Standardized CPUE of the Indian Ocean black marlin (*Istiompax indica*) caught by Japanese longline fisheries.

Hirotaka Ijima
ijima@affrc.go.jp

National Research Institute of Far Seas Fisheries, Fishery Research Agency,
Shimizu, Shizuoka, Japan

Highlights

1. Using Japanese longline logbook data, I calculated the standardized CPUE of the Indian Ocean black marlin (*Istiompax indica*). In this study, I changed the methodology of the previous study such as analysis period, analysis area, and statistical model for CPUE standardization.

2. I used the Bayesian Information Criterion (BIC) and likelihood ratio test for the model selection. As a result, the zero-inflated negative binomial with mixed effect model was selected.

3. Using the Pearson residuals, I diagnosed goodness fit for the selected model, while there is no typical trend against explanatory variables. Thus, it was considered that the selected model was well estimated.

4. After 2010, the operation area of Japanese longline has been shrunk. To evaluate the shrinking effect, I compared the two time-period standardized CPUE (1994-2010 and 1994-2017). There is no substantial difference between the two indices. However, CPUE after 2011 may not represent the biomass trend of the Indian Ocean Black marlin.

Abstract

To calculate standardized CPUE of the Indian Ocean black marlin, I analyzed Japanese longline logbook data. In this study, I changed three points of standardization methodology of the previous study. 1) I used shorter period datasets (1994-2017) because Japanese logbook was changed around 1994 and datasets of early period includes large uncertainty such as species definition. 2) I used a different area definition considering size distribution. 3) I used the zero-inflated negative binomial distribution rather than delta log-normal model because catch number is countable data. I also addressed model selection and validation. The selected model was well estimated, but there is substantial uncertainty after 2011.

Introduction

The IOTC Working Party on Billfish (WPB) conducted the stock assessment of black marlin (*Istiompax indica*) in the Indian Ocean. In this stock assessment, production models such as ASPIC and BSPM were used (Yokoi and Nishida 2016, Andrade 2016). Japanese longline CPUE is one of the most critical indices in the Black marlin stock assessment in the Indian Ocean. However, there are three issues in the previous CPUE analysis (Yokoi et al., 2016). Firstly, it is better not to use early period logbook data, because Japanese logbook format changed around 1994 and early period logbook data includes a lot of uncertainty such as species discrimination. Secondly, Yokoi et al., 2016 used core area where Japanese longliners have operated a lot, however, it is better to consider size-dependent area definition for CPUE standardization (Ochi et al., 2016). Because, usually, juvenile catch affects CPUE trend mainly. Finally, a discrete probability distribution such as Poisson distribution should be applied for GLM standardization, because log-normal CPUE lose the impact of fishing effort. To consider random effect is also essential for CPUE standardization (Ijima 2017 a, b). Here, I defined the size-based analysis area and explored the CPUE standardization for the Indian Ocean Black marlin using negative binomial GLM, the negative binomial GLMM, and the zero-inflated negative binomial GLMM.

Material and Methods

Data sets

I used Japanese longline logbook data for the CPUE standardization. Japanese logbook data has been reported the detail of the operation of the longline fishery that resolution is 1x1 grid scale. The format of the logbook was changed around 1994 and early period data includes large uncertainties. For instance, black marlin might be misunderstood as blue marlin in the old datasets. Thus, I used later period logbook data for this analysis. The fishing ground of Japanese longline spread throughout the Indian Ocean from the 1990s to the 2000s, but after 2010, due to the influence of pirates, the fishing ground has shrunk rapidly (Figure 1). There is large zero catch area in the Southeast Indian Ocean (Figure 1). CPUE shows low value and distribute throughout in the Northwest to Southwest area mainly (Figure 1). In all period, large-size black marlin has appeared around 10S (Figure 2). Considering such as spatial CPUE and size information, I defined the analysis area (Figure 3). In this area, Japanese longliners have caught similar size black marlin in all time.

Statistical models

I used Zero-Inflated negative binomial glmm (ZINB-GLMM) for the CPUE standardization. The explanatory variables of fixed part are the year, quarter, and gear. The area and fleet effect treat as random effect. Considering the random effect is appropriate because there are a lot of variables for the vessel name and 5x5 area effect. The random effect model can also remove the pseudo replication by vessel and operating area. The changes in gear configuration show two modes (Figure 4), thus I defined two gear type as shallow or deep sets. In this study, all explanatory variables were treated as the categorical variables. Almost sets could not catch Black marlin (Figure. 5). Thus, it needs to use the Zero-Inflated model because the Zero-Inflated model can estimate "true" zero catch.

To treat such technical needs, I used R software package glmmTMB. glmmTMB can estimate the parameters of the complex model using the Template Model Builder (Brooks et al., 2017). To select an appropriate statistical model, I also considered simpler model (e.g., GLM and GLMM) and I evaluated these models using Bayesian information criterion (BIC) and likelihood ratio test. Finally, I calculated the standardized black marlin CPUE using the R software package lsmeans (Lenth 2016). Lsmeans needs additional R script from the GitHub (https://github.com/glmmTMB) to apply glmmTMB. The ZINB-GLMM is

$$Catch_i \sim ZINB(\pi_i, \mu_i, k),$$

$$E(Catch_i) = \mu_i(1 - \pi_i),$$

$$\mathrm{var}(Catch_i) = (1 - \pi_i)\mu_i(1 + \pi_i\mu_i + \mu_i/k),$$

$$\log(\mu_i) = \beta_0 + \mathbf{X}_i\boldsymbol{\beta} - \log(hooks_i) + v_i + a_i,$$

$$\mathrm{logit}(\pi_i) = \gamma_0 + \mathbf{Z}_i\boldsymbol{\gamma},$$

$v_i \sim N(0, \sigma^2_{vessel}),\ a_i \sim N(0, \sigma^2_{area}),$

here π_i is the probability of zero catch by set i. π_i is estimated by logit link function that the variable matrix is $\mathbf{Z}_i$ and the covariate vector is $\boldsymbol{\gamma}$ respectively. $Catch_i$ is the catch number by set i. μ_i is expected catch number by set i. k is the dispersion parameter. The link function was used for log function. β_0 is the intercept, $\mathbf{X_i}$ is the matrix of variables, $\boldsymbol{\beta}$ is the covariates vectors, and hooks denote the hooks/1000 of the operation respectively. I applied the random effect for vessel name and 5x5 area (v_i and a_i) in catch model. I did not use the interaction for all models to avoid overfitting. However, there might be a correlation between year-area or year-fleet. To treat annual interactions for the vessels, I addressed the random slope GLMM but, this model did not converge.

Result and Discussion

I selected ZINB-GLMM that BIC and AIC were the lowest between other candidate models (Table 1). The deviance of the selected model is also significantly low (Table 1). The standardized CPUE of the Indian Ocean black marlin is different from nominal CPUE (Table 2, Figure 6 (a)). Especially, standardized CPUE dropped between 1994 and 2000 and after 2010, the standardized CPUE below the nominal CPUE (Figure 6(a)). To evaluate the shrinking effect of Japanese longliners, I calculated the two time-period standardized CPUE (1994-2010 and 1994-2017). There is no substantial difference between the two indices. However, CPUE after 2011 may not represent the biomass trend of the Indian Ocean Black marlin because of lack of time-spatial operational data.

I calculated Pearson residuals and summarized by each variable for the model validation (Figure 6 (b)-(e), Figure 7). There are no typical trends in Pearson residuals (Figure 6 (b)-(e), Figure 7). Thus, it was considered that the selected model was well estimated. If there are some trends spatially, it needs to address the geostatistical model, but in this study, there is no trend. However, the geostatistical model might be useful for CPUE standardization because the geostatistical model is a strong tool for time-spatial datasets that includes a lot of lack such as current Japanese longline logbook data.

References

Yokoi, H., and Tom, N. 2016. Stock assessments of black marlin (*makaira indica*) in the Indian Ocean using A Stock-Production Model Incorporating Covariates (ASPIC). IOTC–2016–WPB14–24_Rev1, 13 pp.

Andrade, H.A. 2016. Preliminary stock assessment of black marlin (*Makaira indica*) caught in the Indian Ocean using a Bayesian state-space production model. IOTC–2016–WPB14–28, 16 pp.

Yokoi, H., Yasuko S., Keisuke S., and Tom N. 2016. Standardization of catch rate for black marlin (*Istiompax indica*) exploited by the Japanese tuna longline fisheries in the Indian Ocean (1971-2015). IOTC–2016–WPB14–19 Rev 1, 17pp.

Ochi, D., Ijima, H., Kinoshita, J., and Kiyofuji H. 2016. New fisheries definition from Japanese longline North Pacific albacore size data. ISC/16/ALBWG-02/03, 11pp.

Ijima, H[a]. 2017 CPUE standardization of the Indian Ocean swordfish (*Xiphias gladius*) by Japanese longline fisheries: Using negative binomial GLMM and zero inflated negative binomial GLMM to consider vessel effect. IOTC-2017-WPB15-19, 32pp.

Ijima, H[b]. 2017. CPUE standardization of the Indian Ocean Striped marlin (*Tetrapturus audax)* by Japanese longline fisheries: Using negative binomial GLMM and zero inflated negative binomial GLMM to consider vessel effect. IOTC–2017–WPB15–31, 32pp.

Brooks, ME., Kristensen, K., van Benthem, KJ., Magnusson, A., Berg, CW., Nielsen, A., Skaug, HJ., M˙achler, M., Bolker, B,M. 2017 Modeling zero-inflated count data with glmmTMB. bioRxiv preprint first posted online. pp 14.

Lenth, RV., 2016 Least-Squares Means: The R Package lsmeans. Journal of Statistical Software. 69-1 pp 33.

Table 1. Deviance table for black marlin CPUE in the North East Indian Ocean by Japanese longline fishery (1994-2017). Bold is the best model selected by BIC.

Models	Df	AIC	BIC	deviance	Chisq	Pr(>Chisq)
yr+offset(log(hooks/1000))	25	108286	108542	108236	-	-
yr+qtr+offset(log(hooks/1000))	28	107851	108139	107795	440.1073	<0.001
yr+qtr+gear+offset(log(hooks/1000))	29	107845	108143	107787	8.7735	0.003
yr+qtr+gear+(1|area)+offset(log(hooks/1000))	30	106597	106905	106537	1249.401	<0.001
yr+qtr+gear+(1|area)+(1|fleet)+offset(log(hooks/1000))	31	104885	105204	104823	1714.1372	<0.001
yr+qtr+gear+(1|area)+(1|fleet)+offset(log(hooks/1000)) yr	55	104826	105391	104716	107.3819	<0.001
yr+qtr+gear+(1|area)+(1|fleet)+offset(log(hooks/1000)) yr+qtr+gear	**59**	**103525**	**104130**	**103407**	**1309.1765**	**<0.001**

Following models were not converge.

yr + qtr + gear + (1|area) + (yr|fleet) + offset(log(hooks/1000)

yr + qtr + gear + (1|area) + (1|fleet) + offset(log(hooks/1000)), yr+qtr

yr + qtr + gear + (1|area) + (1|fleet) + offset(log(hooks/1000)) yr + qtr + gear + (1|area)

yr + qtr + gear + (1|area) + (1|fleet) + offset(log(hooks/1000)), yr + qtr + gear + (1|fleet)

yr + qtr + gear + (1|area) + (1|fleet) + offset(log(hooks/1000)), yr + qtr + gear + (1|area) + (1|fleet)

Table 2. Standardized CPUE of Black marlin caught by Japanese longline in the Indian Ocean (1994-2017). Standardized CPUE was calculated by the least square means.

Year	Nominal CPUE	Standardized CPUE	Lower	Upper
1994	0.057	0.173	0.138	0.217
1995	0.042	0.107	0.085	0.134
1996	0.034	0.087	0.070	0.108
1997	0.053	0.086	0.072	0.104
1998	0.082	0.107	0.091	0.126
1999	0.082	0.110	0.092	0.131
2000	0.048	0.054	0.044	0.065
2001	0.042	0.043	0.035	0.053
2002	0.033	0.045	0.037	0.056
2003	0.056	0.065	0.054	0.078
2004	0.056	0.058	0.048	0.069
2005	0.047	0.041	0.035	0.049
2006	0.096	0.066	0.057	0.077
2007	0.090	0.061	0.053	0.072
2008	0.074	0.047	0.040	0.055
2009	0.080	0.057	0.047	0.068
2010	0.161	0.091	0.075	0.110
2011	0.145	0.083	0.069	0.101
2012	0.145	0.096	0.080	0.116
2013	0.146	0.101	0.083	0.124
2014	0.124	0.068	0.056	0.083
2015	0.154	0.072	0.059	0.087
2016	0.198	0.097	0.079	0.119
2017	0.178	0.085	0.070	0.104

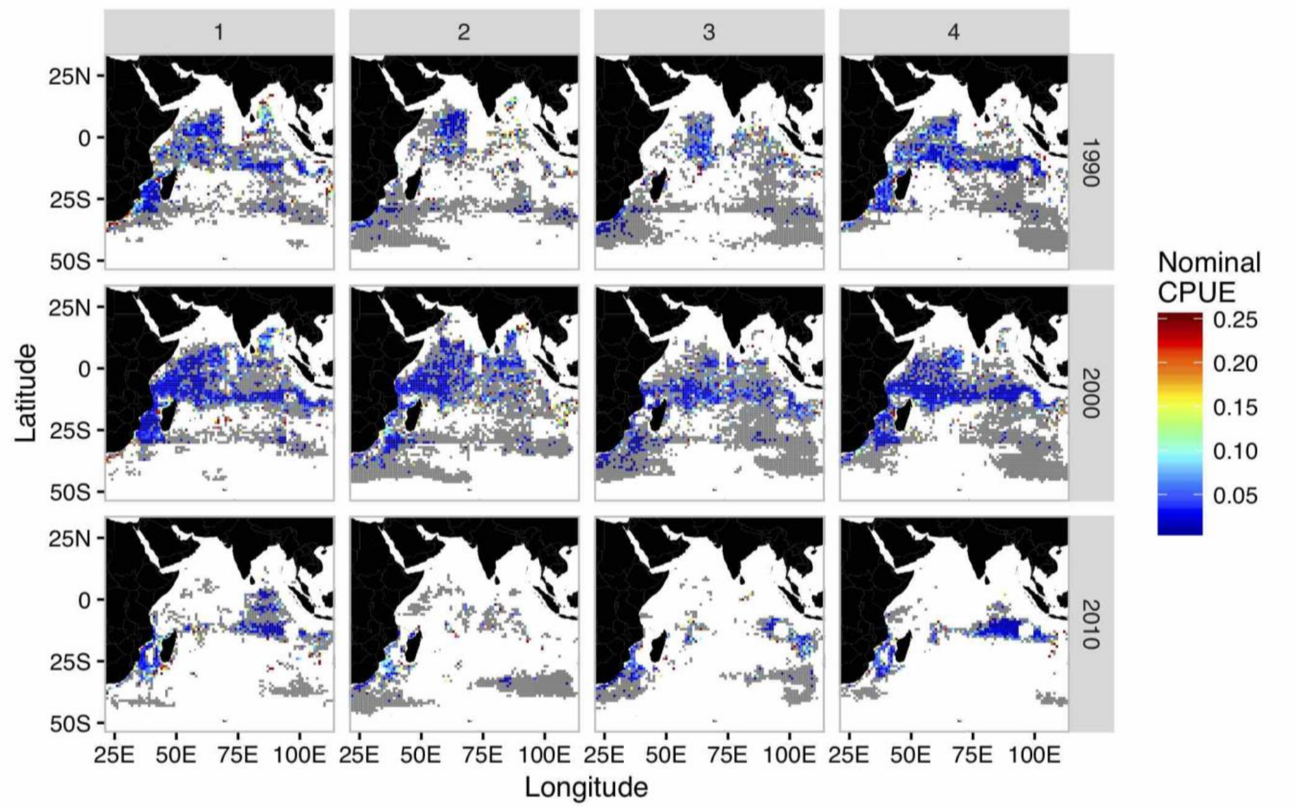

Figure 1. Changes in time-spatial CPUE of Black marlin in the Indian Ocean.

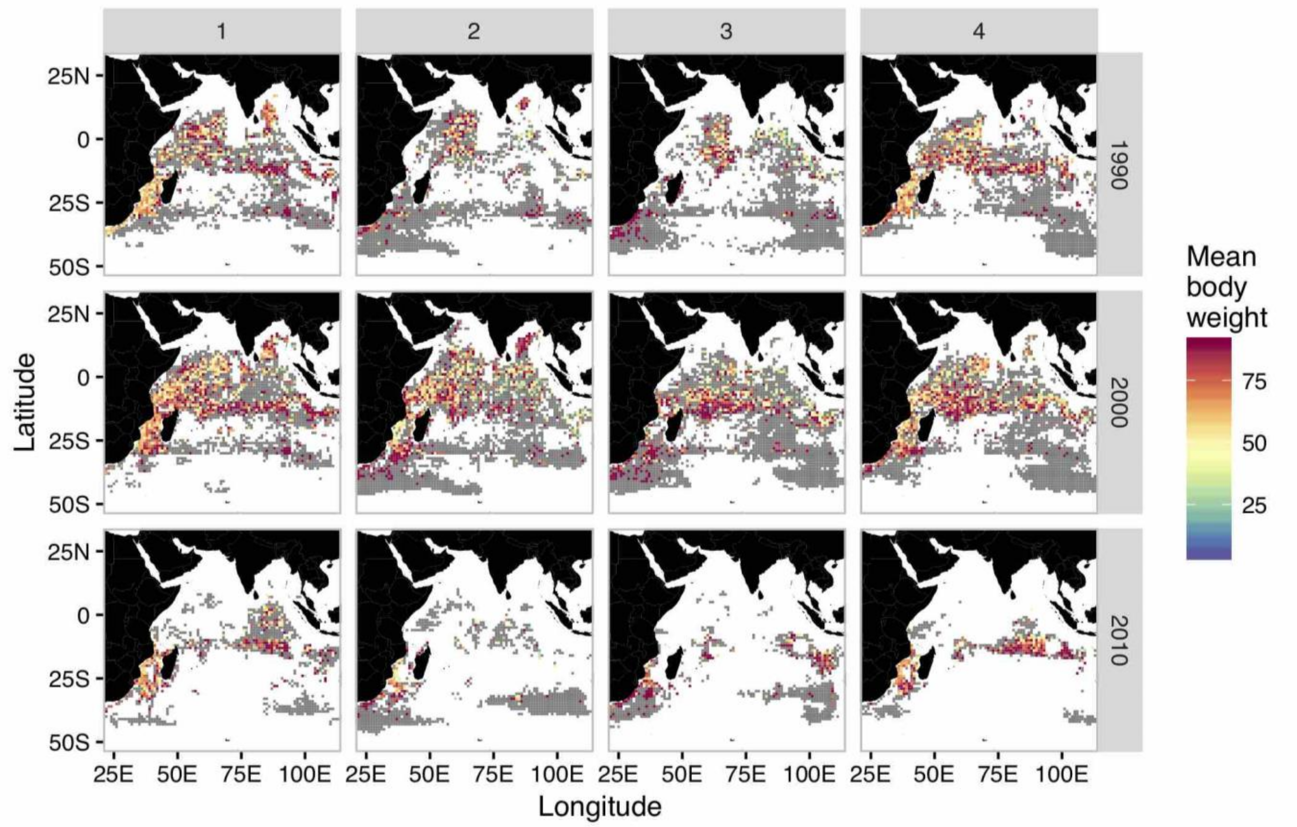

Figure 2. Changes in time-spatial mean body weight of Black marlin caught by Japanese longline vessel.

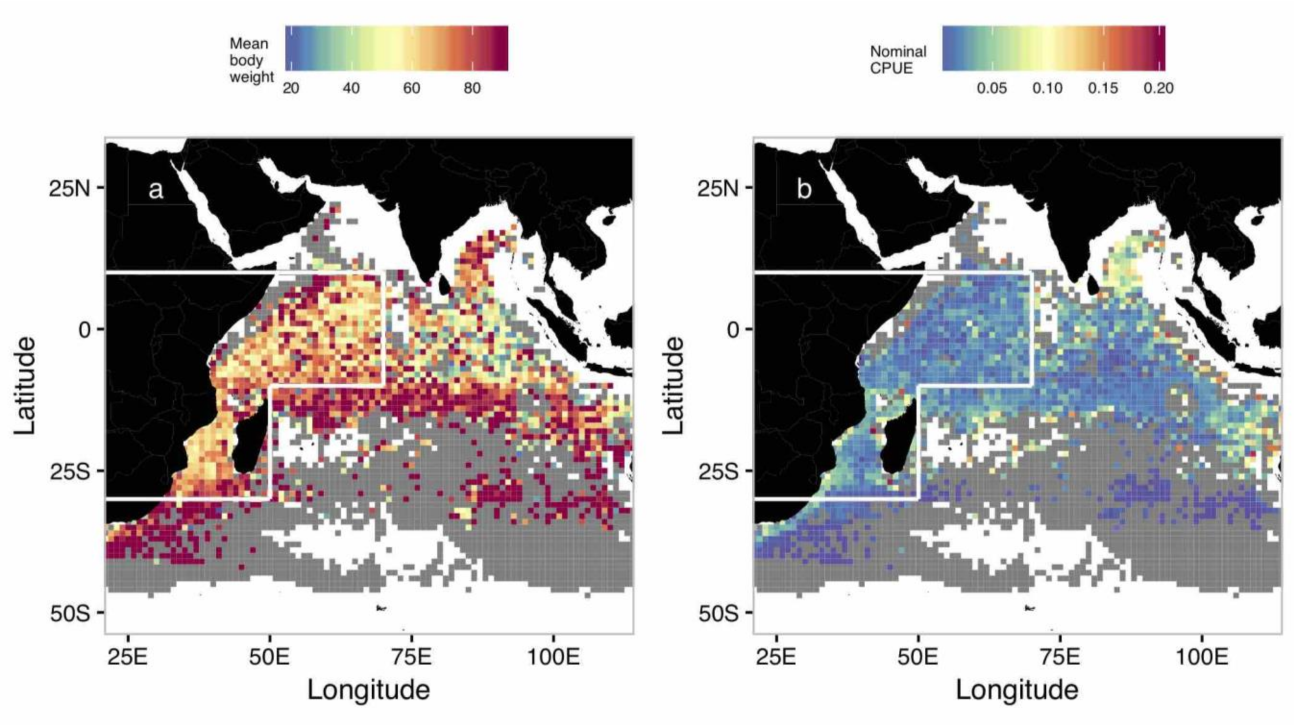

Figure 3. The analysis area for CPUE standardization (inside of white line). I defined the analysis area that a) mean body weight and b) CPUE show a similar trend.

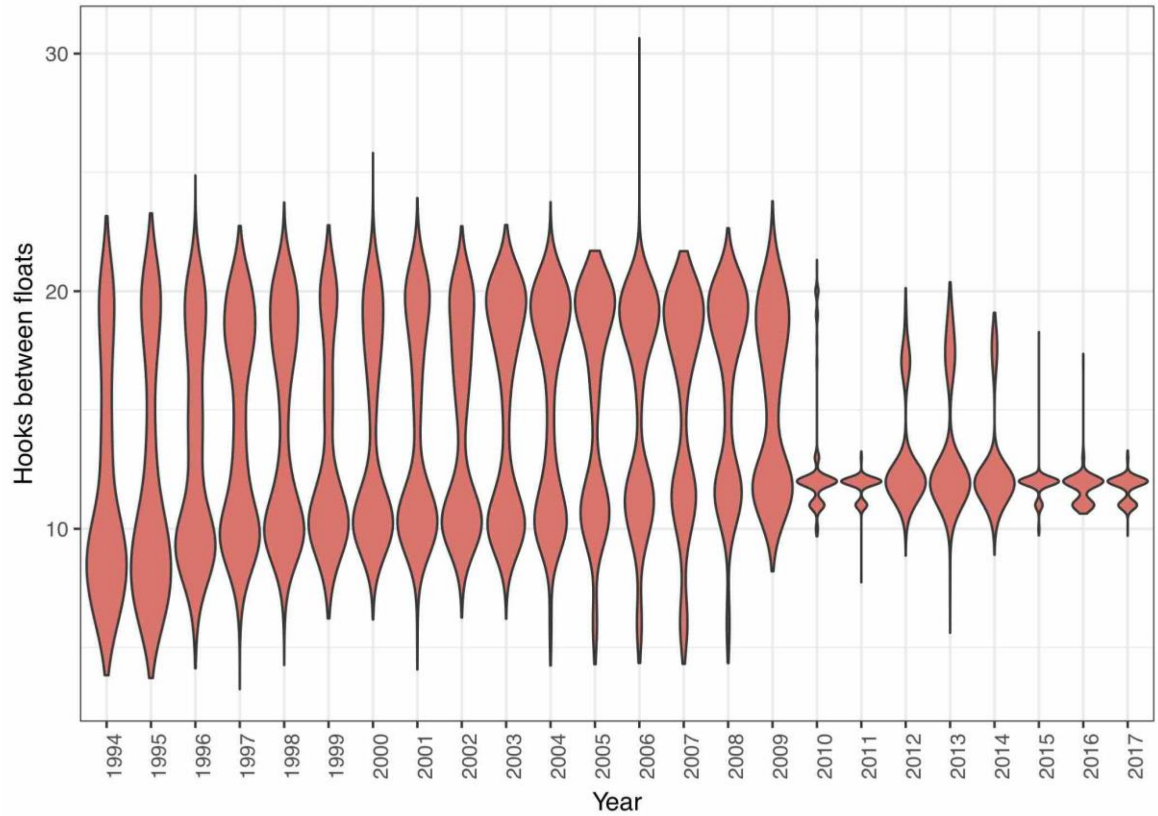

Figure 4. The historical change of hooks between floats. I set two type gear configurations (deep or shallow sets) that boundary is fifteen hooks between floats.

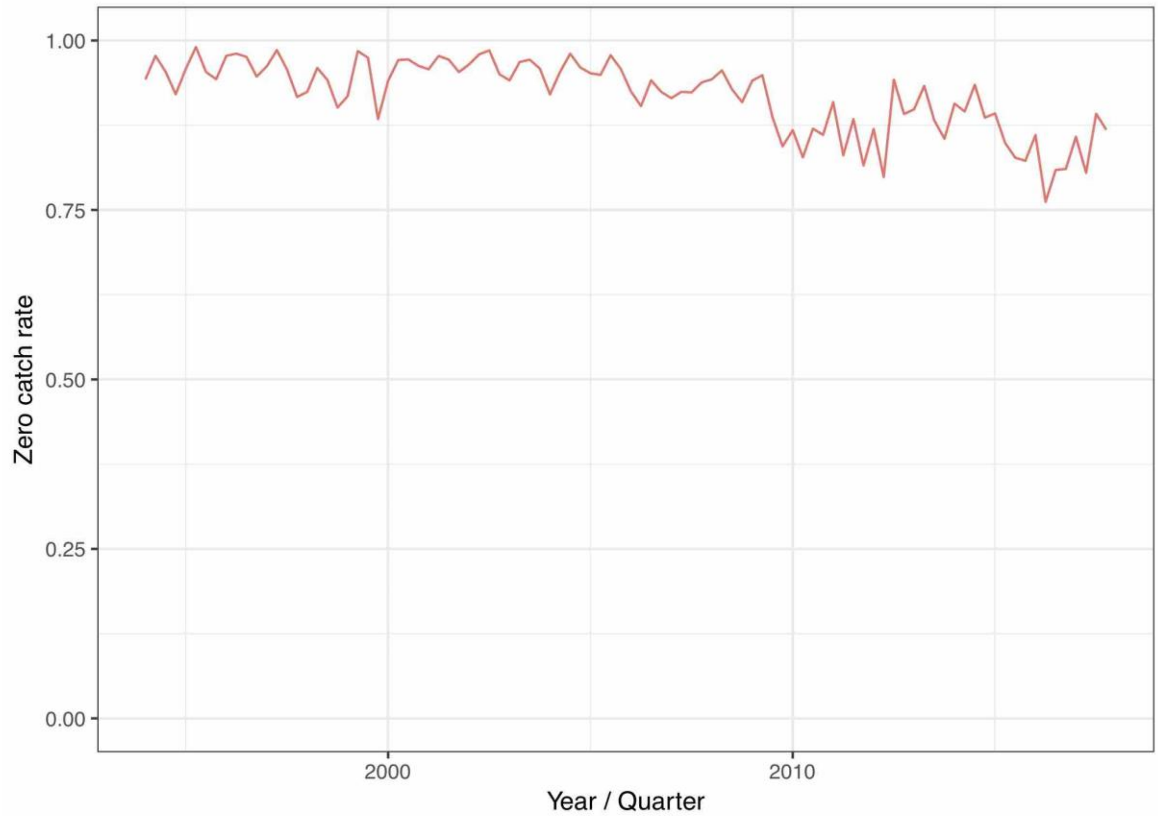

Figure 5. The historical change of zero catch ratio of black marlin.

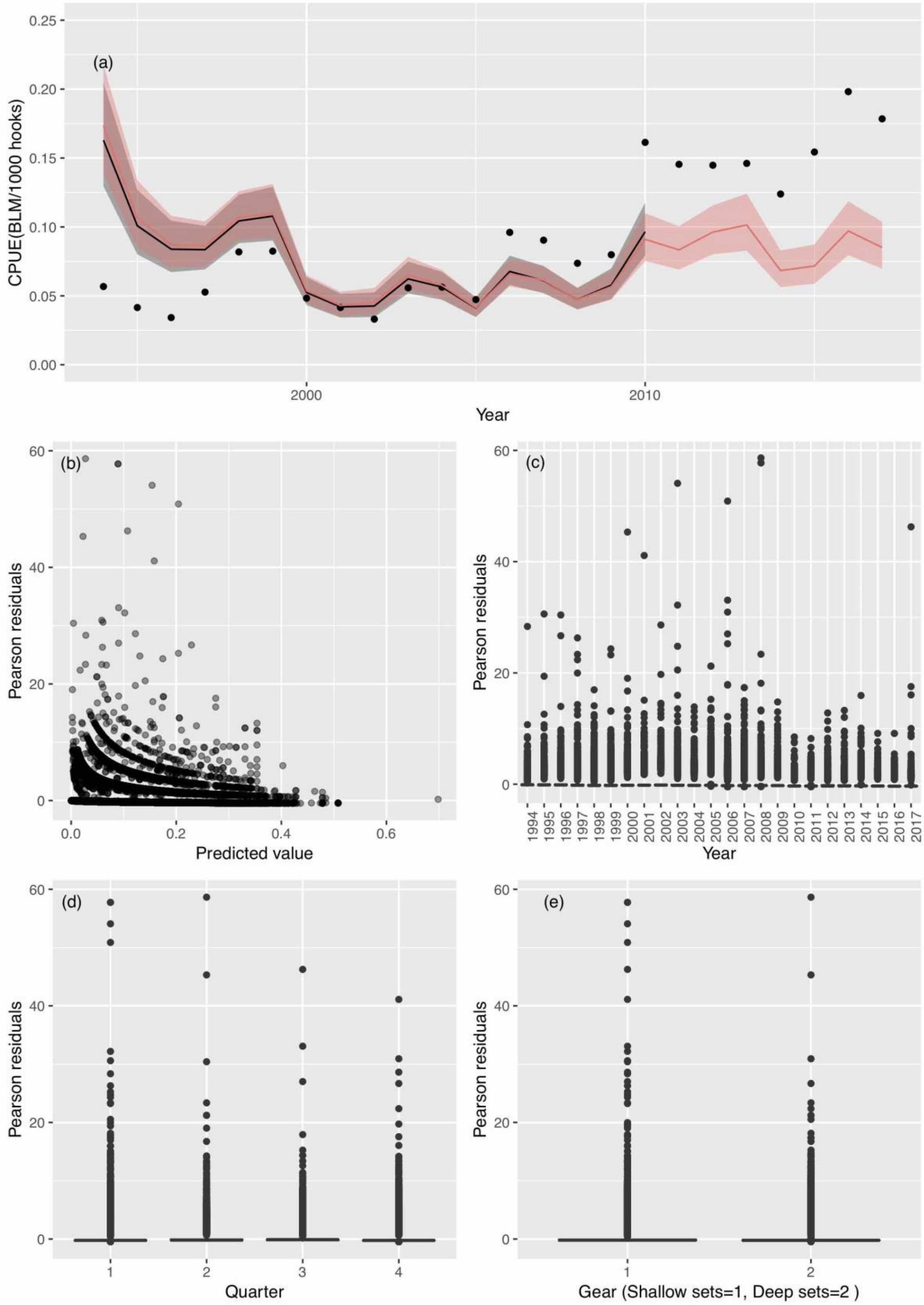

Figure 6. The result of CPUE standardization analysis of black marlin caught by Japanese longline vessel. (a) The comparison between nominal and standardized CPUE. Red line denotes standardized CPUE, Points are nominal CPUE, and filled areas denote 95% confidence interval of standardized CPUE. Black line and filled area show the result using different period data (1994-2010). (b)-(e) The trend of the Pearson residuals.

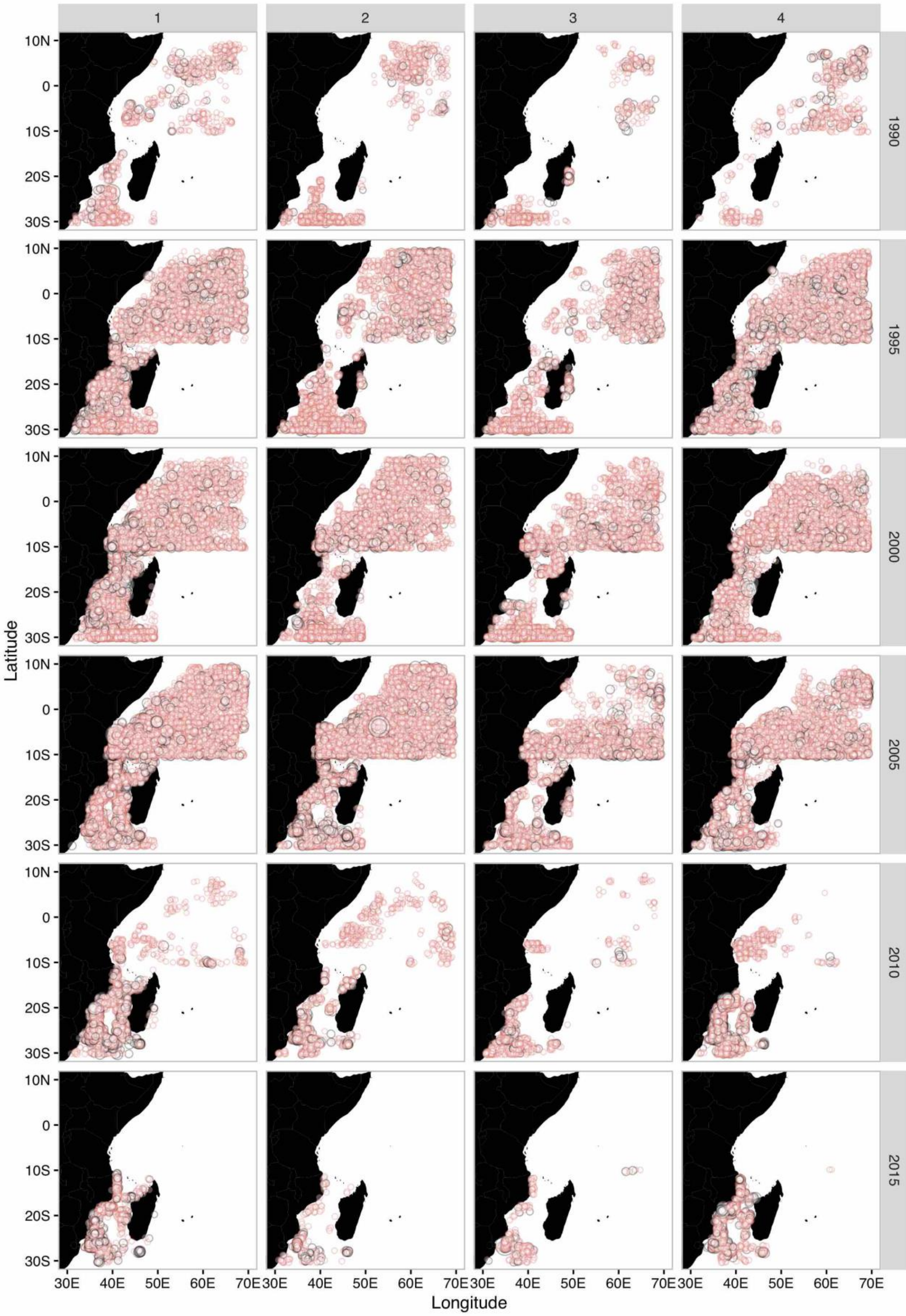

Figure 7. Time spatial change of Pearson residuals. Red circles are positive residuals, and black circles are minus residuals. Size of circle correspond to magnitude of Pearson residuals.