CCSBT-ESC/2008/29
(ESC Agenda item 8)

Fisheries New Zealand
Tini a Tangaroa

Investigation of potential CPUE model improvements for the primary index of Southern Bluefin Tuna abundance

New Zealand Fisheries Assessment Report 2020/##

S.D. Hoyle

ISSN 1179-5352 (online)
ISBN XXXX (online)

July 2020

New Zealand Government

Requests for further copies should be directed to:

Publications Logistics Officer
Ministry for Primary Industries
PO Box 2526
WELLINGTON 6140

Email: brand@mpi.govt.nz
Telephone: 0800 00 83 33
Facsimile: 04-894 0300

This publication is also available on the Ministry for Primary Industries websites at:
http://www.mpi.govt.nz/news-and-resources/publications
http://fs.fish.govt.nz go to Document library/Research reports

TABLE OF CONTENTS

EXECUTIVE SUMMARY

Hoyle, S.D. (2020). Investigation of potential CPUE model improvements for the primary index of Southern Bluefin Tuna abundance.

Draft New Zealand Fisheries Assessment Report 2020/##. 33 p.

Indices of southern bluefin tuna abundance are used by the Commission for the Conservation of Southern Bluefin Tuna in both the stock assessment and the management procedure. In 2019, the Base CPUE model produced an index value for 2018 that was identified as anomalous. This research explored reasons for the high estimate and found that increasing effort concentration produced increasingly sparse coverage in the catch and effort dataset. This sparse coverage led to unstable predictions from the 'Base' GLM model in strata without observed CPUE. A new prediction diagnostic was developed based on the number of extreme values predicted by the model, and a set of new standardisation models was developed. Spatio-temporal smoothing in a generalised additive model using the R package *mgcv* provided more stable predictions for areas with sparse data and fitted the data better as measured by the AIC. A model was recommended for the purposes of the 11th Operating Model and Management Procedure Technical Meeting. Further work to develop the model was recommended.

1. INTRODUCTION

Indices of southern bluefin tuna (*Thunnus maccoyii*) (SBT) abundance are used by the Commission for the Conservation of Southern Bluefin Tuna (CCSBT) in both the 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' standardisation model. Next, two indices are derived from this model using prediction under two alternative weighting schemes, the constant squares (CS) and variable squares (VS) models. These two indices are then recombined using alternative weightings; in recent years these have been denoted as the W0.8 and W0.5 indices (Nishida & Tsuji 1998, Itoh & Takahashi 2019).

The Base model used to generate the CS and VS indices is a linear regression model that includes categorical variables for all spatial and temporal effects, along with three interaction terms as follows:

$$\log(cpue + 0.2) \sim yf + mf + areaf + latf + cpue.bet + cpue.yft + mf*areaf + yf*latf + yf*areaf$$

Here the parameters *yf*, *mf*, *latf*, and *areaf* are categorical variables (factors) representing year, month, statistical area, and latitude respectively. The catch per unit effort (CPUE) of bigeye tuna (*cpue.bet*) and yellowfin tuna (*cpue.yft*) are calculated as catch per thousand hooks and fitted as continuous variables. For southern bluefin tuna, *cpue* is SBT catch per thousand hooks. A constant of 0.2 is added to all *cpue* records to avoid taking the logarithm of zero. The value of 0.2 is approximately 10% of the mean *cpue*, which has been found to minimise the bias due to this adjustment of the catch rate (Campbell et al. 1996; Campbell 2004). Interaction terms that involve the year effect are sometimes ignored but, if substantial, may lead to a biased index (Maunder & Punt 2004). In this case they are justified by clear differences among statistical areas in catch rate trends through time. Procedures for calculating the index in these circumstances are detailed by Campbell (2015).

In recent years CPUE standardisation methods have given more consideration to spatial and temporal correlations (Nishida & 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 spatio-temporal smoothers within generalised additive models 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 categorical variables and linear models that generate the W0.8 and W0.5 indices.

However, in 2019, the Base CPUE model produced an index value for 2018 that was identified as anomalous (Itoh & Takahashi 2019), and the 2020 update generated a similarly unrealistic index value for 2018, with some concern about the 2019 estimate.

This paper explores reasons why the 2019 Base model produced high estimates for recent years and investigates the potential of generalised additive models (GAMs) (Hastie & Tibshirani 1990) that include spatio-temporal smoothers to provide a more reliable SBT abundance index.

2. METHODS

2.1 Data preparation

These analyses were based on a slightly different dataset from the Base model, because the dataset used in that analysis is only available to Japanese scientists (Itoh & Takahashi 2019). The available dataset was sufficiently similar to the primary dataset to provide useful insights. The main differences between the two datasets are listed below.

- 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 as operational (set by set) data (but is aggregated for the Base analysis) whereas the available dataset is aggregated.

The data file 'CPUEInputs_2020_January.txt', available from the private area of the CCSBT website, was used for the analysis. 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.

The following processes were then applied 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 10 000 hooks. Include latitudes north of 50° S.
- Create numeric *catch* variable, the sum of catches of all SBT 4+ and older.
- Create categorical *llf* variable, indicating 5° square that combines latitude and longitude.
- Create categorical *areaf* variable, which merges statistical area 4 with 5 and statistical area 6 with 7.
- Create categorical variables *yf*, *latf*, and *mf*, for year, latitude, and month.
- Adjust numeric longitude variable (*lon*) 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 adjusted longitude variable was from -95 to 260.
- Create numeric *cpue* variable = catch per 1000 hooks.
- Remove a single outlier with *cpue* > 120.

2.2 Characterise data

The data were investigated to identify how effort and catches have changed through time.

Effort was plotted by 5° square, and temporally by two-month periods, for each 5 years since 1985, to explore the spatial distribution of effort through time and by season.

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

$\log(cpue + 0.2) \sim yf + mf + latf + mf*latf$

2.3 Data coverage and Base model estimates

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

2.4 Check relevance of inference from available dataset

The available dataset was standardised using a model similar to the Base model, but without the parameters for other species. This is the simplified Base model:

$$\log(cpue + 0.2) \sim yf + mf + areaf + latf + mf*areaf + yf*latf + yf*areaf,$$

where all parameters are categorical variables.

As for the Base model, predicted catch rates were generated for each combined stratum of *yf*, *mf*, *areaf*, and *latf*.

To compare predictions from the Base (with primary dataset) and simplified Base (with available dataset) models, each set of predictions by year was normalised by the means across years and summed, and then all were plotted on the same figure.

2.5 Changes to the Base model

The fits of the simplified Base model and alternative approaches using categorical data were explored based on Akaike Information Criterion (AIC) and proportion of the deviance explained.

The simplified Base model included three of the six possible two-way interactions. A fourth interaction term (*mf***latf*) was added to produce the 'Base Plus' model. The effects on model fit of dropping each of these four interactions terms were examined. A full two-way model was produced by including all 6 possible interactions terms, and the effect of dropping each of these in turn was examined. Finally, the effect on model fit of adding each of the four possible three-way interaction terms to the Base Plus model was assessed.

2.6 Spatio-temporal smoothers

Generalised additive models are generalised linear models that replace or augment the linear function with an additive function that may include smoothing parameters and use a local scoring algorithm to estimate the smoothing parameters. As implemented in the *mgcv* package (Wood 2011), they are a flexible and powerful tool for data modelling. Under *mgcv* a variety of smoothers and other additive functions are available. Methods for visualising outputs are provided in the package *mgcViz* (Fasiolo et al. 2020).

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

The approach used for the simplified Base models was replicated and then other approaches were explored that might explain more variability while retaining parameter identifiability.

As with the previous analyses, modelling was carried out in R (R Core Team 2019) using the generalised additive modelling packages *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. The variables *areaf*, *latf*, *llf*, *mf*, and *yf* denote factors (categorical variables) for statistical area, latitude, 5° cells of latitude and longitude, month, and year. The variables *lon*, *lat*, *mn*, and *yr* denote continuous variables for longitude, latitude, month, and year. Interaction terms between continuous variables were fitted with the tensor product function *te()*, and individual continuous variables were fitted with the smooth function *s()* using the default thin-plate regression spline.

Each model was fitted with the *mgcv* setting 'gamma = 2' to reduce the effective sample size. CPUE standardisation models in *mgcv* can suffer from excess variability in the smoothers. Fisheries data are often over-dispersed due to dependencies among sets and strata, because individual data represent the combination of multiple sets by the same vessels, and unmodelled effects, such as environmental patterns and fish behaviour, 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 and Base Plus models were replicated in *mgcv* using factor variables as before.

Simplified Base:

$$\log(cpue + 0.2) \sim yf + mf + latf + areaf + mf:areaf + yf:latf + yf:areaf$$

Base plus:

$$\log(cpue + 0.2) \sim yf + mf + latf + areaf + mf:areaf + yf:latf + yf:areaf + mf:latf$$

Next, various smooth functions were used to replace and augment the factors (Table 1). A variety of two-way and three-way interactions were trialled. A four-way interaction with all the continuous variables was also explored.

The model gam13 (Table 1) was designed to be a smoothed version of model 'Base Plus', replacing the factor *areaf* with *te(lon, lat).* The factor *areaf* was replaced with *lon* in interaction terms so that, for example, *yf*areaf* became *te(yr, lon)*.

Model assumptions were checked using the *mgcv* function *gam.check()* 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(*hooks ~ s(residual)*), where *hooks* is the effort (number of hooks set) in the stratum.

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 by the mean of the yearly estimates, to give an index with mean of 1. Because ocean areas change with latitude, and some spatial cells include land, abundance prediction methods will need to be correspondingly adjusted in the future.

To explore the effect on the index of each model component, a progressive series of models that built up the components of the model recommended for inference (gam11) was fitted, as follows. Indices derived from these models were plotted in sequence.

1. *yf*
2. + *te*(*lon*, *lat*, k = c(40,4))
3. + te*(mn*, *lat*, k = c(6,4))
4. + *te*(*lon*, *mn*, k = c(10, 5))
5. + *te*(*yr*, *lat*, k = c(20, 4))
6. + *te*(*yr*, *mn*, k = c(20, 5))
7. + *te*(*lat*, *lon*, *mn*, k = c(4,15, 6))
8. + *te*(*lat*, *lon*, *yr*, k = c(4,10, 9))

Table 1: 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 four columns show the estimated degrees of freedom (df), the AIC, delta AIC, and the percent deviance explained.

Label	Factors	Smooth terms								4-way	df	AIC	δAIC	Deviance explained
Base											252.0	9845.3	2340	61.7%
Base plus	.+mf:latf										267.0	9530.7	2025	64.7%
Base_noYrAr	.-yf:areaf										156.0	10061.8	2556	57.9%
glmmYrAr											199.1	9893.7	2388	60.3%
gam 2	yf+mf	lon,lat									95.1	9522.9	2017	61.8%
gam 3	yf+mf							mn,lon,lat			152.4	8562.3	1057	70.3%
gam 4	yf									All	322.1	8064.9	559	75.5%
gam 5	yf+llf									All	297.7	7886.2	381	76.3%
gam 6	yf+llf		mn,lat							All	274.4	7856.4	351	76.2%
gam 7	yf	lon,lat	mn,lat							All	220.8	7839.1	333	75.7%
gam 8	yf	lon,lat	mn,lat	lon,mn						All	248.9	7758.5	253	76.4%
gam 9	yf	lon,lat	mn,lat	lon,mn	yr,lat			mn,lon,lat	lat,lon,yr		262.2	7701.7	196	76.9%
gam 10	yf	lon,lat	mn,lat	lon,mn	yr,lat	yr,lon	yr,mn				195.8	8048.5	543	74.2%
gam 11	yf	lon,lat	mn,lat	lon,mn	yr,lat		yr,mn	lat,lon,mn	lat,lon,yr		267.1	7593.4	88	77.5%
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	0	78.1%
gam 13	yf+mf	lon,lat	mn,lat	lon,mn	yr,lat	yr,lon					217.8	7968.3	463	74.9%

To identify the causes of the changes at each stage, individual smoothers from these models were plotted and compared with the distributions of effort and CPUE with respect to the parameters in the smoothers.

To identify the effects of data sparsity on the candidate standardisation models, indices of catch rate through time were created for each model by latitude *latf* and by statistical area *areaf*.

2.7 Extreme prediction diagnostic

A new prediction reliability diagnostic was developed (the extreme prediction diagnostic); this is based on the consistency of model predictions by year-month-statistical 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 more variable and influential than low extreme predictions given that errors are lognormally distributed. The x15 filtered values are used to generate the index, and this makes them more relevant than the full range of cells. Counts of extreme values in strata with no observed catch ('gap') are provided for interest, but all strata are used to generate the index.

3. RESULTS

3.1 Data characterisation

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

Statistical areas 4 and 5 had a relatively high proportion of records with zero catch, whereas statistical areas 6, 7, and 9 had very low proportions. 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.

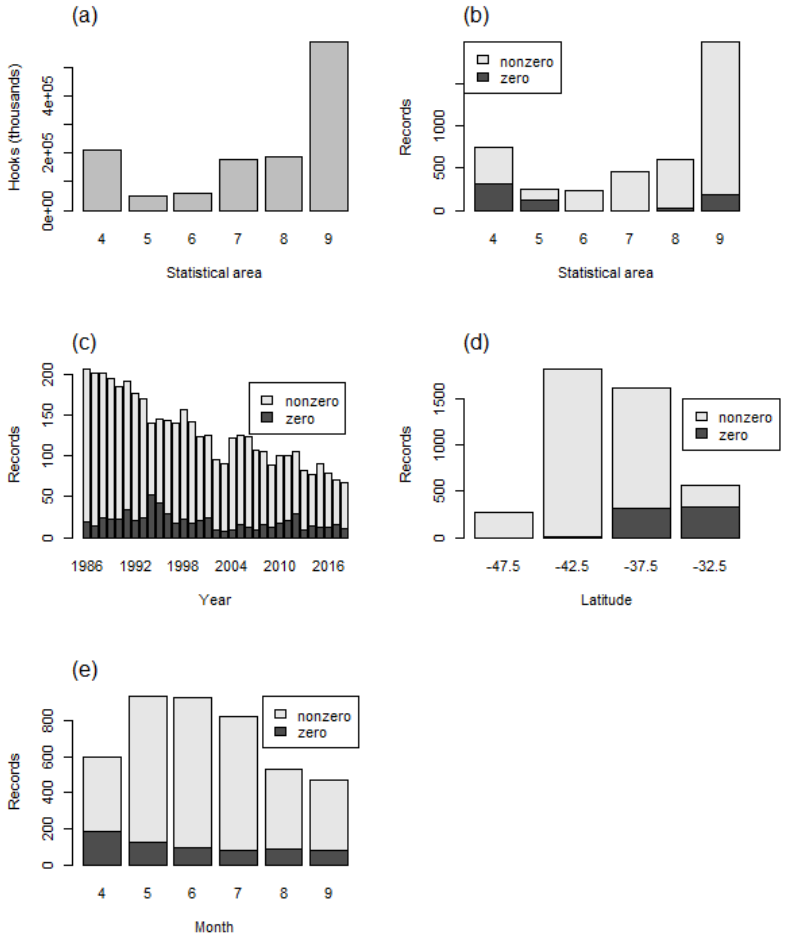

Figure 1: Distribution of effort and data records. (a) Total hooks per statistical area; (b) Total records per statistical area, including both zero and nonzero catches; (c) number of records per year, including both zero and nonzero catches; (d) number of records per latitude band, including both zero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

The spatial distribution of effort changed in consistent ways throughout the year (Figure 2). In April and May there was effort in the west (statistical area 9) and east (statistical areas 4, 5, 6, 7), but little effort in statistical area 8. Later in the year, effort moved north within these regions and into statistical area 8. Similar seasonal fishing patterns between statistical areas 8 and 9 were seen for the Korean fleet (Appendix Figure A1) (Hoyle et al. 2019). If these changes in effort distribution reflect SBT catch rates, they suggest interactions between month and latitude, and between month and longitude or statistical area.

Long-term changes in effort occurred, with less effort in all statistical areas during the 2015–2018 period, apart from an increase in statistical area 8 in April-May.

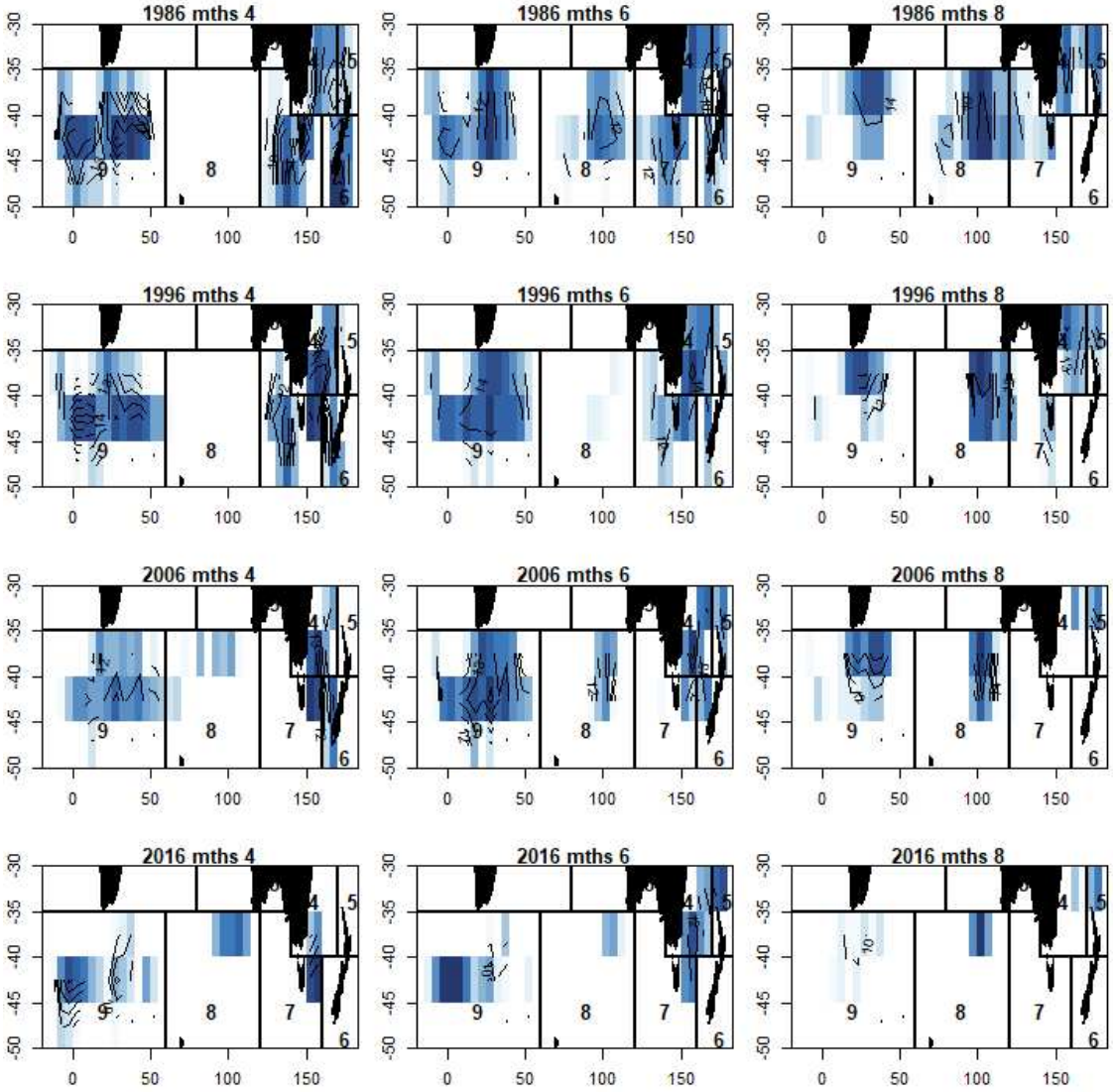

Figure 2: Maps of total effort per two-month period (across), summed over 10-year periods (down) starting in 1986. Darker blue indicates higher effort. Black indicates land.

Results of standardising catch rates separately by statistical area showed catch rates in statistical area 6+7 increasing rapidly from 2007 and stabilising at a high level in 2010 and catch rates in statistical area 9 increasing rapidly from about 2008 to a peak in 2015 (Figure 3). Catch rates in statistical area 4+5 showed no consistent change through time but did spike in 2018. Catch rates in statistical 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 statistical areas but substantially increased standardised catch rates in statistical area 8 from 2010.

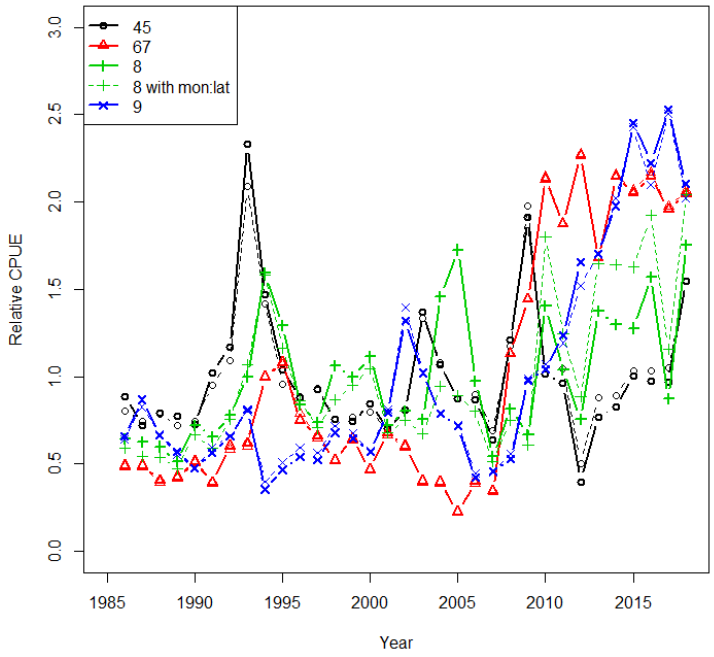

Figure 3: Standardised CPUE indices for individual statistical areas.

3.2 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 2). 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 effect of back-transformation.

The increasing number of empty strata is consistent with figure 1b from Itoh & Takahashi (2019), which shows operations with SBT aged 4+ concentrated into a steadily reducing number of 5° cells and 1° cells through time, with minima for both 1° and 5° cells in 2018 (Figure A2). 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 in the Korean data.

The Base model (predictions from parameters provided by Tomoyuki Itoh, NRIFSF, Japan) predicted very high catches per thousand hooks in 2018 in statistical area 8, with the highest predictions all in strata with no reported effort (Table 2). Of the 18 month-latitude strata in statistical area 8, over 40 fish per thousand hooks were predicted in four strata and almost 20 fish in two strata.

Table 2: 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. Anomalously high values are highlighted in orange. (Continued on next page.)

			1990					2010			2017			2018
stat area	mon	lat5	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

stat area	mon	lat5	hooks	cpue	1990 Pred cpue	hooks	cpue	2010 Pred cpue	hooks	cpue	2017 Pred cpue	hooks	cpue	2018 Pred cpue
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
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

The Base model and the simplified Base model gave very similar indices (Figure 4), 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.

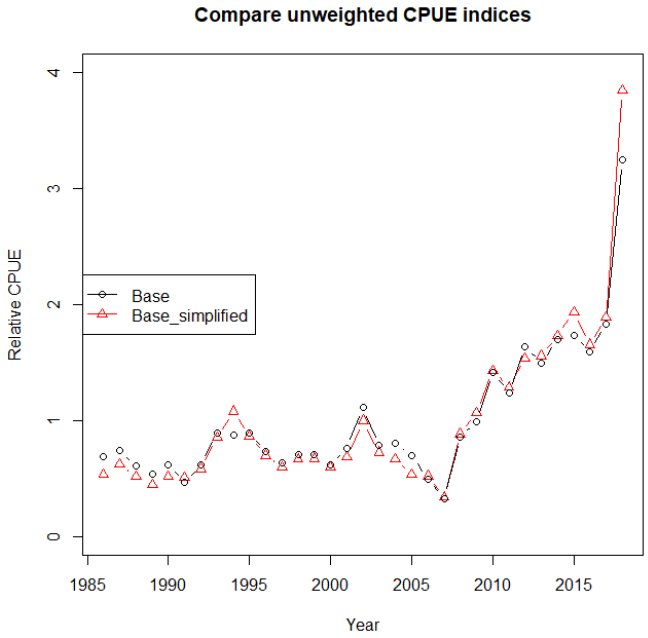

Figure 4: Comparison of the CPUE indices between the predictions generated from the Base parameters and the predictions generated from standardised data. For each, the predicted CPUE estimates are summed across all strata without weighting.

3.3 Changes to the simplified Base model

Removing any of the parameters and interactions included in the Base plus model resulted in poorer fit as measured by the AIC (Table 3). The *mf* * *latf* interaction term (not included in the simplified Base model) had the most effect on the AIC of all the interactions, suggesting that this interaction should be considered in future analyses. This interaction term is consistent with the observed movement patterns of the fleet, which fishes further north within each statistical area later in the season (Figure 2).

The *yf* * *mf* and *areaf* * *latf* interactions (also not included in the Base model) improved the AIC but by the smallest amount of all possible two-way interactions (Table 4).

Three of the four three-way interaction terms improved the AIC of the model, with the most impact coming from *mf* * *latf* * *areaf* (Table 5).

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 over-dispersed, which can make the AIC oversensitive and lead to overfitting. Prediction reliability must also be considered when including additional terms, given the increasing concentration of the effort and the consequent sparse data.

Table 3: Changes in the deviance, degrees of freedom, and AIC as a result of dropping parameters from the Base plus model.

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

	DF	Deviance	AIC
<none>		2049.0	9530.7
mf:areaf	15	2137.6	9682.1
yf:latf	96	2245.9	9731.7
yf:areaf	96	2258.6	9756.0
mf:latf	15	2220.7	9845.3

Table 4: Changes in the 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 + areaf + latf + mf*areaf + yf*latf + yf*areaf + mf*latf + yf*mf + areaf*latf

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

Table 5: Changes in the degrees of freedom and AIC as a result of adding three-way interactions to the following model.

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

3.4 Spatio-temporal smoothers

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

The best 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 (see Table 1). Given these results, the models gam2 to gam8 were dropped, 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 (Table 7) 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 poor performance with extreme predictions, models gam10 and gam12 were dropped. Model gam13, designed to be similar to the Base model but with smoothers, had lower AIC by 1877, predicted 5 high values rather than 19, and explained 74.9% of deviance rather than 61.7%. The model selected as the best was gam11, which had AIC better again by 375, predicted 2 high values, and explained 77.5% of the deviance.

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 A3). 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 A4).

Spatio-temporal 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 (Figures A5, A6, and A7).

Predictions from model gam11 across the spatial domain showed changing spatial distribution by month and year (Figure 5).

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(lon, lat, k = c(40,4)) + te(mn, lat, k = c(6,4)) + te(lon, mn, k = c(10, 5)) + te(yr, lat, k = c(20, 4)) + te(yr, mn, k = c(20, 5)) + te(lat, lon, mn, k = c(4,15, 6)) + te(lat, lon, yr, 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, lon, mn, k = c(4,15, 6))
7859.9	266.5	1.8	te(lat, lon, yr, k = c(4,10, 9))

Table 7: Extreme prediction diagnostic showing the number of year-month-latitude-statistical 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). The version of the extreme prediction diagnostic preferred for inference is in bold.

Limit	Cell filter	Strata	Models							
			Simp	Base	glmm	gam	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

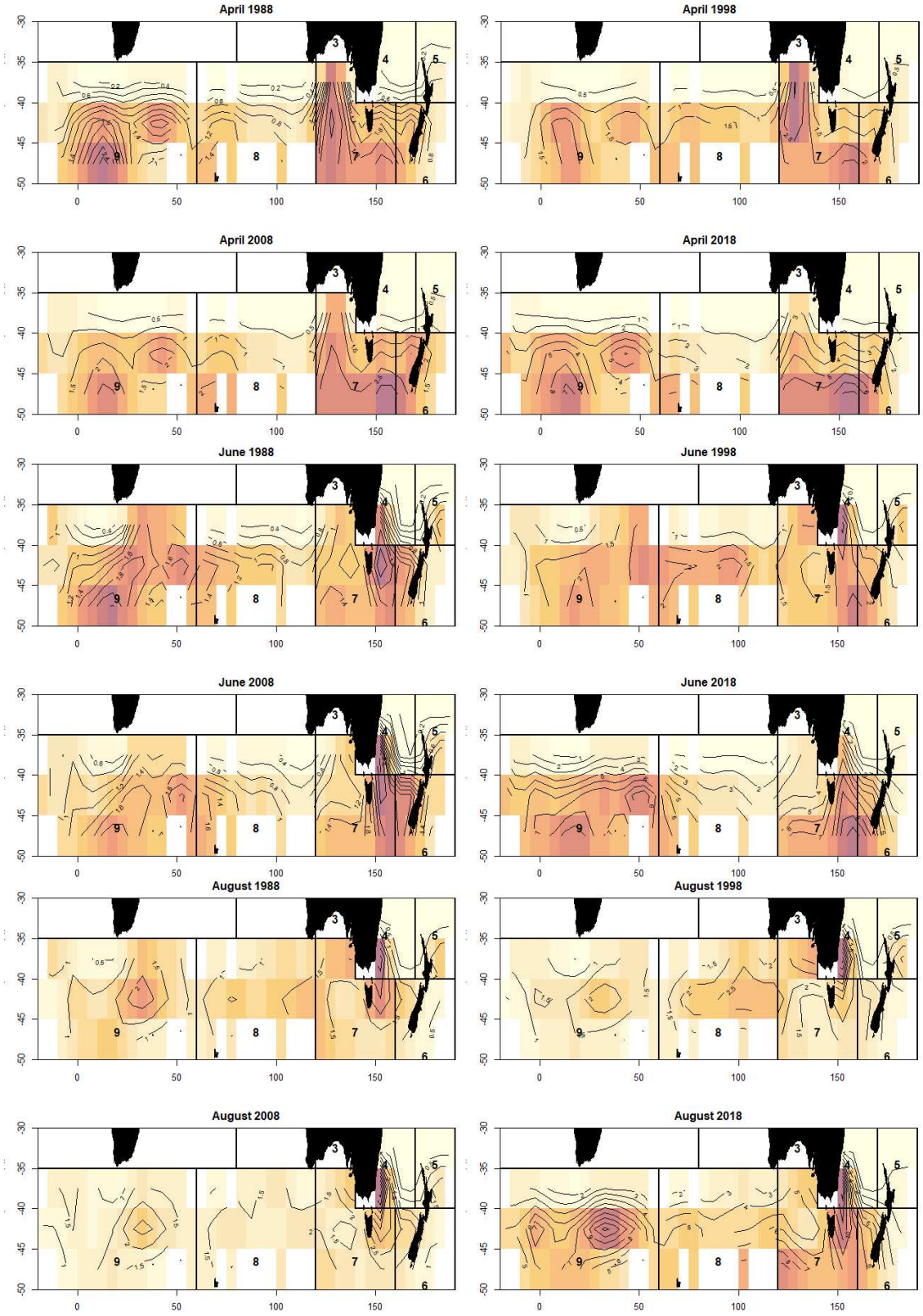

Figure 5: Predicted catch rates from model gam11 for statistical areas 4 to 9 for the years 1988, 1998, 2008, and 2018, in April, June, and August.

Indices from the three 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 6).

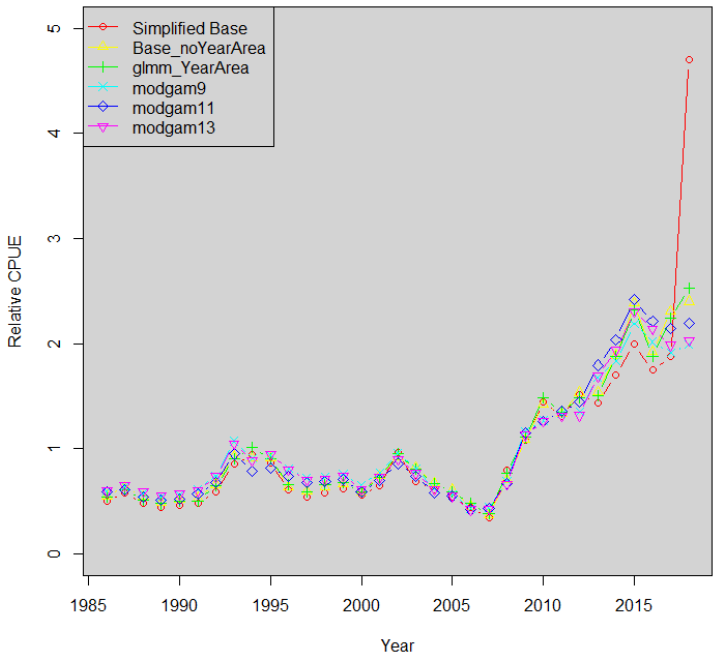

Figure 6: Mean annual predicted CPUE after x15 filtering from the models simplified base, base_noYearArea, glmm_YearArea, gam9, gam11, and gam13.

Progressive changes in the CPUE are shown as the model terms are added (Figure 7). Much of the change in trend since 2015 is associated with fishing location, with a large change in the third row of the progressive plot when te(lat, lon) is introduced. In 2015 there was much more effort in the 37.5° S latitude band and in the west than in 2018, and the spatial smoother expected this north-western effort to have lower catch rates (Figure 8). When the effort moved south the model explained the increase in catch rates with the spatial effect.

When the year-latitude smoother is introduced (row 6 of the progressive plot, Figure 7), the contribution of the year effect increases after about 2013. This smoother increases the expected CPUE in the south in 2015, compared with the expected CPUE in the south in 2010 (Figure 9), which is consistent with the increasing observed CPUE in the south (Figure 10).

When the year-month smoother is introduced (row 7 of the progressive plot) the contribution of the year effect after about 2013 increases further. This smoother reduces the expected CPUE in April and May compared with June and July, in 2015 relative to 2010 (Figure 9). This is consistent with the increasing difference between April-May and June-July during this period (Figure 10).

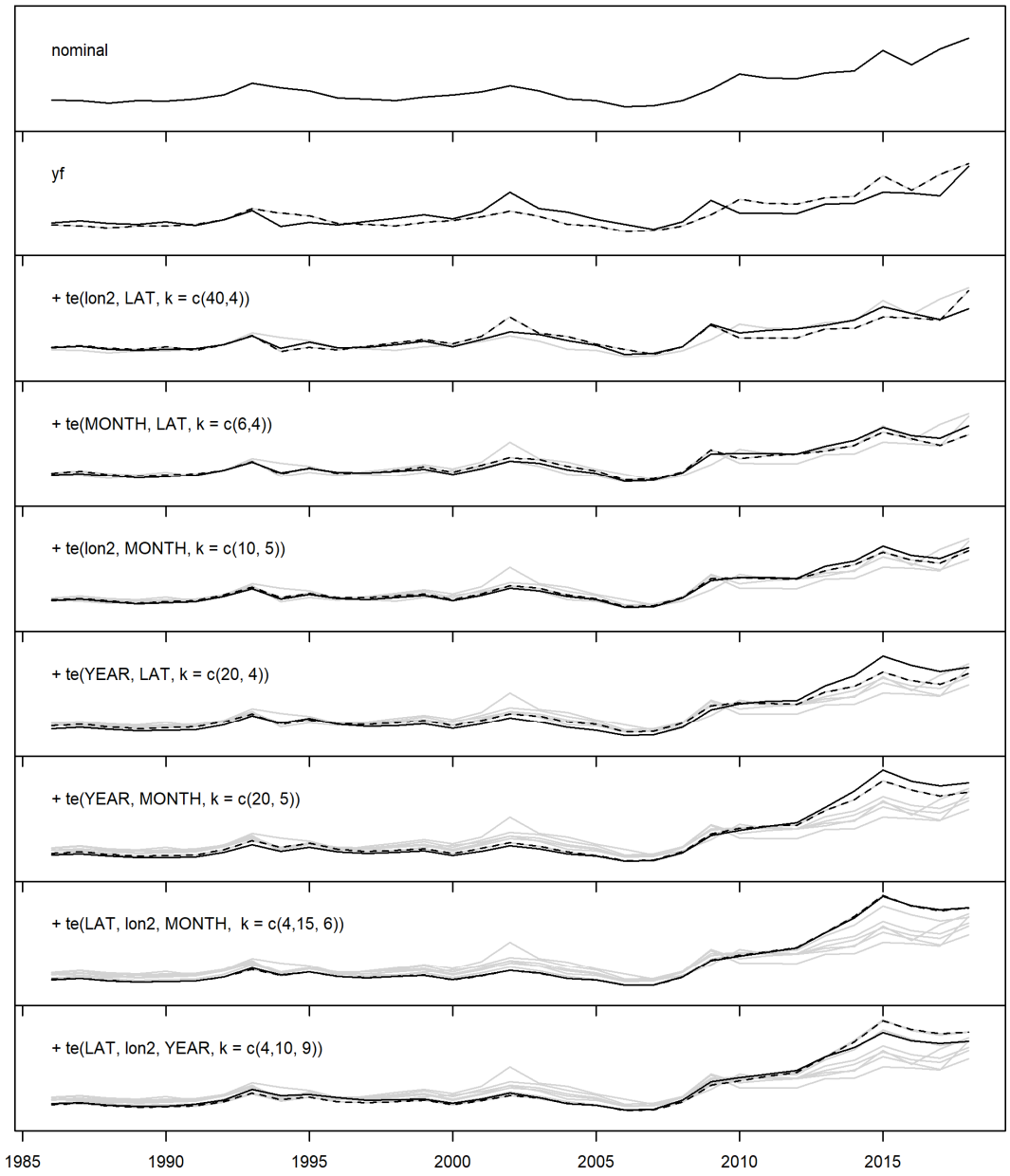

Figure 7: Progressive changes in the index as each term of the gam11 model is added, starting with the nominal CPUE and the simplest model [log(CPUE + 0.2) ~ yf] in the top two rows. Each row of the plot includes the model specified in that row (black), the model from the row above (dashed line), and other models in rows above (grey).

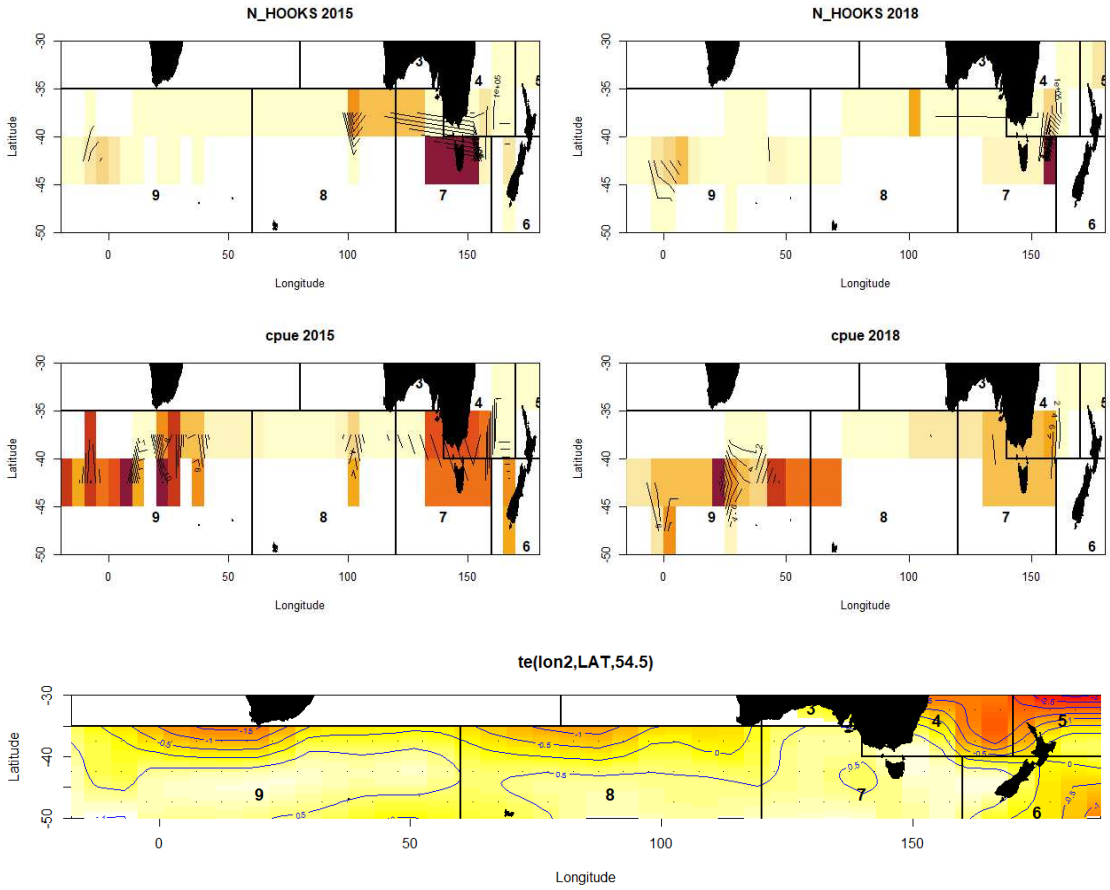

Figure 8: Density plot of effort (above) and CPUE (middle) in 2015 (left) and 2018 (right). Darker colours indicate higher effort or CPUE. Below is the spatial smoother from the model log(CPUE + 0.2) ~ yf + te(lon, lat).

Predictions by latitude and by statistical area were much more stable than those from the factor-based models (Figure 11). High variability associated with sparse data was apparent at the northern and southern latitudes -32.5 and -47.5, and in statistical areas 8, 4+5, and 6+7. Note, however, that statistical areas with sparse data were given low weight by the constant squares and variable squares algorithm and by x15 filtering, which greatly limited their impact on the resulting indices. These figures nevertheless illustrate how spatio-temporal smoothing can stabilise estimates when data are sparse.

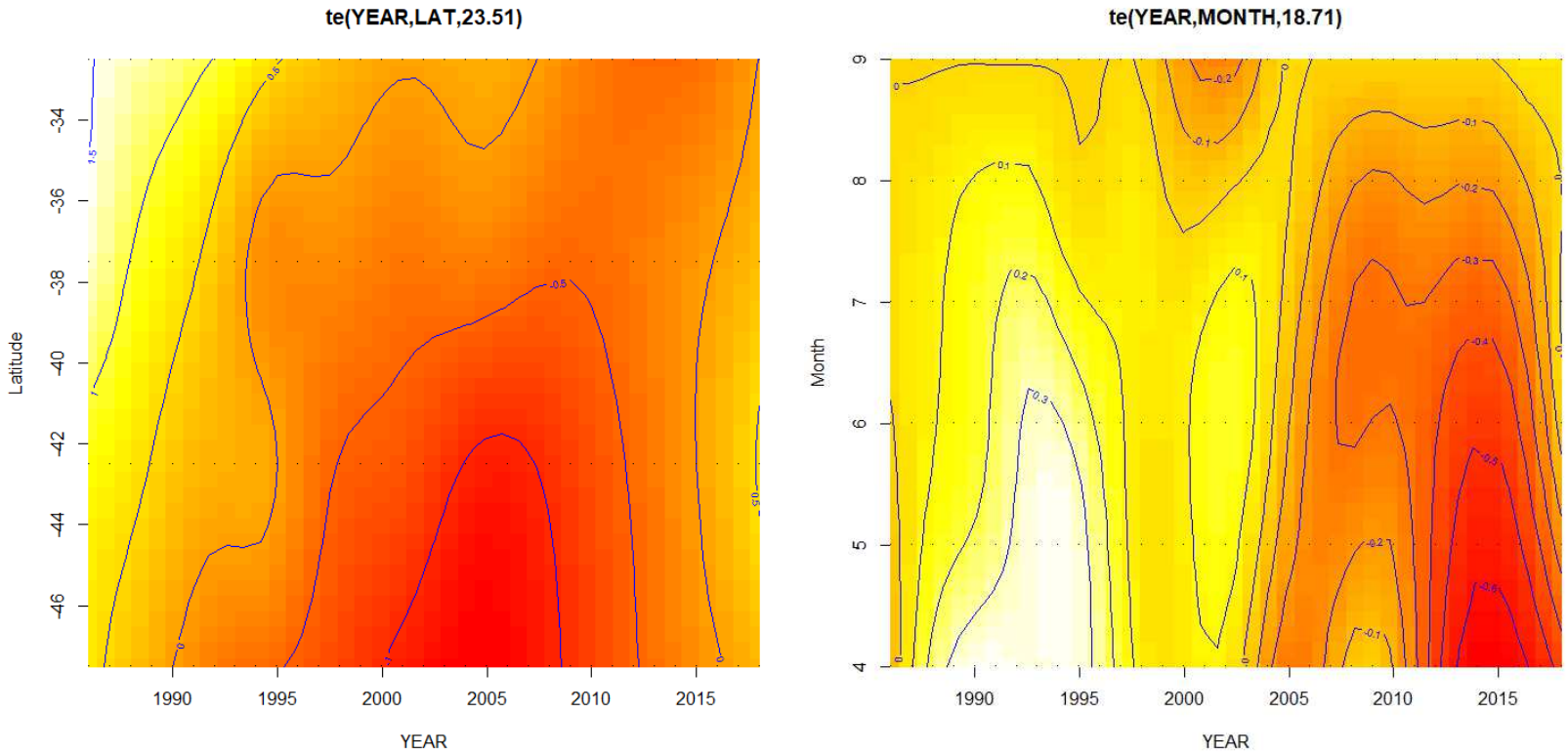

Figure 9: (Left) Smoother on yr and lat (left) from the fifth model in the progressive series: log(cpue + 0.2) ~ yf + te(lon, lat, k = c(40, 4)) + te(mn, lat, k = c(6, 4)) + te(lon, mn, k = c(10, 5)) + te(yr, lat, k = c(20, 4)). (Right) Smoother on yr and mn from the sixth model in the progressive series: log(cpue + 0.2) ~ yf + te(lon, lat, k = c(40, 4)) + te(mn, lat, k = c(6, 4)) + te(lon, mn, k = c(10, 5)) + te(yr, lat, k = c(20, 4)) + te(yr, mn, k = c(20, 5)).

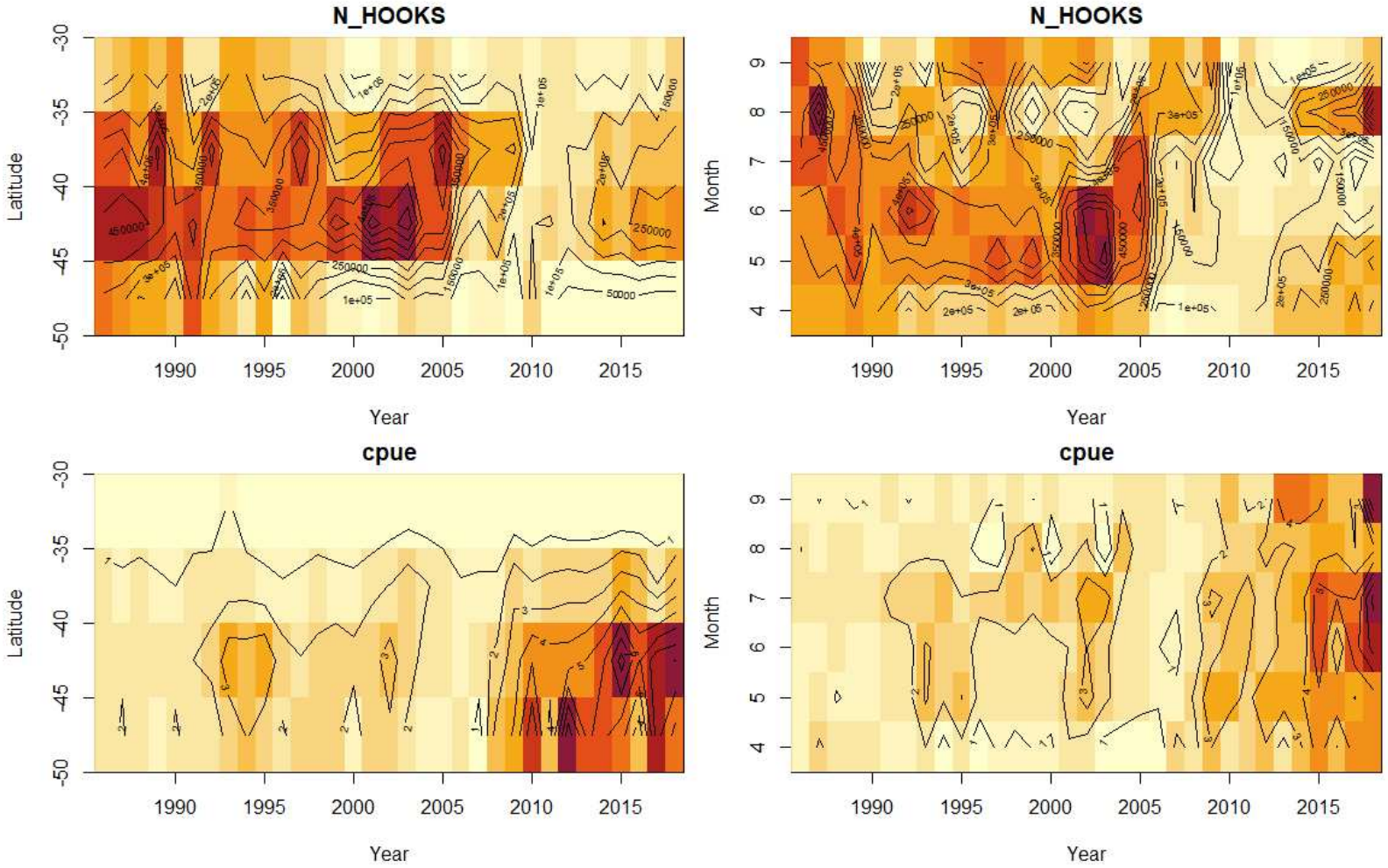

Figure 10: Density plots of effort in number of hooks (above) and CPUE (below) by year and latitude (left) and by year and month (right). Darker colours indicate higher effort or CPUE.

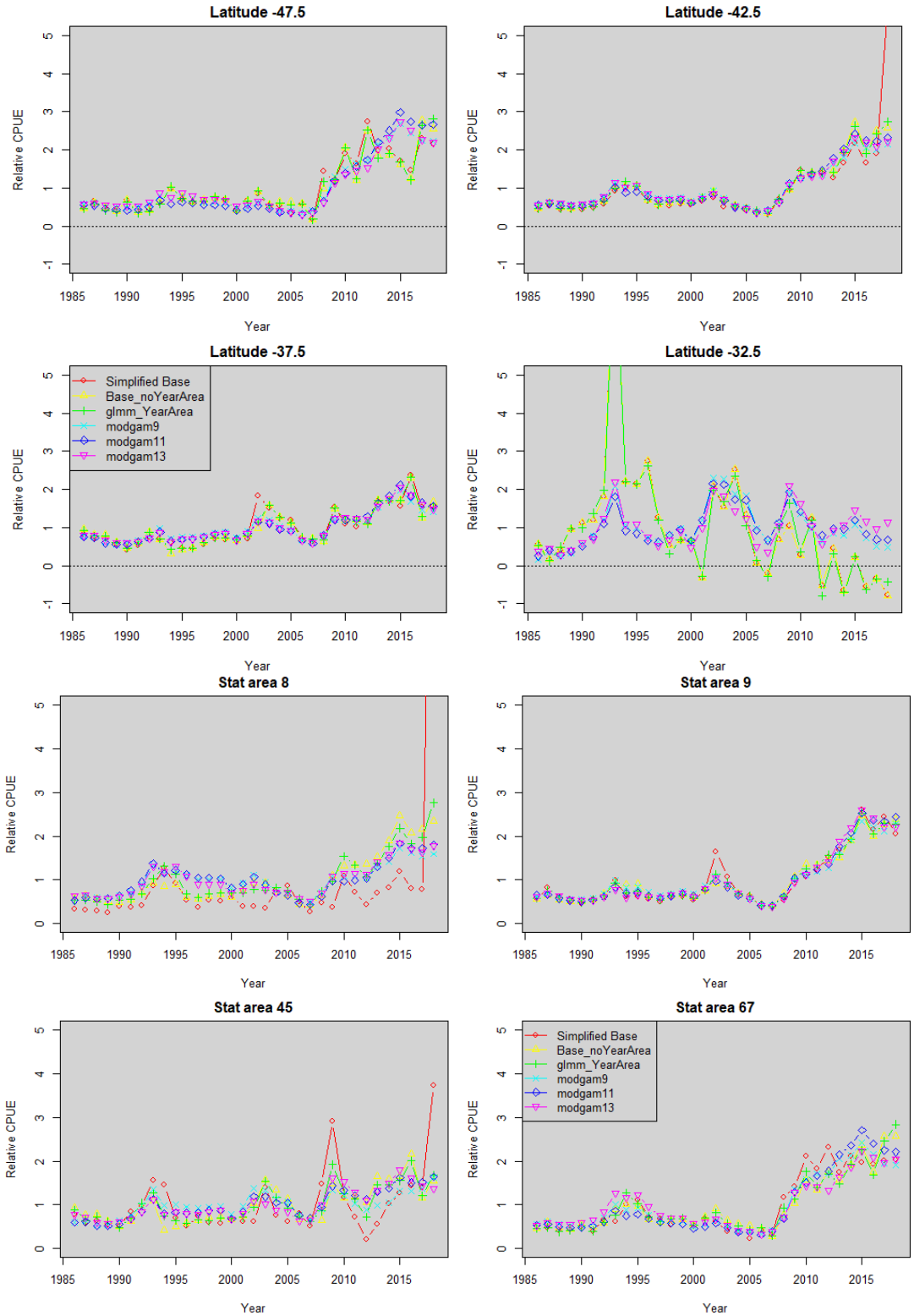

Figure 11: Mean annual predicted CPUE by latitude (above) and statistical area (below) after x15 filtering from the models simplified base, base_noYearArea, glmm_YearArea, gam9, gam11, and gam13.

4. DISCUSSION

This study has shown that the anomalous 2018 result (Itoh & Takahashi 2019) was likely caused by sparse data associated with increasing effort concentration. Generalised additive models with spatio-temporal smoothers were able to address this problem. Such models may also provide more accurate indices of abundance by taking into account factors that cannot be included in the current SBT CPUE standardisation model.

Effort included in the primary SBT index has been increasingly concentrated, with a relatively steady decline since 2006 in cells fished, and an increase from 2010 to 2018 in effort per cell. These changes have been associated with restrictive SBT catch quotas, subsequent increases in fish abundance, and associated changes in fishing behaviour. Changes in technology that facilitate fish finding may also have played a role. Future changes in effort concentration are unpredictable but, if the extent of concentration remains close to or above current levels, problems with prediction behaviour by the Base model are likely to be repeated.

The reasons for the increasing concentration of fishing effort, and its implications, need to be understood. The increasing concentration is very marked in both the Japanese and Korean fleets. As well as causing the analytical problems that motivated this study, it may affect the reliability of CPUE as an index of abundance in ways that spatio-temporal smoothing does not resolve. 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 therefore be appropriate to retain an approach that includes aspects of the VS method.

There may also be concerns about changing stock distribution due, for example, to the effects of climate change on oceanic currents and temperatures. Contraction and expansion would be easier to detect if modelling at the 1° cell scale, which would be straightforward with smoothers. Currently there are only 4 x 5° latitude bands in the model, and changes in distribution must be large to be detectable. It would also be easier to detect changes in distribution if data were available from more vessels, which could be accomplished by including other fleets in the analysis, and perhaps including non-core vessels. Either of these additions to the dataset would make it more important to consider the effects of targeting and vessel-specific catchability.

Spatio-temporal smoothers take advantage of Tobler's (1970) 'First Law of Geography', that 'everything is related to everything else, but near things are more related than distant things', a principle that can also be applied to time. Spatio-temporal smoothers implemented in a GAM with *mgcv* fitted the catch and effort data better than the categorical variables used in the Base linear model or any of the non-smoother alternatives, while using fewer parameters.

Model fits were compared using the AIC which when used with CPUE standardisation tends to select the more highly parameterised model and results in overfitting. This is not a failing of AIC but due to the incorrect model assumption that records are independent. This is particularly problematic with CPUE standardisation of operational longline data that include time series of daily sets by the same vessels. However, it remains a concern with aggregated data. With aggregated data, these short-term considerations are less significant, but there are still dependencies within the data that are not accounted for by the model. These dependencies include factors affecting catchability such as vessel effects, oceanographic features, and targeting and reporting behaviour. Factors affecting fish distribution can also introduce dependence among strata and add process uncertainty to the relationship between CPUE and abundance. Use of alternative fitting criteria such as the Bayesian information criterion (BIC) raises the threshold for model selection but does not address the cause of the problem. A method commonly used in CPUE standardisations is to include only model components that explain at least 1% of the deviance. That approach was used to check the smoothers selected in this study for model gam11, and all such model components were included.

Similarly, lack of independence usually affects the confidence intervals estimated for CPUE indices. Generalised Estimating Equations can help by estimating the dependencies, but these approaches are not always practical and not often used. To adjust the smooth fitting process for lack of independence, effective sample size was reduced by half by setting the gamma parameter to 2.

A major advantage of sharing parameters with spatio-temporal smoothing is the ability to better represent the biological features of the population. The Base model does not include a month by latitude interaction term, although effort tends to consistently move further north later in the season, as sea surface temperatures cool during the austral winter. Adding this interaction term to the simplified Base model improved the AIC but is likely to worsen prediction problems by increasing the number of strata. The models with smoothers were able to include these effects and greatly improve the fit to the data, while at the same time predicting fewer extreme values.

The 'extreme prediction diagnostic' approach used here is a new approach for assessing the reliability and utility of model predictions. The upper limit is particularly relevant to CPUE modelling of target species, because more fishing effort can be expected in strata where catch rates are higher, so strata that are informed by less information may be expected to have lower catch rates. Predicted catch rates that exceed the maximum observed are therefore likely to be unrealistic. Currently, the diagnostic is a tool for comparing models but there are no established criteria for what is acceptable. This is likely to remain subjective and will depend on the dataset and species being modelled. Nevertheless, simulation would be useful to explore the behaviour of the diagnostic.

Poor prediction beyond the range of the data is a well-known problem that occurs with both factors and smoothers. Prediction quality was managed with the criterion of at least 15 records per 5° cell, and by screening models with the extreme value diagnostic. Prediction behaviour varies among types of smoother, and *mgcv* provides many alternatives. Interaction-only tensor products using *ti()* were explored for some terms instead of using full tensor products with *te()* for every smooth term, but resulted in many more extreme predictions.

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: consideration of vessel-specific fishing power which in many fisheries varies considerably and tends to increase through time; adjusting for targeting; the use of hurdle or zero-inflated models to deal with zero-catch strata instead of adding a constant; and the effects of quotas on vessel behaviour and catch rates, particularly within-season.

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

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 spatio-temporal 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 standardisation methods found that 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, and therefore it 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.

5. ACKNOWLEDGMENTS

This work was funded by Fisheries New Zealand and completed as Milestone 2, Objective 1 of Fisheries New Zealand project SEA2019-33. Darcy Webber, Jim Ianelli, Ana Parma, Doug Butterworth, and the CCSBT CPUE working group provided ideas and discussions. Thanks to Tomoyuki Itoh and Japan for their collaboration.

6. 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(1)*: 331–361.

Campbell, R.; Tuck, G.; Tsuji, S.; Nishida, T. (1996). Indices of abundance for southern bluefin tuna from analysis of fine-scale catch and effort data. *Second CCSBT Scientific Meeting*, Hobart, Australia, August 26 – September 6 1996.

Campbell, R.A. (2004). CPUE standardisation and the construction of indices of stock abundance in a spatially varying fishery using general linear models. *Fisheries Research 70(2–3)*: 209–227.

Campbell, R.A. (2015). Constructing stock abundance indices from catch and effort data: Some nuts and bolts. *Fisheries Research 161*: 109–130.

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.

Fasiolo, M.; Nedellec, R.; Goude, Y.; Wood, S.N. (2020). Scalable visualization methods for modern generalized additive models. *Journal of computational and Graphical Statistics 29(1)*: 78–86.

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. *Fisheries Research 213*: 75–93.

Hastie, T.J.; Tibshirani, R.J. (1990). *Generalized additive models*. CRC Press.

Hoyle, S.; Lee, S.I.; Kim, D.N. (2019). Data exploration and CPUE standardization for the Korean southern bluefin tuna longline fishery (1996-2018), CCSBT-ESC/1909/39. *Extended Scientific Committee for the Twenty Fourth Meeting of the Scientific Committe*, Cape Town, South Africa.

Itoh, T.; Takahashi, N. (2019). Update of the core vessel data and CPUE for southern bluefin tuna in 2019. *CCSBT Extended Scientific Committee*.

Maunder, M.N.; Punt, A.E. (2004). Standardizing catch and effort data: a review of recent approaches. *Fisheries Research 70(2–3)*: 141–159. 10.1016/j.fishres.2004.08.002

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. *Fisheries Research 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. *Fisheries Research 70(2–3)*: 265–274. 10.1016/j.fishres.2004.08.008

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, July 23–31, 1998.

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

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 Journal of Marine Science 72(5)*: 1297–1310.

Tobler, W.R. (1970). A computer movie simulating urban growth in the Detroit region. *Economic Geography 46(sup1)*: 234–240.

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.

7. APPENDIX

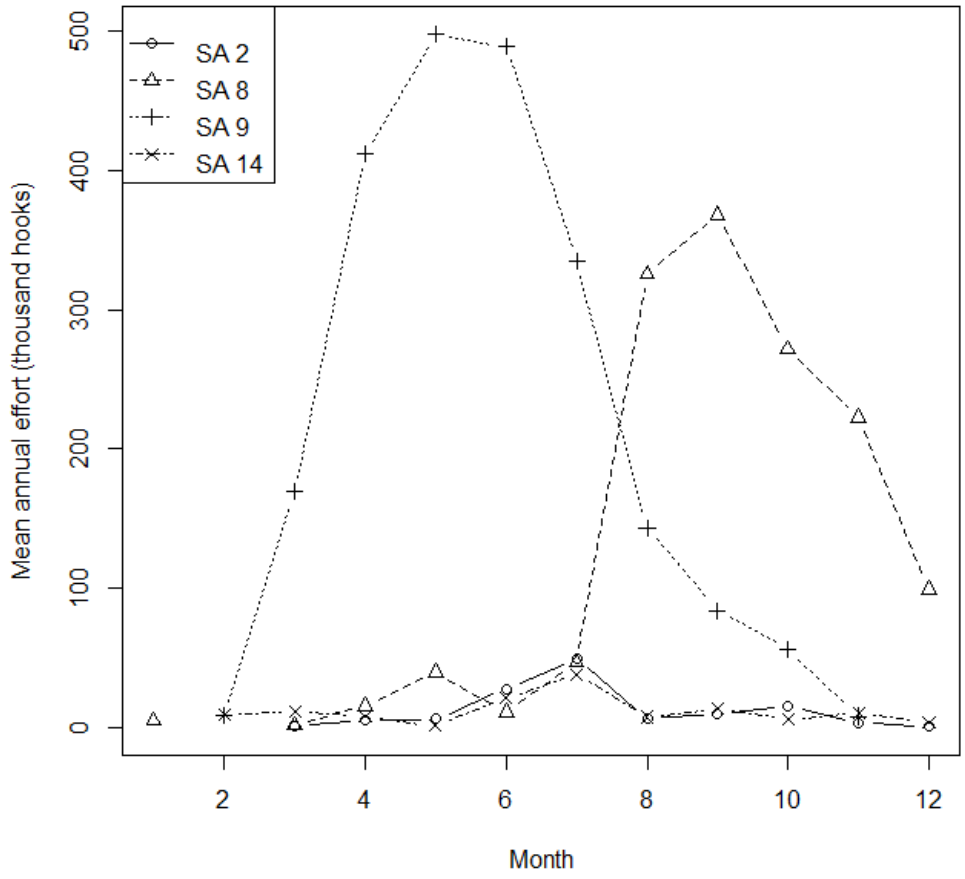

Figure A1: Monthly effort per CCSBT statistical area by the Korean longline fleet, from Hoyle et al. (2019).

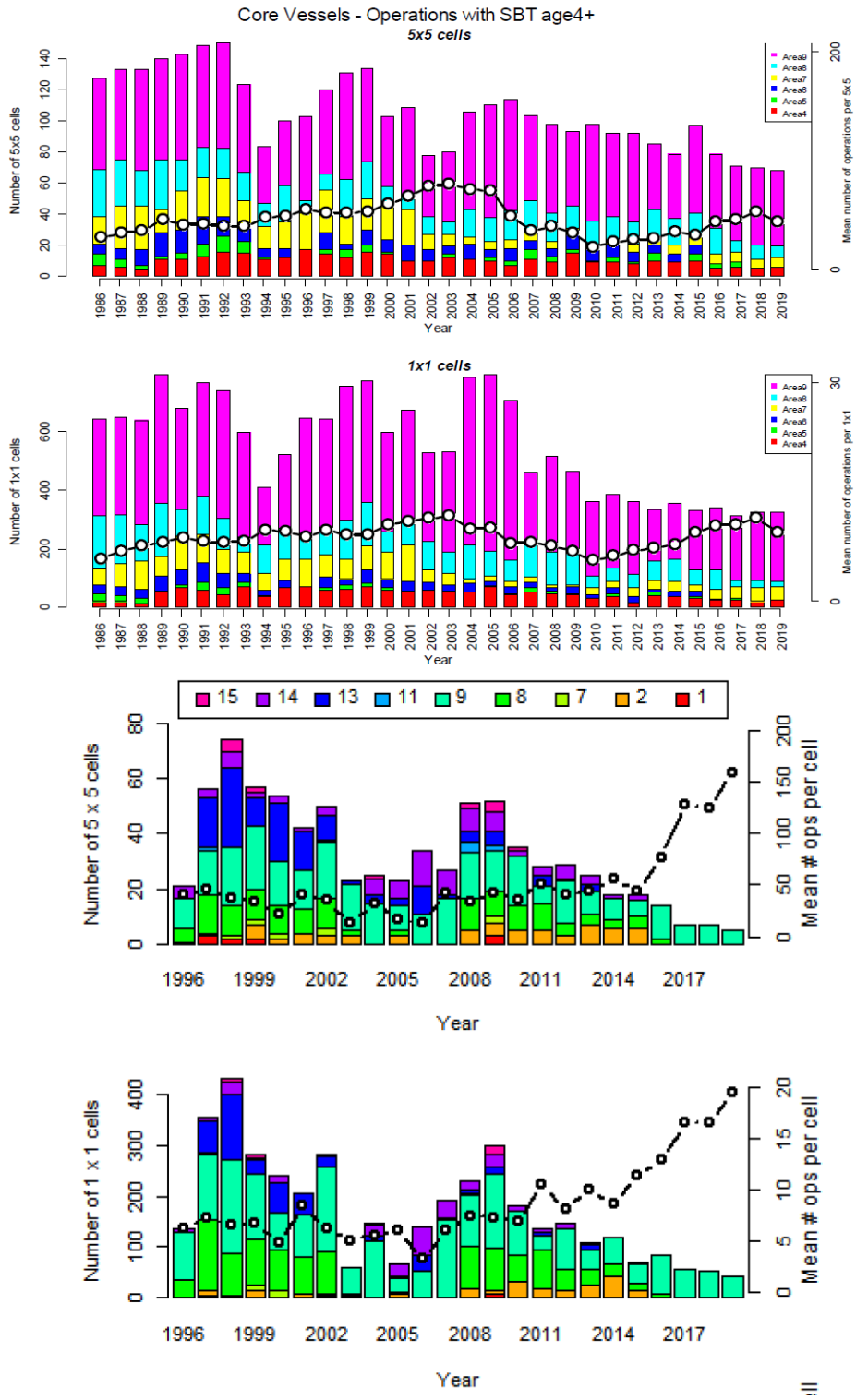

Figure A2: Number of cells fished by Japanese core vessels (above) and Korean vessels (below), copied from CCSBT-ESC/1909/BGD05 and CCSBT-ESC/1909/39. (Upper plot for each fleet): The bars represent the number of major cells (5x5° by month) fished by CCSBT statistical area and year, see left y-axis. 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.

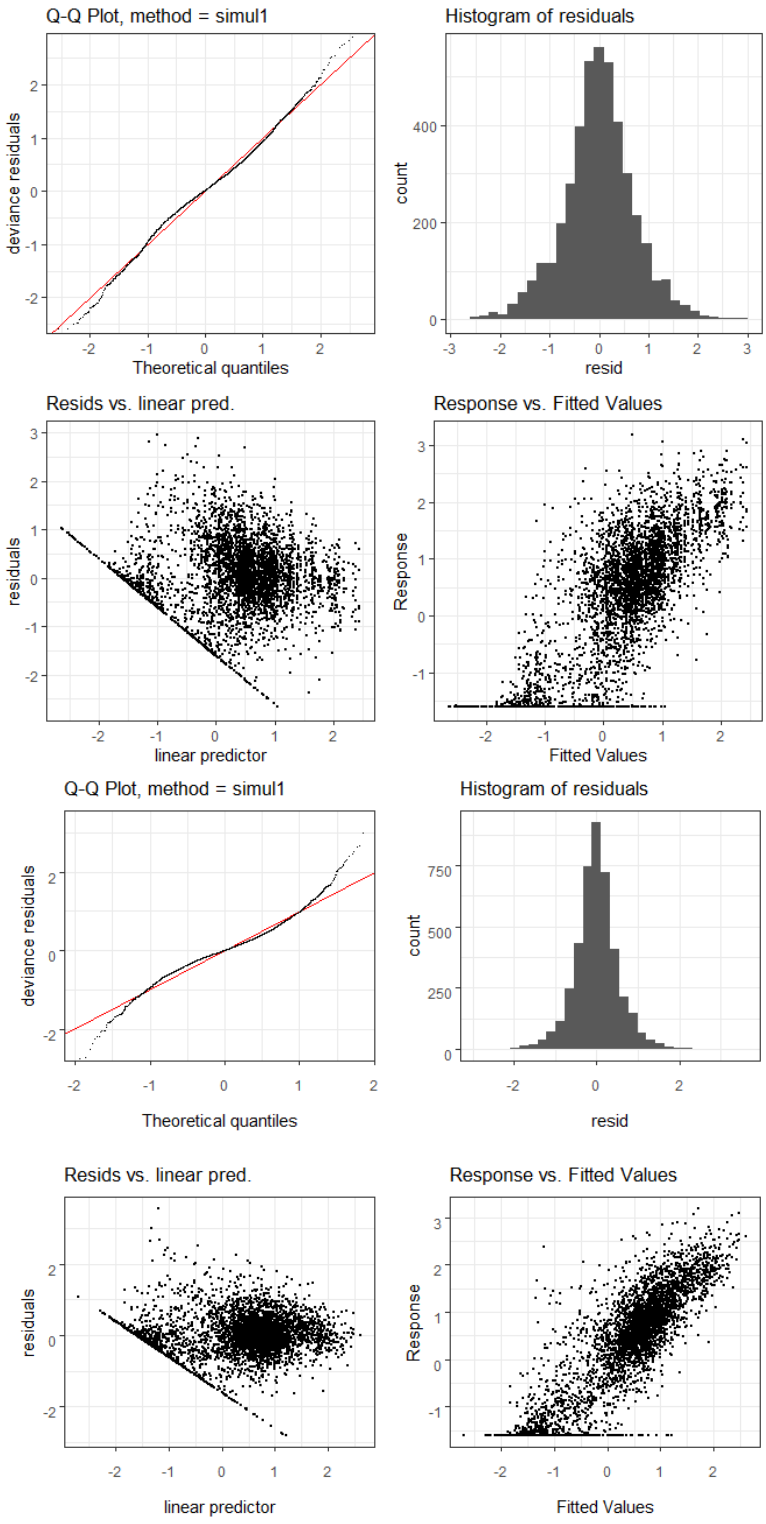

Figure A3: Residual checking plots for the simplified Base model (above) and the gam11 model (below).

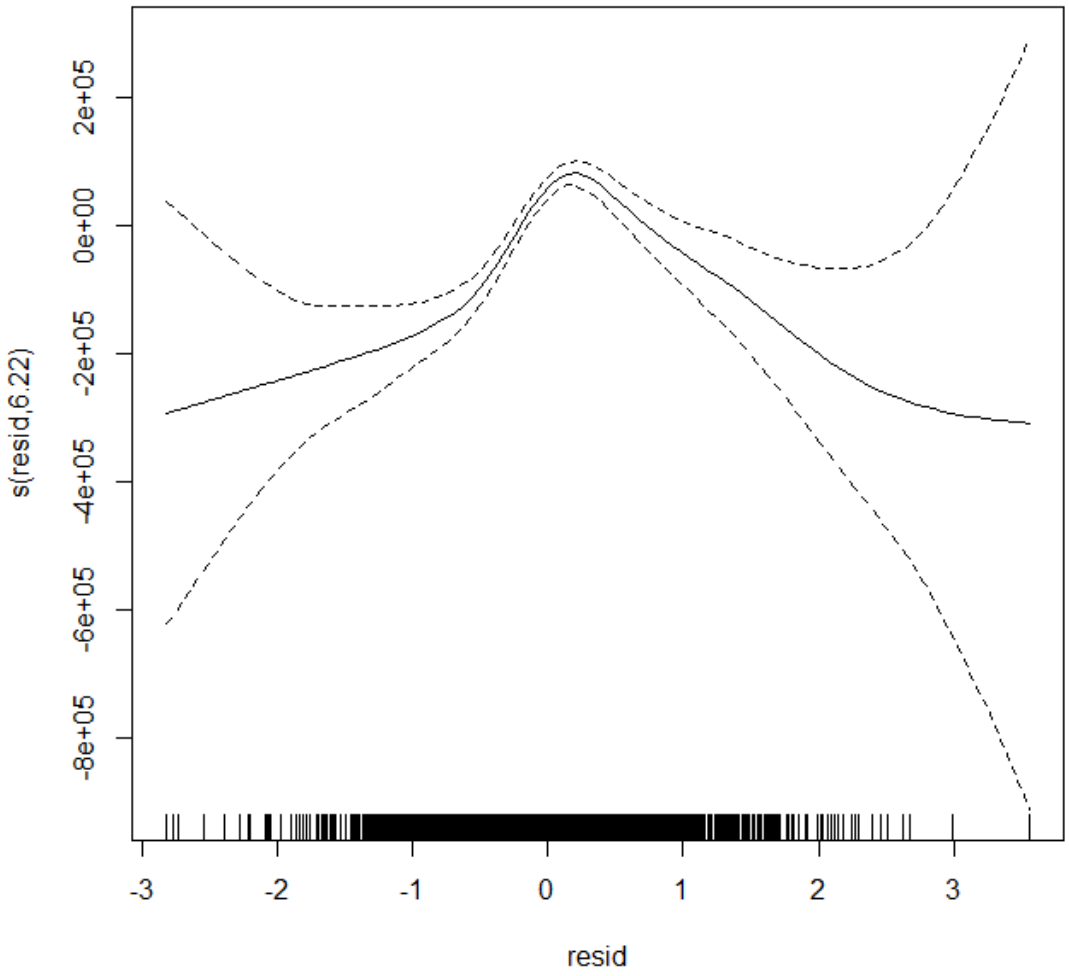

Figure A4: Relationship between residuals from the model gam11 and the effort (hooks) in the stratum, showing that higher variability (the tails of the residual distribution) is associated with less effort in a stratum.

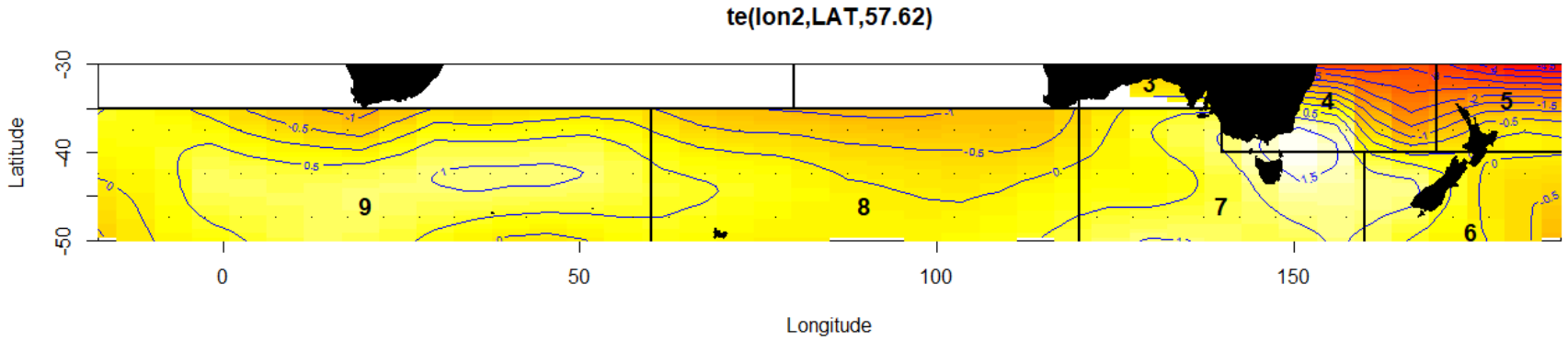

Figure A5: Two-way lon:lat smoother from model gam11.

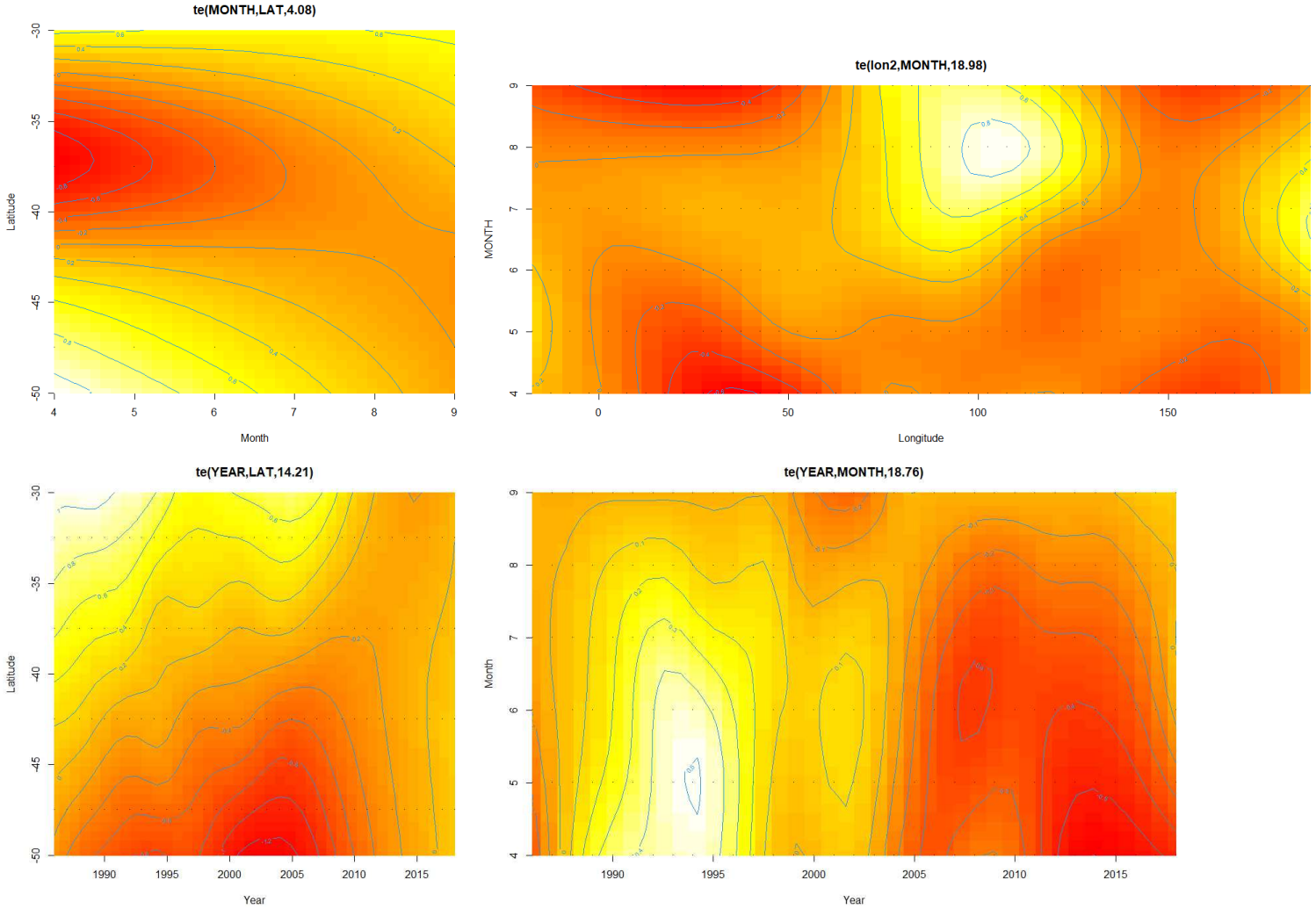

Figure A6: Smoothers from model gam11: mn:lat, lon:mn, yr:lat, and yr:mn.

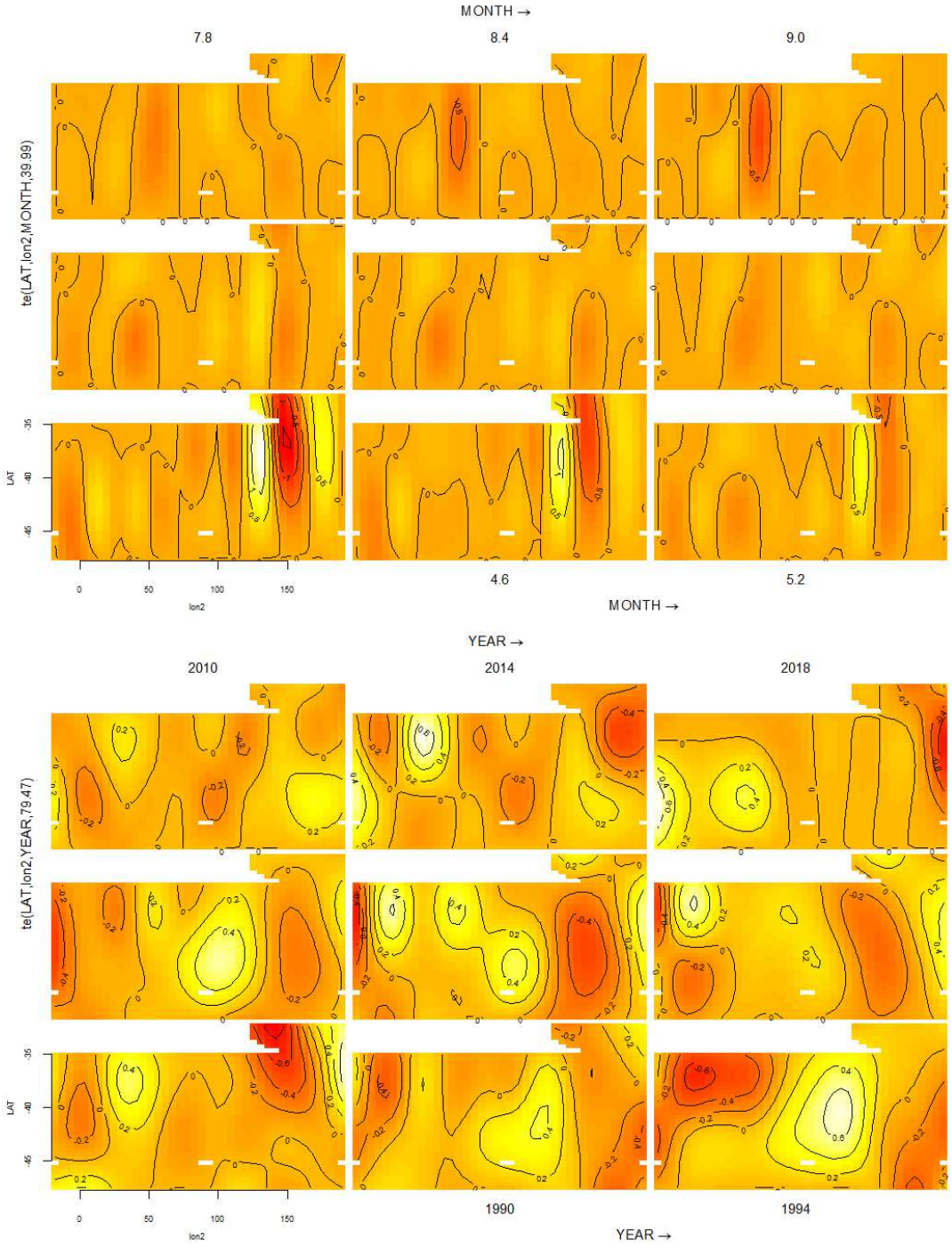

Figure A7: Three-way interaction smoothers from model gam11: lat:lon:mn (above) and lat:lon:yr (below).