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

Simon Hoyle

Hoyle Consulting, simon.hoyle@gmail.com.

Working Paper prepared for the 12th Operating Model and Management Procedure Technical Meeting, CCSBT, 20-24 June 2022, Hobart, Australia.

Abstract

This paper reports work to validate approaches used in the newly developed primary CPUE index of SBT abundance, based on generalized additive models (GAMs) with spatiotemporal smoothers, and a delta lognormal approach. Map time series showed that the temporal and spatial distributions of both fishing effort and the highest catch rates have changed between 1986 and 2020, while the spatial and temporal extents of fishing effort have declined. Simulated data were generated from the best models fitted to the aggregated dataset and used to explore the effectiveness of different model configurations for dealing with these changing distributions. The principal GAM models produced unbiased estimates with the simulated data, while GLM models and less flexible GAM smoothers provided biased indices. Manipulating the simulated dataset to produce a large rapid change in fish distribution resulted in moderately biased indices. Increasing the effort concentration through time to focus effort on areas with higher CPUE also resulted in estimation bias, particularly at the end of the time series when concentration was greatest. This bias may be due to loss of information from the dataset rather than model failure, and it may be helpful to increase the information via models that include data from other fleets as well as Japan. In general, GAM models provided less biased indices than either a GAM equivalent to the variable squares method (GAM_VS) or a combined model (w0.8) approach.

Introduction

The CPUE standardization methods used for SBT are being updated because of problems with recent estimates, particularly an anomalously high value in 2018 (CCSBT, 2020). The main reason for these analytical problems was identified as increasing aggregation of fishing effort, together with a method that relied on data availability in all strata, leading to sparse data and consequent parameter estimation problems. Analyses in 2020 and 2021 (Hoyle, 2021; Hoyle, 2020) developed an alternative approach using generalized additive models (GAMs) implemented with the R package *mgcv* (Wood, 2011). Data were fitted with multi-dimensional smoothers which share information among adjacent values of continuous variables. These approaches were developed using the aggregated datasets held by CCSBT, and further development was required to apply them to the operational data held by Japan (CCSBT, 2021).

These analyses followed previous work (Itoh and Takahashi, 2019; Nishida and Tsuji, 1998) in assuming that the uncertainty distribution was lognormal after adding a constant to the CPUE values. This constant was defined as 10% of the mean CPUE across the whole dataset (Campbell et al., 1996; Campbell, 2004). This historically common 'lognormal constant' approach has been superseded (Bellego and Pape, 2019; Maunder and Punt, 2004) by methods that either permit zero observations, or use a hurdle process to separately model the probability of nonzero catch and catch rates in nonzero catches (Lo et al., 1992). This hurdle (delta lognormal) approach was developed and applied in 2021 (Hoyle, 2021).

Work for 2022 involved determining whether there was an ongoing need for the variable squares (VS) approach, given that the flexibility of the GAM allows it to track and account for changes in fish and effort distribution through time. This involved mapping in GAM predictions (SE or CVs by month) to evaluate possible threshold approaches for dropping cells by year for a VS diagnostic/sensitivity test. There was also a need to prepare operational data simulations to test alternative GAM models and advise on a subset of models to test on real data. However, realistic operational data simulations could not be completed in the time available. Instead, the fitted models were used to generate simulated datasets that in many characteristics were similar to the observed aggregated data. These datasets were then used to explore the reliability of alternative approaches to fitting the data. In addition, the simulated data were manipulated to explore effects on model results of a) increasing effort concentration, and b) large changes in fish and effort distribution.

Methods

Input data

These analyses were based on a different dataset from that used for primary analyses (e.g., Itoh & Takahashi 2019), which is only available to Japanese scientists. The available dataset was sufficiently similar to the primary dataset to provide useful insights, but conclusions should be reconfirmed using the primary dataset. The main differences between the two datasets are listed below.

- The primary dataset is available as operational (set by set) data (but may be aggregated for the main analysis) whereas the available dataset is aggregated.
- 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 data file 'CPUEInputs_2022_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.

Error distribution assumption

Using the variables and interactions previously selected (Hoyle, 2021; Hoyle, 2020), models were run using the delta lognormal approach.

Models started with the same sets of variables on the right-hand side of the equation (RHS). The *mgcv* package uses the offered terms and initial basis dimension (*k*) as a starting point for a search. The *k* parameter sets the upper limit on the degrees of freedom associated with a smooth s, while for a *te* or *ti* tensor product smooth the upper limit is the product of the *k* values for each marginal smooth.

Models were fitted using either the generalized cross-validation (GCV), maximum likelihood (ML), or restricted maximum likelihood (REML). Delta models were fitted using ML to compare models with different number of terms, and REML in final models to estimate the smooth terms. Positive models were fitted using GCV to compare models with different number of terms, and REML in final models to estimate the smooth terms. Positive models were unstable when fitted with ML.

Wood (2011) recommends that models with multiple levels of interactions should specify main effects using either *s()* or *ti()* and interaction terms with *ti()*. Models were fitted using *ti()* for all terms.

RHS = yf + ti(lon2, k=40) + ti(LAT, k=4) + ti(MONTH, k = 6) + ti(lon2, LAT, k = c(40,4)) + ti(MONTH, LAT, k = c(6,4)) + ti(lon2, MONTH, k = c(10, 5)) + ti(YEAR, LAT, k = c(20, 4)) + ti(YEAR, MONTH, k = c(20, 5)) + ti(lon2, YEAR, k =c(10, 9)) + ti(LAT, lon2, MONTH, k = c(4,15, 6)) + ti(LAT, lon2, YEAR, k = c(4,10, 9))

As a final step the model was specified using a 'shrinkage' version of the cubic spline smooth (bs = "cs"), which can penalise a curve to zero and effectively eliminate it from the model.

The lognormal model used *log(cpue)* as the response, with identity link and gaussian error distribution, while the binomial model used '*cpue* != 0' as the response, with logit link function. Binomial models added effort to the RHS formula above, to account for the effect of effort on the probability of non-zero catch in a stratum. Effort was included as a spline rather than a straight line or offset, to allow for potential nonlinearity in the relationship.

catch > 0 ~ RHS + s(log(N_HOOKS))

Table 1: Settings used in mgcv to compare models with different distributions.

Distribution	Family	Dataset	response	Link function	Likelihood
Binomial (DLN)	Binomial	all	cpue > 0	logit	REML
Lognormal (DLN)	Gaussian	nonzero	log(cpue)	identity	ML

delta11 <- gam(cpue > 0 ~ formula, data = a, gamma = 2, method = 'REML', family = binomial)

pos11 <- bam(log(cpue) ~formula, data = apos, gamma = 2, method = 'REML')

Model diagnostics

To explore how data aggregation may affect residuals and consistency with model assumptions, patterns in the residuals from the lognormal constant model were explored by fitting generalized additive models. First, the absolute value of residuals was modelled as a function of effort: abs(resid) ~ s(n_HOOKS), and the relationship was plotted. Next, CPUE predicted from the fitted model in strata with effort was modelled as a function of effort after adjusting for year: pred_CPUE ~ s(N_HOOKS) + yf. Finally, residual variance was modelled as a function of predicted CPUE after adjusting for year: abs(residuals) ~ s(predicted CPUE) + yf.

Models fits were initially explored using the standard R diagnostic plots as implemented in the mgcViz package (Fasiolo et al., 2020). These include a Q-Q plot with theoretical quantiles based on simulated residuals, a histogram of residuals, a scatter plot of predicted values versus residuals, and a scatter plot of predicted versus observed values.

Further diagnostic analyses were carried out using the DHARMa package (Hartig, 2020). These diagnostics involve simulating residuals from the fitted model and comparing the distributions of simulated and observed residuals, with a number of plots and statistical tests available.

Q-Q plots were generated using simulated residuals, and each Q-Q plot also ran a uniformity test (Kolmogorov-Smirnov test for overall uniformity of the residuals) and a dispersion test (a simulation-based test for over/under-dispersion). For further details see Fasiolo et al. (2020).

Boxplots of residuals were plotted for each categorical variable, to compare the relationship between the quantile distribution and the predicted response. For each boxplot tests were run to check whether residuals within each group were distributed according to model assumptions (multiple Kolmogorov-Smirnov tests, with adjustment of p-values for multiple testing), and whether the variance between groups was heterogeneous (using a Levene test). Residuals were also plotted against the smoothed year effect and model predictions, and tests carried out to detect whether quantiles deviated from the expected values. For further details of these tests see Fasiolo et al. (2020). Finally, the median residuals were plotted against rank-transformed model predictions.

Preparation of indices

The density of each stratum (year by month by grid cell) was predicted based on the centre of the cell. Ocean areas of all grid cells were calculated (Hoyle and Langley, 2020). Stratum abundances were predicted by multiplying stratum density by ocean area.

A time series of predicted abundance was calculated by summing predicted abundances across strata for each year. Abundance indices were obtained by dividing the abundances for each year by the mean of all years.

Spatial cells were included in the dataset if represented by at least 15 nonzero observations.

The GAM_VS indices were prepared in the same way as the standard GAM indices, except that predicted abundances for all month x spatial cell x year strata without effort were removed from the time series before summing across strata for each year.

R code

All R code is available at the github repository https://github.com/hoyles/R_ccsbt_cpue. Please email simon.hoyle@gmail.com to request permission to access the repository.

Testing the need to retain the VS method

In the past, the VS method has been applied to account for possible situations in which areas with high fish densities change from year to year, but catch rates remain high because fishing effort follows the high densities of fish. Estimates from the VS approach include only the predictions from strata with fishing effort.

Maps

In order to explore how effort concentration has changed through time, along with the distributions of catch rates and estimation uncertainty, five map series were generated. These were annual plots of a) the number of month x spatial cell strata fished; b) predicted proportions of non-zero strata; c) expected catch rates in positive sets; d) uncertainty in the predicted proportions of non-zero strata; and e) uncertainty in expected catch rates in positive sets.

Estimation bias and time area interactions

New simulated datasets were generated from the fitted models delta15g02s and pos15g02s using the simulate.glm() function from the R package mgcViz (Fasiolo et al., 2020). Performances of different estimation model configurations were compared for their ability to reproduce the 'true' year effect trend. Alternative estimation models (Table 4) included the following:

a) The standard (original) smoothers (delta15g02s and pos15g02s).
b) Reduced-flexibility smoothers with lower starting k values (see below).
c) The smoothers used in the operational data analysis, which lack 3-way interactions for the delta component and have some higher starting k values for the lognormal component.
d) The GLM approach.

Simulated scenarios

Three sets of scenarios were simulated to represent alternative data scenarios, to investigate the ability of the constant-squares gam and its VS equivalent to estimate indices under different types of challenge.

The first scenario 'South' represented a situation in which the distribution of the population changed substantially through time. The distribution change was simulated to occur once, with complete transition in a single year. In this scenario the latitude of all simulated catch and effort records for years after 2000 was moved south by 5°.

The second and third scenarios 'contract20' and 'contract40' represented situations in which effort concentration increased through time at a faster rate than in the existing dataset. These scenarios simulated effort concentration of x% by removing from the dataset all effort strata with CPUE lower than the bottom x^{th} percentile. Percentiles were calculated within each year and based on the number of strata, without considering the number of hooks in each stratum.

In the contract20 scenario, effort concentration began in 1986 at 0%, and increased linearly to reach 20% in the year 2020. Similarly, the contract40 scenario saw effort concentration increasing linearly from 0% in 1986 to 40% in 2020.

Results

Maps

Maps of the data coverage (Figure 1) show that the distribution of fishing in the western statistical area 9 near South Africa has been somewhat variable from year to year. For example, in every year from 1986 to 2001 fishing effort was reported in the grid cell immediately south of South Africa in every month from April to September. From 2002 the proportion of fished strata in the west at the 37.5 S latitude decreased, and by 2017 the great majority of fished strata in the west were at the 42.5 S latitude.

In statistical area 8 to the west of Australia, fishing activity diminished from 1986 to 1998. From about 2006 there was an increase in the number of fished strata in the north and a reduction in the south, which continued until 2020. Statistical areas 4 to 7 to the south and east of Australia saw a marked and sustained reduction in the number of fished strata.

Maps of the delta component (Figure 2) show a high proportion of positive events in the south, with non-zero catch rates in the vast majority of strata to the south of 40S.

Maps of the positive component (Figure 3) in western area 9 show higher mean estimates further south until about 2000, when the highest catch rates begin to occur further north in the 42.5S latitude. However, after about 2013 higher catch rates begin to be observed in the southwest of area 9 at the 47.5 latitude. Further east, in areas 4 to 7, catch rates in positive catches are consistently estimated to be highest close to Australia in the South Australia bight and northeast of Tasmania.

Maps of uncertainty in the delta component (Figure 4) show almost no uncertainty in the south, since zero strata occurred very rarely. Uncertainty estimates elsewhere have increased in recent years, consistent with the reduction in the number of fished strata.

Maps of uncertainty in the positive component (Figure 5) similarly show higher levels of uncertainty in recent years as the number of fished strata has declined. In general, uncertainty is lowest at the 42.5 S latitude where the number of fished strata is greatest.

Estimation bias and time area interactions

The distributions of indices based on predictions from models fitted to 50 simulated datasets were compared. Predictions from all of the GAM models were much less variable than predictions from the GLM (Figure 6).

Models using standard k-value settings for the smoothers, and the adjusted settings adopted for the operational data analyses, provided unbiased estimates of the indices on which the simulations were based (Figures 6 and 7). Models using less flexible smoothers, with lower k-values that limited the number of parameters available to the cubic splines, provided more biased estimates of the indices, with small negative bias in the first 15 years of the timeseries and quite a large positive bias in the last 6 years of the time series. The GLM also provided negatively biased indices at the start of the time series, but biases at the end of the series were more variable, with some positive and some negative.

In all cases estimates based on the VS method were positively biased for the first 8 years of the time series and very negatively biased for 4 out of the last 5 years. Predictions based on combining the standard prediction with the VS predictions were also negatively biased at the end of the time series, and generally performed worse than the standard GAM predictor.

Simulated scenarios

Indices generated from the 'South' scenario, in which distribution changed substantially and rapidly in the year 2000 in the middle of the time series, were positively biased before and negatively biased after 2000. Biases in the GAM_VS predictions were considerably more variable, sometimes better and sometimes worse than the GAM predictions, but were particularly biased in the most recent 5 years. Combined predictions w0.8 and w0.5 were also slightly worse than the GAM predictions for most of the last 5 years.

Indices generated from the 'Contract 20' and 'Contract 40' scenarios were affected by the higher levels of effort concentration, with the GAM indices more biased than the standard indices in the last 2 (Contract 20) or 6 (Contract 40) years. In both cases the GAM_VS indices were much more biased than the GAM indices. For 'Contract 20' the combined w0.8 indices were inferior to the GAM indices for most of the period, and for 3 of the last 5 years, although slightly better in the last 2 years. However, for 'Contract 40' the w0.8 indices were better than the GAM indices for most of the period, and particularly during the last 5 years when effort concentration was highest.

Discussion

The distributions of both fish and effort changed through the time series, in addition to varying seasonally. The increasing sparsity of the fishing distribution was very marked, with far fewer fished strata in 2020 than in 1986. The factors motivating the effort contraction are less well understood. Initially the contraction would been driven by reduced catch quotas. More recently, catch rates that increased faster than quotas are likely to have allowed vessels to catch their quotas with less effort. The contribution of technology improvements is unclear but may have increased vessels' ability to identify and target areas of higher abundance, thus increasing catch rates and reducing effort.

All GAM models performed much better than the GLM at fitting the data simulated from a similar GAM model, with less bias and greater precision in the estimated indices. Caution is warranted given that a model will always be capable of fitting to data simulated from the same model, so the GAM has an advantage over the GLM. Nevertheless, the better performance is consistent with the much better fits to the original dataset.

Similarly, the less flexible GAM models performed consistently worse than the standard GAM and the GAM optimised to fit the operational data, indicating the importance of giving the GAM sufficient flexibility to capture the patterns in the data. As well as having considerably less flexible initial k values than the standard GAM starting models, their main effects were slightly less flexible than the estimated models, since the degrees of freedom (df) for a tensor spline is equal to k-1. It was unclear which smooth model components were most important for improving the model fit and avoiding bias. However, some components were clearly unimportant, since the operational data approach fitted the data as well as or better than the standard approach, despite omitting all 3-way interactions from the delta model.

The flexibility of the GAM model makes it less affected than the GLM by the increasing concentration of the Japanese effort through time. The curvature of the smoothers means that effort can increase in a spatial cell where catch rates are consistently higher without affecting expected catch rates elsewhere. In contrast, better targeting within a statistical area may elevate expected catch rates

throughout the area for the less flexible GLM. Similarly, the less flexible version of the GAM, with less ability to localise strata with higher catch rates, showed positive bias at the end of the time series as effort became more concentrated.

The ability of the GAM and GAM_VS models to adjust for changing fish and effort distributions was explored using the 'South' scenario. By introducing a transition point and moving all effort south by 5 degrees in the year 2000, this scenario tested the flexibility of the year by latitude interaction term. Although the simulation results showed some bias, it was not large in comparison to the scale of the abundance trend, and simulated indices generally followed the true trend. Distribution changes observed in the data appear to occur more slowly than the one simulated here.

The 'Contract 20' and 'Contract 40' scenarios were used to explore how effort concentration may be affecting indices. Increasing concentration into strata with higher catch rates did appear to introduce bias, with a positive bias at the end of the time series when concentration was highest. Thus, the GAM models were not fully able to account for patterns that did not occur in the original model. However, this may be a data problem rather than a modeling problem. The increasing concentration into areas of higher catch rates reduces the available information about CPUE in areas with lower catch rates. Similarly, increasing effort concentration in the real world may be introducing bias into the indices due to lack of information about trends in areas with low catch rates. One way to address this problem may be to develop joint models that include data from other fleets as well as the Japanese fleet. Expanding the dataset in this way may reduce the problematic effects of effort concentration. Such analyses could be carried out using aggregated data. Comparing these joint indices with Japanese indices developed using the same models may help to indicate whether effort concentration is affecting the Japanese operational data indices.

The variable squares (VS) approach was originally designed to address this type of issue, but in these simulations appears to perform poorly. It is particularly badly affected by increasing effort concentration, because its abundance estimates, based on summing cells that include fishing effort, are negatively biased by reducing the number of strata. In this dataset, the number of fished strata has been reduced by the effects of quota controls on fishing, rather than by the lack of availability that is implicitly assumed by the VS approach.

All data in these simulations were aggregated to 5° cell by 1-month strata, so abundance patterns and effort concentration at finer scales such as the 1 ° cell could not be explored. There may be abundance patterns at these finer spatial scales, and vessels may have improved through time in their ability to identify and target them, but these data do not allow us to explore the issue.

In conclusion, the standard GAM models appear adequate to estimate abundance trends and adjust for distribution changes of fish and fishing effort. They perform better in most situations than models using the VS approach. However, there is some risk that increasing concentration of Japanese effort into fewer strata has introduced bias into the indices, or that this may occur or worsen in future with further effort concentration. It may therefore be helpful to expand the dataset to include effort from other fleets.

Bibliography

Bellego, C.; Pape, L.-D. Dealing with logs and zeros in regression models. Série des Documents de Travail; 2019

Campbell, R.; Tuck, G.; Tsuji, S.; Nishida, T. Indices of abundance for southern bluefin tuna from analysis of fine-scale catch and effort data. Second CCSBT Scientific Meeting. Hobart, Australia: Commission for the Conservation of Southern Bluefin Tuna; 1996

Campbell, R.A. CPUE standardisation and the construction of indices of stock abundance in a spatially varying fishery using general linear models. Fisheries Research. 70:209-227; 2004

CCSBT. Report of the Twenty Fifth Meeting of the Scientific Committee. Online: Commission for the Conservation of Southern Bluefin Tuna; 2020

CCSBT. Report of the Twenty Sixth Meeting of the Scientific Committee. Online: Commission for the Conservation of Southern Bluefin Tuna; 2021

Fasiolo, M.; Nedellec, R.; Goude, Y.; Capezza, C.; Wood, S.N. mgcViz: Visualisations for generalized additive models. Computer software] https://CRAN R-project org/package= mgcViz; 2020

Hartig, F. DHARMa: residual diagnostics for hierarchical (Multi-Level/Mixed) regression models. 2020

Hoyle, S. Potential CPUE model improvements for the primary index of Southern Bluefin Tuna abundance. CCSBT Extended Scientific Committee for the 26th Meeting of the Scientific Committee. Online; 2021

Hoyle, S.D. Investigation of potential CPUE model improvements for the primary index of Southern Bluefin Tuna abundance, CCSBT-ESC/2008/29. CCSBT Extended Scientific Committee 25. online; 2020

Hoyle, S.D.; Langley, A.D. Scaling factors for multi-region stock assessments, with an application to Indian Ocean tropical tunas. Fisheries Research. 228:105586; 2020

Itoh, T.; Takahashi, N. 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; 2019

Lo, N.C.H.; Jacobson, L.D.; Squire, J.L. Indices of relative abundance from fish spotter data based on delta-lognormal models. Canadian Journal of Fisheries and Aquatic Sciences. 49:2515-2526; 1992

Maunder, M.N.; Punt, A.E. Standardizing catch and effort data: a review of recent approaches. Fisheries Research. 70:141-159; 2004

Nishida, T.; Tsuji, S. 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; 1998

Wood, S.N. Fast stable restricted maximum likelihood and marginal likelihood estimation of semiparametric generalized linear models. Journal of the Royal Statistical Society: Series B (Statistical Methodology). 73:3-36; 2011

Tables

Table 2: Specifications for 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.

Label	**Factors**	**Smooth terms**								**4-way**
Base										
Base plus	.+mf:latf									
Base_noYrAr	.-yf:areaf									
glmmYrAr										
gam 2	yf+mf	lon,lat								
gam 3	yf+mf							mn,lon,lat		
gam 4	yf									All
gam 5	yf+llf									All
gam 6	yf+llf		mn,lat							All
gam 7	yf	lon,lat	mn,lat							All
gam 8	yf	lon,lat	mn,lat	lon,mn						All
gam 9	yf	lon,lat	mn,lat	lon,mn	yr,lat			mn,lon,lat	lat,lon,yr	
gam 10	yf	lon,lat	mn,lat	lon,mn	yr,lat	yr,lon	yr,mn			
gam 11	yf	lon,lat	mn,lat	lon,mn	yr,lat		yr,mn	lat,lon,mn	lat,lon,yr	
gam 12	yf	lon,lat	mn,lat	lon,mn	yr,lat	yr,lon	yr,mn	lat,lon,mn	lat,lon,yr	
gam 13	yf+mf	lon,lat	mn,lat	lon,mn	yr,lat	yr,lon				
gam 14	yf	lon,lat	mn,lat	lon,mn	yr,lat	yr,lon,mn	yr,mn	lat,lon,mn	lat,lon,yr	

Table 3: Specifications for models fitted with smooth terms specified using ti(). The factors column reports all variables included as categorical variables. Smooth terms include one-way terms and two-way and three-way interactions. All models include categorical variables for year (yf) and one-way smooth terms for longitude (lon), month (mn), and latitude (lat). Within interactions year is specified as a smooth term (yr). Terms marked 0 were included but effectively removed via shrinkage.

Model	Factors	lon, lat	mn, lat	lon, mn	yr, lat	yr, lon	yr, mn	yr, lon, mn	lat, lon, mn	lat, lon, yr	lat, mn, yr
9	yf	x	x	x	x	x			x	x	
10	yf	x	x	x	x	x	x				
11	yf	x	x	x	x	x	x		x	x	
13	yf+mf	x	x	x	x	x					
14	yf	x	x	x	x	x	x	x	x	x	
15	yf	x	x	x	x	x	x	x	x	x	x
15s	yf	x	0	x	x	x	x	x	0	x	x
16	yf	x		x	x	x	x	x		x	x

Table 4: Values of k (initial degrees of freedom) applied to each component of the delta and lognormal models, under the standard, reduced, and operational data scenarios, along with (Est.) the estimated degrees of freedom in the models used to simulate the data. Note that the 'Est.' value for each smoother is for all variables in the smoother combined.

	Delta				**Lognormal**			
	Est.	**Std.**	**Red.**	**Oper.**	**Est.**	**Std.**	**Red.**	**Oper.**
ti(lon2)	12.7	40	13	20	20.0	40	20	20
ti(LAT)	2.1	4	3	4	2.9	4	3	4
ti(MONTH)	3.7	6	4	5	4.2	6	4	6
ti(lon2,LAT)	6.1	40, 4	3,3	10,4	35.3	40, 4	20,3	20,4
ti(MONTH,LAT)	0	6,4	3,3	5,4	6.0	6,4	3,3	6,4
ti(lon2,MONTH)	4.8	10,5	3,3	10,5	7.8	10,5	4,3	20,6
ti(YEAR,LAT)	4.0	20,4	3,3	10,4	8.3	20,4	5,3	20,4
ti(YEAR,MONTH)	6.1	20,5	3,3	10,5	9.9	20,5	4,3	20,6
ti(lon2,YEAR)	0.9	10,9	3,3	10,10	13.7	10,9	4,3	20,20
ti(LAT,lon2,MONTH)	0	3,3,3	3,3,3		15.0	4,15,6	3,6,3	4,20,6
ti(YEAR,lon2,MONTH)	8.2	3,3,3	3,3,3		2.8	3,3,3	3,3,3	15,15,6
ti(LAT,lon2,YEAR)	82.8	4,10,9	3,3,3		34.0	4,10,9	3,5,3	4,15,15
ti(LAT,MONTH,YEAR)	1.1	4,5,9	3,3,3		5.7	4,5,9	3,3,3	4,6,20
ti(log(N_HOOKS))	1.5	10	3	10				

Figures

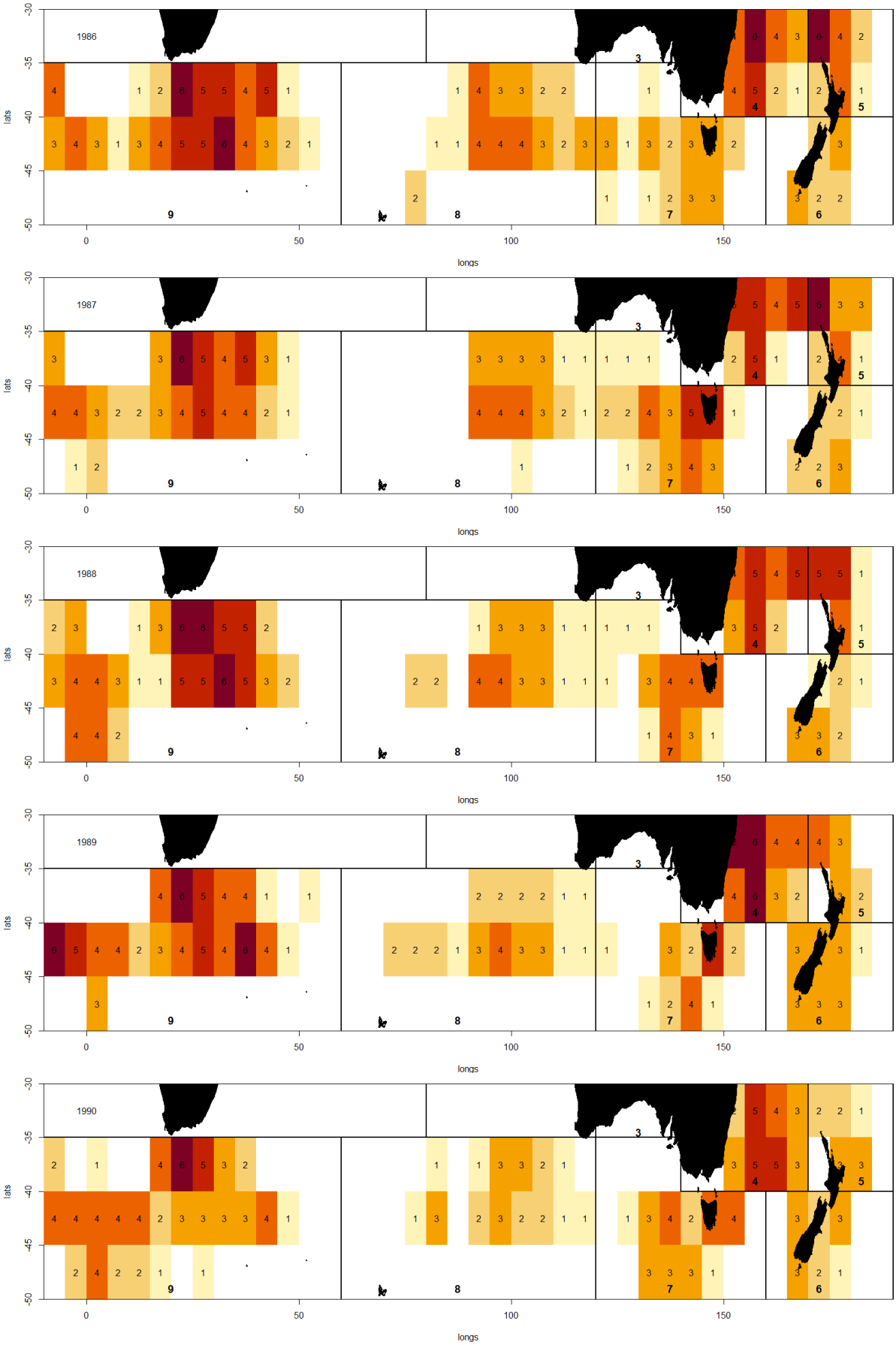

Figure 1: Data coverage by year, as a count of the number of monthly strata with reported effort per cell.

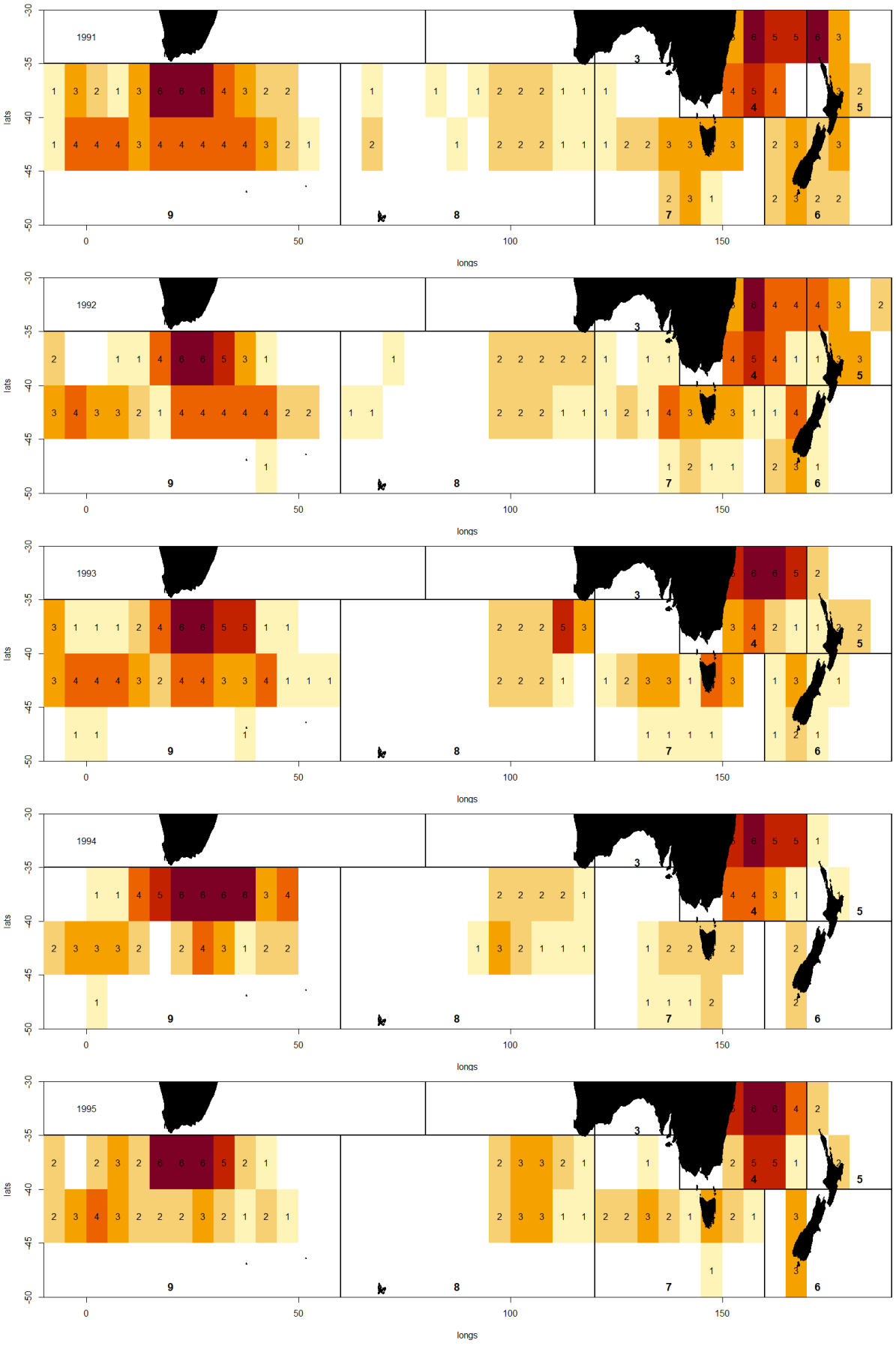

Figure 1 (cont.): Data coverage by year, as a count of the number of monthly strata with reported effort per cell.

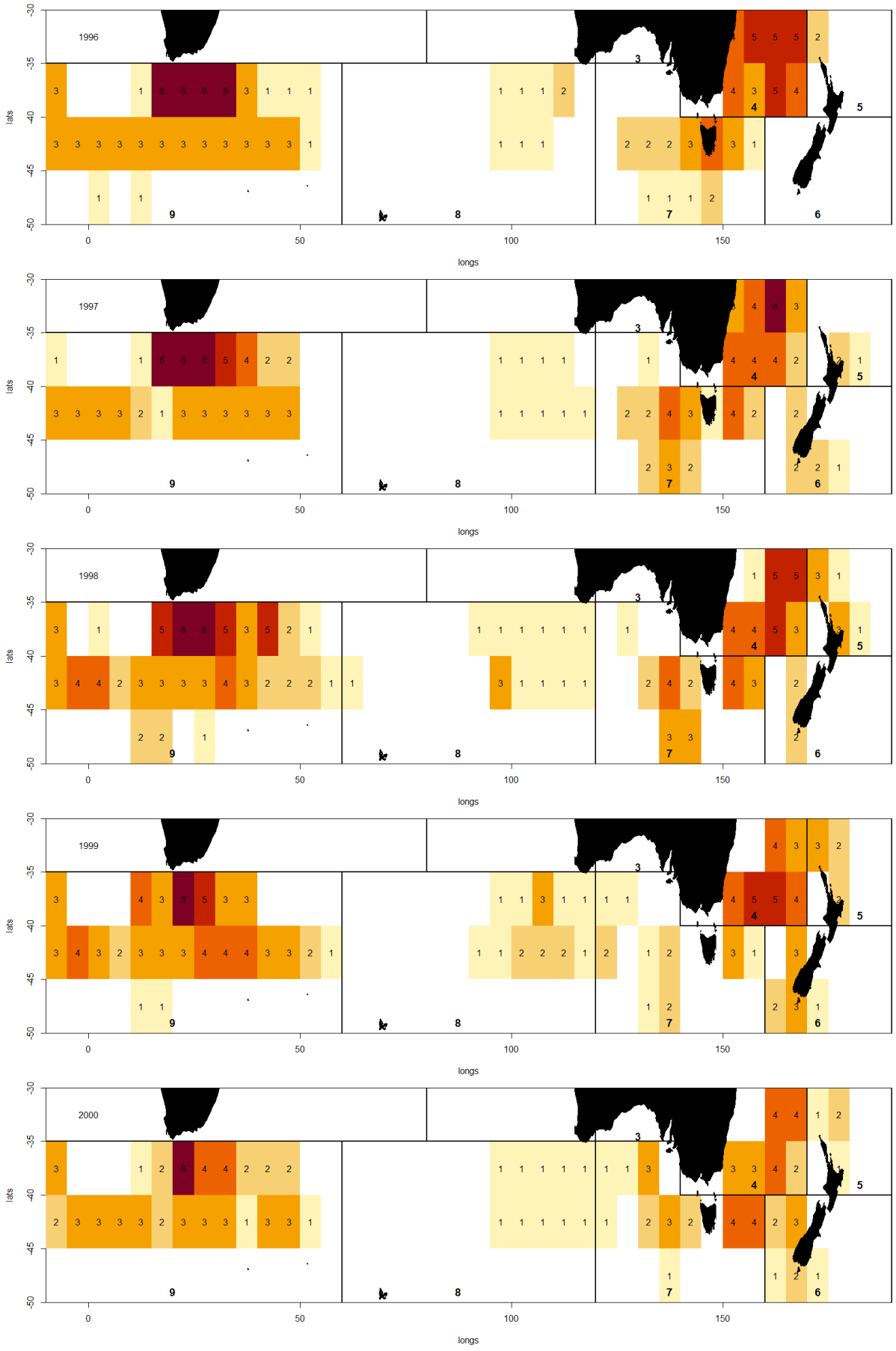

Figure 1 (cont.): Data coverage by year, as a count of the number of monthly strata with reported effort per cell.

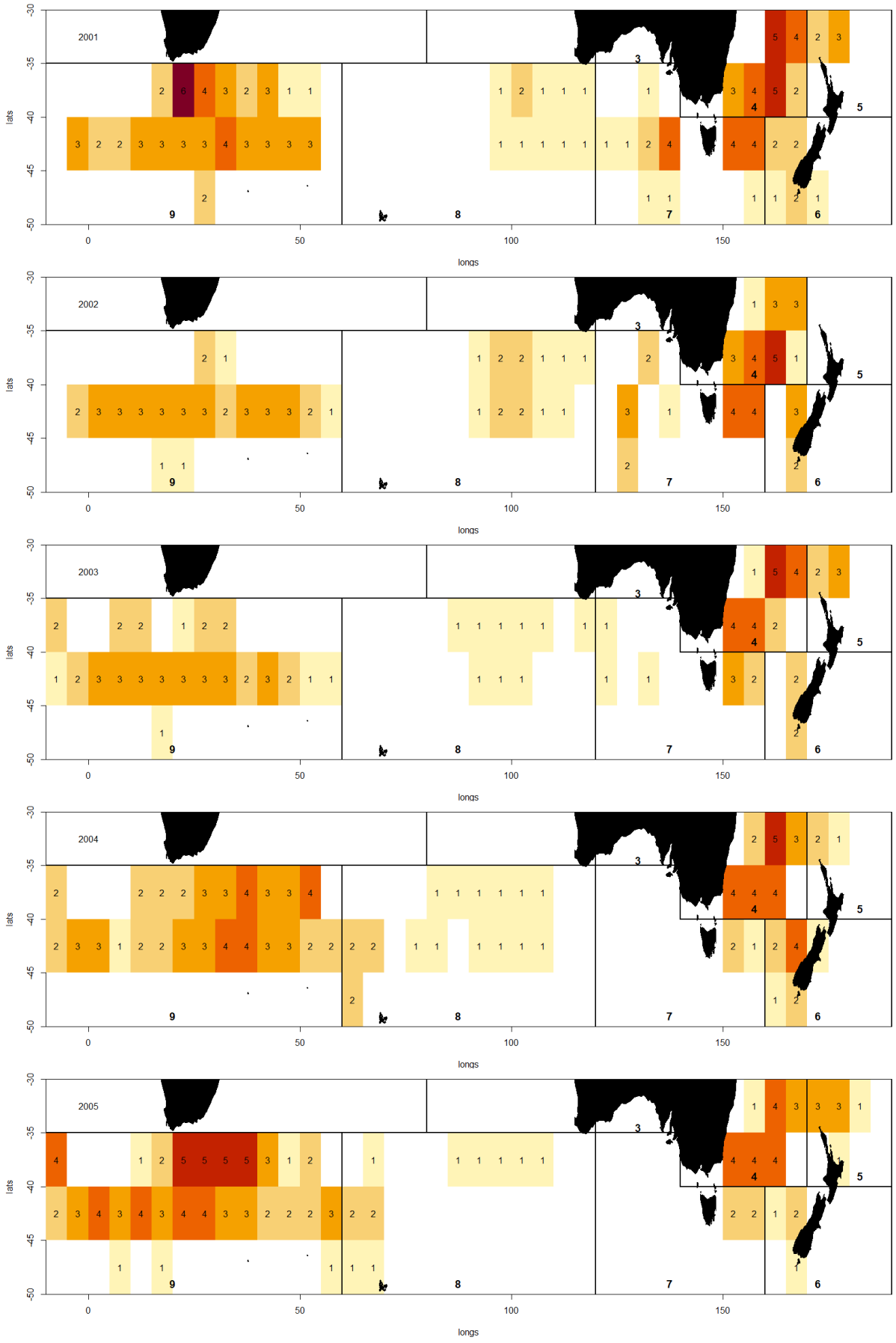

Figure 1 (cont.): Data coverage by year, as a count of the number of monthly strata with reported effort per cell.

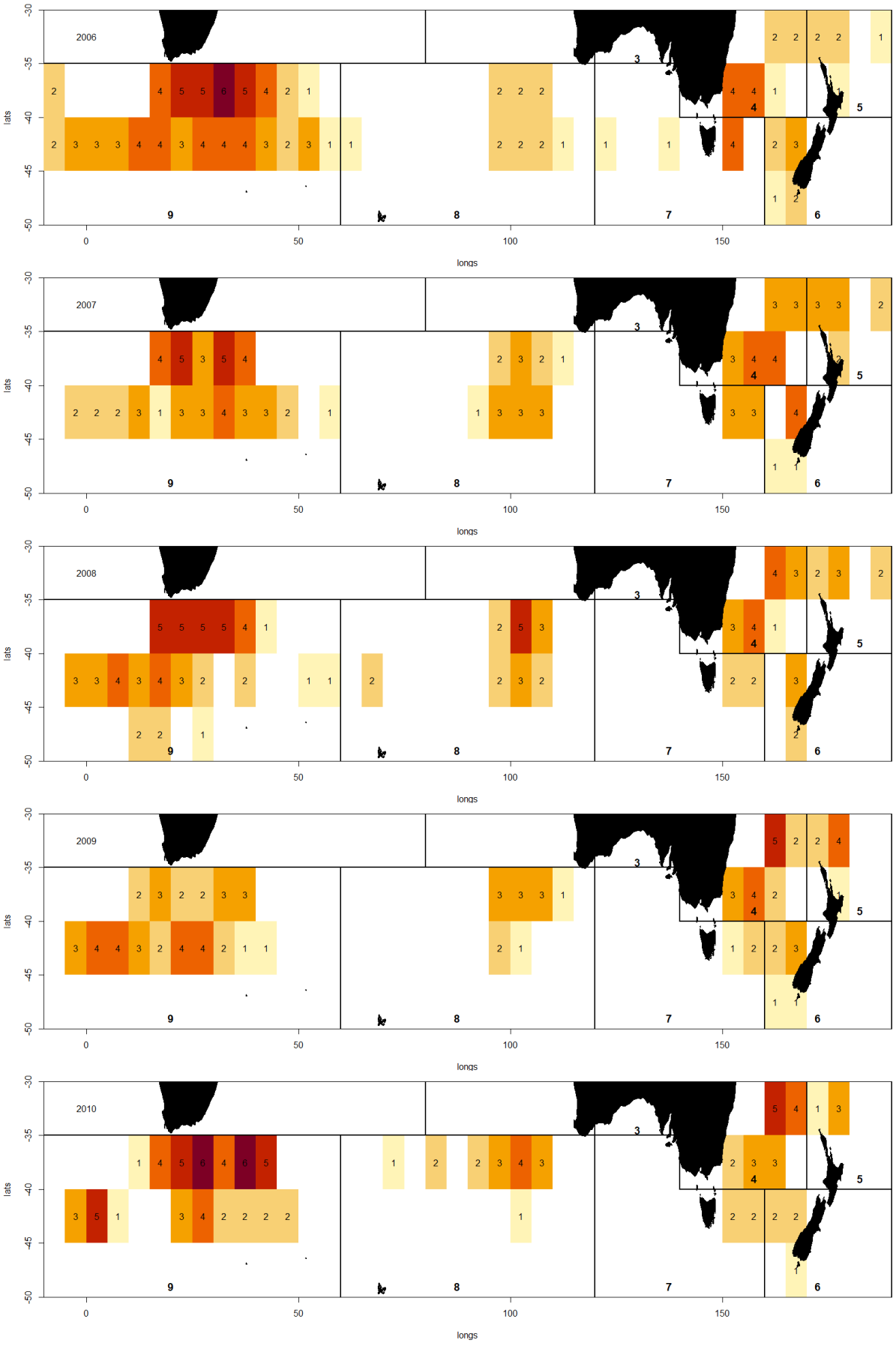

Figure 1 (cont.): Data coverage by year, as a count of the number of monthly strata with reported effort per cell.

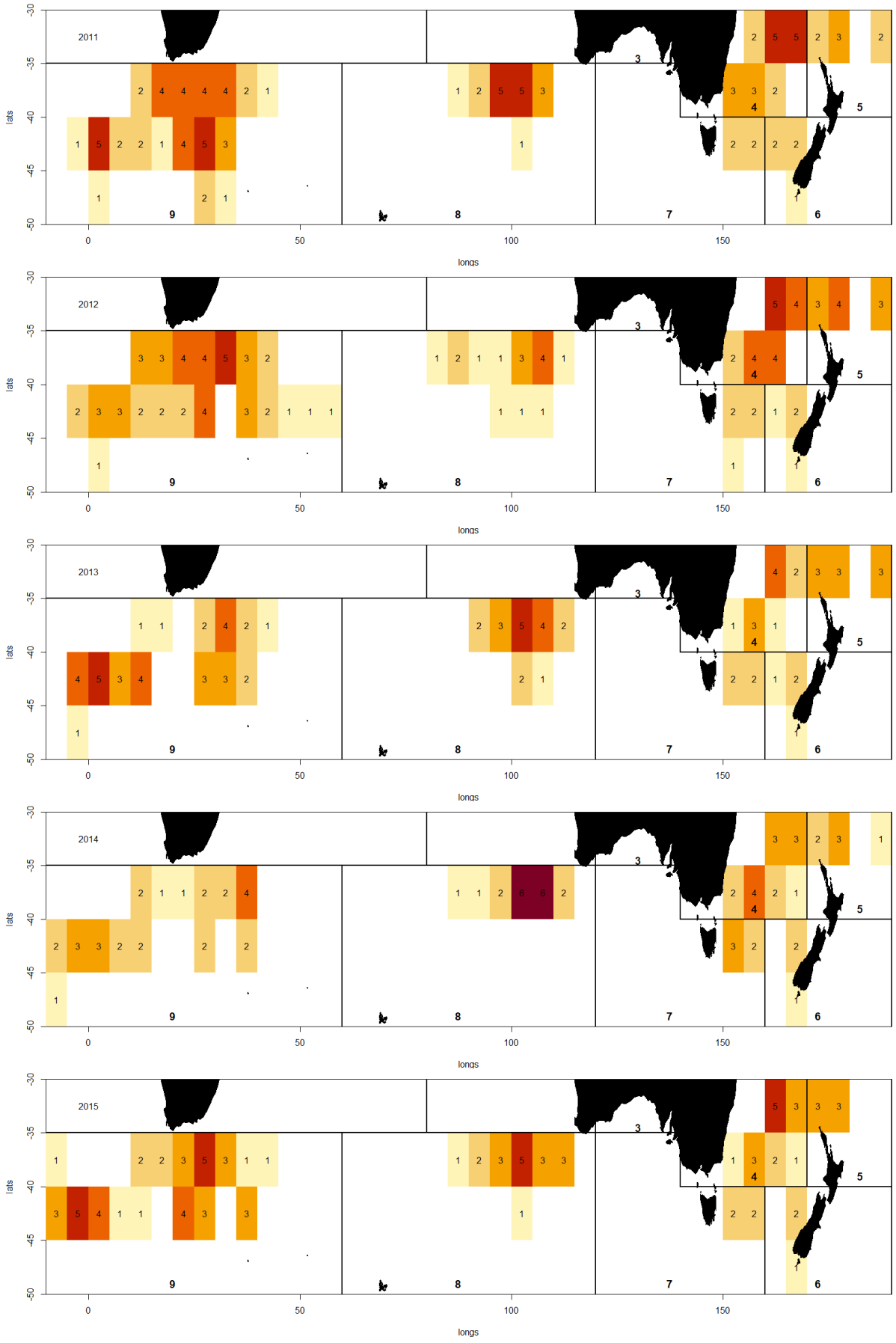

Figure 1 (cont.): Data coverage by year, as a count of the number of monthly strata with reported effort per cell.

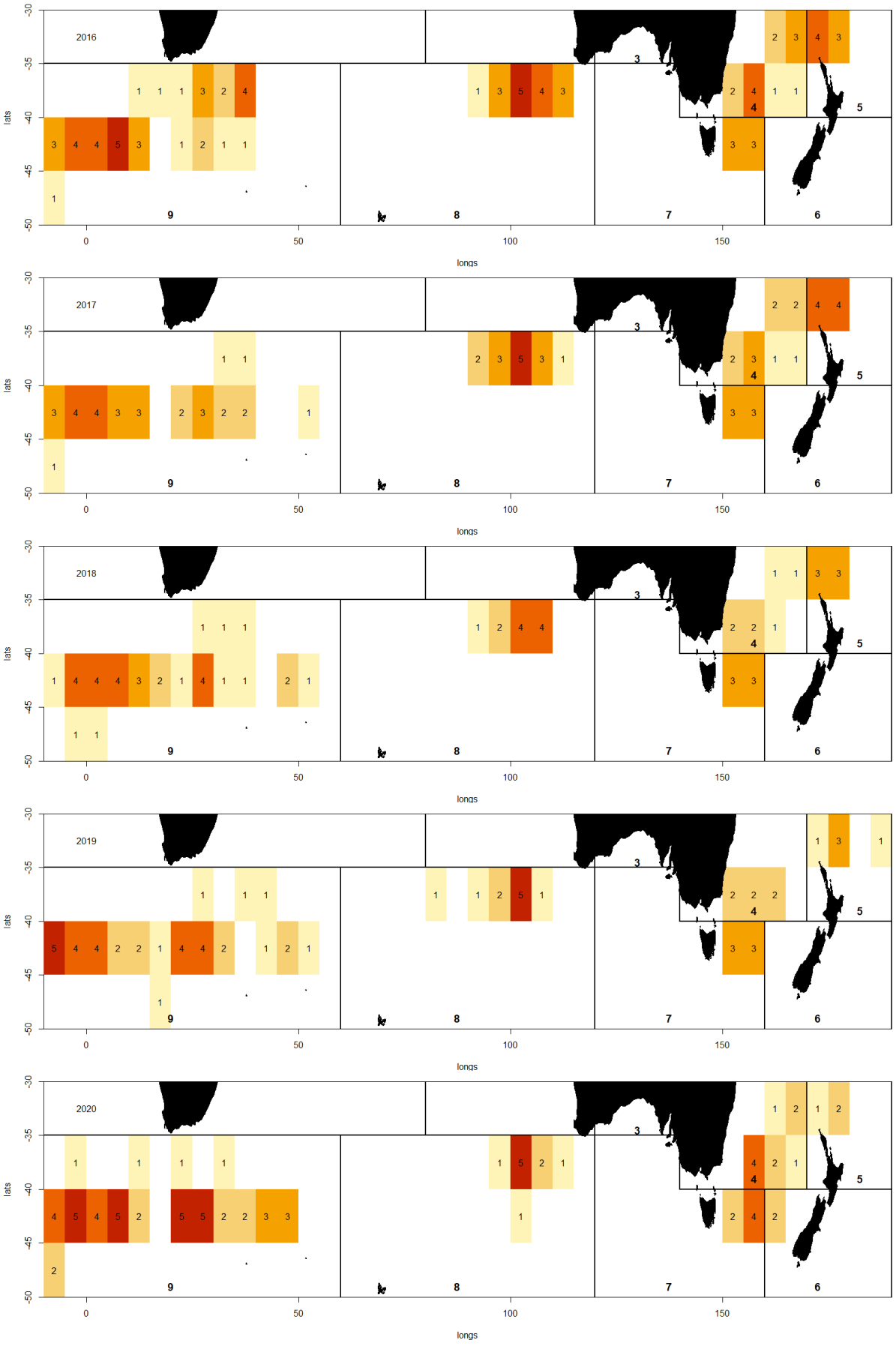

Figure 1 (cont.): Data coverage by year, as a count of the number of monthly strata with reported effort per cell.

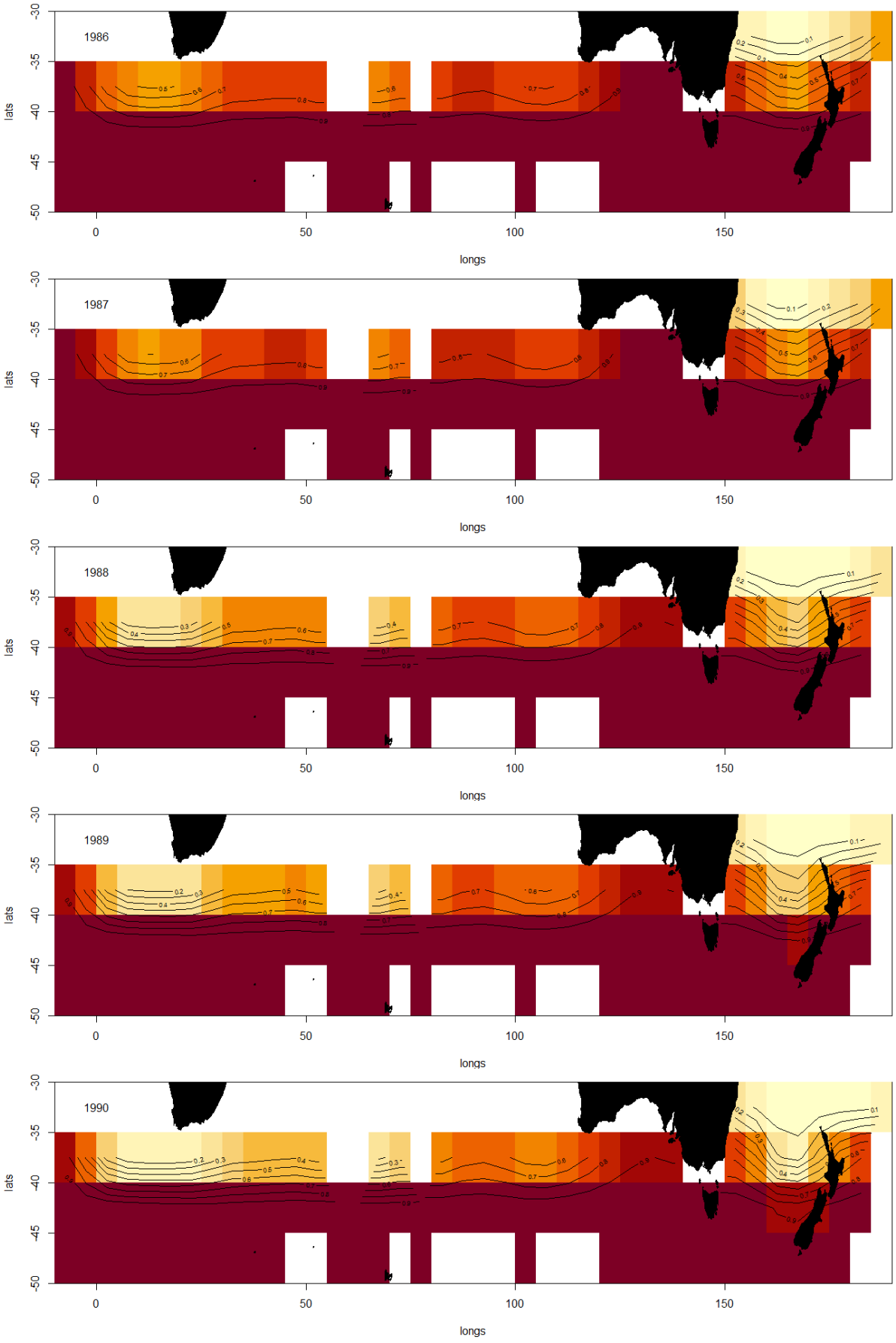

Figure 2: Estimates for binomial component of model.

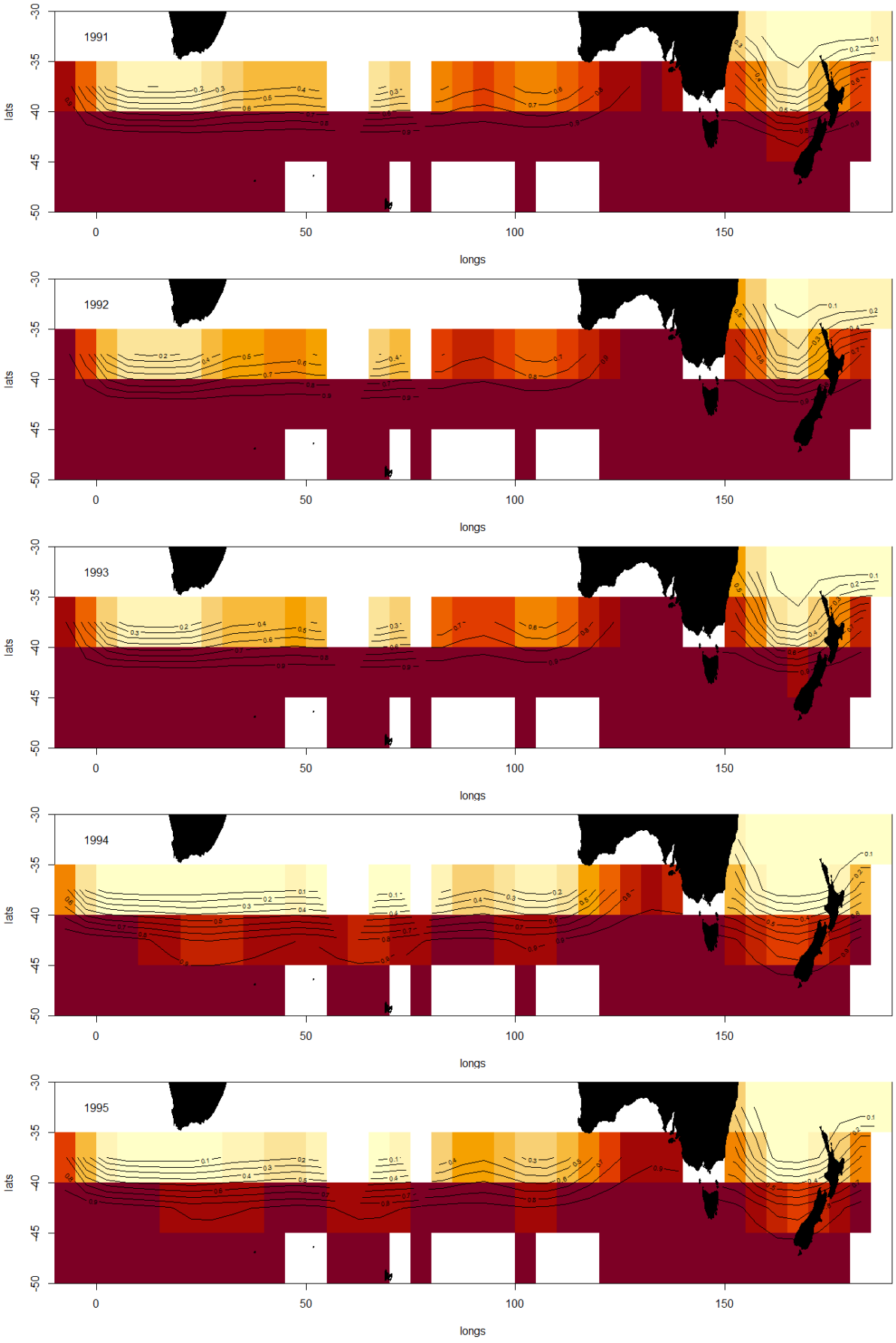

Figure 2 (cont.) Estimates for binomial component of model.

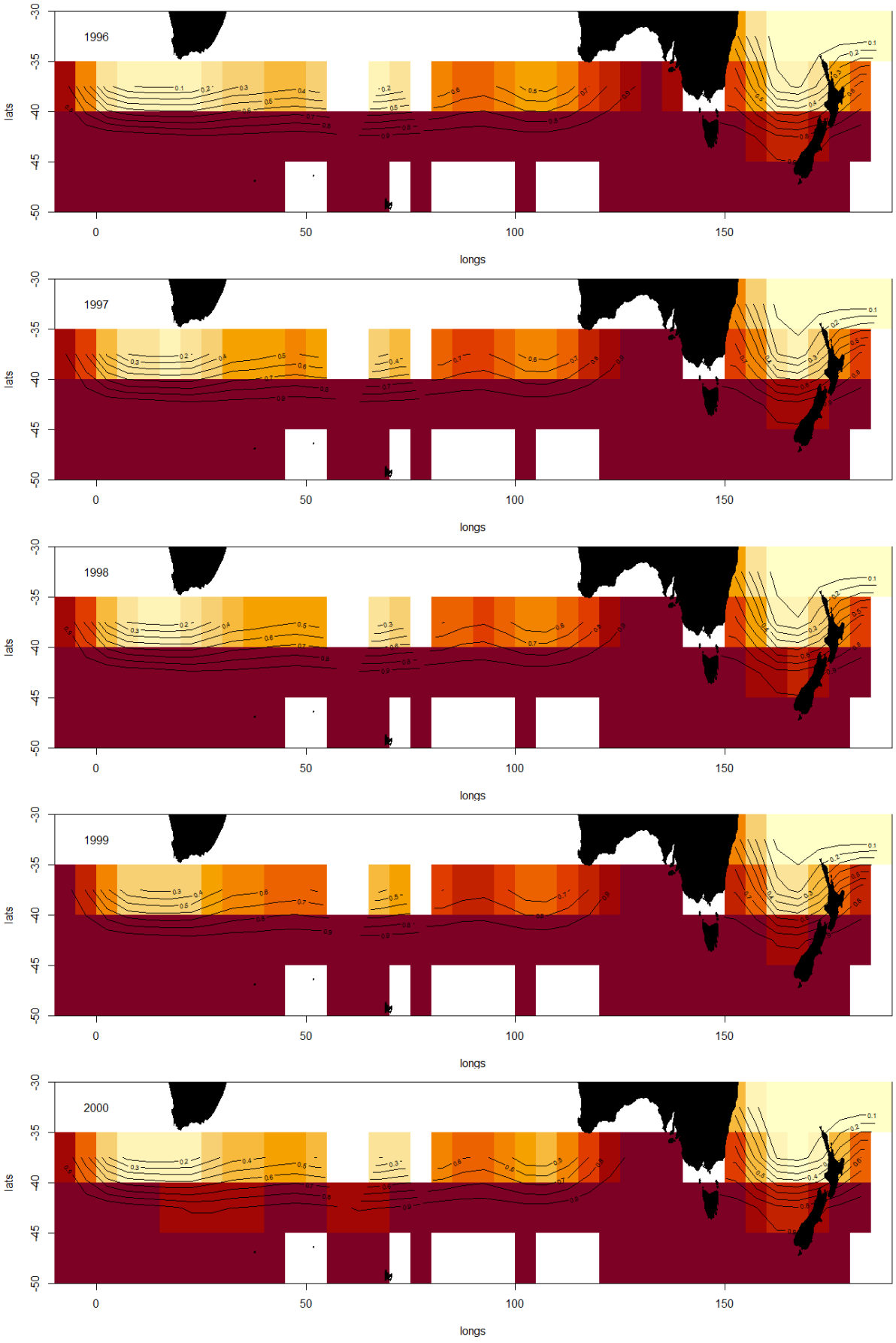

Figure 2 (cont.) Estimates for binomial component of model.

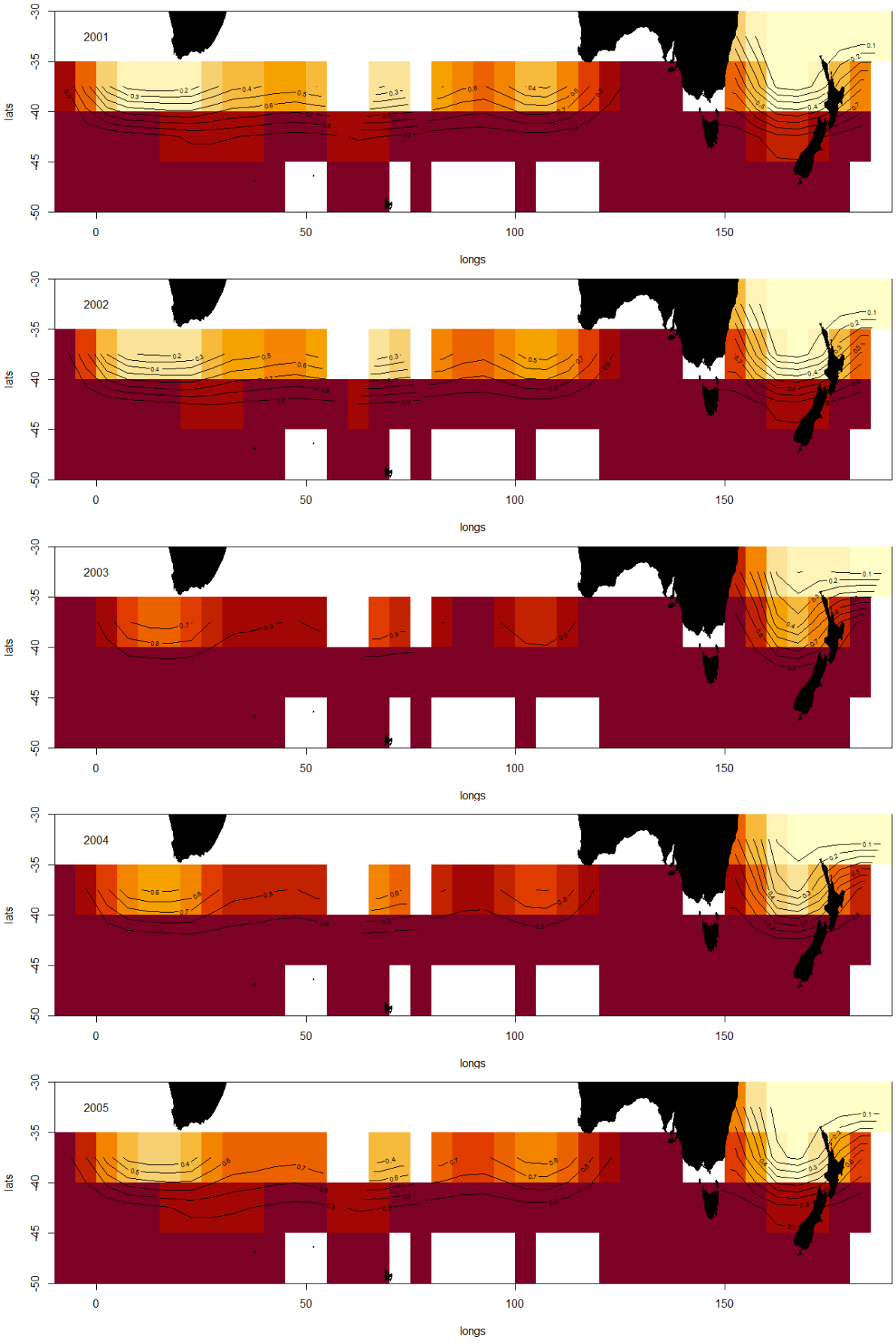

Figure 2 (cont.) Estimates for binomial component of model.

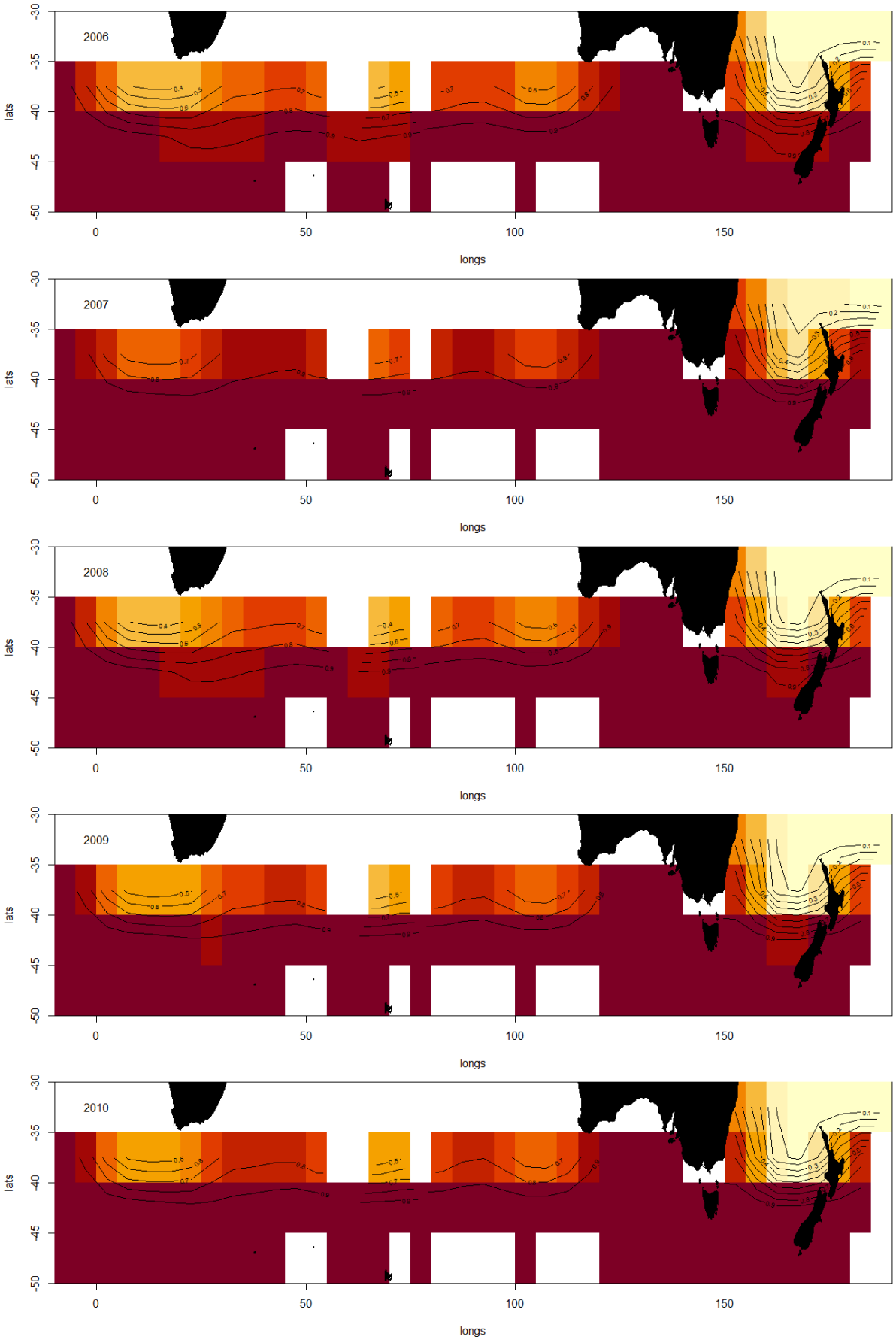

Figure 2 (cont.) Estimates for binomial component of model.

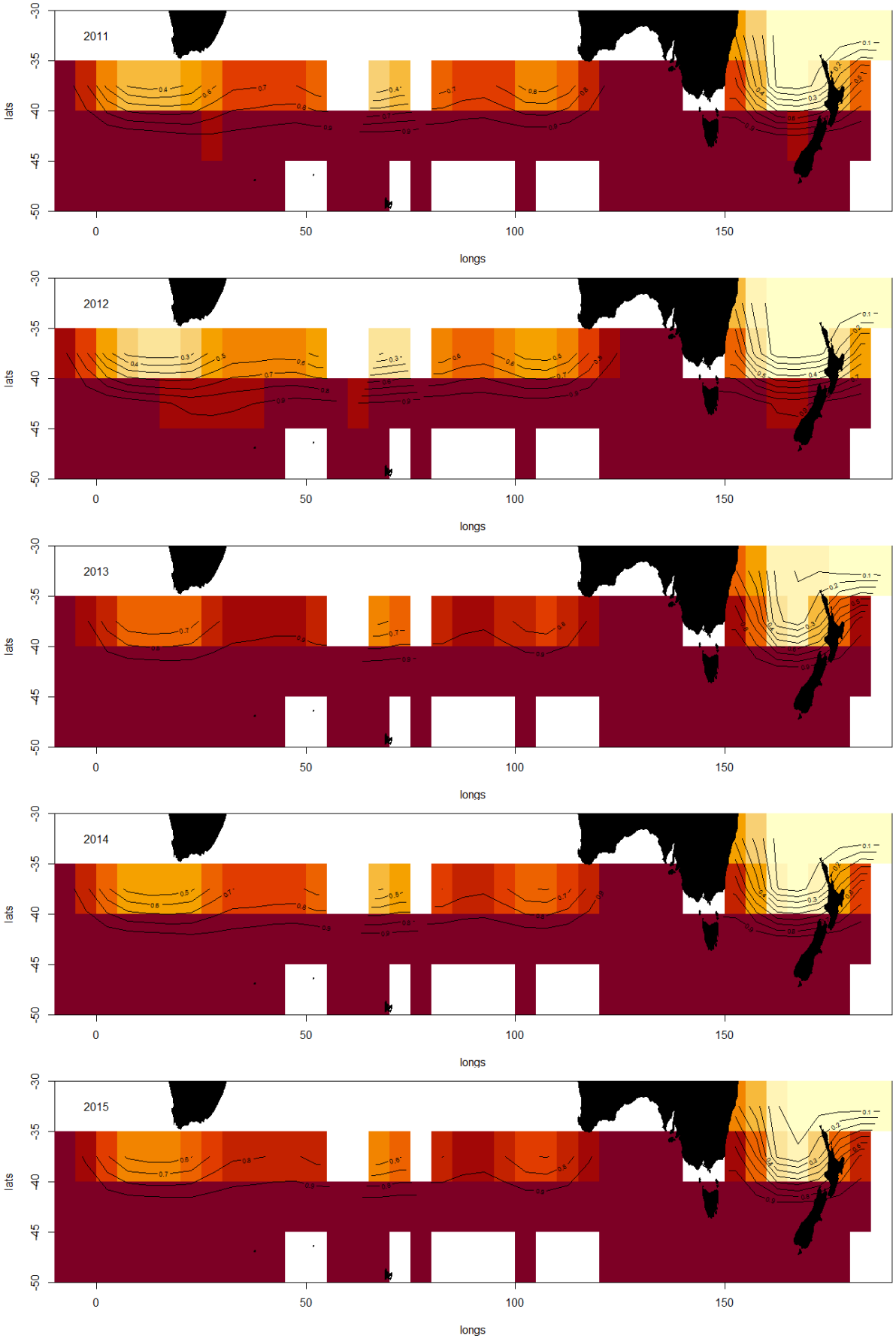

Figure 2 (cont.) Estimates for binomial component of model.

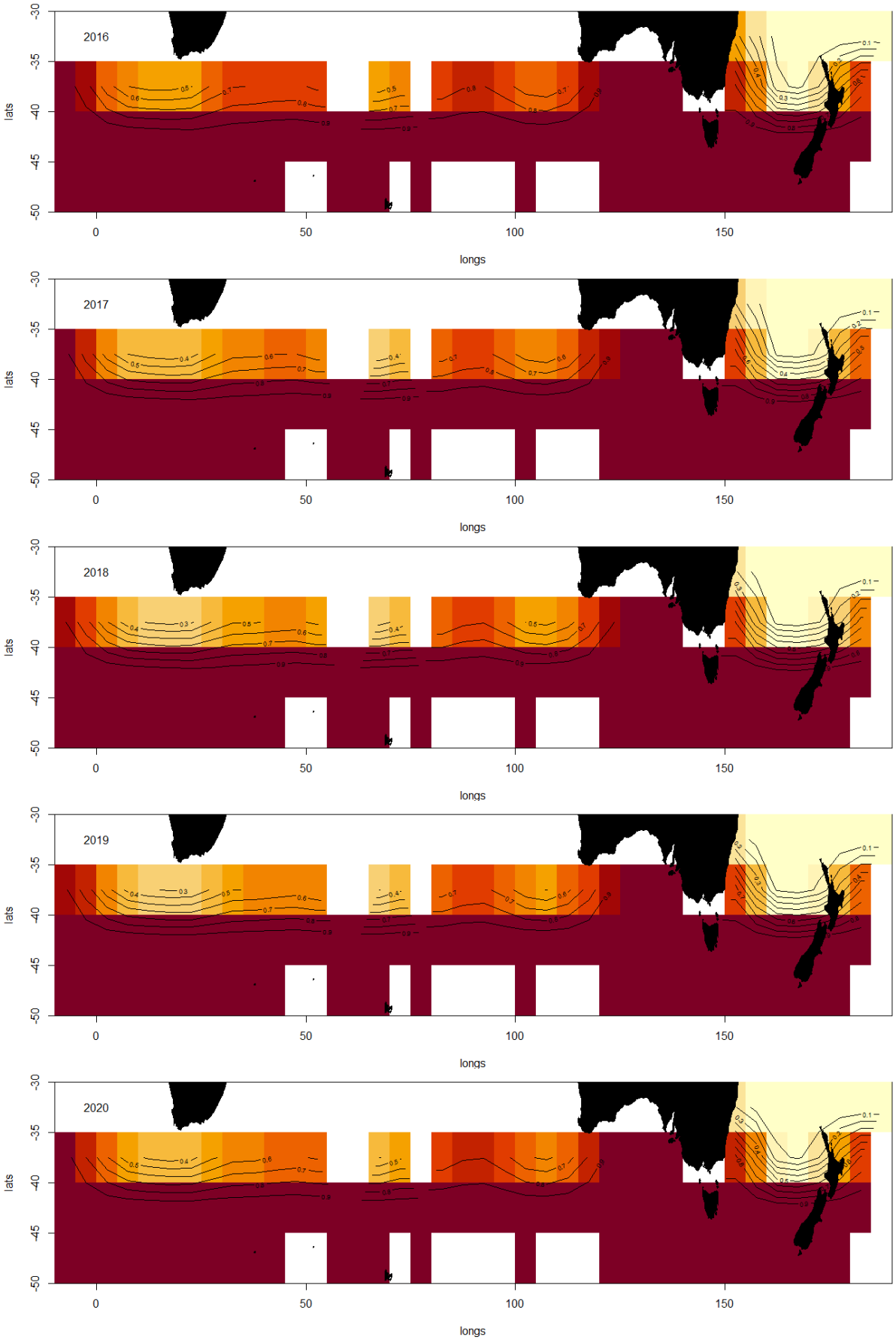

Figure 2 (cont.) Estimates for binomial component of model.

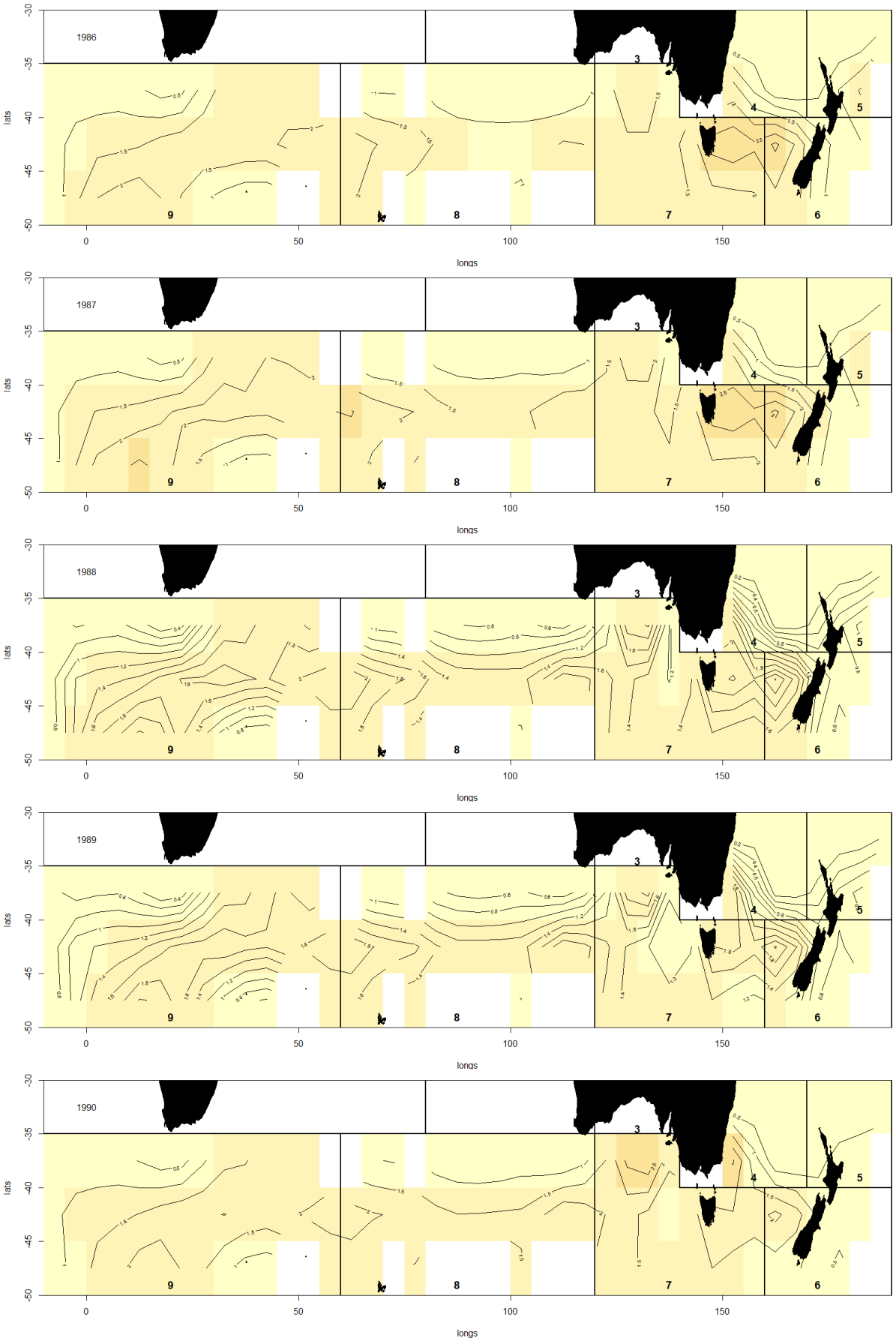

Figure 3: Estimates for the positive component of the model.

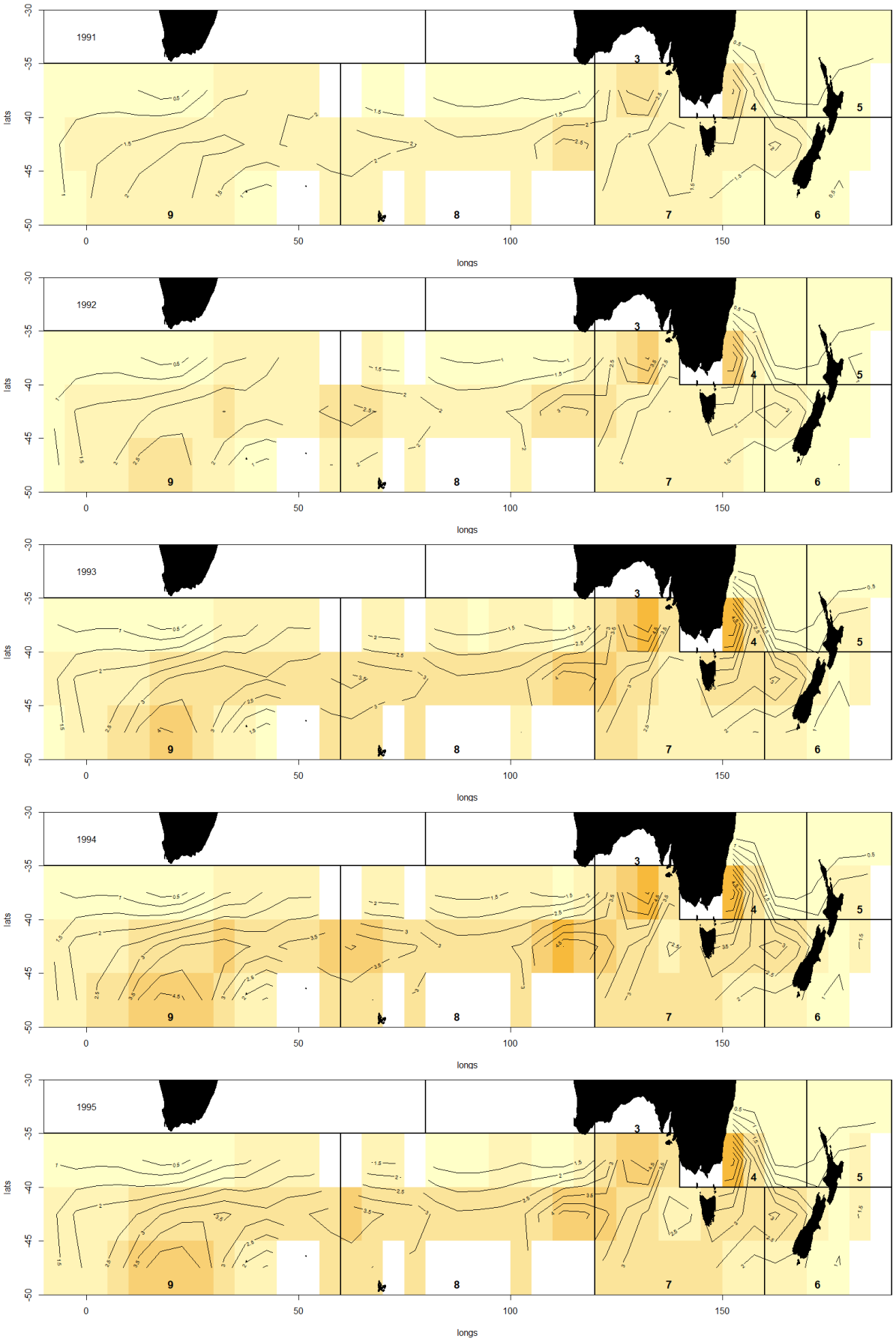

Figure 3 (cont.): Estimates for the positive component of the model.

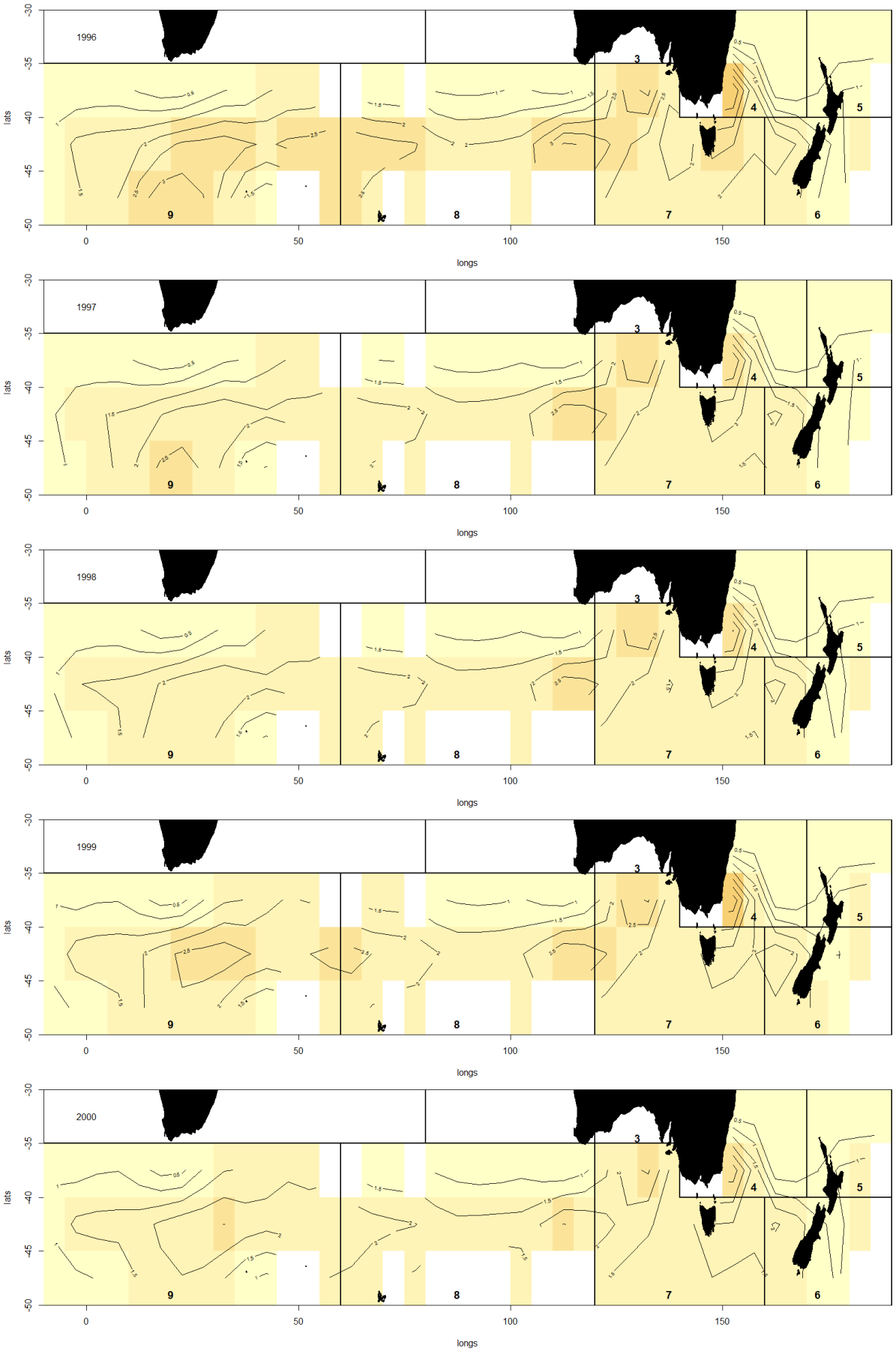

Figure 3 (cont.): Estimates for the positive component of the model.

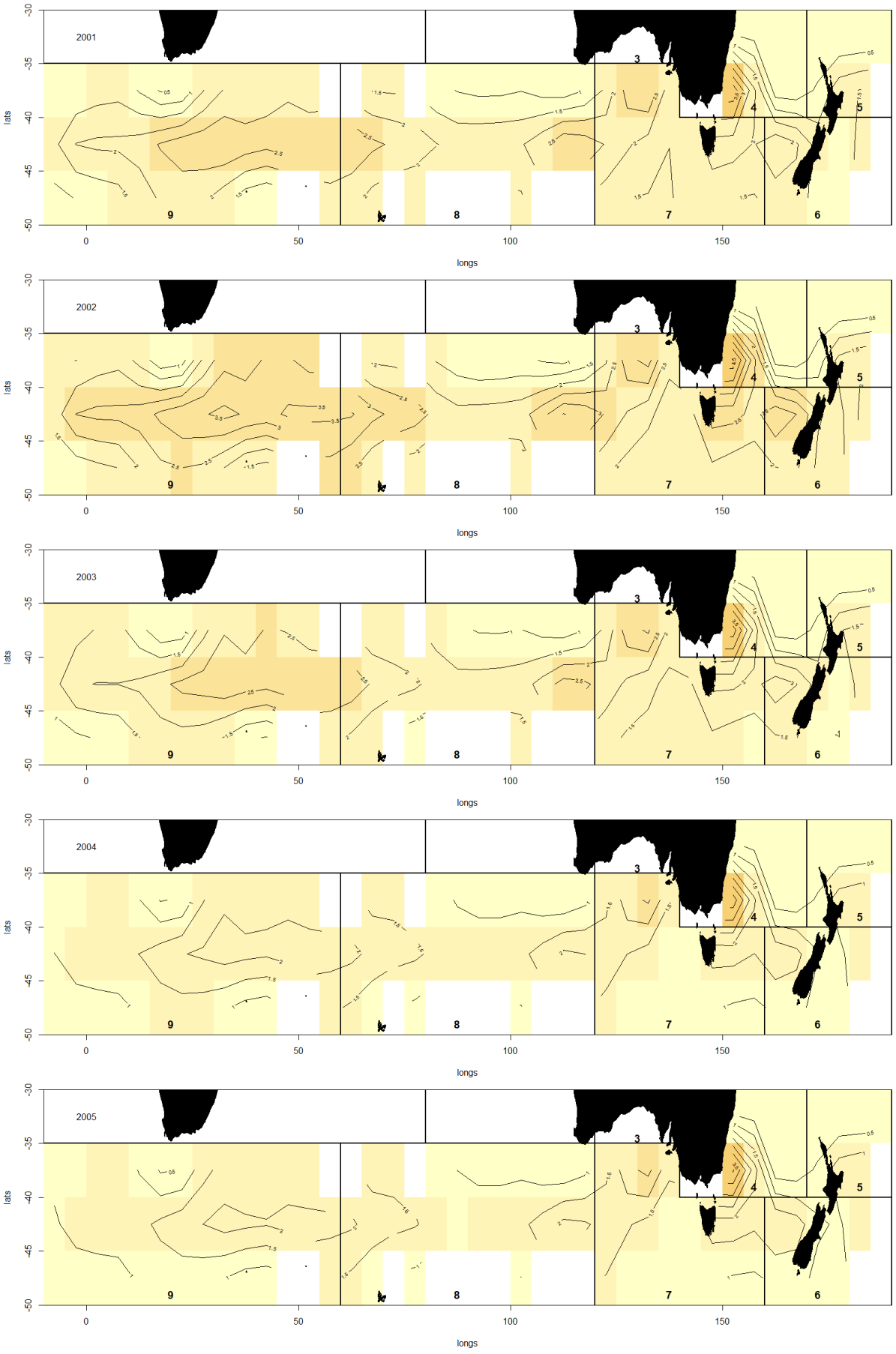

Figure 3 (cont.): Estimates for the positive component of the model.

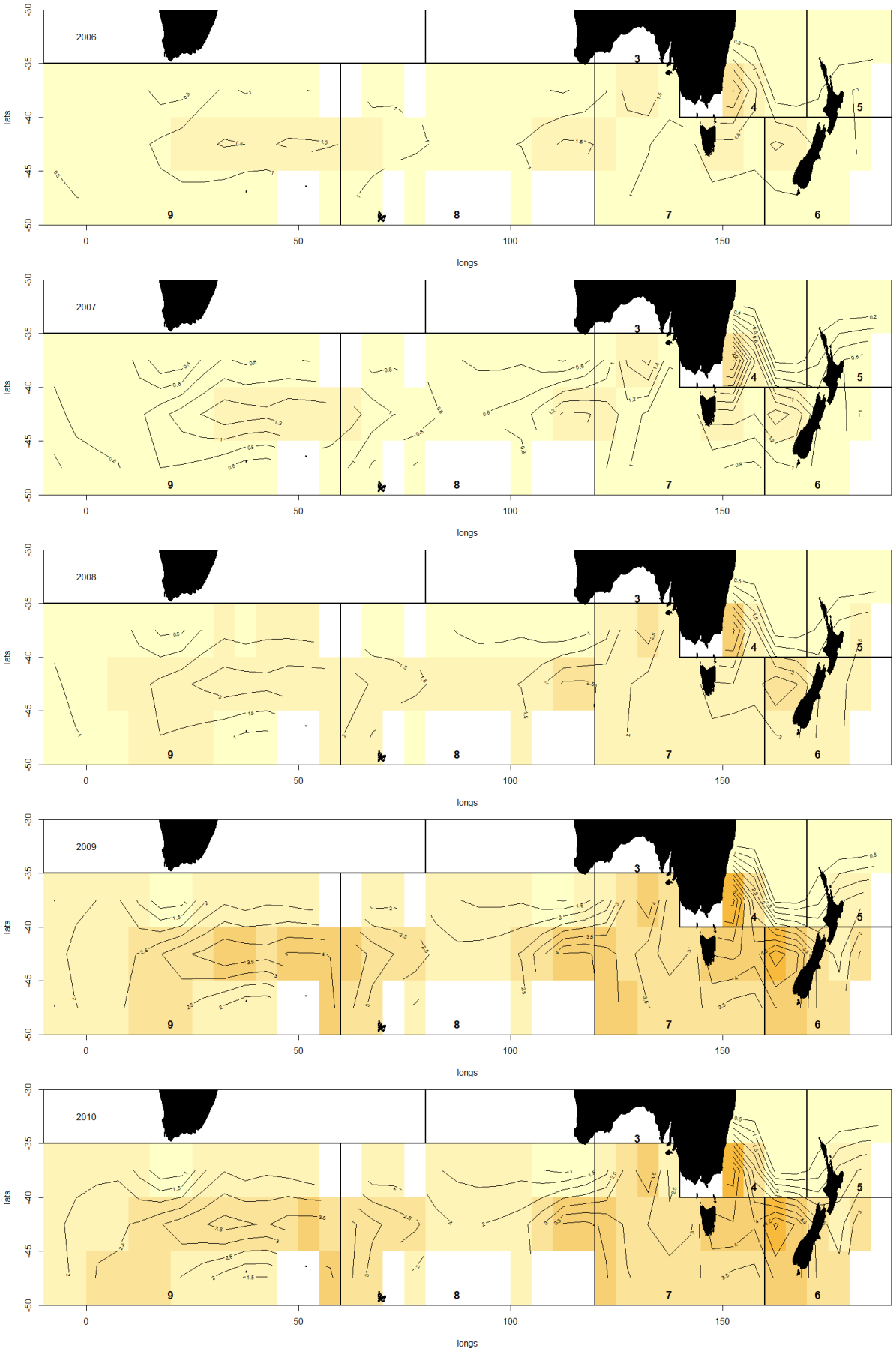

Figure 3 (cont.): Estimates for the positive component of the model.

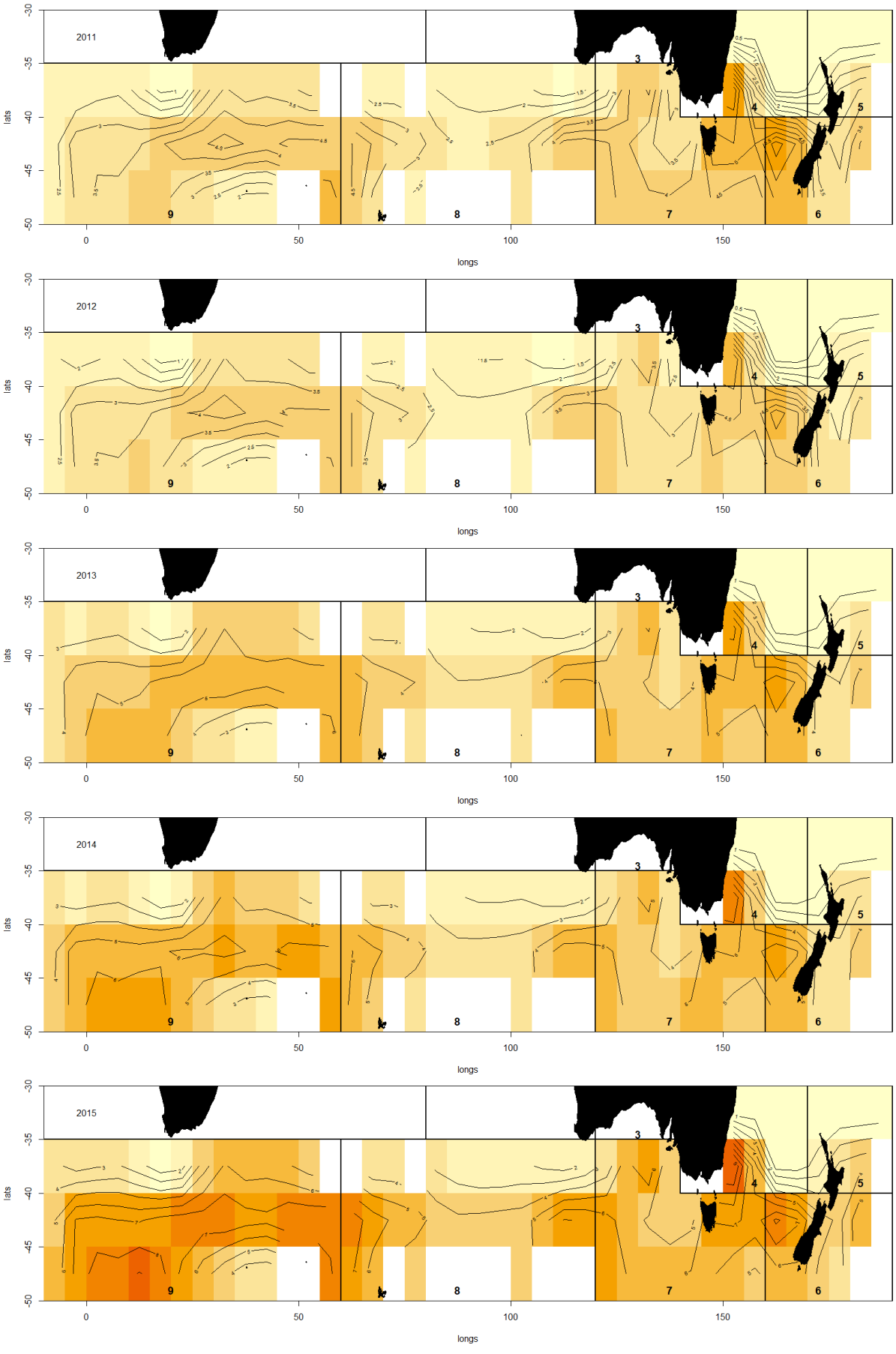

Figure 3 (cont.): Estimates for the positive component of the model.

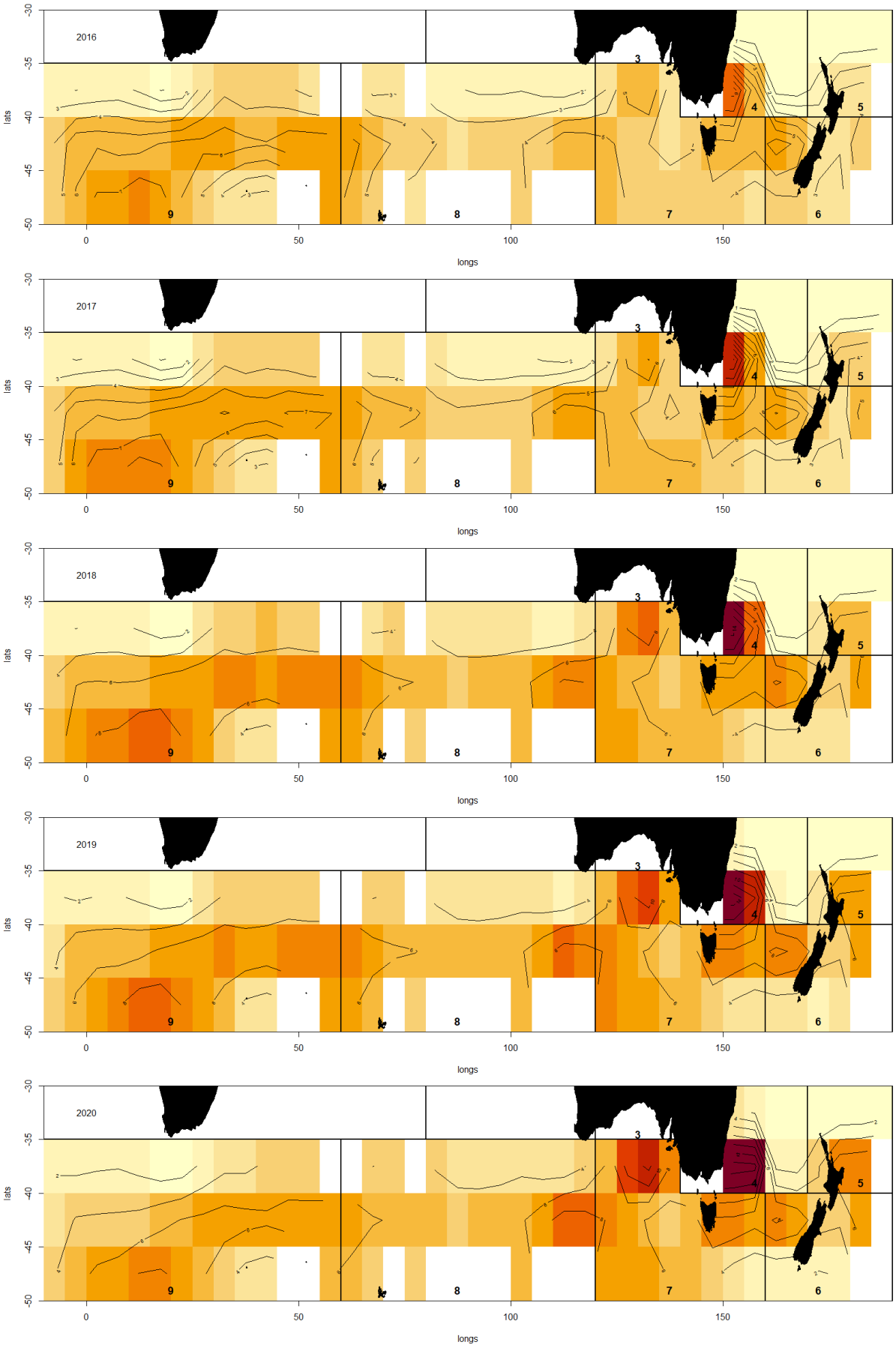

Figure 3 (cont.): Estimates for the positive component of the model.

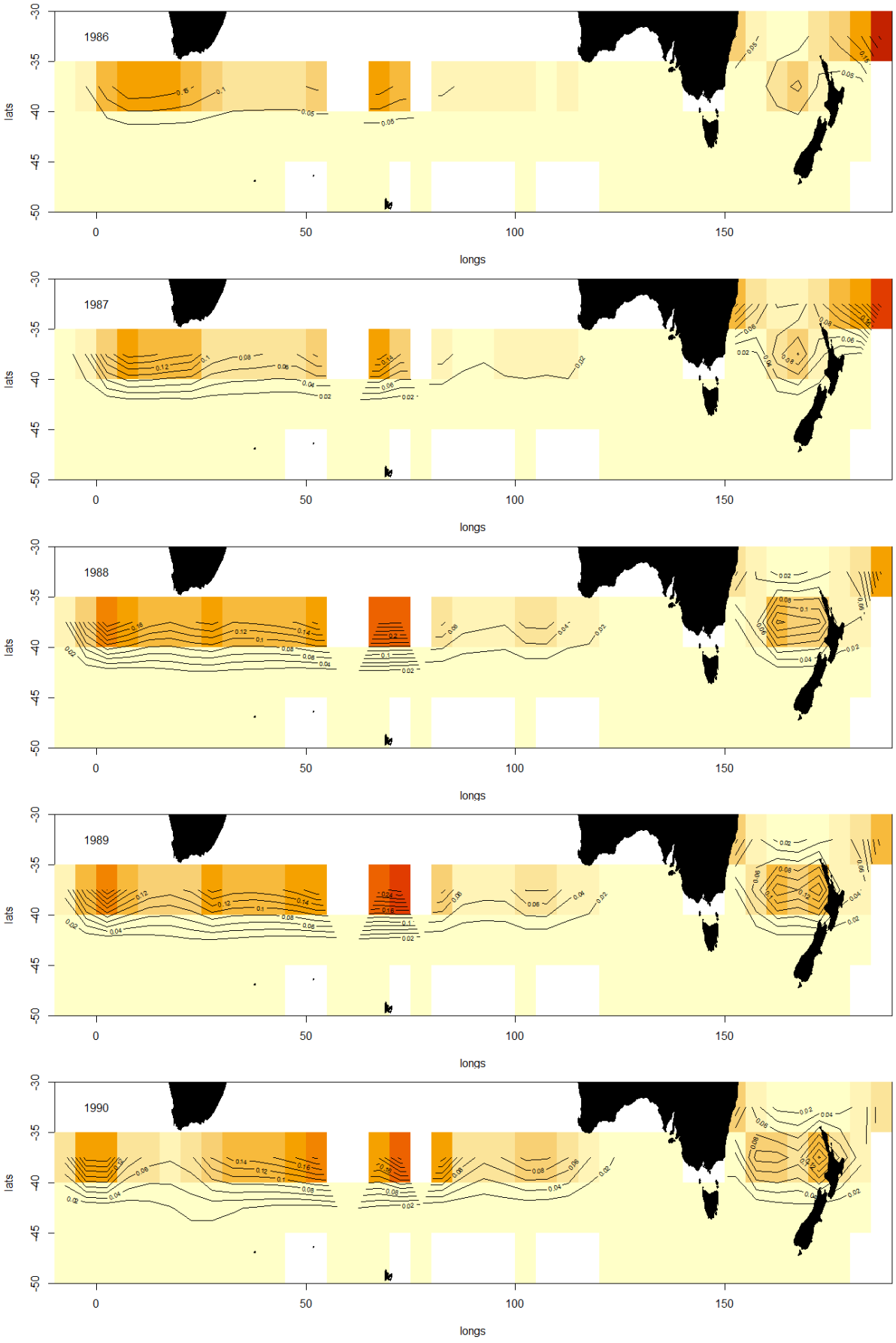

Figure 4: Uncertainty estimates for binomial component of model.

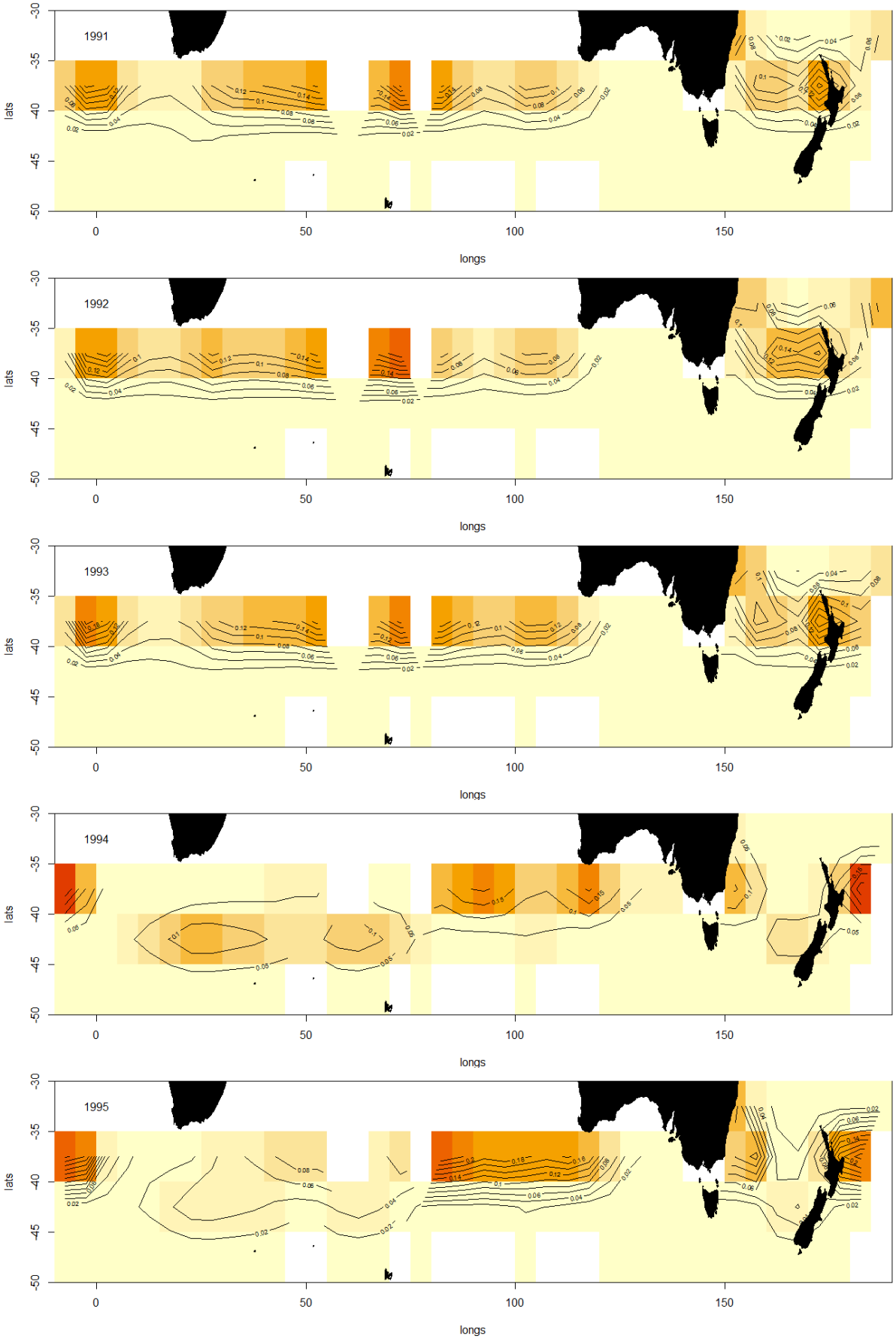

Figure 4 (cont.): Uncertainty estimates for binomial component of model.

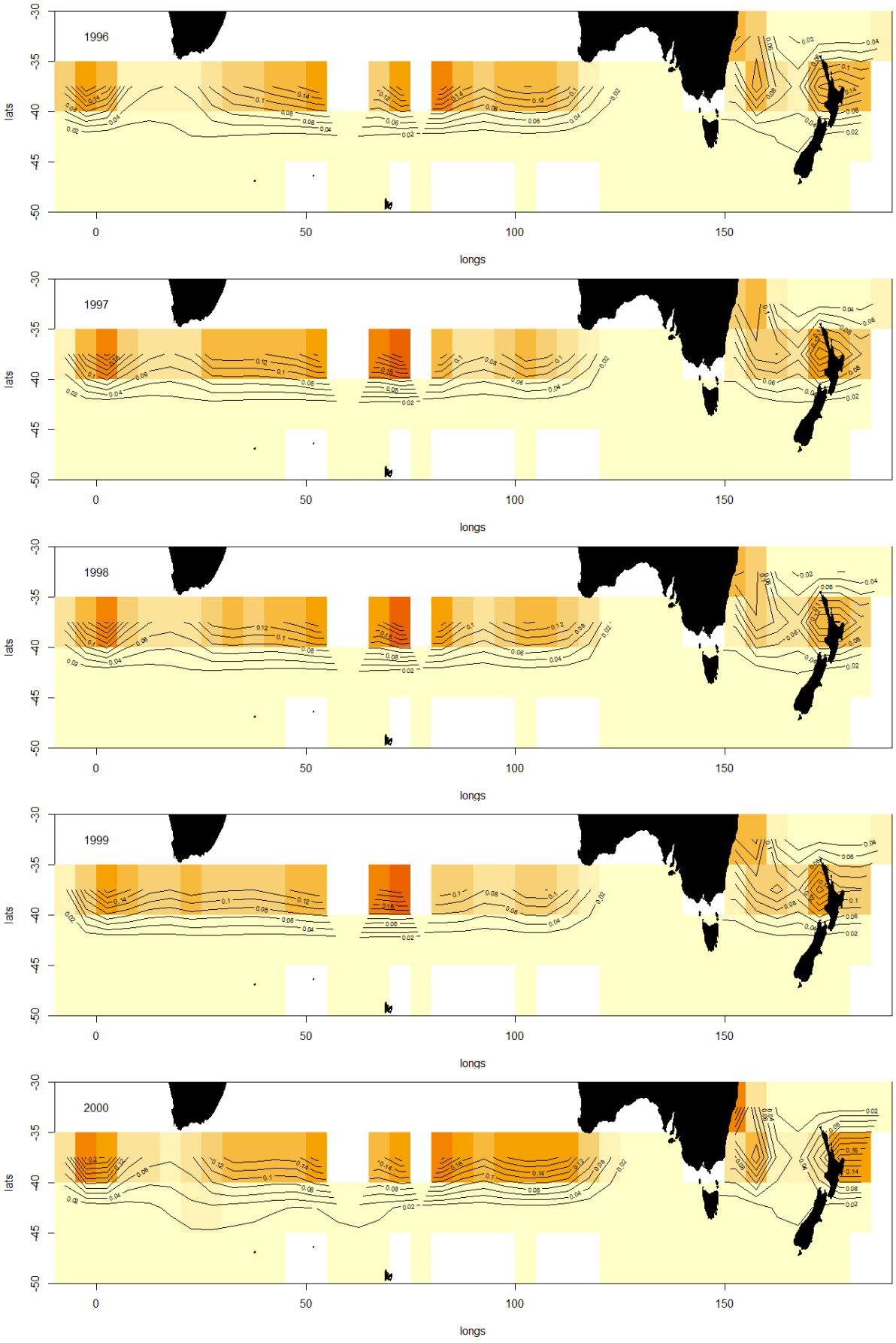

Figure 4 (cont.): Uncertainty estimates for binomial component of model.

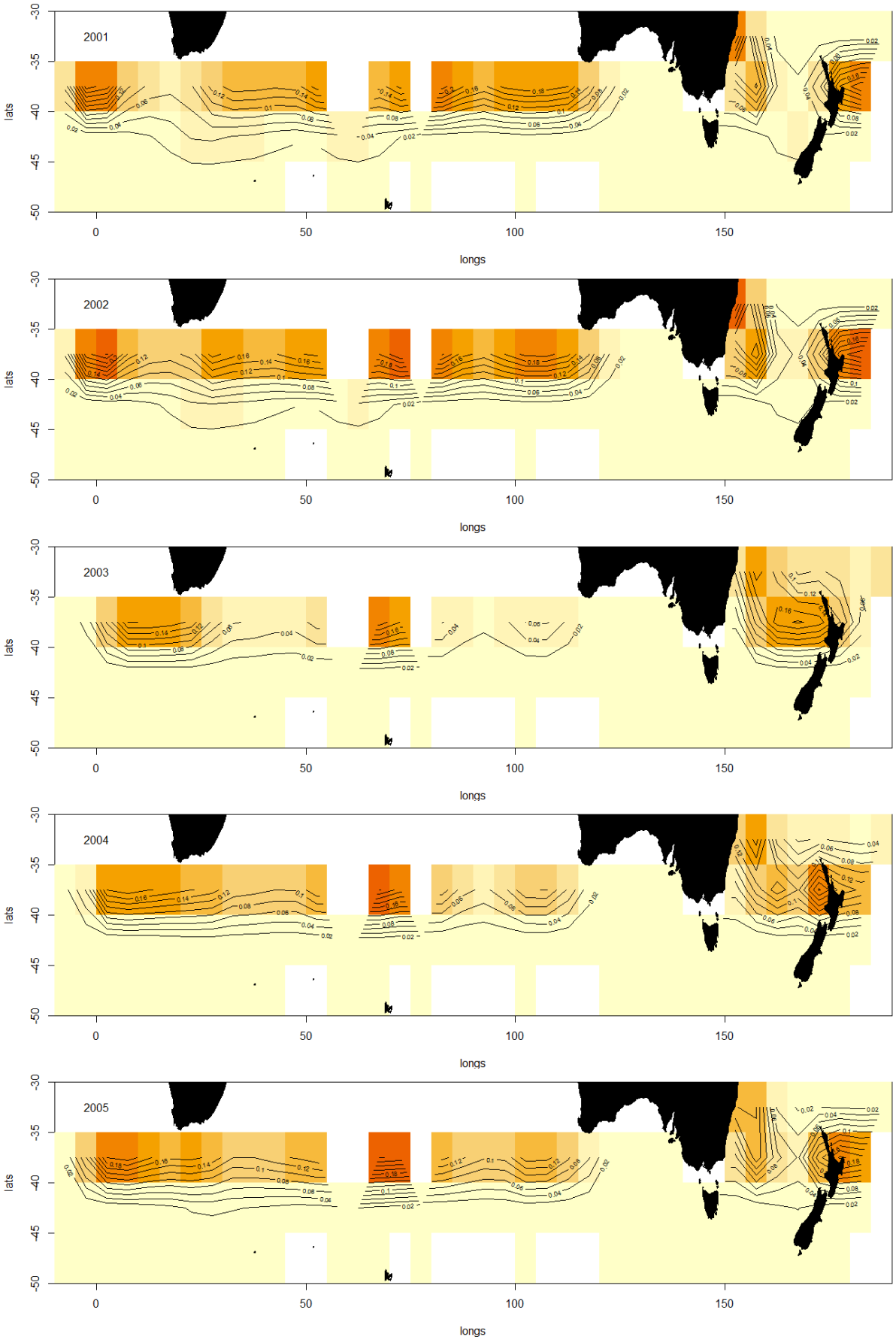

Figure 4 (cont.): Uncertainty estimates for binomial component of model.

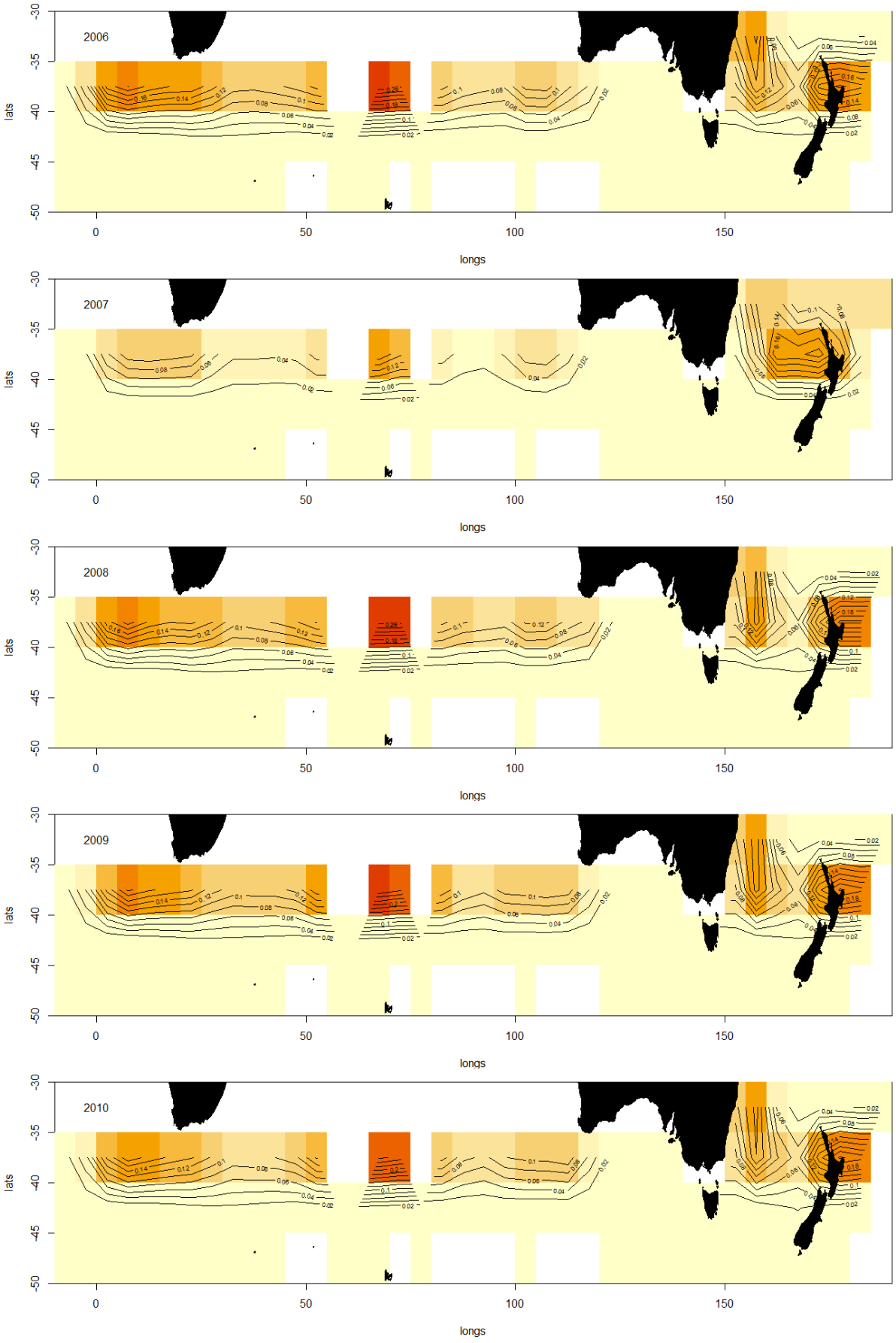

Figure 4: Uncertainty estimates for binomial component of model.

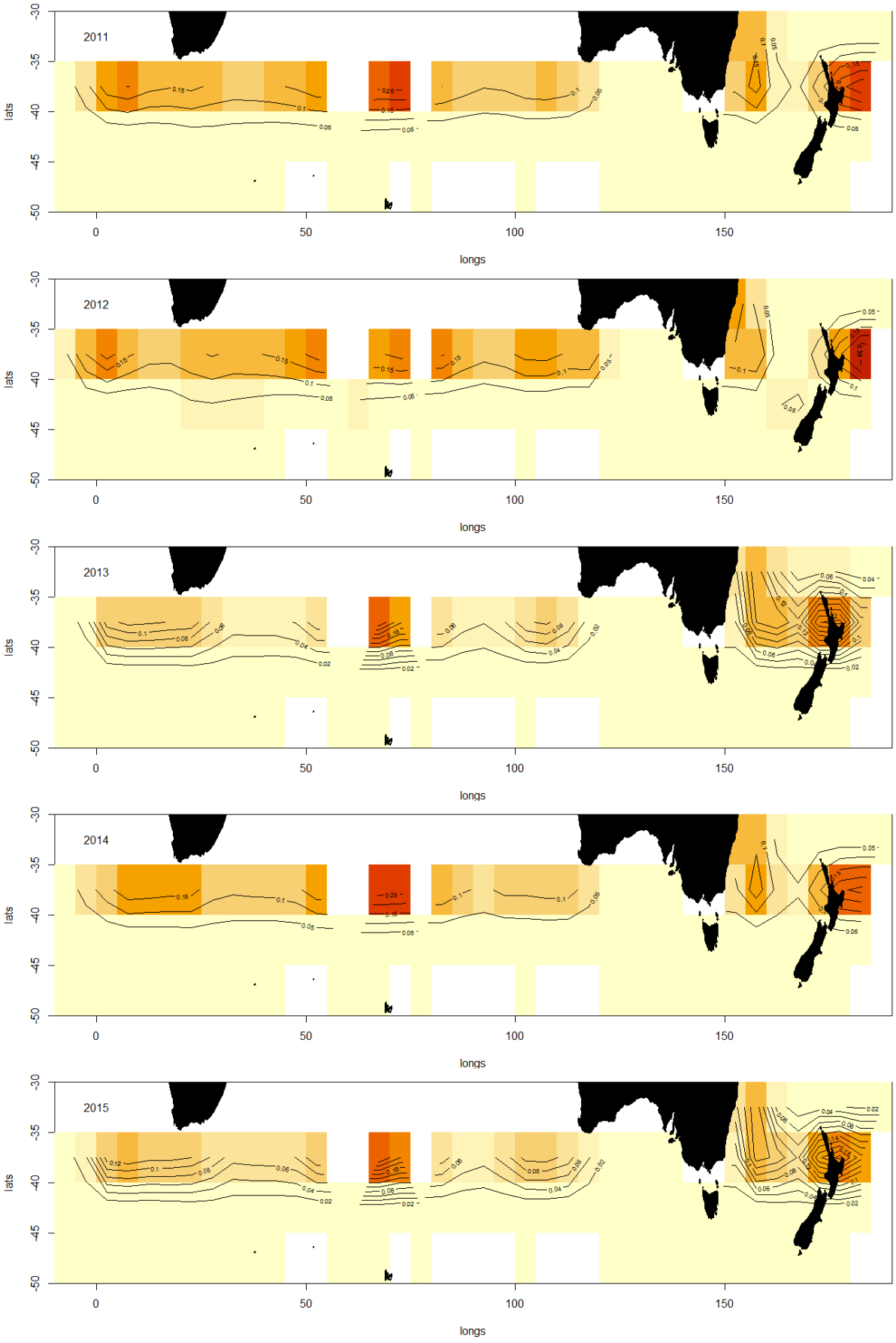

Figure 4 (cont.): Uncertainty estimates for binomial component of model.

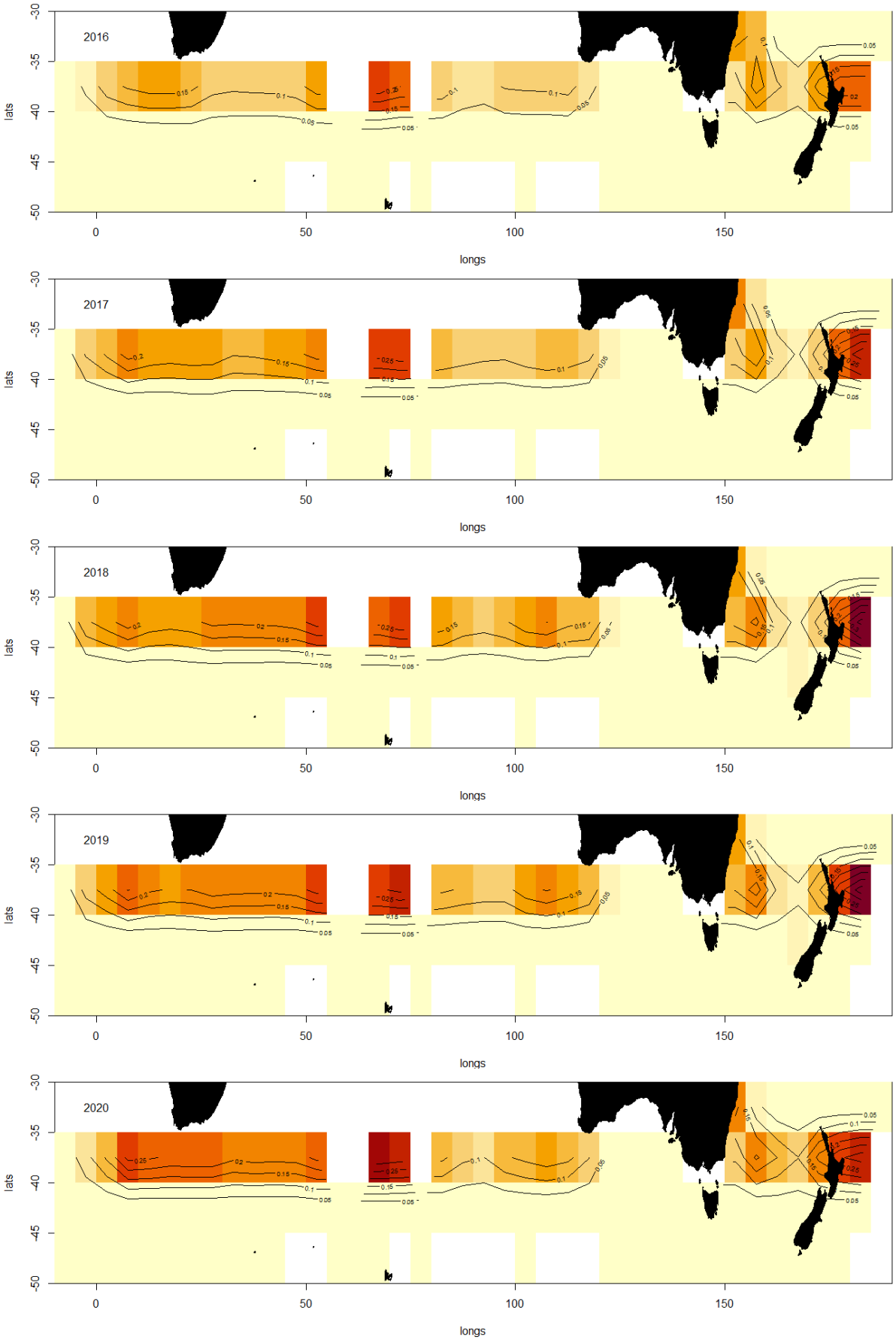

Figure 4 (cont.): Uncertainty estimates for binomial component of model.

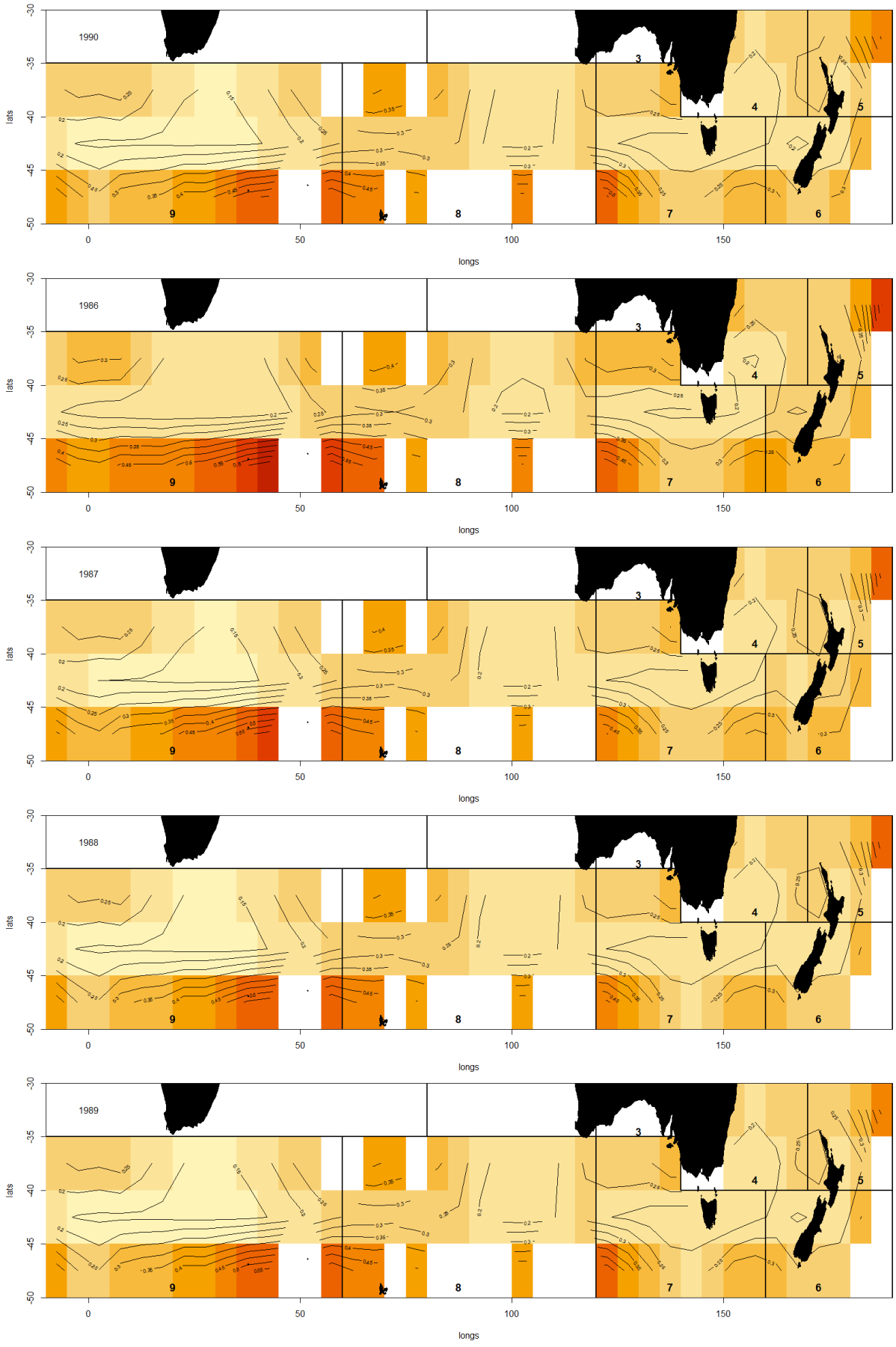

Figure 5: Variance estimates for lognormal component of model.

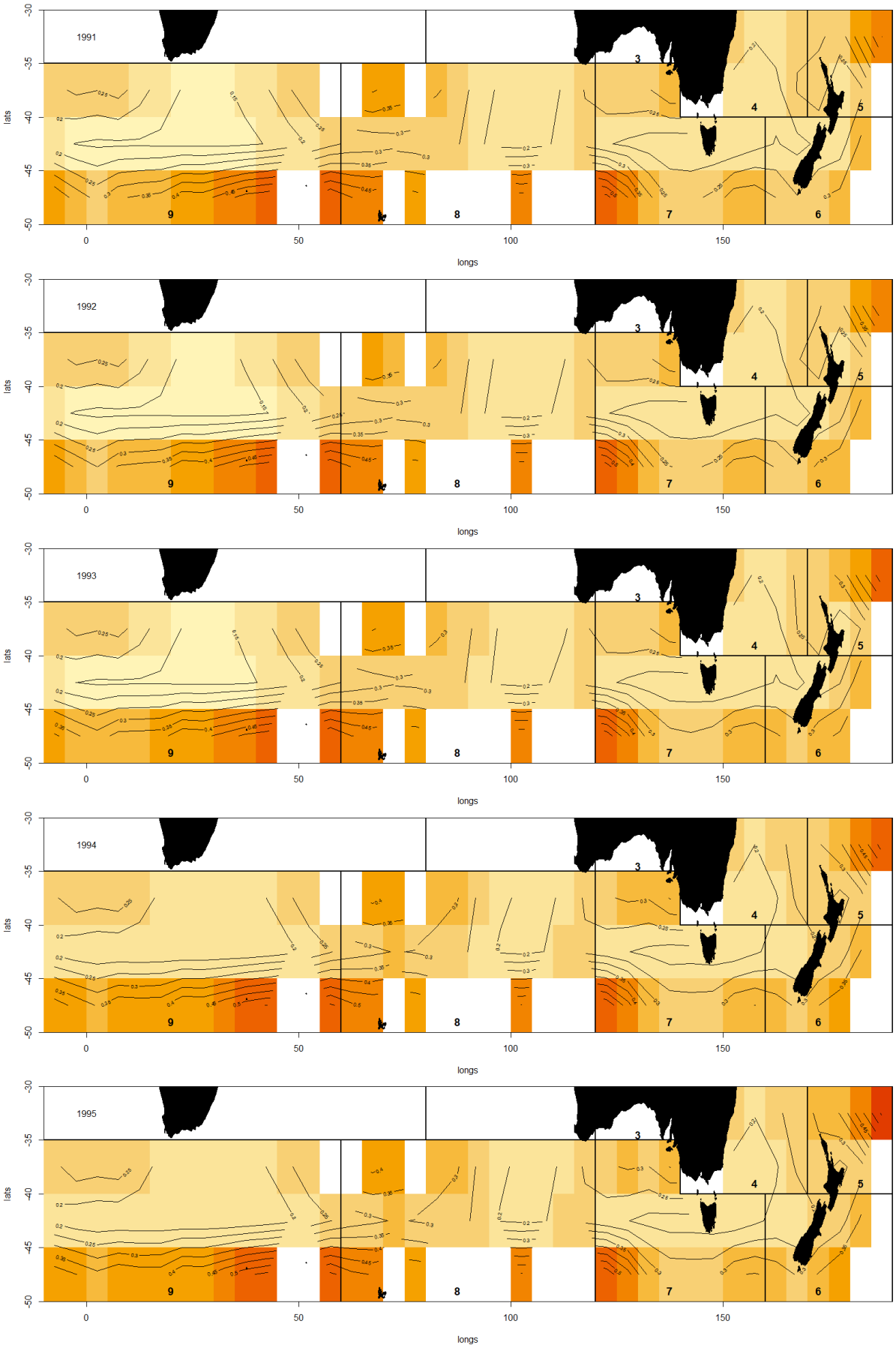

Figure 5 (cont.): Uncertainty estimates for lognormal component of model.

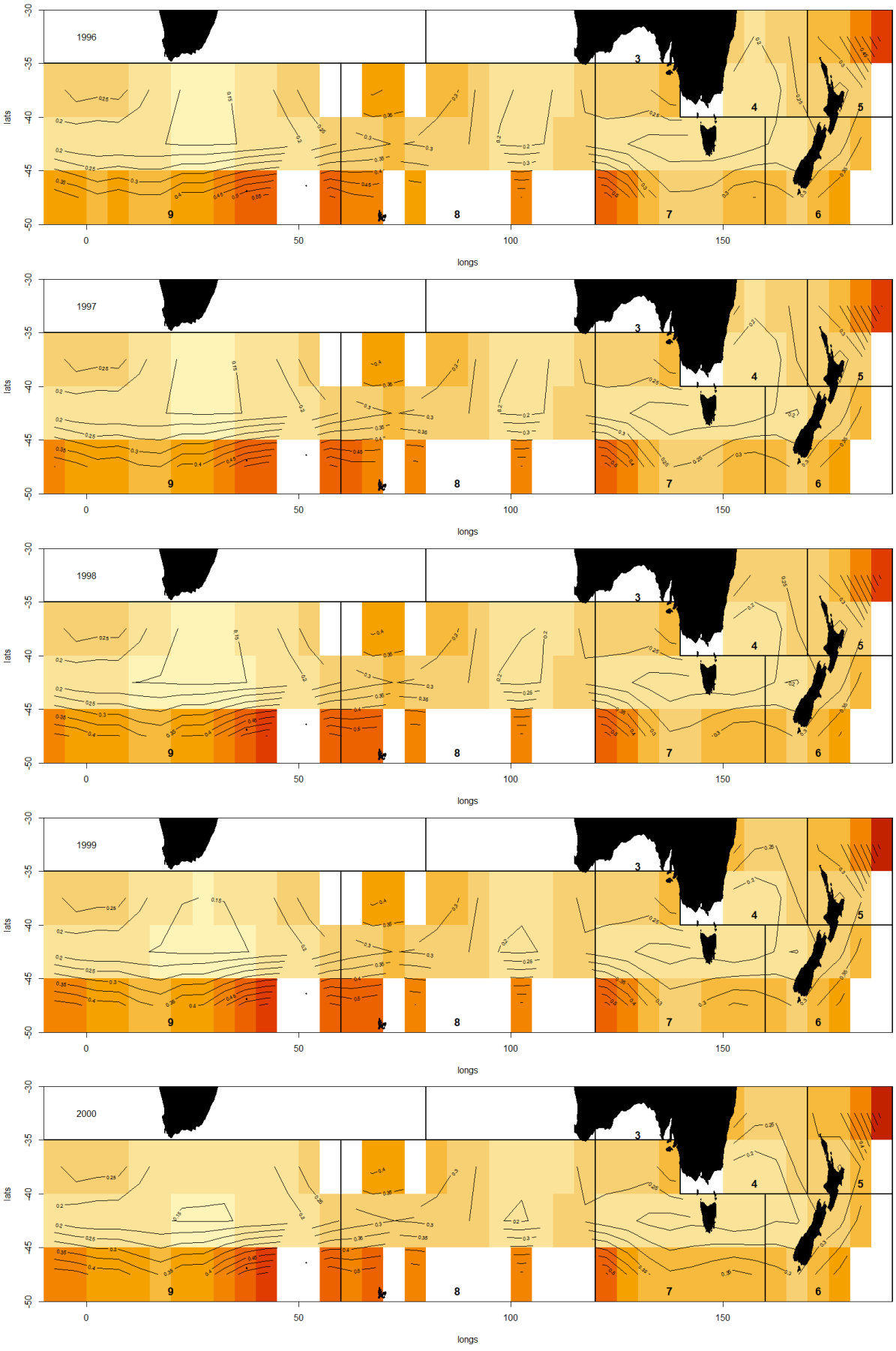

Figure 5 (cont.): Uncertainty estimates for lognormal component of model.

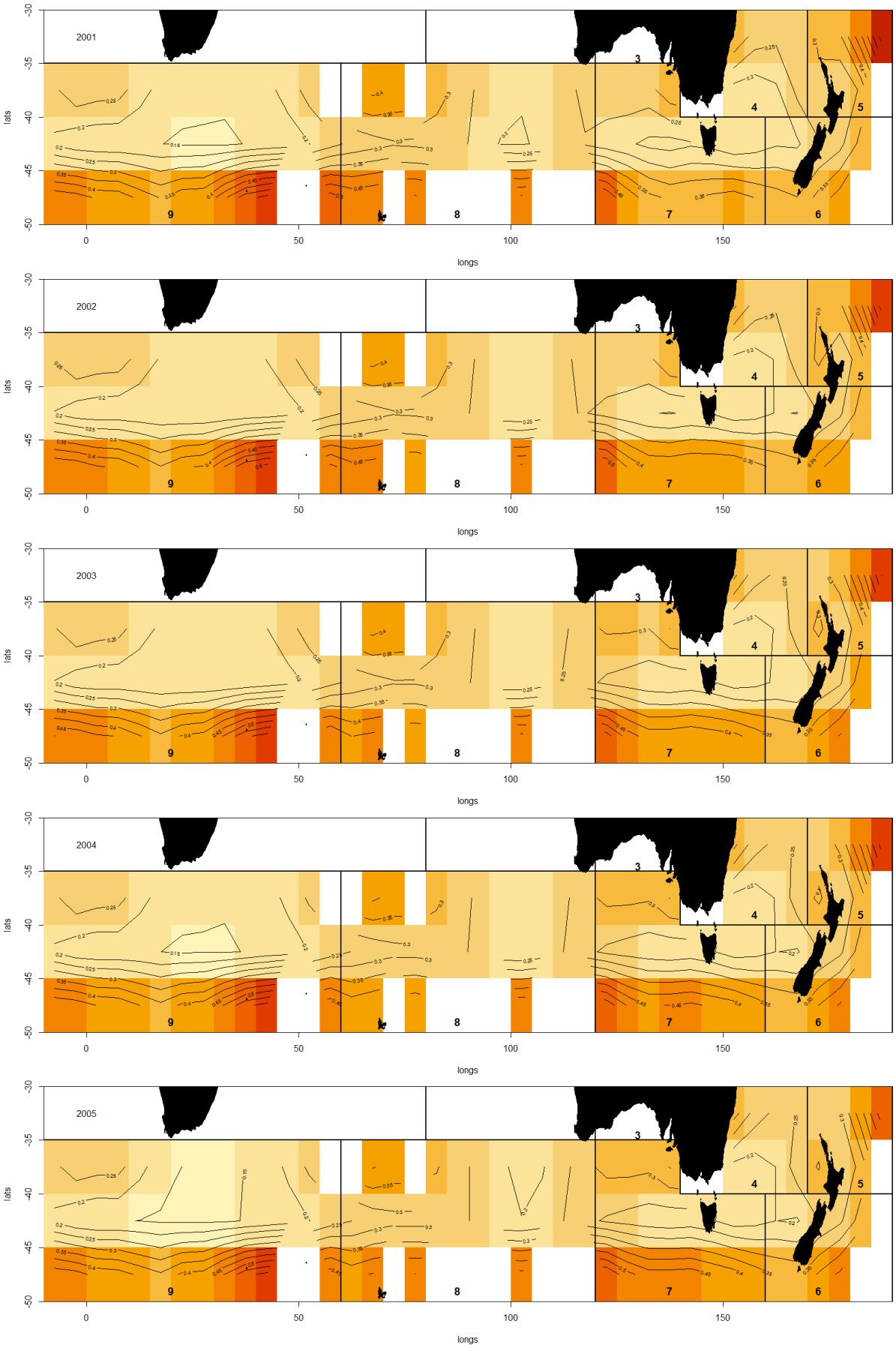

Figure 5 (cont.): Uncertainty estimates for lognormal component of model.

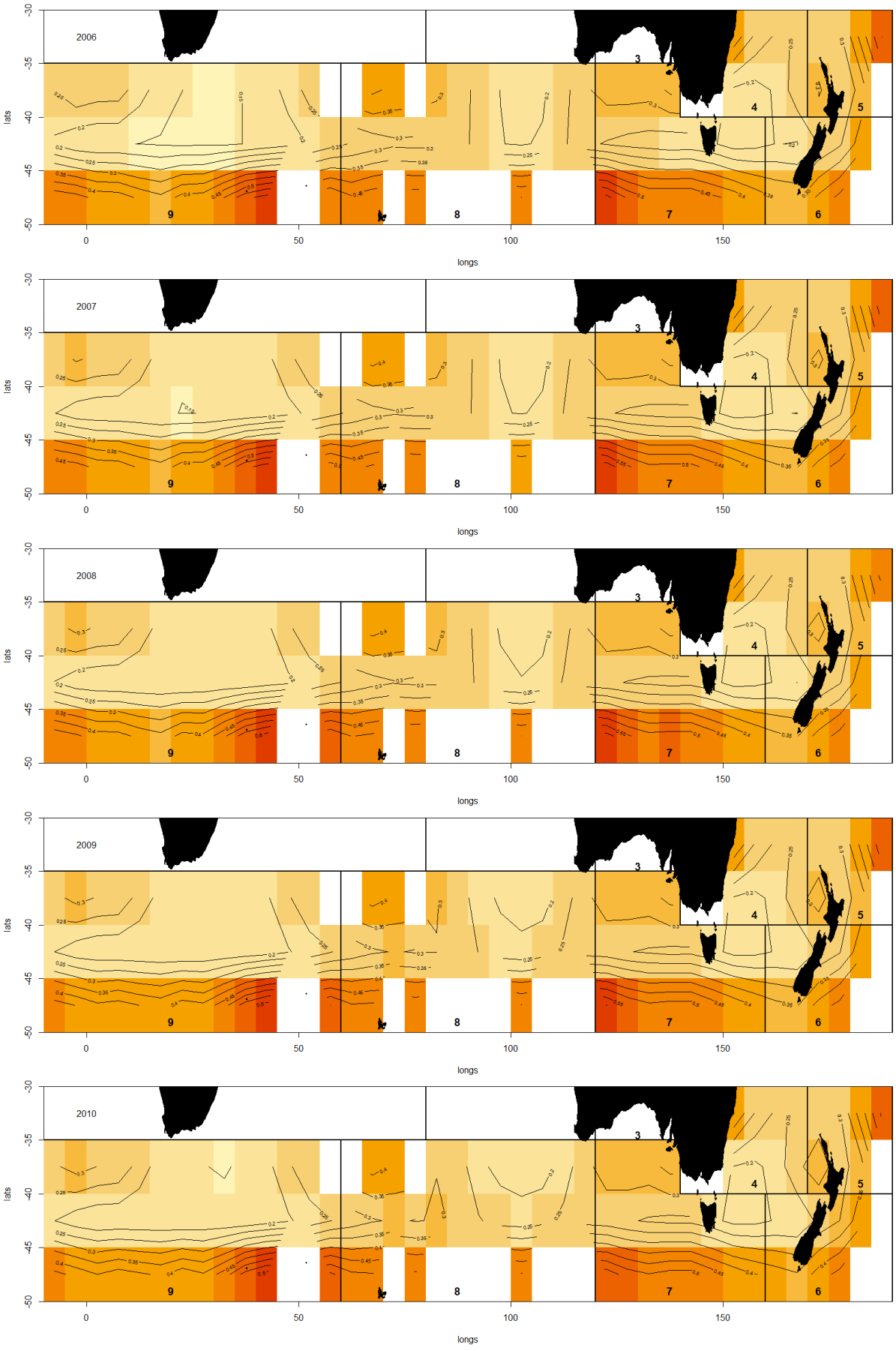

Figure 5 (cont.): Uncertainty estimates for lognormal component of model.

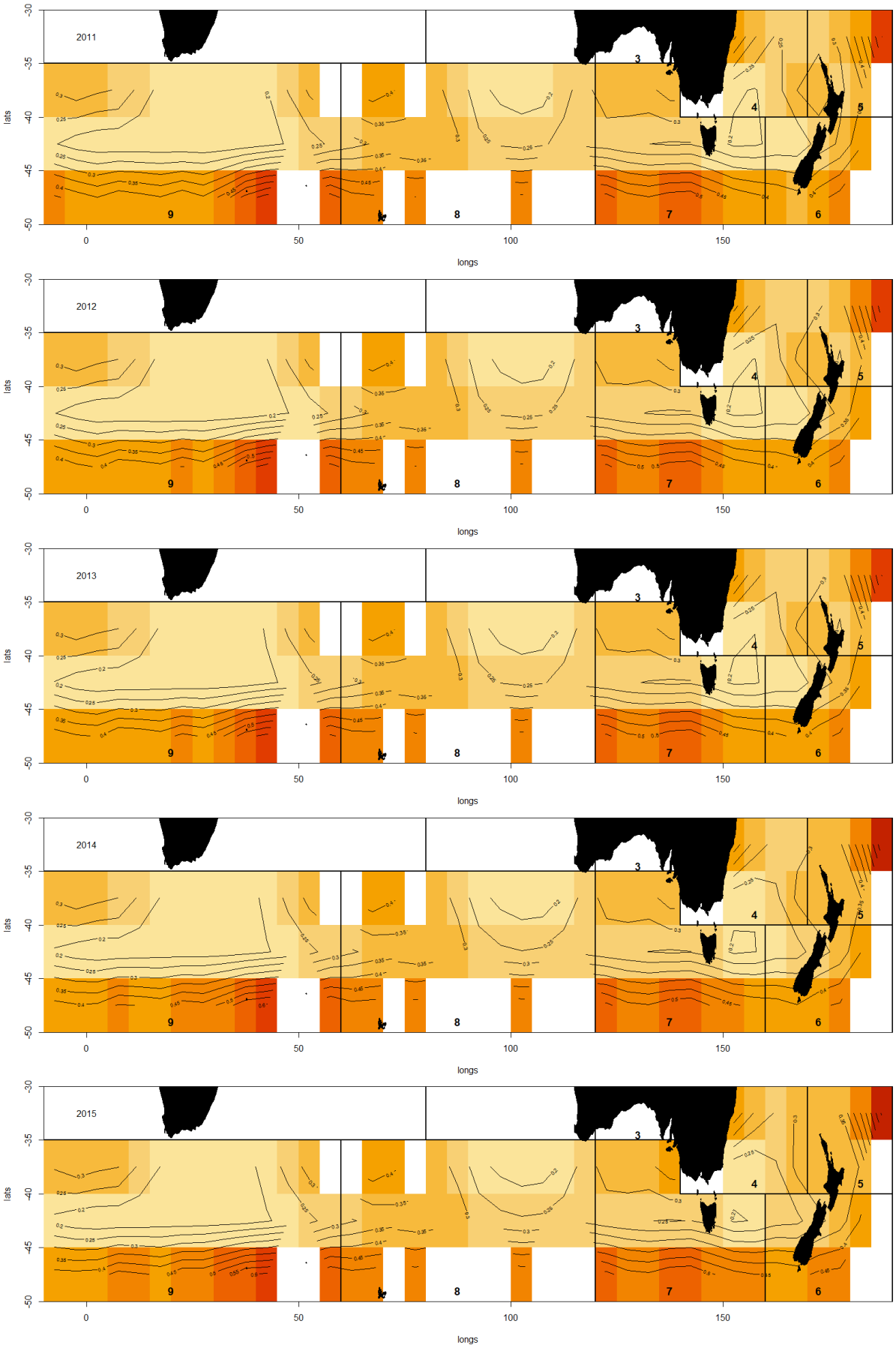

Figure 5 (cont.): Uncertainty estimates for lognormal component of model.

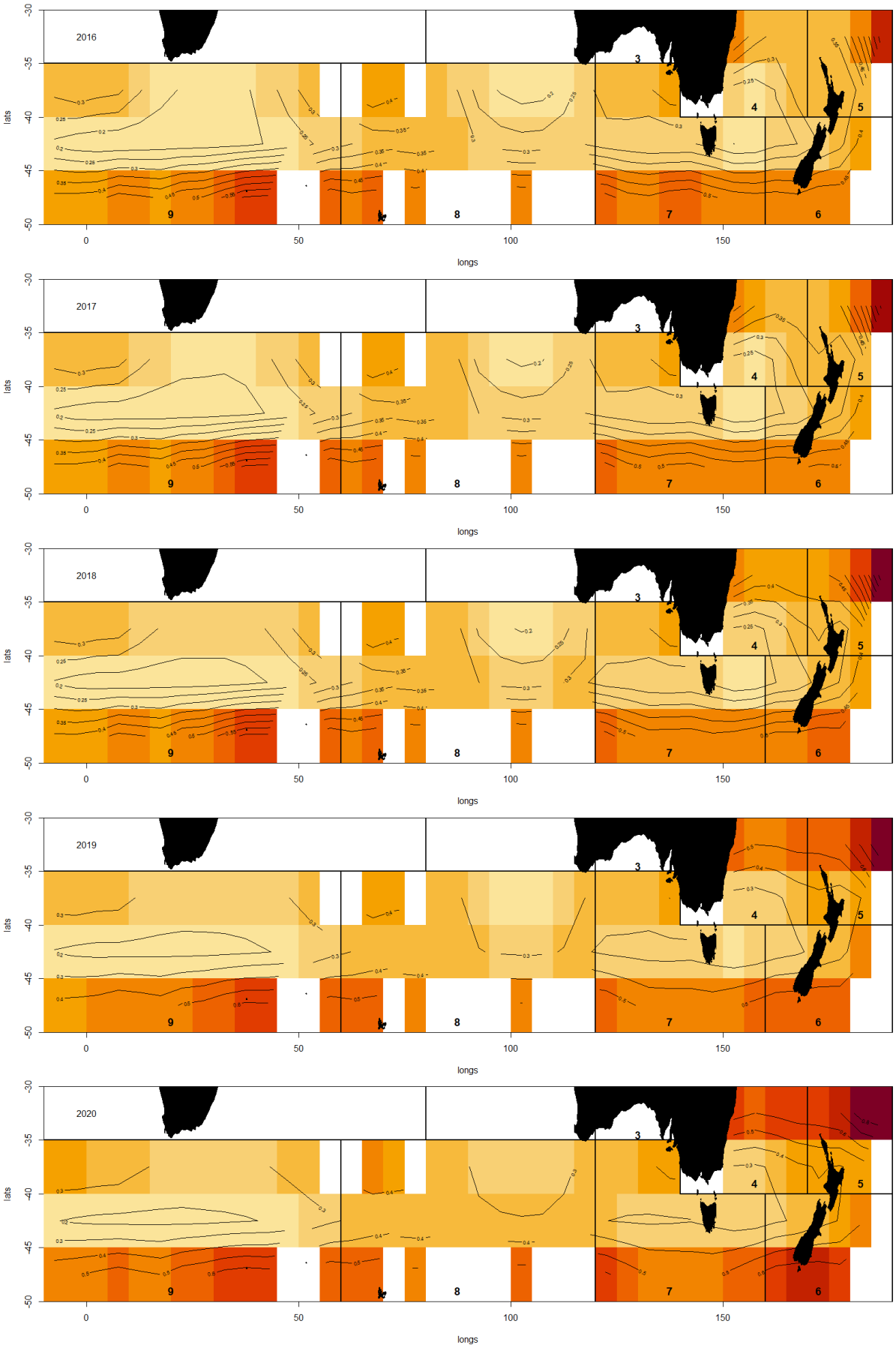

Figure 5 (cont.): Uncertainty estimates for lognormal component of model.

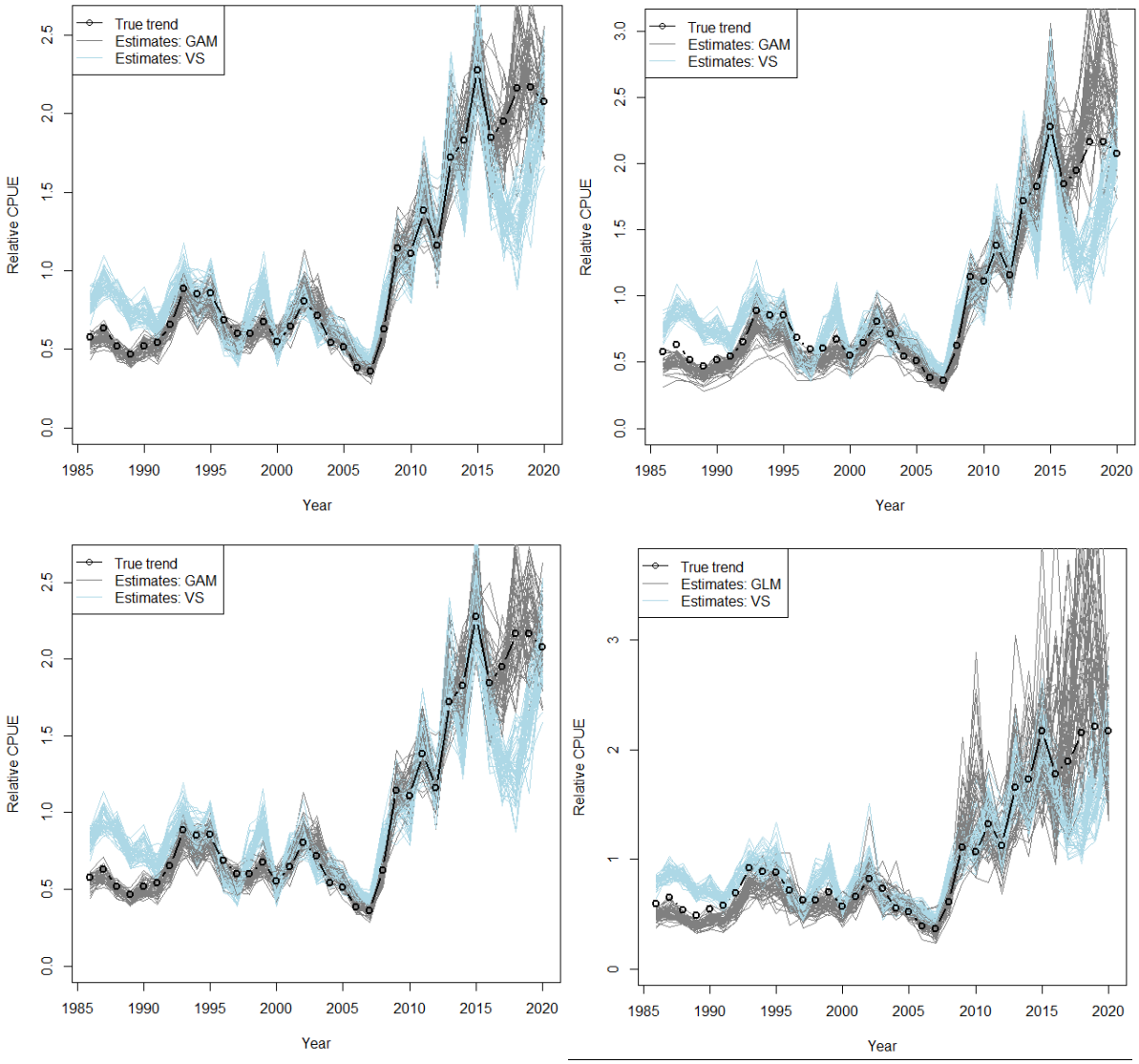

Figure 6: Estimated CS and VS indices from four sets of models. The first three are GAMs fitted using standard smoothers (top left), reduced smoothers (top right), and the smoothers applied to the operational data (bottom left), and the final model (bottom right) is the GLM used until recently. All models are fitted to data simulated from the models delta15g02 and pos15g02.

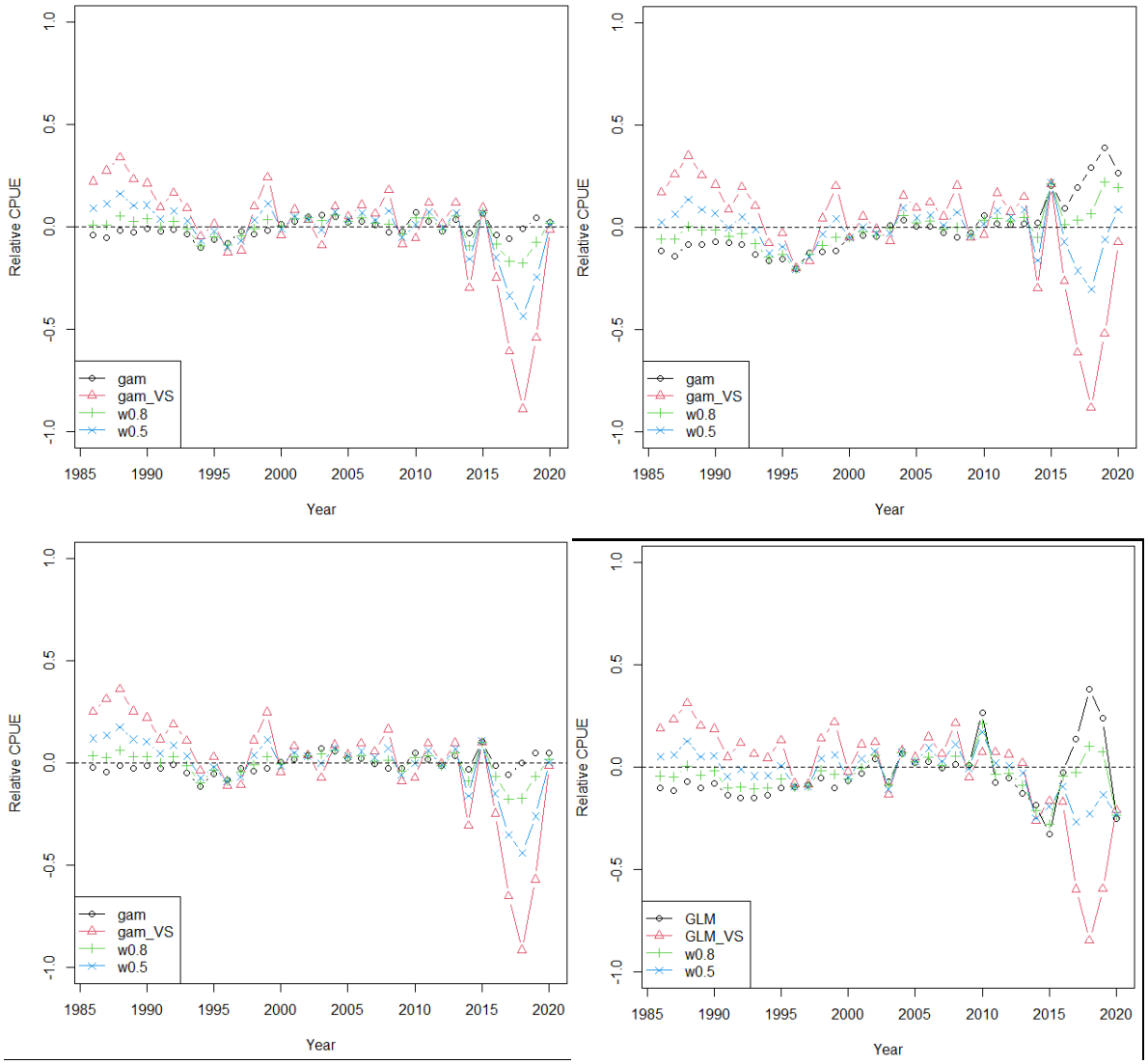

Figure 7: Bias in the estimated CS (gam), VS (gam_VS) and combined (w0.5 and w0.8) indices from four sets of models. The first three are GAMs fitted using standard smoothers (top left), reduced smoothers (top right), and the smoothers applied to the operational data (bottom left), and the final model (bottom right) is the GLM used until recently. Biases are calculated by subtracting the true value of the relative abundance from the median of the simulation estimates for each year.

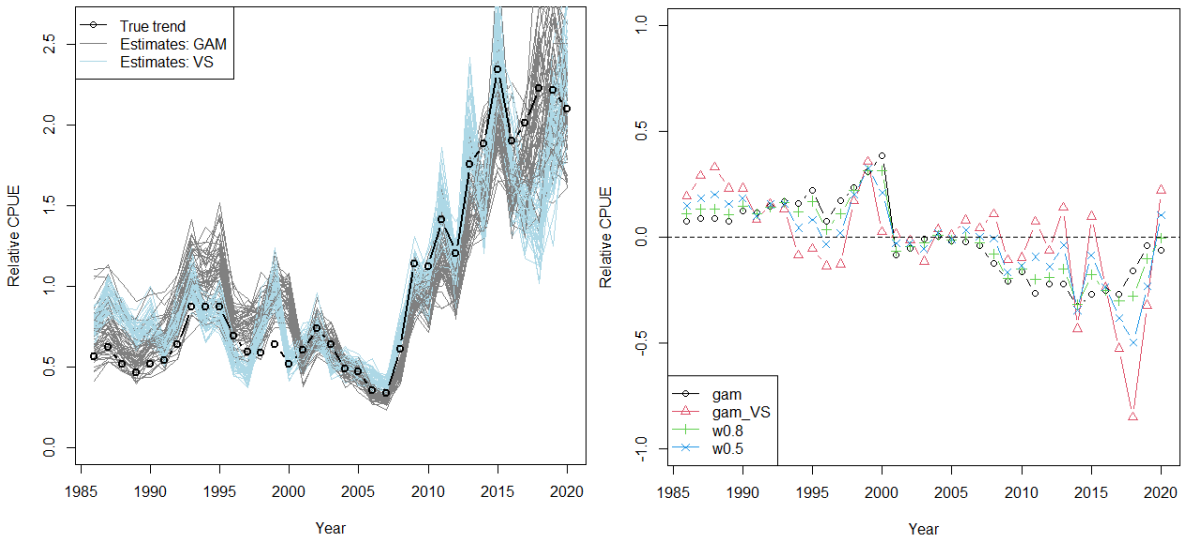

Figure 8: Effects of the 'South' dataset on estimated CS and VS indices fitted to simulated data, compared with the true trend (left), and median bias in the estimated CS (gam), VS (gam_VS) and combined (w0.5 and w0.8) indices (right).

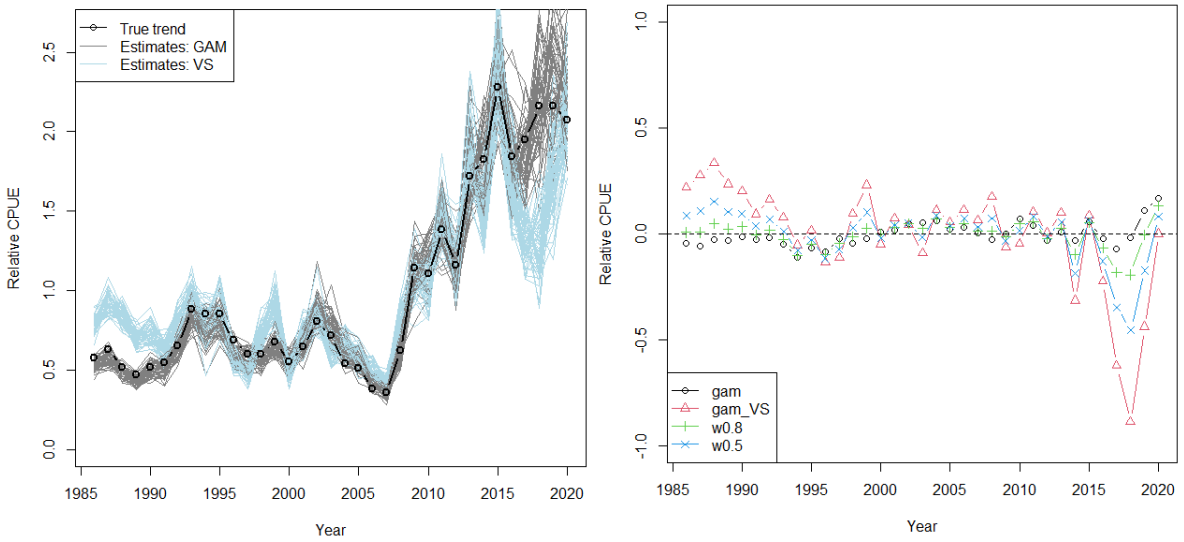

Figure 9: Effects of the 'Contract 20' dataset on estimated CS and VS indices fitted to simulated data, compared with the true trend (left), and median bias in the estimated CS (gam), VS (gam_VS) and combined (w0.5 and w0.8) indices (right).

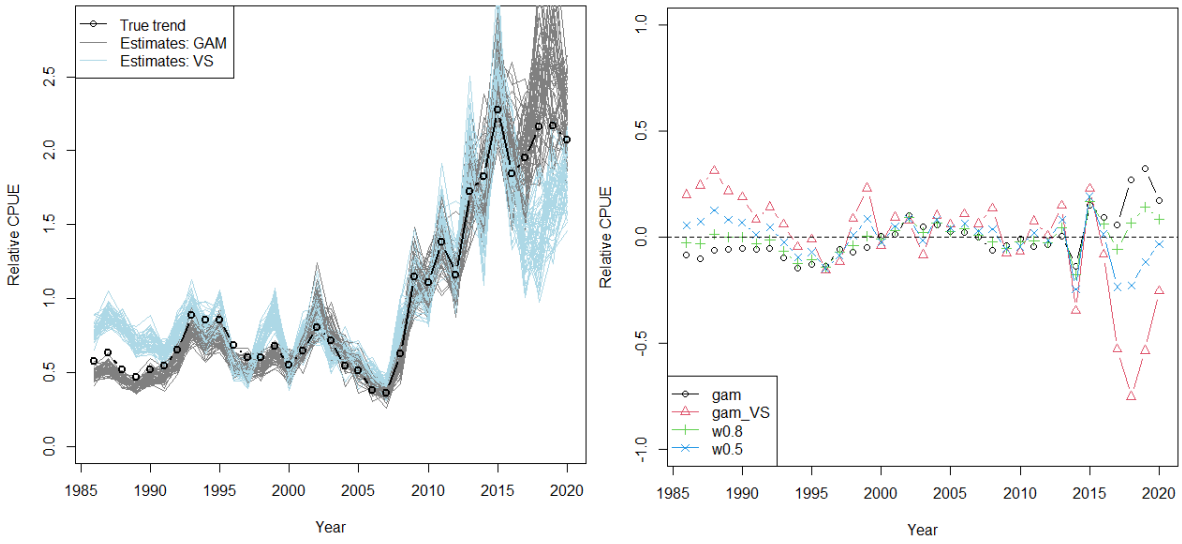

Figure 10: Effects of the 'Contract 40' dataset on estimated CS and VS indices fitted to simulated data, compared with the true trend (left), and median bias in the estimated CS (gam), VS (gam_VS) and combined (w0.5 and w0.8) indices (right).