Exploratory analyses for primary CCSBT CPUE index

Simon Hoyle

Introduction

Indices of southern bluefin tuna abundance are used by the CCSBT in both the CCSBT stock assessment (Butterworth et al., 2003) and the management procedure (CCSBT26-2019). To develop the primary indices used for both purposes, data are fitted using the 'Base' standardization model. Next, two indices are derived from this model using prediction under two alternative weighting schemes, the constant squares and variable squares models. These two indices are then recombined using two alternative weightings into the W0.8 and W0.5 indices (Nishida and Tsuji, 1998; Itoh and Takahashi, 2019).

The Base model used to generate the constant squares (CS) and variable squares (VS) indices is a linear model that includes categorical variables for all spatial and temporal effects, along with three interaction terms:

log(cpue + 0.2) ~ year + month + area + latitude + cpue.bet + cpue.yft + month*area + year*latitude + year*area

In recent years CPUE standardization methods have given more consideration to spatial and temporal correlations (Nishida and Chen, 2004; Chambers, 2014a; Grüss et al., 2019). Many of these methods use the correlations among adjacent areas to estimate parameters more efficiently. Approaches using spatiotemporal smoothers within generalized additive models (GAMs) have been explored for SBT (Chambers, 2013; Chambers, 2014a; Chambers, 2014b), but to date the primary CPUE index has continued to be based on the W0.8 and W0.5 indices.

In 2019 the Base CPUE model produced an index value for 2018 that was identified as anomalous (Itoh and Takahashi, 2019), and the 2020 update generated similarly unrealistic index values for both 2018 and 2019.

This paper explores reasons why the linear model-based index is generating high estimates for recent years and investigates the potential of models that include spatiotemporal smoothers to provide a more reliable index.

These analyses are based on a slightly different dataset from the Base model, since the dataset used in that analysis is not available to CCSBT scientists (Itoh and Takahashi, 2019). The available dataset is sufficiently similar to provide useful insights. The main differences are listed below. We explore consistency between results from the two datasets.

- The primary dataset uses a set of core vessels that have high SBT catches for at least 3 years, whereas the available dataset includes data from all vessels.
- The primary dataset includes catches of bigeye and yellowfin tuna, but the available dataset does not.
- The primary dataset is available at the operational level (but is aggregated for the Base analysis) whereas the available dataset is aggregated.

Methods

Data preparation

We downloaded the file 'CPUEInputs_2020_January.txt' from the CCSBT website. These data are aggregated by year, month, and 5° latitude and longitude, with catches reported by age-class based on spatially and temporally stratified size sampling.

We applied the following processes to the dataset:

- Filter to include effort from 1986 to 2018, with DATA_CODE 'COMBINED', in statistical areas 4 to 9, and months 4 to 9. Include strata with more than 10000 hooks. Include latitudes north of 50°S.
- Create numeric catch variable, the sum of catches of all SBT 4+ and older.
- Create categorical latlong5 variable indicating 5° square that combines latitude and longitude.
- Create categorical s_area variable, which merges statistical area 4 with 5 and statistical area 6 with 7.
- Create categorical variables for year, latitude, longitude, and month.
- Create numeric lon2 variable by adding 360 to all longitudes between -180 and -100, to provide continuity across the spatial domain of the fishery. Longitudes are recorded as -180 to 180 and so the range of the lon2 variable was from -95 to 260.
- Create numeric CPUE variable = catch per 1000 hooks.
- Remove outlier with CPUE > 120.

Characterize data

We explored the data to identify how effort and catches have changed through time.

To explore the spatial distribution of effort through time and by season, we plotted effort by 5° square, and temporally by two-month period, for each 5 years since 1985.

To explore spatial and temporal changes through time, and the possibility that trends have varied by area, we modeled CPUE separately by statistical area, using in each case both a main effects model and a model that included a month by latitude interaction term.

log(cpue + 0.2) ~ year + month + lat + month:lat

Data coverage and Base model estimates

We examined the availability of data, and how data gaps in space and time might affect the Base model. We generated predicted catch rates from the Base model for each combined stratum of year, month, stat area and latitude, by predicting from the parameters of the Base model provided by the analyst (Tomoyuki Itoh, personal communication).

Check relevance of inference from available dataset

Next we standardized the available dataset using a model similar to the Base model, but without the parameters for other species. We call this the simplified Base model.

log(cpue + 0.2) ~ year + month + area + lat + month*area + year*lat + year*area

All parameters are categorical variables.

As for the Base model, we generated predicted catch rates for each combined stratum of year, month, stat area and latitude.

To compare predictions from the Base (with primary dataset) and simplified Base (with available dataset) models, we summed each set of predictions by year, normalized by the means across years, and plotted them on the same figure.

Changes to the Base model

We explored the fit of the simplified Base model and alternative approaches in the categorical data paradigm.

The simplified Base model includes three of the six possible two-way interactions. We examined the effects on model fit of including additional two-way and three-way interactions.

Extreme prediction diagnostic

A new diagnostic of prediction reliability was developed based on the consistency of model predictions by year-month-area-latitude stratum with the range of observed CPUE throughout the time series. The max diagnostic counts the number of predictions higher than the maximum observed CPUE in the same year, and the min diagnostic counts the number of observations lower than the minimum observed CPUE in the same year. These extreme counts are also classified based on whether there is an observed CPUE in the same stratum, and whether the stratum is included in results that are based on the x15 criterion.

The preferred version of the extreme prediction diagnostic is based on high values, x15 filtering, and all strata. High extreme predictions tend to be much more variable and influential than low values when errors are lognormally distributed. The x15 filtered values are used to generate the index, so are more relevant than the full range of cells. Estimates for gap-only strata are provided for interest, but all strata are used to generate the index.

Spatiotemporal smoothers

With data increasingly sparse both spatially and seasonally, spatiotemporal smoothers allow for correlations between adjacent spatial and temporal cells, and reduce the number of parameters being estimated.

Initially we replicated the approach used for the simplified Base models, and then explored approaches that would explain more variability while retaining parameter identifiability.

Modeling was carried out in R (R Core Team, 2019) using the generalized additive modelling package mgcv (Wood, 2011).

All models were fitted to the same dataset. Each variable was modelled either as a categorical (factor) variable, as in the simplified Base model, or as a continuous variable, which is necessary when using smoothers. Factors are indicated here with the suffix *f*.

Each model was fitted with the mgcv setting ‘gamma = 2’ to reduce the effective sample size. CPUE standardization models in mgcv can suffer from excess variability in the smoothers. Fishery data are often over-dispersed due to dependencies among sets and strata, since individual data represent the combination of multiple sets by the same vessels, and unmodelled effects such as environmental patterns and fish behavior can lead to overdispersion at multiple spatial and temporal scales. These dependencies reduce the amount of independent information in the data, which can be allowed for by reducing the effective sample size. Setting gamma to 2 was an ad hoc choice and further exploration is warranted.

Initially the simplified Base was replicated in mgcv.

glm1: <- gam: log(cpue + 0.2) ~ yf + mf + latf + sf + mf:sf + yf:latf + yf:sf

glm2: <- gam: log(cpue + 0.2) ~ yf + mf + latf + sf + mf:sf + yf:latf + yf:sf + mf:latf

Next we explored various smoothers (tensor splines), which are listed in Table 5. Changes include replacing the latitude factor with a two-dimensional spatial smoother, and adding various two-way and three-way interactions. A four-way interaction with all the continuous variables was also explored. The model gam13 was designed to be a smoothed version of model 'Base Plus', replacing area with te(longitude, latitude), and with longitude in interaction terms.

Model assumptions were checked using the gam.check() function or the equivalent check.gamViz function from the mgcViz package.

Reasons for skewed residuals were checked by examining the relationship of effort to residual size, using a smoothing spline in mgcv: gam(N_HOOKS ~ s(residual)).

Individual smoothers were plotted from the model with the best combination of low AIC and lack of extreme / unrealistic predictions.

Indices from each model were obtained by predicting catch rates in all spatial cells that were fished in at least 15 temporal strata (known as 'x15' filtering). Catch rates in these spatial cells were predicted for all years and months. Based on the unreliable assumption that all spatial cells have the same ocean area, catch rates were summed for each year and divided through by the mean of the yearly estimates, to give an index with mean of 1. Since ocean areas change with latitude, and some spatial cells include land, abundance prediction methods will in future need to be correspondingly adjusted.

Distribution checking

Fits to the following distributions were checked by running the simplified Base gam model with the following distribution assumptions and examining QQ plots using the gam.check() function.

1. Lognormal(cpue + 0.2)
2. Gamma(cpue + 0.2)
3. quasiPoisson(catch)
4. Poisson(catch)
5. Binomial(catch > 0)
6. Lognormal (nonzero catch)
7. Gamma(nonzero catch)
8. Tweedie(catch)

Results

Data characterization

The area with most effort (hooks set) was area 9, while areas 5 and 6 had relatively little effort. Areas 4, 7 and 8 had intermediate effort levels, with approximately 200 million hooks set from 1986 to 2018 (Figure 1).

Areas 4 and 5 had a relatively high proportion of records with zero catch, while areas 6, 7, and 9 had a very low proportion of zeroes. High proportion of zeroes in areas 4 and 5 particularly. The proportions of records with zero catch did not change substantially through time, though there was a small peak in the early 1990s.

By latitude, most of the effort was between 35° and 45° S. Proportions of zeroes were much higher further north, with over half the strata reporting zero SBT north of 35°S.

Effort by month was highest in May, June and July. Proportions of zeroes were highest in April and declined through the season.

The spatial distribution of effort changed in consistent ways through the year (Figure 2). In April and May there was effort in the west (area 9) and east (areas 4, 5, 6, 7) but little effort in area 8. Later in the year the effort would move north within these regions and also into area 8. Similar seasonal fishing patterns between areas 8 and 9 have been seen for the Korean fleet (Figure 3). If these changes in effort distribution reflect SBT catch rates, they suggest interactions between month and latitude, and between month and longitude / area.

There were also some long-term changes in effort, with less effort in all areas during the 2015-2018 period, apart from an increase in area 8 in April-May.

Results of standardizing catch rates separately by statistical area showed catch rates in area 67 increasing rapidly from 2007 and stabilizing at a high level in 2010, and catch rates in area 9 increasing rapidly from about 2008 to a peak in 2015 (Figure 4). Catch rates in area 45 showed no consistent change through time but did spike up in 2018. Catch rates in area 8 were variable throughout but have been higher on average since 2010 than in the previous 10 years. Including a month:latitude interaction had little effect on most areas but substantially raised standardized catch rates in area 8 since 2010.

Data coverage and Base model estimates

The number of empty strata (by year, statistical area, month and latitude) increased progressively from 1990 to 2018 (Table 1). Predicted CPUE in most years was reasonably consistent with the observed CPUE in the same strata, apart from a few negative values due to low predictions and the back-transformation exp(pred) – 0.2.

The increasing number of empty strata is consistent with Figure 1b from Itoh and Takahashi (2019), which shows operations with SBT aged 4+ concentrating into a steadily reducing number of 5° cells and 1° cells through time, with minima for both 1° and 5° cells in 2018 (Figure 5). At the same time the effort per cell has increased since 2010, so that effort in 2018 was perhaps twice as concentrated as in 2010. An even stronger pattern of increasingly concentrated effort is apparent for Korea.

The Base model (predictions from parameters provided by Itoh-san) predicted very high catches per thousand hooks in 2018 in area 8, with the highest predictions all in strata with no reported effort (Table 1). Of the 18 month-latitude strata in area 8, over 40 fish per thousand hooks were predicted in four strata and almost 20 fish in two strata.

The Base model (predictions from parameters provided by Itoh-san) and the simplified Base model gave very similar indices (Figure 6), suggesting that approaches that improve results for the simplified Base model are also likely to work for the Base model. The primary and 'available' datasets are sufficiently similar to allow the available dataset to stand in for the primary dataset for exploratory analyses regarding some issues.

Changes to the simplified Base model

An interaction term not included in the simplified Base model (month * latitude) had the most effect on the AIC of all interactions (Table 3) so we included it in some further analyses, and suggest that this interaction should be considered in other analyses. This interaction term is consistent with the

observed movement patterns of the fleet, which fishes further north within each region later in the season (Figure 2).

The other two interaction terms not included in the Base model (year * month and area * latitude) improved the AIC but were ranked 5th and 6th among all possible two-way interactions.

Three of the four three-way interaction terms improved the AIC of the model, with the most impact coming from month*lat*area (Table 4).

These trials suggest that including additional terms in the simplified Base model is likely to improve the fit to the data. The proportion of deviance explained should also be considered when adding terms, because most CPUE datasets are overdispersed which makes AIC oversensitive and can lead to overfitting. Prediction reliability must also be considered when including additional terms, given the increasing concentration of the effort and consequent sparse data.

Spatiotemporal smoothers

All models with spatiotemporal smoothers fitted the data with lower AIC values than the simplified Base model (Table 5). Where checked, most of the smooth terms explained over 1% of the deviance (Table 6).

The model fits according to AIC were (best to worst) gam12, gam11, gam9, gam8, gam7, gam6, gam5, gam13, gam10, gam4, gam3, gam2, Base plus, simplified Base, glmm_YrArea, Base_noYrArea (Table 5). Given these results the models gam2 to gam8 were dropped so as to focus on the best-fitting gams. The model Base plus was also dropped to focus on models more relevant to the discussion, but it was nevertheless the best-fitting factor-based model.

The extreme prediction diagnostic based on high values showed the best performance across all cells for model gam11, followed by model gam13, gam9, Base_noYearArea, glmm_YearArea, gam12, simplified Base, and gam10. After removing cells not included in index prediction due to x15 filtering, the sequence from best to worst became gam9, gam11, Base_noYearArea, gam13, glmm_YearArea, gam10, gam12, simplified Base.

Given their worse performance with extreme predictions, models gam10 and gam12 were dropped.

Residual diagnostic plots showed relatively normal distributions, although the tails of the gams with better fit to the data did not follow the expected distribution in the tails (Figure 7). This appears to be because the strata of aggregated data with lower sample sizes (fewer hooks set) are more variable than the strata with more effort, as demonstrated by fitting a gam to the relationship between residual size (x-axis) and effort (y-axis) (Figure 8).

Spatiotemporal smoothers from model gam11 showed relatively smooth catch rate patterns across space, but it is difficult to interpret individual plots in biological terms given their interactions, and the fact that the overall effect is the aggregate of all components (Figure 9, Figure 10, and Figure 11).

Predictions from model gam11 across the spatial domain showed changing spatial distribution by month and year (Figure 12).

Indices from the 3 smoothed models were relatively similar to the factor models for most of the period, but a little less variable through time, and lacking the anomalously high value for 2018 that motivated this study (Figure 13).

Predictions by latitude and by statistical area were much more stable than those from the factor-based models (Figure 15). High variability associated with sparse data was apparent at the northern

and southern latitudes -32.5 and -47.5, and in statistical areas 8, 45, and 67. Note however that areas with sparse data are given low weight by the constant squares and variable squares algorithm and by x15 filtering, which greatly limits their impact on the resulting indices. These figures nevertheless illustrate how spatiotemporal smoothing can stabilize estimates when data are sparse.

Distribution assumptions

Probably the best QQ plot was obtained with Tweedie distribution, p=1.5 (Figure 15).

Discussion

This study has shown that models that use spatiotemporal smoothers can address problems with sparse data caused by increasing effort concentration. Such models may also provide more accurate indices of abundance by taking into account factors not included in the current model.

Further work is needed to improve these preliminary gam models, which can be done with both the primary dataset and with the dataset available here. Issues to consider include allowing for the different ocean areas of the spatial cells, and examining how different data weighting/filtering methods affect results. Possible data filtering changes include adjusting the required number of records per stratum from 15, and changing the stratification in the filter to lat-long-month rather than lat-long. Alternative values of the 'gamma' adjustment to effective sample size should be considered, as should alternative initial smoothness values (assigned with the k parameter) in each of the tensor spline smoothers.

Other and likely more important and influential issues include (a non-exhaustive list) the use of hurdle or zero-inflated models to deal with zero-catch strata instead of adding a constant; consideration of vessel-specific fishing power which in many fisheries varies considerably and tends to increase through time; adjusting for targeting; and the effects of quotas on vessel behavior and catch rates, particularly within-season.

Given that sparse data is causing problems, including data from other fleets is likely to be helpful if it helps to fill in some of the gaps.

These gam models used tensor splines (te) to fit smoothers, but mgcv provides various other options that should be explored. Interaction-only tensor splines (ti) were explored but resulted in many more extreme predictions.

The increasing concentration of fishing effort is very marked in both the Japanese and Korean fleets. As well as causing the analysis problems that motivated this study, it may affect the reliability of CPUE as an index of abundance in ways that spatiotemporal smoothing does not resolve. The reasons for this change, and their implications, need to be understood. If effort is concentrating because the fleet is getting better at finding fish, this would also tend to increase the average observed catch rate. In this situation the CS method and the gam with spatial smoothers may produce a hyperstable CPUE index. It may be appropriate to retain an approach that includes aspects of the VS method.

There may be concerns about changing stock distribution, e.g. associated with climate change. Contraction and expansion would be easier to detect if modelling at 1° cell scale, which would be straightforward with smoothers. Currently there are only 4 x 5° latitude bands in the model, and distribution change must be large to be detectable. It would also be easier to detect distribution change if data were available from more vessels, i.e. by including other fleets in the analysis, and

perhaps including non-core vessels. These approaches would make it even more important to consider target change and vessel-specific catchability.

The residual distributions were distinctly non-normal, perhaps mostly because the dataset was aggregated and catch rates for strata with low effort were more variable than those with more effort. This is potentially problematic because effort is likely to be higher in strata with higher catch rates. Issues like this are complex, and simulation may be the best approach for addressing them.

Preliminary analysis and QQ plots suggested that the Tweedie model may give the best fit to the observed residual distributions. The fit of the lognormal (cpue + 0.2) model was reasonable for the Base model but less so for gams with spatiotemporal smoothers.

Using spatial smoothers in gams with mgcv has successfully addressed problems due to sparse data. Gams are very efficient for data exploration and applying a variety of statistical methods but are only one of the potential approaches available. A recent comparison of standardization methods found that the VAST (Thorson et al., 2015) performed slightly better than gam-based approaches (Grüss et al., 2019). VAST also has potential to include multiple categories in a model, so can model size and catch rate data jointly (e.g. Maunder et al., 2020). This approach has potential to avoid the age slicing currently used to generate the 4+ dataset, which introduces some error.

Acknowledgments

Fisheries New Zealand for funding this study. Darcy Webber, Tomoyuki Itoh, Jim Ianelli, Ana Parma, Doug Butterworth and the CPUE working group for ideas, discussions, and sharing analysis results.

References

Butterworth, D.; Ianelli, J.; Hilborn, R., 2003. A statistical model for stock assessment of southern bluefin tuna with temporal changes in selectivity. African Journal of Marine Science. 25, 331-361.

Chambers, M., 2013. A generalised additive model for southern bluefin tuna catch per unit effort (CPUE). CCSBT-ESC/1309/13 (Rev. 1). CCSBT Extended Scientific Committee, Canberra, Australia

Chambers, M., 2014a. A CPUE index based on a GAMM. CCSBT-ESC/1409/09. CCSBT Extended Scientific Committee, Auckland, New Zealand

Chambers, M., 2014b. A CPUE model with interactions as random effects. CCSBT-ESC/1409/10. CCSBT Extended Scientific Committee, Auckland, New Zealand

Grüss, A.; Walter III, J.F.; Babcock, E.A.; Forrestal, F.C.; Thorson, J.T.; Lauretta, M.V.; Schirripa, M.J., 2019. Evaluation of the impacts of different treatments of spatio-temporal variation in catch-per-unit-effort standardization models. Fish. Res. 213, 75-93.

Itoh, T.; Takahashi, N., 2019. Update of the core vessel data and CPUE for southern bluefin tuna in 2019. CCSBT Extended Scientific Committee: Commission for the Conservation of Southern Bluefin Tuna: CCSBT-ESC/1909/BGD 05.

Maunder, M.N.; Thorson, J.T.; Xu, H.; Oliveros-Ramos, R.; Hoyle, S.D.; Tremblay-Boyer, L.; Lee, H.H.; Kai, M.; Chang, S.-K.; Kitakado, T., 2020. The need for spatio-temporal modeling to determine catch-per-unit effort based indices of abundance and associated composition data for inclusion in stock assessment models. Fish. Res. 229, 105594.

Nishida, T.; Chen, D.-G., 2004. Incorporating spatial autocorrelation into the general linear model with an application to the yellowfin tuna (Thunnus albacares) longline CPUE data. Fish. Res. 70, 265-274.

Nishida, T.; Tsuji, S., 1998. Estimation of abundance indices of southern bluefin tuna (*Thunnus maccoyii*) based on the coarse scale Japanese longline fisheries data (1969-97). Fourth CCSBT Scientific Meeting. Shimizu, Shizuoka, Japan

R Core Team, 2019. R: A language and environment for statistical computing, version 3.3.1. Vienna, Austria: R Foundation for Statistical Computing.

Thorson, J.T.; Shelton, A.O.; Ward, E.J.; Skaug, H.J., 2015. Geostatistical delta-generalized linear mixed models improve precision for estimated abundance indices for West Coast groundfishes. ICES J. Mar. Sci. 72, 1297-1310.

Wood, S.N., 2011. Fast stable restricted maximum likelihood and marginal likelihood estimation of semiparametric generalized linear models. J. R. Stat. Soc. B. 73, 3-36.

Tables

Table 1:Observed numbers of hooks and catch rates in the aggregated Japanese dataset, and approximate predicted catch rates in the same strata generated by the Base model. Hooks are in thousands and catch rates in SBT per thousand hooks.

			1990			2010			2017			2018		
stat area	**mon**	**lat 5**	**hooks**	**cpue**	**Pred cpue**	**hooks**	**cpue**	**Pred cpue**	**hooks**	**cpue**	**Pred cpue**	**hooks**	**cpue**	**Pred cpue**
8	4	-45	0		0.14	0		0.54	0		1.01	0		5.17
8	4	-40	0		0.10	0		0.63	0		1.08	0		5.55
8	4	-35	0		0.03	196	0.46	0.20	879	0.01	0.02	534	0.06	0.30
8	5	-45	0		0.35	0		0.98	0		1.73	0		8.35
8	5	-40	0		0.29	0		1.12	0		1.84	0		8.96
8	5	-35	0		0.17	427	0.14	0.44	545	0.04	0.16	226	0.25	0.60
8	6	-45	0		0.99	0		2.37	0		4.00	0		18.39
8	6	-40	245	1.70	0.85	0		2.66	0		4.23	0		19.71
8	6	-35	0		0.60	304	1.38	1.19	17	0.00	0.58	0		1.54
8	7	-45	0		2.38	0		5.37	0		8.91	0		40.16
8	7	-40	2211	1.79	2.09	0		6.02	0		9.42	0		43.03
8	7	-35	430	1.68	1.53	0		2.81	35	2.64	1.48	0		3.57
8	8	-45	0		2.73	0		6.14	0		10.17	0		45.73
8	8	-40	979	2.55	2.40	186	4.27	6.88	0		10.75	0		49.00
8	8	-35	1113	2.16	1.77	663	3.49	3.23	1841	5.27	1.72	1857	7.19	4.09
8	9	-45	0		2.04	0		4.64	0		7.71	0		34.83
8	9	-40	33	1.99	1.79	0		5.20	0		8.15	0		37.32
8	9	-35	101	0.79	1.30	43	4.08	2.41	476	8.52	1.26	544	7.96	3.07
9	4	-45	258	2.90	1.42	0		2.22	0		5.18	77	5.64	5.24
9	4	-40	3418	2.19	1.24	799	4.16	2.50	1201	8.07	5.48	1603	6.69	5.63
9	4	-35	491	0.00	0.89	202	0.81	1.11	0		0.79	20	0.00	0.31
9	5	-45	1082	1.96	1.79	0		2.76	25	7.39	6.38	0		6.45
9	5	-40	4075	1.44	1.56	1202	3.51	3.10	2070	6.35	6.74	2741	6.38	6.93
9	5	-35	279	0.27	1.13	282	1.88	1.40	23	2.39	1.02	26	0.00	0.42
9	6	-45	205	1.42	1.96	0		3.02	0		6.94	0		7.02
9	6	-40	4167	1.54	1.71	870	4.56	3.39	1171	8.62	7.34	1572	9.21	7.54
9	6	-35	1555	0.62	1.25	602	1.48	1.54	0		1.12	0		0.48
9	7	-45	21	1.53	2.53	0		3.87	0		8.83	0		8.94
9	7	-40	1746	1.31	2.22	366	5.37	4.34	433	8.67	9.34	381	10.97	9.59
9	7	-35	3952	0.98	1.63	267	2.35	2.00	0		1.47	0		0.65
9	8	-45	0		2.79	0		4.27	0		9.71	0		9.83
9	8	-40	0		2.46	17	1.22	4.78	27	7.18	10.26	0		10.54
9	8	-35	1500	1.10	1.81	392	2.13	2.21	0		1.63	0		0.74
9	9	-45	0		2.32	0		3.56	0		8.16	0		8.25
9	9	-40	0		2.04	0		4.00	0		8.62	0		8.86
9	9	-35	17	0.00	1.49	51	3.08	1.83	0		1.34	0		0.59
45	4	-35	49	0.00	0.23	313	0.09	0.39	16	0.36	0.30	0		1.35
45	4	-30	16	0.00	-0.06	45	0.00	-0.12	0		-0.13	0		-0.13
45	5	-35	149	0.00	0.93	1133	3.02	1.34	294	4.32	1.12	234	5.77	3.89
45	5	-30	98	0.00	0.16	11	0.00	0.01	0		-0.01	0		-0.02
45	6	-35	3346	1.18	1.15	507	8.63	1.65	468	2.97	1.38	718	6.61	4.70
45	6	-30	746	0.01	0.23	340	0.00	0.05	144	0.00	0.02	147	0.00	0.02
45	7	-35	3610	2.17	1.87	0		2.64	0		2.22	0		7.32
45	7	-30	2964	0.40	0.46	311	0.00	0.18	339	0.03	0.14	524	0.00	0.14
45	8	-35	315	1.19	1.97	0		2.78	0		2.34	0		7.69
45	8	-30	1699	0.25	0.49	137	0.04	0.20	364	0.10	0.16	419	0.00	0.15
45	9	-35	0		2.44	0		3.42	0		2.89	0		9.40
45	9	-30	0		0.64	88	0.05	0.29	25	0.00	0.24	0		0.23
67	4	-45	1795	2.62	1.95	0		6.00	0		5.61	0		6.02
67	4	-40	3644	2.68	1.71	969	5.57	6.72	2088	5.02	5.93	2005	4.38	6.46

			1990			2010			2017			2018		
stat area	**mon**	**lat 5**	**hooks**	**cpue**	**Pred cpue**	**hooks**	**cpue**	**Pred cpue**	**hooks**	**cpue**	**Pred cpue**	**hooks**	**cpue**	**Pred cpue**
67	4	-35	0		1.24	0		3.15	0		0.87	0		0.38
67	5	-45	1288	2.22	2.23	167	6.79	6.79	0		6.35	0		6.80
67	5	-40	5918	1.68	1.96	422	9.69	7.60	1983	5.09	6.71	2094	6.93	7.30
67	5	-35	0		1.43	0		3.58	0		1.01	0		0.45
67	6	-45	1163	1.24	2.22	0		6.77	0		6.34	0		6.79
67	6	-40	2950	1.51	1.95	96	5.55	7.58	334	10.15	6.70	340	9.65	7.29
67	6	-35	0		1.42	0		3.57	0		1.01	0		0.45
67	7	-45	0		2.32	0		7.06	0		6.61	0		7.08
67	7	-40	124	1.53	2.04	0		7.90	0		6.99	0		7.60
67	7	-35	0		1.49	0		3.72	0		1.06	0		0.48
67	8	-45	0		1.50	0		4.70	0		4.39	0		4.71
67	8	-40	29	1.05	1.31	0		5.26	0		4.64	0		5.06
67	8	-35	0		0.94	0		2.45	0		0.65	0		0.26
67	9	-45	0		1.29	0		4.10	0		3.83	0		4.11
67	9	-40	0		1.13	0		4.60	0		4.05	0		4.42
67	9	-35	0		0.80	0		2.12	0		0.54	0		0.20

Table 2: Changes in deviance, degrees of freedom and AIC as a result of dropping parameters from the Base_simplified model.

`log(cpue + 0.2) ~ yf + mf + sf + latf + mf:sf + yf:latf + yf:sf + mf:latf.`

	Df	Deviance	AIC
<none>		2049	9530.7
mf:sf	15	2137.6	9682.1
yf:latf	96	2245.9	9731.7
yf:sf	96	2258.6	9756
mf:latf	15	2220.7	9845.3

Table 3: Changes in deviance, degrees of freedom and AIC as a result of dropping parameters from the model with all possible two-way interactions.

log(cpue + 0.2) ~ yf + mf + sf + latf + mf * sf + yf * latf + yf * sf + mf * latf + yf * mf + sf * latf

	Df	Deviance	AIC
<none>		1859.6	9443.3
mf:sf	15	1909.7	9527.2
yf:latf	96	2030.4	9627.6
yf:sf	96	2070.5	9711.4
mf:latf	15	2000.1	9725.3
yf:mf	160	2014.4	9465.7
sf:latf	4	1891.5	9508.1

Table 4: Changes in degrees of freedom and AIC as a result of adding three-way interactions to the following model.

log(cpue + 0.2) ~ yf + mf + sf + latf + mf * sf + yf * latf + yf * sf + mf * latf

Interaction term	df	AIC
-	267	9530.748
month*lat*area	285	9379.198
year*month*area	764	9453.124
year*lat*area	332	9461.132
year*month*lat	753	9690.653

Table 5: Models run using mgcv. The factors column reports all variables included as categorical variables. Smooth terms include two-way, three-way and four-way ('All') interactions. The last two columns show the estimated degrees of freedom (df) and the AIC.

Label	**Factors**	**Smooth terms**								**4-way**	**df**	**AIC**
Base											252	9845.3
Base plus	.+mf:latf										267	9530.7
Base_noYrAr	.-yf:area										156	10061.8
glmmYrAr											199.1	9893.7
gam 2	yf+mf	lon,lat									95.1	9522.9
gam 3	yf+mf							mn,lon,lat			152.4	8562.3
gam 4	yf									All	322.1	8064.9
gam 5	yf+ll									All	297.7	7886.2
gam 6	yf+ll		mn,lat							All	274.4	7856.4
gam 7	yf	lon,lat	mn,lat							All	220.8	7839.1
gam 8	yf	lon,lat	mn,lat	lon,mn						All	248.9	7758.5
gam 9	yf	lon,lat	mn,lat	lon,mn	yr,lat			mn,lon,lat	lat,lon,yr		262.2	7701.7
gam 10	yf	lon,lat	mn,lat	lon,mn	yr,lat	yr,lon	yr, mn				195.8	8048.5
gam 11	yf	lon,lat	mn,lat	lon,mn	yr,lat	yr,mn		lat,lon,mn	lat,lon,yr		267.1	7593.4
gam 12	yf	lon,lat	mn,lat	lon,mn	yr,lat	yr,lon	yr,mn	lat,lon,mn	lat,lon,yr		283.8	7505.7
gam 13	yf+mf	lon,lat	mn,lat	lon,mn	yr,lat	yr,lon					217.8	7968.3

Table 6: For each three-way smooth term in model gam 11, the percentage of deviance explained, and the effect on the AIC of dropping the term from the model. The model is specified as:
log(cpue + 0.2) ~ yf + te(lon2, LAT, k = c(40,4)) + te(MONTH, LAT, k = c(6,4)) + te(lon2, MONTH, k = c(10, 5)) + te(YEAR, LAT, k = c(20, 4)) + te(YEAR, MONTH, k = c(20, 5)) + te(LAT, lon2, MONTH, k = c(4,15, 6)) + te(LAT, lon2, YEAR, k = c(4,10, 9))

AIC	delta AIC	Change in % Deviance explained	Smoother dropped
7593.4	0.0		
8048.2	454.8	3.4	te(LAT, lon2, MONTH, k = c(4,15, 6))
7859.9	266.5	1.8	te(LAT, lon2, YEAR, k = c(4,10, 9))

Table 7: Extreme prediction diagnostic showing the number of year-month-latitude-area stratum predictions that either exceed the maximum observed values in any stratum per year (max) or predict less than zero (min). Results are reported for cell filtering methods that include all cells (full) or only those with at least 15 records observed (x15). Results are also reported both for all strata (all) and for only those strata with no observed CPUE in the year of the prediction (gap).

Limit	Cell filter	Strata	Models							
			Simp Base	Base NoYrAr	glmm YA	gam 9	gam 10	gam 11	gam 12	gam 13
max	full	all	34	9	14	4	38	0	19	2
max	full	gap	30	9	13	4	38	0	18	2
max	**x15**	**all**	**19**	**4**	**6**	**1**	**11**	**2**	**12**	**5**
max	x15	gap	15	4	5	0	11	1	11	4
min	full	all	95	92	90	47	36	40	36	56
min	full	gap	40	39	38	21	18	17	15	21
min	x15	all	88	91	88	44	32	39	33	48
min	x15	gap	33	38	36	24	17	19	17	18

Figures

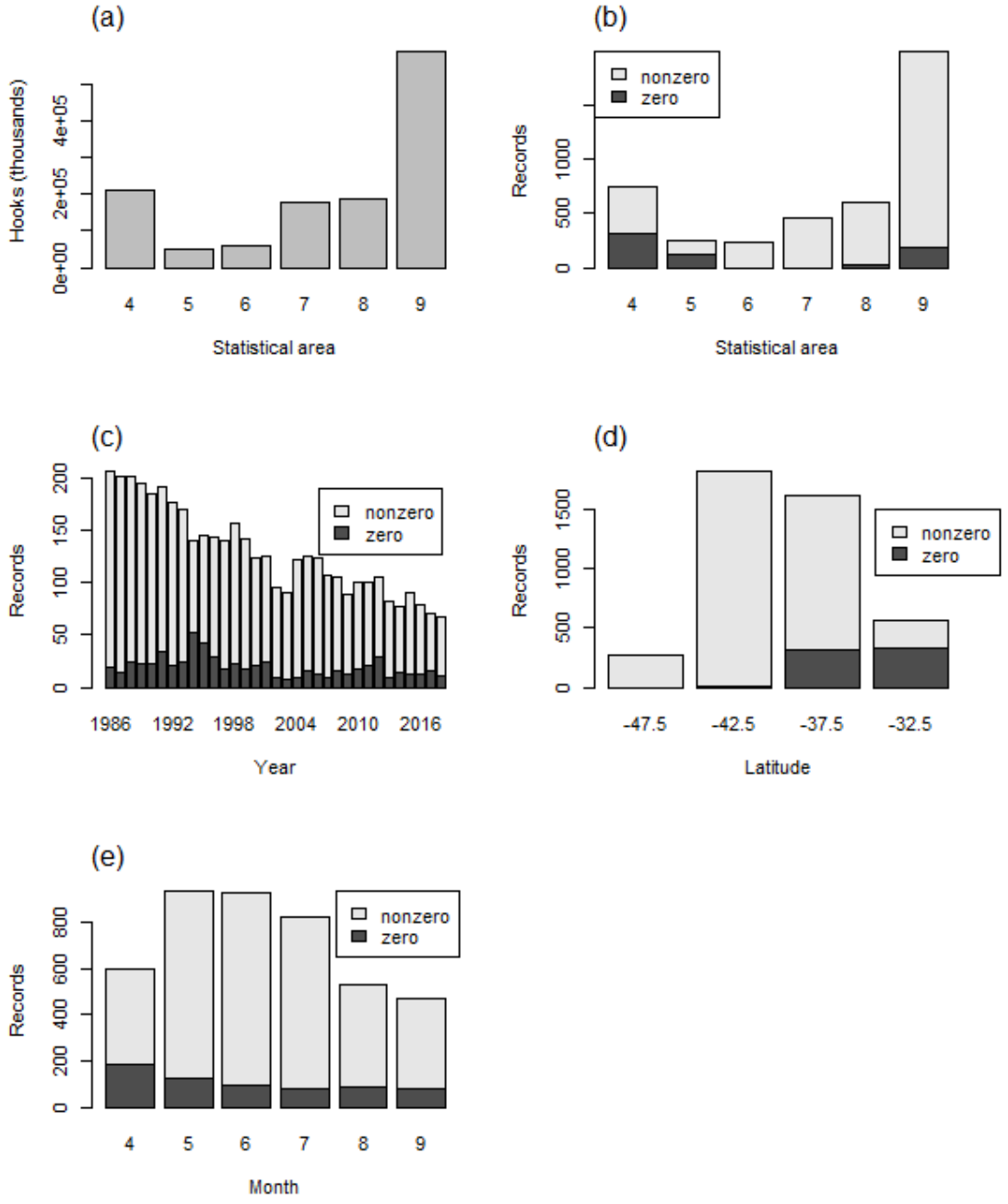

Figure 1: Distribution of effort and data records. (a) Total hooks per statistical area; (b) Total records per statistical area, including both zero and nonzero catches; (c) number of records per year, including both zero and nonzero catches; (d) number of records per latitude band, including both zero and nonzero catches; (e) number of records per month, including both zero and nonzero catches.

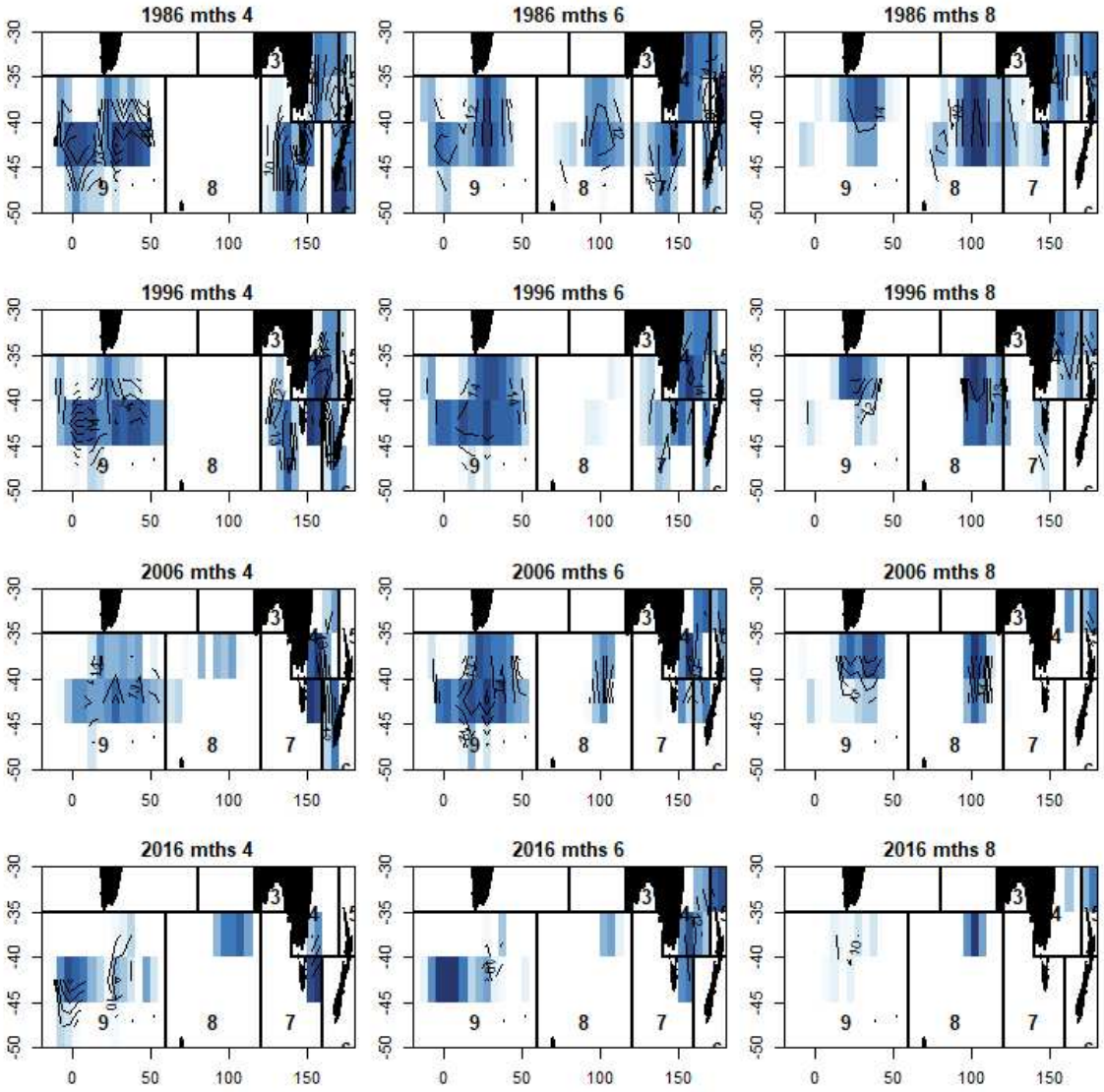

Figure 2: Maps of total effort per two-month period (across), summed over 10-year periods (down) starting in 1986. Darker blue indicates higher effort.

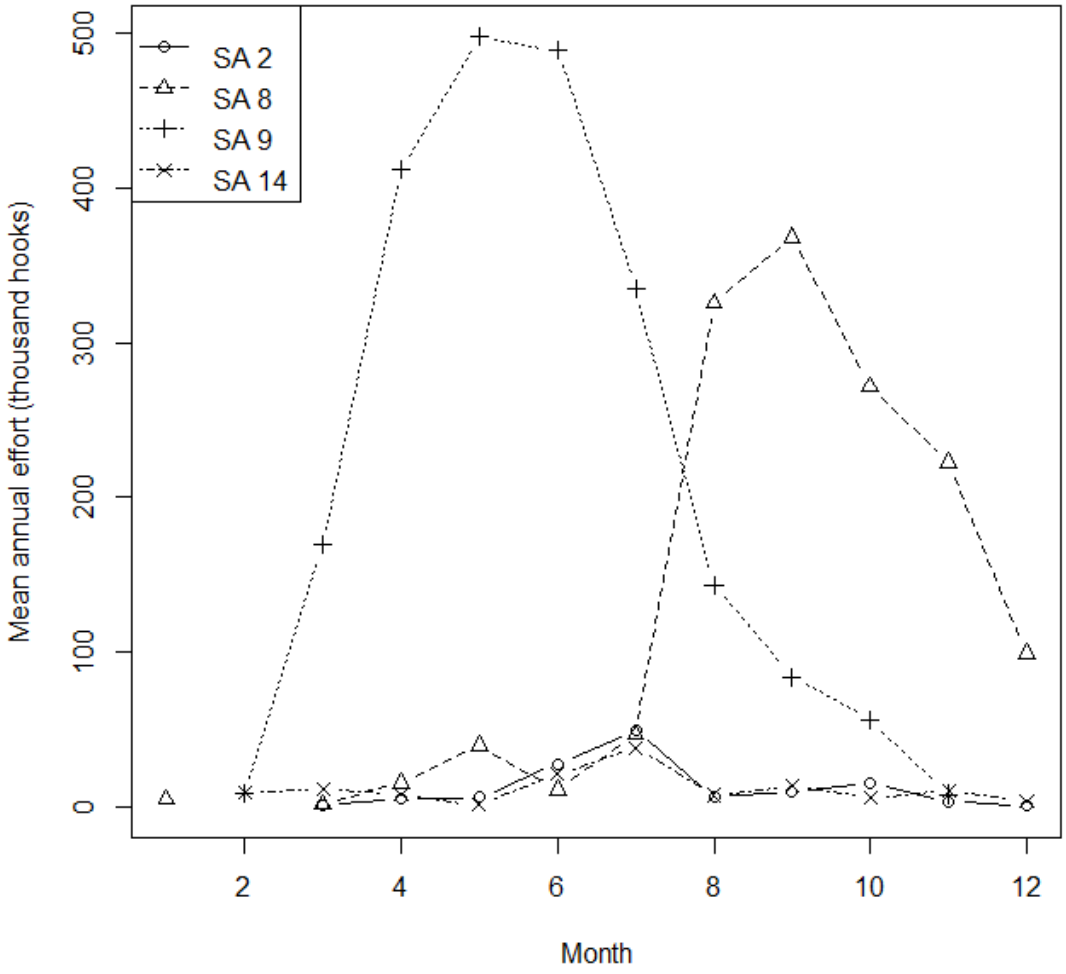

Figure 3: Monthly effort per CCSBT statistical area by the Korean longline fleet, from Hoyle et al 2019.

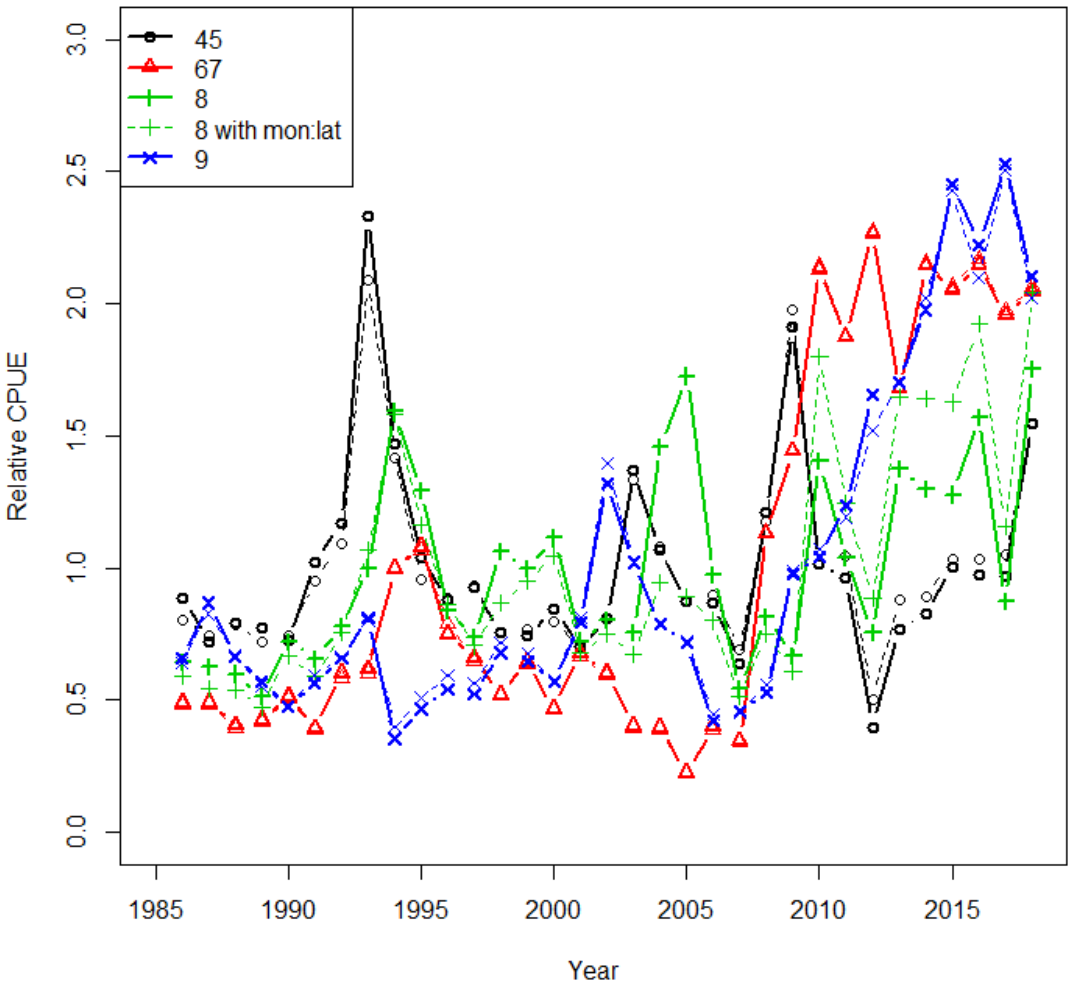

Figure 4: Standardized CPUE indices for individual statistical areas.

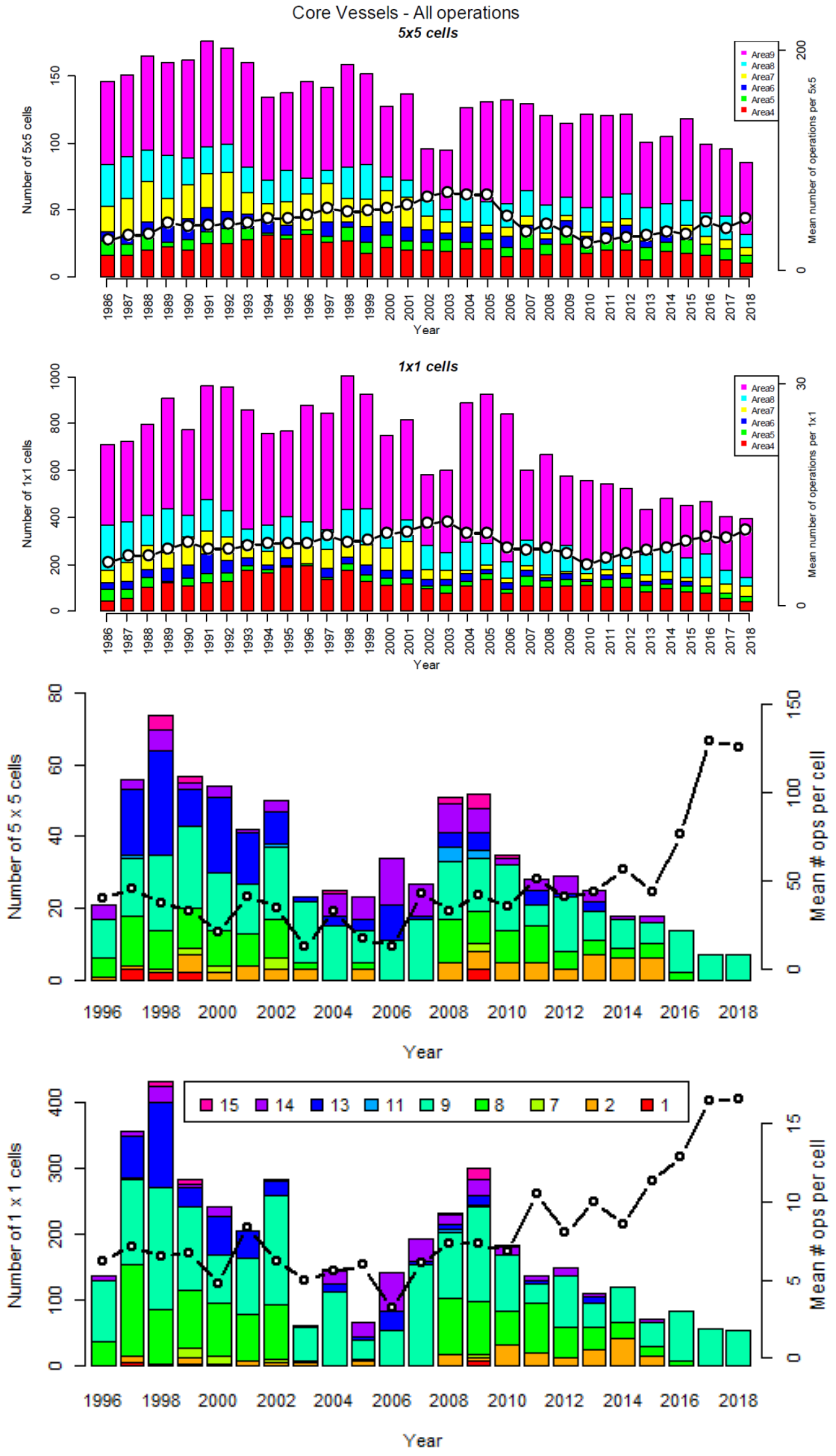

Figure 5: Number of cells fished by Japanese core vessels (above) and Korean vessels (below), copied from CCSBT-ESC/1909/BGD05 and CCSBT-ESC/1909/39. (Upper plot for each fleet) The bars represent the number of major cells (5x5° by month) fished by CCSBT statistical area and year, see left y-axis. The line represents the mean annual operations per cell, see right y-axis. (Lower plot for each fleet) As for upper plot, but with minor cells (1x1° by month) instead of major cells. The colours represent the statistical areas and differ between the Japanese and Korean figures.

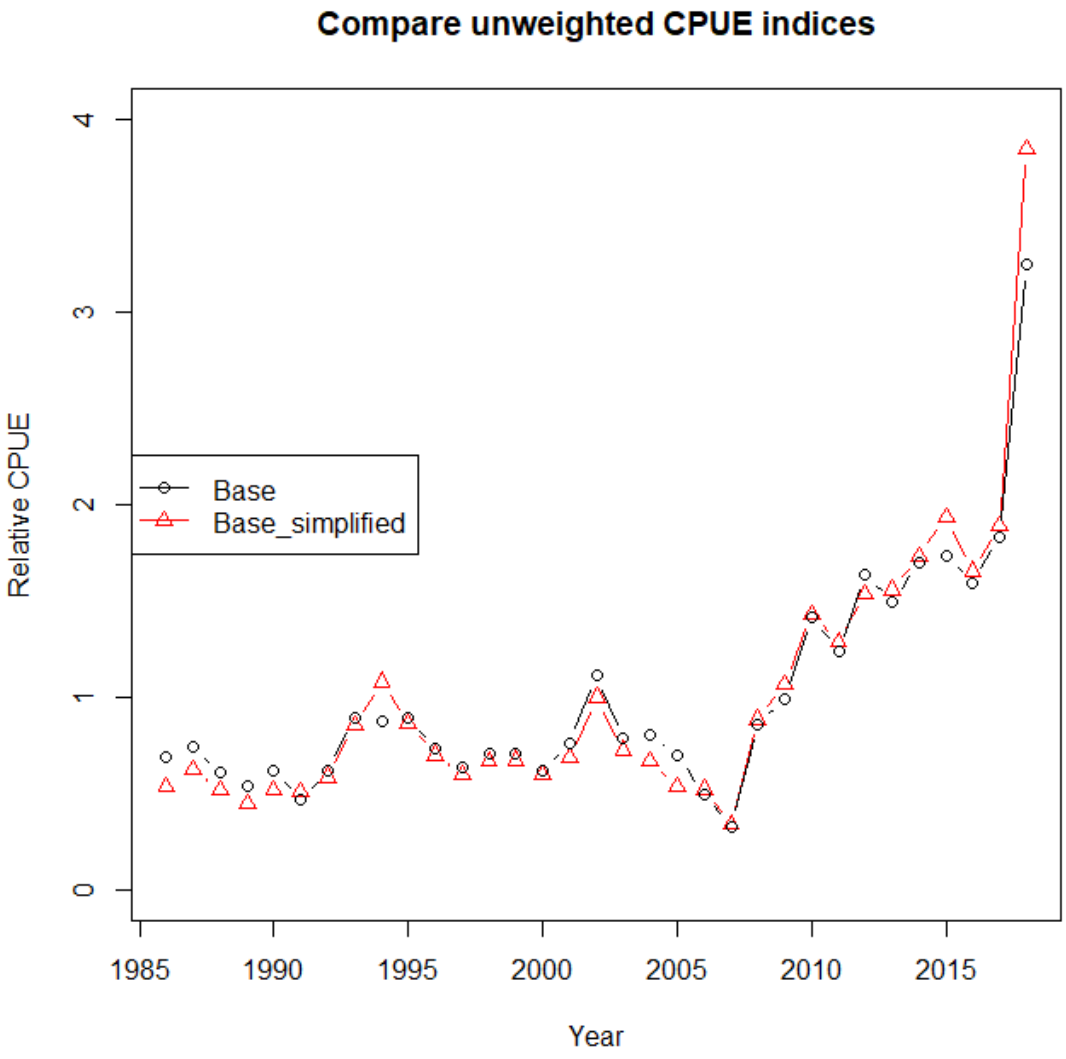

Figure 6: Compare the CPUE indices between the predictions generated from the Base parameters and the predictions from standardized data. In each case the predicted CPUE estimates are summed across all strata without weighting.

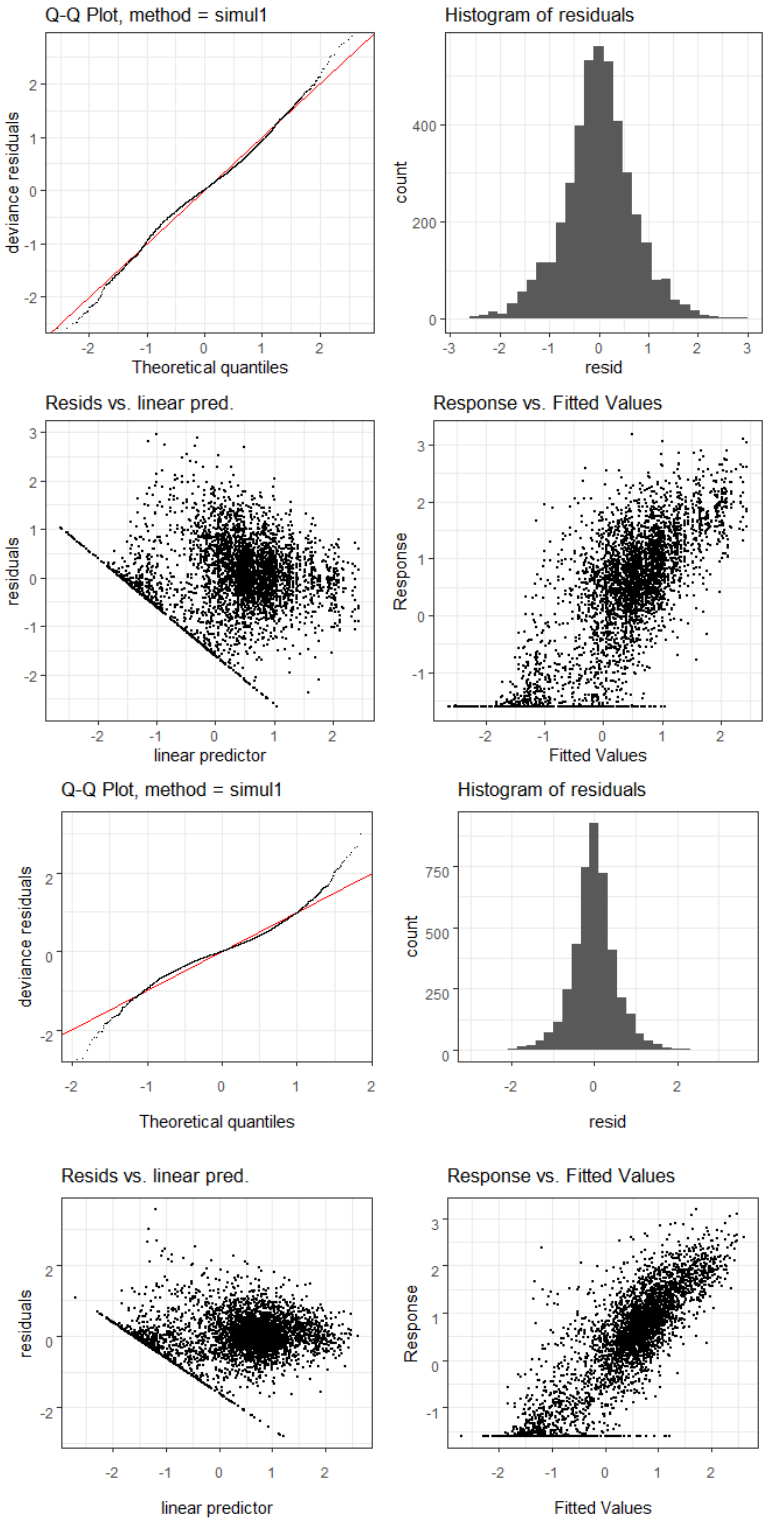

Figure 7: Residual checking plots for the simplified base model (above) and the gam11 model (below).

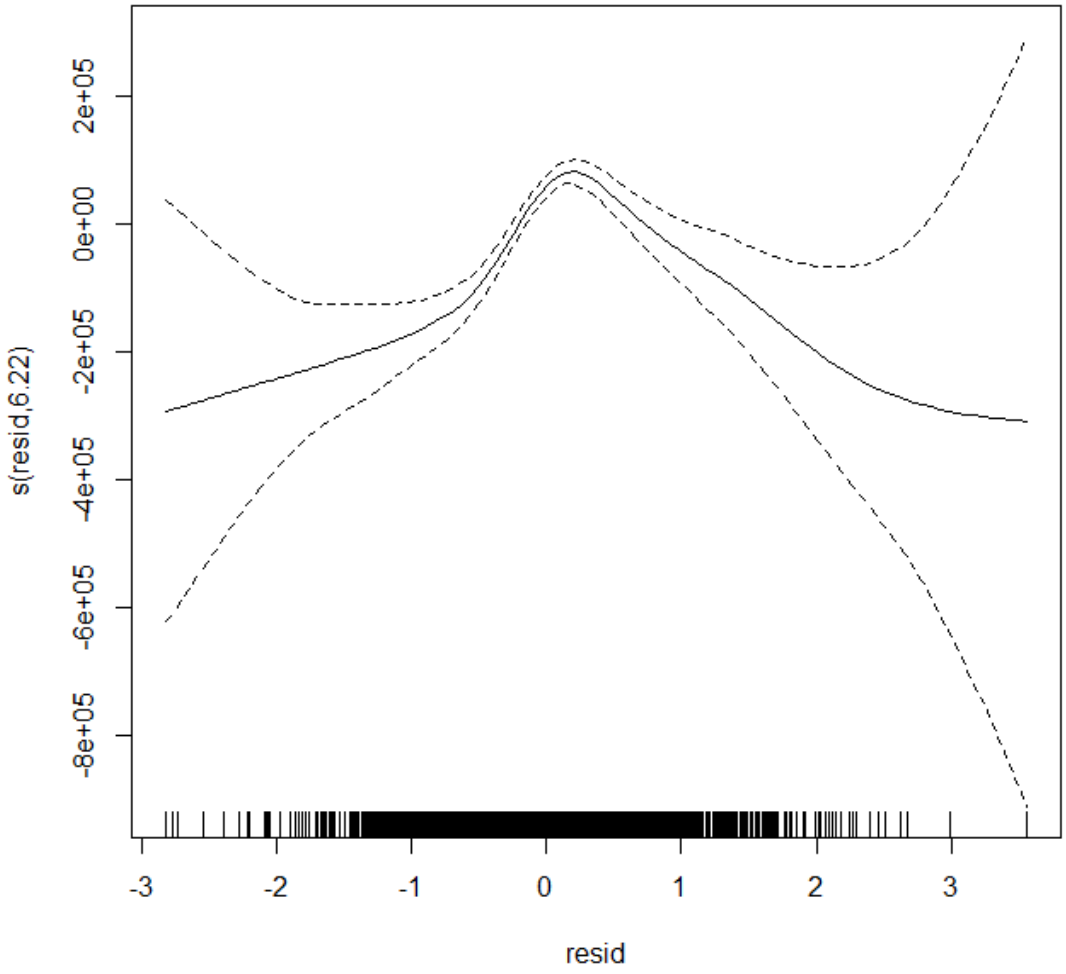

Figure 8: Relationship between residuals from the model gam11 and the effort in the stratum, showing that higher variability (the tails of the residual distribution) is associated with less effort in a stratum.

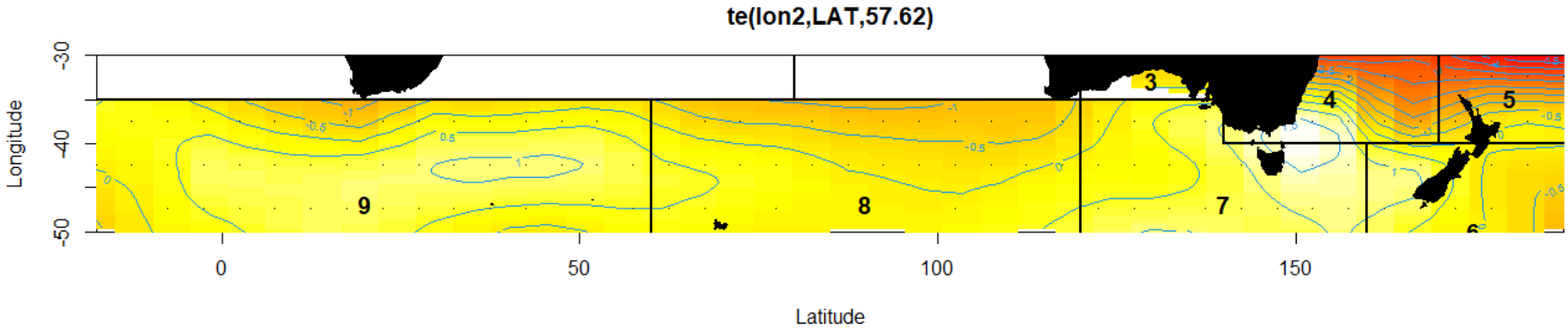

Figure 9: Two-way long:lat smoother from model gam11.

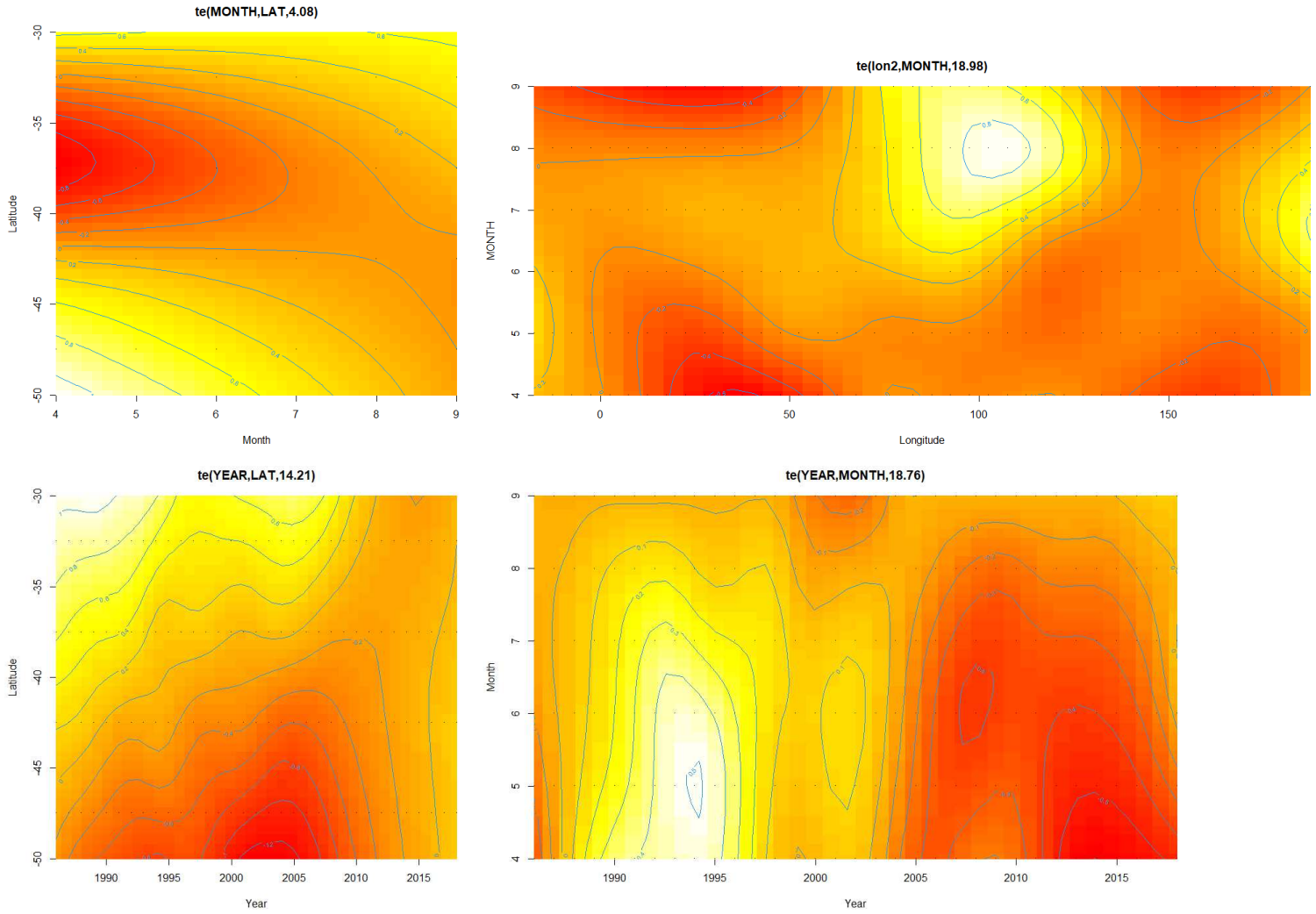

Figure 10: Smoothers from model gam11: month:lat, longitude:month, year:latitude, and year:month.

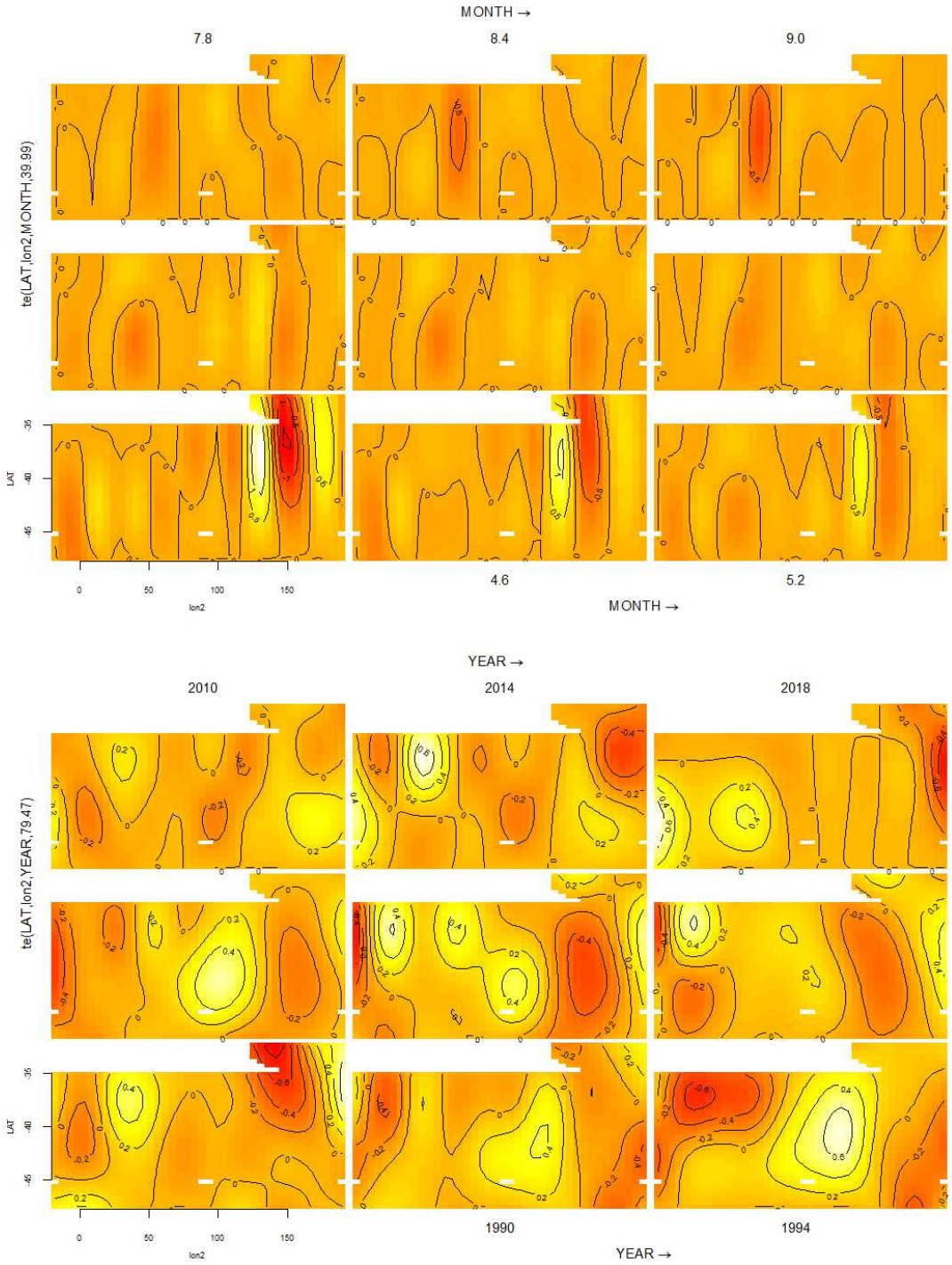

Figure 11: Smoothers from model gam11 Three-way interaction smoothers: lat:long:month (above) and lat:long:year (below).

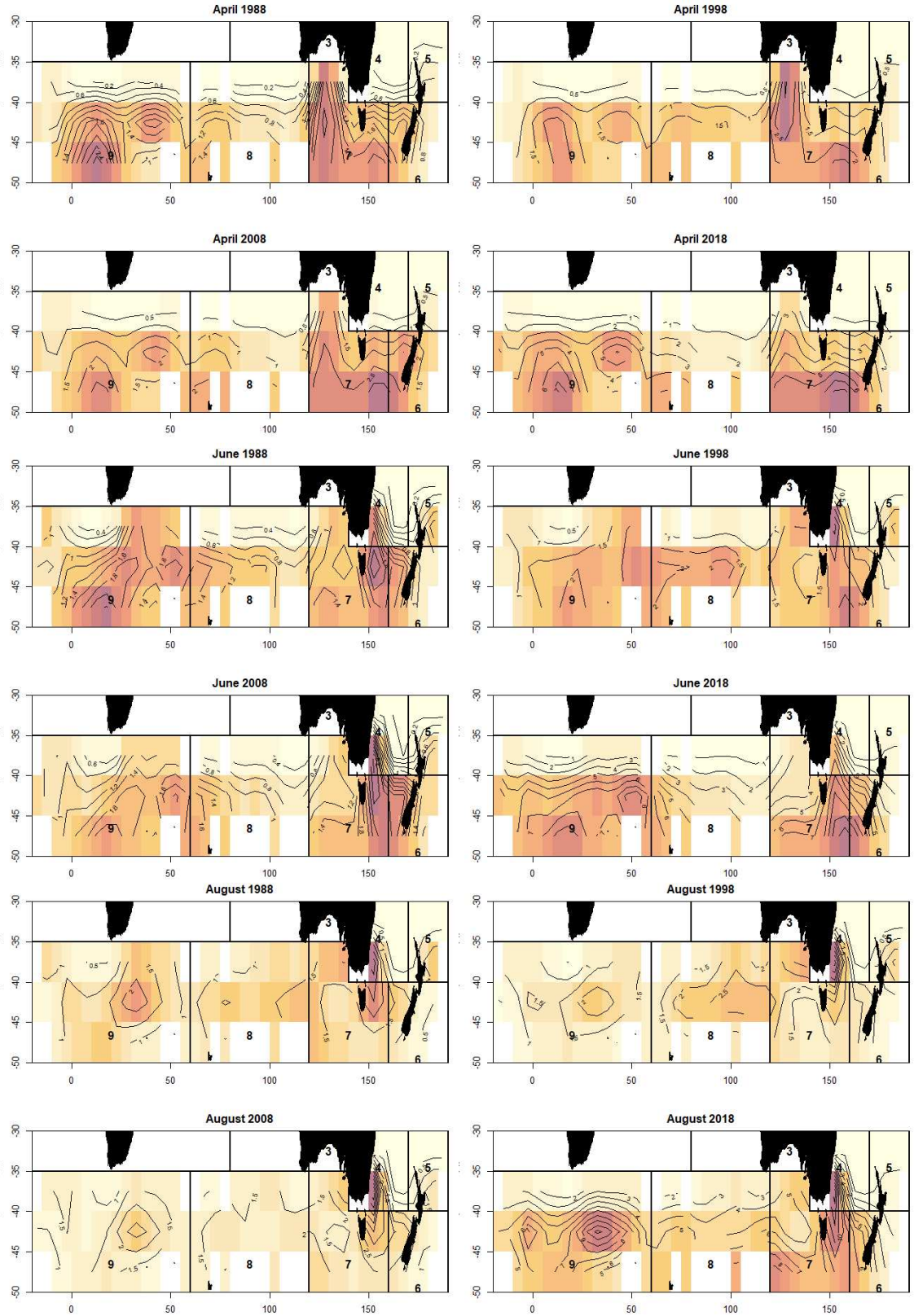

Figure 12: Predicted catch rates from model gam1 for regions 4 to 9 for the years 1988, 1998, 2008 and 2018, in April, June, and August.

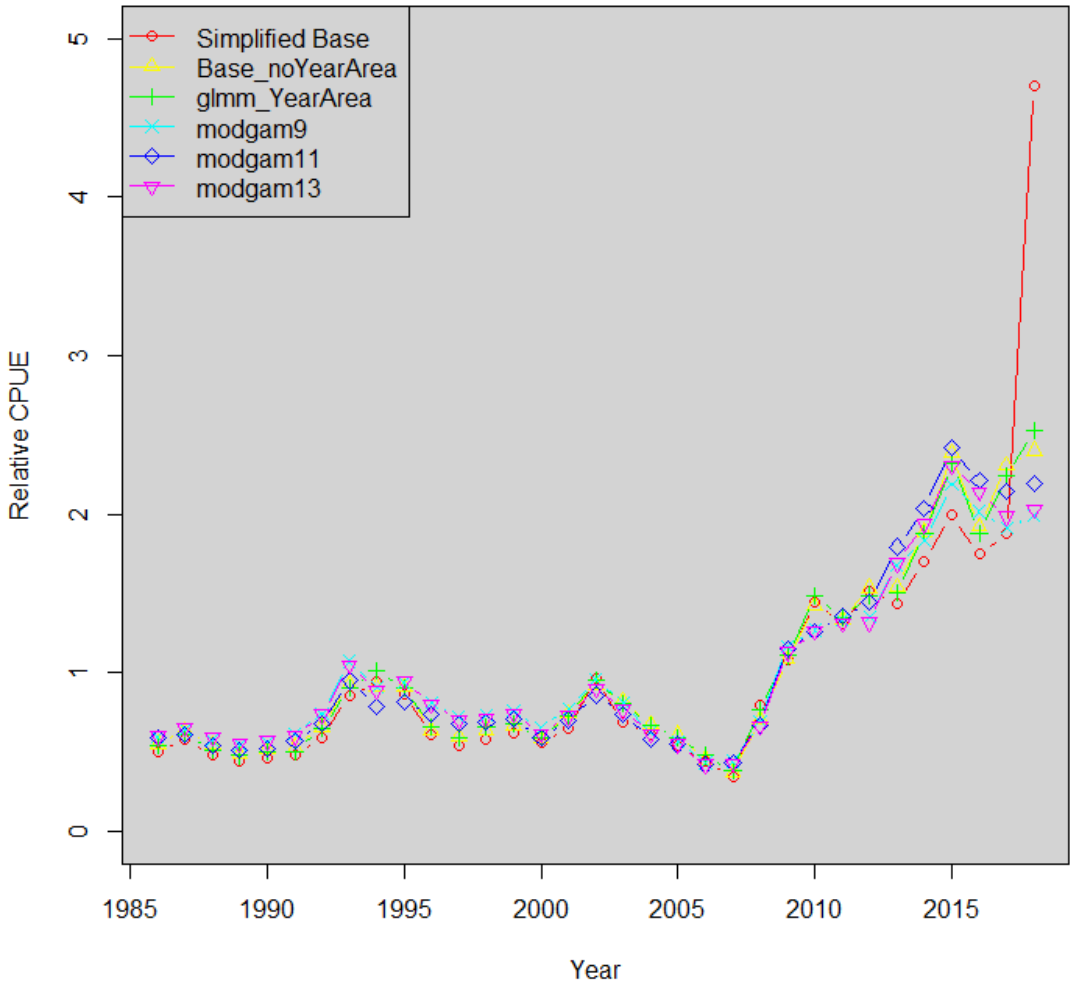

Figure 13: Mean annual predicted CPUE after x15 filtering from the models simplified base, base_noYearArea, glmm_YearArea, gam9, gam11, and gam13.

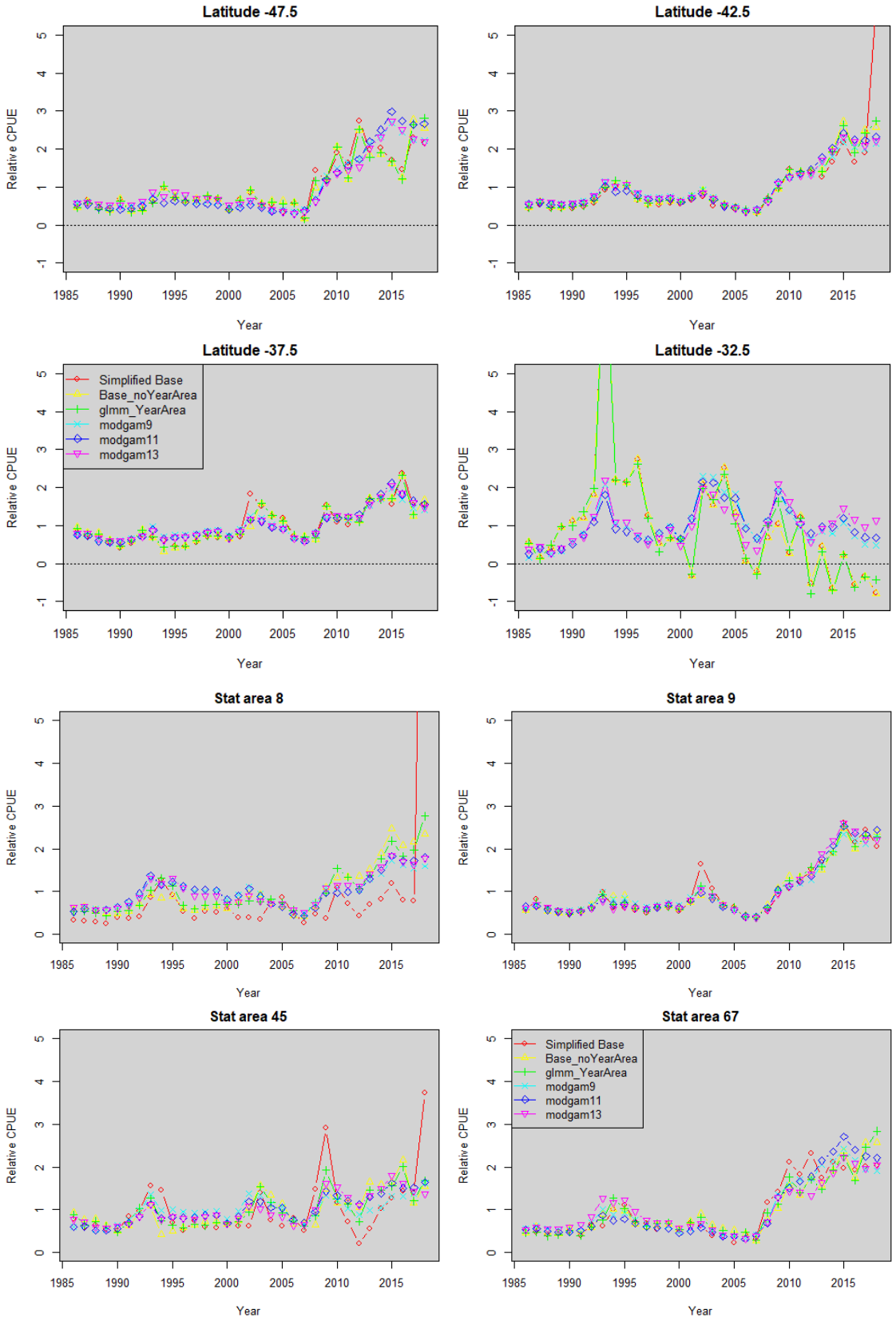

Figure 14: Mean annual predicted CPUE by latitude (above) and statistical area (below) after x15 filtering from the models simplified base, base_noYearArea, glmm_YearArea, gam9, gam11, and gam13.

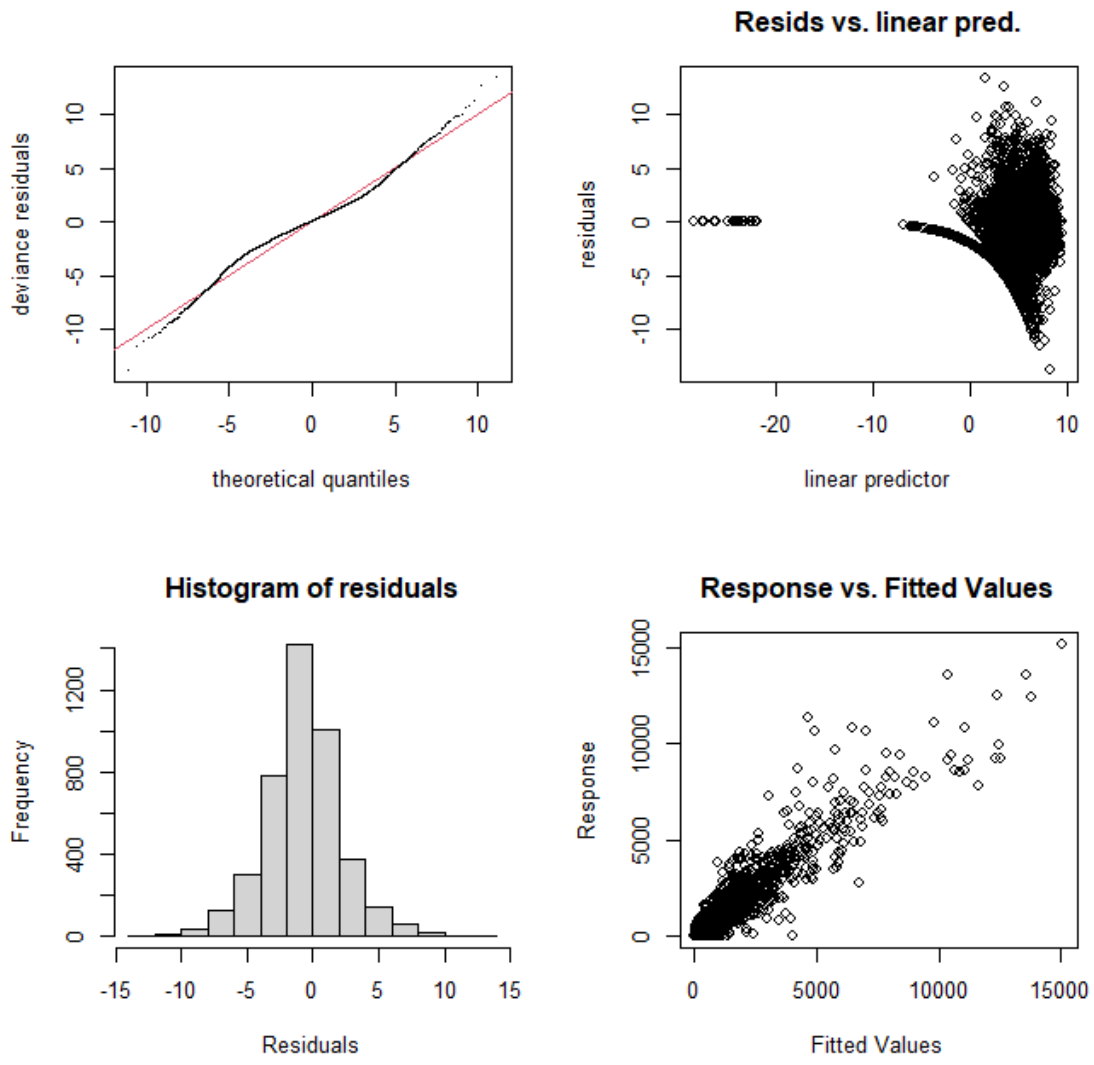

Figure 15:Diagnostic plots for the simplified Base model with residuals assumed to follow the Tweedie distribution with p=1.5.