NPFC-2020-SSC PS06-WP12

**Standardized CPUE of Pacific saury (*Cololabis saira*) caught**
**by the Japanese stick-held dip net fishery up to 2019**

Midori HASHIMOTO, Miyako NAYA, Shin-Ichiro NAKAYAMA, Taiki FUJI,
Satoshi SUYAMA, and Kazuhiro OSHIMA

*Fisheries Researches Institute, Japan Fisheries Research and Education Agency*

**Summary**
We updated the standardized catch-per-unit-effort (CPUE) of Pacific saury caught by the Japanese stick-held dip net fishery up to 2019 for the stock assessment under the framework of NPFC. CPUE was explained by a generalized linear model (GLM) incorporating explanatory variables such as year, month, fishing area, vessel size, and sea surface temperature. The standardized CPUE in 2019 decreased to the lowest level since 1994.

# Introduction

In the stock assessment on Pacific saury (*Cololabis saira* hereafter PS) under the framework of the North Pacific Fisheries Commission (NPFC), it was assessed that stock biomass fell to the lowest value since 1980 in 2017 and then increased in 2018 (NPFC-2019-SSC PS04-Final Report). Since the latest status of PS stock would be informative for next PS stock assessment, we updated the standardized CPUE of PS by the Japanese stick-held dip net (hereafter SHDN) fishery up to 2019.

# Method

Standardization of CPUE for PS was conducted according to the standardization protocol (Annex E in NPFC-2019-SSC PS05-Final Report) updated in the 5th meeting of Small Scientific Committee on Pacific Saury (see **Appendix I**).

1. Commercial fishery data sources

Catch data of the Japanese SHDN fishery for PS was obtained through the landing survey deployed at major six landing ports. Procedure of the landing survey and types of information in the obtained data were described in the previous working paper (Suyama et al., 2018). The landing survey data includes information on date, fishing position (longitude and latitude), catch in weight (mt), number of hauls, *in situ* sea surface temperature (SST) measured using an on-board thermometer, and size of the fishing vessels (GRT). CPUE was defined as catch in weight per number of hauls in a fishing operation. In 2019, the survey data obtained from 1,519 fishing operations carried out from August to December were used for CPUE standardization and their total catch in weight accounted for 52.7% of Japanese total landing of PS. One record with zero catch was eliminated from data, because fishing operations were basically conducted only when the fish schools were detected. Fishing ground of the Japanese SHDN fishery for PS was divided into five subareas based on oceanographic characteristics (**Fig. 1**). Features of each subarea are described by Suyama et al. (2018). Because fishing ground expanded eastward in 2019, total area of Area V increased.

1. Statistical method
2. Model specification

A generalized linear model (GLM) was used to standardize CPUE. Factors of year, month, fishing area, GRT of fishing vessels and SST were incorporated as explanatory variables. CPUE varied annually and monthly with hitting its peak around October (**Fig. 2**). There observed differences in CPUE among categories for fishing area, vessel size and SST. The correlation matrix for these explanatory variables is shown in **Fig. 3**. Full model was given as:

ln(*CPUE*) = Intercept + *Year* + *Month* + *Area* + *Grt* + *Sst* + two-way interactions + *ε*,

where *Year*, *Month* and *Area* are categorical variables, composed of 26 years (1994–2019), 5 months (August–December) and 5 subareas (I–V), respectively (**Table 1**). Vessel size was divided into 10 (*Grt1*) or 5 (*Grt2*) categories at intervals of 20 or 40 mt, respectively. SST was divided into 12 (*Sst1*) or 5 (*Sst2*) categories at intervals of 1 or 3 °C, respectively. Parameter *ε* denotes an error term with *ε* ~ *N*(0, *σ*2).

In Japanese SHDN fisheries, no fishing operation in December was occurred in some years and spatial allocation of fishing efforts has varied across years (**Fig.4**). Re-stratification was therefore conducted for explanatory variables other than year used for two-way interactions (Month.int, Area.int, Grt.int, and Sst.int) in order to avoid no observation in any stratum (**Table 1**). Order of interaction terms in full model was determined based on their interpretability in terms of habitat suitability and fishing strategy (**Table 2**). For example, interaction between *Year* and *Grt* is essential and was placed in early order, because CPUE for large vessels equipped with larger nets greatly increased along with increase in stock biomass compared to CPUE for small vessels (left panel in **Fig.5**). The other interaction terms with *Grt* were put at the end, because CPUE for large vessels slightly increased in the main fishing season and main fishing ground while CPUE for small vessels did not increase (middle and right panels in **Fig.5**).

1. Model selection and diagnostics

We employed a Bayesian information criterion (BIC) to measure the predictive ability and select the best model. The optimal categorizations regarding vessel size and SST were determined through model selections. For model diagnostics, the percent deviation explained was calculated in addition to Q-Q plot and residual plots.

1. Calculation of standardized CPUE

Time series of standardized CPUE was estimated using the best GLM. We first generated a data that was composed of all combinations of explanatory variables and then predicted annual values of ln(*CPUE*) for area *a* (ln(*CPUE*)*y,a*). Finally annual standardized CPUE were calculated as the area-weighted mean of (CPUE)*y,a*:

*CPUEy* = Σ*a*{ exp(ln(*CPUE*)*y,a*) × (A*a* / ΣA) },

where A*a* indicates an area of area *a*. Coefficient of variation and 95% confidential intervals were calculated by bootstrap resampled residuals with 1000 replications. The standardized CPUE was compared with nominal CPUE (annual mean of CPUE).

# Results and discussion

1. Model selection

After conducting a backward step-wise model selection (**Table 3**), following model was selected as the best GLM:

ln(CPUE) = Intercept + *Year* + *Month* + *Grt1* + *Sst2* + *Year:Month.int* + *Year:Area.int* + *Year:Grt.int* + *Month.int:Area.int* + *Month.int:Grt.int* + *Area.int:Grt.int* + *ε*.

The same model was selected, comparing with the previous best GLM when using data up to 2018 (Hashimoto et al. 2019). BIC value and percent deviance explained were 90,807 and 51.8%, respectively. Analysis of deviance (Type III tests) indicated that all selected explanatory variables were significant at a significant level of <0.05 (**Table 4**). Q-Q plot and residuals distribution indicated residuals were distributed normally around 0, even though long tails were observed at the both ends (**Fig. 6**). Furthermore, there found no tendencies in residuals across years. It is concluded that CPUE were appropriately modeled using the selected explanatory variables.

1. Year trend of standardized CPUE

Annual standardized CPUE derived from the best GLM showed a similar trend with nominal CPUE (**Fig. 7**). An apparent different trend in 2007 were derived by the interaction between Year and Area. The CPUE in the terminal year decreased to the lowest level since 1994.

 **Reference**
Hashimoto M, Naya M, Nakayama S, Fuji T, Suyama S and Oshima K (2019) Standardized CPUE of Pacific saury (*Cololabis saira*) caught by the Japanese stick-held dip net fishery up to 2018. NPFC-2019-SSC PS05-WP06.

Suyama S, Kidokoro H, Naya M, Hashimoto M and Vijai D (2018) Standardization of CPUE data of Pacific saury (*Cololabis saira*) caught by the Japanese stick-held dip net fishery during 1994 to 2017. NPFC-2018-SSC PS03-WP05.

Small Scientific Committee on Pacific Saury. (2019) 4th Meeting Report. NPFC-2019-SSC PS04-Final Report. 48 pp. (Available at www.npfc.int)

Small Scientific Committee on Pacific Saury (2019) 5th Meeting Report. NPFC-2019-SSC PS05-Final Report. 44 pp. (Available at www.npfc.int)

# Table and figures

**Table 1** Summary of explanatory variables in GLM.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Variables | Cases | Number of categories | Detail | Note |
| Year | *Year* | 26 | 1994–2019 |  |
| Month | *Month* | 5 | August–December |  |
|  | *Month.int* | 4 | Aug, Sep, Oct, Nov + Dec | for interaction terms |
| Area | *Area* | 5 | I–V | see **Fig. 1** |
|  | *Area.int* | 3 | I+V, II+IV, III | for interaction terms |
| Vessel tonnage | *Grt1* | 10 | Grt＜20 tons, 20≦Grt＜40, …, 180≦Grt＜200 | at intervals of 20 tons |
|  | *Grt2* | 5 | Grt＜40 tons, 40≦Grt＜80, …, 160≦Grt＜200 | at intervals of 40 tons |
|  | *Grt.int* | 3 | Grt＜80 tons, 80≦Grt＜160, 160≦Grt＜200 | for interaction terms |
| Sea surface temperature | *Sst1* | 12 | Sst＜10°C, 10≦Sst＜11, …, 20≦Sst | at intervals of 1 °C |
|  | *Sst2* | 5 | Sst＜10°C, 10≦Sst＜13, …, 19≦Sst | at intervals of 3 °C |
|  | *Sst.int* | 4 | Sst＜13°C, 13≦Sst＜16, …, 19≦Sst | for interaction terms |

**Table 2** Order based on interpretation of interaction terms.

|  |  |  |
| --- | --- | --- |
| Interaction terms | Order in the full model | Possible interpretation |
| *Year:Month.int* | 1 | Main fishing season differs among years |
| *Year:Area.int* | 2 | Main fish distribution (high density area) and/or main fishing ground (large effort) differs among years |
| *Year:Grt.int* | 3 | Annual catchability differs among vessel sizes |
| *Year:Sst.int* | - | Little difference in suitable water temperature among years |
| *Month.int:Area.int* | 4 | Main fish distribution/fishing ground differs among months |
| *Month.int:Grt.int* | 7 | Little difference in monthly catchability among vessel sizes |
| *Month.int:Sst.int* | 5 | Suitable water temperature differs among months |
| *Area.int:Grt.int* | 8 | Little difference in spatial catchability (e.g., accessibility to fishing ground) among vessel sizes |
| *Area.int:Sst.int* | 6 | Suitable water temperature differs among fishing areas |
| *Grt.int:Sst.int* | - | Little difference in suitable water temperature among vessel sizes |

**Table 3** Results of back-ward model selection.

|  |  |  |  |
| --- | --- | --- | --- |
| No. | Models | BIC | ΔBIC |
| 1 (Full) | Year + Month + Area + Grt1 + Sst2 + Year:Month.int + Year:Area.int + Year:Grt.int + Month.int:Area.int + Month.int:Sst.int + Area.int:Sst.int + Month.int:Grt.int + Area.int:Grt.int | 90852 | 0 |
| 2 | Full model - Area.int:Grt.int | 90879 | 28 |
| 3 | Full model - Month.int:Grt.int | 90855 | 4 |
| 4 | Full model - Area.int:Sst.int | 90811 | -41 |
| 5 | Full model - Area.int:Sst.int - Month.int:Sst.int | 90811 | -41 |
| 6 | Full model - Area.int:Sst.int - Month.int:Sst.int – Month.int:Area.int | 90855 | 3 |
| 7 | Full model - Area.int:Sst.int - Month.int:Sst.int –Year:Grt.int | 90824 | -27 |
| 8 | Full model - Area.int:Sst.int - Month.int:Sst.int –Year:Area.int | 90926 | 74 |
| 9 | Full model - Area.int:Sst.int - Month.int:Sst.int –Year:Month.int | 92108 | 1256 |
| 10 | Full model - Area.int:Sst.int - Month.int:Sst.int - Sst2 | 90916 | 65 |
| 11 | Full model - Area.int:Sst.int - Month.int:Sst.int - Grt1 | 91301 | 449 |
| **12 (Best)** | **Full model - Area.int:Sst.int - Month.int:Sst.int - Area** | **90807** | **-45** |
| 13 | Full model - Area.int:Sst.int - Month.int:Sst.int - Area - Month | 91033 | 181 |
| 14 (Null) | Year | 103552 | 12700 |

**Table 4** Analysis of deviance table (Type III tests) for the best GLM with minimum BIC.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | *SS* | *Df* | *F* | *Pr*(>*F*) | *Signif. codes* |
| Year | 488.2 | 25 | 40.25 | < 2.2e-16 | \*\*\* |
| Month | 114.3 | 1 | 235.65 | < 2.2e-16 | \*\*\* |
| Grt1 | 274.1 | 7 | 80.69 | < 2.2e-16 | \*\*\* |
| Sst2 | 67.1 | 4 | 34.55 | < 2.2e-16 | \*\*\* |
| Year:Month.int | 1086.5 | 75 | 29.86 | < 2.2e-16 | \*\*\* |
| Year:Area.int | 338.4 | 52 | 13.41 | < 2.2e-16 | \*\*\* |
| Year:Grt.int | 266.7 | 50 | 10.99 | < 2.2e-16 | \*\*\* |
| Month.int:Area.int | 47.7 | 6 | 16.40 | < 2.2e-16 | \*\*\* |
| Month.int:Grt.int | 33.4 | 6 | 11.49 | 7.00E-13 | \*\*\* |
| Area.int:Grt.int | 32.1 | 4 | 16.55 | 1.47E-13 | \*\*\* |
| Residuals | 20196.4 | 41624 |  |  |  |

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

**Table 5** Nominal and standardized CPUE of Japanese stick-held dip net fishery for Pacific saury from 1994 to 2019.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Year | Nominal CPUE (metric ton/haul) | Standardized CPUE by GLM | CV (%) | Lower limit of 95% CI | Upper limit of 95% CI |
| 1994 | 5.38 | 3.07 | 4.04 | 2.84 | 3.32 |
| 1995 | 4.41 | 2.16 | 7.56 | 1.85 | 2.50 |
| 1996 | 2.40 | 1.67 | 5.03 | 1.50 | 1.83 |
| 1997 | 4.77 | 3.74 | 14.47 | 2.79 | 4.95 |
| 1998 | 1.44 | 1.07 | 4.11 | 0.98 | 1.16 |
| 1999 | 1.45 | 0.80 | 4.17 | 0.73 | 0.87 |
| 2000 | 2.18 | 1.43 | 4.67 | 1.30 | 1.56 |
| 2001 | 3.18 | 2.12 | 5.88 | 1.89 | 2.38 |
| 2002 | 1.93 | 1.17 | 6.60 | 1.02 | 1.31 |
| 2003 | 3.21 | 2.19 | 4.58 | 2.00 | 2.38 |
| 2004 | 3.65 | 2.61 | 4.05 | 2.41 | 2.82 |
| 2005 | 6.63 | 4.47 | 4.29 | 4.10 | 4.85 |
| 2006 | 6.03 | 4.09 | 4.61 | 3.73 | 4.48 |
| 2007 | 7.81 | 3.89 | 4.54 | 3.57 | 4.26 |
| 2008 | 7.81 | 5.02 | 4.51 | 4.59 | 5.46 |
| 2009 | 4.60 | 3.73 | 4.71 | 3.44 | 4.10 |
| 2010 | 2.73 | 1.55 | 3.69 | 1.44 | 1.67 |
| 2011 | 4.45 | 2.40 | 4.27 | 2.21 | 2.61 |
| 2012 | 3.65 | 2.39 | 4.43 | 2.18 | 2.59 |
| 2013 | 3.04 | 1.44 | 3.94 | 1.33 | 1.55 |
| 2014 | 5.42 | 2.52 | 3.87 | 2.35 | 2.72 |
| 2015 | 2.65 | 1.34 | 4.58 | 1.22 | 1.48 |
| 2016 | 2.82 | 1.52 | 6.07 | 1.36 | 1.72 |
| 2017 | 1.40 | 1.08 | 4.43 | 0.99 | 1.18 |
| 2018 | 2.96 | 1.45 | 3.79 | 1.35 | 1.56 |
| 2019 | 1.39 | 0.68 | 6.34 | 0.60 | 0.77 |



**Fig.1** Area definition applied for CPUE standardization in this study.



**Fig.2** Relationship between CPUE and each factor (Year, Month, Area, vessel size and SST).



**Fig.3** Correlation matrix of used explanatory variables.



**Fig.4** Annual changes in monthly fishing ground of Japanese stick-held dip net fishery for Pacific saury from 1994 to 2019.



**Fig.5** Relation to CPUE for variables in interaction terms.



**Fig.6** Q-Q plot, histogram of residuals and residual plots across years for the best GLM.



**Fig.7** Scaled nominal CPUE and annual scaled standardized CPUE when using catch data up to 2019. Gray zone indicates 95% confidence intervals of standardized CPUE.

**Appendix I** Checklist for the CPUE standardization protocol.

|  |  |  |
| --- | --- | --- |
| (1) | Conduct a thorough literature review to identify key factors (i.e., spatial, temporal, environmental, and fisheries variables) that may influence CPUE values; | Yes (see Suyama et al. (2018)) |
| (2) | Determine temporal and spatial scales for data grouping for CPUE standardization; | Yes (**Table 1**) |
| (3) | Plot spatio-temporal distributions of fishing efforts and catch to evaluate spatio-temporal patterns of fishing effort and catch; | Yes (**Fig. 4**) |
| (4) | Calculate correlation matrix to evaluate correlations between each pair of those variables; | Yes (**Fig. 3**) |
| (5) | Identify potential explanatory variables based on (1)-(4) as well as interaction terms to develop full model for the CPUE standardization; | Yes (**Table 1**) |
| (6) | Fit candidate statistical models to the data (e.g., GLM, GAM, Delta-lognormal GLM, Neural Networks, Regression Trees, Habitat based models, and Statistical habitat based models); | Yes (GLM) |
| (7) | Evaluate the models using methods such as likelihood ratio, AIC/BIC and cross validation; | Yes (BIC) |
| (8) | Evaluate if distributional assumptions are satisfied and if there is a significant spatial/temporal pattern of residuals in CPUE standardization modeling; | Yes (**Fig. 6**) |
| (9) | Extract yearly standardized CPUE and standard error by a method that is able to account for spatial heterogeneity of effort, such as least squares mean or expanded grid. If the model includes area and the size of spatial strata differs or the model includes interactions between time and area, then standardized CPUE should be calculated with area weighting for each time step. Model with interactions between area and season or month requires careful consideration on a case by case basis; | Yes (see **Method** section) |
| (10) | Recommend a time series of yearly standardized CPUE and associated uncertainty; | Yes (**Table 4**) |
| (11) | Plot nominal and standardized CPUEs over time; | Yes (**Fig.7**) |