Japanese Longline CPUE Standardization (1979-2019) for black marlin (*Makaira indica*) in the Indian Ocean using Bayesian hierarchical spatial model

Kenji Taki*, Hirotaka Ijima, and Mikihiko Kai
*takisan@affrc.go.jp

Fisheries Resources Institutes (FRI), Japan Fisheries Research and Education Agency

Summary

To estimate a historical trajectory of black marlin stock abundance in the Indian Ocean, we standardized the CPUE of black marlin caught by Japanese longliners for 1979-2019. We defined the same area of analysis based on the spatial distribution of the mean body weight as Ijima (2018), and divided the time-period into two periods, 1979-1993 and 1994-2019. In this analysis, we applied Bayesian hierarchical spatial models. Since the catch data is countable and characterize by many zeros, we used zero-inflated Poisson generalized linear mixed model (ZIP-GLMM). All analyses were performed using R, specifically the R-INLA package. The INLA procedure, in accordance with the Bayesian approach, calculates the marginal posterior distribution of all random effects and then estimates parameters involved in the model. We applied a half Cauchy distribution as a prior for the random effect. Best model was selected from multiple models using Widely Applicable Bayesian Information Criterion (WAIC) for the defined area in each period. Gradual annual decline trend with interannual variation were generally observed for the standardized CPUEs during 1979-1993, while stable annual trends were observed for the standardized CPUEs during 1994-2019. The 95% credible intervals were wider due to the inclusion of spatial effect as compared to the previous non-spatial model for 1994-2017 (Ijima 2018), while the point estimates of the standardized CPUE trends were similar. To reduce the uncertainty in the estimation of the standardized CPUEs, selecting an appropriate catchability (q), applying the state space model and/or latent variable model will be essential to improve the stock assessment of this species.

1. Introduction

The IOTC Working Party on Billfish (WPB) conducted the stock assessment of black marlin (*Makaira indica*) in the Indian Ocean. In this stock assessment, production models such as ASPIC and BSPM (Yokoi and Nishida 2016, Andrade 2016) were used. Ijima (2018) standardized the CPUE of black marlin caught by Japanese longliners in the Indian Ocean using zero-inflated Negative Binomial generalized linear mixed model (ZINB-GLMM) without considering the spatial random effect. It is generally thought that the abundance indices of Japan are very critical for the stock assessment.

Integrated nested Laplace approximations (INLA) methodology and its powerful application to the modelling of complex datasets has recently been introduced to wider nontechnical audience (Illian et al. 2013). As opposed to Markov Chain Monte Carlo (MCMC) simulations, INLA uses an approximation for inference and hence avoids the intense computational demands, convergence, and mixing problems that sometimes encountered by MCMC algorithms (Rue and Martino 2007). Additionally, R-INLA includes the stochastic partial differential equations (SPDE) approach (Lindgren et al. 2011) which is another statistical development. This approach enables us to model spatial random effect (Gaussian random field, GRFs) and to construct flexible fields that are better adept to handle datasets with complex partial structure (Lindgren 2013). This is often the case with fisheries data, since fishermen tend to aggregate particular fishing grounds, resulting in clustered spatial patterns and a lack of data at large regions. Together, these new statistical methods and their implementation in R allows scientists to fit considerably faster and more reliably complex spatiotemporal model (Rue et al. 2009, Cosandey-Godin et al. 2015).

The aim of this paper is to estimate the annual trends in abundance indices of black marlin (*Makaira indica*) caught by Japanese longliners in the Indian Ocean from 1979 to 2019. A zero-inflated Bayesian hierarchical approach is applied in consideration with spatial changes in the fishery and the species.

2. Materials and methods

Data sets

Japanese longline logbook data was used for the CPUE standardization of black marlin in the Indian Ocean. The logbook data has information about the resolution of fishing location at 1 x 1 degree grid scale. We used the data from 1979 onwards because the number of hooks between floats and the vessel name, which largely affect the CPUE standardization, are completely available since then. We divided the time-period into two periods, 1979-1993 and 1994-2019, as the gear configuration of Japanese longline fishery such as number of hooks between floats had drastically changed in the early period of 1990s. At the same time, the quality and quantity of logbook data were improved by adding new items to the logsheet as well. We defined the same area of the analysis as Ijima (2018), considering spatial CPUE and body weight information (Figure 1). In this area, Japanese longliners tended to catch similar body weight of black marlin in all time. Japanese longliners have operated throughout the Indian Ocean from the 1990s to the 2000s, but in 2010s, the fishing ground was shrunk rapidly (Figure 2). There are two main reasons for that the influence of pirates in the northwest

Indian Ocean, and the target shift of fishermen to southern bluefin tuna in the Southern Indian Ocean. The target shift makes it difficult to catch black marlins staying frequently in the shallower depths.

Statistical models

We applied Bayesian hierarchical spatial models, but we did not directly consider the spatiotemporal effects in the model because this approach is computationally intensive and the Widely Applicable Bayesian Information Criterion (WAIC; Watanabe, 2012) did not differed so much between spatial and spatiotemporal models in the preliminary analysis. Since the catch data is countable and characterized by many zeros (Figure 3), we used a zero-inflated Poisson GLMM (ZIP-GLMM). The zero-inflated model is useful because it can estimate "true" zero catch. As an alternative way, it is possible to use ZINB-GLMM) but we did not use the model because the ZINB tended to cause underdispersion (Ijima and Kanaiwa, 2019).

The explanatory variables of fixed effect are year (yr) and quarter (Jan-Mar, Apr-Jun, Jul-Sep, Oct-Dec; qtr), and those of random effect are area (5 x 5 degree scale; latlon), month (month), vessel (jp_name), and gear configuration (number of hooks between floats; hpb). The hpb increased remarkably in the early period of 1990s in the defined area (Figure 4). Most variables were treated as the categorical variable, but the autoregressive model (AR1) was applied to year effect for two spatial models to consider the autocorrelation. The latest SPDE models using AR1 tended to show smaller WAIC as compared to those using year as fixed effect (e.g., Ijima and Koike 2020). The use of these random effects in the model seems more appropriate to raise the accuracy of estimation (Ijima and Kanaiwa 2019). The random effects are also expected to remove the pseudo-replication by each effect (vessel, gear configuration, month, and area).

All analyses were performed using R, specifically the R-INLA package. The INLA procedure, in accordance with the Bayesian approach, calculates the marginal posterior distribution of all random effects and parameters involved in the model. We applied a half Cauchy distribution as a prior for the random effect. We plot latent spatial field to indicate the expected CPUE distribution. Best candidate model was selected based on WAIC for the defined area in each period.

3. Result and discussion

We compared the WAIC among seven different structurer's models for the defined area and each period (Table 1). The best model (yellow marker) was selected based on the lowest WAIC.

The predicted CPUE was higher in the northwestern coastal part in the defined area during 1979-1993, while that was lower for the same part during 1994-2019 (Figure 5).　The annual predicted CPUE showed a gradual decline trend for 1979-1993, while that showed no apparent trend was observed for 1994-2019. (Figure 6, Table 2).

Figure 7 showed a comparison of annual changes in standardized CPUE between present and previous studies (Ijima 2018) for the defined area. The annual trends in the point estimates were almost similar between them. However, the 95% credible intervals for the present models were wider than that for the previous model.

We will improve the accuracy of the stock assessments for black marlin in the Indian Ocean through exploring the appropriate catchability (q) in the model, applying the state space model (e.g., Yin et al. 2019) and/or latent variable model (e.g., Warton et al. 2015) in the future.

4. References

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.

Cosandey-Godin, A., Krainski, E. T., Worm, B. and Flemming, J. M. 2015. Applying Bayesian spatiotemporal models to fisheries bycatch in the Canadian Arctic. Can. J. Fish. Aquat. Sci. 72: 186–197. dx.doi.org/10.1139/cjfas-2014-0159.

Ijima, H. 2018. Standardized CPUE of the Indian Ocean black marlin (*Makaira indica*) caught by Japanese longline fisheries. IOTC-2018-WPB16-26.

Ijima, H. and Kanaiwa, M. 2019. Japanese longline CPUE of the striped marlin (*Kajikia audax*) in the WCNPO. ISC/19/BILLWG-01/07.

Ijima, H. and Koike, H. 2020. Preliminary analysis for the CPUE standardization of the Pacific blue marlin using Japanese longline logbook and the R software package R-INLA. ISC/20/BILLWG-03/02.

Illian, J.B., Martino, S., Sørbye, S.H., Gallego-Fernández, J.B., Zunzunegui, M., Paz Esquivias, M., and Travis, J.M.J. 2013. Fitting complex ecological point process models with integrated nested Laplace approximation. Methods Ecol. Evol. 4(4): 305–315. doi:10.1111/2041-210x.12017.

IOTC. 2014. Report of the 12th session of the IOTC working party on billfish. IOTC-2014-WPB-R [E].

Lindgren, F., and Rue, H. 2013. Bayesian spatial and spatio-temporal modelling with R-INLA. Journal of Statistical Software. [In press.] Rue, H., and Martino, S. 2007. Approximate Bayesian inference for hierarchical Gaussian Markov random field models. J. Stat. Plann. Infer. 137(10): 3177–3192. doi:10.1016/j.jspi.2006.07.016.

Lindgren, F., Rue, H., and Lindström, J. 2011. An explicit link between Gaussian fields and Gaussian Markov random fields: the stochastic partial differential equation approach. J. R. Stat. Soc. Ser. B Stat. Methodol. 73(4): 423–498. doi:10.1111/j.1467-9868.2011.00777.x.

Rue, H., and Martino, S. 2007. Approximate Bayesian inference for hierarchical Gaussian Markov random field models. J. Stat. Plann. Infer. 137(10): 3177–3192. doi:10.1016/j.jspi.2006.07.016.

Rue, H., Martino, S., and Chopin, N. 2009. Approximate Bayesian inference for latent Gaussian models by using integrated nested Laplace approximations. J. R. Stat. Soc. Ser. B Stat. Methodol. 71(2): 319–392. doi:10.1111/j.1467-9868. 2008.00700.x.

Warton, D, I., Blanchet, F. G., O'Hara, R. B., Ovaskainen, O., Taskinen, S., Walker, S. C., and Hui, F. K.C. 2015. So many variables: joint modeling in community ecology. Trends in Ecology and Evolution. 30 (12): 766-779.

Watanabe, S. 2013. A widely applicable Bayesian information criterion. Journal of Machine Learning Research 14: 867-897.

Yin, Y., Aeberhard, W. H., Smith, S. J., and Flemming, J. M. 2019. Identifiable state-space models: A case study of the Bay of Fundy sea scallop fishery. Canadian Journal of Statistics 47 (1): 27-45.

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.

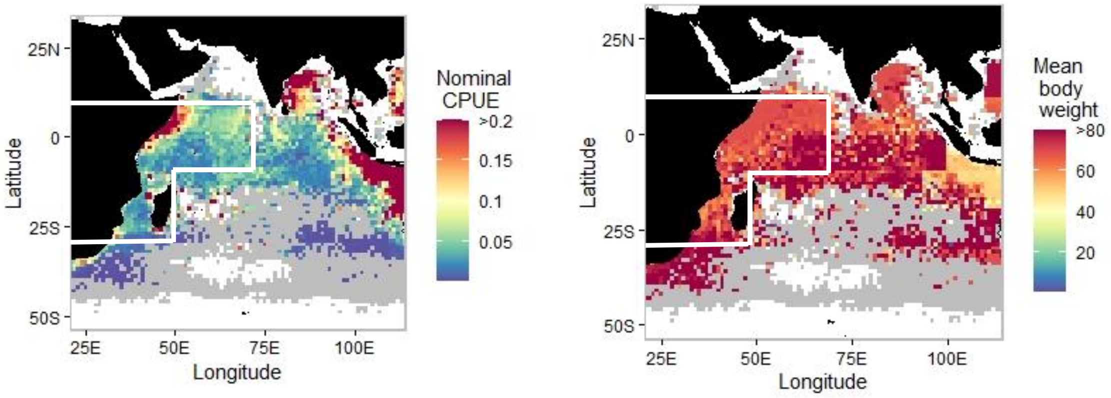

Figure 1. Spatial distributions of nominal CPUE and mean body weight for black marlin caught by Japanese longliners in the Indian Ocean. The area used for CPUE standardization (water inside of white line) in the present study.

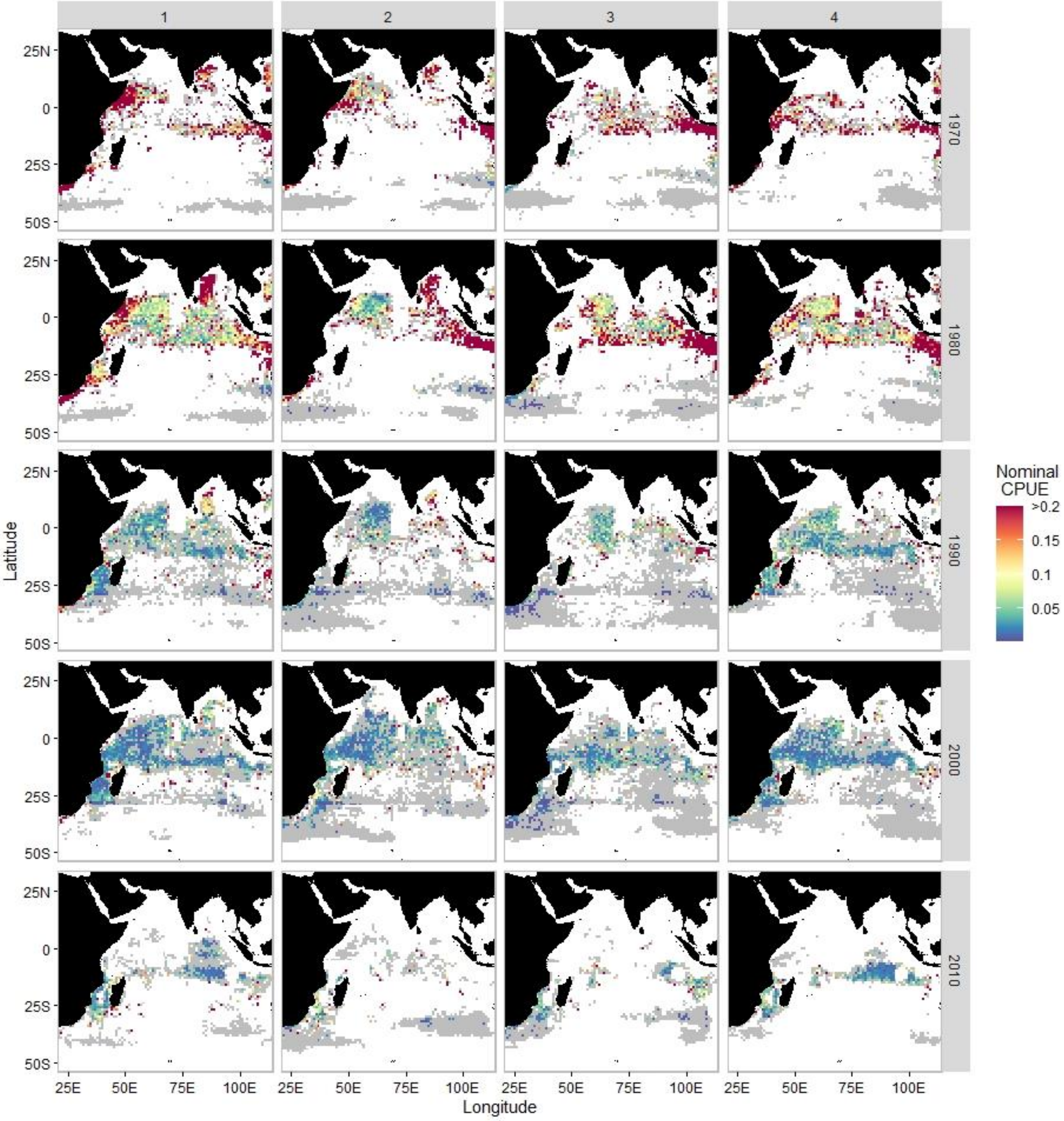

Figure 2. Spatial-temporal (seasonal and decadal) changes in the nominal CPUE for black marlin caught by Japanese longliners in the Indian Ocean.

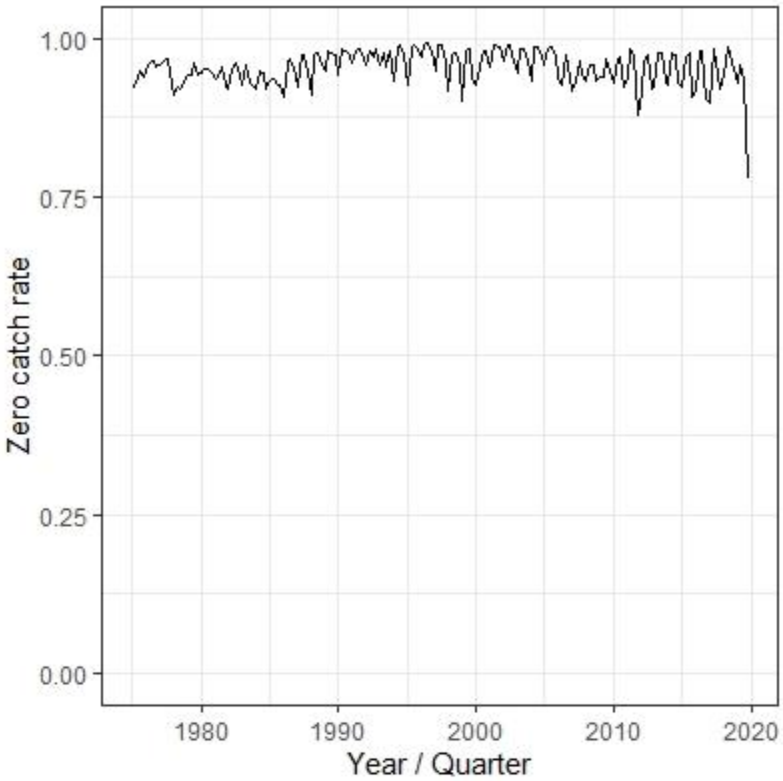

Figure 3. Annual changes in zero catch ratio of black marlin caught by Japanese longliners in the defined area of the Indian Ocean.

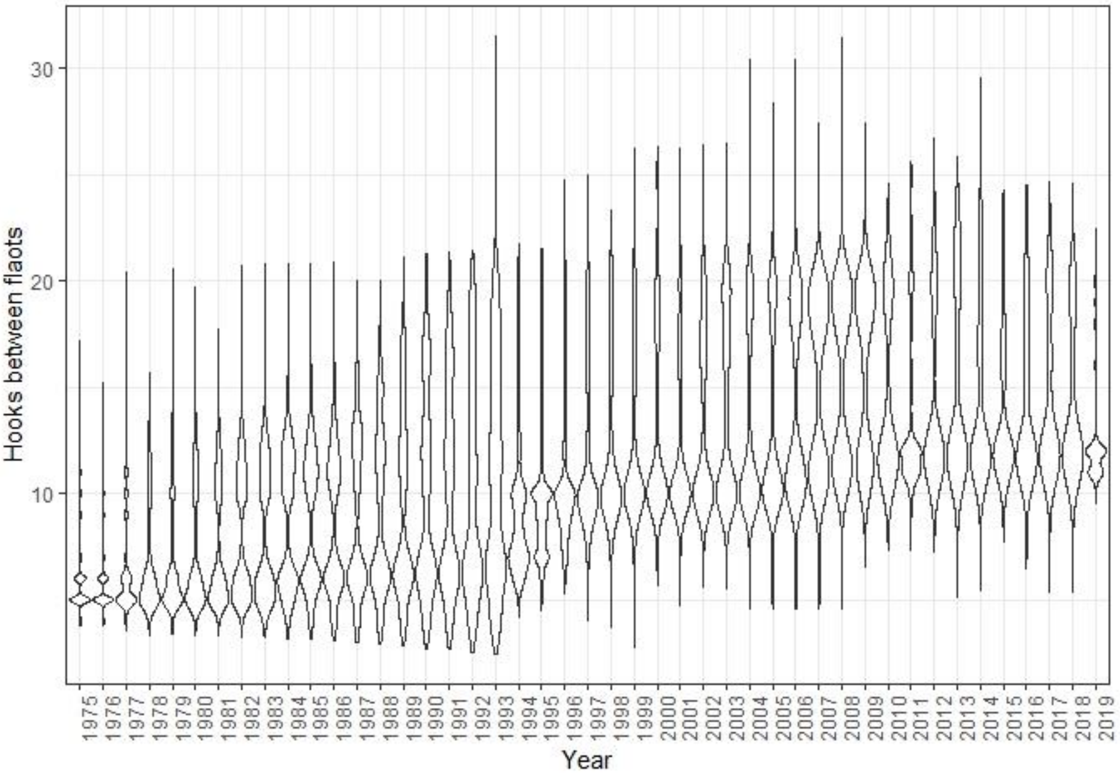

Figure 4. Historical changes in the gear configuration (number of hooks between floats) in the defined area of the Indian Ocean.

Table 1. Seven models and their WAIC values for two time periods of four areas. Selected models corresponded to those with the smallest values yellow-highlighted.

	1979-1993	1994-2019
m_null = inla (swo~1,data=d,offset=log(d$hooks/1000),family="poisson")	57409	78669
m_glm = inla (swo~yr + latlon,data=d,offset=log(d$hooks/1000),family="poisson")	$>10^{18}$	$>10^{18}$
m_glmm = inla (swo~yr + qtr + f(latlon,model="iid",hyper=hcprior) + f(jp_name,model="iid")+f(hpb,model="iid"),data=d,offset=log(d$hooks/1000),family="poisson")	51434	70963
m_zip_glmm = inla (swo~yr + qtr + f(latlon,model="iid") + f(jp_name,model="iid"), data=d,offset=log(d$hooks/1000),family="zeroinflatedpoisson1")	51070	68831
m_spde = inla (swo~0 + intercept + yr + qtr + f(hpb,model="iid") + f(jp_name,model="iid") + f(w,model=spde), data=inla.stack.data(StackFit),offset=log(d$hooks/1000),family="poisson")	51063	68418
m_spde2 = inla (swo~0 + intercept + f(yr,model="ar1") + f(month,model="iid",hyper=hcprior) + f(hpb,model="iid",hyper=hcprior) + f(jp_name,model="iid",hyper=hcprior) + f(w,model=spde), data=inla.stack.data(StackFit2),offset=log(d$hooks/1000),family="poisson")	50948	68414
m_zip_spde2 = inla (swo~0 + intercept + f(yr,model="ar1") + f(month,model="iid",hyper=hcprior) + f(hpb,model="iid") + f(jp_name,model="iid") + f(w,model=spde), data=inla.stack.data(StackFit2),offset=log(d$hooks/1000),family="zeroinflatedpoisson1")	50483	66646

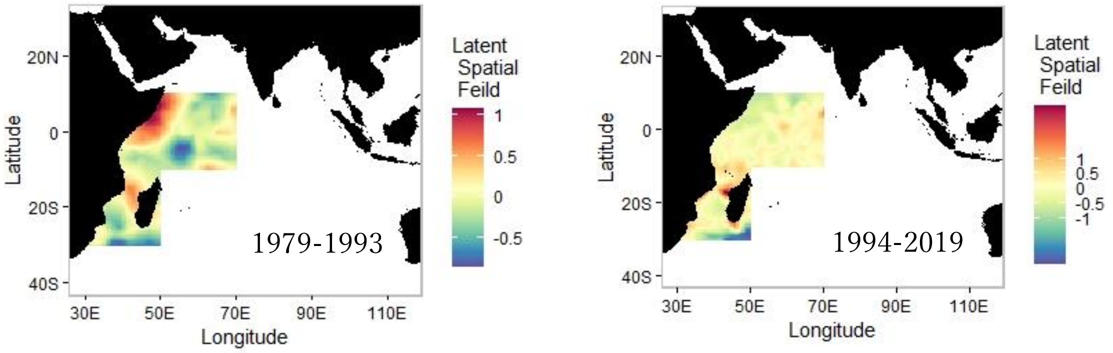

Figure 5. Spatial distribution in standardized CPUE (mean latent spatial field) of black marlin for two periods in the defined area of the Indian Ocean.

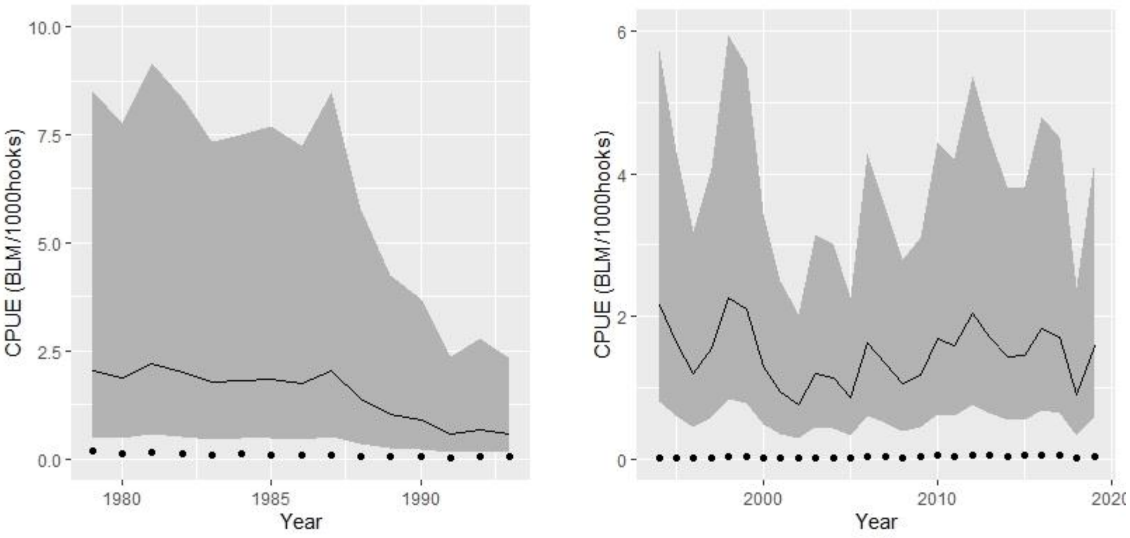

Figure 6. Historical changes in the CPUEs of black marlin for two periods in the defined area of the Indian Ocean. Thin line and filled points denote point estimates of standardized and nominal CPUE. Gray shadows denote 95% credible intervals. Note that the scale of y-axis is different between right and left figures.

Table 2. Nominal and standardized CPUEs of black marlin for two periods; 1979-93 and 1994-2019 in the defined area of the Indian Ocean.

year	nominal	Standardized	2.50%	97.50%	year	nominal	Standardized	2.50%	97.50%
1979	0.19	2.05	0.52	8.52	1994	0.02	2.16	0.81	5.72
1980	0.12	1.87	0.47	7.77	1995	0.02	1.63	0.61	4.31
1981	0.14	2.21	0.56	9.16	1996	0.01	1.21	0.45	3.18
1982	0.11	2.02	0.51	8.36	1997	0.02	1.54	0.58	4.07
1983	0.09	1.77	0.45	7.35	1998	0.03	2.26	0.85	5.94
1984	0.10	1.81	0.46	7.51	1999	0.03	2.09	0.79	5.52
1985	0.09	1.85	0.47	7.68	2000	0.02	1.30	0.49	3.44
1986	0.08	1.75	0.44	7.24	2001	0.01	0.94	0.35	2.49
1987	0.08	2.04	0.52	8.47	2002	0.01	0.77	0.29	2.02
1988	0.06	1.39	0.35	5.79	2003	0.02	1.19	0.45	3.14
1989	0.04	1.03	0.26	4.26	2004	0.02	1.14	0.43	3.01
1990	0.04	0.89	0.22	3.68	2005	0.01	0.86	0.32	2.25
1991	0.03	0.57	0.14	2.37	2006	0.03	1.63	0.61	4.29
1992	0.04	0.67	0.17	2.79	2007	0.03	1.34	0.50	3.52
1993	0.04	0.56	0.14	2.33	2008	0.02	1.06	0.40	2.79
					2009	0.02	1.18	0.44	3.11
					2010	0.05	1.68	0.63	4.44
					2011	0.04	1.59	0.60	4.20
					2012	0.04	2.04	0.77	5.38
					2013	0.04	1.71	0.64	4.53
					2014	0.04	1.44	0.54	3.80
					2015	0.04	1.45	0.54	3.82
					2016	0.06	1.82	0.68	4.80
					2017	0.05	1.70	0.64	4.50
					2018	0.02	0.90	0.34	2.39
					2019	0.03	1.58	0.59	4.23

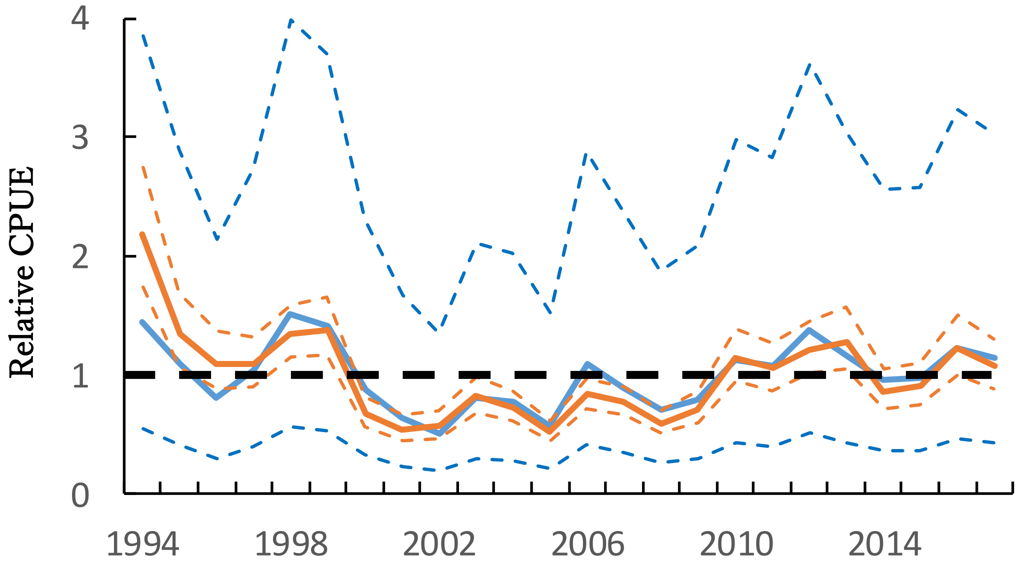

Figure 7. Comparison of annual standardized CPUE of black marlin (relative to its mean value for 1994-2017; horizontal broken line) between present (blue lines) and previous (Ijima 2018) studies (orange lines) for the defined area of the Indian Ocean. Solid and broken lines denote point estimates of standardized CPUE and its 95% credible intervals.