Japanese Longline CPUE Standardization (1979-2019) for striped marlin (*Tetrapturus audax*) 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 striped marlin stock abundance in the Indian Ocean, we standardized the CPUE of striped marlin caught by Japanese longliners for 1979-2019. We separated the logbook data into four areas (NW, NE, SW, SE) based on the IOTC area definition, 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 characterized by many zeros, we used zero-inflated Poisson generalized linear mixed model (ZIP-GLMM). All analyses were performed using R, specifically R-INLA package. The INLA procedure, in accordance with the Bayesian approach, calculated the marginal posterior distribution of all random effects and then estimated 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 each area in each period. Gradual annual decline trends with interannual variations were generally observed for all the standardized CPUEs. 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 a stock assessment of striped marlin (*Tetrapturus audax*) in the Indian Ocean. In the stock assessment, Ijima (2018) standardized the CPUE caught by Japanese longliners using a 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 striped marlin caught by Japanese longliners in the Indian Oceans 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 in the CPUE standardization for striped 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 also separated the Indian Ocean into four areas (NW, NE, SW, and SE) based on the IOTC area definition as Ijima (2018) (Figure 1). Japanese longliners have operated the four areas 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 NW Indian Ocean, and the target shift of fishermen to southern bluefin tuna in the Southern Indian Ocean (SW, and SE). The target shift makes it difficult to catch stripe marlins staying frequently in the shallower depths.

Statistical models

We applied Bayesian hierarchical spatial models in the present study, 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 degrees grid 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 four areas (Figure 4). Most variables were treated as 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 the 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 each area in each period.

3. Result and discussion

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

Northwest

The predicted CPUE was higher in the northwestern part in this area for both periods (Figure 5). The annual standardized CPUE showed a gradual decline trend in interannual variation for both periods (Figure 6, Table 2).

Northeast

The predicted CPUE was higher in the northwestern part in this area for both periods (Figure 7). The annual standardized CPUE showed a gradual decline trend for both periods (Figure 8, Table 3).

Southwest

The predicted CPUE was higher in the northern part in this area nearby Madagascar for 1979-1993 (Figure 9). The annual standardized CPUE showed a gradual decline trend for 1979-1993, while the no apparent trend was observed for 1994-2019 (Figure 10, Table 4). For the model of latter period, non-spatial model (m_zip_glmm) was selected as the best model (Table 1).

Southeast

The model could not provide reasonable outputs for both periods due to a low area coverage of catch data (Figure 2).

Figure 11 showed a comparison of annual changes in standardized CPUE between present and previous studies (Ijima, 2018) for three areas (NW, NE, and SW). The annual trends in point estimates are almost similar between them for each area. However, the 95% credible intervals for the present models were much wider than those for the previous models due to adding the spatial random effect to the present models except for SW.

We will improve the accuracy of stock assessment for stripe 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 future work.

4. References

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 striped marlin (*Tetrapturus audax*) caught by Japanese longline fishery: Update analysis between 1994 and 2017. IOTC-2018-WPB16-25.

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.

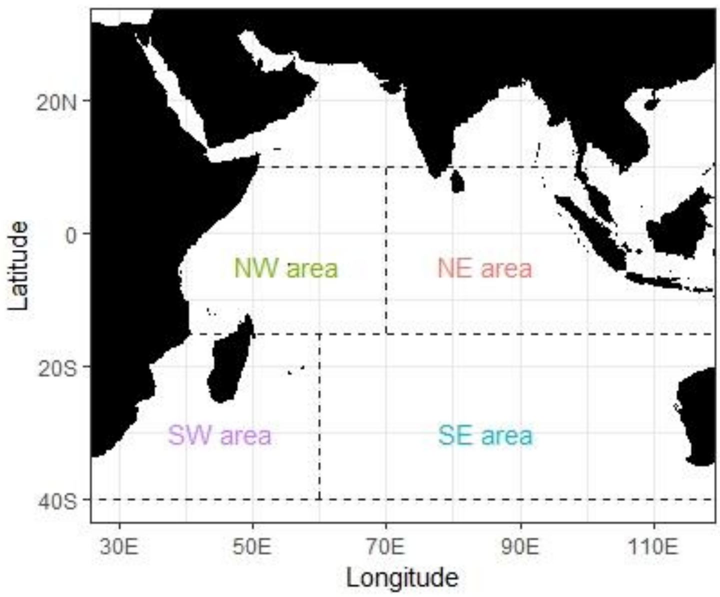

Figure 1. Four areas used in the analysis of CPUE standardization for the striped marlin in the Indian Ocean, where set in the 9th session of the IOTC working party on billfish (IOTC 2014).

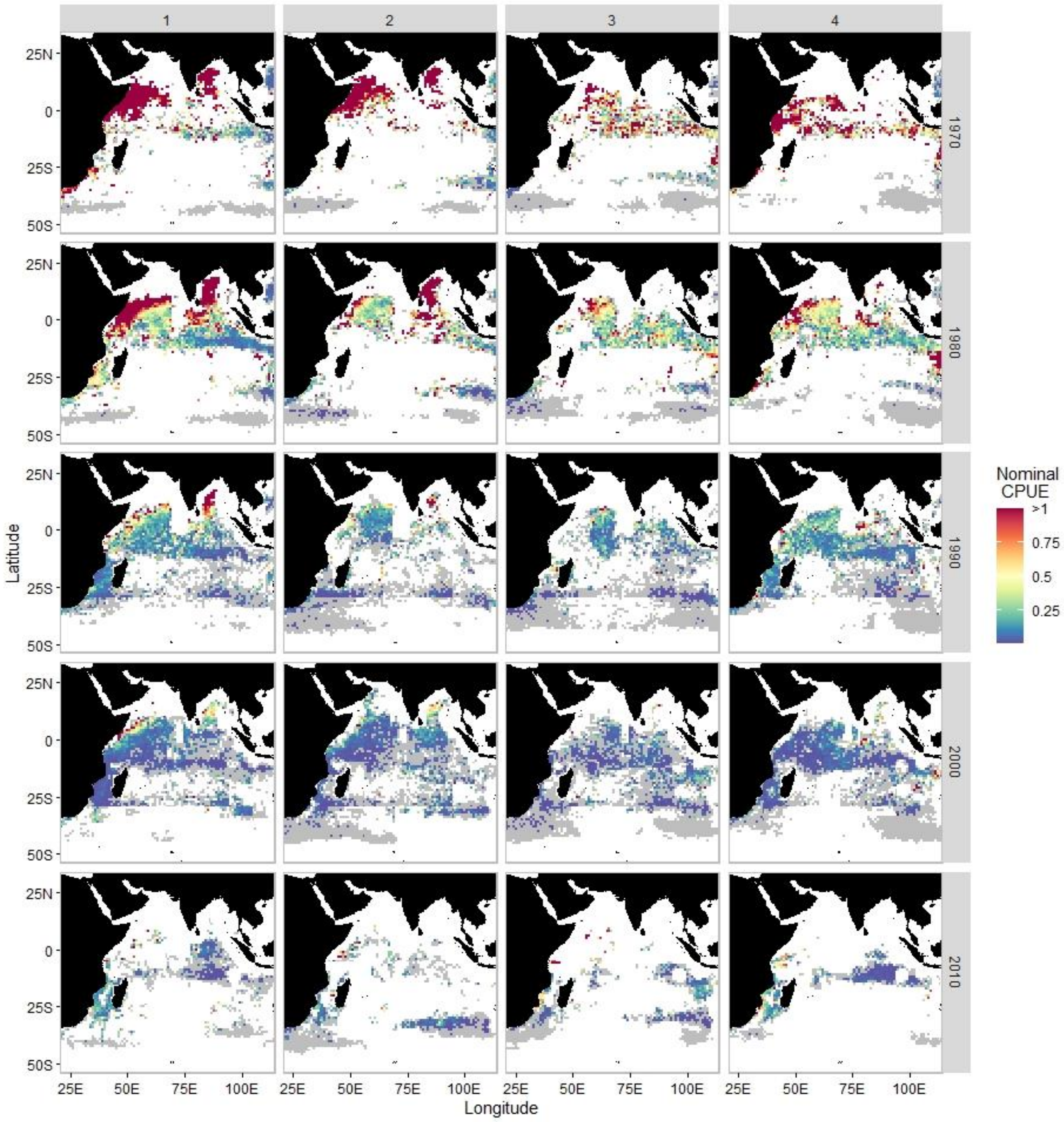

Figure 2. Spatial-temporal (seasonal and decadal) changes in the nominal CPUE for striped marlin caught by Japanese longliners in the Indian Ocean.

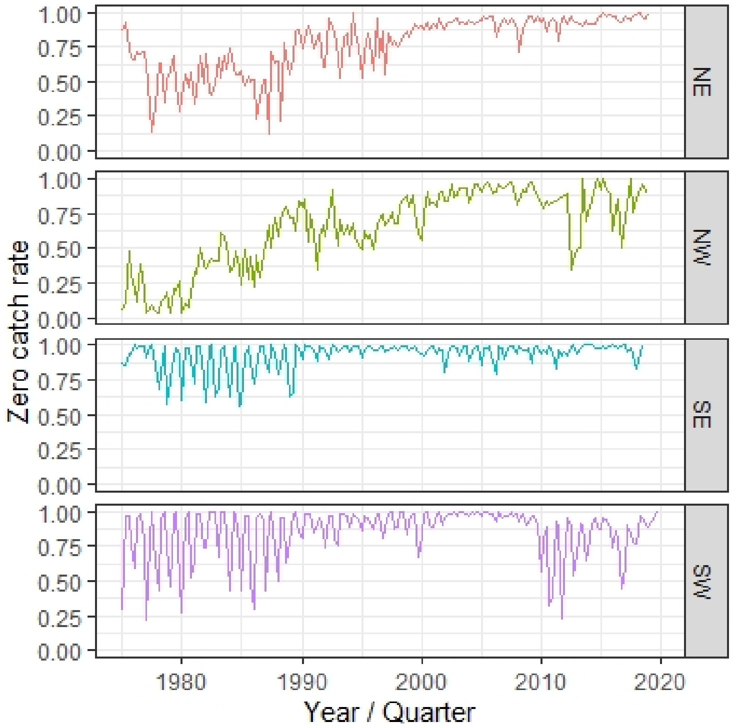

Figure 3. Annual changes in zero catch ratio of striped marlin caught by Japanese longliners in four areas of the Indian Ocean.

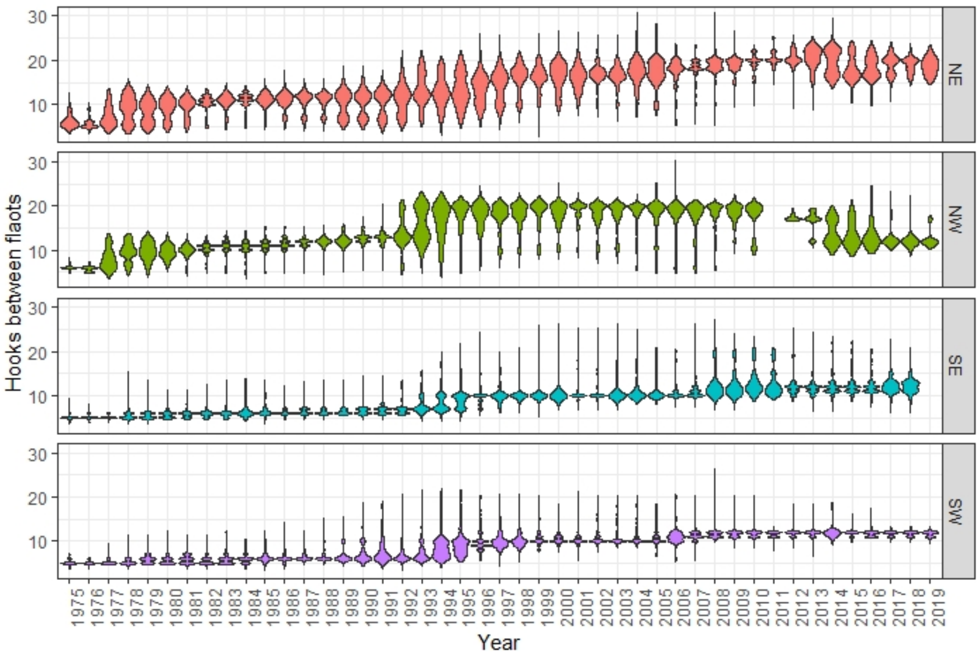

Figure 4. Historical changes in the gear configuration (number of hooks between floats) in four areas of the Indian Ocean.

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.

	NW(79-93)	NW(94-19)	NE(79-93)	NE(94-19)	SW(79-93)	SW(94-19)	SE(79-93)	SE(94-19)
m_null = inla (swo~1,data=d,offset=log(d$hooks/1000),family="poisson")	201027	115320	130501	76137	75465	145975	80367	41239
m_glm = inla (swo~yr + latlon,data=d,offset=log(d$hooks/1000),family="poisson")	$>10^{18}$	84879	$>10^{18}$	$>10^{18}$	$>10^{18}$	$>10^{15}$	$>10^{16}$	$>10^{18}$
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$hooks/1000),family="poisson")	120125	79532	55169	42319	27958	75272	15995	29495
m_zip_glmm = inla (swo~yr + qtr + f(latlon,model="iid") + f(jp_name,model="iid"), data=d,offset=log(d$hooks/1000),family="zeroinflatedpoisson1")	118005	76220	53744	41253	31514	69069	22312	24705
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$hooks/1000),family="poisson")	117431	78207	52839	41571	26490	71988	15372	28672
m_spde2 = inla (swo~0 + intercept + f(yr,model="ar1") + f(month,model="iid",hyper=hcprior) + f(hpb,model="iid",hyper=hcprior) + f(jp_name,model="iid",hyper=hcprior) + f(w,model=spde), data=inla.stack.data(StackFit2),offset=log(d$hooks/1000),family="poisson")	116423	77583	52300	41411	25790	70143	15015	28034
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")	114551	74679	51143	40361	—	—	—	—

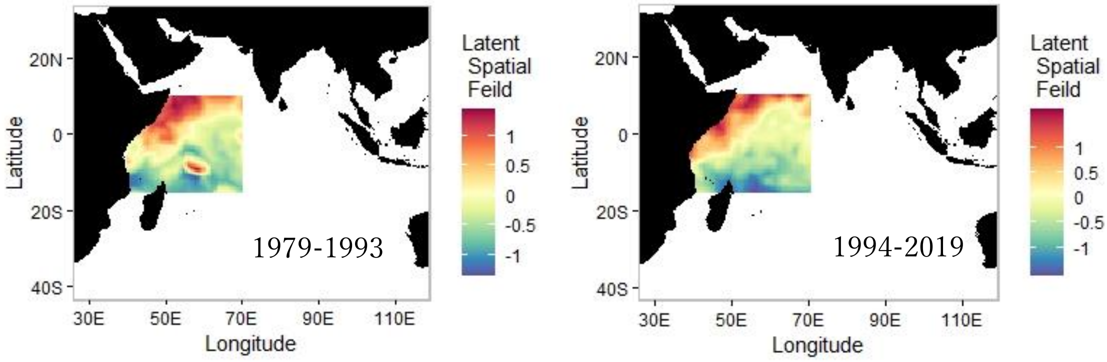

Figure 5. Spatial distribution in standardized CPUE (mean latent spatial field) of striped marlin for two periods in the **Northwest** area in the Indian Ocean.

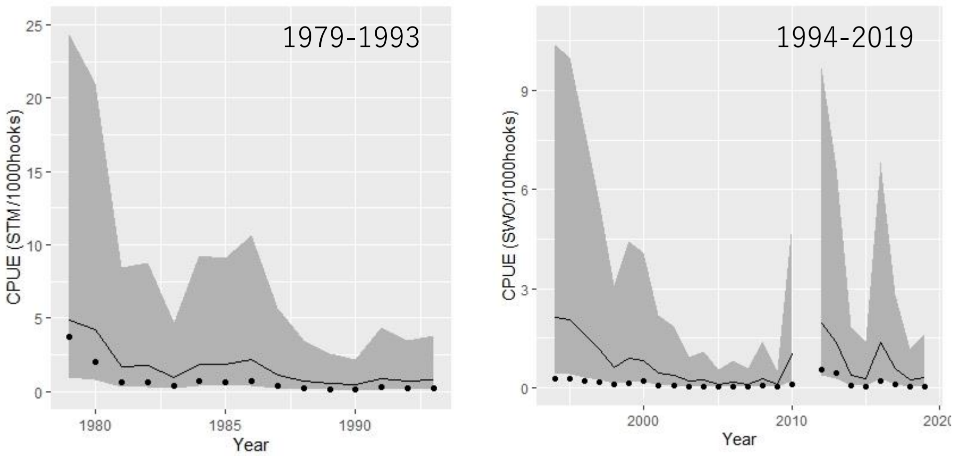

Figure 6. Historical changes in the standardized CPUEs of striped marlin for two periods in the **Northwest** area in the Indian Ocean. Thin line and filled point denote point estimates of standardized and nominal CPUEs, respectively. Grey shadow denotes 95% credible interval. Note that the scale of y-axis is different between right and left figures.

Table 2. Nominal and standardized CPUEs of striped marlin for two periods of 1979-93 and 1994-2019 in the **Northwest** area in the Indian Ocean.

year	nominal	Standardized	2.50%	97.50%	year	nominal	Standardized	2.50%	97.50%
1979	3.68	4.90	0.93	24.34	1994	0.29	2.13	0.42	10.38
1980	1.97	4.23	0.80	21.03	1995	0.28	2.05	0.41	10.01
1981	0.66	1.70	0.32	8.47	1996	0.20	1.61	0.32	7.85
1982	0.61	1.76	0.33	8.77	1997	0.16	1.15	0.23	5.62
1983	0.40	0.95	0.18	4.71	1998	0.11	0.63	0.13	3.09
1984	0.69	1.86	0.35	9.27	1999	0.15	0.91	0.18	4.42
1985	0.62	1.84	0.35	9.12	2000	0.20	0.84	0.17	4.09
1986	0.72	2.15	0.41	10.66	2001	0.05	0.45	0.09	2.21
1987	0.39	1.15	0.22	5.69	2002	0.06	0.38	0.08	1.86
1988	0.20	0.70	0.13	3.50	2003	0.03	0.19	0.04	0.94
1989	0.16	0.52	0.10	2.57	2004	0.04	0.23	0.05	1.11
1990	0.12	0.43	0.08	2.16	2005	0.02	0.11	0.02	0.54
1991	0.33	0.87	0.17	4.35	2006	0.03	0.17	0.03	0.83
1992	0.19	0.70	0.13	3.47	2007	0.02	0.12	0.02	0.57
1993	0.20	0.77	0.15	3.83	2008	0.06	0.29	0.06	1.41
					2009	0.02	0.10	0.02	0.51
					2010	0.11	1.03	0.20	5.02
					2011	NA	NA	NA	NA
					2012	0.56	1.97	0.39	9.68
					2013	0.43	1.36	0.27	6.68
					2014	0.07	0.38	0.07	1.86
					2015	0.03	0.28	0.05	1.40
					2016	0.20	1.39	0.27	6.85
					2017	0.09	0.57	0.11	2.82
					2018	0.04	0.23	0.04	1.16
					2019	0.03	0.31	0.06	1.66

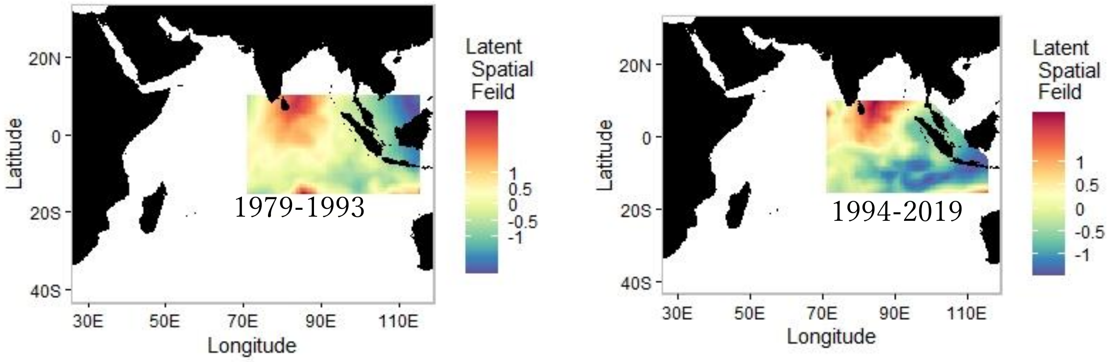

Figure 7. Spatial distribution in standardized CPUE (mean latent spatial field) of striped marlin for two periods in the **Northeast** area of the Indian Ocean.

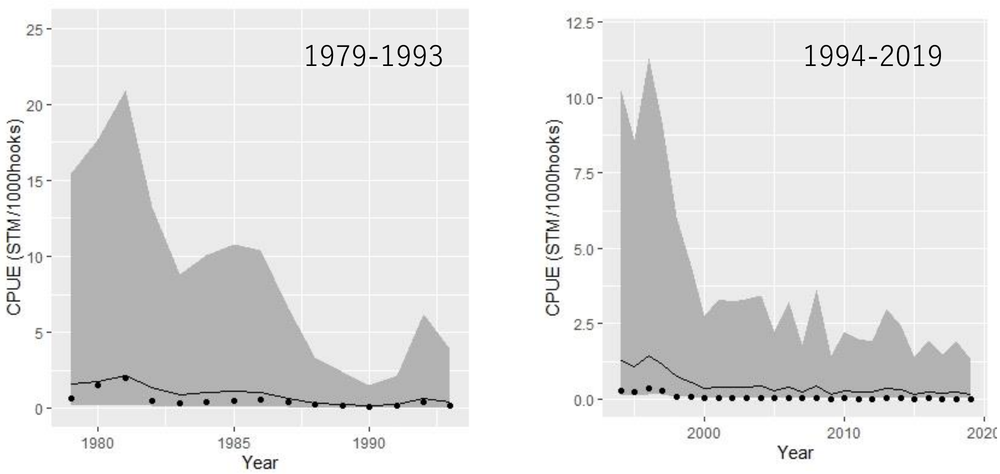

Figure 8. Historical changes of in the standardized CPUEs of striped marlin for two periods in the **Northeast** area of the Indian Ocean. Thin line and filled point denote point estimates of predicted CPUE and nominal CPUE, respectively. Gray shadow denotes 95% credible interval. Note the scale of y-axis is different between right and left figures.

Table 3. Nominal and standardized CPUEs of striped marlin for two periods of 1979-93 and 1994-2019 in the **Northeast** area of the Indian Ocean

year	nominal	Standardized	2.50%	97.50%	year	nominal	Standardized	2.50%	97.50%
1979	0.66	1.57	0.16	15.45	1994	0.29	1.28	0.16	10.24
1980	1.51	1.80	0.18	17.65	1995	0.25	1.08	0.13	8.60
1981	1.99	2.13	0.21	20.91	1996	0.36	1.42	0.18	11.33
1982	0.49	1.35	0.14	13.27	1997	0.26	1.15	0.14	9.18
1983	0.35	0.90	0.09	8.84	1998	0.10	0.75	0.09	6.02
1984	0.37	1.02	0.10	10.04	1999	0.06	0.56	0.07	4.43
1985	0.53	1.10	0.11	10.81	2000	0.05	0.35	0.04	2.77
1986	0.55	1.06	0.11	10.40	2001	0.05	0.41	0.05	3.30
1987	0.44	0.67	0.07	6.57	2002	0.04	0.41	0.05	3.24
1988	0.22	0.34	0.03	3.38	2003	0.03	0.41	0.05	3.30
1989	0.21	0.25	0.02	2.41	2004	0.03	0.43	0.05	3.42
1990	0.12	0.15	0.02	1.49	2005	0.03	0.28	0.03	2.23
1991	0.17	0.22	0.02	2.15	2006	0.03	0.41	0.05	3.24
1992	0.42	0.63	0.06	6.18	2007	0.02	0.23	0.03	1.80
1993	0.18	0.40	0.04	3.92	2008	0.05	0.46	0.06	3.63
					2009	0.01	0.18	0.02	1.42
					2010	0.03	0.28	0.03	2.24
					2011	0.02	0.25	0.03	1.99
					2012	0.01	0.24	0.03	1.90
					2013	0.02	0.37	0.05	2.98
					2014	0.02	0.31	0.04	2.44
					2015	0.01	0.17	0.02	1.40
					2016	0.02	0.25	0.03	1.97
					2017	0.01	0.18	0.02	1.47
					2018	0.01	0.24	0.03	1.91
					2019	0.00	0.16	0.02	1.32

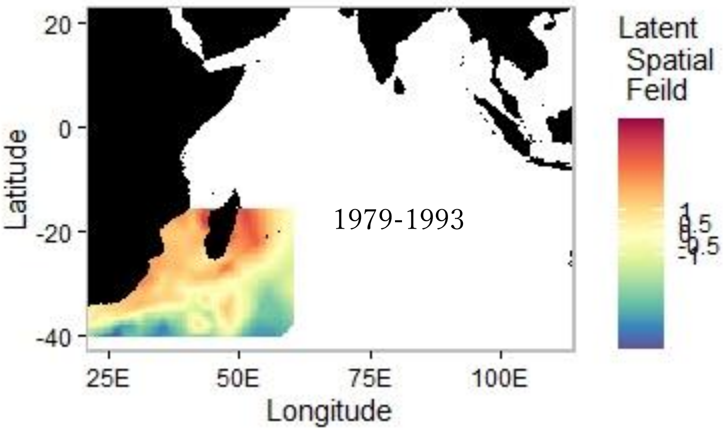

Figure 9. Spatial distribution in standardized CPUE (mean latent spatial field) of striped marlin for two periods in the **Southwest** area of the Indian Ocean.

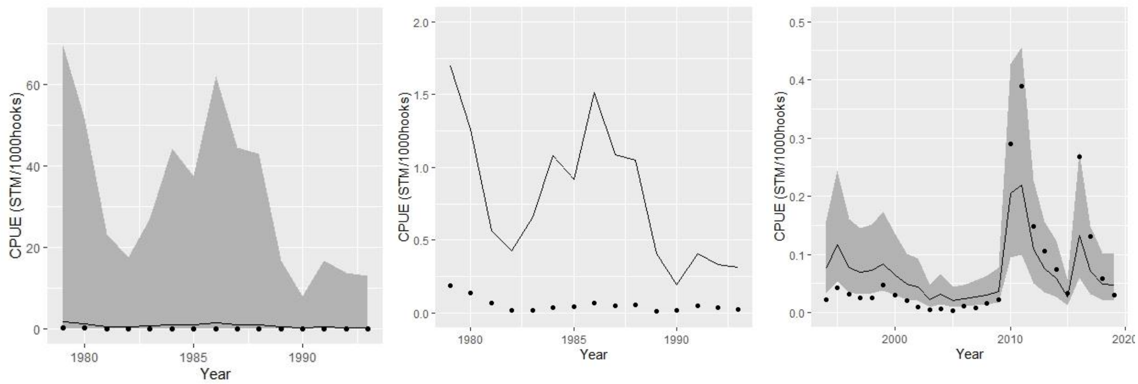

Figure 10. Historical changes in the standardized CPUEs of striped marlin for two periods in the **Southwest** area of the Indian Ocean. Thin line and filled point denote point estimates of predicted and nominal CPUE, respectively. Gray shadow denotes 95% credible interval. Middle figure indicates the point estimates of standardized CPUEs with scale down for y-axis of the left figure during 1979-1993. Also note that the scale of y-axis of right figure is different from the left figure.

Table 4. Nominal and standardized CPUEs of striped marlin for two periods of 1979-93 and 1994-2019 in the **Southwest** area in the Indian Ocean.

year	nominal	Standardized	2.50%	97.50%	year	nominal	Standardized	2.50%	97.50%
1979	0.19	1.70	0.04	69.50	1994	0.02	0.08	0.03	0.16
1980	0.14	1.26	0.03	51.61	1995	0.04	0.12	0.05	0.24
1981	0.07	0.57	0.01	23.24	1996	0.03	0.08	0.04	0.16
1982	0.02	0.43	0.01	17.46	1997	0.03	0.07	0.03	0.15
1983	0.02	0.66	0.02	27.10	1998	0.03	0.07	0.03	0.15
1984	0.04	1.08	0.02	44.27	1999	0.05	0.08	0.04	0.17
1985	0.04	0.92	0.02	37.52	2000	0.03	0.07	0.03	0.14
1986	0.07	1.52	0.03	61.94	2001	0.02	0.05	0.02	0.10
1987	0.05	1.09	0.02	44.55	2002	0.01	0.04	0.02	0.09
1988	0.06	1.05	0.02	42.92	2003	0.00	0.02	0.01	0.05
1989	0.02	0.41	0.01	16.68	2004	0.01	0.03	0.01	0.07
1990	0.02	0.19	0.00	7.92	2005	0.00	0.02	0.01	0.04
1991	0.05	0.41	0.01	16.61	2006	0.01	0.02	0.01	0.05
1992	0.04	0.33	0.01	13.66	2007	0.01	0.03	0.01	0.06
1993	0.03	0.32	0.01	12.97	2008	0.02	0.03	0.01	0.06
					2009	0.02	0.04	0.02	0.08
					2010	0.29	0.21	0.09	0.43
					2011	0.39	0.22	0.10	0.46
					2012	0.15	0.11	0.05	0.23
					2013	0.11	0.08	0.03	0.16
					2014	0.07	0.06	0.03	0.12
					2015	0.03	0.03	0.01	0.06
					2016	0.27	0.13	0.06	0.28
					2017	0.13	0.07	0.03	0.15
					2018	0.06	0.05	0.02	0.10
					2019	0.03	0.05	0.02	0.10

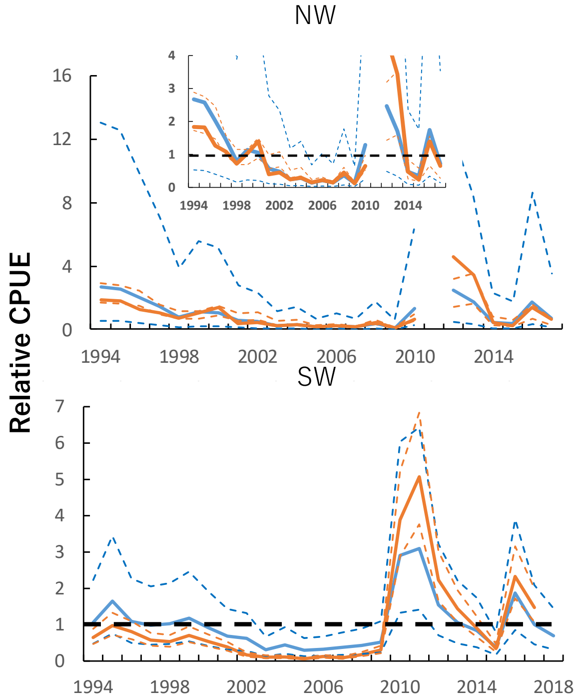

Figure 16. Comparison of annual standardized CPUE of striped marlin (relative to its mean value for 1994 – 2017; horizontal broken line) between present (blue lines) and previous (Ijima 2018) studies (orange lines) for three areas in the Indian Ocean. Solid and broken lines denote point estimates of standardized CPUE and its 95% credible intervals. Insets in NW and NE denote figures of scale down for y-axis.