NPFC-2021-SSC PS07-WP02

# 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). 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 marked by such factors as Year and Month with observations on a daily scale from 1994 to 2020. The oldest data is stored in TINRO from the period of 1994 to 1999. The data from the period of 2000-2015 was provided to the NPFC as an Attachment to Annex 1 of NPFC-2017-AR (Annual report by Russia). It includes daily catch records for each vessel. We added all saury catches reported in 2016–2020 years to and provided by DER of Federal government-financed institution "Centre of Fishery Monitoring and Communications". Then we applied a filter on August-November and Russian flag on vessels. Many of 207 vessels did not report saury catches for several years and their catches were not targeted; therefore, to avoid problems with vessel coefficient estimation we selected the most active 52 vessels, which in fact reported at least 229 days of saury catches. Finally, the selection included 24281 reports made by those 52 vessels. They took part in saury fishery at least for 5 years. We did not 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. The main problem is that we do not know exact amount of catch for each operation, because catch reports were submitted in sum for each day by each vessel though they did 4-8 and sometimes even more sets per night which could be far away from each other. Positions aggregated for the Appendices were selected on time of report submission around 11 pm local time.

## Correlations between each pair of predictors and response

We do not 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 – 24281 daily reports (Fig. 1).

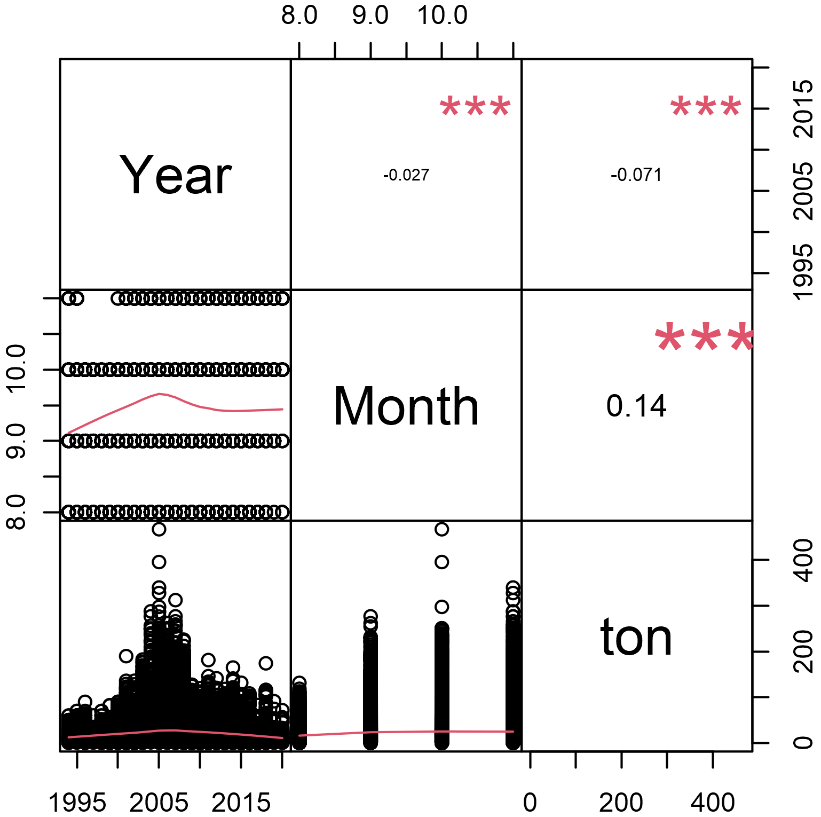


Figure 1 – Pearson’s correlations, where “\*\*\*” indicates *p* < 0.001, “\*\*” – *p* < 0.01,   
“\*” – *p* < 0.05 and “∙” – *p* < 0.1

## 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 do not have officially reported values for SST as well as other possible predictors, so we will 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.

## Fit candidate statistical models to the data

We did not 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 will 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 awfully 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.* Thus, we fit Gamma instead of Tweedie and finally compared it to Gaussian distribution with the same link function (logarithmic). Full model with Gamma distribution was the best.

## 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 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, highest explained deviance, and best performance in 100 cross-validation runs (see Tables 1 and 2). 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 |  | 0.84 | 0.12 | 13.9 | 211953 | 28 |
| 2 |  | 0.83 | 0.13 | 15.2 | 211571 | 31 |
| 3 |  | 0.79 | 0.16 | 18.9 | 211091 | 105 |
| 4 |  | 0.73 | 0.20 | 23.1 | **210152** | 156 |
| 5 |  | 0.76 | 0.17 | 19.4 | 210701 | 82 |

*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*).

Table 2 – statistical properties of 100 cross-validation runs on 20% of out of bag test sets

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| GLM No | Average | SD | Median | MAD | IQR |
| Error measure: RMSE | | | | |
| 1 | 31.0859 | 0.7322 | 31.0444 | 0.7812 | 1.0519 |
| 2 | 30.9029 | 0.7245 | 30.8564 | 0.7302 | 1.0116 |
| 3 | 30.5000 | 0.7077 | 30.4667 | 0.7399 | 0.9547 |
| 4 | **29.8255** | **0.6908** | **29.7978** | **0.6596** | **0.8991** |
| 5 | 30.2286 | 0.7000 | 30.1830 | 0.7220 | 0.9755 |
|  | Error measure: R2 | | | | |
| 1 | 0.1220 | 0.0074 | 0.1220 | 0.0077 | 0.0105 |
| 2 | 0.1323 | 0.0079 | 0.1319 | 0.0085 | 0.0118 |
| 3 | 0.1549 | 0.0089 | 0.1546 | 0.0091 | 0.0126 |
| 4 | **0.1919** | **0.0096** | **0.1911** | **0.0104** | **0.0133** |
| 5 | 0.1698 | 0.0088 | 0.1688 | 0.0089 | 0.0117 |
|  | Error measure: MAE | | | | |
| 1 | 21.0804 | 0.2815 | 21.0843 | 0.2977 | 0.3798 |
| 2 | 20.9282 | 0.2790 | 20.9366 | 0.2758 | 0.3803 |
| 3 | 20.4850 | 0.2709 | 20.5093 | 0.2862 | 0.4032 |
| 4 | **19.8632** | **0.2648** | **19.8858** | **0.2794** | **0.3622** |
| 5 | 20.2902 | 0.2714 | 20.3232 | 0.2650 | 0.3685 |

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

Gamma distribution suited very well to capture overdispersion and we do not see patterns in the residuals (Fig. 2). The rank of the final model is less than 1 (155/159) therefore huge number of parameters do not make our model saturated one. The same model with Gaussian distribution (and log link) had problems with ML estimation, but after REML run it explained less deviance (20.4%) then with Gamma (23.1%). Therefore we will continue to use Gamma distribution in the selected GLM as we did all the previous years.

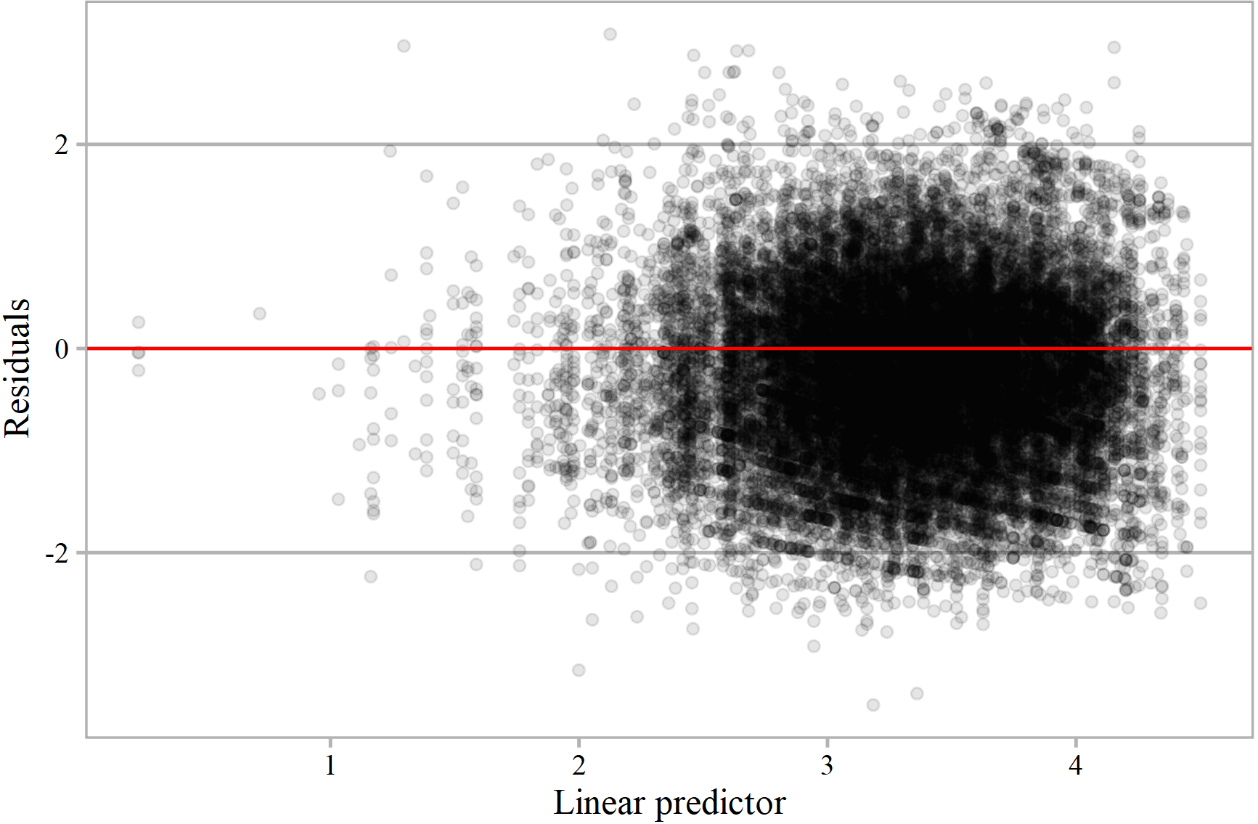


Figure 2 – Residuals of GLM No 4 versus linear predictor

Residual autocorrelation was checked by generalized Durbin-Watson (D-W) statistics, which bootstrapped p-values were higher than 0.37 by years at lags from 1 to 5 years (Table 3).

Table 3 – Autocorrelation, D-W statistic, and its significance by different lags in time

|  |  |  |  |
| --- | --- | --- | --- |
| lag | Autocorrelation | D-W Statistic | bootstrapped p-value |
| 1 | 0.1203 | 1.66 | 0.376 |
| 2 | 0.0886 | 1.65 | 0.478 |
| 3 | -0.0591 | 1.75 | 0.784 |
| 4 | -0.1729 | 1.90 | 0.744 |
| 5 | -0.1231 | 1.63 | 0.868 |

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).

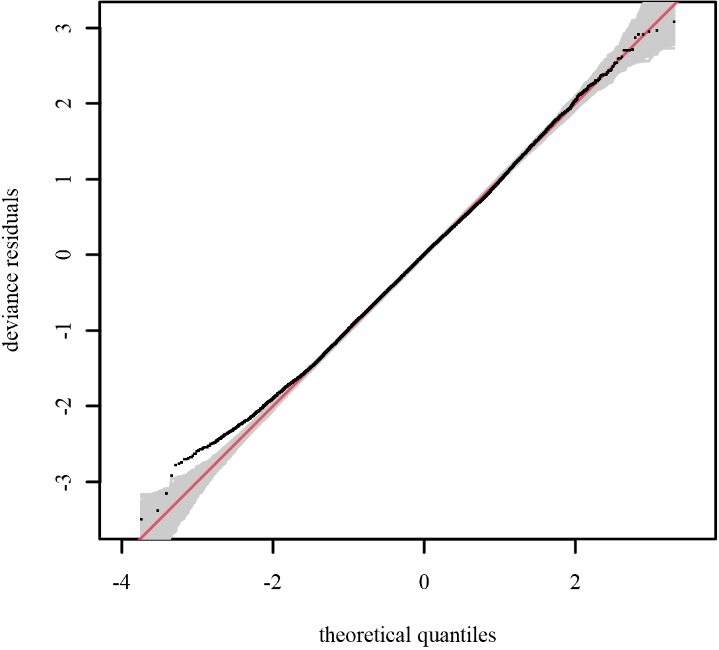
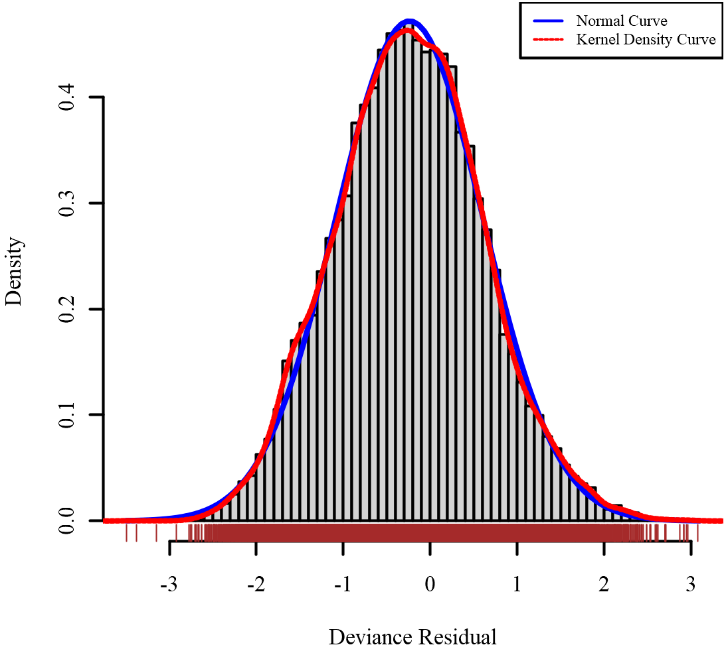


Figure 3 – Quantile-quantile plot (left) conditional on the fitted model coefficients and scale parameter with 1000 simulated quantiles of the residual distribution and histogram (right) of residuals from GLM No 4 on a link scale

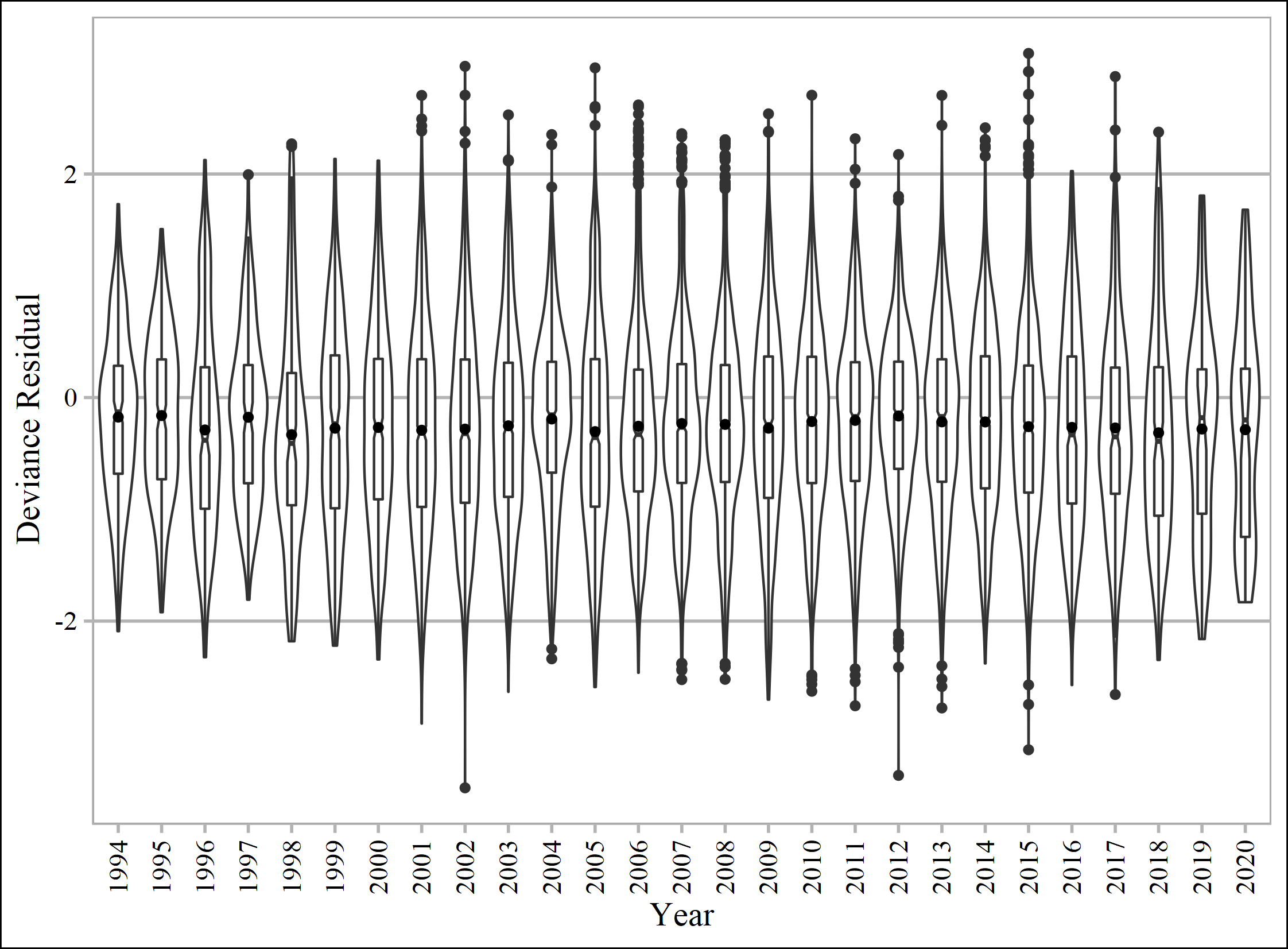


Figure 4 – Distribution of deviance residuals by years, where notched boxes show IQR and dots inside them show the average and outside show the values higher than 1.5\*IQR

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

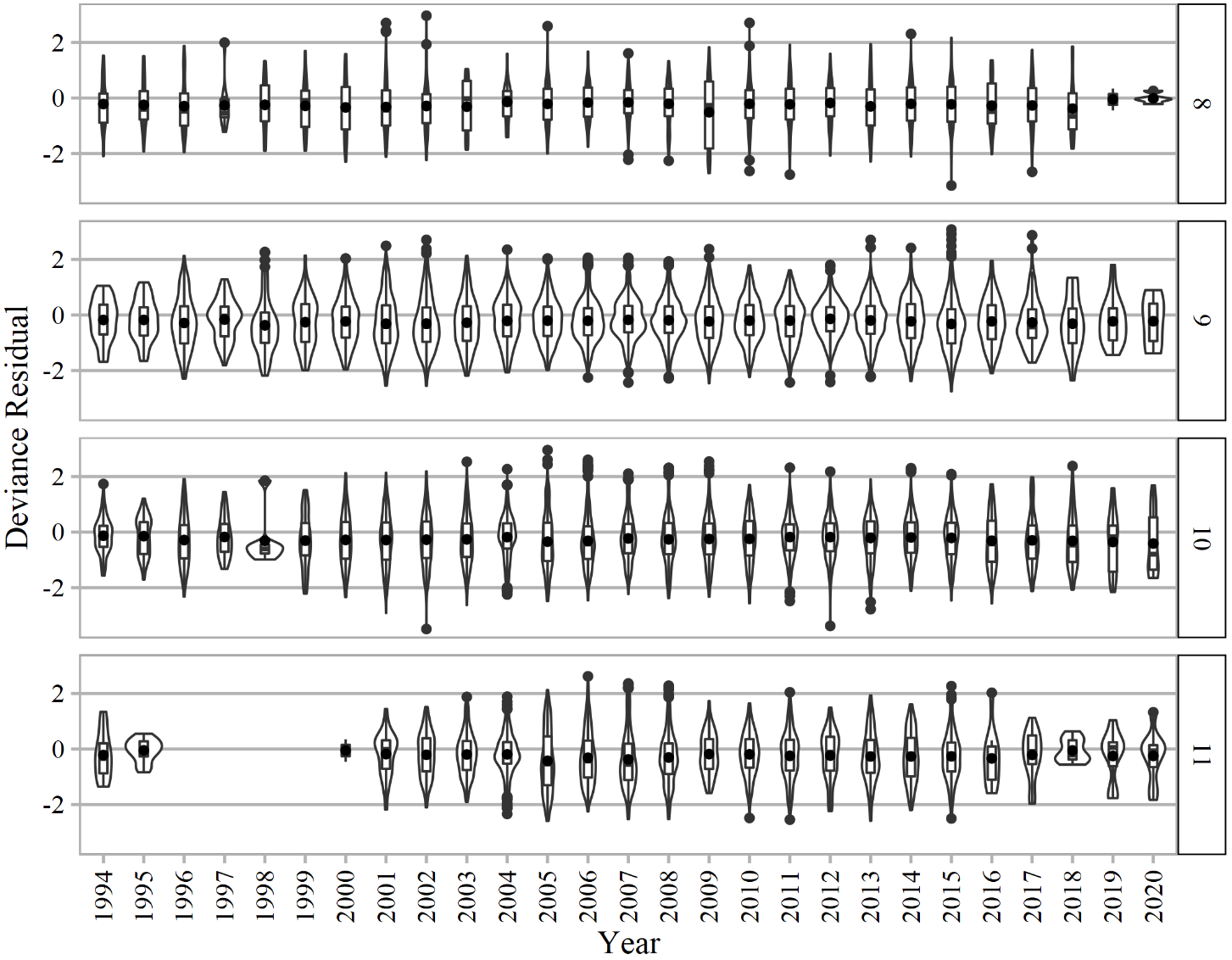


Figure 5 – Distribution of deviance residuals by years given month (numbers in the box to the right of each plot), where boxes show IQRs and dots inside them show the average and outside show the values higher than 1.5\*IQR

The highest influence (Bentley et al., 2012) was found for the interaction term and Idves in 2020 (Fig. 6). The most part of the selected vessels did not take part in saury fishing in 2019 and there was only 1 vessel in 2020 most of the time, while the second vessel in 2020 fished occasionally and it did not pass our filter. Thus, index of abundance may not be accurate enough in 2019-2020 (Fig. 7).

## Extract yearly standardized CPUE and standard error

We could continue to increase the number of predictors, e.g. to include power of vessels. But many parameters, describing vessels are causally 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.



Figure 6 – Influence of interaction term () and Idves



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

## A time series of yearly standardized CPUE and associated uncertainty

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 (1994-2020), unique identifiers of vessels (52 levels). Then we predicted catches using GLM No 4 and summarized them. Summary statistics of predicted values is given below by years (Appendix 4, Table 1) as well as statistics of raw data used for standardization (Appendix 4, Table 2) and the same tables described from data transformed with natural logarithm (Appendix 4, Tables 3 and 4 respectively). We show all the uncertainty in the predictions with violin and boxplots in Fig. 8 as we did in our previous working papers. Figure 9 shows all of 100 runs by different GLMS during cross-validation.

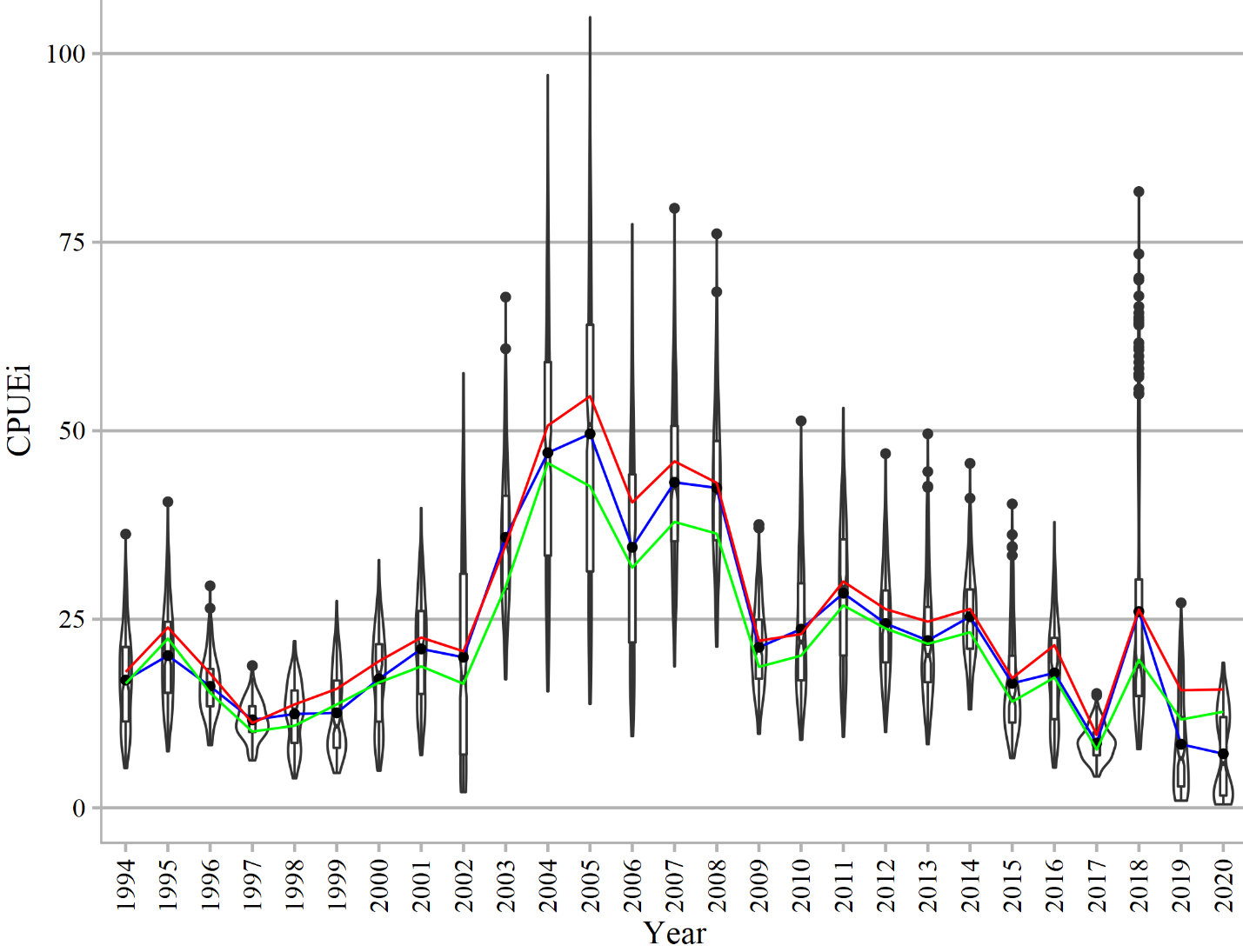


Figure 8 – Violin and box plots for catches per day per vessel (CPUEi) predicted from GLM No 4, where blue line connects means of predictions, red line connects means of raw catch values (tons per day) used for standardization, while green line connects trimmed means of original catch values

Most of the estimates in GLM4 are below the estimates of other GLMs in 2019 and 2020. Though on its own GLM4 has not so high variation in predictions for 2019 and 2020 (Fig. 10).

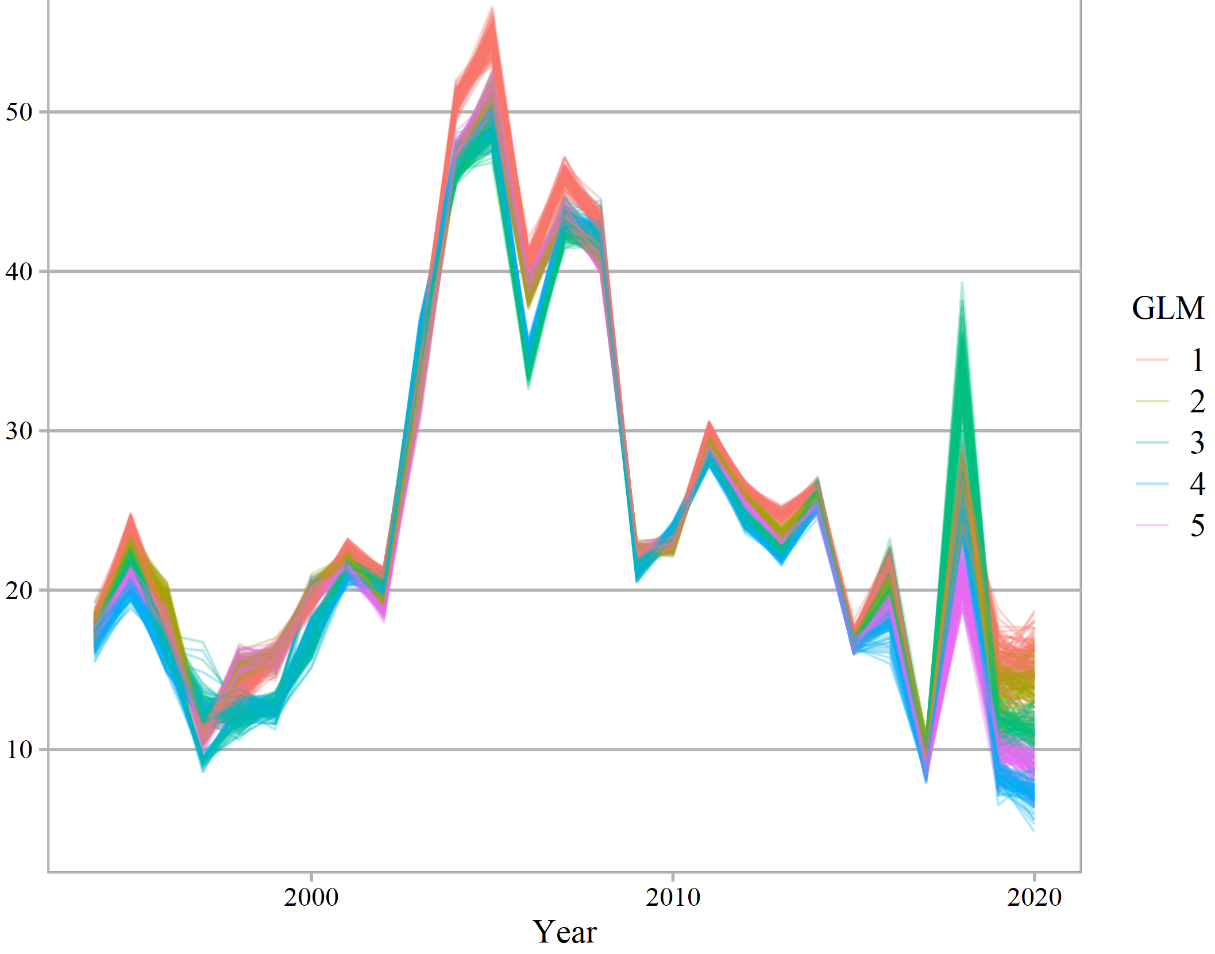


Figure 9 – Estimates from expanded grids which were used for predictions by different GLMs

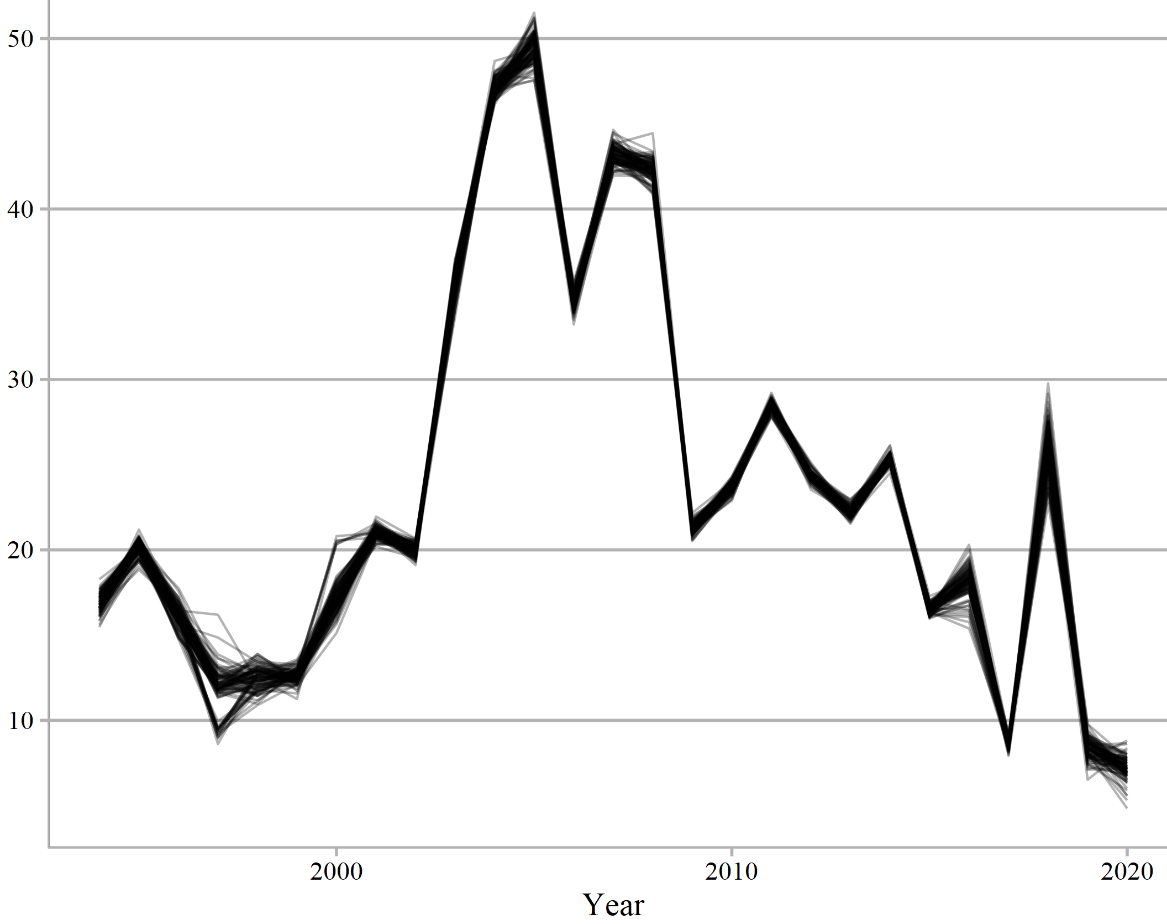


Figure 10 – Mean timeseries from expanded grid which was used for predictions by different GLMs

Descriptive statistics for all of 100 runs of GLM4 is given in the table 4. We recommend using the mean estimates from the table 4.

Table 4 – Statistics of 100 mean predictions on expanded grids from GLM4, which were trained using 80% resamples each time

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Year | Median | Mean | SD | Quantile 0.025 | Quantile 0.975 |
| 1994 | 16.923 | 16.890 | 0.531 | 15.837 | 17.755 |
| 1995 | 20.158 | 20.150 | 0.435 | 19.297 | 20.796 |
| 1996 | 16.216 | 16.154 | 0.625 | 14.848 | 17.215 |
| 1997 | 11.895 | 11.739 | 1.364 | 9.008 | 13.775 |
| 1998 | 12.546 | 12.486 | 0.667 | 11.161 | 13.772 |
| 1999 | 12.632 | 12.605 | 0.410 | 11.773 | 13.390 |
| 2000 | 17.122 | 17.310 | 1.038 | 15.833 | 20.484 |
| 2001 | 21.038 | 21.048 | 0.337 | 20.343 | 21.673 |
| 2002 | 20.067 | 20.009 | 0.347 | 19.325 | 20.586 |
| 2003 | 35.840 | 35.762 | 0.703 | 34.119 | 36.893 |
| 2004 | 47.061 | 47.104 | 0.495 | 46.317 | 48.026 |
| 2005 | 49.471 | 49.500 | 0.819 | 47.810 | 51.180 |
| 2006 | 34.589 | 34.567 | 0.452 | 33.689 | 35.551 |
| 2007 | 43.169 | 43.208 | 0.535 | 42.197 | 44.275 |
| 2008 | 42.339 | 42.311 | 0.585 | 40.983 | 43.296 |
| 2009 | 21.231 | 21.258 | 0.313 | 20.621 | 21.888 |
| 2010 | 23.674 | 23.681 | 0.307 | 23.049 | 24.221 |
| 2011 | 28.473 | 28.492 | 0.295 | 27.859 | 29.009 |
| 2012 | 24.377 | 24.361 | 0.286 | 23.787 | 24.966 |
| 2013 | 22.161 | 22.196 | 0.302 | 21.684 | 22.860 |
| 2014 | 25.378 | 25.370 | 0.277 | 24.903 | 25.819 |
| 2015 | 16.525 | 16.517 | 0.263 | 16.036 | 17.022 |
| 2016 | 18.253 | 18.171 | 0.869 | 16.108 | 19.575 |
| 2017 | 8.600 | 8.591 | 0.255 | 8.103 | 9.075 |
| 2018 | 26.213 | 26.055 | 1.310 | 23.469 | 28.505 |
| 2019 | 8.409 | 8.393 | 0.586 | 7.202 | 9.415 |
| 2020 | 7.173 | 7.188 | 0.669 | 5.575 | 8.505 |

## Plot nominal and standardized CPUEs over time

Nominal CPUE is shown in Fig. 8 and Appendix IV (Table 2). Standardized CPUE from Table 4 is compared to Standardized CPUE from “NPFC-2020-SSC PS06-WP04 Pacific saury CPUE standardization\_Russia” on original scale in Fig. 11 and after mean centering on a logarithmic scale and exponentiating back in Fig. 12.

Figure 11 – Standardized CPUEs in tons per day per vessel up to 2019 and 2020 dashed black line shows nominal CPUE

Figure 12 – Mean centered on a log scale and exponentiated back CPUEs up to 2019 and 2020 dashed black line shows nominal CPUE

**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

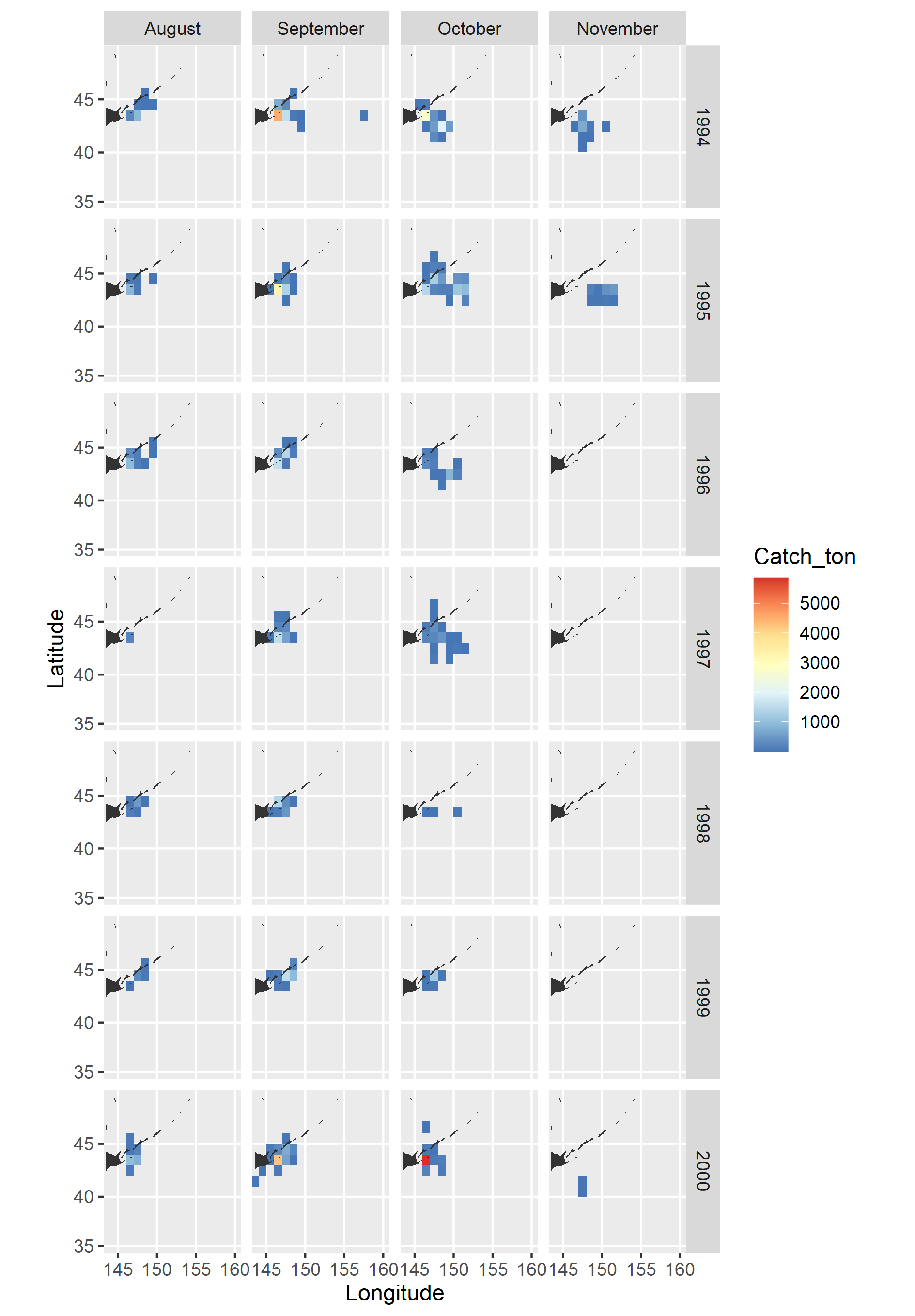
Parin, N. V. 1960. The range of the saury (Cololabis saira Brev.-Scombresocidae, Pices) and effects of oceanographic features on its distribution. *Proc. Acad. Sci. USSR*, *130*(3), 649–652.

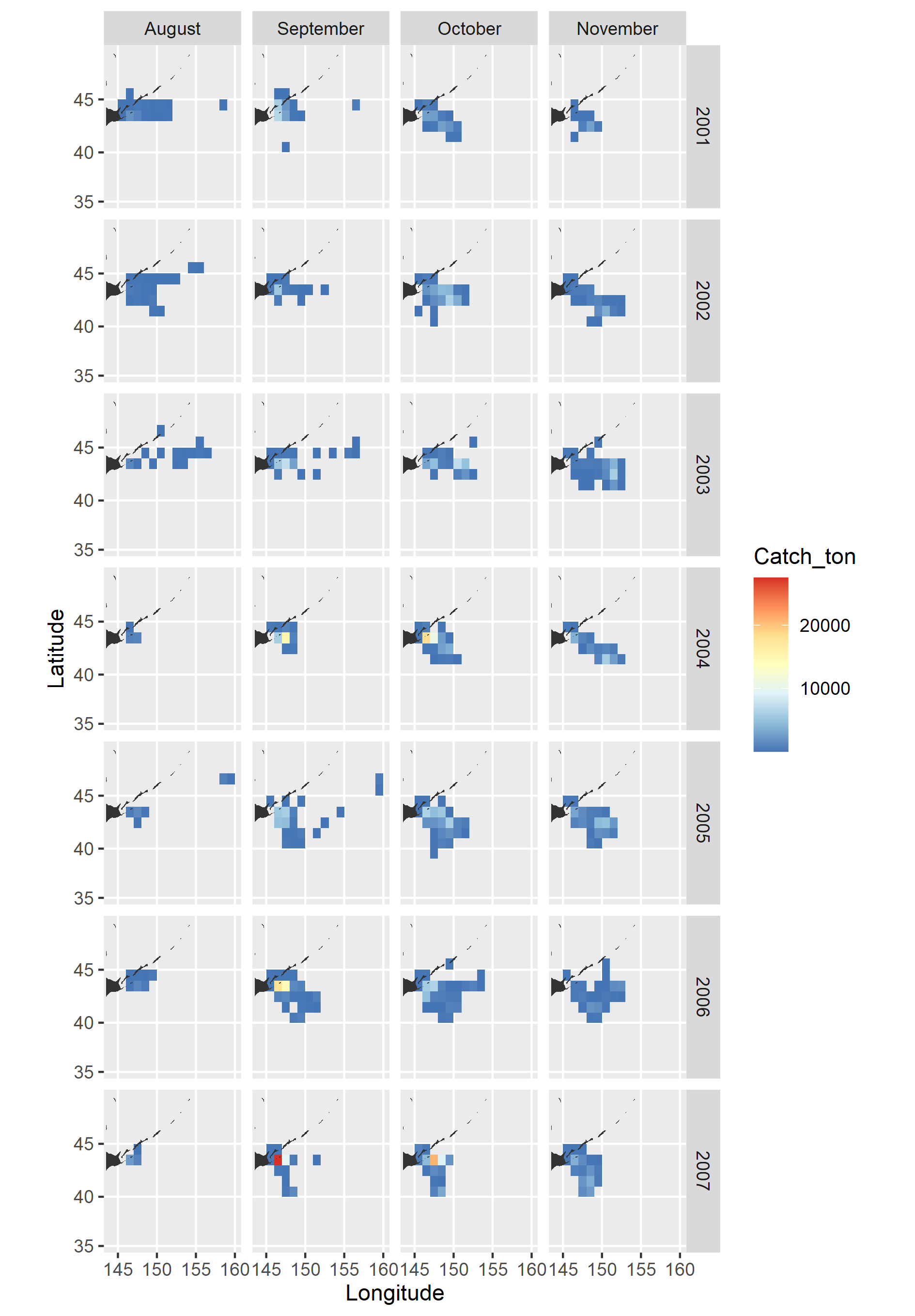
Wood, S. N. (2011). Fast stable restricted maximum likelihood and marginal likelihood estimation of semiparametric generalized linear models. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, *73*(1), 3–36. https://doi.org/10.1111/j.1467-9868.2010.00749.x

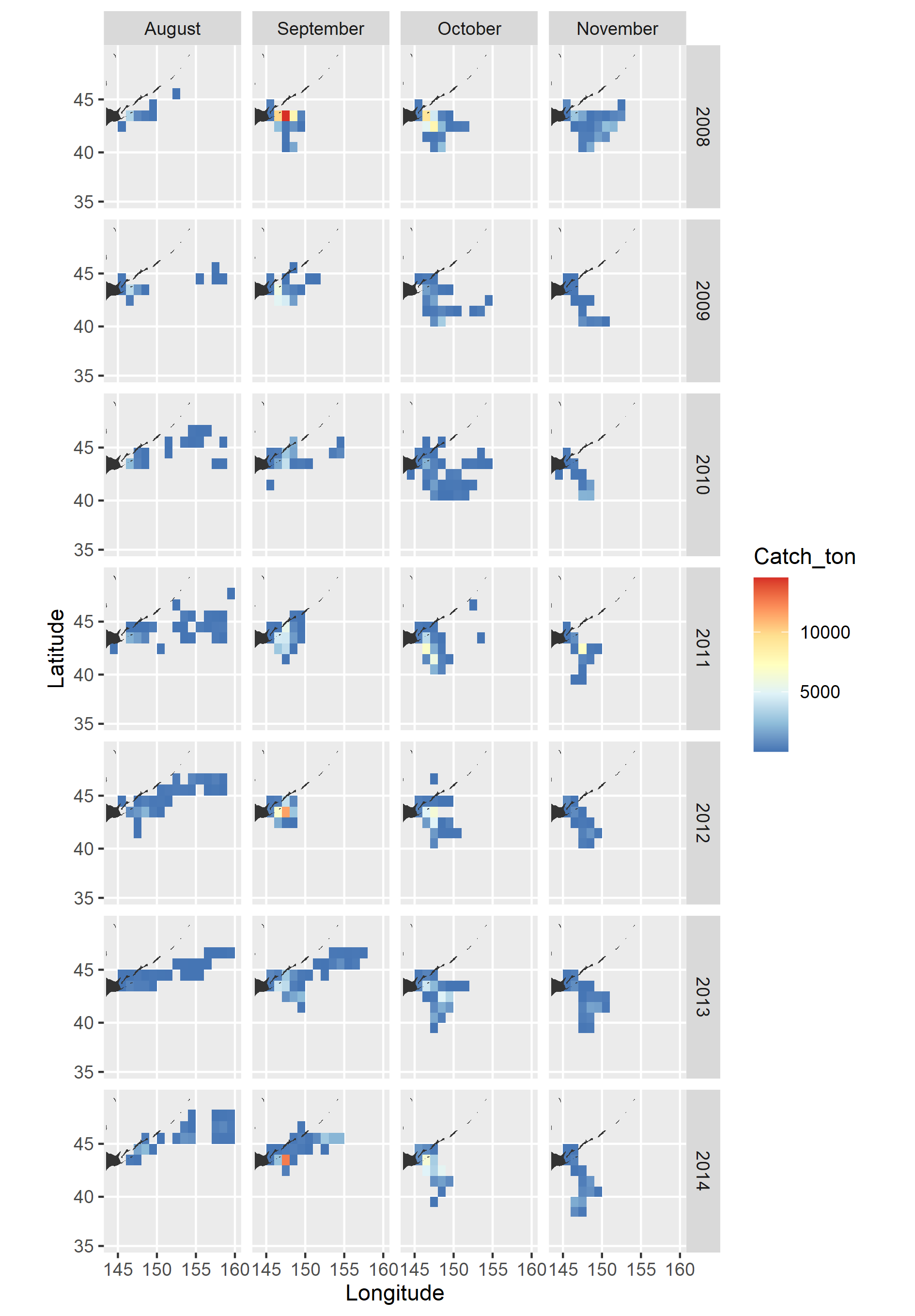
# Appendix I: Checklist for the CPUE standardization protocol

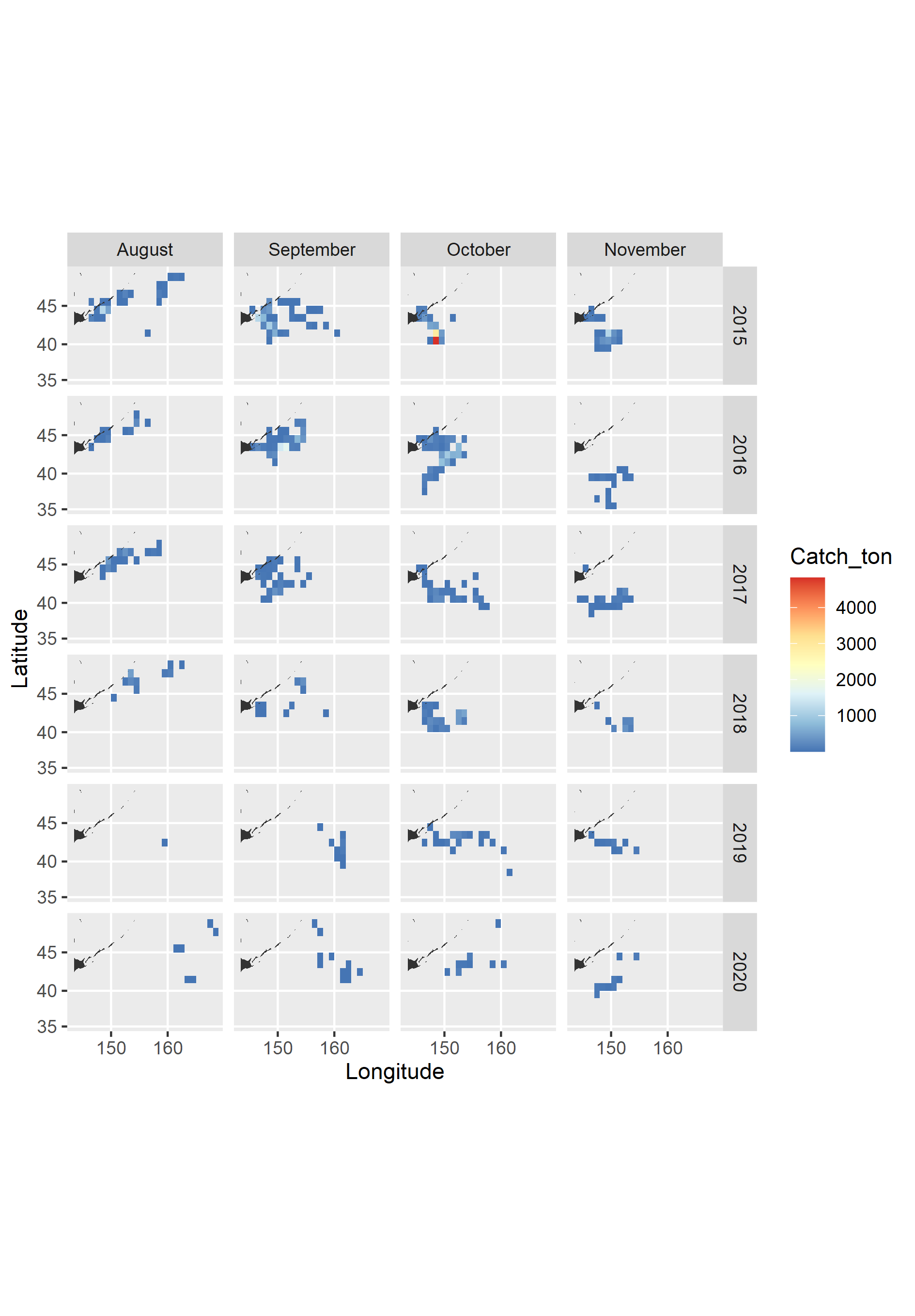
|  |  |  |
| --- | --- | --- |
| (1) | Conduct a thorough literature review to identify key factors (i.e., spatial, temporal, environmental, and fisheries variables) that may influence CPUE values; | Yes |
| (2) | Determine temporal and spatial scales for data grouping for CPUE standardization; | Yes 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) as well as interaction terms to develop full model for the CPUE standardization; | Yes |
| (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, *see* Table 1 and CV, *see* Table 2) |
| (8) | Evaluate if distributional assumptions are satisfied and if there is a significant spatial/temporal pattern of residuals in CPUE standardization modeling; | Yes, for distribution assumption and temporal term (*see* Fig. 2 – 5 and Table 2), and No, for spatial, because there is no exact spatial information. |
| (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, expanded grid was used (*see* Table 4). Spatial grouping is not available. |
| (10) | Recommend a time series of yearly standardized CPUE and associated uncertainty. | Yes (*see* Table 4) |

**Appendix II:** Monthly catches of Russian stick-held dip net fishery for Pacific saury from 1994 to 2020

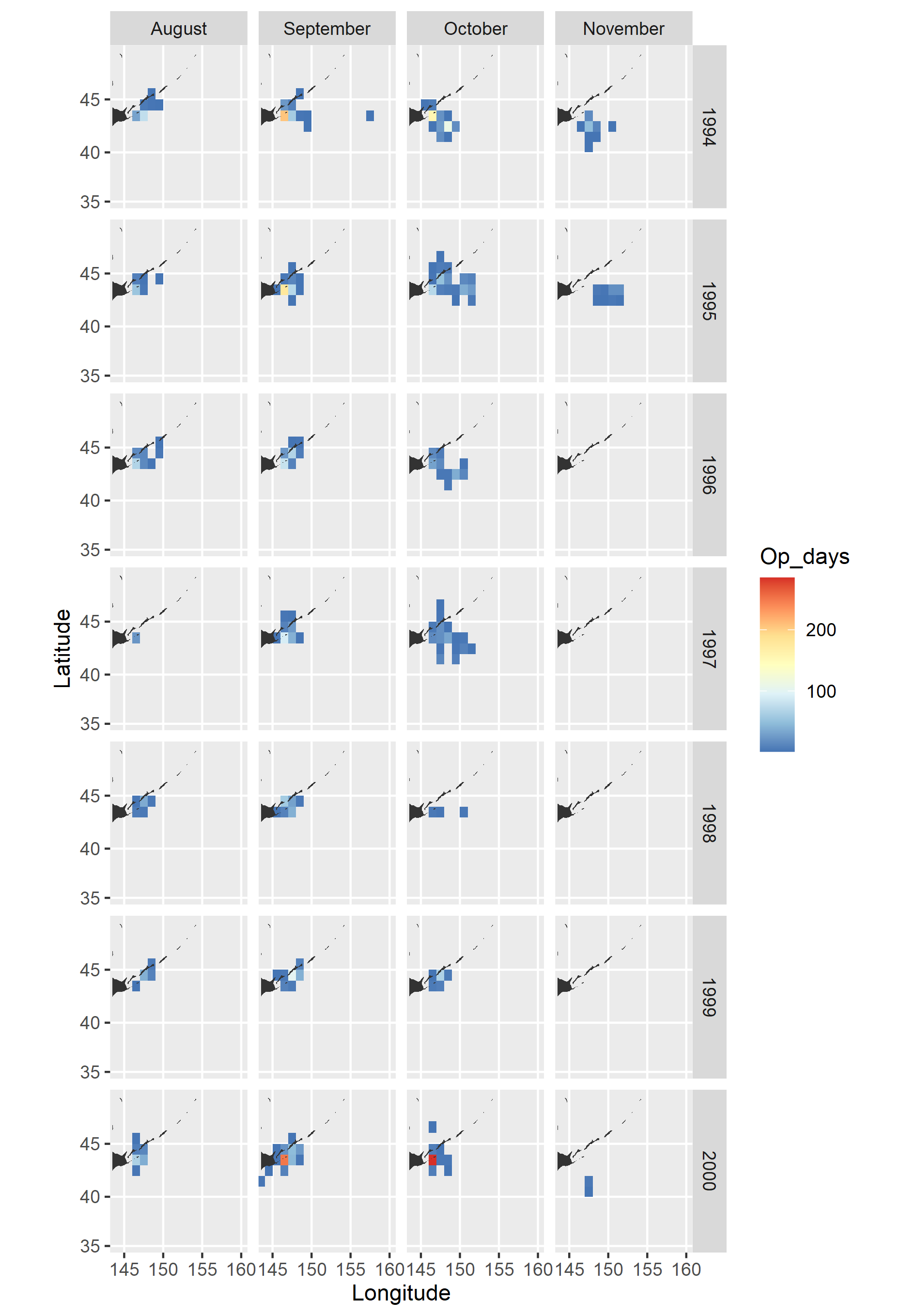


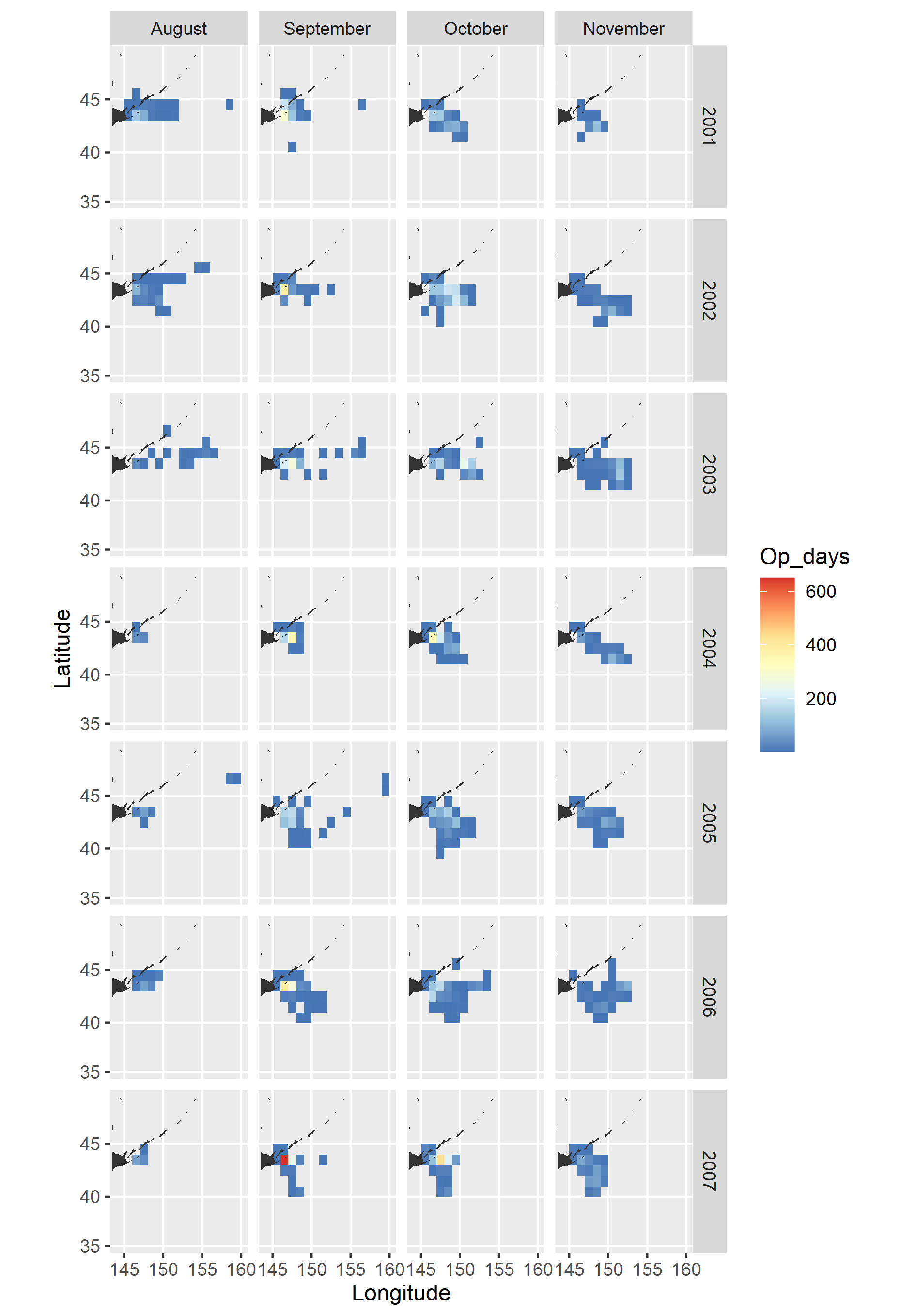


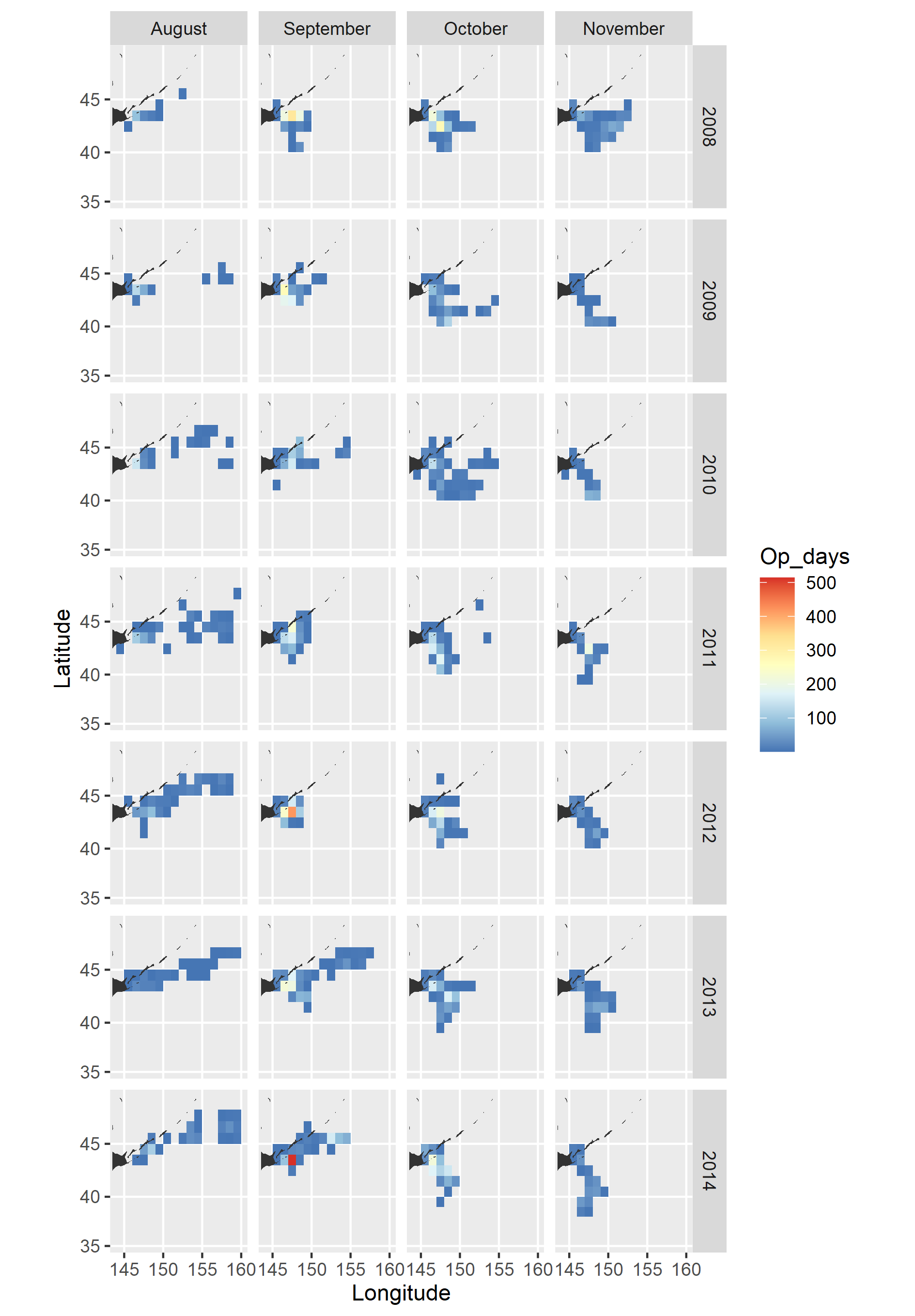


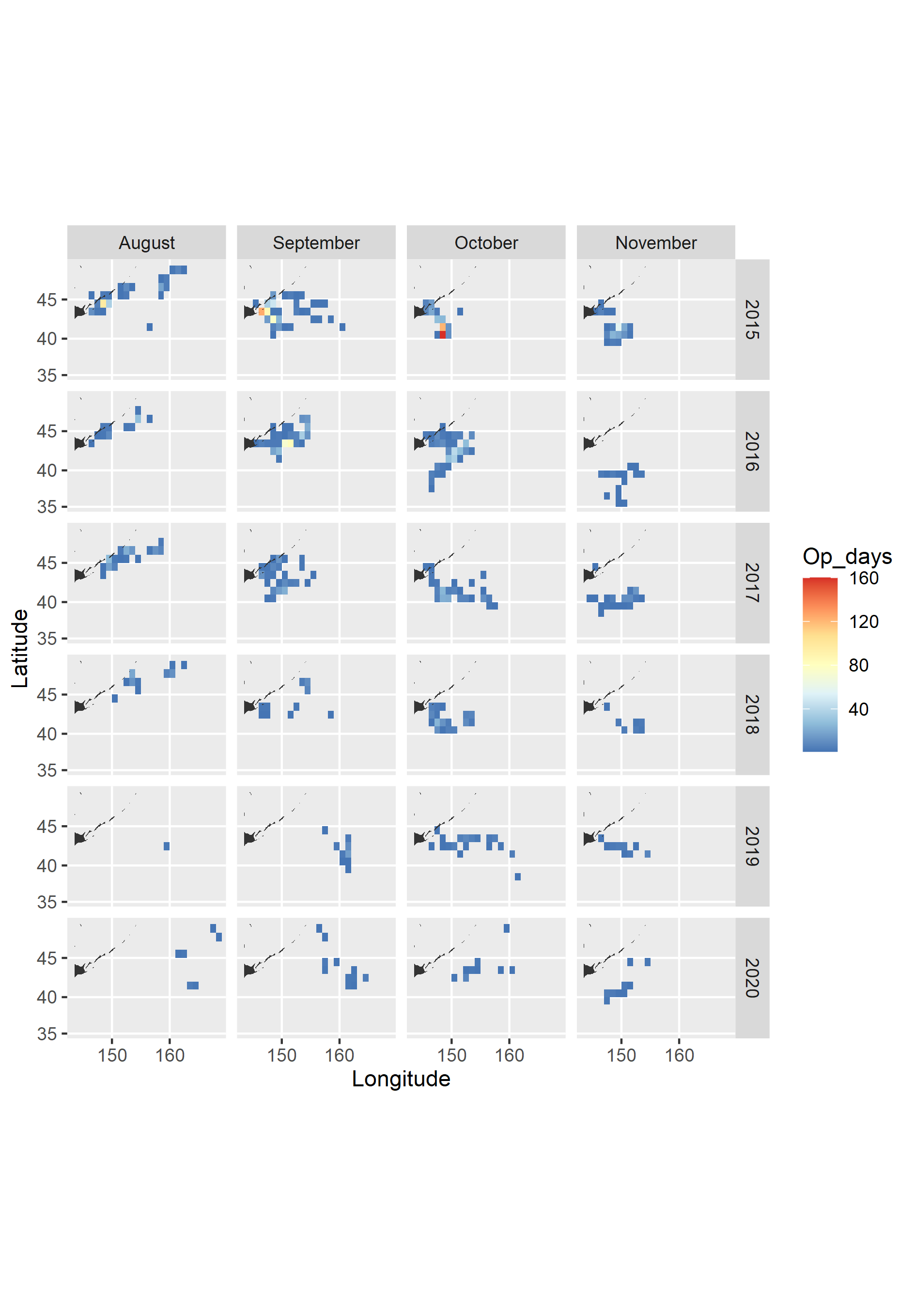


**Appendix III:** Monthly sum of fishing days of Russian stick-held dip net fishery for saury from 1994 to 2020









**Appendix IV:** descriptivestatistics of predicted and original CPUE

Table 1 – Summary statistics of predicted 208 values (52 Idves\*4 Month) for each year

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Year | Mean | SD | SE | Median | Trimmed mean by  (by 10%) | MAD | Min | Max | Skew | Kurtosis |
| 1994 | 16.942 | 6.293 | 0.436 | 16.598 | 16.677 | 7.238 | 5.292 | 36.294 | 0.303 | -0.441 |
| 1995 | 20.192 | 6.436 | 0.446 | 19.937 | 19.914 | 7.033 | 7.501 | 40.572 | 0.359 | -0.328 |
| 1996 | 16.115 | 3.884 | 0.269 | 15.935 | 15.971 | 3.674 | 8.320 | 29.434 | 0.397 | 0.099 |
| 1997 | 11.668 | 2.536 | 0.176 | 11.552 | 11.660 | 2.539 | 6.285 | 18.841 | 0.140 | -0.232 |
| 1998 | 12.445 | 4.214 | 0.292 | 12.965 | 12.473 | 5.035 | 3.914 | 22.115 | -0.090 | -0.861 |
| 1999 | 12.599 | 5.297 | 0.367 | 10.806 | 12.253 | 5.727 | 4.635 | 27.418 | 0.497 | -0.860 |
| 2000 | 17.096 | 6.240 | 0.433 | 17.859 | 17.112 | 7.283 | 4.925 | 32.843 | -0.097 | -0.877 |
| 2001 | 21.067 | 7.054 | 0.489 | 21.469 | 20.964 | 8.252 | 7.009 | 39.741 | 0.082 | -0.740 |
| 2002 | 19.966 | 13.664 | 0.947 | 17.282 | 18.971 | 18.509 | 2.090 | 57.638 | 0.431 | -0.834 |
| 2003 | 35.880 | 9.252 | 0.641 | 34.846 | 35.444 | 9.378 | 17.045 | 67.717 | 0.477 | 0.120 |
| 2004 | 47.077 | 16.761 | 1.162 | 46.949 | 46.468 | 18.859 | 15.458 | 97.167 | 0.270 | -0.560 |
| 2005 | 49.587 | 20.220 | 1.402 | 50.822 | 48.979 | 23.280 | 13.787 | 104.821 | 0.182 | -0.823 |
| 2006 | 34.552 | 14.375 | 0.997 | 34.551 | 33.911 | 16.539 | 9.503 | 77.393 | 0.299 | -0.559 |
| 2007 | 43.155 | 11.398 | 0.790 | 42.500 | 42.780 | 11.549 | 18.767 | 79.498 | 0.309 | -0.229 |
| 2008 | 42.420 | 10.007 | 0.694 | 41.814 | 42.126 | 9.564 | 21.373 | 76.107 | 0.335 | 0.015 |
| 2009 | 21.291 | 5.456 | 0.378 | 21.064 | 21.099 | 5.803 | 9.824 | 37.565 | 0.333 | -0.222 |
| 2010 | 23.699 | 8.635 | 0.599 | 21.986 | 23.077 | 9.097 | 9.020 | 51.301 | 0.617 | -0.298 |
| 2011 | 28.456 | 9.327 | 0.647 | 29.248 | 28.422 | 10.937 | 9.433 | 53.002 | -0.008 | -0.779 |
| 2012 | 24.436 | 7.033 | 0.488 | 24.068 | 24.139 | 7.093 | 10.069 | 46.968 | 0.380 | -0.241 |
| 2013 | 22.185 | 7.934 | 0.550 | 20.220 | 21.459 | 7.422 | 8.438 | 49.592 | 0.817 | 0.321 |
| 2014 | 25.345 | 5.966 | 0.414 | 25.021 | 25.159 | 5.828 | 13.064 | 45.668 | 0.353 | 0.062 |
| 2015 | 16.508 | 6.930 | 0.481 | 14.509 | 15.678 | 5.582 | 6.600 | 40.290 | 1.040 | 0.419 |
| 2016 | 17.894 | 6.752 | 0.468 | 18.159 | 17.685 | 7.601 | 5.336 | 37.887 | 0.193 | -0.588 |
| 2017 | 8.591 | 2.166 | 0.150 | 8.436 | 8.514 | 2.248 | 4.156 | 15.155 | 0.352 | -0.171 |
| 2018 | 26.023 | 17.000 | 1.179 | 18.780 | 23.328 | 7.546 | 7.783 | 81.700 | 1.319 | 0.504 |
| 2019 | 8.429 | 6.343 | 0.440 | 6.554 | 7.730 | 6.660 | 0.955 | 27.178 | 0.794 | -0.356 |
| 2020 | 7.168 | 5.614 | 0.389 | 5.885 | 6.877 | 7.465 | 0.462 | 19.251 | 0.243 | -1.489 |

Table 2 – 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 | Kur-tosis |
| 1994 | 248 | 18.022 | 13.202 | 0.838 | 15.000 | 16.477 | 11.861 | 0.500 | 60.000 | 0.936 | 0.144 |
| 1995 | 236 | 23.913 | 16.100 | 1.048 | 20.000 | 22.439 | 14.826 | 1.000 | 70.000 | 0.700 | -0.384 |
| 1996 | 225 | 17.824 | 16.592 | 1.106 | 12.000 | 15.257 | 11.861 | 0.500 | 90.000 | 1.393 | 1.564 |
| 1997 | 149 | 11.252 | 8.451 | 0.692 | 10.000 | 10.112 | 7.413 | 1.000 | 60.300 | 1.934 | 6.682 |
| 1998 | 128 | 13.713 | 14.749 | 1.304 | 9.550 | 10.890 | 8.228 | 0.500 | 70.000 | 1.886 | 3.362 |
| 1999 | 222 | 15.781 | 14.371 | 0.965 | 10.000 | 13.714 | 10.897 | 0.500 | 82.700 | 1.545 | 3.213 |
| 2000 | 483 | 19.445 | 17.501 | 0.796 | 15.000 | 16.616 | 14.826 | 0.500 | 90.000 | 1.561 | 2.554 |
| 2001 | 1041 | 22.578 | 22.420 | 0.695 | 15.500 | 18.757 | 15.567 | 0.100 | 190.000 | 2.048 | 6.308 |
| 2002 | 1165 | 20.759 | 22.681 | 0.665 | 12.600 | 16.454 | 13.936 | 0.020 | 136.100 | 1.948 | 4.249 |
| 2003 | 1268 | 34.820 | 32.201 | 0.904 | 24.450 | 29.305 | 23.054 | 0.400 | 182.500 | 1.703 | 3.179 |
| 2004 | 1439 | 50.683 | 39.278 | 1.035 | 44.100 | 45.712 | 33.757 | 1.700 | 288.000 | 1.808 | 5.608 |
| 2005 | 1265 | 54.574 | 59.198 | 1.664 | 34.000 | 42.626 | 32.914 | 1.000 | 467.500 | 2.063 | 5.131 |
| 2006 | 1568 | 40.487 | 43.398 | 1.096 | 25.100 | 31.847 | 23.392 | 0.884 | 262.000 | 2.337 | 6.243 |
| 2007 | 1857 | 45.946 | 43.519 | 1.010 | 31.991 | 37.910 | 28.156 | 0.500 | 312.500 | 2.104 | 5.328 |
| 2008 | 1756 | 43.094 | 39.639 | 0.946 | 29.882 | 36.356 | 27.453 | 0.333 | 226.000 | 1.880 | 4.053 |
| 2009 | 1391 | 22.153 | 20.576 | 0.552 | 16.500 | 18.694 | 14.989 | 0.229 | 127.435 | 1.779 | 3.765 |
| 2010 | 1262 | 23.040 | 19.401 | 0.546 | 17.686 | 20.182 | 15.567 | 0.110 | 120.000 | 1.466 | 2.469 |
| 2011 | 1512 | 30.004 | 23.657 | 0.608 | 24.006 | 26.836 | 20.278 | 0.191 | 181.400 | 1.419 | 2.766 |
| 2012 | 1727 | 26.342 | 19.391 | 0.467 | 22.200 | 23.776 | 16.309 | 0.035 | 141.500 | 1.412 | 2.620 |
| 2013 | 1471 | 24.690 | 20.417 | 0.532 | 19.500 | 21.740 | 17.050 | 0.200 | 118.100 | 1.320 | 1.650 |
| 2014 | 1880 | 26.372 | 21.833 | 0.504 | 20.419 | 23.283 | 18.083 | 0.404 | 165.000 | 1.460 | 2.809 |
| 2015 | 978 | 17.194 | 17.336 | 0.554 | 11.692 | 14.065 | 11.190 | 0.019 | 131.700 | 2.134 | 6.110 |
| 2016 | 465 | 21.562 | 21.977 | 1.019 | 13.891 | 17.286 | 13.858 | 0.421 | 108.309 | 1.848 | 3.324 |
| 2017 | 301 | 9.660 | 10.362 | 0.597 | 6.086 | 7.723 | 5.761 | 0.085 | 83.013 | 2.681 | 10.944 |
| 2018 | 148 | 26.281 | 31.810 | 2.615 | 12.850 | 19.612 | 13.591 | 0.261 | 174.267 | 1.948 | 3.730 |
| 2019 | 50 | 15.586 | 19.038 | 2.692 | 7.789 | 11.713 | 9.090 | 1.005 | 91.455 | 2.080 | 4.590 |
| 2020 | 46 | 15.700 | 17.597 | 2.595 | 9.447 | 12.748 | 11.622 | 0.804 | 71.757 | 1.434 | 1.262 |

Table 3 – Summary statistics of predicted 208 values for each year on a Log scale

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Year | Mean | SD | SE | Median | Trimmed mean by  by 10% | MAD | Min | Max | Skew | Kurtosis |
| 1994 | 2.754 | 0.404 | 0.028 | 2.809 | 2.774 | 0.415 | 1.666 | 3.592 | -0.481 | -0.462 |
| 1995 | 2.952 | 0.335 | 0.023 | 2.993 | 2.964 | 0.332 | 2.015 | 3.703 | -0.372 | -0.342 |
| 1996 | 2.750 | 0.246 | 0.017 | 2.768 | 2.758 | 0.228 | 2.119 | 3.382 | -0.281 | -0.081 |
| 1997 | 2.432 | 0.226 | 0.016 | 2.447 | 2.445 | 0.224 | 1.838 | 2.936 | -0.444 | -0.053 |
| 1998 | 2.454 | 0.386 | 0.027 | 2.562 | 2.486 | 0.358 | 1.365 | 3.096 | -0.719 | -0.328 |
| 1999 | 2.443 | 0.431 | 0.030 | 2.380 | 2.448 | 0.533 | 1.534 | 3.311 | -0.052 | -1.087 |
| 2000 | 2.759 | 0.425 | 0.029 | 2.883 | 2.794 | 0.373 | 1.594 | 3.492 | -0.742 | -0.405 |
| 2001 | 2.985 | 0.370 | 0.026 | 3.067 | 3.008 | 0.387 | 1.947 | 3.682 | -0.586 | -0.379 |
| 2002 | 2.664 | 0.910 | 0.063 | 2.850 | 2.719 | 0.898 | 0.737 | 4.054 | -0.561 | -1.013 |
| 2003 | 3.547 | 0.262 | 0.018 | 3.551 | 3.553 | 0.265 | 2.836 | 4.215 | -0.227 | -0.134 |
| 2004 | 3.783 | 0.385 | 0.027 | 3.849 | 3.801 | 0.389 | 2.738 | 4.576 | -0.458 | -0.480 |
| 2005 | 3.808 | 0.459 | 0.032 | 3.928 | 3.836 | 0.454 | 2.624 | 4.652 | -0.529 | -0.659 |
| 2006 | 3.445 | 0.463 | 0.032 | 3.542 | 3.469 | 0.462 | 2.252 | 4.349 | -0.502 | -0.605 |
| 2007 | 3.729 | 0.274 | 0.019 | 3.749 | 3.739 | 0.269 | 2.932 | 4.376 | -0.364 | -0.154 |
| 2008 | 3.719 | 0.242 | 0.017 | 3.733 | 3.728 | 0.237 | 3.062 | 4.332 | -0.324 | -0.071 |
| 2009 | 3.025 | 0.264 | 0.018 | 3.048 | 3.033 | 0.267 | 2.285 | 3.626 | -0.319 | -0.185 |
| 2010 | 3.099 | 0.367 | 0.025 | 3.090 | 3.103 | 0.424 | 2.199 | 3.938 | -0.070 | -0.685 |
| 2011 | 3.287 | 0.366 | 0.025 | 3.376 | 3.314 | 0.361 | 2.244 | 3.970 | -0.656 | -0.312 |
| 2012 | 3.153 | 0.298 | 0.021 | 3.181 | 3.163 | 0.299 | 2.309 | 3.849 | -0.318 | -0.262 |
| 2013 | 3.038 | 0.352 | 0.024 | 3.007 | 3.036 | 0.357 | 2.133 | 3.904 | -0.003 | -0.389 |
| 2014 | 3.204 | 0.241 | 0.017 | 3.220 | 3.213 | 0.232 | 2.570 | 3.821 | -0.311 | -0.066 |
| 2015 | 2.724 | 0.394 | 0.027 | 2.675 | 2.711 | 0.407 | 1.887 | 3.696 | 0.287 | -0.573 |
| 2016 | 2.804 | 0.420 | 0.029 | 2.899 | 2.829 | 0.394 | 1.674 | 3.635 | -0.564 | -0.475 |
| 2017 | 2.118 | 0.259 | 0.018 | 2.132 | 2.126 | 0.261 | 1.425 | 2.718 | -0.299 | -0.180 |
| 2018 | 3.087 | 0.562 | 0.039 | 2.933 | 3.049 | 0.415 | 2.052 | 4.403 | 0.664 | -0.627 |
| 2019 | 1.790 | 0.894 | 0.062 | 1.880 | 1.819 | 1.020 | -0.046 | 3.302 | -0.331 | -1.058 |
| 2020 | 1.466 | 1.155 | 0.080 | 1.755 | 1.537 | 1.196 | -0.773 | 2.958 | -0.451 | -1.307 |

Table 4 – 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 | 248 | 2.574 | 0.882 | 0.056 | 2.708 | 2.632 | 0.791 | -0.693 | 4.094 | -0.705 |
| 1995 | 236 | 2.898 | 0.816 | 0.053 | 2.996 | 2.957 | 1.028 | 0.000 | 4.248 | -0.636 |
| 1996 | 225 | 2.414 | 1.052 | 0.070 | 2.485 | 2.460 | 1.157 | -0.693 | 4.500 | -0.405 |
| 1997 | 149 | 2.157 | 0.758 | 0.062 | 2.303 | 2.178 | 0.787 | 0.000 | 4.099 | -0.260 |
| 1998 | 128 | 2.088 | 1.092 | 0.097 | 2.255 | 2.114 | 1.079 | -0.693 | 4.248 | -0.203 |
| 1999 | 222 | 2.273 | 1.112 | 0.075 | 2.303 | 2.356 | 1.100 | -0.693 | 4.415 | -0.585 |
| 2000 | 483 | 2.528 | 1.033 | 0.047 | 2.708 | 2.586 | 1.028 | -0.693 | 4.500 | -0.532 |
| 2001 | 1041 | 2.601 | 1.132 | 0.035 | 2.741 | 2.668 | 1.138 | -2.303 | 5.247 | -0.584 |
| 2002 | 1165 | 2.423 | 1.222 | 0.036 | 2.534 | 2.477 | 1.285 | -3.912 | 4.913 | -0.505 |
| 2003 | 1268 | 3.113 | 1.014 | 0.028 | 3.197 | 3.160 | 1.061 | -0.916 | 5.207 | -0.453 |
| 2004 | 1439 | 3.594 | 0.911 | 0.024 | 3.786 | 3.673 | 0.806 | 0.531 | 5.663 | -0.800 |
| 2005 | 1265 | 3.448 | 1.131 | 0.032 | 3.526 | 3.489 | 1.135 | 0.000 | 6.147 | -0.355 |
| 2006 | 1568 | 3.221 | 1.021 | 0.026 | 3.223 | 3.239 | 1.045 | -0.123 | 5.568 | -0.194 |
| 2007 | 1857 | 3.422 | 0.960 | 0.022 | 3.465 | 3.458 | 0.920 | -0.693 | 5.745 | -0.452 |
| 2008 | 1756 | 3.334 | 1.029 | 0.025 | 3.397 | 3.407 | 0.918 | -1.100 | 5.421 | -0.724 |
| 2009 | 1391 | 2.626 | 1.105 | 0.030 | 2.803 | 2.714 | 0.982 | -1.474 | 4.848 | -0.802 |
| 2010 | 1262 | 2.729 | 1.024 | 0.029 | 2.873 | 2.809 | 0.946 | -2.207 | 4.787 | -0.849 |
| 2011 | 1512 | 3.040 | 0.961 | 0.025 | 3.178 | 3.122 | 0.870 | -1.655 | 5.201 | -0.911 |
| 2012 | 1727 | 2.975 | 0.852 | 0.020 | 3.100 | 3.034 | 0.782 | -3.352 | 4.952 | -0.917 |
| 2013 | 1471 | 2.807 | 1.008 | 0.026 | 2.970 | 2.885 | 0.969 | -1.609 | 4.772 | -0.790 |
| 2014 | 1880 | 2.885 | 0.977 | 0.023 | 3.016 | 2.956 | 0.959 | -0.906 | 5.106 | -0.660 |
| 2015 | 978 | 2.345 | 1.111 | 0.036 | 2.459 | 2.406 | 1.086 | -3.963 | 4.881 | -0.720 |
| 2016 | 465 | 2.577 | 1.062 | 0.049 | 2.631 | 2.604 | 1.065 | -0.865 | 4.685 | -0.305 |
| 2017 | 301 | 1.769 | 1.079 | 0.062 | 1.806 | 1.814 | 1.063 | -2.465 | 4.419 | -0.472 |
| 2018 | 148 | 2.581 | 1.257 | 0.103 | 2.553 | 2.605 | 1.335 | -1.343 | 5.161 | -0.205 |
| 2019 | 50 | 2.081 | 1.208 | 0.171 | 2.052 | 2.062 | 1.449 | 0.005 | 4.516 | 0.094 |
| 2020 | 46 | 2.064 | 1.280 | 0.189 | 2.242 | 2.067 | 1.581 | -0.218 | 4.273 | -0.059 |