NPFC-2020-SSC PS06-WP04

# CPUE standardization for the Pacific saury Russian catches in the Northwest Pacific Ocean

Vladimir KULIK1, Oleg KATUGIN1 and Aleksei BAITALIUK1,2

1 *Pacific branch of the Russian Federal Research Institute of Fisheries and Oceanography «VNIRO» («TINRO»)*

2 *Russian Federal Research Institute of Fisheries and Oceanography «VNIRO»*

Russian Vessel Monitoring System (VMS) records changes in position of all vessels, that officially enter the Russian EEZ, but the catch and place of each operation were not provided in daily electronic reports (DER). TINRO has got information about daily catches of Russian vessels in the period of 1994 — 2002 and foreign vessels in the Russian EEZ after 2002. We deleted foreign catches from 2003 to avoid duplication of the other Member’s catches in CPUE estimation.

## Literature review

Geographic range of the Pacific saury (*Cololabis saira*) in the Pacific Ocean extends from Japan eastward to the Gulf of Alaska and further southward down to Mexico, and the species distribution patterns depend on environmental conditions (Parin 1960). The timing, abundance, and geographic distribution of fish aggregations are associated with sea surface temperature (SST) (Huang et al. 2007). In open waters of the Pacific Ocean, saury forms aggregations in the early summer, and these aggregations experience intra-seasonal changes (Baitaliuk et al. 2013). Pacific saury prefers SST from 12 to 18.5°C, with significant monthly variability (Tseng et al. 2011).

Unfortunately, we could not find data from *in situ* measurements of SST and other potentially useful hydrographic measurements in Russian VMS records as well as exact place and volume of each operation.

## Temporal and spatial scales for data grouping for CPUE standardization

The data is divided on the time scale by such factors as Year and Month.

We used a subset of August-November Russian operations from the dataset that was provided to the NPFC as an Attachment to Annex 1 of NPFC-2017-AR Russia. Then we inserted saury catches from August-November Russian operations in 2016–2019 years from DER of Federal government-financed institution "Centre of Fishery Monitoring and Communications". Thus, we have got 35846 daily reports of saury catches in 1994–2019. Many of 207 vessels didn’t report saury catches for several years and their catches were not targeted; therefore, to avoid problems with vessel coefficient estimation we selected top half (104) of vessels, which in fact reported at least 101 days of saury catches. Finally, the selection included 31909 reports made by vessels that took part in saury fishery at least for 2 years.

We didn’t use preselected regions for spatial grouping, because the most intensive fishing activity of Russian vessels occurred inside 1 fishing subzone of the Russian EEZ and interaction term between Year and Month factors captures spatial differences due to temporal migration patterns.

## Spatio-temporal distributions of fishing efforts and catch

Spatio-temporal distributions of fishing efforts and catch are provided annually to the NPFC at monthly level of grouping (*see* Appendix II and III), but not daily as the catches which were used here.

## Correlations between each pair of predictors and response

We don’t have continuous variables as predictors, because all of them – Year, Month and their interaction term as well as vessel identifiers are presented as factors. Nevertheless, we transformed back to numeric values such factors as years and months just for demonstration purpose.

There were no strong correlations, but all of them were highly significant due to huge number of selected observations – 31909 daily reports (Fig. 1).

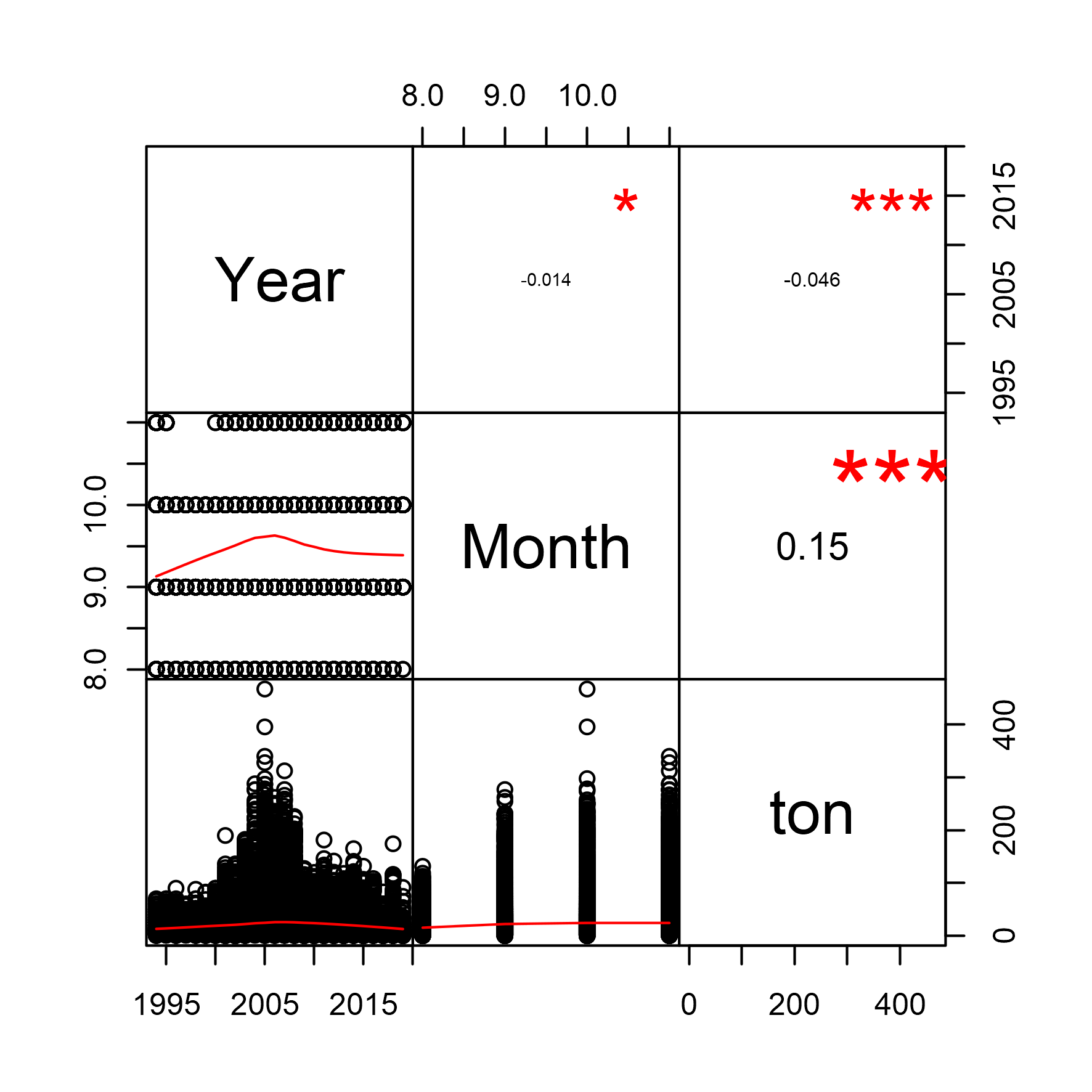


Figure 1 – Pearson’s correlations, where “\*\*\*” indicates *p* < 0.001, “\*\*” – *p* < 0.01,   
“\*” – *p* < 0.05 and “∙” – *p* < 0.1

## Potential explanatory variables based on (1) – (4) to develop full model for the CPUE standardization

According to published information (see Literature review), months could be good predictors. Therefore, the full model may include them. We don’t have officially reported values for SST as well as other possible predictors, so we’ll use only those additional factors, that we can find in our dataset – Month and vessel unique identifiers (Idves). So, the full model will include Month and Idves as additional factors. Year to year difference in the pattern of spatial distribution by Month could be captured by interaction term of Month given Year, therefore it will be also included in the full model.

## Statistical assumptions in the full models

We didn’t include 0 catches, because such “catches” are just DER about positions of vessels during non-fishing operations. So, catches can be positive only and we’ll need a logarithmic link. We checked overdispersion using optimization of power parameter (*p*) in Tweedie family in mgcv package for the full model and found out that it is very close to the possible boundary of 2 (*p* = 1.99). It means that Gamma family may be the best candidate (Wood 2011), because Compound Poisson-Gamma model, which is a member of the Tweedie family, is approximately equal to Gamma model when *p = 2.*

## Select and evaluate the models

We used Schwarz's Bayesian criterion (BIC). BIC is stricter than AIC (Akaike 1974) to the number of parameters in the fitted model. If AIC is calculated as -2‧log-likelihood + *k*‧*d*, where *d* represents the number of parameters in the fitted model, and *k* = 2 for the usual AIC, then for BIC *k* = log(*n*), where *n* is the number of observations. Thus, it can help us to avoid overfitting.

All models were tuned in mgcv package using maximum likelihood for selection based on BIC. We used Generalized linear models, or GLMs. Common part of GLMs, which were used, can be expressed as follows:

where — is the link function (natural logarithm here), which establishes the connection between the linear predictor, *η*, and the mean of the distribution, *µ*, in such way, that the inverse of link function equals to the expectation *E* of catches *Y* given the group of observations (*t*) from catches (*y*) in tons per day distributed according to the variance function with scale parameter . The variance function was from Gamma (Г) exponential family. Therefore, GLMs distinguished only by linear predictor and scale (Table 1). The best GLM contains linear predictor No 4, because it has the lowest BIC and highest explained deviance (see Table 1). Final estimates of the coefficients for GLM No 4 were found using restricted maximum likelihood as it is recommended (Wood 2011).

Table 1 – Statistical properties of converged GLMs by linear predictors

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| No |  |  | Adj. R2 | Dev. expl.  % | BIC | *df* |
| 1 |  | 0.86 | 0.12 | 13.2 | 275588.1 | 27 |
| 2 |  | 0.84 | 0.13 | 14.9 | 274906.0 | 30 |
| 3 |  | 0.80 | 0.16 | 18.6 | 274036.1 | 101 |
| 4 |  | 0.73 | 0.20 | 23.3 | **272961.8** | 204 |
| 5 |  | 0.77 | 0.18 | 19.8 | 273823.5 | 133 |

*β*0 – intercept, – coefficient of *i*-th year (*yeari*), – coefficient of *i*-th month (*monthi*), – coefficient of *i*-th unique ID of a vessel (*Idvesi*).

## Evaluate if distributional assumptions are satisfied and if there is a consistent spatial/temporal distribution of residuals in CPUE standardization modeling

Gamma distribution suited very well to capture overdispersion and we don’t see patterns in residuals (Fig. 2). The rank of the final model is less than 1 (203/207) therefore huge number of parameters do not make our model saturated one.

The same model with Gaussian distribution (and log link) had higher BIC (306251.4) and explained less deviance (21.0%) therefore we will continue to use Gamma distribution in the selected GLM.

Изображение выглядит как снимок экрана

Автоматически созданное описание

Figure 2 – Residuals of GLM No 4 versus linear predictor

Histogram and quantile-quantile Plot show that normality assumption about residuals is not violated (Fig. 3). Deviance residuals overlapped by inter-quartile ranges (IQR) by years (Fig. 4).

Изображение выглядит как карта

Автоматически созданное описаниеИзображение выглядит как карта

Автоматически созданное описание

Figure 3 – Quantile-quantile plot (left) conditional on the fitted model coefficients and scale parameter with 1000 simulated quantiles of the residual distribution and histogram (right) of residuals from GLM No 4 on a link scale

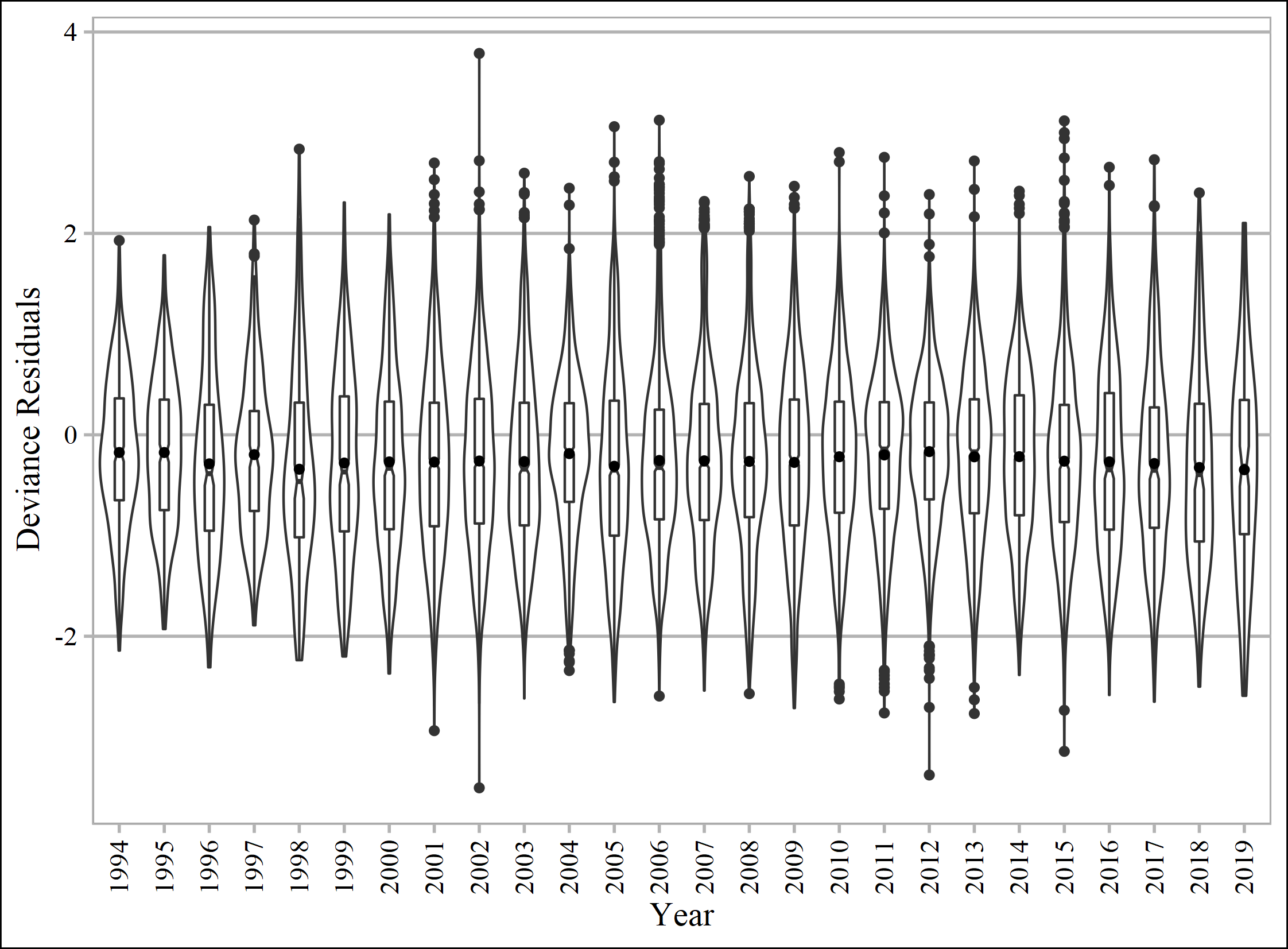


Figure 4 – Distribution of deviance residuals by years, where notched boxes show IQR and dots inside them show the average and outside show the values higher than 1.5\*IQR

Residual autocorrelation was checked by generalized Durbin-Watson statistics and their bootstrapped p-values were higher than 0.08 by years at lags from 1 to 5 years.

Deviance residuals overlapped by inter-quartile ranges (IQR) by years given month also (Fig. 5).

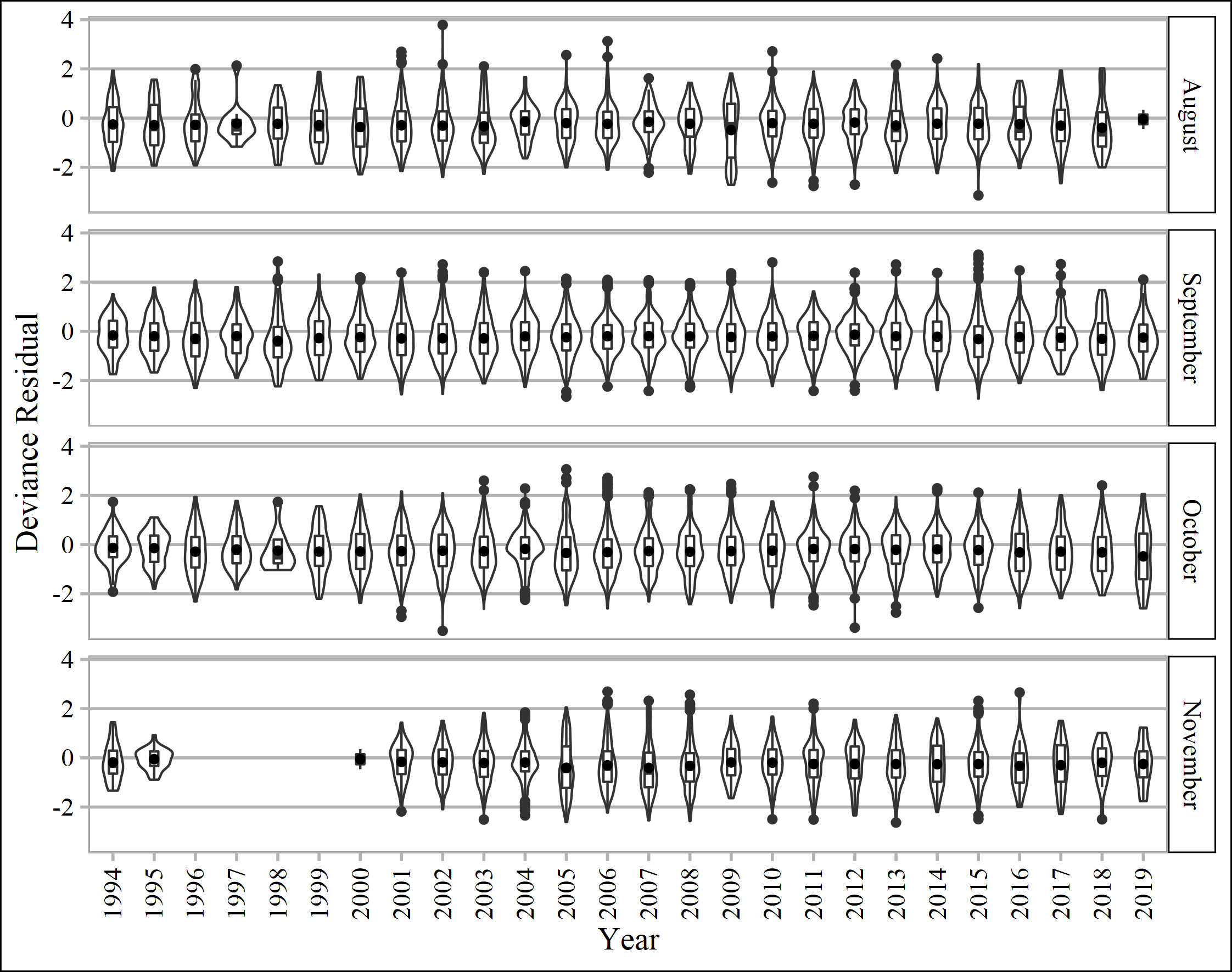


Figure 5 – Distribution of deviance residuals by years given month, where boxes show IQRs and dots inside them show the average and outside show the values higher than 1.5\*IQR

The highest influence (Bentley et al., 2012) was found for the interaction term and Idves in 2019 (Fig. 6). The most part of the selected vessels didn’t take part in saury fishing in 2019, therefore the least of them influenced strongly on the Year coefficient estimation (Fig. 7).



Figure 6 – Influence of interaction term () and Idves



Figure 7 – Influence measured for each Idves by Year (circles)

## The optimal model to estimate yearly standardized CPUE

We could continue to increase the number of predictors, e.g. to include power of vessels. But many parameters, describing vessels that could be used for further including into the model are directly linked with each vessel (or its Idves). So, we decided to stop including other covariates to avoid overfitting. Therefore, our optimal model is GLM No 4.

## Nominal and standardized CPUE over time

Interaction term of factor Month given factor Year complicates the use of Year’s coefficients as indices of abundance in GLM No 4. To overcome this difficulty, we expanded a grid which included all used levels of months (from August to November), years (1994-2019), unique identifiers of vessels (104 levels). Then we predicted catches using GLM No 4 and summarized them (Fig. 8).

Изображение выглядит как длинный, белый, сидит, висит

Автоматически созданное описание

Figure 8 – Violin and box plots for catches per day (CPUEi) predicted from GLM No 4, where blue line connects means of predictions, red line connects means of raw catch values (tons per day) used for standardization, while green line connects trimmed means of raw catch values

Summary statistics of predicted values is given below by years (Table 2) as well as statistics of raw data used for standardization (Table 3) and the same tables described from data transformed with natural logarithm (Tables 4 and 5 respectively).

Table 2 – Summary statistics of predicted 416 values (104 Idves\*4 Month) for each year

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Year | Mean | SD | SE | Median | Trimmed mean by  (by 10%) | MAD | Min | Max | Skew | Kurtosis |
| 1994 | 17.7 | 7.06 | 0.35 | 17.0 | 17.2 | 7.94 | 5.1 | 42.3 | 0.57 | -0.05 |
| 1995 | 20.4 | 7.17 | 0.35 | 19.7 | 20.0 | 7.63 | 6.6 | 45.4 | 0.52 | -0.03 |
| 1996 | 15.0 | 4.29 | 0.21 | 14.6 | 14.7 | 3.99 | 6.7 | 31.5 | 0.71 | 0.60 |
| 1997 | 11.3 | 2.75 | 0.14 | 11.1 | 11.2 | 2.65 | 5.5 | 20.6 | 0.39 | 0.06 |
| 1998 | 12.4 | 4.42 | 0.22 | 12.6 | 12.3 | 5.01 | 3.6 | 25.8 | 0.18 | -0.53 |
| 1999 | 11.2 | 5.08 | 0.25 | 9.7 | 10.8 | 5.38 | 3.6 | 26.7 | 0.60 | -0.60 |
| 2000 | 16.1 | 6.09 | 0.30 | 16.3 | 16.0 | 6.95 | 4.4 | 33.3 | 0.18 | -0.58 |
| 2001 | 20.9 | 7.12 | 0.35 | 21.0 | 20.6 | 8.05 | 6.5 | 43.3 | 0.29 | -0.39 |
| 2002 | 19.6 | 12.87 | 0.63 | 16.8 | 18.6 | 15.85 | 2.6 | 59.2 | 0.54 | -0.66 |
| 2003 | 28.5 | 11.55 | 0.57 | 27.1 | 27.4 | 11.31 | 8.9 | 70.8 | 0.82 | 0.36 |
| 2004 | 43.9 | 17.07 | 0.84 | 42.2 | 43.0 | 19.29 | 12.6 | 101.2 | 0.47 | -0.23 |
| 2005 | 46.4 | 20.24 | 0.99 | 45.2 | 45.2 | 24.32 | 11.8 | 112.6 | 0.47 | -0.37 |
| 2006 | 32.3 | 13.19 | 0.65 | 31.1 | 31.4 | 14.45 | 9.0 | 78.4 | 0.59 | 0.00 |
| 2007 | 40.7 | 11.51 | 0.56 | 40.0 | 40.2 | 12.63 | 15.7 | 79.6 | 0.42 | -0.03 |
| 2008 | 41.0 | 10.35 | 0.51 | 40.4 | 40.5 | 9.96 | 18.8 | 77.9 | 0.45 | 0.17 |
| 2009 | 20.6 | 5.67 | 0.28 | 20.3 | 20.3 | 5.84 | 8.5 | 39.1 | 0.47 | 0.07 |
| 2010 | 22.3 | 8.94 | 0.44 | 20.2 | 21.5 | 8.39 | 7.6 | 54.4 | 0.80 | 0.17 |
| 2011 | 26.1 | 8.79 | 0.43 | 26.4 | 25.9 | 10.07 | 8.1 | 52.9 | 0.24 | -0.44 |
| 2012 | 23.2 | 7.01 | 0.34 | 22.6 | 22.8 | 7.38 | 8.9 | 47.9 | 0.56 | 0.12 |
| 2013 | 20.9 | 8.08 | 0.40 | 20.1 | 20.4 | 8.54 | 6.3 | 49.8 | 0.62 | 0.11 |
| 2014 | 23.8 | 6.19 | 0.30 | 23.1 | 23.5 | 5.80 | 11.2 | 47.0 | 0.54 | 0.36 |
| 2015 | 15.3 | 6.64 | 0.33 | 13.4 | 14.5 | 5.17 | 5.5 | 40.6 | 1.14 | 0.89 |
| 2016 | 15.9 | 6.86 | 0.34 | 14.9 | 15.3 | 7.11 | 4.6 | 40.2 | 0.76 | 0.11 |
| 2017 | 8.3 | 2.07 | 0.10 | 8.2 | 8.2 | 2.04 | 3.8 | 15.3 | 0.42 | 0.10 |
| 2018 | 21.0 | 10.74 | 0.53 | 17.1 | 19.4 | 6.64 | 7.2 | 61.6 | 1.31 | 1.06 |
| 2019 | 6.6 | 4.50 | 0.22 | 5.5 | 6.2 | 5.45 | 0.8 | 20.5 | 0.58 | -0.61 |

Table 3 – Summary statistics of raw catch (metric tons per day per vessel) values used for CPUE standardization

|  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Year | n | Mean | SD | SE | Median | Trimmed mean by  10% | MAD | Min | Max | Skew | Kurtosis |
| 1994 | 437 | 18.7 | 13.54 | 0.65 | 15.0 | 17.1 | 13.34 | 0.5 | 70.0 | 1.02 | 0.69 |
| 1995 | 389 | 22.3 | 15.65 | 0.79 | 20.0 | 20.7 | 14.83 | 1.0 | 70.0 | 0.79 | -0.13 |
| 1996 | 284 | 17.5 | 16.33 | 0.97 | 12.0 | 15.0 | 11.86 | 0.5 | 90.0 | 1.37 | 1.47 |
| 1997 | 211 | 12.6 | 10.47 | 0.72 | 10.0 | 10.9 | 7.41 | 1.0 | 60.3 | 1.82 | 4.02 |
| 1998 | 154 | 14.5 | 15.72 | 1.27 | 10.0 | 11.4 | 8.90 | 0.5 | 88.0 | 2.05 | 4.57 |
| 1999 | 277 | 14.9 | 14.13 | 0.85 | 10.0 | 12.8 | 10.38 | 0.5 | 82.7 | 1.66 | 3.59 |
| 2000 | 711 | 18.3 | 16.95 | 0.64 | 12.4 | 15.4 | 11.27 | 0.5 | 90.0 | 1.63 | 2.74 |
| 2001 | 1556 | 22.2 | 21.49 | 0.55 | 15.3 | 18.6 | 15.20 | 0.1 | 190.0 | 2.03 | 6.07 |
| 2002 | 2008 | 19.6 | 20.85 | 0.47 | 12.4 | 15.7 | 13.12 | 0.0 | 136.1 | 2.01 | 4.92 |
| 2003 | 1744 | 31.2 | 31.46 | 0.75 | 20.0 | 25.4 | 20.31 | 0.4 | 182.5 | 1.88 | 3.94 |
| 2004 | 1693 | 49.3 | 38.69 | 0.94 | 42.0 | 44.3 | 32.32 | 0.8 | 288.0 | 1.90 | 6.14 |
| 2005 | 1603 | 52.4 | 58.68 | 1.47 | 30.5 | 40.3 | 30.38 | 0.3 | 467.5 | 2.12 | 5.25 |
| 2006 | 1842 | 39.1 | 41.35 | 0.96 | 25.0 | 31.1 | 22.54 | 0.4 | 262.0 | 2.41 | 6.95 |
| 2007 | 2268 | 44.6 | 44.02 | 0.92 | 30.3 | 36.2 | 28.91 | 0.5 | 312.5 | 2.08 | 5.03 |
| 2008 | 2217 | 41.8 | 39.13 | 0.83 | 29.4 | 35.1 | 27.28 | 0.3 | 226.0 | 1.85 | 3.94 |
| 2009 | 1618 | 22.0 | 20.35 | 0.51 | 16.2 | 18.6 | 15.05 | 0.2 | 127.4 | 1.73 | 3.53 |
| 2010 | 1428 | 22.0 | 19.13 | 0.51 | 16.6 | 19.0 | 14.78 | 0.1 | 120.0 | 1.54 | 2.71 |
| 2011 | 2091 | 28.8 | 22.15 | 0.48 | 24.1 | 26.1 | 20.35 | 0.2 | 181.4 | 1.39 | 2.99 |
| 2012 | 2329 | 25.7 | 18.62 | 0.39 | 21.9 | 23.4 | 16.23 | 0.0 | 141.5 | 1.37 | 2.65 |
| 2013 | 2127 | 24.2 | 19.93 | 0.43 | 19.2 | 21.4 | 17.42 | 0.2 | 118.1 | 1.25 | 1.40 |
| 2014 | 2414 | 26.2 | 21.24 | 0.43 | 20.4 | 23.4 | 18.46 | 0.4 | 165.0 | 1.36 | 2.52 |
| 2015 | 1133 | 16.6 | 16.90 | 0.50 | 11.1 | 13.6 | 11.09 | 0.0 | 131.7 | 2.15 | 6.28 |
| 2016 | 601 | 21.6 | 21.78 | 0.89 | 13.7 | 17.5 | 14.06 | 0.4 | 108.3 | 1.74 | 2.91 |
| 2017 | 447 | 10.5 | 11.09 | 0.52 | 6.7 | 8.4 | 6.65 | 0.1 | 83.0 | 2.30 | 7.23 |
| 2018 | 216 | 24.5 | 30.06 | 2.05 | 11.3 | 18.4 | 12.20 | 0.3 | 174.3 | 1.89 | 3.62 |
| 2019 | 111 | 12.0 | 15.81 | 1.50 | 5.1 | 8.7 | 5.75 | 0.1 | 91.5 | 2.49 | 7.34 |

Table 4 – Summary statistics of predicted 416 values for each year on a link (Log) scale

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Year | Mean | SD | SE | Median | Trimmed mean by  (by 10%) | MAD | Min | Max | Skew | Kurtosis |
| 1994 | 2.79 | 0.02 | 0.42 | 2.83 | 2.81 | 0.44 | 1.63 | 3.74 | -0.30 | -0.53 |
| 1995 | 2.95 | 0.02 | 0.37 | 2.98 | 2.96 | 0.39 | 1.89 | 3.82 | -0.29 | -0.40 |
| 1996 | 2.67 | 0.01 | 0.29 | 2.68 | 2.67 | 0.28 | 1.90 | 3.45 | -0.06 | -0.27 |
| 1997 | 2.39 | 0.01 | 0.25 | 2.40 | 2.40 | 0.24 | 1.70 | 3.02 | -0.24 | -0.31 |
| 1998 | 2.45 | 0.02 | 0.40 | 2.54 | 2.47 | 0.40 | 1.27 | 3.25 | -0.58 | -0.34 |
| 1999 | 2.31 | 0.02 | 0.47 | 2.27 | 2.31 | 0.59 | 1.27 | 3.29 | -0.03 | -1.08 |
| 2000 | 2.70 | 0.02 | 0.42 | 2.79 | 2.73 | 0.43 | 1.47 | 3.51 | -0.59 | -0.42 |
| 2001 | 2.98 | 0.02 | 0.37 | 3.04 | 3.00 | 0.39 | 1.87 | 3.77 | -0.48 | -0.33 |
| 2002 | 2.71 | 0.04 | 0.78 | 2.82 | 2.75 | 0.93 | 0.97 | 4.08 | -0.36 | -1.10 |
| 2003 | 3.27 | 0.02 | 0.40 | 3.30 | 3.27 | 0.44 | 2.19 | 4.26 | -0.05 | -0.55 |
| 2004 | 3.70 | 0.02 | 0.41 | 3.74 | 3.72 | 0.45 | 2.53 | 4.62 | -0.36 | -0.52 |
| 2005 | 3.73 | 0.02 | 0.47 | 3.81 | 3.75 | 0.51 | 2.47 | 4.72 | -0.37 | -0.67 |
| 2006 | 3.39 | 0.02 | 0.43 | 3.44 | 3.40 | 0.46 | 2.20 | 4.36 | -0.30 | -0.53 |
| 2007 | 3.66 | 0.01 | 0.29 | 3.69 | 3.67 | 0.30 | 2.76 | 4.38 | -0.30 | -0.25 |
| 2008 | 3.68 | 0.01 | 0.26 | 3.70 | 3.69 | 0.25 | 2.94 | 4.36 | -0.22 | -0.29 |
| 2009 | 2.99 | 0.01 | 0.28 | 3.01 | 2.99 | 0.30 | 2.14 | 3.67 | -0.23 | -0.29 |
| 2010 | 3.03 | 0.02 | 0.40 | 3.00 | 3.03 | 0.46 | 2.02 | 4.00 | 0.02 | -0.68 |
| 2011 | 3.20 | 0.02 | 0.36 | 3.28 | 3.22 | 0.38 | 2.09 | 3.97 | -0.51 | -0.31 |
| 2012 | 3.10 | 0.02 | 0.31 | 3.12 | 3.11 | 0.33 | 2.19 | 3.87 | -0.19 | -0.34 |
| 2013 | 2.97 | 0.02 | 0.40 | 3.00 | 2.98 | 0.43 | 1.84 | 3.91 | -0.25 | -0.47 |
| 2014 | 3.14 | 0.01 | 0.26 | 3.14 | 3.14 | 0.25 | 2.42 | 3.85 | -0.16 | -0.27 |
| 2015 | 2.64 | 0.02 | 0.41 | 2.60 | 2.63 | 0.41 | 1.71 | 3.70 | 0.30 | -0.52 |
| 2016 | 2.67 | 0.02 | 0.44 | 2.70 | 2.68 | 0.50 | 1.53 | 3.69 | -0.10 | -0.67 |
| 2017 | 2.08 | 0.01 | 0.26 | 2.11 | 2.09 | 0.26 | 1.35 | 2.73 | -0.23 | -0.30 |
| 2018 | 2.93 | 0.02 | 0.46 | 2.84 | 2.91 | 0.41 | 1.98 | 4.12 | 0.50 | -0.51 |
| 2019 | 1.59 | 0.04 | 0.83 | 1.71 | 1.63 | 0.97 | -0.23 | 3.02 | -0.36 | -1.11 |

Table 5 – Summary statistics of raw catch values used for CPUE standardization on a log scale

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Year | Mean | SD | SE | Median | Trimmed mean by  10% | MAD | Min | Max | Skew | Kurtosis |
| 1994 | 2.61 | 0.90 | 0.04 | 2.71 | 2.68 | 0.76 | -0.69 | 4.25 | -0.84 | 0.76 |
| 1995 | 2.80 | 0.86 | 0.04 | 3.00 | 2.87 | 1.03 | 0.00 | 4.25 | -0.64 | -0.04 |
| 1996 | 2.40 | 1.05 | 0.06 | 2.49 | 2.44 | 1.12 | -0.69 | 4.50 | -0.39 | -0.28 |
| 1997 | 2.22 | 0.82 | 0.06 | 2.30 | 2.24 | 0.94 | 0.00 | 4.10 | -0.17 | -0.37 |
| 1998 | 2.14 | 1.11 | 0.09 | 2.30 | 2.18 | 1.03 | -0.69 | 4.48 | -0.29 | -0.32 |
| 1999 | 2.21 | 1.10 | 0.07 | 2.30 | 2.27 | 1.10 | -0.69 | 4.42 | -0.46 | -0.46 |
| 2000 | 2.46 | 1.02 | 0.04 | 2.52 | 2.51 | 1.08 | -0.69 | 4.50 | -0.42 | -0.19 |
| 2001 | 2.63 | 1.07 | 0.03 | 2.73 | 2.69 | 1.07 | -2.30 | 5.25 | -0.56 | 0.32 |
| 2002 | 2.40 | 1.18 | 0.03 | 2.51 | 2.45 | 1.23 | -3.91 | 4.91 | -0.49 | 0.20 |
| 2003 | 2.95 | 1.07 | 0.03 | 3.00 | 2.98 | 1.14 | -1.05 | 5.21 | -0.30 | -0.30 |
| 2004 | 3.57 | 0.91 | 0.02 | 3.74 | 3.64 | 0.80 | -0.22 | 5.66 | -0.79 | 0.59 |
| 2005 | 3.38 | 1.17 | 0.03 | 3.42 | 3.42 | 1.10 | -1.21 | 6.15 | -0.39 | 0.02 |
| 2006 | 3.20 | 1.02 | 0.02 | 3.22 | 3.23 | 1.03 | -1.04 | 5.57 | -0.30 | 0.10 |
| 2007 | 3.35 | 1.02 | 0.02 | 3.41 | 3.39 | 1.00 | -0.69 | 5.75 | -0.44 | 0.19 |
| 2008 | 3.27 | 1.09 | 0.02 | 3.38 | 3.35 | 0.95 | -1.10 | 5.42 | -0.75 | 0.55 |
| 2009 | 2.62 | 1.09 | 0.03 | 2.78 | 2.71 | 1.01 | -1.47 | 4.85 | -0.75 | 0.58 |
| 2010 | 2.67 | 1.03 | 0.03 | 2.81 | 2.74 | 0.94 | -2.21 | 4.79 | -0.75 | 0.70 |
| 2011 | 3.01 | 0.96 | 0.02 | 3.18 | 3.10 | 0.86 | -1.66 | 5.20 | -0.98 | 1.28 |
| 2012 | 2.95 | 0.86 | 0.02 | 3.09 | 3.02 | 0.78 | -3.35 | 4.95 | -0.95 | 2.09 |
| 2013 | 2.78 | 1.02 | 0.02 | 2.96 | 2.86 | 0.99 | -1.61 | 4.77 | -0.76 | 0.49 |
| 2014 | 2.88 | 0.98 | 0.02 | 3.02 | 2.96 | 0.98 | -0.91 | 5.11 | -0.69 | 0.18 |
| 2015 | 2.31 | 1.11 | 0.03 | 2.41 | 2.36 | 1.10 | -3.96 | 4.88 | -0.64 | 0.97 |
| 2016 | 2.58 | 1.07 | 0.04 | 2.62 | 2.61 | 1.13 | -0.87 | 4.69 | -0.28 | -0.33 |
| 2017 | 1.83 | 1.10 | 0.05 | 1.90 | 1.88 | 1.08 | -2.47 | 4.42 | -0.43 | 0.16 |
| 2018 | 2.44 | 1.35 | 0.09 | 2.43 | 2.49 | 1.37 | -1.34 | 5.16 | -0.29 | -0.31 |
| 2019 | 1.68 | 1.42 | 0.14 | 1.63 | 1.77 | 1.39 | -2.51 | 4.52 | -0.50 | 0.11 |

REFERENCES

Akaike, H. (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, *19*(6), 716–723. https://doi.org/10.1109/TAC.1974.1100705

Baitaliuk, A. A., Orlov, A. M., & Ermakov, Y. K. (2013). Characteristic features of ecology of the Pacific saury Cololabis saira (Scomberesocidae, Beloniformes) in open waters and in the northeast Pacific ocean. *Journal of Ichthyology*, *53*(11), 899–913. https://doi.org/10.1134/S0032945213110027

Bentley, N., Kendrick, T. H., Starr, P. J., & Breen, P. A. (2012). Influence plots and metrics: tools for better understanding fisheries catch-per-unit-effort standardizations. *ICES Journal of Marine Science*, *69*(1), 84–88. https://doi.org/10.1093/icesjms/fsr174

Huang, W.-B., Lo, N. C. H., Chiu, T.-S., & Chen, C.-S. (2007). Geographical Distribution and Abundance of Pacific Saury, Cololabis saira (Brevoort) (Scomberesocidae), Fishing Stocks in the Northwestern Pacific in Relation to Sea Temperatures. *Zoological Studies*, *46*(6), 705–716. Retrieved from http://zoolstud.sinica.edu.tw/Journals/46.6/705.pdf

Parin, N. V. 1960. The range of the saury (Cololabis saira Brev.-Scombresocidae, Pices) and effects of oceanographic features on its distribution. *Proc. Acad. Sci. USSR*, *130*(3), 649–652.

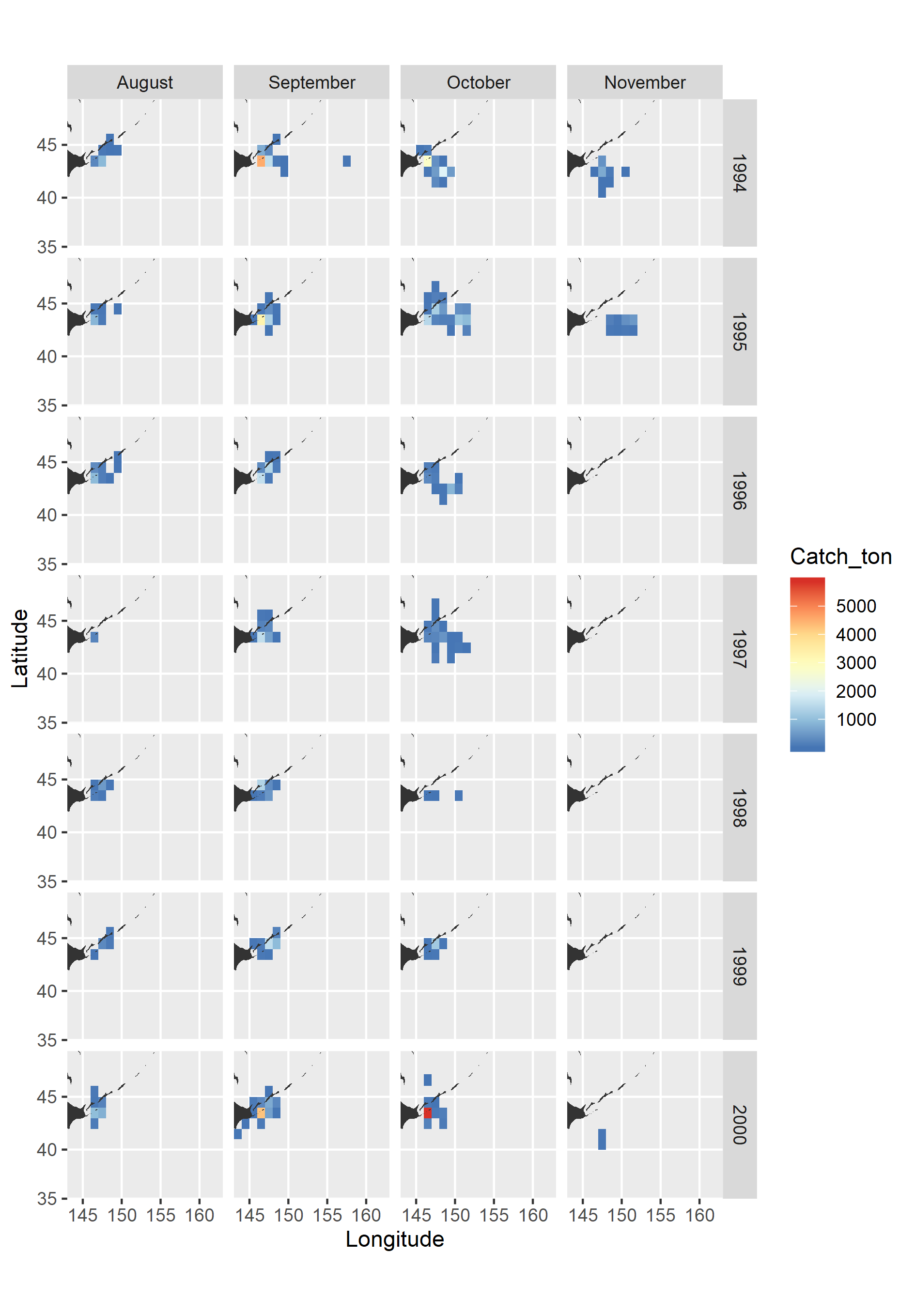
Tseng, C. Te, Sun, C. L., Yeh, S. Z., Chen, S. C., Su, W. C., & Liu, D. C. (2011). Influence of climate-driven sea surface temperature increase on potential habitats of the Pacific saury (Cololabis saira). *ICES Journal of Marine Science*. https://doi.org/10.1093/icesjms/fsr070

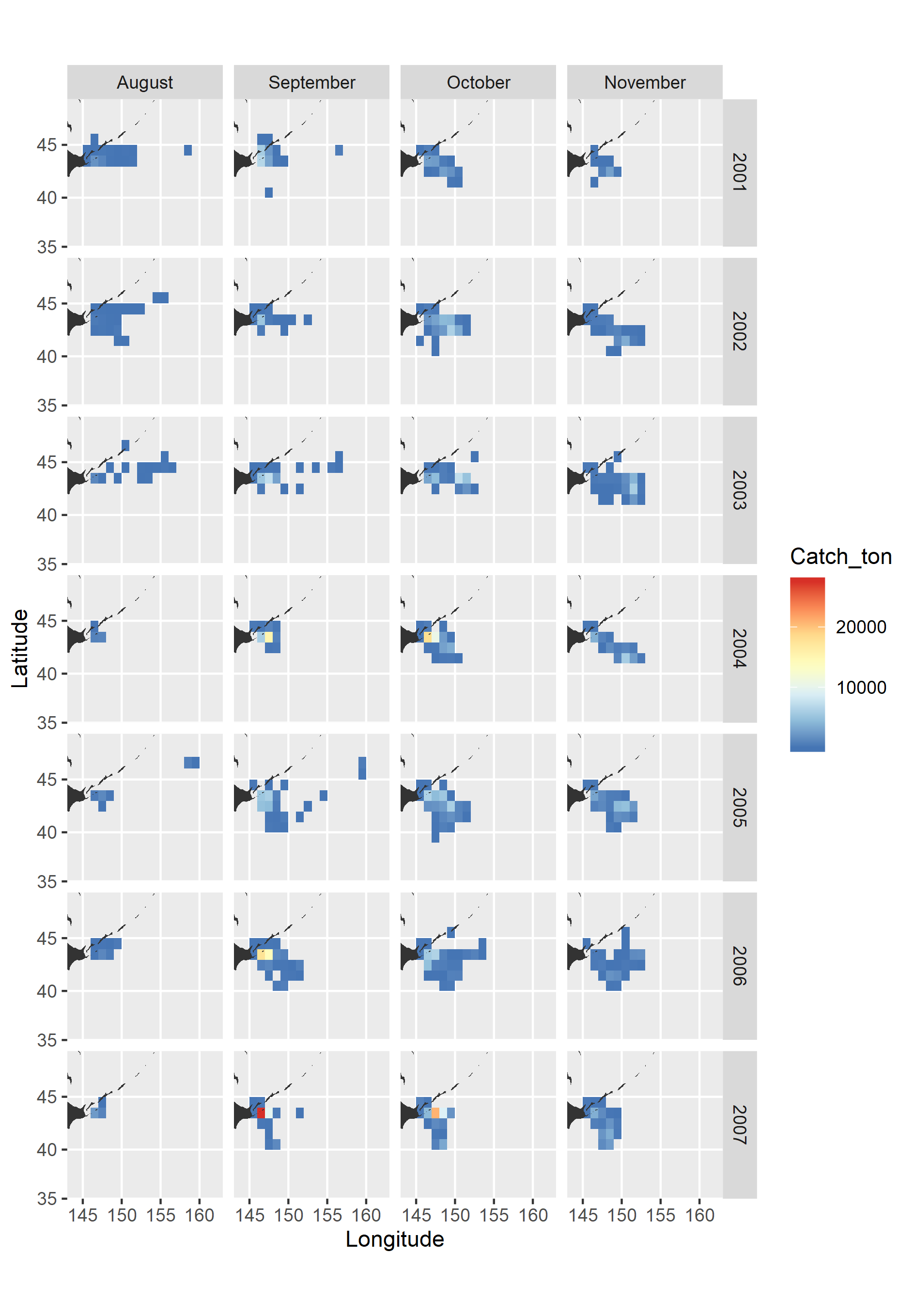
Wood, S. N. (2011). Fast stable restricted maximum likelihood and marginal likelihood estimation of semiparametric generalized linear models. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, *73*(1), 3–36. https://doi.org/10.1111/j.1467-9868.2010.00749.x

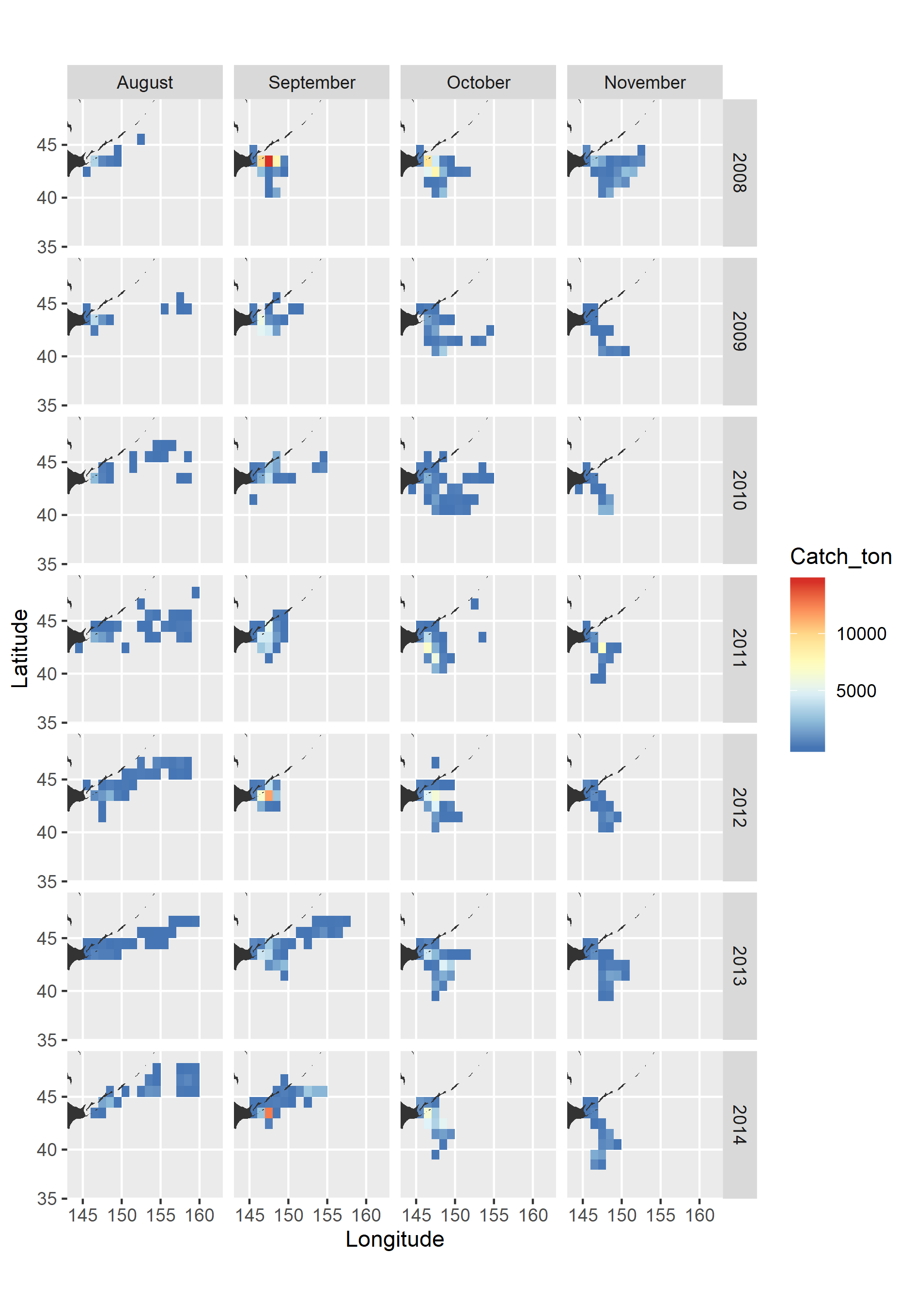
# Appendix I Checklist for the CPUE standardization protocol.

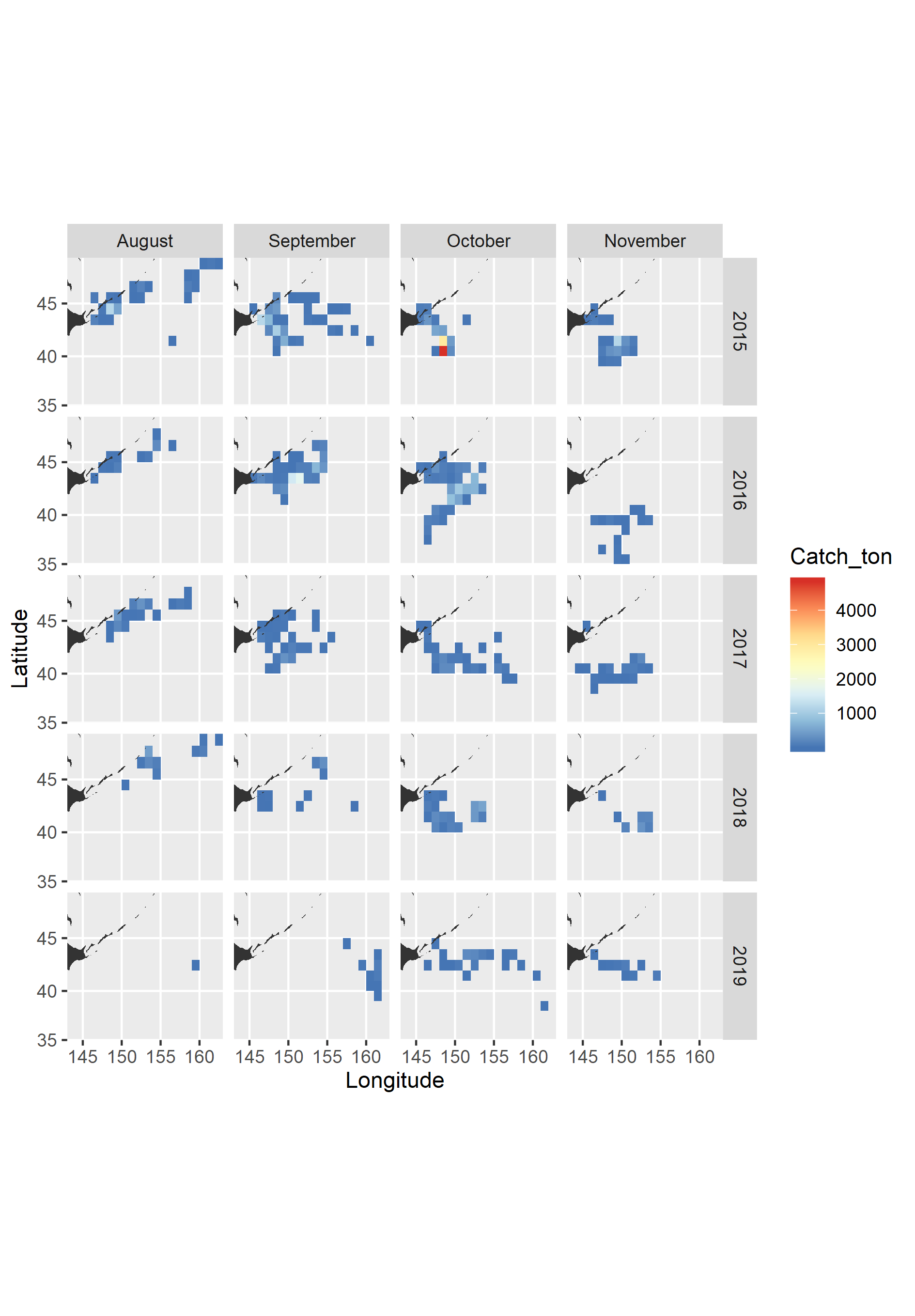
|  |  |  |
| --- | --- | --- |
| (1) | Conduct a thorough literature review to identify key factors (i.e., spatial, temporal, environmental, and fisheries variables) that may influence CPUE values; | Yes |
| (2) | Determine temporal and spatial scales for data grouping for CPUE standardization; | Yes and No  (exact spatial information is not available at the same level of details as the catches are) |
| (3) | Plot spatio-temporal distributions of fishing efforts and catch to evaluate spatio-temporal patterns of fishing effort and catch; | Yes (*see* Appendix II and III) |
| (4) | Calculate correlation matrix to evaluate correlations between each pair of those variables; | Yes (*see* Fig. 1) |
| (5) | Identify potential explanatory variables based on (1)-(4) to develop full model for the CPUE standardization; | Yes |
| (6) | Make statistical assumptions on the full models and fit the data to the assumed statistical models (i.e., GLM, GAM, Delta-lognormal GLM, Neural Networks, Regression Trees, Habitat based models, and Statistical habitat based models); | Yes (GLM) |
| (7) | Select and evaluate the models using methods such as likelihood ratio, AIC, BIC or cross validation; | Yes (BIC) |
| (8) | Evaluate if distributional assumptions are satisfied and if there is a consistent spatial/temporal distribution of residuals in CPUE standardization modeling; | Yes (*see* Fig. 2 – 5) |
| (9) | Determine the optimal model to estimate yearly standardized CPUE and their associated uncertainty. | Yes (*see* Table 1) |
| (10) | Plot nominal and standardized CPUEs over time. | Yes (*see* Fig. 8) |

**Appendix II** Annual changes in monthly catches of Russian stick-held dip net fishery for Pacific saury from 1994 to 2019









**Appendix III** Annual changes in monthly sum of fishing days of Russian stick-held dip net fishery for Pacific saury from 1994 to 2019

