IOTC-2024-WPB22-18

Japanese Longline CPUE Standardization (1979-2022) for striped marlin (*Tetrapturus audax*) in the Indian Ocean using Bayesian hierarchical spatial model

Takayuki Matsumoto, Kenji Taki, Hirotaka Ijima and Mikihiko Kai

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

Summary

The CPUE of striped marlin caught by Japanese longliners during 1979-2022 was standardized. Area definition is the same as that in the previous studies. Time-period was divided into two, 1979-1993 and 1994-2022. Bayesian hierarchical spatial models were applied. Considering high zero catch ratio, zero-inflated Poisson generalized linear mixed model (ZIP-GLMM) was used with the R-INLA package. Best model was selected from multiple models mainly using Widely Applicable Bayesian Information Criterion (WAIC). Gradual annual decline trend with interannual variation were generally observed for the standardized CPUEs. The trends of CPUEs were similar to those for the previous study.

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) and Taki et al. (2021) standardized 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 Japanese longliners are very critical for the stock assessment.

Integrated nested Laplace approximations (INLA) methodology and its powerful application to the modelling of complex datasets have recently been introduced to a 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 are 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 and Rue 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 allow scientists to fit considerably faster and more reliably complex spatiotemporal models (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 2022 for the stock assessment of this species using the same method as that in the previous study (Taki et al. 2021). 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 1994-2022, as the gear configuration of Japanese longline fishery such as number of hooks between floats and gear material had drastically changed in the early 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), as with the previous study (Taki et al., 2021). Japanese longliners operated in the four areas from the 1990s to the 2000s, but in the 2010s, the fishing ground was shrunk rapidly (Figure 2). There are two main reasons for that the influence of piracy activities 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 ID (vessel name; jp_name), and gear configuration (number of hooks between floats; hpb). The hpb increased remarkably in the early 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 a latent spatial field to indicate the expected CPUE distribution. Best candidate model was selected based on WAIC and if the results are reasonable (i.e. credible interval for CPUE is not too broad).

3. Result and discussion

We compared the WAIC among eight different structure’s models for each area and period (Table 1). The models selected are highlighted by yellow marker. Basically, the same models as those in the previous study (Taki et al., 2021) have been selected.

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 near Madagascar for 1979-1993 (Figure 9). The annual standardized CPUE showed a declining trend with fluctuation for 1979-1993, while the no apparent trend was observed for 1994-2022 (Figure 10, Table 4). For the model of the 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), as with the previous study (Taki et al., 2021).

Figure 11 shows a comparison of annual changes in standardized CPUE between present and previous (Taki et al., 2021) studies for three areas (NW, NE, and SW). The annual trends in point estimates are similar between them for each area, although some difference is observed for SW area. The trend of CPUE differed among areas, although some similarity was observed between NW and SW areas.

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.

Taki, K., Ijima, H. and Kai, M. 2021. Japanese Longline CPUE Standardization (1979-2019) for striped marlin (*Tetrapturus audax*) in the Indian Ocean using Bayesian hierarchical spatial model. IOTC-2021-WPB19-25.

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

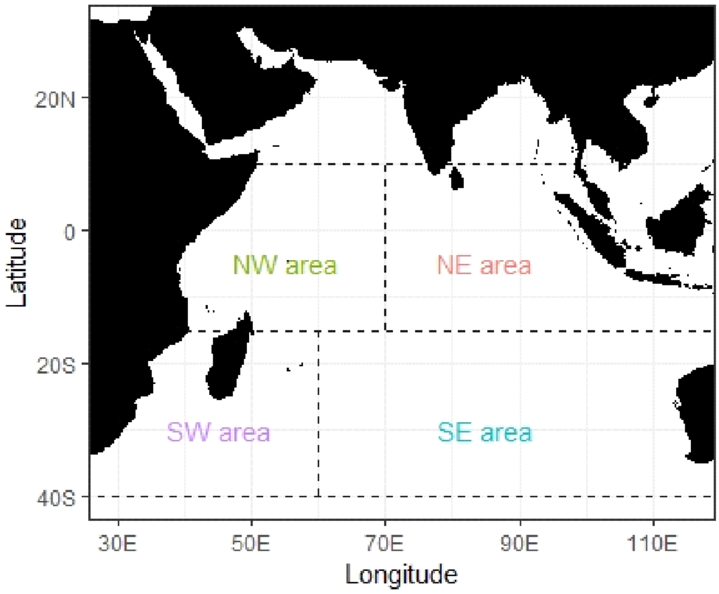


Figure 1. Four areas used in the analysis of CPUE standardization for the striped marlin in the Indian Ocean, which were set in the 9th session of the IOTC working party on billfish (IOTC 2014).

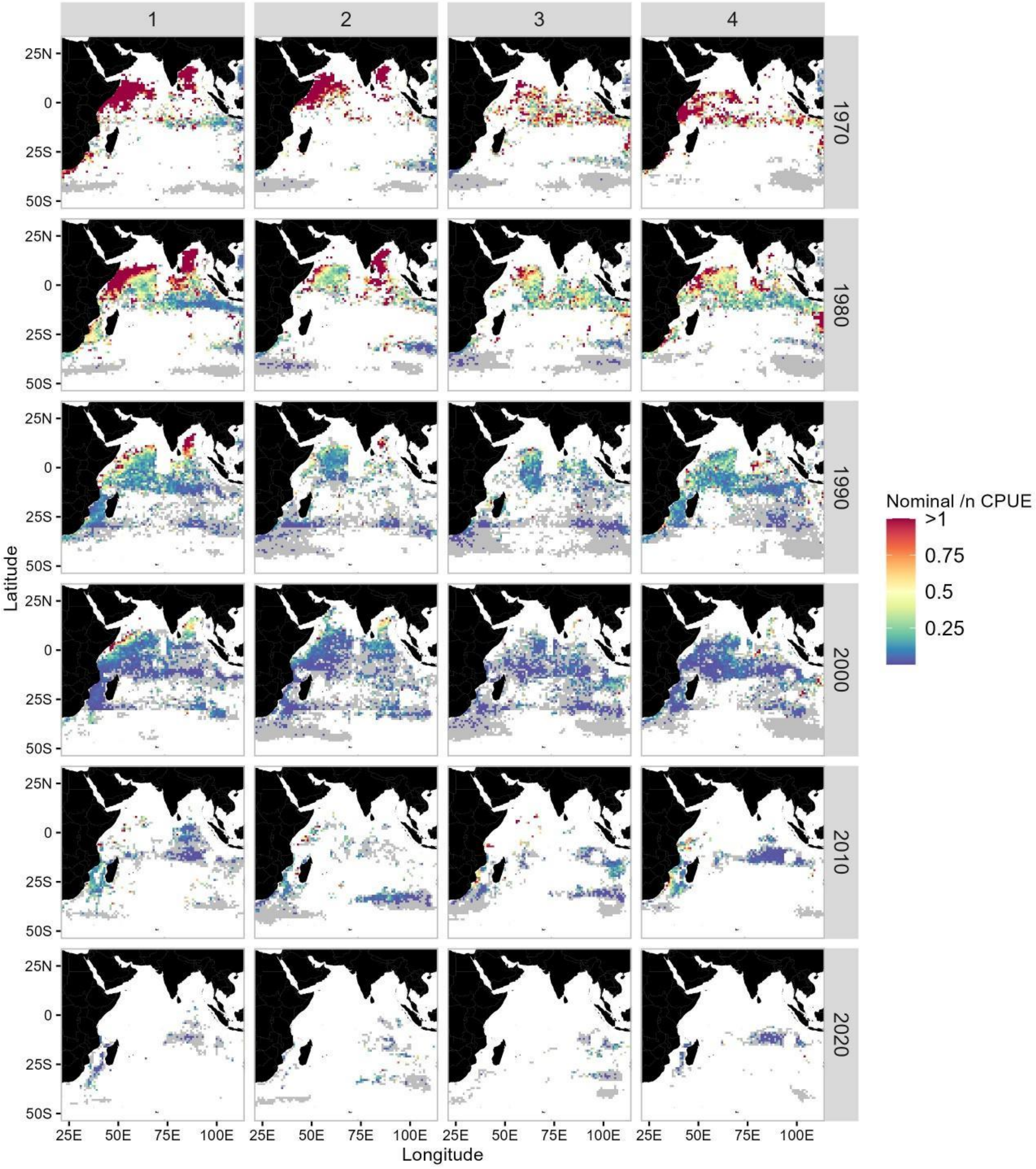


Figure 2. Spatial-temporal (seasonal and decadal) changes in the nominal CPUE for striped marlin caught by Japanese longliners in the Indian Ocean. 1: Jan-Mar, 2: Apr-Jun, 3: Jul-Sep, 4: Oct-Dec.

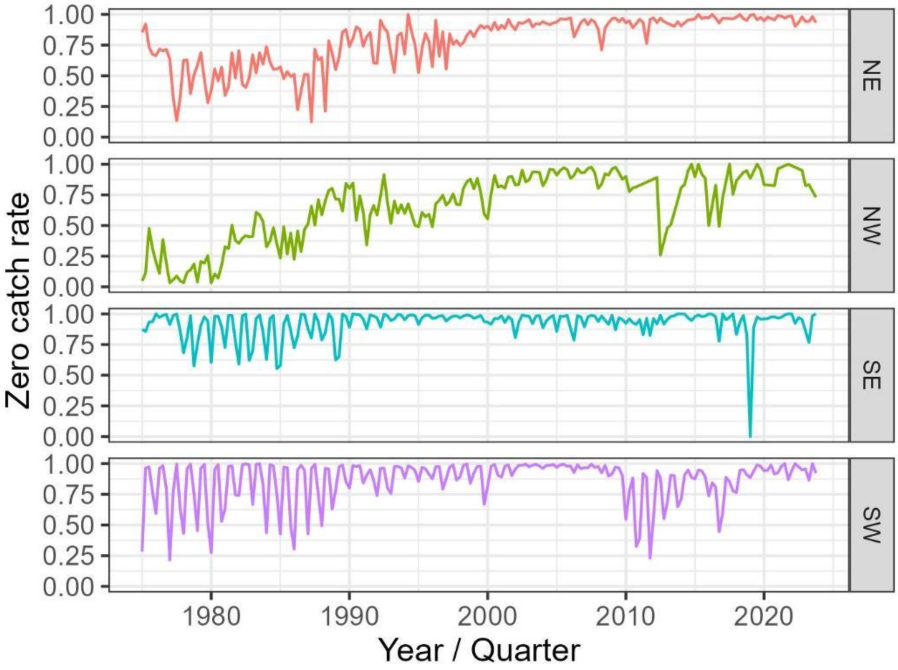


Figure 3. Annual changes in zero catch ratio of striped marlin caught by Japanese longliners in four areas of the Indian Ocean.

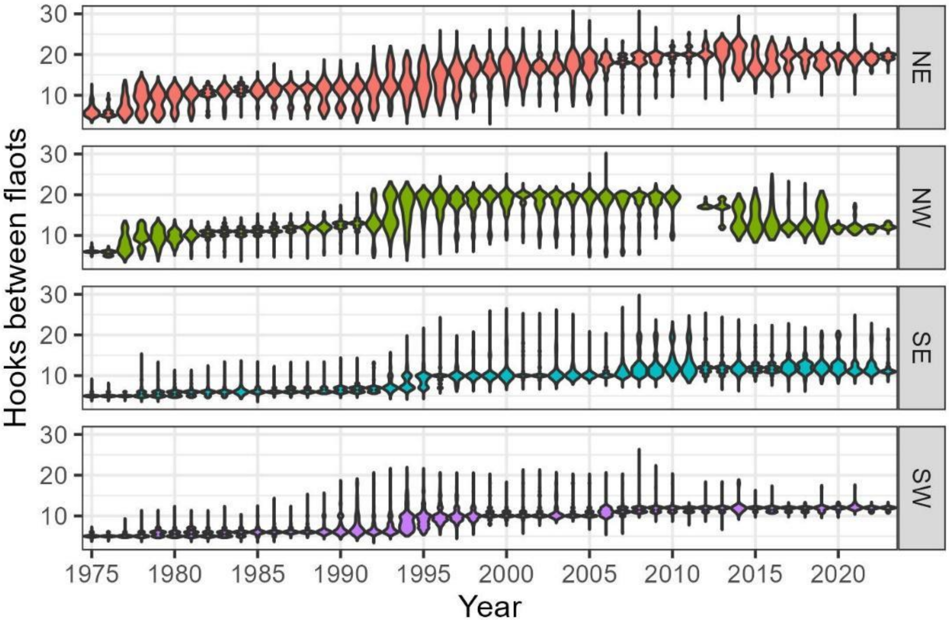


Figure 4. Historical changes in the gear configuration (number of hooks between floats) in four areas of the Indian Ocean. Vertical range of the plots shows the range of the data, and width shows frequency of the data.

Table 1. The models examined for the analyses. Selected models are yellow-highlighted. Note: the models with smallest WAIC were not always selected because the models with unreasonable results were eliminated.

Model*	NW(1979-1993)	NW(1994-2022)	NE(1979-1993)	NE(1994-2022)	SW(1979-1993)	SW(1994-2022)
m_null = inla (stm~1, data=d,offset=log(hooks/1000), family="poisson")	201027	107706	130501	67632	75465	119913
m_glm = inla (stm~yr + qtr + latlon, data=d,offset=log(hooks/1000), family="poisson")	$>10^{18}$	97609	$>10^{18}$	74014	$>10^{18}$	89482
m_glmm = inla (stm~yr + qtr + f(latlon,model="iid", hyper=hcprior) + f(jp_name,model="iid")+f(hpb,model="iid"), data=d,offset=log(hooks/1000), family="poisson")	120125	93199	55169	44672	27958	138939
m_zip_glmm = inla (stm~yr + qtr + f(latlon,model="iid") + f(jp_name,model="iid"), data=d, offset=log(hooks/1000), family="zeroinflatedpoisson1")	118005	88045	53744	42797	31514	113672
m_spde = inla (stm~0 + intercept + yr + qtr + f(hpb,model="iid") + f(jp_name,model="iid") + f(w,model=spde), data=inla.stack.data(StackFit), offset=log(hooks/1000), family="poisson")	117431	91993	52839	44436	26490	107455
m_spde2 = inla (stm~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(hooks/1000),family="poisson")	116423	85376	52300	44829	25790	100041
m_zip_spde = inla (stm~0 + intercept + yr + qtr + f(hpb,model="iid") +f(jp_name,model="iid") + f(w,model=spde), data=inla.stack.data(StackFit), offset=log(hooks/1000), family="zeroinflatedpoisson1")	115434	81534	51533	42057	-	94528
m_zip_spde2 = inla (stm~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(hooks/1000),family="zeroinflatedpoisson1")	114551	80501	51143	42350	-	87719

* stm: catch of striped marlin in number, hooks: number of hooks, yr: year, qtr: quarter, latlon: 5 x 5 degree latitude and longitude, hpb: number of hooks between floats, jp_name: vessel ID (vessel name), iid: Gaussian random effects, ar1: auto-regressive model of order 1, spde: stochastic partial differential equations, hyper: hyperparameters, hcprior: halfcauchy prior, family: likelihood family, d: catch and effort data set used in the program code. StackFit, StackFit2: stacked data for INLA.

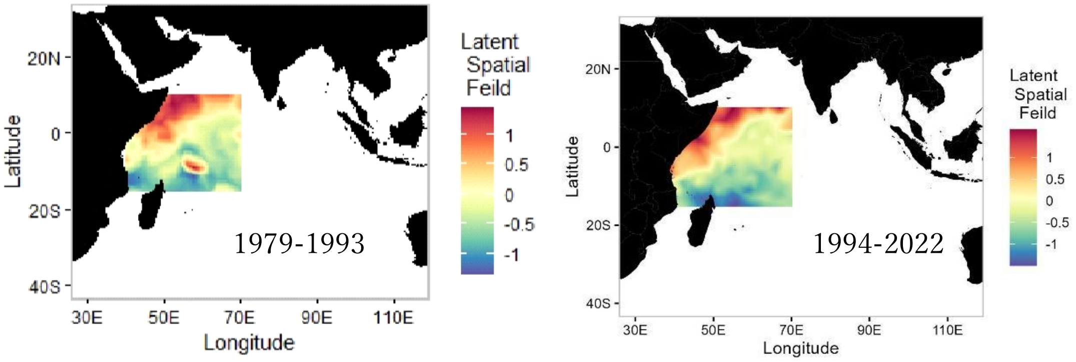


Figure 5. Spatial distribution in standardized CPUE (mean latent spatial field) of striped marlin for two periods in the **Northwest** area in the Indian Ocean**.**

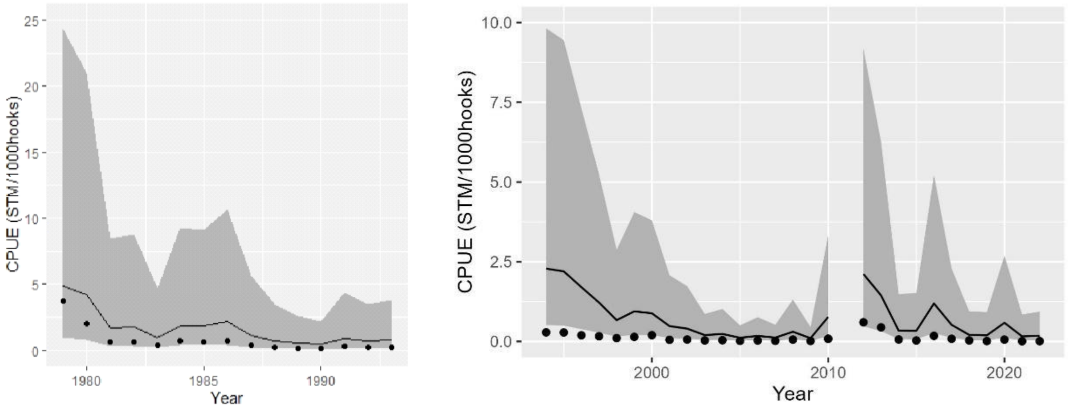


Figure 6. Historical changes in the standardized CPUEs of striped marlin for two periods in the **Northwest** area in the Indian Ocean. Thin line and filled point denote point estimates of standardized and nominal CPUEs, respectively. Grey shadow denotes 95% credible interval. Note that the scale of 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 1994-2022 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	0.288	2.285	9.820	0.519
1980	1.97	4.23	0.80	21.03	1995	0.283	2.199	9.448	0.499
1981	0.66	1.70	0.32	8.47	1996	0.196	1.704	7.310	0.387
1982	0.61	1.76	0.33	8.77	1997	0.166	1.222	5.245	0.278
1983	0.40	0.95	0.18	4.71	1998	0.106	0.669	2.875	0.152
1984	0.69	1.86	0.35	9.27	1999	0.145	0.945	4.060	0.215
1985	0.62	1.84	0.35	9.12	2000	0.198	0.885	3.800	0.201
1986	0.72	2.15	0.41	10.66	2001	0.053	0.483	2.076	0.109
1987	0.39	1.15	0.22	5.69	2002	0.063	0.404	1.738	0.092
1988	0.20	0.70	0.13	3.50	2003	0.034	0.201	0.862	0.046
1989	0.16	0.52	0.10	2.57	2004	0.040	0.237	1.020	0.054
1990	0.12	0.43	0.08	2.16	2005	0.020	0.117	0.505	0.027
1991	0.33	0.87	0.17	4.35	2006	0.031	0.177	0.761	0.040
1992	0.19	0.70	0.13	3.47	2007	0.020	0.122	0.524	0.028
1993	0.20	0.77	0.15	3.83	2008	0.061	0.306	1.316	0.069
					2009	0.016	0.105	0.452	0.024
					2010	0.086	0.769	3.330	0.173
					2011	NA	NA	NA	NA
					2010	0.086	0.769	3.330	0.173
					2012	0.602	2.116	9.184	0.476
					2013	0.444	1.444	6.258	0.325
					2014	0.064	0.339	1.485	0.076
					2015	0.032	0.335	1.520	0.073
					2016	0.177	1.193	5.226	0.264
					2017	0.087	0.523	2.291	0.116
					2018	0.034	0.208	0.943	0.045
					2019	0.014	0.194	0.922	0.040
					2020	0.060	0.587	2.692	0.122
					2021	0.008	0.161	0.849	0.030
					2022	0.011	0.181	0.936	0.033

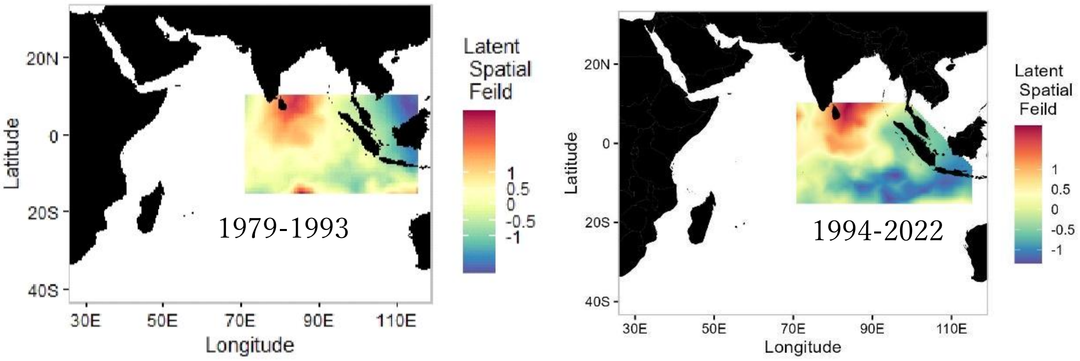


Figure 7. Spatial distribution in standardized CPUE (mean latent spatial field) of striped marlin for two periods in the **Northeast** area of the Indian Ocean.

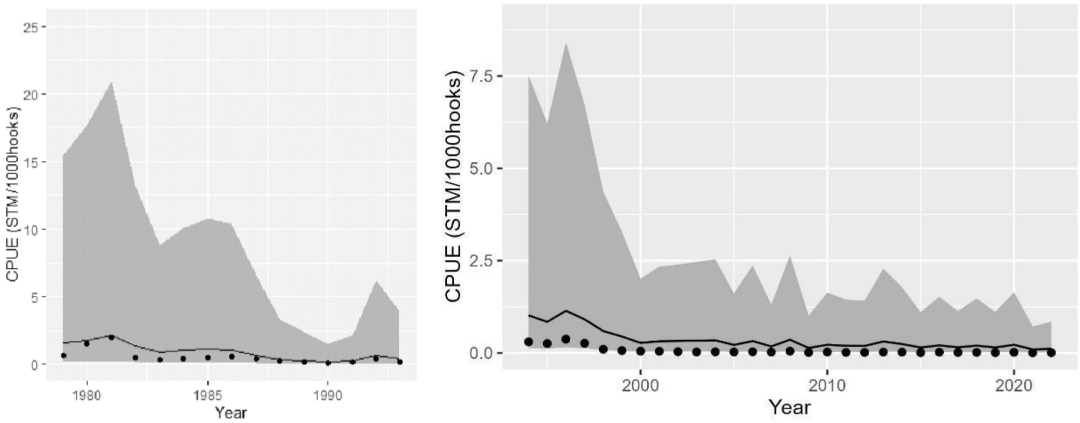


Figure 8. Historical changes of in the standardized CPUEs of striped marlin for two periods in the **Northeast** area of the Indian Ocean. Thin line and filled point denote point estimates of predicted CPUE and nominal CPUE, respectively. Gray shadow denotes 95% credible interval. Note the scale of y-axis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

Table 3. Nominal and standardized CPUEs of striped marlin for two periods of 1979-93 and 1994-2022 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	0.3016	1.0167	7.5055	0.1372
1980	1.51	1.80	0.18	17.65	1995	0.2547	0.8430	6.2210	0.1138
1981	1.99	2.13	0.21	20.91	1996	0.3734	1.1414	8.4187	0.1540
1982	0.49	1.35	0.14	13.27	1997	0.2634	0.9112	6.7200	0.1230
1983	0.35	0.90	0.09	8.84	1998	0.0990	0.5925	4.3711	0.0800
1984	0.37	1.02	0.10	10.04	1999	0.0642	0.4448	3.2812	0.0600
1985	0.53	1.10	0.11	10.81	2000	0.0483	0.2724	2.0094	0.0368
1986	0.55	1.06	0.11	10.40	2001	0.0453	0.3171	2.3407	0.0428
1987	0.44	0.67	0.07	6.57	2002	0.0356	0.3241	2.3918	0.0437
1988	0.22	0.34	0.03	3.38	2003	0.0259	0.3336	2.4683	0.0449
1989	0.21	0.25	0.02	2.41	2004	0.0304	0.3425	2.5337	0.0461
1990	0.12	0.15	0.02	1.49	2005	0.0254	0.2176	1.6153	0.0292
1991	0.17	0.22	0.02	2.15	2006	0.0326	0.3216	2.3732	0.0433
1992	0.42	0.63	0.06	6.18	2007	0.0226	0.1776	1.3119	0.0239
1993	0.18	0.40	0.04	3.92	2008	0.0491	0.3581	2.6393	0.0482
					2009	0.0131	0.1363	1.0078	0.0184
					2010	0.0250	0.2214	1.6357	0.0298
					2011	0.0151	0.1944	1.4444	0.0260
					2012	0.0145	0.1908	1.4141	0.0256
					2013	0.0221	0.3099	2.2853	0.0417
					2014	0.0215	0.2414	1.7813	0.0325
					2015	0.0095	0.1485	1.1046	0.0199
					2016	0.0194	0.2057	1.5227	0.0276
					2017	0.0128	0.1524	1.1329	0.0204
					2018	0.0149	0.1987	1.4757	0.0265
					2019	0.0088	0.1485	1.1108	0.0198
					2020	0.0169	0.2218	1.6443	0.0296
					2021	0.0059	0.0964	0.7206	0.0129
					2022	0.0082	0.1140	0.8466	0.0151

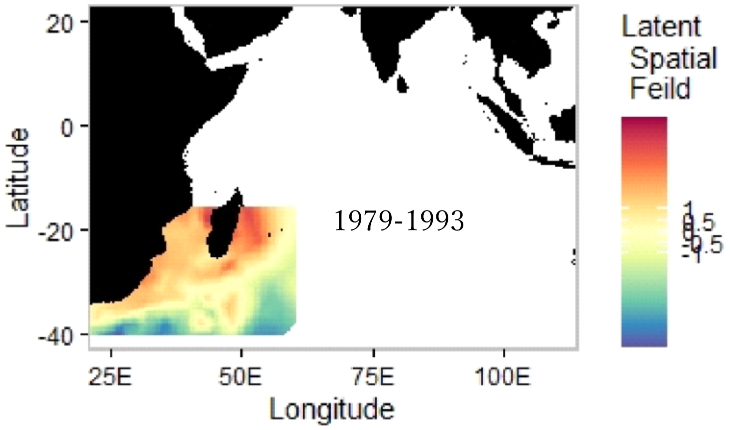


Figure 9. Spatial distribution in standardized CPUE (mean latent spatial field) of striped marlin for 1979-1993 in the **Southwest** area of the Indian Ocean.

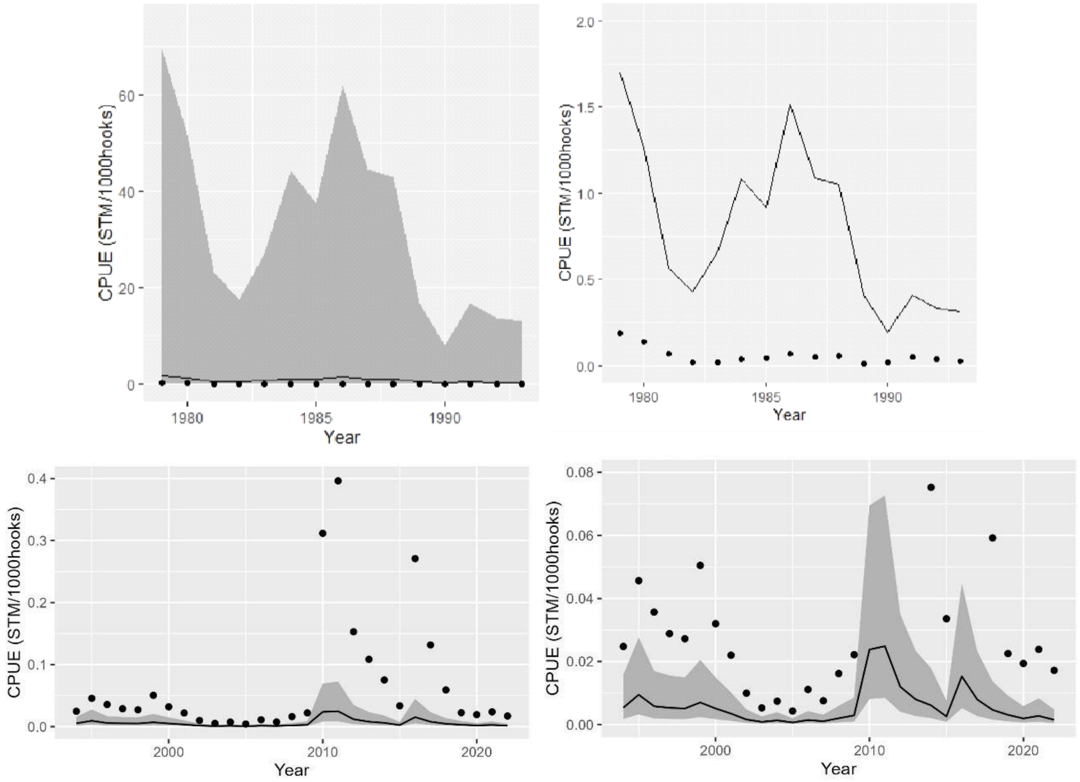
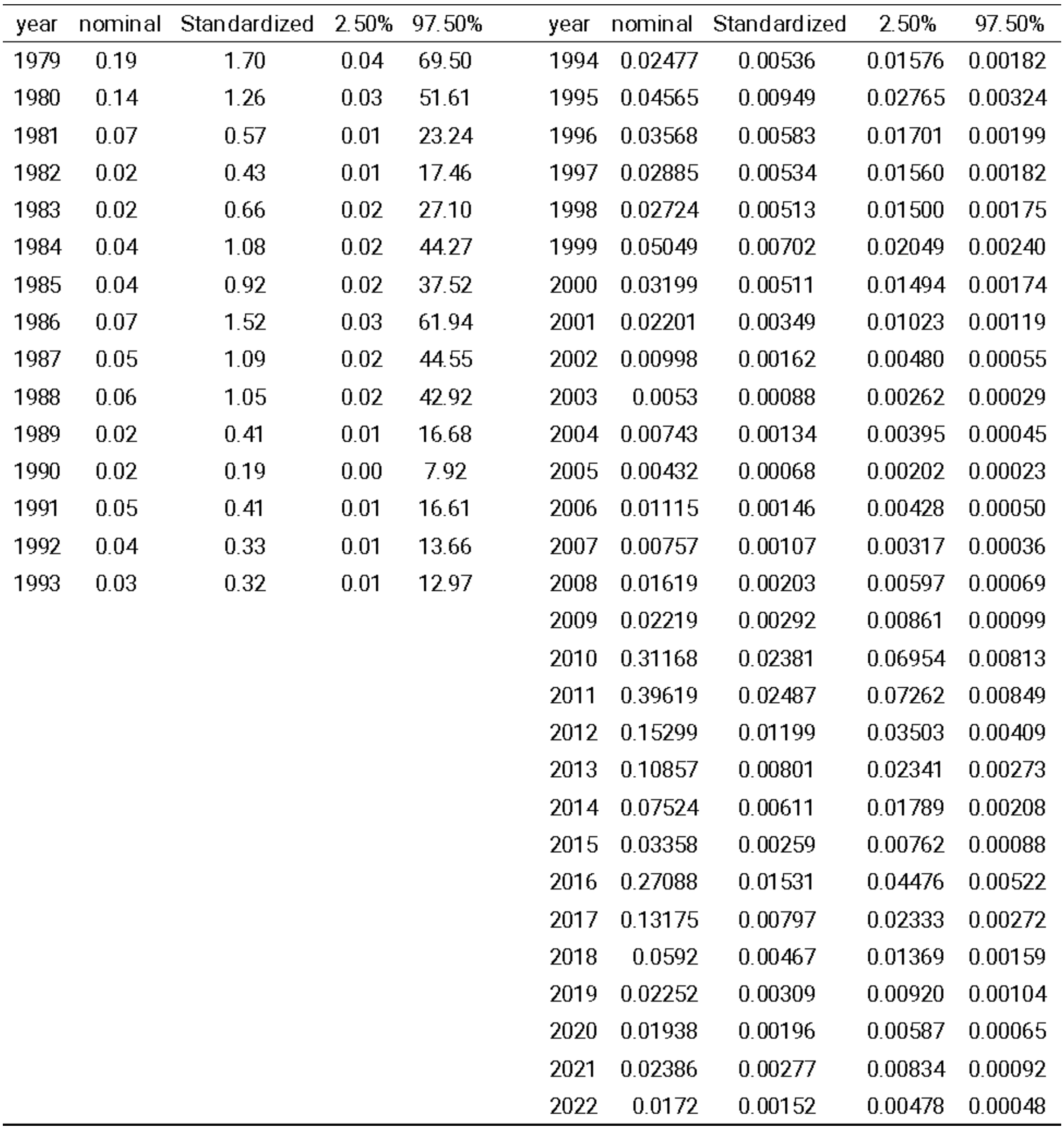


Figure 10. Historical changes in the standardized CPUEs of striped marlin for two periods in the **Southwest** area of the Indian Ocean. Thin line and filled point denote point estimates of predicted and nominal CPUE, respectively. Gray shadow denotes 95% credible interval. Right figures indicate with scale down for y-axis of the left figures. Also note that the scale of y-axis of between upper and lower figures differs.

Table 4. Nominal and standardized CPUEs of striped marlin for two periods of 1979-93 and 1994-2022 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	0.02477	0.00536	0.01576	0.00182
1980	0.14	1.26	0.03	51.61	1995	0.04565	0.00949	0.02765	0.00324
1981	0.07	0.57	0.01	23.24	1996	0.03568	0.00583	0.01701	0.00199
1982	0.02	0.43	0.01	17.46	1997	0.02885	0.00534	0.01560	0.00182
1983	0.02	0.66	0.02	27.10	1998	0.02724	0.00513	0.01500	0.00175
1984	0.04	1.08	0.02	44.27	1999	0.05049	0.00702	0.02049	0.00240
1985	0.04	0.92	0.02	37.52	2000	0.03199	0.00511	0.01494	0.00174
1986	0.07	1.52	0.03	61.94	2001	0.02201	0.00349	0.01023	0.00119
1987	0.05	1.09	0.02	44.55	2002	0.00998	0.00162	0.00480	0.00055
1988	0.06	1.05	0.02	42.92	2003	0.0053	0.00088	0.00262	0.00029
1989	0.02	0.41	0.01	16.68	2004	0.00743	0.00134	0.00395	0.00045
1990	0.02	0.19	0.00	7.92	2005	0.00432	0.00068	0.00202	0.00023
1991	0.05	0.41	0.01	16.61	2006	0.01115	0.00146	0.00428	0.00050
1992	0.04	0.33	0.01	13.66	2007	0.00757	0.00107	0.00317	0.00036
1993	0.03	0.32	0.01	12.97	2008	0.01619	0.00203	0.00597	0.00069
					2009	0.02219	0.00292	0.00861	0.00099
					2010	0.31168	0.02381	0.06954	0.00813
					2011	0.39619	0.02487	0.07262	0.00849
					2012	0.15299	0.01199	0.03503	0.00409
					2013	0.10857	0.00801	0.02341	0.00273
					2014	0.07524	0.00611	0.01789	0.00208
					2015	0.03358	0.00259	0.00762	0.00088
					2016	0.27088	0.01531	0.04476	0.00522
					2017	0.13175	0.00797	0.02333	0.00272
					2018	0.0592	0.00467	0.01369	0.00159
					2019	0.02252	0.00309	0.00920	0.00104
					2020	0.01938	0.00196	0.00587	0.00065
					2021	0.02386	0.00277	0.00834	0.00092
					2022	0.0172	0.00152	0.00478	0.00048

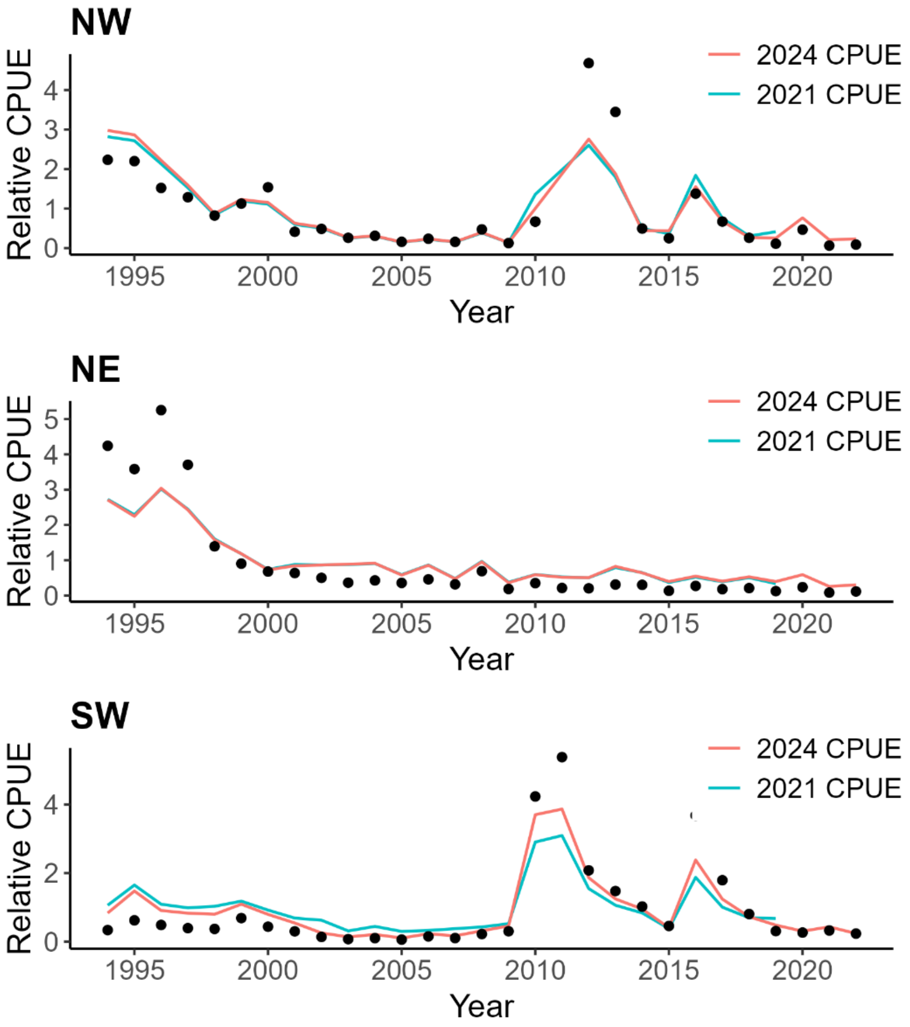


Figure 11. Comparison of annual standardized CPUE of striped marlin (relative to its mean value for 1994-2019) between present (red line) and previous (blue line , Taki et al., 2021) studies for the defined area of the Indian Ocean. Black dots denote nominal CPUE.