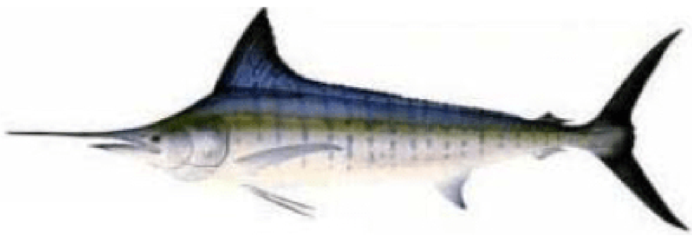
Japanese longline CPUE of the striped marlin (*Kajikia audax*) in the WCNPO. [1]

*Hirotaka Ijima and **Minoru Kanaiwa
E-mail:ijima@affrc.go.jp

*National Research Institute of Far Seas Fisheries, Fisheries Research and Education Agency
5-7-1, Orido, Shimizu, Shizuoka, 424-8633, Japan
**Mie University
1577, Kurima-Machiyacho, Tsu, Mie, 514-8507, Japan



[1]This working paper was submitted to the ISC Billfish Working Group Intercessional Workshop, 14-21 January 2019, held in the Pacific Islands Fisheries Science Center of the National Marine Fisheries Service, Hawaii USA.

Abstract

We standardized CPUE of striped marlin caught by Japanese longline vessel in the WCNPO. We standardized CPUE for two fleets that depended on the analysis results of the finite mixture model. Also, these two fleets were divided into two time series in the year when the format of Japanese logbook statistics changed (early series: 1976-1993, late series: 1994 - 2017). In this analysis, we tested three candidate models for goodness-of-fit to the CPUE data: (i) negative binomial distribution GLMM (NB), (ii) zero-inflated Poisson GLMM (ZIP), and (iii) zero-inflated negative binominal GLMM (ZINB). For the model diagnostics, we used Bayesian information criterion (BIC), model deviance explained and randomized quantile residuals. For all 4 fleets, the lowest BIC was produced by the NB model and followed by the ZIP. The most complicated ZINB model did not converge or became under dispersed and as a result was eliminated as a candidate model. Randomized quantile residuals showed generally good dispersion in both candidate models (NB and ZIP). There is a possibility that the operational pattern of Japanese longline fishery in the later period had gradually changed. However, there were no data to confirm such a change. To address this issue, we attempted to estimate the true zero catch using the ZIP model. However, the ZIP model could not incorporate the change of fishing operation entirely. In future work, it may be necessary to consider a more complex hierarchical random effects model.

Introduction

In the last stock assessment, the ISC Billfish working group (BILLWG) conducted the striped marlin stock assessment of the Western Central North Pacific ocean using Stock Synthesis 3 (SS3)(ISC 2015). In the SS 3, Fleet is defined for each selectivity of different fishing gear. Japanese longline fishery was classified into thirteen fisheries according to the finite mixture model analysis result. For the stock assessment, it is necessary to prepare catch amount, length frequency data, and standardized CPUE for each fleet. In this document, we explored several regression models to standardize four-time series CPUE that are the area 3 in quarter one of the early period, area 3 in quarter one of the late period, the area 1 in quarter one of the early period, area 1 in quarter one of the late period (Figure.1). These fleets represent the biomass trend of juvenile and young-adult striped marlin.

Material and methods

To standardize striped marlin CPUE, we used Japanese logbook data. We used data from 1976 on which includes the vessel names. Also, we divided the Japanese logbook into two-time series that are 1976-1993 and 1994-2017, because the observe method was changed in 1993.

To select an appropriate statistical model, we compared and examined multiple the GLMM (Table.1−4). In the GLMM, we treated a vessel name as a random effect because the random effect raises the accuracy of the model (Ochi, Ijima, Kinoshita, and Kiyofuji 2016). As well as the effect of the vessel, we organized the area by a grid and treated it as a variable effect because a pesudopreplication of marlin is occurring. For the probability density distribution of the response variable, we examined Poisson distribution (PO), Negative Binomial distribution (NB), Zero-inflated Poisson distribution (ZIP) and Zero-inflated Negative Binomial distribution (ZINB).

We considered the fixed effect of explanatory variables that are the effect of the year, the effect of the moth and the effect of the gear. The definition of the gear types is deep sets of more than eight hooks between floats and shallow sets less than that. We do not use the interaction between explanatory variables to avoid overfitting. If the impact of interaction is significant, hierarchical modeling will be required. The parameters of all models were estimated by maximum likelihood estimation using R software package "glmmTMB" (Bolker 2016).

BIC selected these multiple models. We also focused on over dispersion of the model in model selection. We used randomized quantile residuals instead of Pearson residuals for model validation (Bai 2018). Explanatory variables aggregate the randomized quantile residuals. Finally, we calculated the least square means using R software package "emmeans".

Result and discussion

As a result of model selection, the smallest BIC was obtained by the simple NB GLMM and ZIP was the secondly small BIC model at all CPUE time series (Table.5−8). Comparing the degree of improvement of the Deviance from the Null model, the late period of Area 1 in quarter 1 has no difference in performance between NB and ZIP model. In the other models, the simple NB model makes Deviance smaller than other models (Table.5−8). Randomized quantile residuals was uniformly distributed in both NB and ZIP in all CPUE time series (Figure.2−9). Comparing the standardized CPUE, there was no substantial difference in trends in the early period, but there was a difference in the trend in the late period (Figure.10). It seems to be related to the fact that the depth of the branch lines in the longline fishery has gradually become deeper since the middle of the 1990s. However, the number of hooks between floats does not represent the depth of gear. For example, the depth of blanch line also determines by the type of line and the vessel speed at throwing the main lines. Focus on the transition of the number of hooks between floats, we can see that the operation pattern of the longline fishery has gradually changed (Figure.11). The zero catch rate is thought to be higher when the depth of the branch line gets deeper. However, such an operation shift was not structurally incorporated in the NB model. Thus, there is a possibility that overfitting to the data is occurring.

References

Bai, W. (2018). *Randomized Quantile Residual for Assessing Generalized Linear Mixed Models with Application to Zero-Inflated Microbiome Data*. Ph. D. thesis, University of Saskatchewan.

Bolker, B. (2016). Getting started with the glmmtmb package.

ISC (2015). Stock assessment update for striped marlin (kajika audax) in the western and central north pacific ocean through 2013.

Ochi, D., H. Ijima, J. Kinoshita, and H. Kiyofuji (2016). New fisheries definition from japanese longline north pacific albacore size data. *ISC/16/ALBWG-02/06*.

Table 1: Candidate of a statistical model for adult CPUE index (quarter 1 in area 1) in the early period.

Model	Distribution	linear regression
m1	PO	stm~yr+month+gear+latlon+(1|jpname)+offset(log(hooks))
m2	PO	stm~yr+month+gear+(1|jpname)+(1|latlon2)+offset(log(hooks))
m3	ZIP	stm~yr+month+gear+(1|jpname)+(1|latlon2)+offset(log(hooks))
		zi~yr+month+gear
m4	NB	stm~yr+month+gear+(1|jpname)+(1|latlon2)+offset(log(hooks))
m5	ZINB	stm~yr+month+gear+(1|jpname)+(1|latlon2)+offset(log(hooks))
		zi~yr+month+gear
m6	ZIP	stm~yr+month+gear+(1|latlon2)+(1|jpname)+offset(log(hooks))
		zi~yr+(1|latlon2)+(1|jpname)

Table 2: Candidate of a statistical model for adult CPUE index (quarter 1 in area 1) in the late period.

Model	Distribution	linear regression
m1	PO	stm~yr+month+gear+latlon+(1|jpname)+offset(log(hooks))
m2	PO	stm~yr+month+gear+(1|jpname)+(1|latlon2)+offset(log(hooks))
m3	ZIP	stm~yr+month+gear+(1|jpname)+(1|latlon2)+offset(log(hooks))
		zi~yr+month+gear
m4	NB	stm~yr+month+gear+(1|jpname)+(1|latlon2)+offset(log(hooks))
m5	ZINB	stm~yr+month+gear+(1|jpname)+(1|latlon2)+offset(log(hooks))
		zi~yr+month+gear
m6	ZIP	stm~yr+month+gear+(1|latlon2)+offset(log(hooks))
		zi~yr+(1|latlon2)
m7	ZIP	stm~yr+month+gear+(1|latlon2)+offset(log(hooks))
		zi~yr+(1|jpname)
m8	ZIP	stm~yr+month+gear+(1|latlon2)+offset(log(hooks))
		zi~yr+(1|jpname)+(1|latlon2)
m9	ZIP	stm~yr+gear+(1|jpname)+(1|latlon2)+offset(log(hooks))
		zi~yr+gear+(1|jpname)+(1|latlon2)
m10	ZIP	stm~yr+month+gear+(1|jpname)+(1|latlon2)+offset(log(hooks))
		zi~yr+(1|jpname)+(1|latlon2)

Table 3: Candidate of a statistical model for juvenile CPUE index (quarter 1 in area 3) in the early period.

Model	Distribution	linear regression
m1	PO	stm~yr+month+gear+latlon+(1|jpname)+offset(log(hooks))
m2	PO	stm~yr+month+gear+(1|jpname)+(1|latlon2)+offset(log(hooks))
m3	ZIP	stm~yr+month+gear+(1|jpname)+(1|latlon2)+offset(log(hooks))
		zi~yr+month+gear
m4	NB	stm~yr+month+gear+(1|jpname)+(1|latlon2)+offset(log(hooks))
m5	ZINB	stm~yr+month+gear+(1|jpname)+(1|latlon2)+offset(log(hooks))
		zi~yr+month+gear
m6	ZIP	stm~yr+month+gear+(1|latlon2)+(1|jpname)+offset(log(hooks))
		zi~yr+gear+(1 |latlon2)+(1|jpname)
m7	ZIP	stm~yr+month+gear+(1|latlon2)+(1|jpname)+offset(log(hooks))
		zi~yr+month+gear+(1|latlon2)+(1|jpname)

Table 4: Candidate of a statistical model for juvenile CPUE index (quarter 1 in area 3) in the late period.

Model	Distribution	linear regression
m1	PO	stm~yr+month+gear+latlon+(1|jpname)+offset(log(hooks))
m2	PO	stm~yr+month+gear+(1 |jpname)+(1|latlon2)+offset(log(hooks))
m3	ZIP	stm~yr+month+gear+(1 |jpname)+(1|latlon2)+offset(log(hooks))
		zi~yr+month+gear
m4	NB	stm~yr+month+gear+(1 |jpname)+(1|latlon2)+offset(log(hooks))
m5	ZINB	stm~yr+month+gear+(1 |jpname)+(1|latlon2)+offset(log(hooks))
		zi~yr+month+gear
m6	ZIP	stm~yr+month+gear+(1|latlon2)+(1|jpname)+offset(log(hooks))
		zi~yr+gear+(1 |jpname)+(1|latlon2)
m7	ZIP	stm~yr+month+gear+(1|latlon2)+(1|jpname)+offset(log(hooks))
		zi~yr+month+gear+(1|jpname)+(1|latlon2)

Table 5: Deviance table of adult CPUE index (quarter 1 in area 1) in the early period.

Model	Df	AIC	BIC	Deviance	Chi sq	Pr	%Chi null	Dispersion
null	1	104564.45	104573.47	104562.45	–	–	–	4.88
m2	23	80570.05	80777.56	80524.05	24038.40	0	0.23	1.45
m4	24	75367.60	75584.14	75319.60	5204.45	0	0.28	1.02
m1	39	81474.38	81826.25	81396.38	0.00	1	0.22	1.56
m6	43	76185.12	76573.07	76099.12	5297.27	0	0.27	0.98
m3	44	77388.20	77785.18	77300.20	0.00	1	0.26	1.09
m5	45	75199.67	75605.67	75109.67	2190.53	0	0.28	0.24

Table 6: Deviance table of adult CPUE index (quarter 1 in area 1) in the late period.

Model	Df	AIC	BIC	Deviance	Chi sq	Pr	%Chi null	Dispersion
null	1	53205.34	53213.98	53203.34	–	–	–	3.43
m2	29	41862.41	42113.07	41804.41	11398.93	0	0.21	1.35
m4	30	39513.08	39772.39	39453.08	2351.33	0	0.26	1.04
m1	46	42168.64	42566.25	42076.64	0	1	0.21	1.43
m6	53	41646.46	42104.57	41540.46	536.18	>0.05	0.22	1.15
m7	53	41009.46	41467.57	40903.46	637.01	0	0.23	1.04
m8	54	40724.53	41191.28	40616.53	286.93	>0.05	0.24	1.01
m9	54	39733.68	40200.44	39625.68	990.84	0	0.26	1.02
m10	55	39718.56	40193.96	39608.56	17.13	>0.05	0.26	1.02
m3	56	40155.49	40639.54	40043.49	0	1	0.25	1.03
m5	57	39356.02	39848.71	39242.02	801.47	>0.05	0.26	0.52

Table 7: Deviance table of juvenile CPUE index (quarter 1 in area 3) in the early period.

Model	Df	AIC	BIC	Deviance	Chi sq	Pr	%Chi null	Dispersion
null	1	268198.43	268207.26	268196.43	–	–	–	6.81
m2	23	183912.14	184115.23	183866.14	84330.29	0	0.31	2.50
m4	24	155969.01	156180.94	155921.01	27945.12	0	0.42	1.04
m3	44	171863.45	172251.98	171775.45	0	1	0.36	1.48
m6	44	171070.90	171459.42	170982.90	792.56	0	0.36	1.46
m5	45	–	–	–	–	–	–	0.13
m7	46	169770.95	170177.13	169678.95	–	–	0.37	1.46
m1	53	191297.53	191765.53	191191.53	0	1	0.29	2.79

Table 8: Deviance table of juvenile CPUE index (quarter 1 in area 3) in the late period.

Model	Df	AIC	BIC	Deviance	Chi sq	Pr	%Chi null	Dispersion
null	1	162139.49	162147.83	162137.49	-	-	-	6.50
m2	29	118960.56	119202.51	118902.56	43234.93	0	0.27	3.35
m4	30	96467.41	96717.70	96407.41	22495.15	0	0.41	1.26
m3	56	106526.48	106993.69	106414.48	0	1	0.34	1.65
m6	56	105784. 0 7	106251.27	105672.07	742.42	0	0.35	1.70
m1	57	121453.51	121929.06	121339.51	0	1	0.25	3.68
m5	57	95445.78	95921.33	95331.78	26007.73	0	0.41	0.26
m7	58	105676.93	106160.82	105560.93	0	1	0.35	1.70

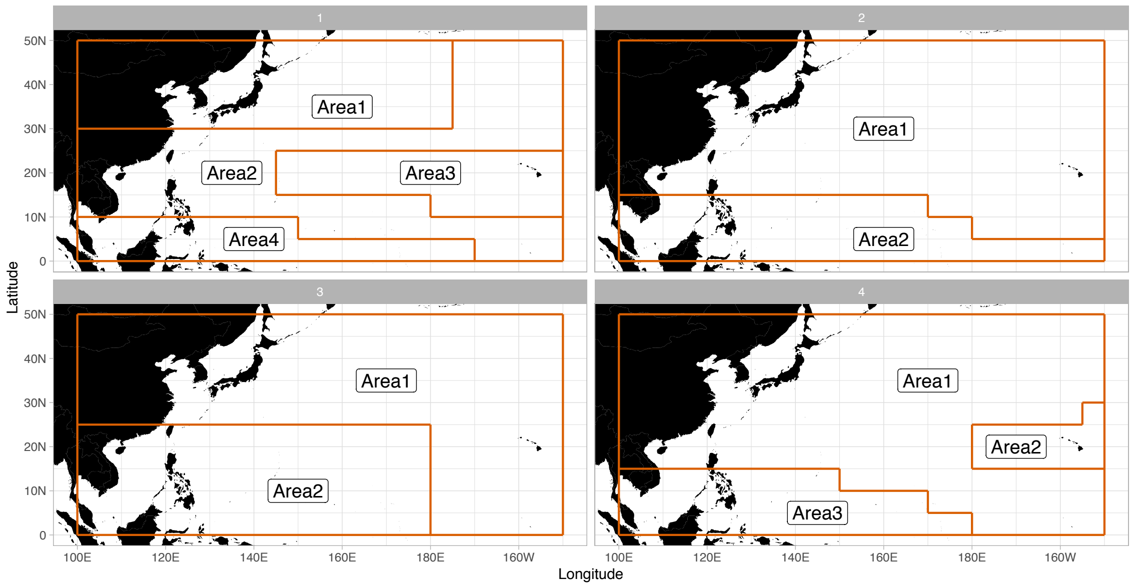


Figure 1: Area-quarterly fleet definition of Japanese longline fishery. we standardized two CPUE time series that are quarter1 on area1and quarter3 on area1. Also, these two fleets were divided two periods as "early" and "late".

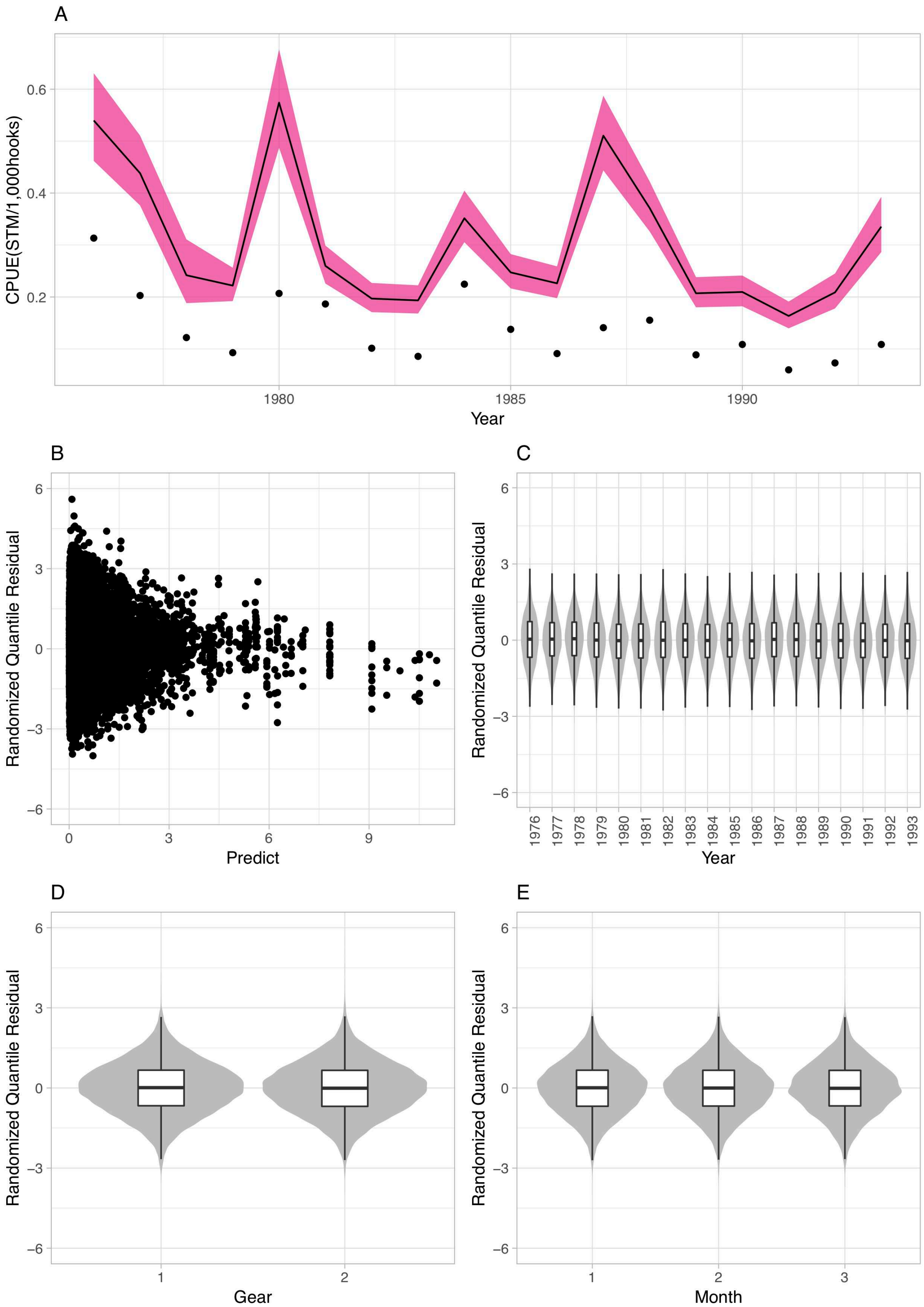


Figure 2: Results of adult CPUE index (quarter1 in area 1) in early period using a negative binomial GLMM. **A**: Historical changes of longline CPUE. Black line is a standardized CPUE, Black points denote nominal CPUE and filled areas is 95% confidence interval. **B**: Plots of randomized quantile residuals versus predicted value. **C-E**: The trends of randomized quantile residuals by covariates.

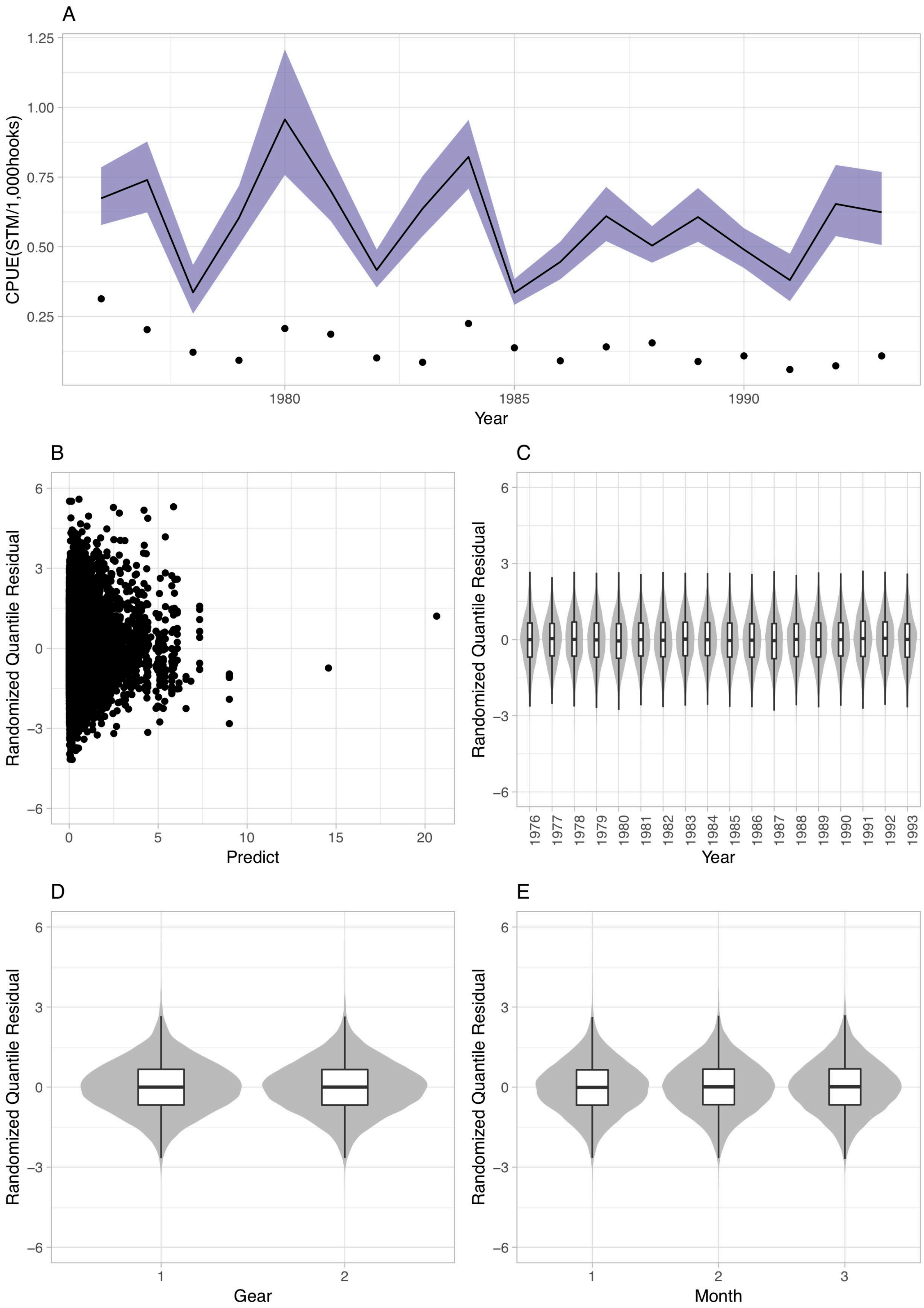


Figure 3: Results of adult CPUE index (quarter1 in area 1) in early period using a zero-inflated poisson GLMM. **A**: Historical changes of longline CPUE. Black line is a standardized CPUE, Black points denote nominal CPUE and filled areas is 95% confidence interval. **B**: Plots of randomized quantile residuals versus predicted value. **C-E**: The trends of randomized quantile residuals by covariates.

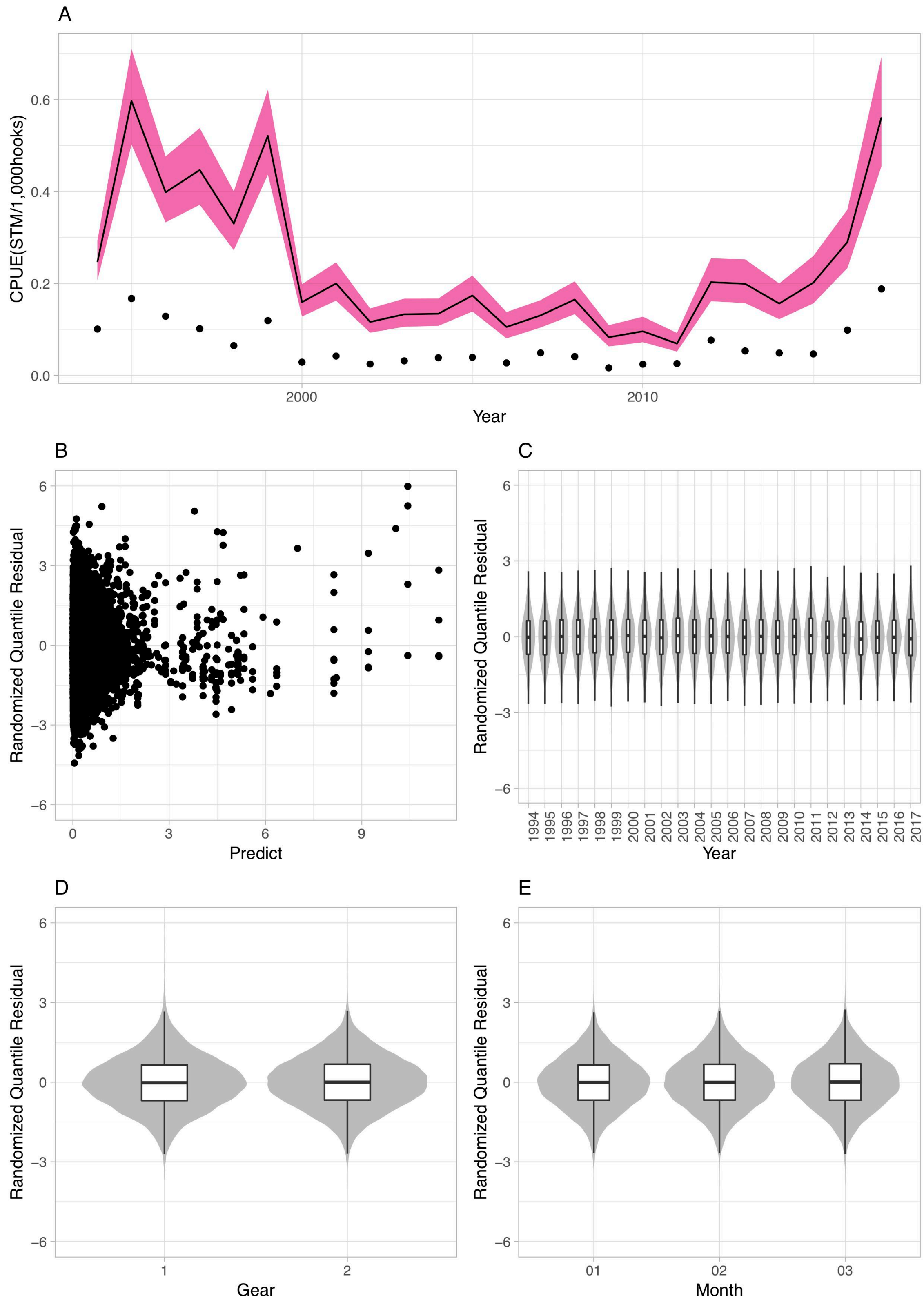


Figure 4: Results of adult CPUE index (quarter1 in area 1) in late period using a negative binomial GLMM. **A**: Historical changes of longline CPUE. Black line is a standardized CPUE, Black points denote nominal CPUE and filled areas is 95% confidence interval. **B**: Plots of randomized quantile residuals versus predicted value. **C-E**: The trends of randomized quantile residuals by covariates.

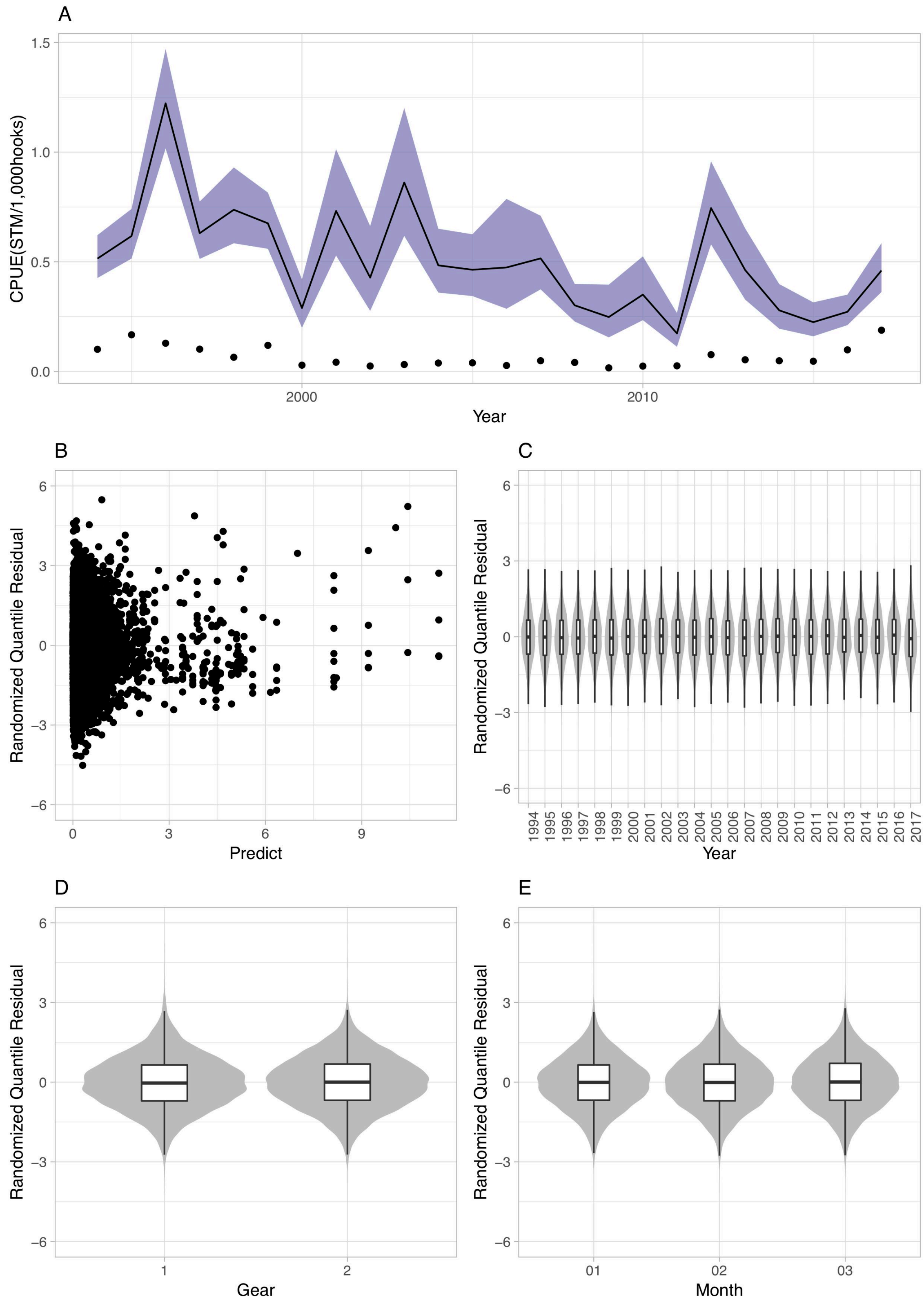


Figure 5: Results of adult CPUE index (quarter1 in area 1) in late period using a zero-inflated poisson GLMM. **A**: Historical changes of longline CPUE. Black line is a standardized CPUE, Black points denote nominal CPUE and filled areas is 95% confidence interval. **B**: Plots of randomized quantile residuals versus predicted value. **C-E**: The trends of randomized quantile residuals by covariates.

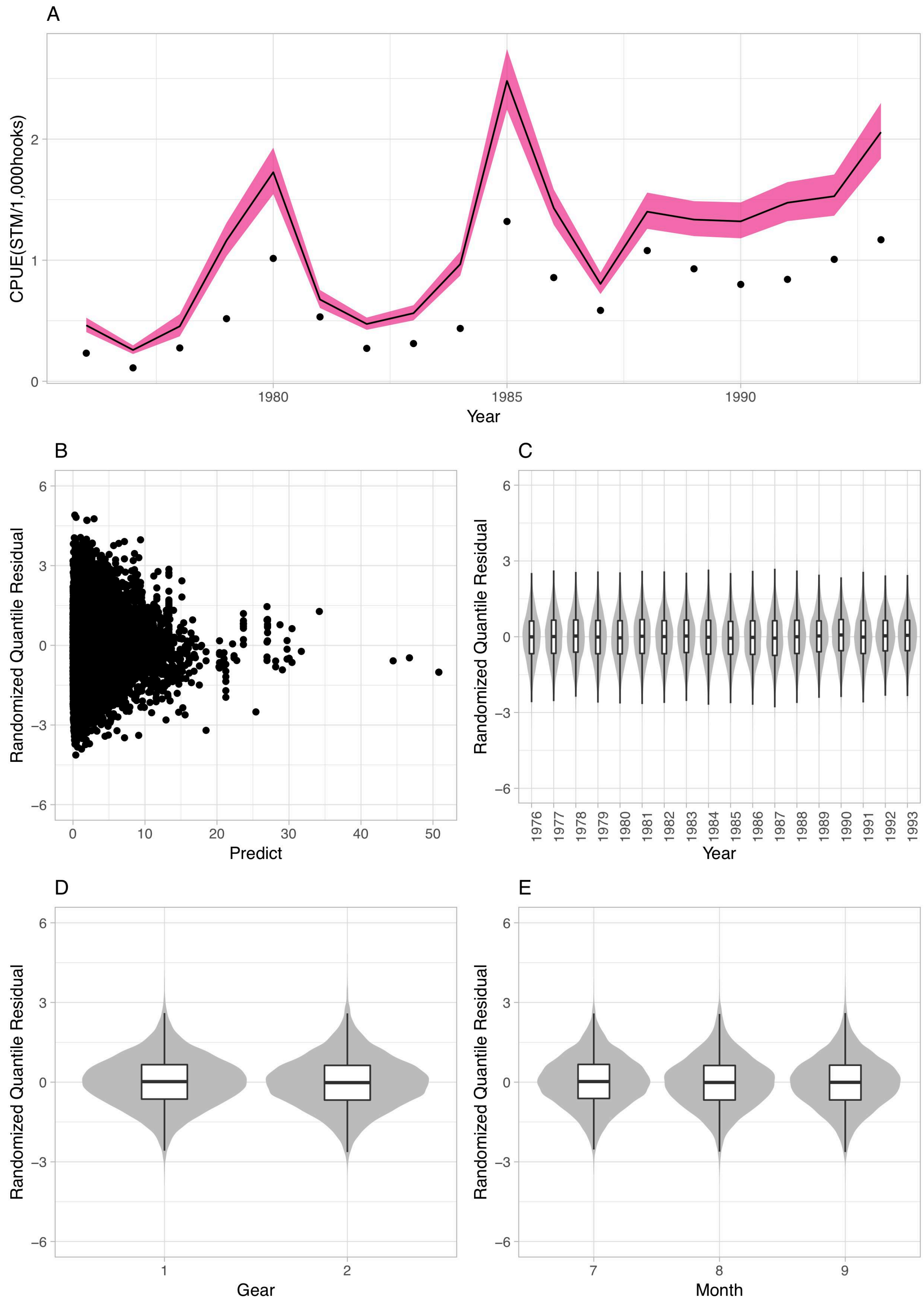


Figure 6: Results of juvenile CPUE index (quarter1 in area 3) in early period using a negative binomial GLMM. **A**: Historical changes of longline CPUE. Black line is a standardized CPUE, Black points denote nominal CPUE and filled areas is 95% confidence interval. **B**: Plots of randomized quantile residuals versus predicted value. **C-E**: The trends of randomized quantile residuals by covariates.

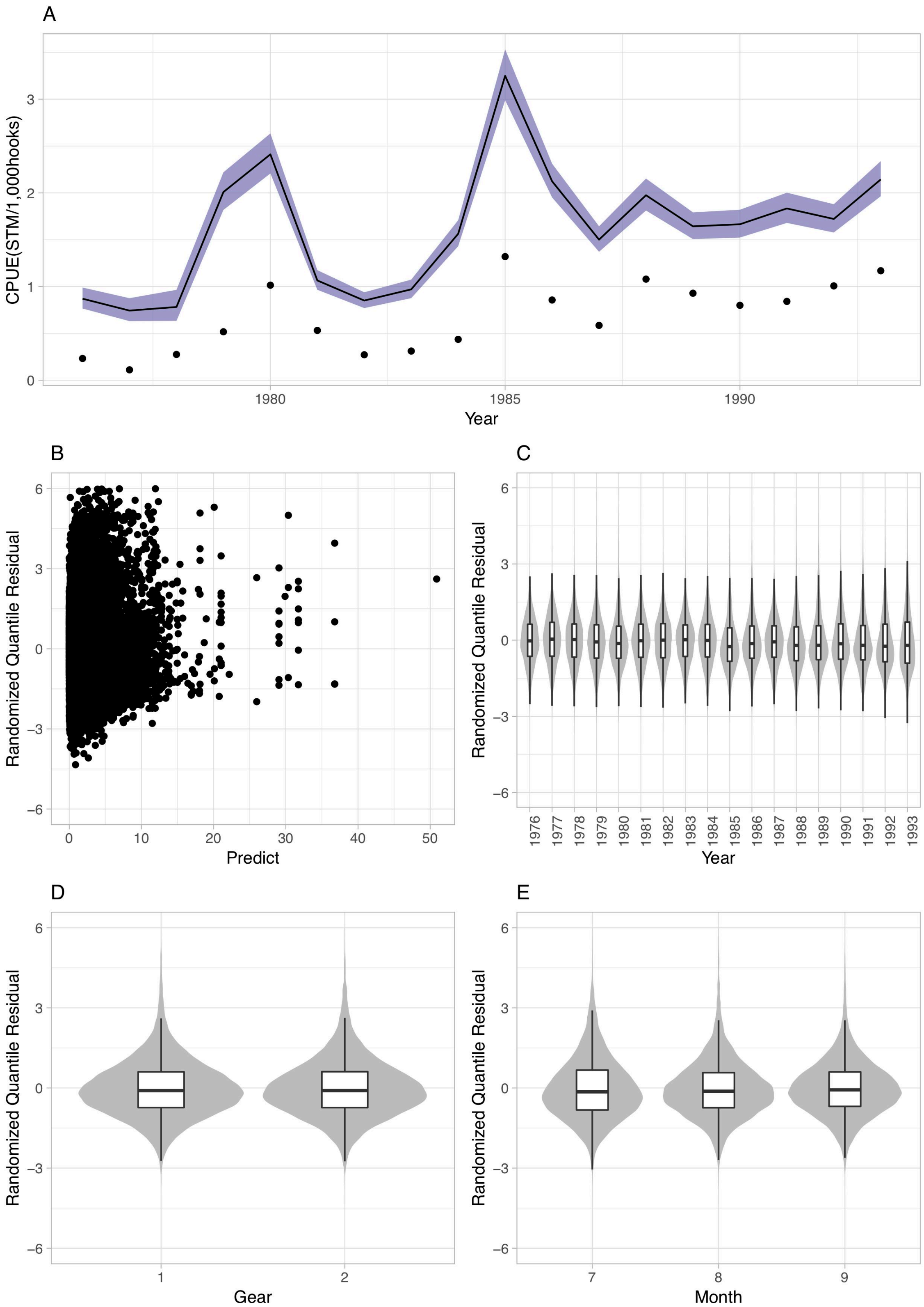


Figure 7: Results of juvenile CPUE index (quarter1 in area 3) in early period using a zero-inflated poisson GLMM. **A**: Historical changes of longline CPUE. Black line is a standardized CPUE, Black points denote nominal CPUE and filled areas is 95% confidence interval. **B**: Plots of randomized quantile residuals versus predicted value. **C-E**: The trends of randomized quantile residuals by covariates.

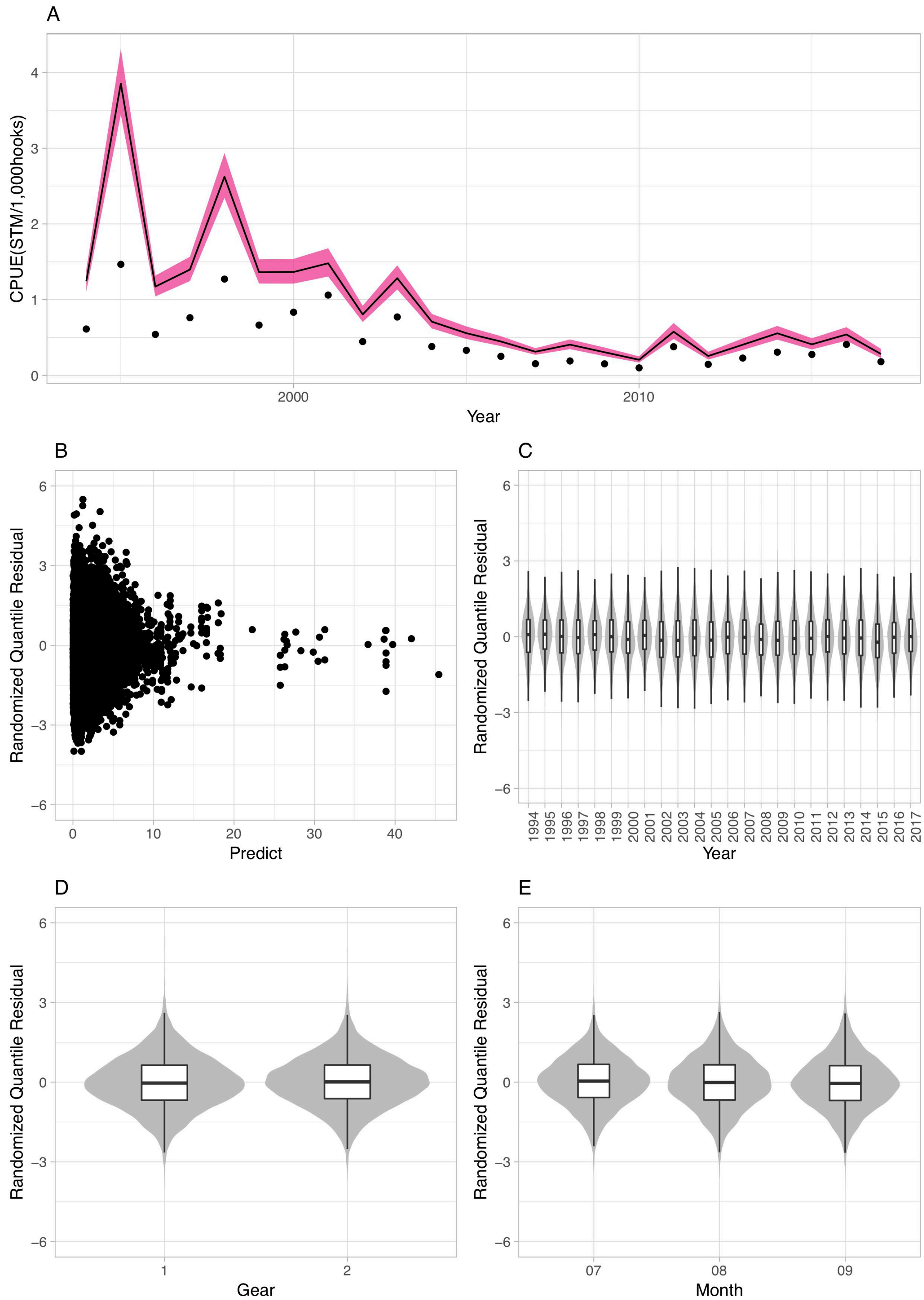


Figure 8: Results of juvenile CPUE index (quarter1 in area 3) in late period using a negative binomial GLMM. **A**: Historical changes of longline CPUE. Black line is a standardized CPUE, Black points denote nominal CPUE and filled areas is 95% confidence interval. **B**: Plots of randomized quantile residuals versus predicted value. **C-E**: The trends of randomized quantile residuals by covariates.

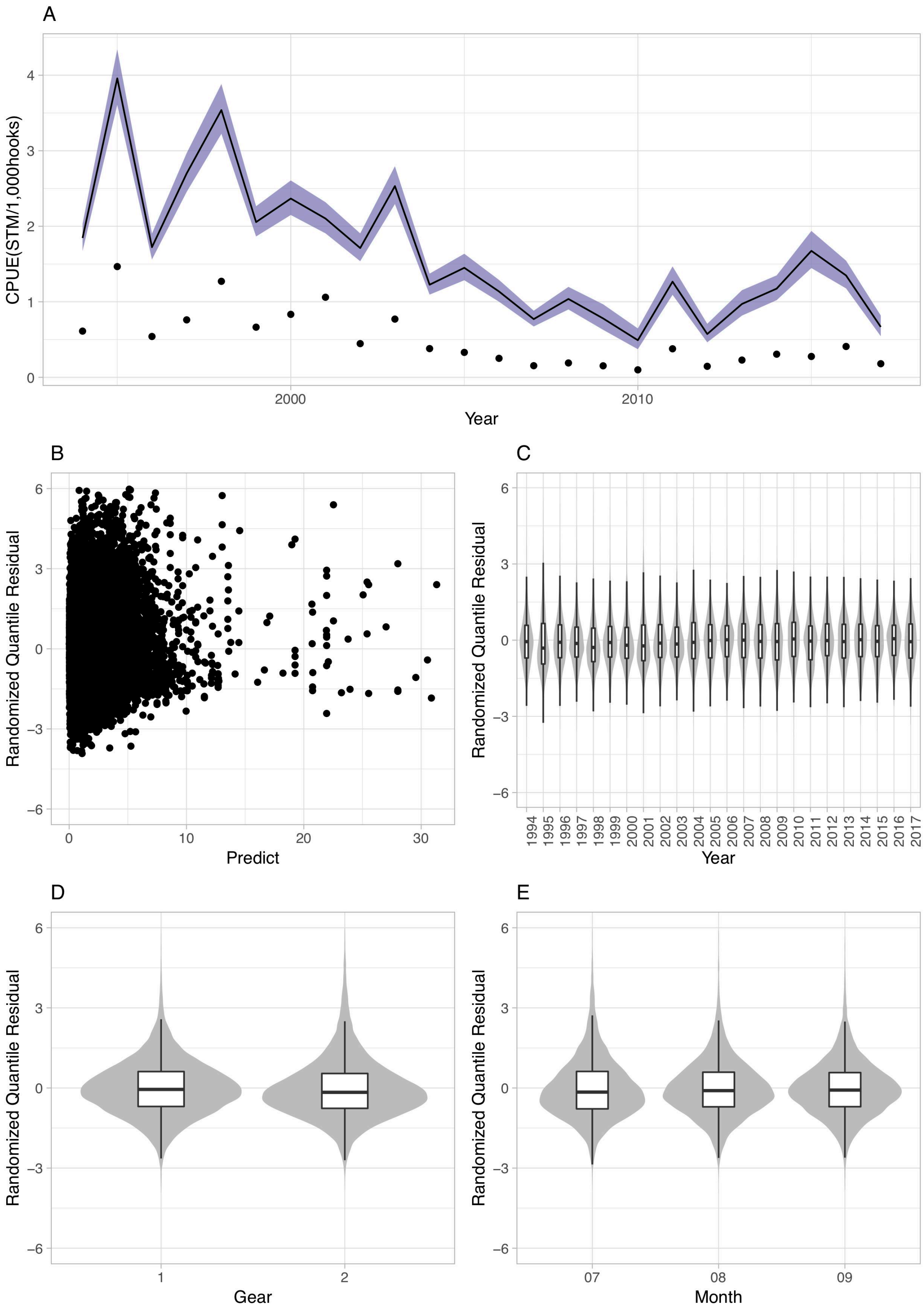


Figure 9: Results of juvenile CPUE index (quarter1 in area 3) in late period using a zero-inflated poisson GLMM. **A**: Historical changes of longline CPUE. Black line is a standardized CPUE, Black points denote nominal CPUE and filled areas is 95% confidence interval. **B**: Plots of randomized quantile residuals versus predicted value. **C-E**: The trends of randomized quantile residuals by covariates.

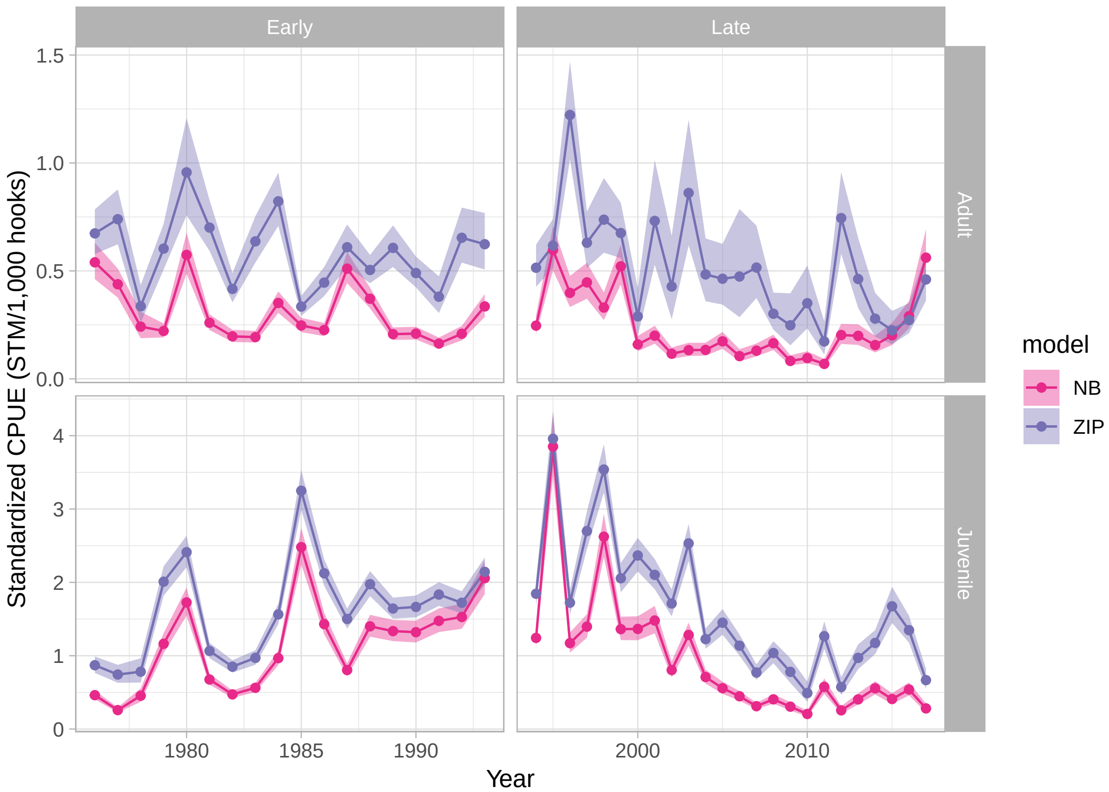
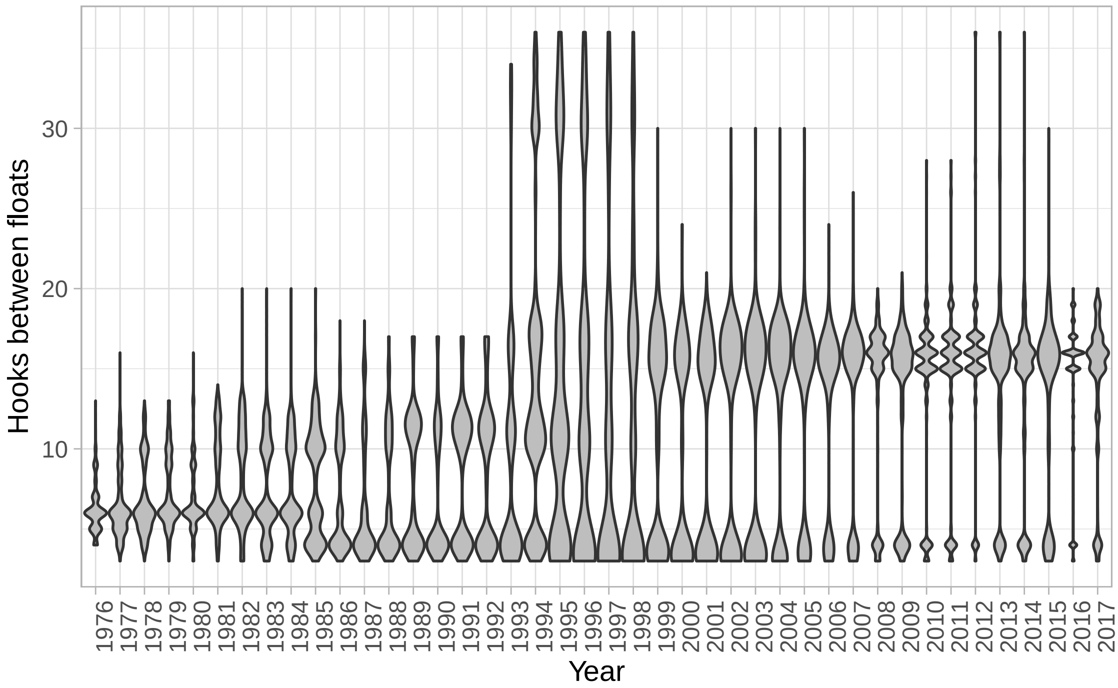


Figure 10: Standardize CPUE of Japanese longline fishery.

A Quarter1 area1 (Adult)



B Quarter3 area1 (Juvenile)

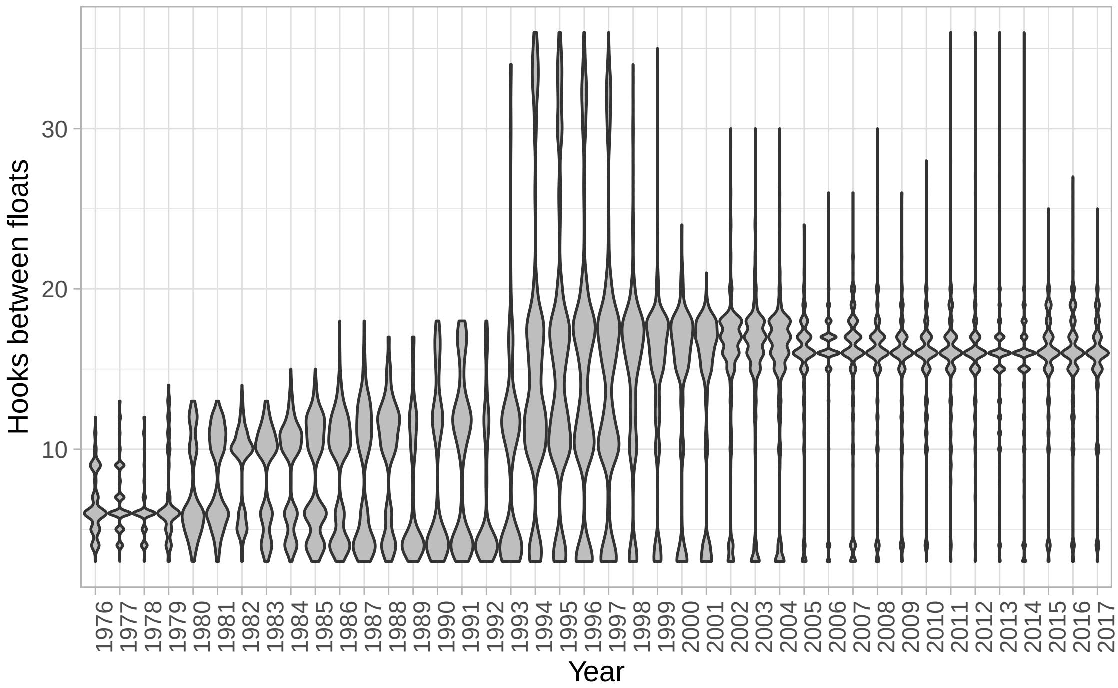


Figure 11: Historical changes of hooks between floats.