CPUE Standardization of Bigeye Tuna, *Thunnus obesus* (Lowe, 1839) from Indonesian Tuna Longline Fishery in the Eastern Indian Ocean

Hety Hartaty[1], Bram Setyadji[1*], Tom Nishida[2], Zulkarnaen Fahmi[1]

[1]Research Institute for Tuna Fisheries, Agency for Marine and Fisheries Research and Human Resources, Jl. Mertasari No. 140, Denpasar, Bali, Indonesia
[2]National Research Institute of Far Seas Fisheries, Shimizu, Shizuoka 424-8633, Japan

Abstract
Bigeye tuna, *Thunnus obesus* (Lowe, 1839) is one of the main target species for Indonesian tuna longline fishery in the Eastern Indian Ocean. The tuna longline fishery has begun since 1978 and around 1980, bigeye tuna started as target when deep longline introduced. However, little is known about its abundance, especially in the north eastern area where is the core fishing ground for Indonesian tuna longline fishery. The objective of the study is to provide a preliminary assessment about the abundance indices of bigeye tuna from Indonesian tuna longline fishery. In this paper, four types of Generalized Linear Model (GLM) was used to standardize the catch per unit effort (CPUE) and to estimate the relative abundance indices, i.e. Zero-inflated Negative Binomial (ZINB), Negative Binomial (NB), Tweedie (TW) and Delta-lognormal (DEL). We used two types of data used in this study; the scientific observer data conducted by Research Institute for Tuna Fisheries (RITF) from 2006 to 2018 and national observer program conducted by Directorate General of Capture Fisheries (DGCF) from 2016-2017. On overall, the abundance of bigeye tuna was depleted quite substantially over the years (almost two-fold from the beginning of observation). ZINB failed to give plausible indices due to convergence problems when areas were included. NB, TW and DEL produced similar trends, especially NB and TW produced almost identical trends. DEL produced higher abundance indices between 2006-2012 and lower prediction afterwards compared to previous two models. NB produced the lowest AIC but TW has the lowest BIC values than others, however, it was suggested that abundance indices by NB is likely the most plausible.
Keywords: bigeye tuna, CPUE standardization, Generalized Linear Models

Introduction

World tuna and tuna-like production in 2016 reached up to 7.5 million ton and around 23% (1.7 million tons) came from Indian Ocean . Indonesia contributes more than 207,010 ton in 2010, rise up 1.84% from previous year. Port of Benoa contributes more than 60% of tuna production in Indonesia (Setyadji et al., 2012). Among tuna and tuna-like species, bigeye tuna (*Thunnus obesus*) is one of the most commercially important species in the Indian Ocean (Fonteneau et al., 2005; Lee et al., 2005; Nugraha et al., 2010; Polacheck, 2006). They are widely distributed from tropical to subtropical waters among 3 major oceans, between 45ºN and 40ºS except the Mediterranean

* Corresponding author at: E-mail: bram.setyadji@gmail.com, Tel.: +62 361 726201, Fax.: +62 361 8497447

Sea (Collette and Nauen, 1983). It also the principal target species of the large longliners from Japan, China, and Taiwan and smaller longliners based in several Indian Ocean Island countries, especially Indonesia (Nootmorn, 2004).

The longline catch of bigeye tuna in the Indian Ocean has increased from approximately 40,000 tons in the late–1980s and soared up to ~170,000 tons in the mid 2000's before declining to less than half in the last 5 years (~90.000 tons) (IOTC-WPTT20, 2018). Indonesia is the largest contributor with average catch around 26.000 tons (27%) during 2013-2017, followed by Taiwan, Seychelles, EU (Spain) and Japan with each proportion 18%, 12%, 12% and 5%, respectively (IOTC-WPTT20, 2018). The nominal CPUE of bigeye tuna of Indonesian longline fishery from 2005-2017 showed a declining trend over the years, and distribution of high CPUE occurred between 5-20 ^{o}S and 30-35 ^{o}S (Hartaty et al., 2018). Bigeye tuna became one of the main targets of Indonesian longline fisheries since the introduction of deep tuna longline in mid 1980's (Sadiyah et al., 2011). The proportion of zero-catches in the scientific observer operational catch and effort data sets from 2005-2018 were considered low (~30%). An indication of targeting from the longline fishing gear. However, despite of its importance, no update on abundance indices of bigeye tuna has been made since the preliminary work from Sadiyah et al. (2012).

The objective of this study was to bridge the research's gap on abundance indices of bigeye tuna, especially in the eastern Indian Ocean. The result hopefully can be involved in stock assessment study in the near future.

Materials and methods

We used the Indonesian scientific observer data from commercial tuna longline vessels based in Benoa Fishing Port, Bali. The observer program started in 2005 through an Australia-Indonesia collaboration (Project FIS/2002/074 of Australian Centre for International Agricultural Research), and since 2010 it has been conducted by the Research Institute for Tuna Fisheries (RITF Indonesia).

The dataset includes information concerning the number of fishes caught by species, the total number of hooks, the number of Hooks Between Floats (HBF), start time of the set, soak time, and geographic position (latitude and longitude) where the longlines deployed into the water. The response variable in the models was nominal catch or CPUE depending on model. Year and quarter were used as a categorical (factor) explanatory variables.

CPUE Standardization

We attempted four GLM models, 3 single structured and 1 two-structured models namely: Negative Binomial (NB), Zero-Inflated Negative Binomial (ZINB), Tweedie (TW), and Delta-lognormal (DELTA) models. Poisson model was not used because of over-dispersion problems, where the variance of the dataset was higher than the mean. Response variable for DELTA were log(CPUE) for positive sub-model and proportion of positive catch for second sub-model. On the other hand, TW, NB and ZINB used number of catch as response variable with effort included in the models as an offset caught (in natural log format). The final models' construction was listed as follow:

1) Tweedie

   $Catch \sim \mu + Year + Quarter + HBF + Start\ Time + Area\ Tree + offset(\log(Hooks)) + \varepsilon$

2) Negative binomial

   $Catch \sim \mu + Year + Quarter + HBF + Start\ Set + Area\ Tree + offset(\log(Hooks)) + \varepsilon$

3) Zero-inflated Negative binomial

   $Catch \sim \mu + Year + Quarter + Moon + HBF + Start\ Set + Soak\ Time + offset(\log(Hooks)) + \varepsilon$

4) Delta-lognormal

   Lognormal model for CPUE of positive catch:

   $\log(CPUE) = \mu + Year + Quarter + Start\ Set + Area\ Tree + \varepsilon^{lognormal}$

   Delta model for presence and absence of catch:

   $\mathrm{PA} = \mu + Year + Quarter + HBF + Moon + Soak\ Time + Start\ Set + Area\ Tree + \varepsilon^{del}$

Where:

a. Year: analyzed between 2005 and 2016;
b. A quarter of the year: 4 categories: 1 = January to March, 2 = April to June, 3 = July to September, 4 = October to December;
c. Area: treated as a categorical variable, area stratification method was applied using GLM-tree approach proposed by Ichinokawa and Brodziak (2010); The algorithm showed that the area divided into 21 categories (Figure 1).
d. Start time of the set: treated as a quantitative variable, the values were rounded to the nearest integer;
e. Soak time: calculated as the time elapsed between the start of setting up the longline and the

start of hauling. Soak time in the model was treated as a continuous variable. Thus, the value was rounded to the nearest integer;

f. The number of hooks between floats: treated as a quantitative variable instead of factor.

g. Moon phase: Moon phase information is available as a daily index of moon fraction for all recorded sets and ranges between 0 and 1 (from new moon to full moon). The moon phase was calculated using lunar package (Lazaridis, 2014). To account for the effect of cyclic behavior, the moon phase was defined by the following function (Sadiyah et al., 2012):

$$Moon = sin(2\pi\ x\ moon\ phase) + cos(2\pi\ x\ moon\ phase)$$

We applied a forward approach to select the explanatory variables and the order 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 the lowest residual deviance was selected as a start. As the second step, the model with the selected variable then received other variables one at a time, and the model with the lowest residual deviance was again selected. The same procedure will be extended until the residual deviance did not decrease as new variables added to the previously selected model. Finally, all main effects and first-order interactions were analyzed and a backward procedure

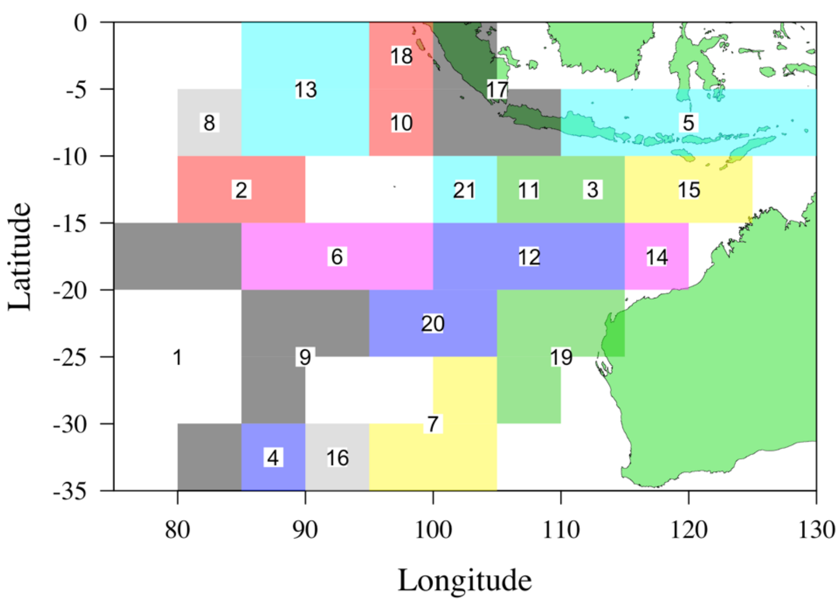

Figure 1. Area stratification used in the analysis based on glm.tree package (Ichinokawa and Brodziak, 2010)

The qualities of the fittings were assessed by comparing the observed frequency distributions of the number of fishes caught to the predicted frequency distribution, as calculated using the selected models. Based on Akaike Information Criterion (AIC) (Akaike, 1974) and Bayesian Information Criterion (BIC) (Schwarz, 1978), Log Likelihood, and R^2 (Lüdecke et al., 2019) the best model was selected. All the statistical analyses were carried out using R software version 3.5.3 (R Core Team, 2019), particularly the package pscl (Jackman, 2017), emmeans (Lenth, 2018), MASS (Venables and Ripley, 2002), Hmisc (Harrell Jr. et al., 2018), and statmod (Giner and Smyth, 2016).

Result and discussion

Descriptive Catch Statistic

Scientific observers and national observers recorded catch and operational data at sea following Indonesian tuna longline commercial vessels from 2006-2018 and 2016-2017, respectively. The combined dataset contained 112 trips, 2984 sets, 3,703 days-at-sea, and more than 3.9 million hooks deployed, respectively (Table 1).

Table 1. Summary of observed fishing effort from Indonesian tuna longline fishery during 2005–2016. Results are pooled and presented by year of observation. Operational parameters are means and standard deviations (in parenthesis).

Year	Trips	Sets	Days at Sea	Total Hooks	Hooks per Set		Hooks per Float	
2006	13	401	401	577,243	1,439.51	(214.9)	11.2	(3.9)
2007	13	265	258	406,135	1,532.58	(326.5)	14.0	(4.4)
2008	15	370	404	483,662	1,307.19	(385.9)	13.0	(4.5)
2009	13	283	288	323,042	1,141.49	(234.7)	12.1	(4.9)
2010	6	165	152	220,394	1,335.72	(457.5)	13.6	(5.2)
2011	3	105	111	110,384	1,051.28	(173.9)	12.0	-
2012	8	198	192	290,265	1,465.98	(559.1)	14.1	(2.3)
2013	7	225	198	252,919	1,124.08	(210.4)	12.7	(2.1)
2014	5	167	265	193,740	1,160.12	(176.9)	15.0	(2.0)
2015	5	148	241	172,463	1,165.29	(145.2)	14.1	(3.2)
2016	8	244	383	324,068	1,314.89	(146.4)	15.2	(6.4)
2017	10	218	489	279,204	1,214.04	(395.3)	17.2	(4.8)
2018	6	195	321	262,856	1,349.98	(242.9)	14.8	(4.5)

High catch rate mainly distributed in eastern Indian Ocean between 0^o-35^o S and 75^o-130^o E, in the area in between south of Indonesian and Australian waters except for below 30^o S. (Figure 2).

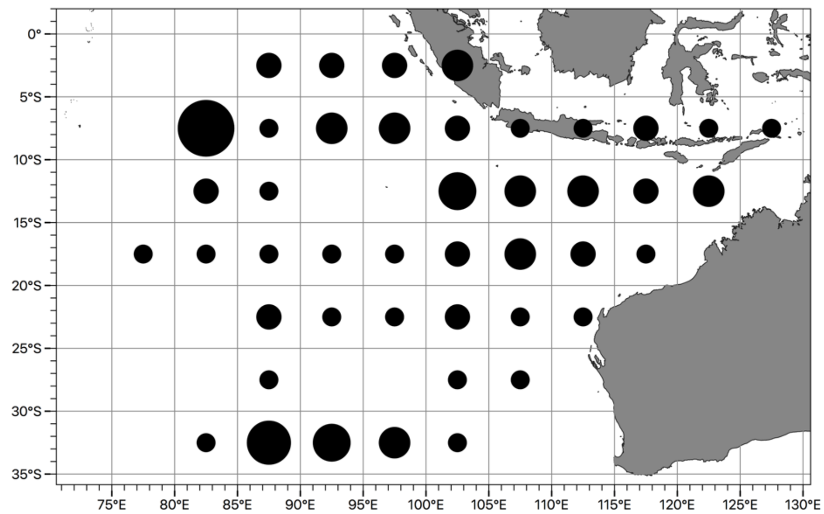

Figure 1. Catch-rate distribution of bigeye tuna from Indonesian longline fleet 2006-2018

Trend of nominal CPUE

In general, the catch rates of bigeye tuna during 2006-2013 were relatively stable (0.21$\pm$0.01 in average), rose substantially to around 0.29$\pm$0.02 and then dropped to merely just 0.11$\pm$0.01 in 2018. The lowest catch rate recorded was in 2016 (0.11$\pm$0.02), as the highest was in 2014 (0.29$\pm$0.02). On the other hand, the proportion of zero catch for bigeye tuna was varied annually between a minimum of 22% in 2012 and a maximum of 43% in 2018 with average proportion around 30% per year (Figure 3).

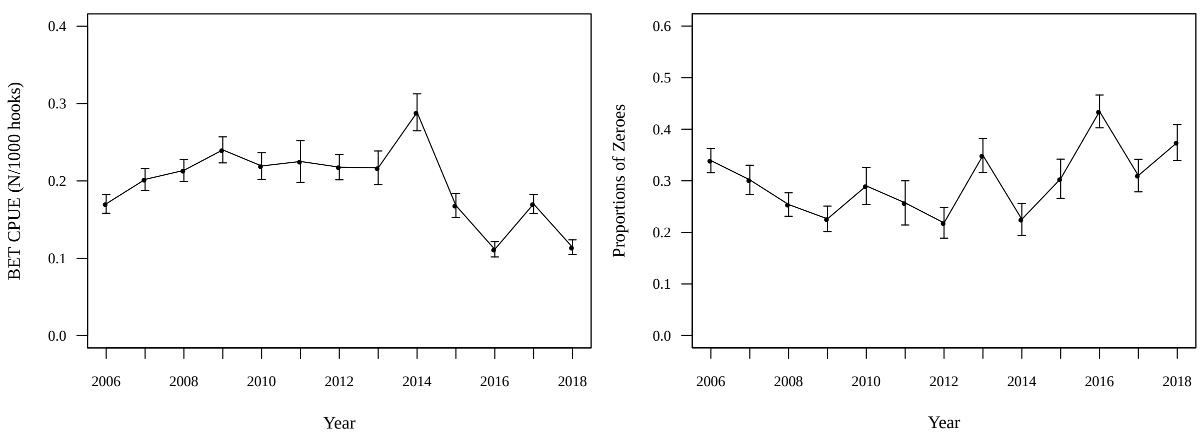

Figure 3. Nominal CPUE series (N/100 hooks) (left panel) and proportion of zero bigeye tuna catches from 2005 to 2018 (right panel). The error bars refer to the standard errors

CPUE standardization

Based on model selection, all effects were remained and statistically significant, except for moon phase. If we rely on AIC, simple negative binomial models (NB) was resulted as the best fit model, on the other hand, if we choose best fitting model based on the lowest BIC then tweedie (TW) should be chosen. However, the highest R^2 was produced by zero-inflated negative binomial model (ZINB). Delta-lognormal model produced a consistent value of R^2 between positive sub-model and proportion of positive catch sub-model (0.20-0.22), even though it should asked whether two structured model can be compared with single structured model (Table 2). Therefore, we decided to choose the simplest model (NB) as the best fit for describing the abundance of bigeye tuna.

Table 2. Summary of indicators as calculated using four model structures: Tweedie (TW), Negative Binomial (NB), Zero-inflated with Negative Binomial (ZINB), Delta-lognormal (DELTA). The terms in the column at left indicate: number of parameters (k), Akaike (AIC) and Bayesian (BIC) Information Criteria, log likelihood (logLik) and R^2.

Indicators	TW	NB	ZINB	DELTA	
				Lognormal	Delta
k	38	38	40	37	38
AIC	11877.06	**11579.37**	12014.95	4468.81	3105.99
BIC	**11801.06**	11813.28	12254.86	4677.47	3333.90
logLik	-5899.53	**-5750.68**	-5966.48	-2196.40	-1515.00
R^2	0.52	0.39	**0.53**	0.22	0.20

All variables were considered as statistically significant to the model, however, Year and AreaTree were the most influential ones in determining the catch rate of bigeye tuna. By contrast, moon phase and soak time were excluded from the model (Table 3). In terms of model validation, NB model seemed adequate for this particular situation with a low quantity of zeros (<30%), as the residual analysis, including the residuals distribution along the fitted values, the QQ plots and the residuals histograms, did not identified any major problems in the models (Figure 4).

Tabel 3. Deviance table of the parameters used for BET CPUE standardization for selected model (NB). Each parameter indicated the degrees of freedom (Df), the deviance (Dev), the residual degrees of freedom (Resid Df), the residual deviance (Resid. Dev), the F test statistic and the significance (p-value).

Parameters	Df	Deviance	Resid. Df	Resid. Dev	Pr(>Chi)	
NULL			2973.0	4284.2		
Year	12	179.8	2961.0	4104.4	0.0000	***
Quarter	3	59.7	2958.0	4044.8	0.0000	***
HBF	1	63.3	2957.0	3981.4	0.0000	***
Start_Set	1	34.2	2956.0	3947.3	0.0000	***
AreaTree	20	719.7	2936.0	3227.6	0.0000	***

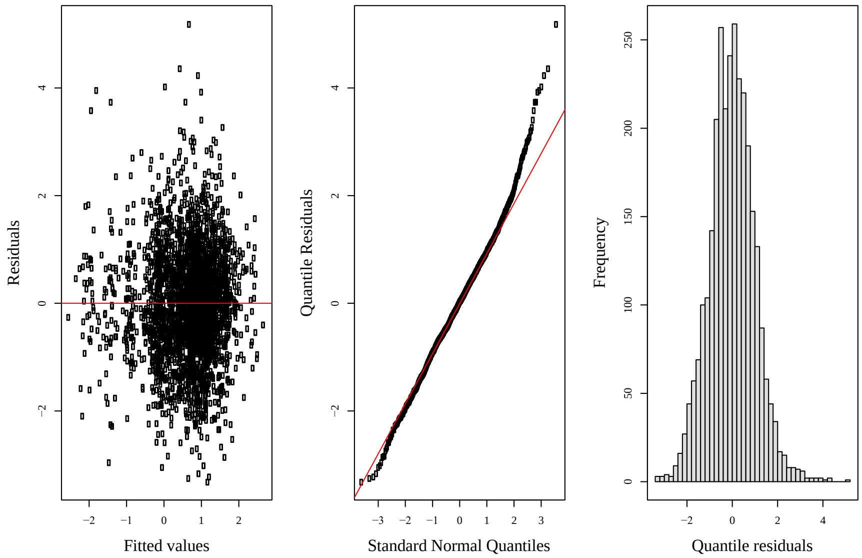

Figure 4. Residual analysis for the final NB model used for the bigeye CPUE standardization in the Indian Ocean. Left to right panel: the residuals along the fitted values on the log scale, the QQPlot and the histogram of the distribution of the residuals.

Estimations of standardized catch rates from four models are shown in Figure 5. On overall, the abundance of bigeye tuna was decreased quite substantially over the years (almost two-fold from the beginning of observation). Zero-inflated Negative Binomial (ZINB) failed to give plausible indices due to converge issue when area was included. Negative Binomial (NB), Tweedie (TW) and Delta-lognormal (DEL) give similar trend lines, however, between NB and TW projected somewhat identical indices. On the other hand, DEL gave higher estimation between 2006-2012 and lower prediction afterwards compared to previous two models. Hence, a NB model was chosen for CPUE standardization since it has the lowest AIC values among models (Figure 6).

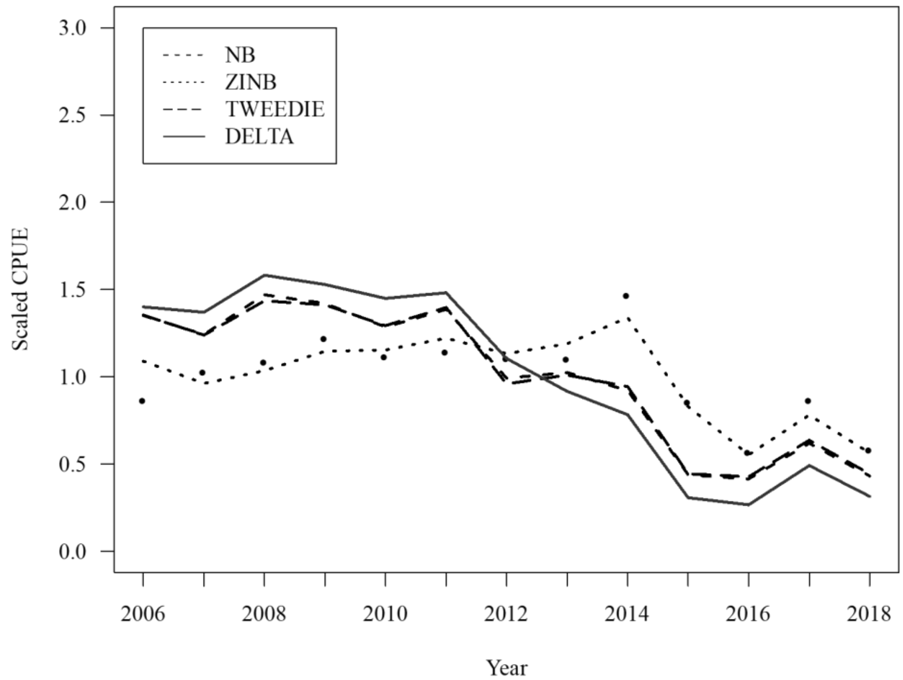

Figure 5. Standardize catch per unit effort (CPUE) of bigeye tuna calculated using various models. Values were scaled by dividing them by their means.

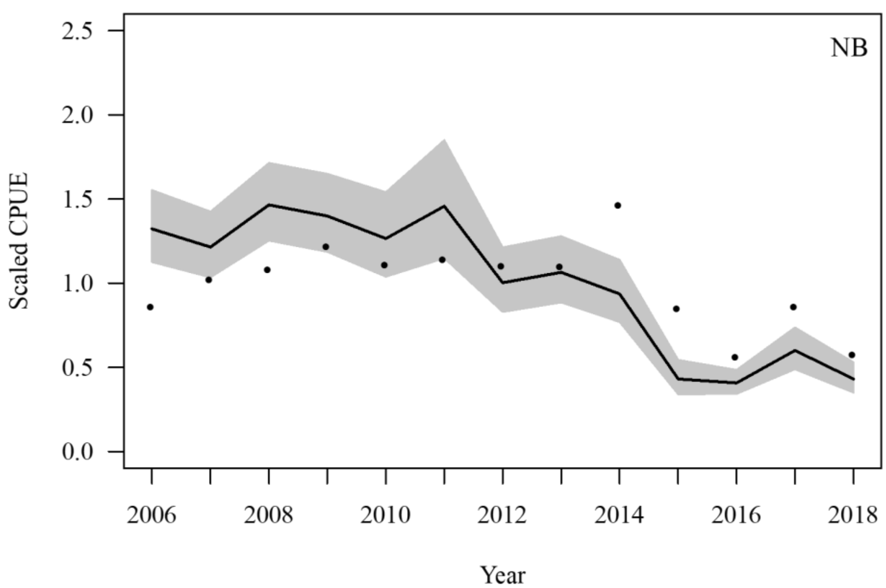

Figure 6. Final graph for standardized catch per unit effort (CPUE) of bigeye tuna calculated using NB model with 95% confidence interval (greyed area). Values were scaled by dividing their means.

Acknowledgement

The Authors would like to thank to all scientific observers of Research Institute for Tuna Fisheries (RITF) and national observers of Directorate General of Capture Fisheries (DGCF) 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 Centre for International Agricultural Research (ACIAR) and the Research Institute for Capture Fisheries (RCCF) for their funding support through research collaboration in the project FIS/2002/074: Capacity Development to Monitor, Analyze and Report on Indonesian Tuna Fisheries. The authors would also like to thank Dr. Humber Andrade, Dr. Rui Coelho and Dr. Sheng-ping Wang for their valuable input on developing the statistical analysis.

References

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

Collette, B.B., Nauen, C.E., 1983. FAO species catalogue. Volume 2. Scombrids of the world. An annotated and illustrated catalogue of tunas, mackerels, bonitos and related species known to date., FAO Fisheries Synopsis. Food and Agriculture Organization of the United Nations, Rome.

Fonteneau, A., Ariz, J., Delgado, A., Pallares, P., Pianet, R., 2005. A comparison of bigeye (Thunnus obesus) stocks and fisheries in the Atlantic, Indian and Pacific Oceans. Col Vol Sci Pap ICCAT 57, 41–66.

Giner, G., Smyth, G.K., 2016. statmod: probability calculations for the inverse Gaussian distribution. arXiv preprint arXiv:1603.06687.

Harrell Jr., F.E., Dupont, C., Others, 2018. Hmisc: Harrell Miscellaneous.

Hartaty, H., Setyadji, B., Fahmi, Z., 2018. Updated information on catch and effort of bigeye tuna (Thunnus obesus) from Indonesian tuna longline fishery. IOTC-2018-WPTT20-28.

Ichinokawa, M., Brodziak, J., 2010. Using adaptive area stratification to standardize catch rates with application to North Pacific swordfish (Xiphias gladius). Fisheries Research 106, 249–260.

IOTC-WPTT20, 2018. Report of the 20th Session of the IOTC Working Party on Tropical Tunas (No. IOTC–2018–WPTT20–R[E]). Indian Ocean tuna Commission (IOTC), Seychelles.

Jackman, S., 2017. pscl: Classes and Methods for R Developed in the Political Science Computational Laboratory. United States Studies Centre, University of Sydney, Sydney, New South Wales, Australia.

Lazaridis, E., 2014. lunar: Lunar Phase & Distance, Seasons and Other Environmental Factors.

Lee, P.-F., Chen, I., Tzeng, W.-N., 2005. Spatial and temporal distribution patterns of bigeye tuna (Thunnus obesus) in the Indian Ocean. Zoological studies 44, 260–270.

Lenth, R., 2018. emmeans: Estimated Marginal Means, aka Least-Squares Means.

Lüdecke, D., Makowski, D., Waggoner, P., 2019. Performance: assessment of regression models performance.

Nootmorn, P., 2004. Reproductive biology of bigeye tuna in the Eastern Indian Ocean. IOTC Proceedings 7, 1–5.

Nugraha, B., Baskoro, M.S., Pane, A.B., Nugroho, E., 2010. Genetic diversity of bigeye tuna (Thunnus obesus) based on mtDNA analysis with the PCR-RFLP technique. Indonesian Fisheries Research Journal 16, 25–32. https://doi.org/10.15578/ifrj.16.1.2010.25-32

Polacheck, T., 2006. Tuna longline catch rates in the Indian Ocean: Did industrial fishing result in a 90% rapid decline in the abundance of large predatory species? Marine Policy 30, 470–482.

QGIS Developer Team, 2018. QGIS Geographic Information System. Open Source Geospatial Foundation Project.

R Core Team, 2019. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria.

Sadiyah, L., Dowling, N., Prisantoso, B.I., 2012. Developing recommendations for undertaking CPUE standardisation using observer program data. Indonesian Fisheries Research Journal 18, 19–33. https://doi.org/10.15578/ifrj.18.1.2012.19-33

Sadiyah, L., Dowling, N., Prisantoso, B.I., 2011. Changes in fishing pattern from surface to deep longline fishing by the Indonesian vessels operating in the Indian Ocean. Ind.Fis.Res.J 17, 87–99.

Schwarz, G., 1978. Estimating the dimension of a model. The annals of statistics 6, 461–464.

Setyadji, B., Bahtiar, A., Nugraha, B., 2012. Analysis of Sex-Ratio By Length Class of Bigeye Tuna (Thunnus obesus) in The Indian Ocean. Widyariset 15, 593–598.

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