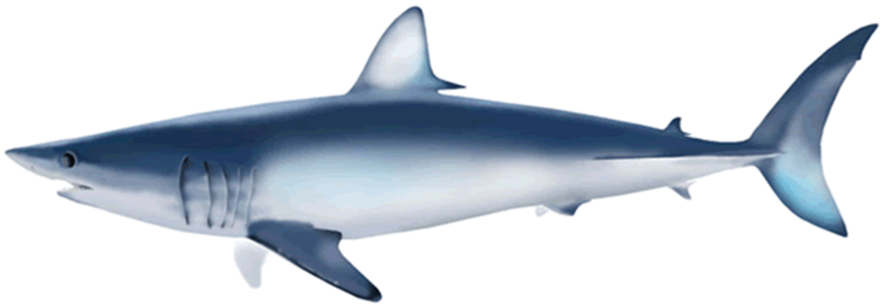
ISC/17/SHARKWG-3/06

Stock Abundance Indices for Mako Shark estimated by observer data of Japanese longline data in the North Pacific Ocean[1]

Minoru Kanaiwa[2], Yasuko Semba[3] and Mikihiko Kai[3]

[2]Mie University, Faculty of Bioresources
1577 Kurimamachiya-cho Tsu, Mie 514-8507, JAPAN
[3]National Research Institute of Far Seas Fisheries, Fishery Research Agency
5-7-1 Orido, Shimizu-ku, Shizuoka 424-8633, JAPAN

Email: kanaiwa@bio.mie-u.ac.jp



.

[1]Working document submitted to the ISC Shark Working Group Workshop, 28 November - 4 December 2017, Shizuoka, JAPAN. **Document not to be cited without author's permission.**

Abstract

In this paper, standardized CPUEs estimated by the data set of observer data of Japanese longline operated in North Pacific Ocean were provided between 2011 and 2016. The optimal statistical model was the generalized additive mixture model which cruse ID was assumed as an random factor and was assumed that the latitude and longitude were followed by 2 dimensional 3 degree spline function. The annual trend estimated by optimal model showed flat.

Introduction

There are several issues to use log book data of commercial catch information to standardize a catch per unit effort as stock abundance indices, e.g. uncertainty and misreporting of recording by fishermen. In contrast observer's data of commercial catch are considered more certain but has different issue, e.g. limitation of operational area and season.

Under such a condition, it is important to compare the trend between standardized results by using logbook and observer data to evaluate the certainty of logbook data. In this paper, the standardized indices are provided to address this issue.

Material and method

data sets

Observer set by set data of Japanese longline operated in North Pacific Ocean are used in this analysis. The observer program was started in 2008 but because of the number of data, the data between 2011 and 2016 were used for this analysis (Table 1).

Table 1. The observed number of sets in each year

year	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017
No.	52	68	155	267	650	665	1353	1841	1149	176

We also clarified the data and the number of used data showed below (Table 2).

Tablue 2. Used number of sets in this analysis

year	2011	2012	2013	2014	2015	2016
No.	267	646	665	1352	1841	1118

The operational area used in this analysis was plotted in Figs. 1. The data gathered around the coast of Japan.

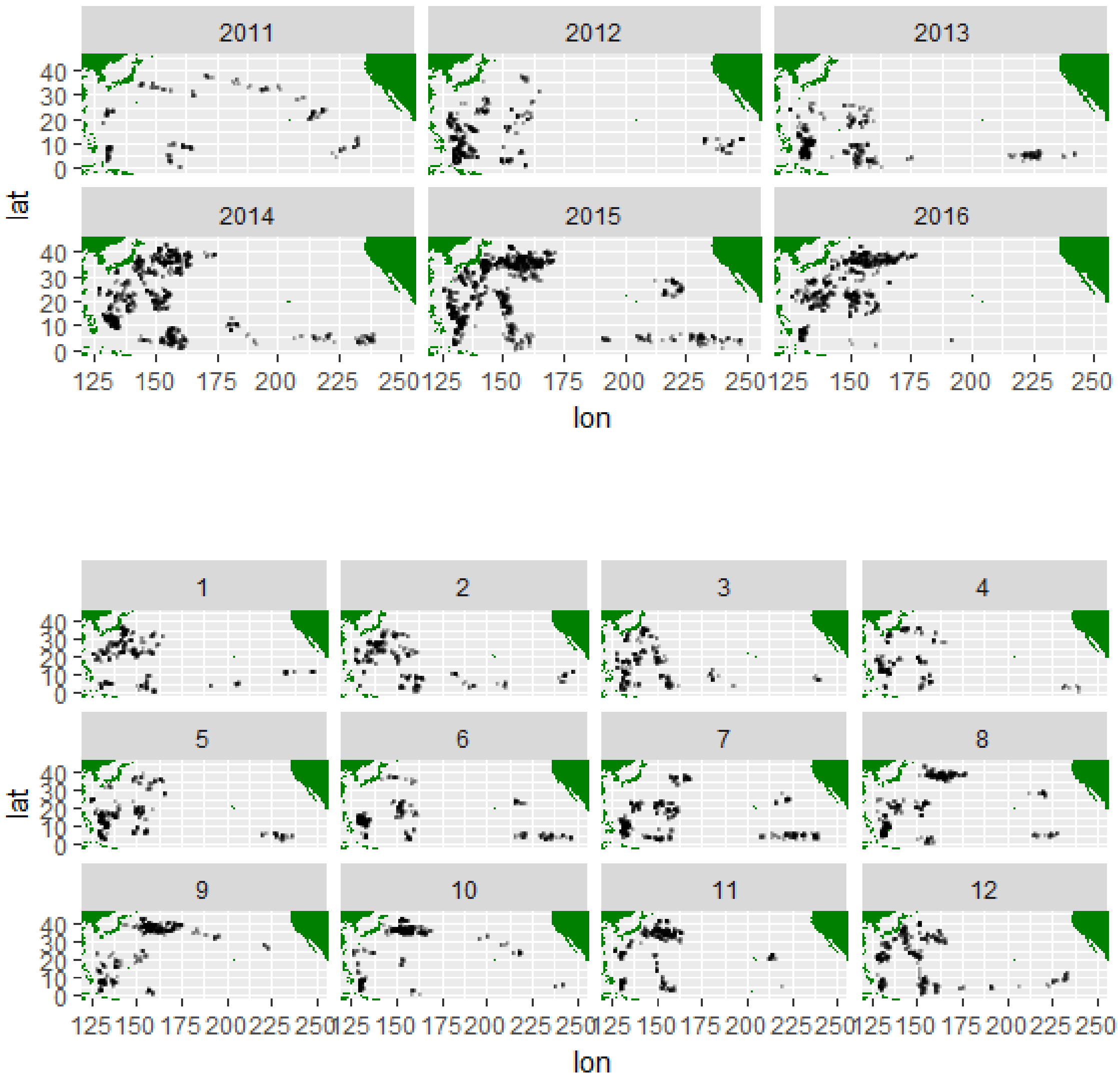


Figure 1. Spatioal distribution of used data. Upper figure shows the distribtuion in each year and lower figure shows one in each month.

Statistical methods

Observed catch number of mako shark was used as a dependent factor and two categories of hooks per baskets (<=4 and >4), year and month were used as independent categorical factors and latitude and longitude were used as independent continuous factors. The logarithm of observed hooks was used as an offset value. Catch number was assumed to be followed by Poisson distribution except zero inflated model.

We tested several models. 1) generalized linear model (GLM), 2), generalized linear mixture model which cruse ID was assumed as random factor (GLMM1), 3) generalized linear mixture model which observer name was assumed as random factor (GLMM2) 4) zero inflated model with Poisson error distribution, 5) generalized additive model which was assumed that the latitude and longitude were followed by 2 dimentional 3 degrees spline function (GAM) 6) generalized additive mixture model which cruse ID was assumed as random factor and was assumed that the latitude and longitude were followed by 2 dimentional 3 degrees spline function (GAMM), 7) same with GAMM but the spline function of latitude and longitude can be changed by year (GAMM by Y), 8) same with GAMM but the spline gunction of latitude and longitude was changed by month (GAMM by M) and 9) generalized addtive model with zero inflated poisson error distribution which was assumed that the latitude and longitude were followed by 2 dimentional 3 degrees spline function (ZI). All models were compared by BIC. All point estimated values were defined as least squared means. All calculations were conducted on R 3.3.3 (R Development Core Team, 2005) and mgcv library was used to calculate GLMMs, GAM, GAMM and ZI (Wood, 2016) .

Results and Discussions

Model comparison was executed by BIC and GAMM was adopted as a most optimal model (Table 3).

Table 3 BIC by each model

	df	BIC
GLM	20	8180.498
GLMM1	21	7161.582
GLMM2	21	7431.192
ZI	53	7744.268
GAM	46	8019.467
GAMM	22	7107.719
GAMM by Y	37	7115.764
GAMM by M	55	7254.836

On comparison among scaled standardized CPUE by GLM, GLMM1, ZI, GAM, GAMM and nominal CPUE, a set of nominal CPUE and ZI and a set of GLM, GAM and GAMM had similar trends, respectively. GLMM1 had different trend especially in 2012 and 2013 (Fig. 2).

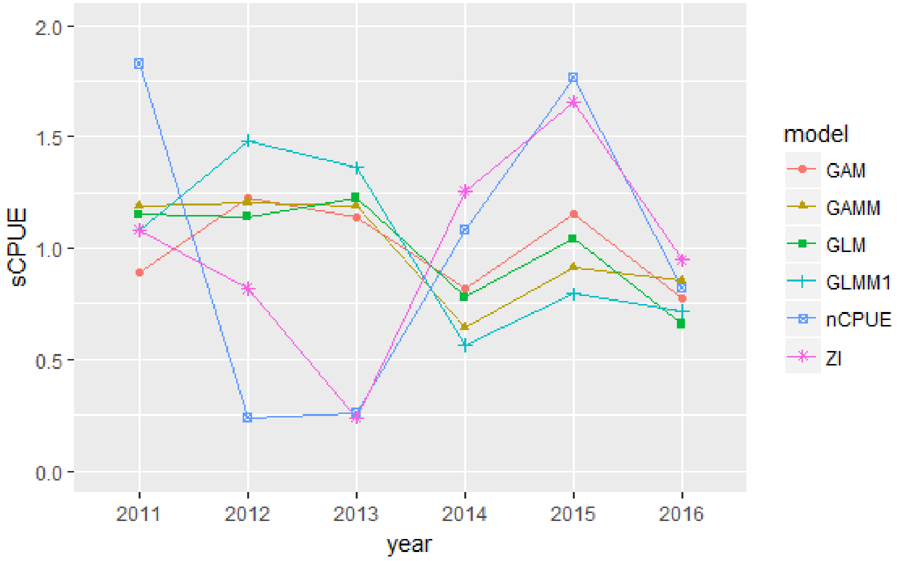


Figure 2. Scaled CPUE estimated by sveral models.

About the optimal model, GAMM, the randomized quantile residuals (Dunn and Smyth, 1996) were calculated and there was no trend by each categorical factors (Fig. 3).

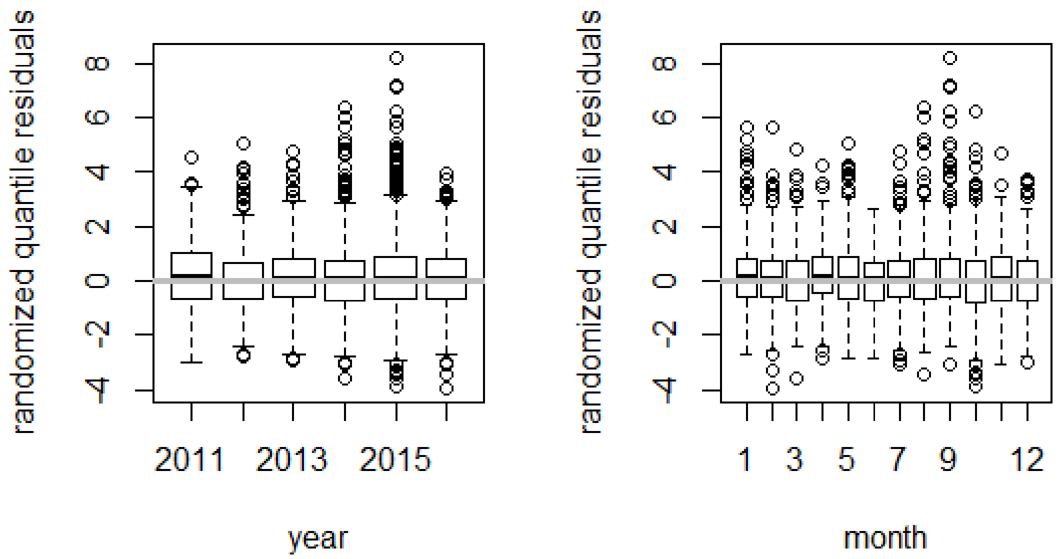


Figure 3. the randomized quantile residuals by GAMM for each categorical factor.

The spatial distribution of estimated CPUE showed higher CPUE on the offshore of Northan Japan (Fig. 4).

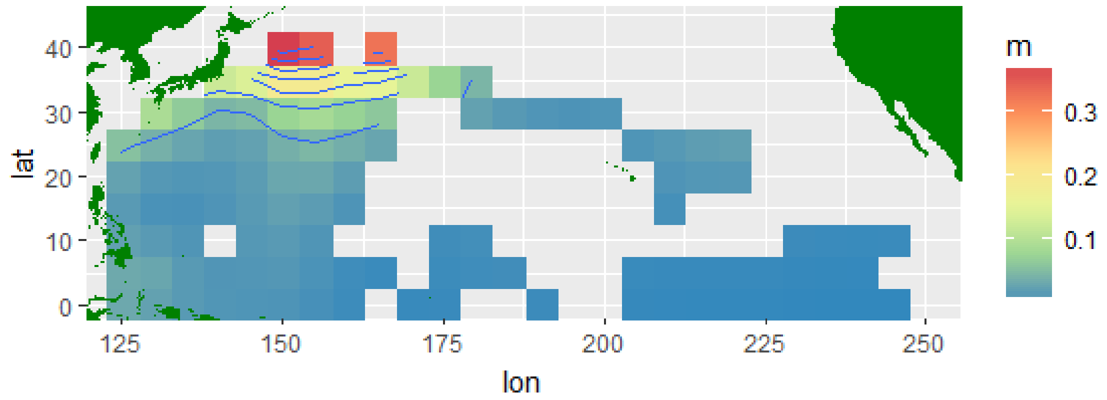


Figure 4. Spatial distribution of estimated CPUE by GAMM.

The estimated annual trend showed basically flat and those 95% confidence interval showed wider range in 2011 and narrower in 2014 (Fig. 5 and Table 4).

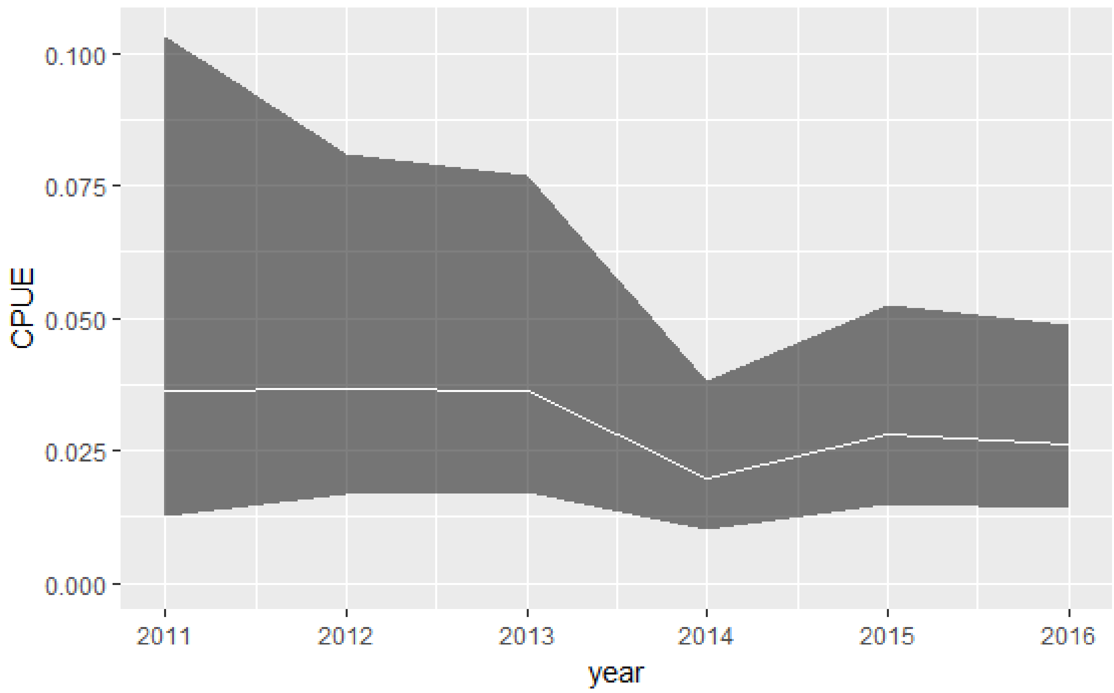


Figure.5 Annual trend with 95% confidence interval estimated by GAMM.

Table 4 nominal CPUE and standardized CPUE and those SE estimated by GAMM

year	nCPUE	sCPUE	se
2011	0.493	0.0364	0.533
2012	0.065	0.0369	0.402
2013	0.070	0.0363	0.385
2014	0.292	0.0197	0.339
2015	0.476	0.0279	0.323
2016	0.222	0.0262	0.318

Refferences

Dunn, P. K. and Smyth, G. K.(1996) Randomized quantile residuals. Journal of Computational and graphical statistics 5 236-244

R Development Core Team (2005). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL http://www.R-project.org.

Wood, S. N. (2016). Just Another Gibbs Additive Modeler: Interfacing JAGS and mgcv. Journal of Statistical Software, 75(7), 1-15. doi:10.18637/jss.v075.i07

Appendix1 summary of mixture part of GAMM

Generalized linear mixed model fit by maximum likelihood (Laplace Approximation) [glmerMod]

Family: poisson (log)

AIC	BIC	logLik	deviance	df.resid
6960.7	7107.7	-3458.4	6916.7	5867

Scaled residuals:

Min	1Q	Median	3Q	Max
-3.3037	-0.3583	-0.1773	-0.1166	20.3752

Random effects:

Groups	Name	Variance	Std.Dev.
cr.id	(Intercept)	0.9436	0.9714
Xr	s(lon,lat)	9.4966	3.0817

Number of obs: 5889, groups: cr.id, 333; Xr, 27

Fixed effects:

	Estimate	Std. Error	z value	Pr(>|z|)	
X(Intercept)	-7.187786	0.453468	-15.851	< 2e-16	***
Xhpb22	-1.941398	0.199754	-9.719	< 2e-16	***
Xyear2012	0.014075	0.488859	0.029	0.977030	
Xyear2013	-0.002973	0.499087	-0.006	0.995247	
Xyear2014	-0.612904	0.451492	-1.358	0.174620	
Xyear2015	-0.266254	0.443242	-0.601	0.548042	
Xyear2016	-0.327583	0.451282	-0.726	0.467903	
Xmonth2	-0.079595	0.121073	-0.657	0.510915	
Xmonth3	-0.614264	0.179609	-3.420	0.000626	***
Xmonth4	-0.534273	0.274890	-1.944	0.051945	.
Xmonth5	-0.245058	0.258802	-0.947	0.343693	
Xmonth6	-0.739093	0.280861	-2.632	0.008500	**
Xmonth7	-1.100747	0.286703	-3.839	0.000123	***
Xmonth8	-1.052117	0.261364	-4.025	5.69e-05	***
Xmonth9	-0.988700	0.243937	-4.053	5.05e-05	***
Xmonth10	-1.083679	0.239012	-4.534	5.79e-06	***

```
Xmonth11           -1.202685     0.243350   -4.942  7.72e-07  ***
Xmonth12           -1.213085     0.262995   -4.613  3.98e-06  ***
Xs(lon,lat)Fx1   0.841965     0.444751     1.893  0.058343  .
Xs(lon,lat)Fx2   0.947761     0.290960     3.257  0.001125  **
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Correlation matrix not shown by default, as p = 20 > 12.
Use print(x, correlation=TRUE)  or
           vcov(x)  if you need it
```

Appendix2 summary of gam part of GAMM

```
Family: poisson
Link function: log

Formula:
c.mako ~ hpb2 + year + month + s(lon, lat) + offset(log(o.hooks))

Parametric coefficients:
                  Estimate  Std. Error  z value  Pr(>|z|)
(Intercept)  -7.187786    0.454730  -15.807   <  2e-16  ***
hpb22           -1.941398    0.199703   -9.721   <  2e-16  ***
year2012         0.014075    0.496482     0.028  0.977383
year2013       -0.002973    0.506068   -0.006  0.995313
year2014       -0.612904    0.457399   -1.340  0.180253
year2015       -0.266254    0.449104   -0.593  0.553278
year2016       -0.327583    0.456735   -0.717  0.473233
month2          -0.079595    0.121115   -0.657  0.511062
month3          -0.614264    0.179711   -3.418  0.000631  ***
month4          -0.534273    0.276551   -1.932  0.053371  .
month5          -0.245058    0.260433   -0.941  0.346724
month6          -0.739093    0.283336   -2.609  0.009093  **
month7          -1.100747    0.288378   -3.817  0.000135  ***
month8          -1.052117    0.262338   -4.011  6.06e-05  ***
month9          -0.988700    0.244264   -4.048  5.17e-05  ***
month10         -1.083679    0.238403   -4.546  5.48e-06  ***
```

month11	-1.202685	0.242934	-4.951	7.40e-07	***
month12	-1.213085	0.265168	-4.575	4.77e-06	***

Signif. codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Approximate significance of smooth terms:

	edf	Ref.df	Chi.sq	p-value	
s(lon,lat)	18.32	18.32	300.9	<2e-16	***

Signif. codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

R-sq.(adj) = 0.382

glmer.ML = 3642.8 Scale est. = 1 n = 5889

Appendix 3 Comparison among scaled CPUEs by using Japanese different data set