# Appendix A – Details of the Individual-Based Model (IBM)

This Appendix describes the equations and assumptions used to generate the length-at-age datasets used in the simulation testing component of this study. The first section describes the generation of the observed data, and the second component details further how spatial variation in fish size at age is introduced. Table A1 provides parameter values used in the study.

# A.1 Generation of age-length data

The IBM is designed to mimic individual variation in growth for an exploited fishery. The model runs for 100 years. Generally, all fish within each simulation are subject to the same baseline life history parameters, with three different growth “Regimes” (defined by distinct parameters of the growth equation, see below) assigned spatial ranges accordingly (see Section A.1.4 Assigning Spatio-temporal Variation, below). Here we detail the growth component of the IBM, code to execute the simulations is available here: <https://github.com/mkapur/sab-growth/blob/master/IBM_master.R>

## A.1.1. Growth in the IBM

The growth module of the IBM is a von Bertalanffy growth function parameterized in terms of *L*1 and *L*2:

= App. Equation 1

where represents the lengths of a fish at ages , and *k* is the growth coefficient. The size of individual *i* at age *a* is defined by its length in the previous year and a growth increment *I* that is lognormal:

App. Equation 2

where ) and = 0.025 for all ages and growth Regimes (Table A1).

## A.1.2 Survival

The composition of the fishery during year *y* includes all surviving fish from recruitment to a maximum age (represented here as a plus group ). After recruitment, all fish are subject to natural mortality *M*, which in consists only of natural mortality (set to 0.25yr-1 for all ages for all years) as there is no fishery, thus fishing mortality (typically denoted *F*) and selectivity are ignored. Because no fishing pressure nor selectivity acted upon the simulated population, we are unconcerned about variation in growth that can either be engendered (over time) or misrepresented by differences in selectivity. An individual surviving the year is simulated by randomly drawing a number *u* from U[0,1] and allowing the individual to survive if this number is less than exp(-*M*).

App. Equation 3

The total population of individuals in the population at year *y* and age *a* is given by:

App. Equation 4

## A.1.3 Recruitment

Recruitment in the IBM is governed by a Beverton-Holt stock-recruitment function (Beverton and Holt, 1957), and a size-based maturity ogive which determines the probability of individual *i’*smaturity in a given year ~~age~~ *y*, . Similarly to survival, this is determined via a random variable *ui* drawn from U[0,1] for each individual. The individual is deemed mature if this variable is less than the expected maturity ogive at that individual’s length in year y .

Recruitment in a given year *Ry*is the sum of the product of the maturity ogive and empirical weights of each individual in that year, which is governed by a deterministic exponential length-weight relationship (Figure A1). The parameters of this relationship were the same for all Regimes. The maturity ogives were also fixed for all regimes, with (the length at 50% maturity) 143.68 cm, and the slope of the ogive at -0.1034. The starting stock spawning biomass (SSB0) is calculated using the sum of the product of each individual’s weight and maturity in year one. Recruitment happens at a point midway into the year, and is subject to variation via a bias-corrected lognormal recruitment deviation . The number of fish in the system is scaled by R0, or the maximum number of recruits per year; for computing efficiency we set this to twelve (see sensitivities on sample size in later sections).

App. Equation 5

App. Equation 6

App. Equation 7

App. Equation 8

App. Equation 9

## A.1.4 Assigning Spatio-temporal Variation

The simulation testing component required generation of datasets that comprised variation in fish length-at-age across space and/or time. To obtain spatial variation in length-at-age, we conducted simulations using one of two growth “Regimes”. Our synthetic populations were designed to mimic the level of variation among *L*1 and *L*2 in the sablefish dataset, which ranged from 10% to 30% between regions; we used a difference in 20% for each of *L*1 and *L*2 to generate our synthetic population. Other parameters were held constant across regimes. Spatial scenarios tested are described in Table 1 of the main text. To simulate spatial zones, fish locations were sampled from a uniform distribution with boundaries specific to a certain growth Regime. In all except Scenario 4, where the break is located at 48° and non-spatial scenarios, the latitude and longitude of fish grown under Regime 1 were sampled independently and at random from a uniform distribution between 0° and 25°; for simulations with spatial variation, fish grown under Regime 2 have latitude and longitude sampled uniformly from 25° to 50°. In Scenario 4, all simulated fish were assigned latitudes sampled independently and at random from a uniform distribution from 0° to 50°. Fish simulated under Regime 2 were assigned longitudes sampled randomly from 0° to 48° and fish simulated under Regime two have longitudes sampled randomly from 48° to 50°, forming a vertical “band” of larger fish in higher longitudes.

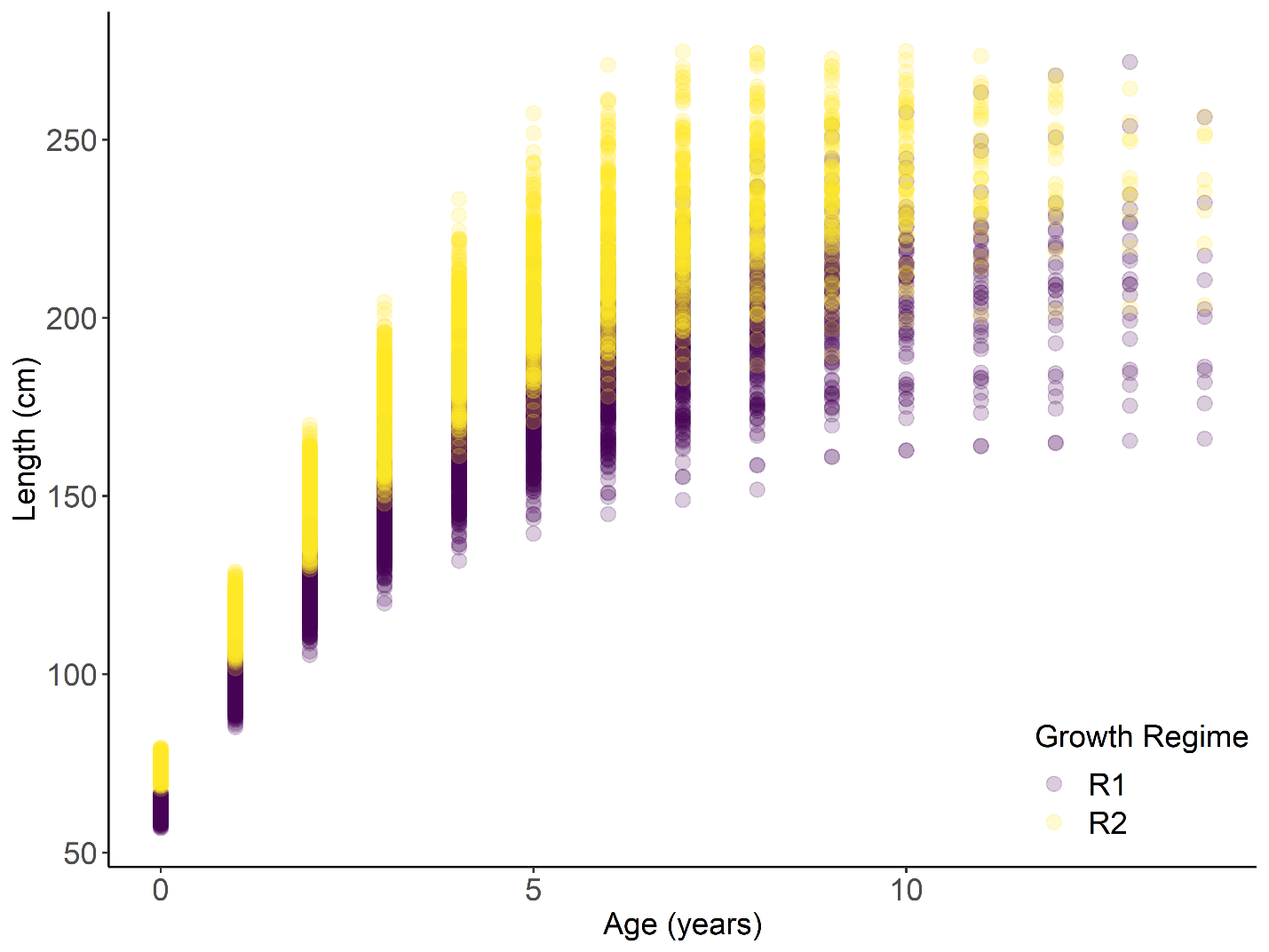


Figure A1. Example growth trajectories from simulated populations. Each circle represents a simulated individual fish’s length and age; colors correspond to the growth regime (i.e., growth curve) under which that fish was generated.

|  |  |  |  |
| --- | --- | --- | --- |
| Module | Parameter | Definition | Value |
| Growth |  | Length at age (cm) | 62 (Regime 1)  74 (Regime 2) |
| Growth |  | Length at age (cm) | 215 (Regime 1)  258 (Regime 2) |
| Growth | k | Growth coefficient (years -1) | 0.25 (Regime 1)  0.25 (Regime 2) |
| Growth |  | Age at (years) | 0 |
| Growth |  | Age at (years) | 15 |
| Growth |  | Lognormal growth error term | 0.025 |
| Growth | *a* | Multiplier of length-weight function (g/cm) | 1.35e-6 |
| Growth | *b* | Exponent of length-weight function | 3.427 |
| Survival | *M* | Natural mortality (yr-1) | 0.25 |
| Recruitment | *r* | Slope of maturity ogive | -0.1034 |
| Recruitment | *L50* | Length at 50% maturity (cm) | 143.68 |
| Recruitment | *h* | Steepness of Beverton-Holt SRR | 0.9 |
| Recruitment | *R0* | Maximum number of recruits per year | 12 |
| Recruitment |  | Variation in recruitment | 0.1 |

Table A1. Parameter symbols, definitions and values used in the simulation study.

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | | | **Original sample size (average # age six fish = 530)** | | **With sample size halved** | | **With sample size reduced by 25%** | |
| **Scenario Number** | **Scenario Description** | **True Break Points** | **Coverage probability for L1, L2** | **Proportion correct latitude, longitude, year** | **Coverage probability for L1, L2** | **Proportion correct latitude, longitude, year** | **Coverage probability for L1, L2** | **Proportion correct latitude, longitude, year** |
| 1 | No spatial breaks | None | 0.95, 0.89 | 0.94, 0.92, 0.61 | 0.96, 0.93 | 0.84, 0.89, 0.63 | 0.96, 0.92 | 0.88, 0.91, 0.71 |
| 2 | Single, spatial break in middle of range, with no overlap and strong contrast | 25° Latitude and 25° Longitude | 0.49, 0.47 | 0.94, 0.96, 0.75 | 0.48, 0.47 | 0.77, 0.65, 0.77 | 0.47, 0.45 | 0.83, 0.87, 0.81 |
| 3 | Some overlap between regions | Between 20° and 25° Latitude | 0.26, 0.24 | 1, 1, 0.85 | 0.29, 0.27 | 0.98, 0.97, 0.8 | 0.27, 0.25 | 0.99, 0.99, 0.87 |
| 4 | Single spatial break at edge of range with no overlap | 48° Longitude | 0.63, 0.59 | 0.75, 0.2, 0.78 | 0.56, 0.52 | 0.75, 0.11, 0.85 | 0.6, 0.57 | 0.73, 0.17, 0.8 |
| 5 | Single temporal break at year 50 (of 100); no spatial variability | None for latitude or longitude; all fish under Regime 1 from years 0 to 49 and Regime 2 thereafter | 0.76, 0.74 | 0.92, 0.93, 0.29 | 0.7, 0.64 | 0.9, 0.9, 0.18 | 0.77, 0.73 | 0.9, 0.9, 0.3 |

Table A2. Summary of true break points, coverage probabilities of the endpoints of the post-aggregation growth curves, and the proportion of simulations which detected the exact breakpoints each or all of the three smoothers. For the overlapping scenario (Scenario 3), spatial breakpoints were considered a match if they fell within the true range. This analysis was repeated for the same datasets with the number of age-six fished reduced by either 50% or 25%.

# A.2 – Additional tables and figures from the GAM-based analysis of sablefish size at age.

|  |  |  |
| --- | --- | --- |
| **Age** | **Sex** | **n** |
| 4 | F | 4366 |
| 4 | M | 3204 |
| 6 | F | 4413 |
| 6 | M | 3404 |
| 10 | F | 2064 |
| 10 | M | 1765 |
| 30 | F | 168 |
| 30 | M | 231 |

Table A3. Number of sablefish at key ages by sex and used in full data fitting.

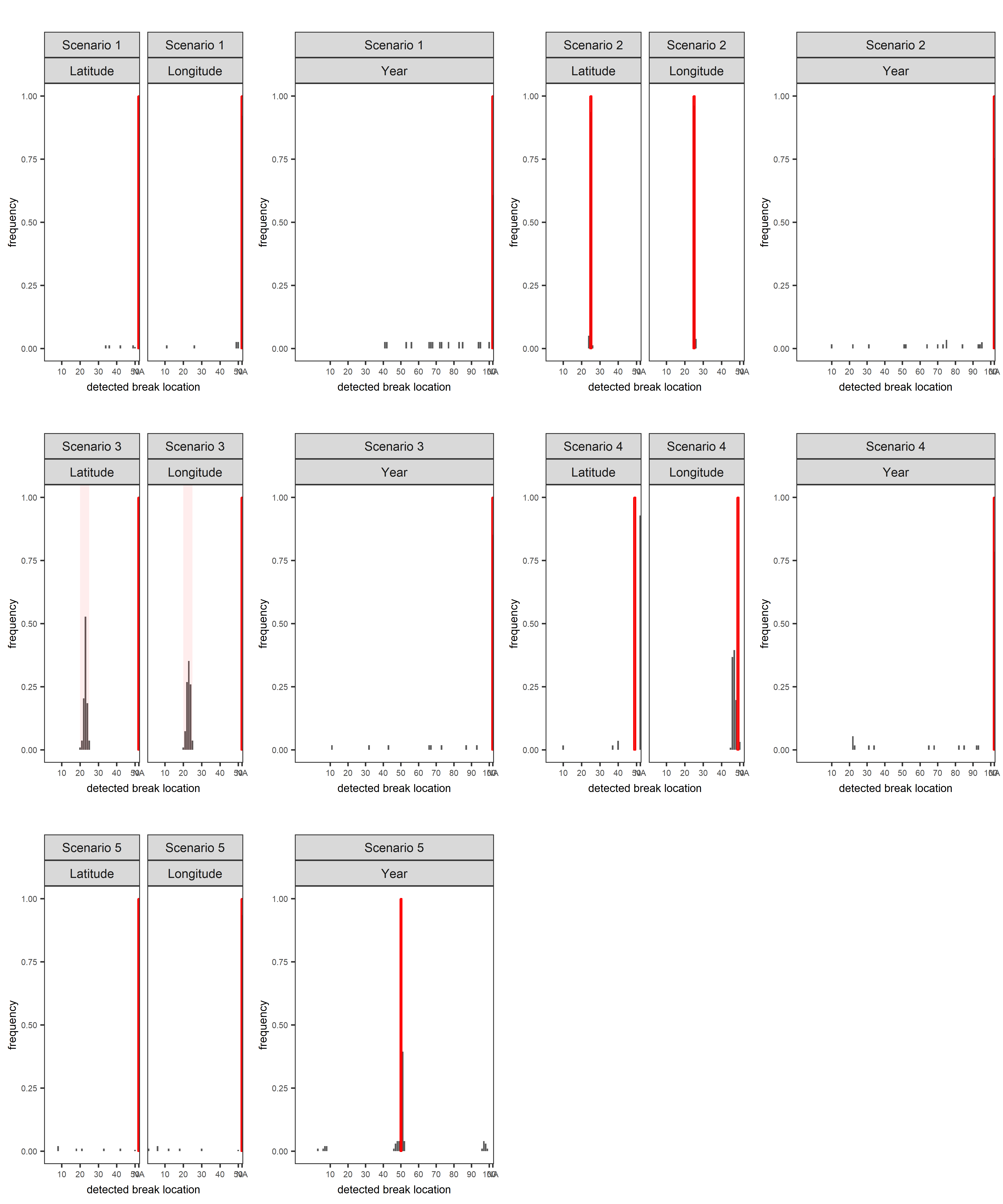


Figure A2 Histogram of detected breakpoints (grey bars) from GAM analysis by scenario. Vertical red bars indicate true breakpoints used to generate synthetic populations. For Scenario 3, the synthetic population overlapped between 20 and 25 degrees latitude and longitude.

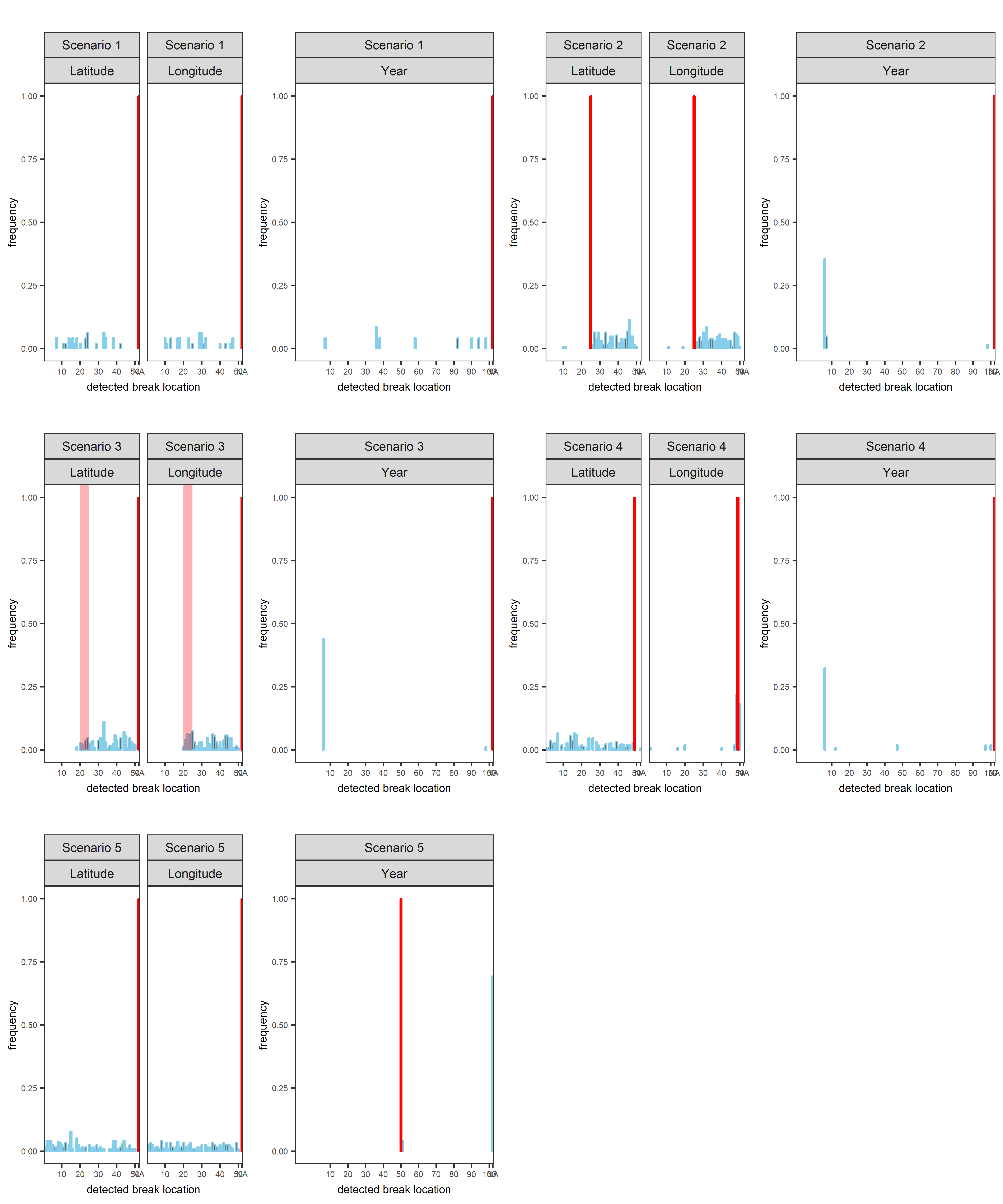


Figure A3 Histogram of detected breakpoints (grey bars) from STARS analysis by scenario. Vertical red bars indicate true breakpoints used to generate synthetic populations. For Scenario 3, the synthetic population overlapped between 20 and 25 degrees latitude and longitude.

Figure A4 through Figure A13 are identical in form to Figure 5 and 6 in main text, which presented results for age six female sablefish. These plots contain results for ages four and thirty for males and females, and age-four males.



Figure A4. Diagnostic plots of best-fit GAM model for female age four sablefish. Clockwise from top left: quantile-quantile plot of deviance residuals; histogram of residuals; observed response values (lengths, in cm) vs predicted values, and model-predicted residuals vs linear predictor. See Supplementary Material for equivalent plots for other key ages and all sexes.



Figure A5 Diagnostic plots of best-fit GAM model for male age four sablefish. Clockwise from top left: quantile-quantile plot of deviance residuals; histogram of residuals; observed response values (lengths, in cm) vs predicted values, and model-predicted residuals vs linear predictor.

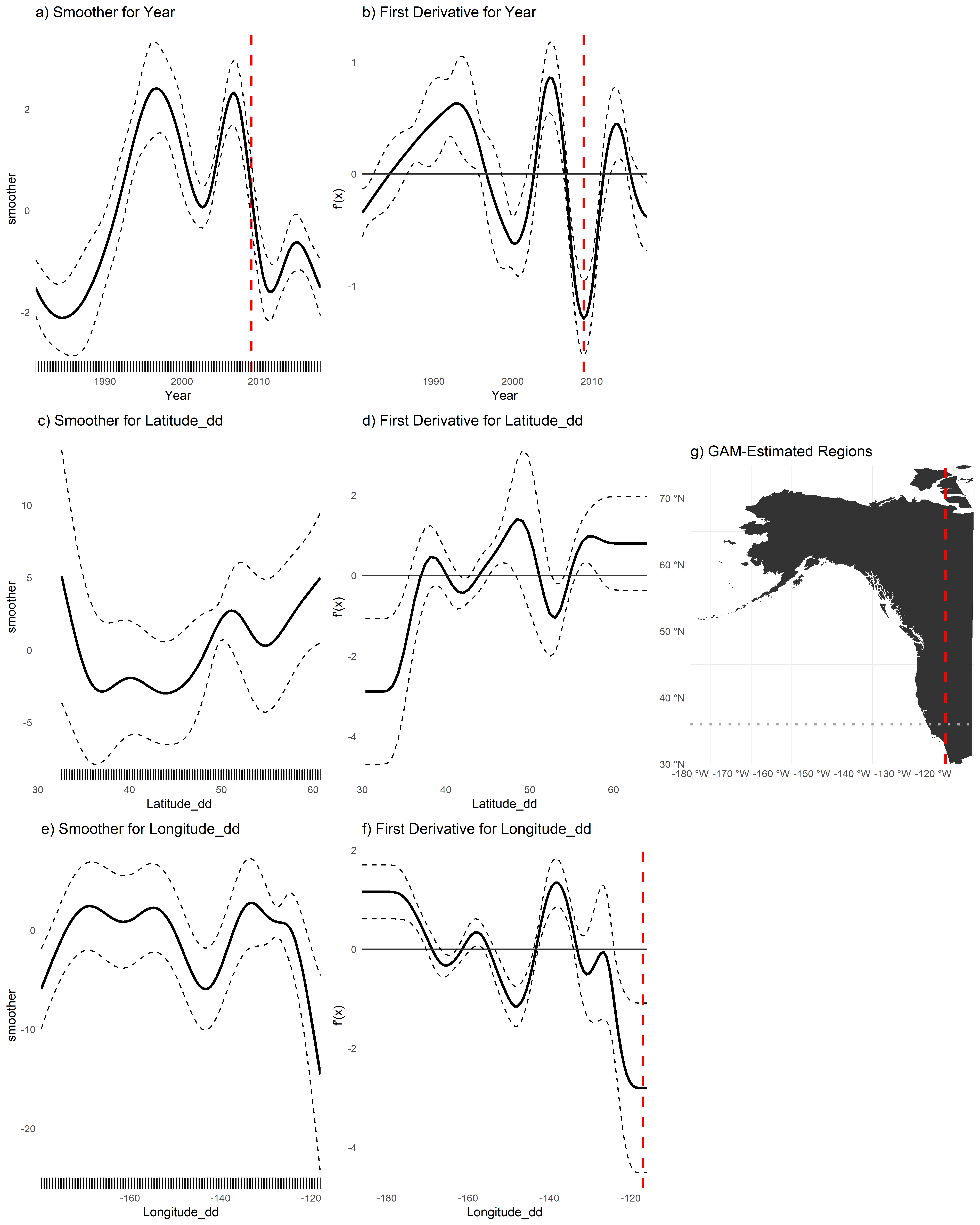


Figure A6 (a,c,e) Plots of smoothers for Year, Latitude, and Longitude, and first derivatives thereof for age-four male sablefish (b,d,f). Red lines indicate latitudes or longitudes that produced the highest first derivative and had a confidence interval that did not include zero. g) map with model-detected breakpoints (red lines).



Figure A7 Diagnostic plots of best-fit GAM model for female age six sablefish. Clockwise from top left: quantile-quantile plot of deviance residuals; histogram of residuals; observed response values (lengths, in cm) vs predicted values, and model-predicted residuals vs linear predictor.

 Figure A8 Diagnostic plots of best-fit GAM model for male age six sablefish. Clockwise from top left: quantile-quantile plot of deviance residuals; histogram of residuals; observed response values (lengths, in cm) vs predicted values, and model-predicted residuals vs linear predictor.

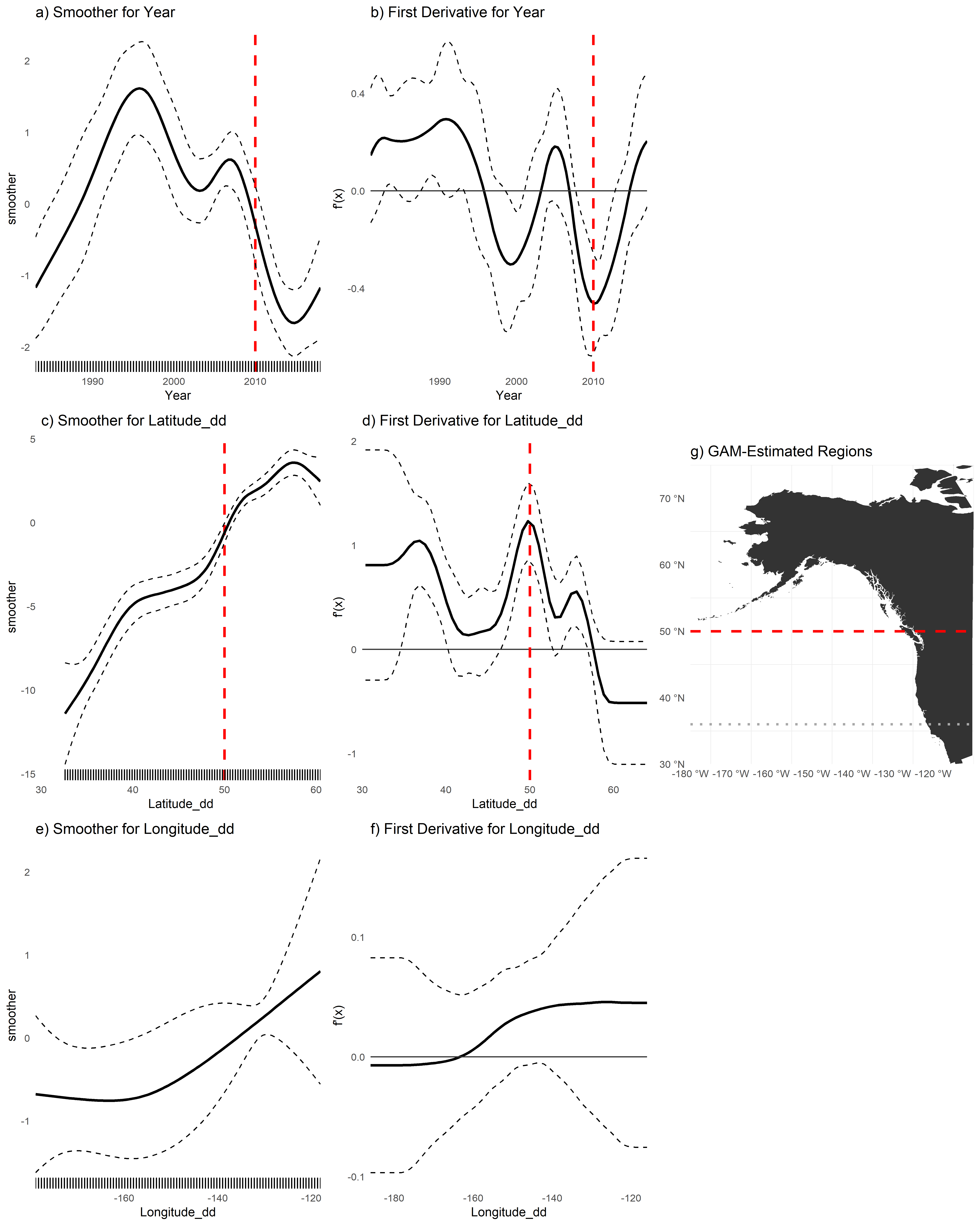
Figure A9 (a,c,e) Plots of smoothers for Year, Latitude, and Longitude, and first derivatives thereof for male age six sablefish (b,d,f). Red lines indicate latitudes or longitudes that produced the highest first derivative and had a confidence interval that did not include zero.g) map with model-detected breakpoints (red lines).

Figure A10 Diagnostic plots of best-fit GAM model for female age thirty sablefish. Clockwise from top left: quantile-quantile plot of deviance residuals; histogram of residuals; observed response values (lengths, in cm) vs predicted values, and model-predicted residuals vs linear predictor.

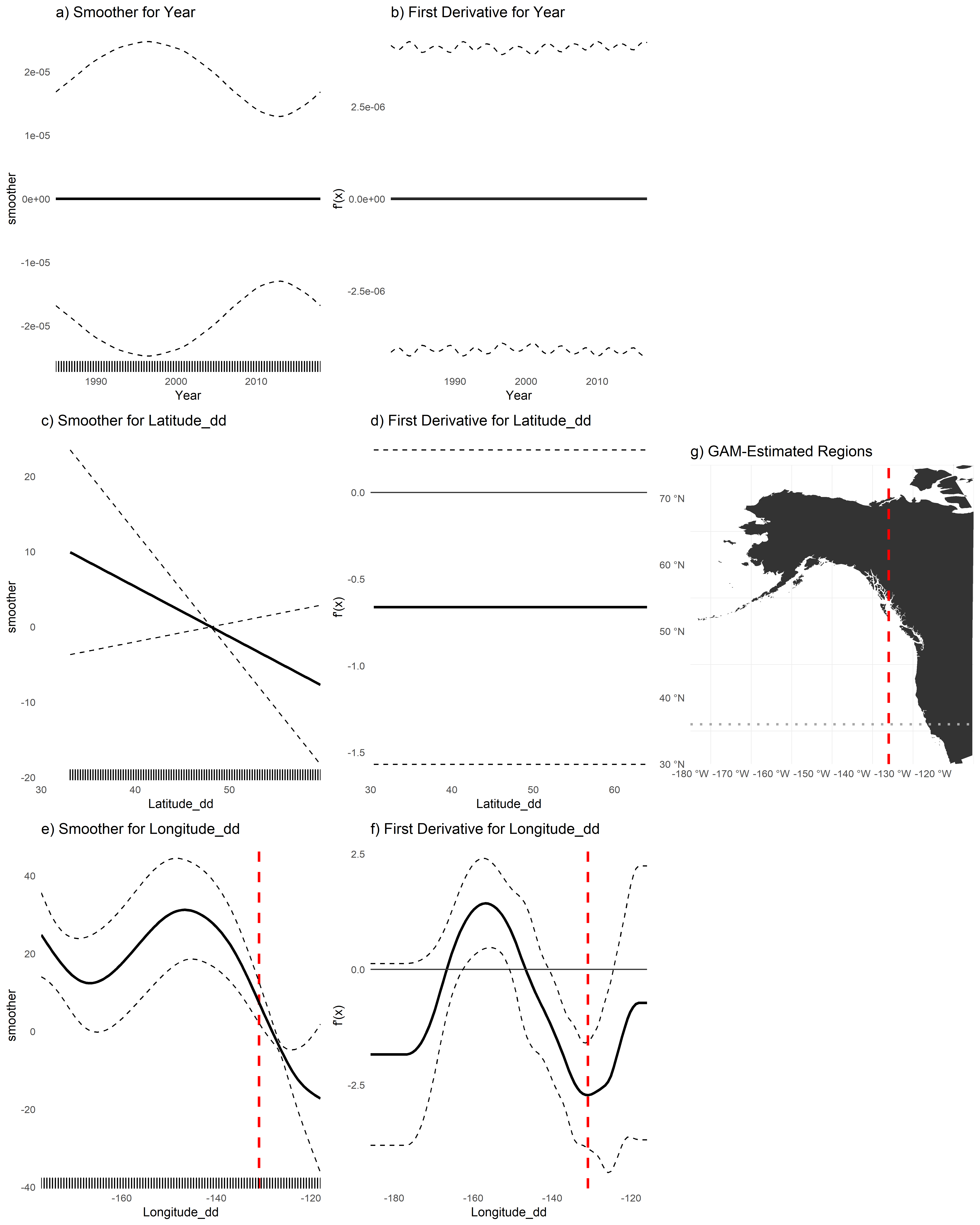
Figure A11 (a,c,e) Plots of smoothers for Year, Latitude, and Longitude, and first derivatives thereof for female age thirty sablefish (b,d,f). Red lines indicate latitudes or longitudes that produced the highest first derivative and had a confidence interval that did not include zero.g) map with model-detected breakpoints (red lines).

Figure A12 Diagnostic plots of best-fit GAM model for male age thirty sablefish. Clockwise from top left: quantile-quantile plot of deviance residuals; histogram of residuals; observed response values (lengths, in cm) vs predicted values, and model-predicted residuals vs linear predictor.

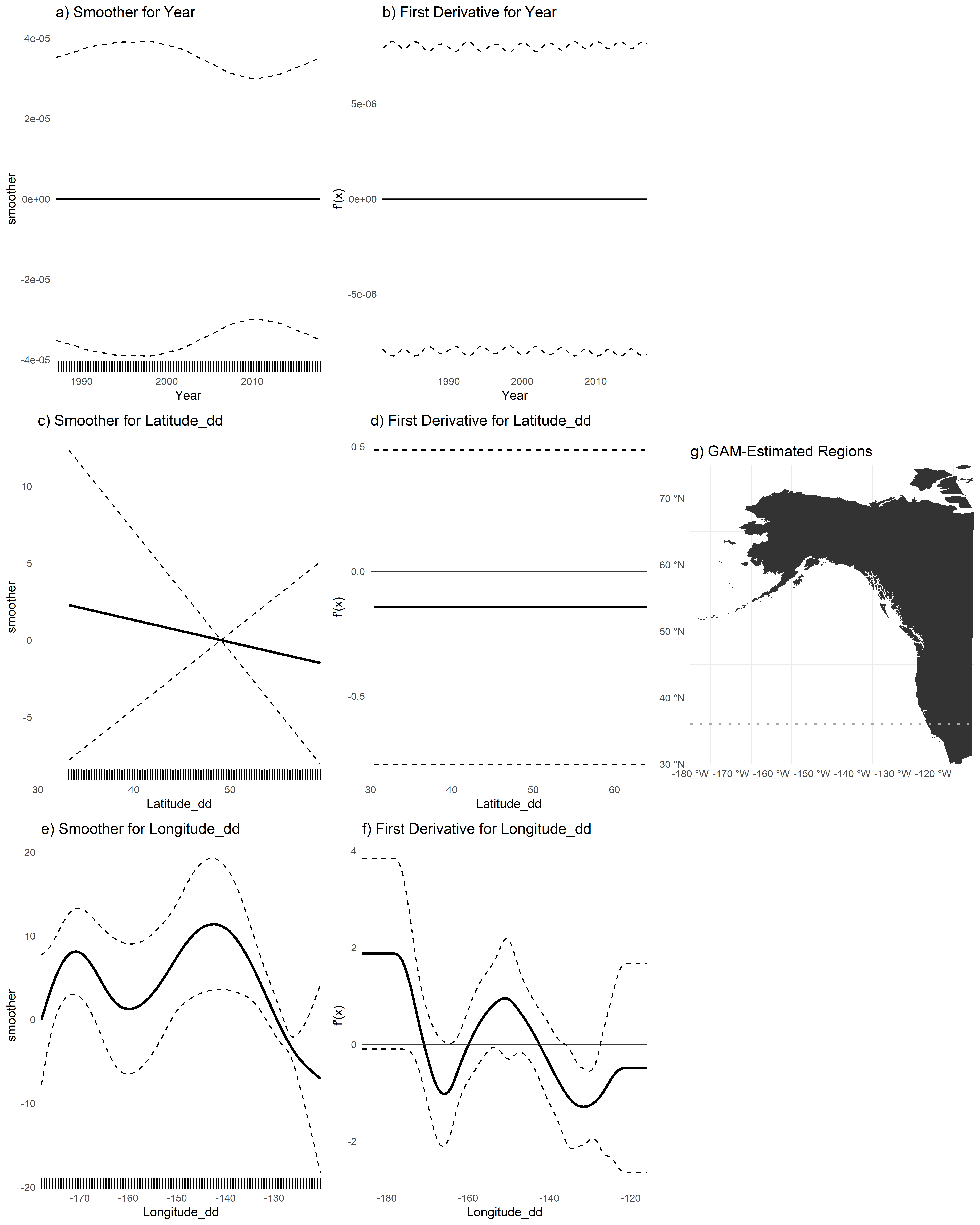
Figure A13 (a,c,e) Plots of smoothers for Year, Latitude, and Longitude, and first derivatives thereof for male age thirty sablefish (b,d,f). Red lines indicate latitudes or longitudes that produced the highest first derivative and had a confidence interval that did not include zero. g) map with model-detected breakpoints (red lines).



Figure A14 *L∞* estimates for the fully stratified, 5-region, 2-period (during and after 2010, and before) and 2-sex model. Bars represent 95% confidence intervals. Strata from the same spatial region and sex that shared overlapping ranges for *L∞* are colored in red and early and late periods were combined within their respective regions and sexes for the subsequent analysis.



Figure A15 Fits of von Bertalanffy growth function (black lines) to data for Phase 1 spatio-temporal aggregation. Points are raw survey data colored by their source. Line types denote whether fit is for early, late or pooled time period.

**

Figure A16 *L∞* estimates for the second (final) aggregation phase, where a time break at 2010 was only applied to both sexes for regions 3, 4 and 5; Regions 1 and 2 use data for all years combined and still estimate sexes separately. Bars represent 95% confidence intervals.

# A.3 STARS Method Results

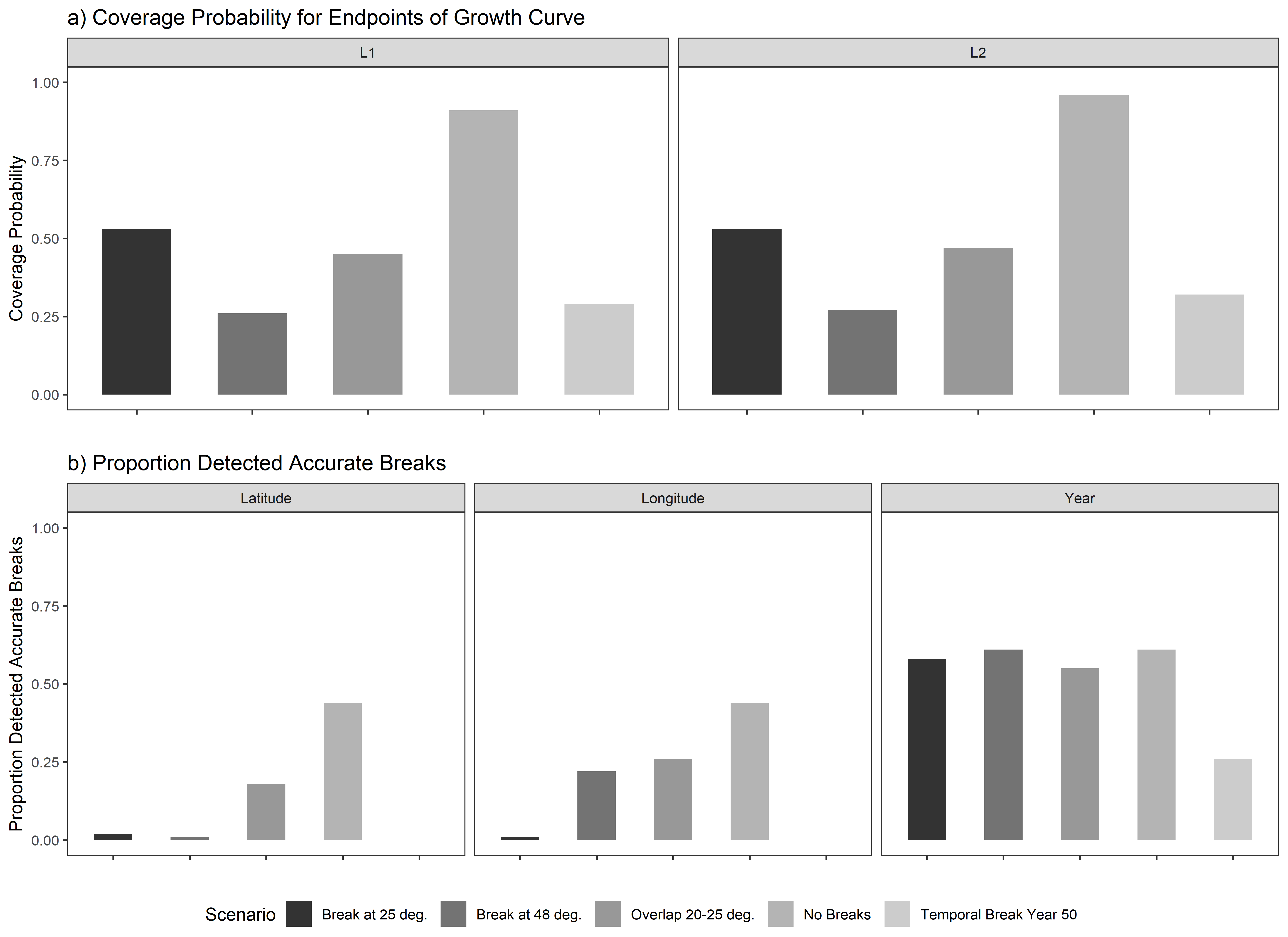


Figure A17. Using the STARS method (Rodionov, 2004) a) coverage probabilities for the endpoints of the growth curve, L1 (left) and L2 (right), and b) proportion of 100 simulations for each spatial scenario wherein the correct latitudinal breaks (left), or longitudinal breaks (center) or yearly break (right) were detected.

# References

Beverton, R.J.H., Holt, S.J., 1957. On the Dynamics of Exploited Fish Populations, Fisheries Investigations Series 2: Sea Fisheries. https://doi.org/10.1007/BF00044132

Rodionov, S.N., 2004. A sequential algorithm for testing climate regime shifts. Geophys. Res. Lett. 31, 2–5. https://doi.org/10.1029/2004GL019448