contributed around 7% (~600 tons/year). Relative abundance indices as calculated based on commercial catches are the input data for several to run stock assessment analyses that provide models to gather information useful information for decision making and fishery management. In this paper a Delta-Lognormal Model (GLM) was used to standardize the catch per unit effort (CPUE) and to calculate estimate relative abundance indices based on the Indonesian longline dataset. Data was collected from August 2005 to December 2021 through scientific observer program. Most of the vessels monitored were based in Benoa Port, Bali. The result showed that year, quarter, latitude and longitude statistically significant and kept in the lognormal model, whereas moon and cluster were excluded. In addition, according to the delta model, targeting effect (cluster) and vertical movement (latitude) played no part on the possibility of catching BUM. Whereas, longitude, year, quarter, and the presence of moon phase were likely the more influential effects. However, high uncertainties seemed as lingering issue, which is inevitable due to low coverage of scientific observer data.

Keywords: abundance indices, stock assessment, Generalized Linear Model (GLM), by-catch

Introduction

Blue marlin *Makaira indica* (Lacépède, 180) is an apex predator, highly migratory species and considered as a non-target species from indonesian industrial and small-scale tuna fishery (Sulistyaningsih et al., 2011; Widodo et al., 2016; Nugraha and Setyadji, 2013). It is a 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-WPB19, 2021). Contribution of blue marlin from Indonesian fleet between 2015-2019 was around 7% (~600 tons) of total catch in Indian Ocean, ranked third after Taiwan, China and Srilanka (IOTC-WPB19, 2021). Results of latest stock assessment

*Corresponding author
Email addresses: bram.setyadji@brin.go.id (Bram Setyadji), DParker@dffe.gov.za (Denham Parker), wsp@mail.ntou.edu.tw (Sheng-Ping Wang)

undertaken in 2017, as calculated based on the Bayesian State-Space Surplus Production model JABBA indicated that, blue marlin stock of the Indian Ocean is overfished and subject to overfishing (IOTC-WPB19, 2021), with a very high chance (87%) of exceeding 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.

Materials and Methods

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 scheduled to be taken over by Directorate General of Capture Fisheries (DGCF) from 2022 onward.

A total of 2,135 set-by-set data span in detail 1x1 degree latitude and longitude grid from January 2006 to August 2021 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.

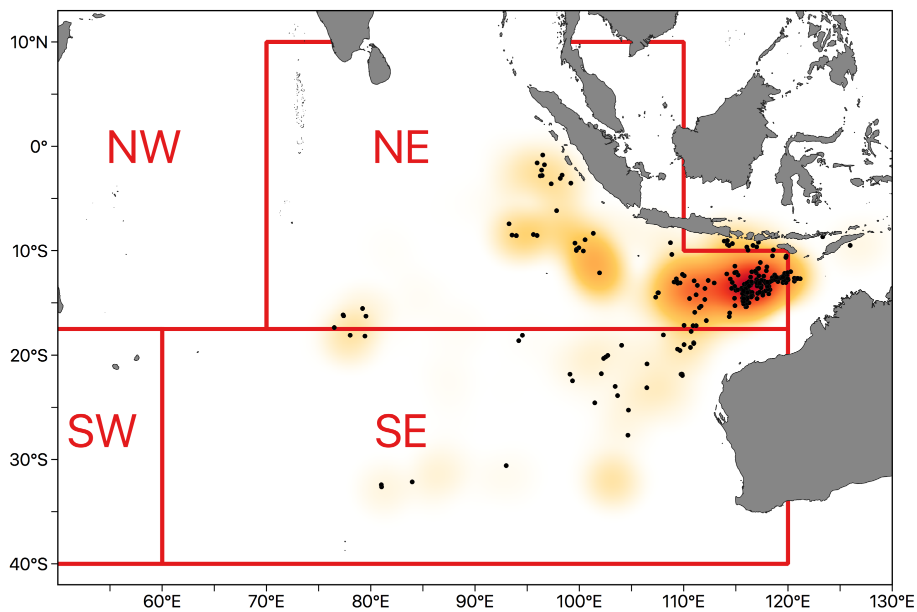

Figure 1. Area stratification used in the analysis (Wang, 2018) based on the aggregation of the relative sizes from nine IOTC statistics areas for swordfish in the Indian Ocean (Nishida and Wang, 2006)

Cluster Analysis

Cluster analysis was performed based on species composition as proposed by He et al. (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 and Prisantoso (2011) suggested to perform two step clustering methods, by using non-hierarchical k-means and followed by agglomerative hierarchical clustering. However, 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).

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 et al., 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, close 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 and criteria above, total number of sets used in the analysis was 2,135 and zero catch ratio were slightly reversed to ~85%. Moreover, the filtering process also intended to find spatial consistency across years of observation.

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 follows:

a. **Year**, set as categorical variable (2006-2021);
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:

Gamma model for CPUE of positive catch:

$$log\ (CPUE) = \mu + Year + Quarter + Cluster + Moon + Lat + Lon + \varepsilon^{lognormal} \tag{1}$$

Delta model for presence and absence of catch:

$$PA = \mu + Year + Quarter + Cluster + Moon + Lat + Lon + \varepsilon^{delta} \tag{2}$$

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:

$$index = e^{\log(CPUE)} \left(\frac{e^{\tilde{P}}}{1 + e^{\tilde{P}}} \right) \tag{3}$$

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 were produced using QGIS version 3.14 (Team, 2020) and the statistical analyses were carried out using R software version 4.2.0 (R Core Team, 2022), particularly the package *emmeans* (Lenth, 2018), and *MASS* (Venables and Ripley, 2002).

Results

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).

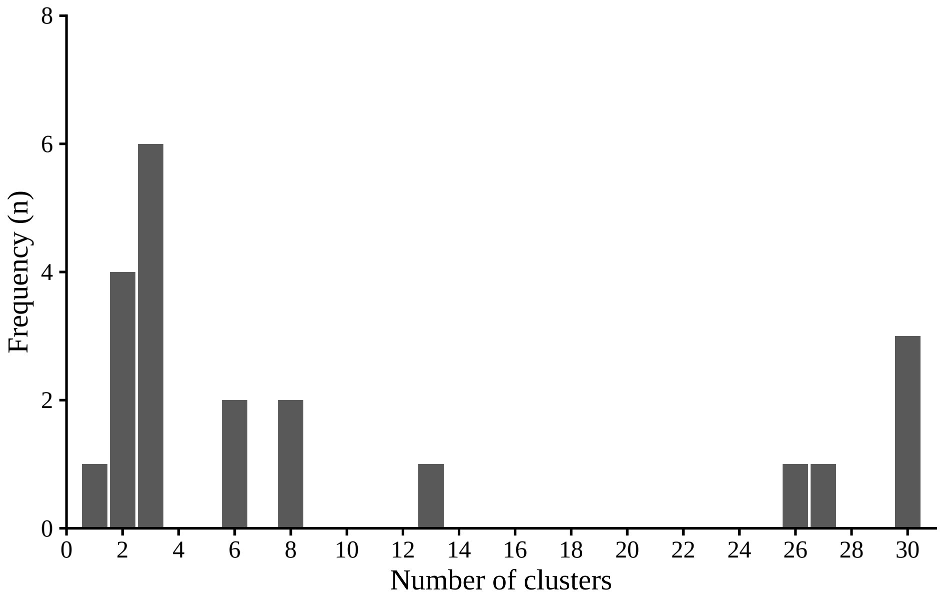

Figure 2. Selection of optimum number of clusters, based on the majority rules.

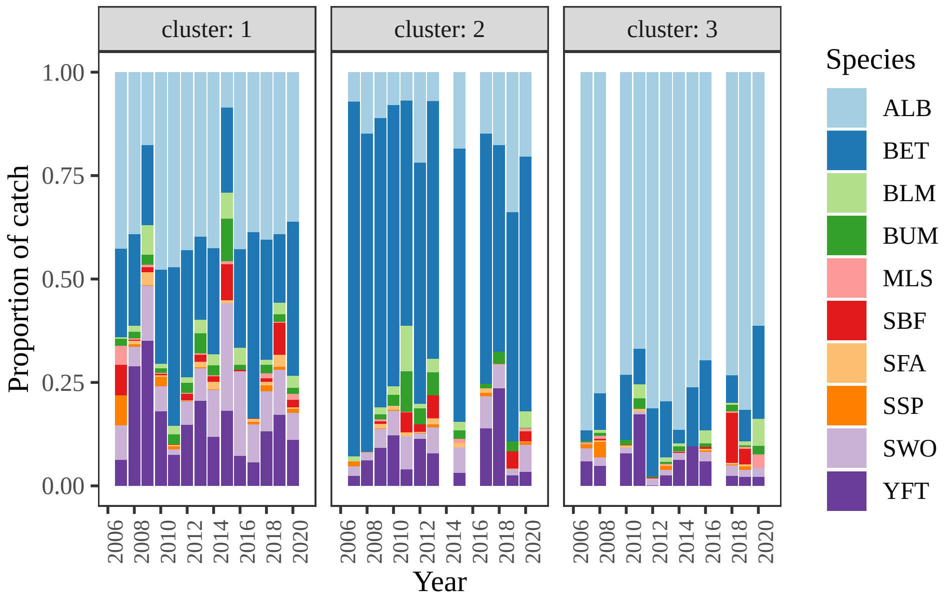

Figure 3. Catch proportions of BUM caught by Indonesian longline fleets operated in the north-eastern Indian Ocean.

Descriptive Statistic

Observers recorded catch and operational data at sea (after cleaning) following Indonesian tuna longline commercial vessels from 2006-2021. The filtered dataset contained 72 trips, 2135 sets, and around 2.8 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 0°-20°S and 75°-125°E (Figure 4).

Table 1. Summary of observed effort from Indonesian tuna longline fishery during 2006–2021. 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

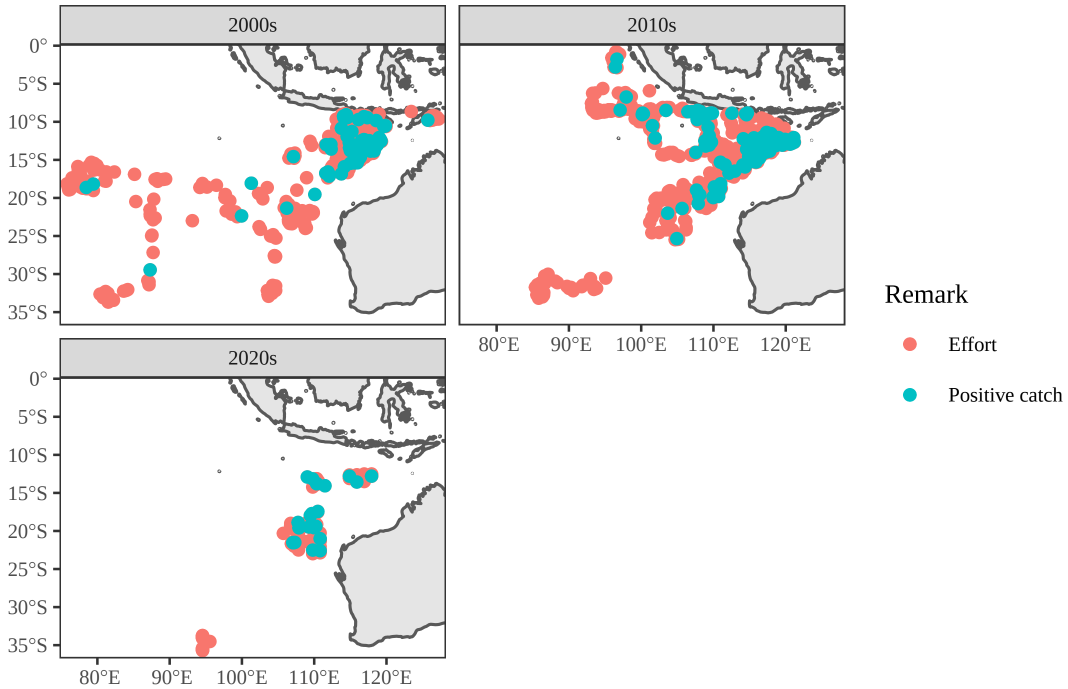

Figure 4. Effort and positive catch distribution of BUM by Indonesian longline fleets operated in the north-eastern Indian Ocean.

CPUE Data Characteristics

In general, the catches of BLM during the last decade were fluctuating, with tendency of declining since 2012. The lowest CPUE recorded was in 2017 (0.04$\pm$0.02), as the highest was in 2012 (0.27$\pm$0.05) (Figure 5). Most of the observations were conducted in the area above 20°S, which belong to the north-eastern Indian Ocean area. In addition, the proportion of zero catch for BUM was quite high. As opposed to nominal CPUE, the trend was varying annually between a maximum of 96% in 2017 and a minimum of 79% in 2012 with average proportion 86%$\pm$0.05 year^{-1} (Figure 5).

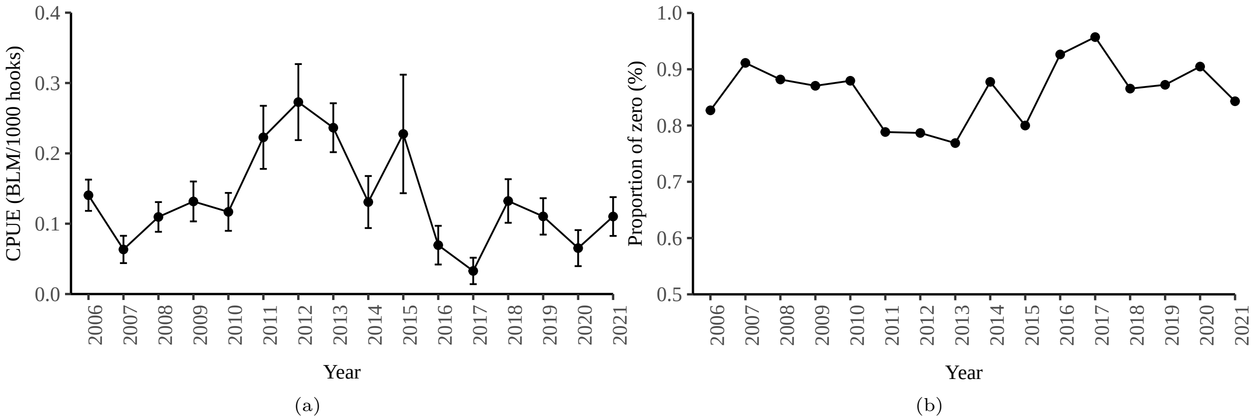

Figure 5. Nominal CPUE series (N/1000 hooks) (left panel) and Proportion of zero-catch-per-set (right panel) for BUM from 2006 to 2021. The error bars refer to the standard errors.

CPUE Standardization

year, quarter, latitude and longitude statistically significant and kept in the lognormal model, whereas moon and cluster were excluded (Table 2). In addition, according to the delta model, targeting effect (cluster) and vertical movement (latitude) played no part on the possibility of catching BUM. Whereas, longitude, year, quarter, and the presence of moon phase were likely the more influential effects (Table 3).

Table 2. The deviance table for selected lognormal model.

	Df	Deviance	Resid. Df	Resid. Dev	F	Pr(>F)
NULL	NA	NA	307	44.2901	NA	NA
Year	15	6.4316	292	37.8584	3.7868	0.0000
Quarter	3	1.2266	289	36.6318	3.6110	0.0138
Lat2	1	1.4932	288	35.1386	13.1878	0.0003
Lon2	1	2.6417	287	32.4969	23.3309	0.0000

Table 3. The deviance table for selected delta model.

	Df	Deviance	Resid. Df	Resid. Dev	Pr(>Chi)
NULL	NA	NA	2134	1761.913	NA
Year	15	40.5196	2119	1721.393	0.0004
Moon	7	21.0024	2112	1700.391	0.0038
Quarter	3	12.0157	2109	1688.375	0.0073
Lon2	1	7.1764	2108	1681.199	0.0074

Overall, the standardized CPUE trend was relatively stable over time with some notably peak in 2011-2013 and 2015. However, high uncertainties seemed as lingering issue, which is inevitable due to low coverage of scientific observer data (Figure 6).

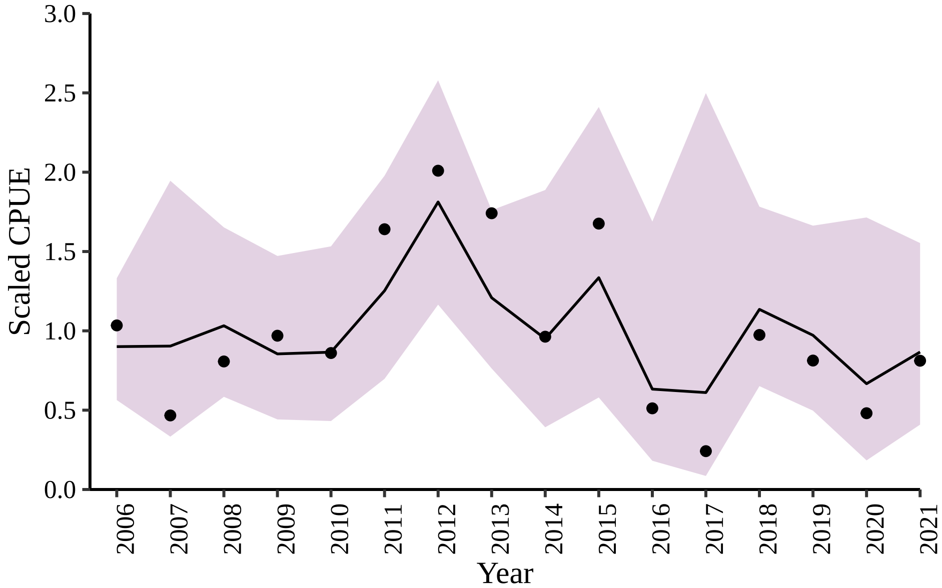

Figure 6. Standardized catch-per-unit-effort (CPUE) calculated using delta-lognormal model. Values were scaled by dividing them by their means.

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, H., 1974. A new look at the statistical model identification. IEEE transactions on automatic control 19, 716–723. doi: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, 123–124.

Charrad, M., Ghazzali, N., Boiteau, V., Niknafs, A., 2014. Nbclust: An r package for determining the relevant number of clusters in a data set. Journal of Statistical Software 61, 1–36.

Collette, B.B., Carpenter, K.E., Polidoro, B.A., Juan-Jordá, M.J., Boustany, A., Die, D.J., Elfes, C., Fox, W., Graves, J., Harrison, L.R., 2011. High value and long life, double jeopardy for tunas and billfishes. Science 333, 291–292. doi:10.1126/science.1208730.

He, X., Bigelow, K.A., Boggs, C.H., 1997. Cluster analysis of longline sets and fishing strategies within the Hawaii-based fishery. Fisheries Research 31, 147–158.

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

IOTC-WPB19, 2021. Report of the 19th Session of the IOTC Working Party on Billfish. Working Party Report IOTC– 2021–WPB19– R[E]_Rev1. Indian Ocean tuna Commission (IOTC). Microsoft Teams Online.

Kaufman, L., Rousseeuw, P.J., 1990. Finding groups in data: An introduction to cluster analysis, in: Probability and Mathematical Statistics. Applied Probability and Statistics. Wiley Series.

Lazaridis, E., 2014. Lunar: Lunar phase & distance, seasons and other environmental factors.

Lenth, R., 2018. Emmeans: Estimated marginal means, aka least-squares means.

Maunder, M.N., Punt, A.E., 2004. Standardizing catch and effort data: A review of recent approaches. Fisheries Research 70, 141–159. doi:10.1016/j.fishres.2004.08.002.

Murtagh, F., Legendre, P., 2014. Ward's hierarchical agglomerative clustering method: Which algorithms implement Ward's criterion? Journal of Classification 31, 274–295. doi:10.1007/s00357-014-9161-z.

Nakamura, I., 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, 65.

Nishida, T., Wang, S.P., 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, B., Setyadji, B., 2013. Kebijakan pengelolaan hasil tangkapan sampingan tuna longline di Samudera Hindia. Jurnal Kebijakan Perikanan Indonesia 5, 67–71. doi:http://dx.doi.org/10.15578/jkpi.5.2.2013.67-71.

R Core Team, 2022. R: A language and environment for statistical computing.

Sadiyah, L., Prisantoso, B.I., 2011. Fishing strategy of the indonesian tuna longliners in Indian Ocean. Indonesian Fisheries Research Journal 17, 29–35. doi:http://dx.doi.org/10.15578/ifrj.17.1.2011.29-35.

Setyadji, B., Andrade, H.A., Proctor, C.H., 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, 119–129. doi:10.4194/1303-2712-v19_2_04.

Sulistyaningsih, R.K., Barata, A., Siregar, K., 2011. Perikanan Pancing Ulur Tuna Di Kedonganan, Bali. Jurnal Penelitian Perikanan Indonesia 17, 185–191. doi:http://dx.doi.org/10.15578/jppi.17.3.2011.185-191.

Team, Q.D., 2020. QGIS Geographic Information System. Technical Report. Open Source Geospatial Foundation Project.

Venables, W.N., Ripley, B.D., 2002. Modern Applied Statistics with S. Fourth ed., Springer, New York.

Wang, S.P., 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, A.A., Prisantoso, B.I., Suprapto, S., 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, 167–173. doi:http://dx.doi.org/10.15578/jppi.18.3.2012.167-173.