ArboMAP: Arbovirus Modeling and Prediction to Forecast Mosquito-Borne Disease Outbreaks

Summary of Model Outputs (v2.1)

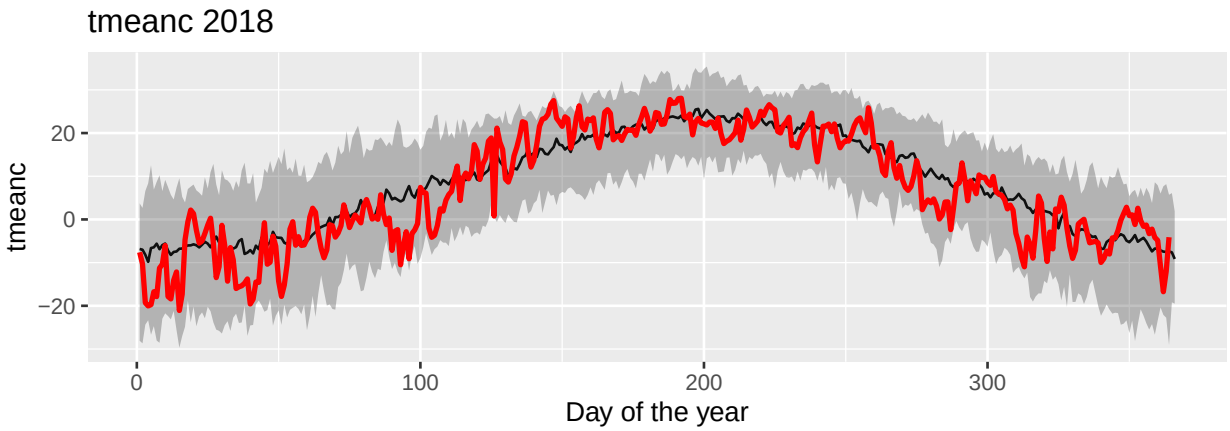
Justin K. Davis and Michael C. Wimberly
(justinkdavis@ou.edu, mcwimberly@ou.edu)
Geography and Environmental Sustainability, University of Oklahoma

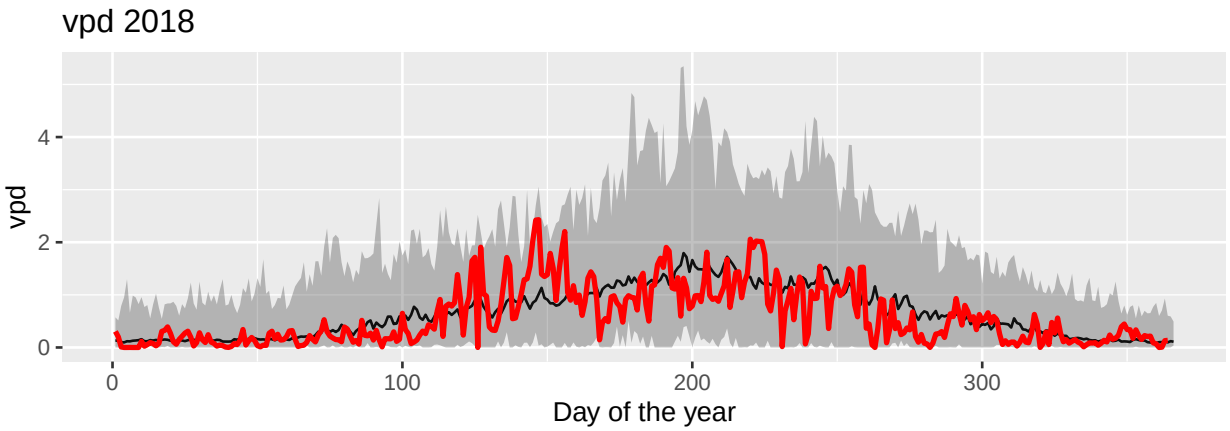
Updated November 16, 2019

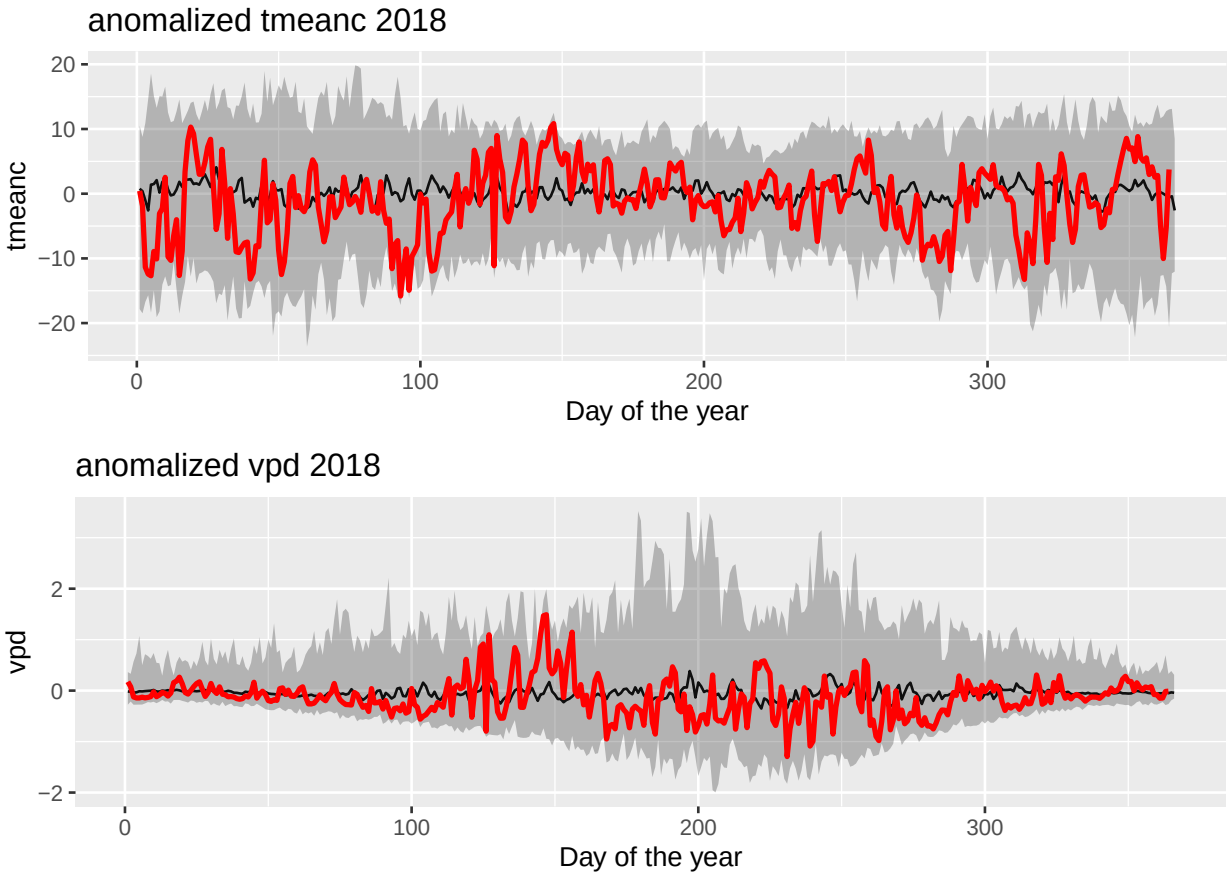
Data used for predictions

Weather data

Weather data from the gridMET data set range from 2000-01-01 to 2018-12-30. Below are graphs of statewide daily averages of tmeanc and vpd. Observations for the current year are in red. Black is the medium from all other years, and the grey band indicates the max/min ever observed. Below this are the anomalized weather indices, from which the weekly averages have been subtracted to show deviations above/below the mean.

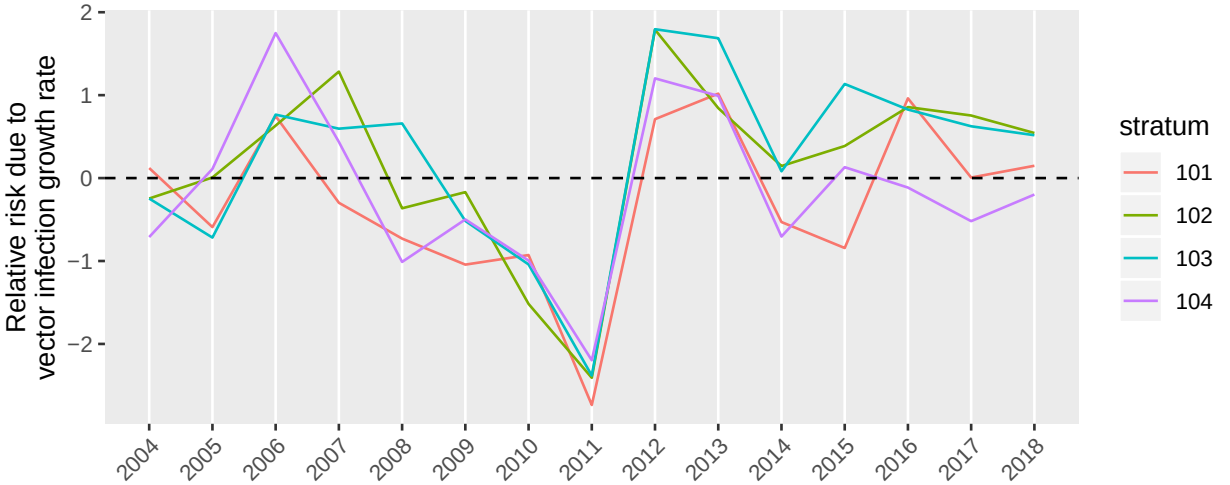


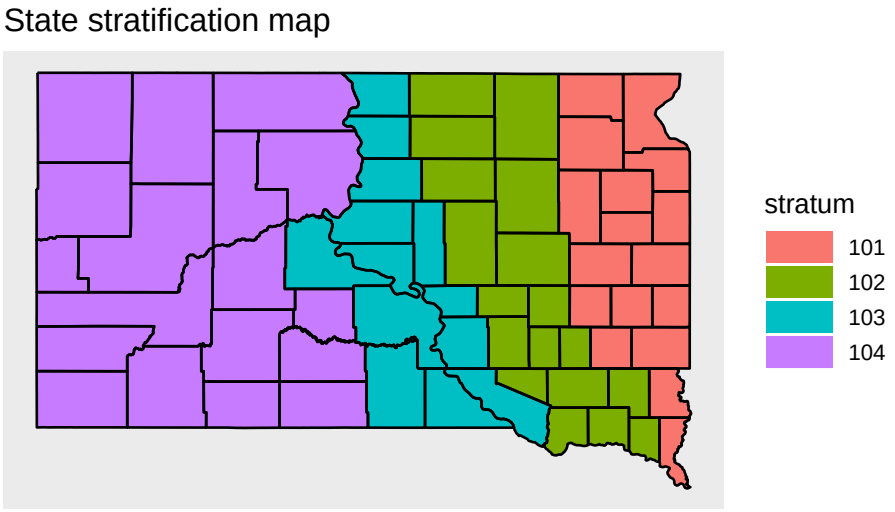




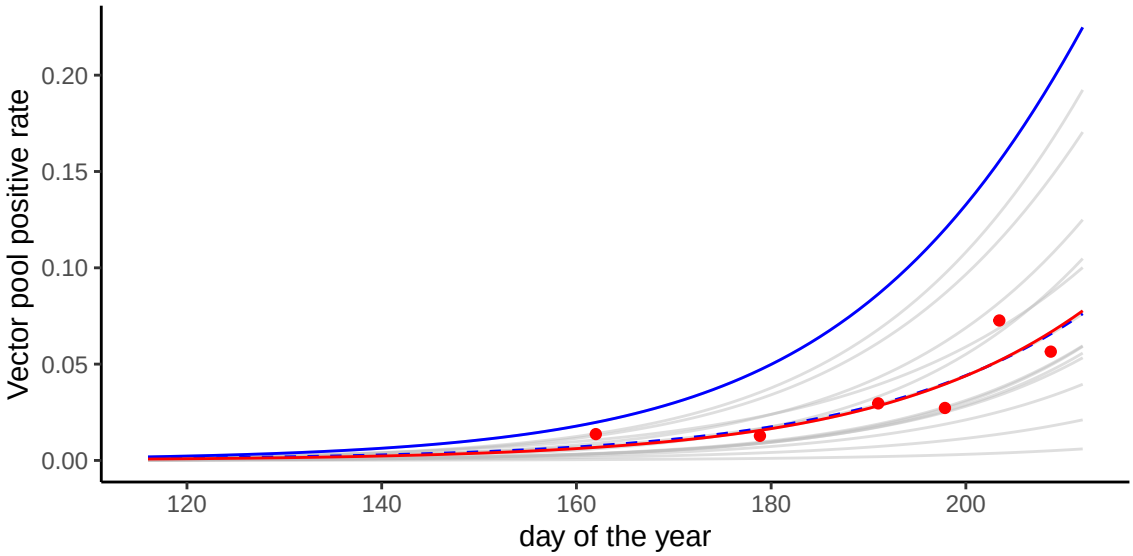
Vector infection data

There are 25791 samples in the vector testing database. For 2018, there are 1965 tested samples, with 67 (3.4%) positive. The estimated risk of human infection due to the early-season vector infection growth rate is shown below. Higher means that the pathogen is spreading more rapidly among vectors, and more human cases should be expected. The regions used for stratification are mapped below; districts are thought to share risk more closely with others in the same strata, although all districts share some level of risk with all other districts.



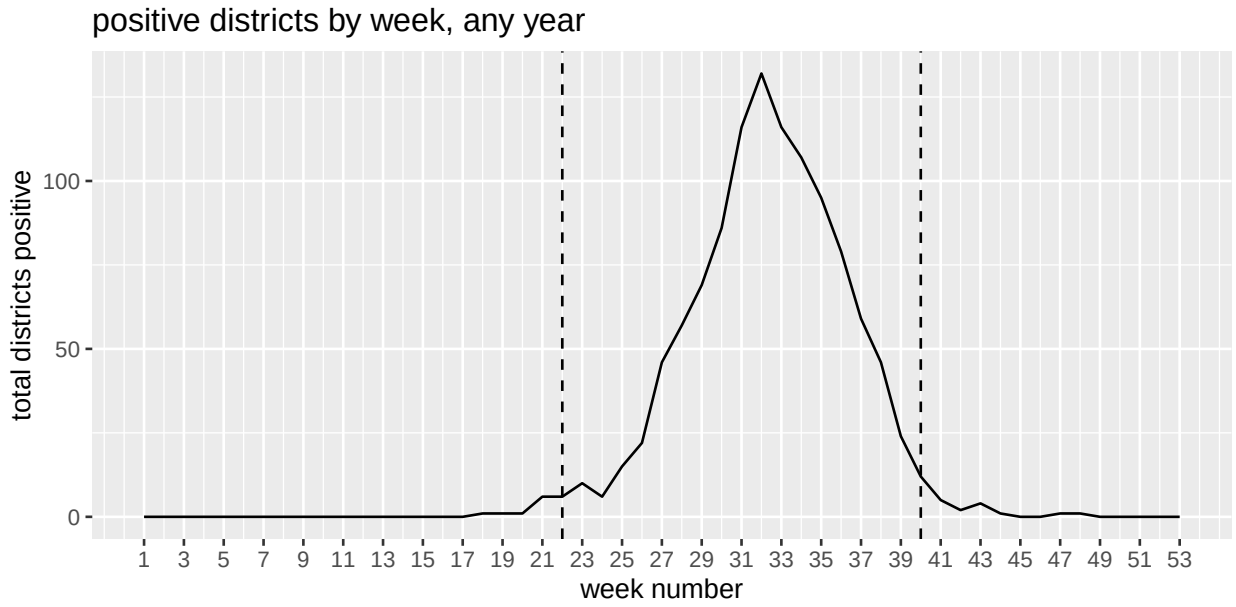


The following graph shows the estimated growth of positive samples for every year (grey), with 2012 (blue) and 2017 (blue, dashed) selected for comparison, and estimates and observations for 2018 (red). The lines are modeled sample positive rates; the actual statewide positive sample rate for 2018 is shown here by grouping observations nearby in time.



Human data

In the graph below, positive districts by week are shown over all years. Only weeks within the two dashed lines (excluding 2% total of the earliest and latest cases) are used in modeling for numerical stability.

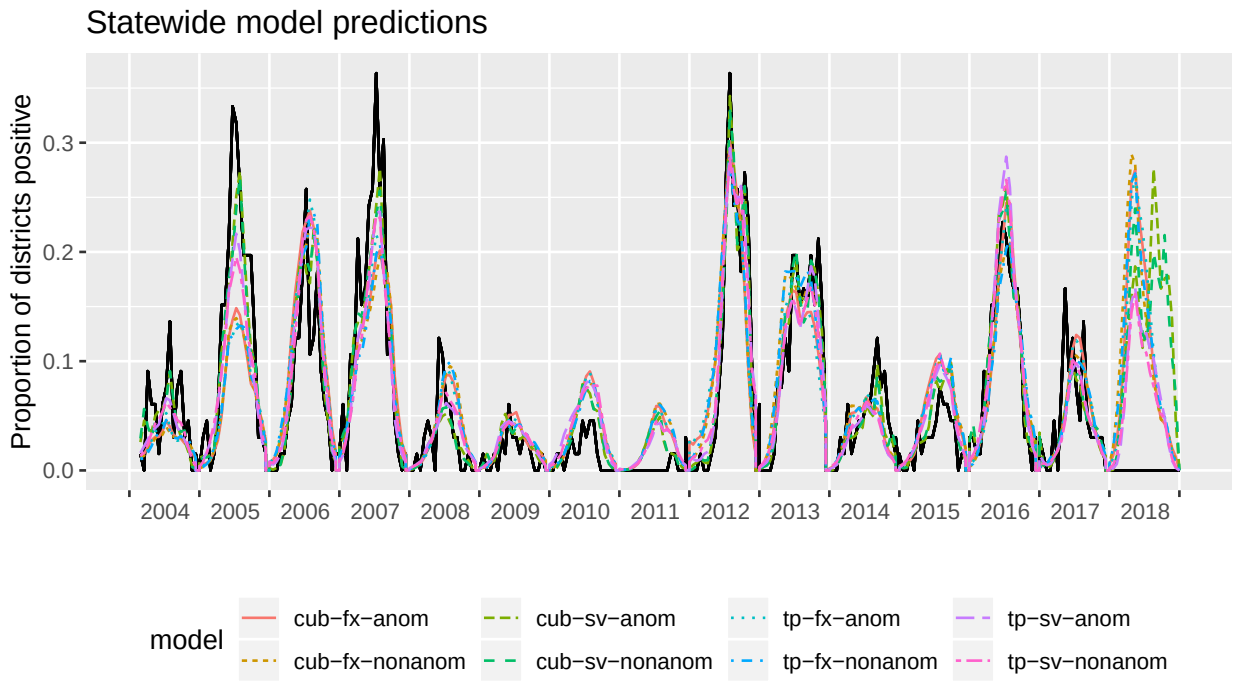


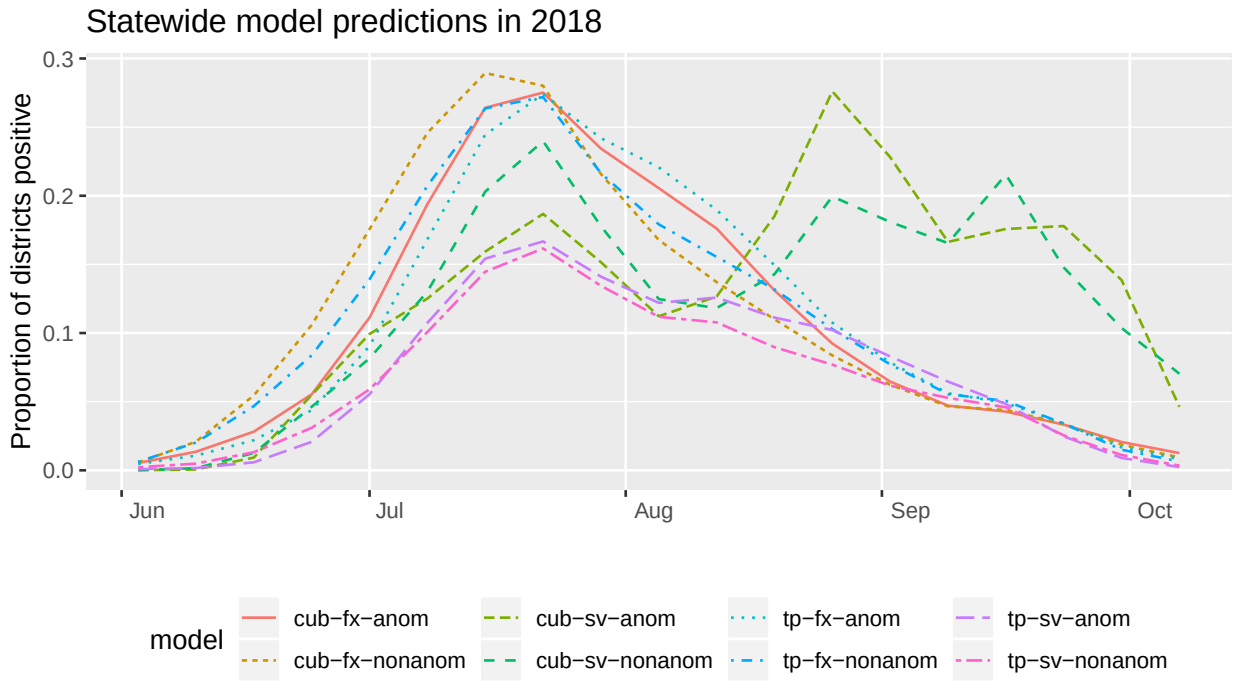
The predictive model of human cases was calibrated using 1486 historical cases, not including any cases from 2018. No cases from 2018 are used to make predictions; the estimates for this year are based solely on weather and vector data. Typically, 92% of a year's cases occur before the end of this week in any given year.

Model results

Statewide trends

The graphs below show observed statewide risk (black) and estimated risk (red) up to Sunday 2018-12-30. Observed risk should be completely zero during the last year - these data are not used in the model, and will only be updated once final human case data are received at the end of the year.





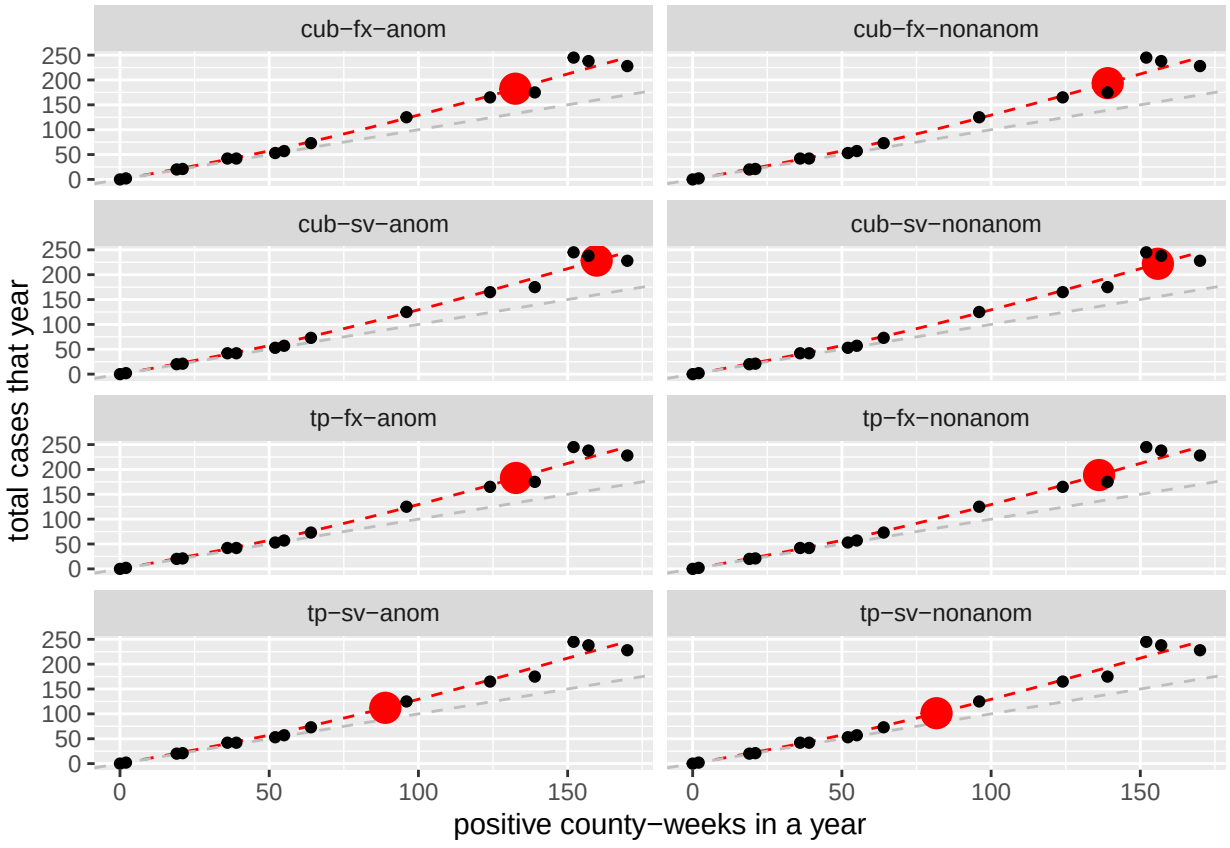
Model fit statistics are shown below.

	model	aic	auc
4	cub-sv-anom	6207.22	0.87
3	cub-sv-nonanom	6215.55	0.87
8	tp-sv-anom	6243.27	0.86
7	tp-sv-nonanom	6269.28	0.86
2	cub-fx-anom	6405.17	0.85

	model	aic	auc
1	cub-fx-nonanom	6410.01	0.85
6	tp-fx-anom	6413.05	0.85
5	tp-fx-nonanom	6416.23	0.85

Estimates for the week in question are shown below.

model	prop. positive	districts positive
cub-fx-anom	4.7%	3.1
cub-fx-nonanom	4.7%	3.1
cub-sv-anom	16.6%	11.0
cub-sv-nonanom	16.5%	10.9
tp-fx-anom	5.6%	3.7
tp-fx-nonanom	5.5%	3.7
tp-sv-anom	6.5%	4.3
tp-sv-nonanom	5.3%	3.5

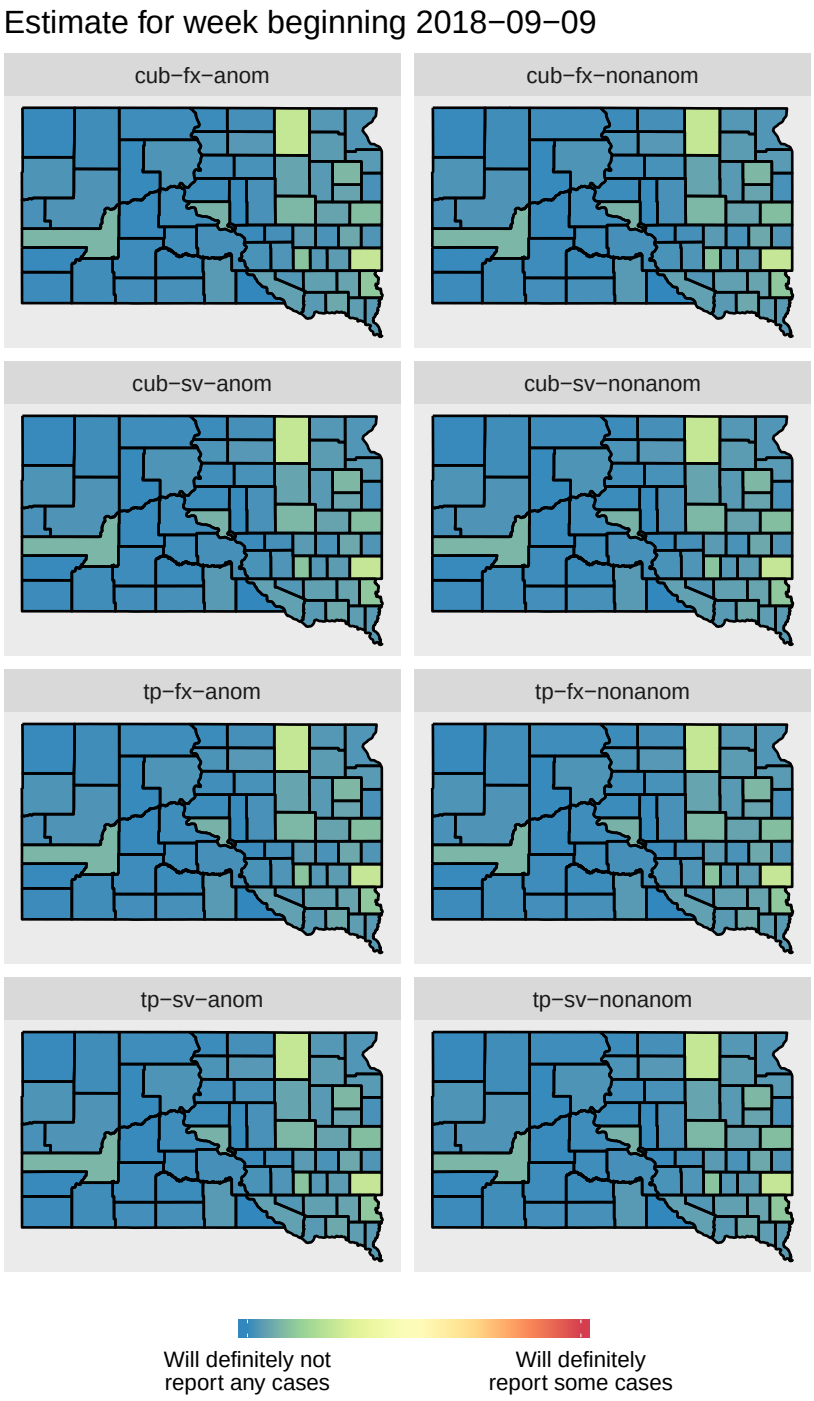


model	est. positives	total est. cases
cub-fx-anom	132.5	182.5
cub-fx-nonanom	139.1	193.6
cub-sv-anom	159.7	228.3
cub-sv-nonanom	155.9	222.0
tp-fx-anom	132.7	182.8

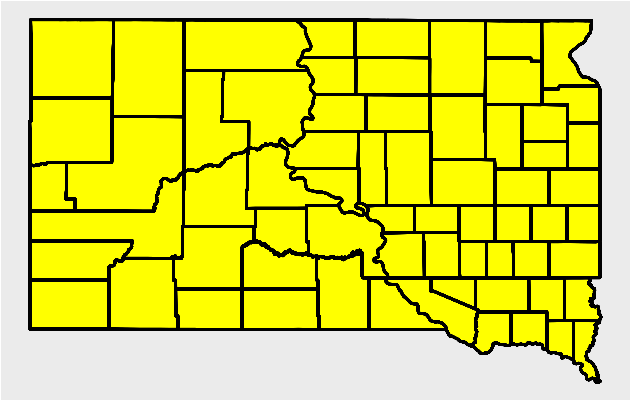
model	est. positives	total est. cases
tp-fx-nonanom	136.2	188.7
tp-sv-anom	88.9	112.2
tp-sv-nonanom	81.7	101.3

Results for 2018-09-09 to 09-15

We visualize the raw estimated risk for 2018-09-15 below. If a district is darkest blue, then we estimate that there should be no human cases reported for this district, during this week. If a district is brightest red, we are certain that there will be at least one human case reported for this district, during this week.



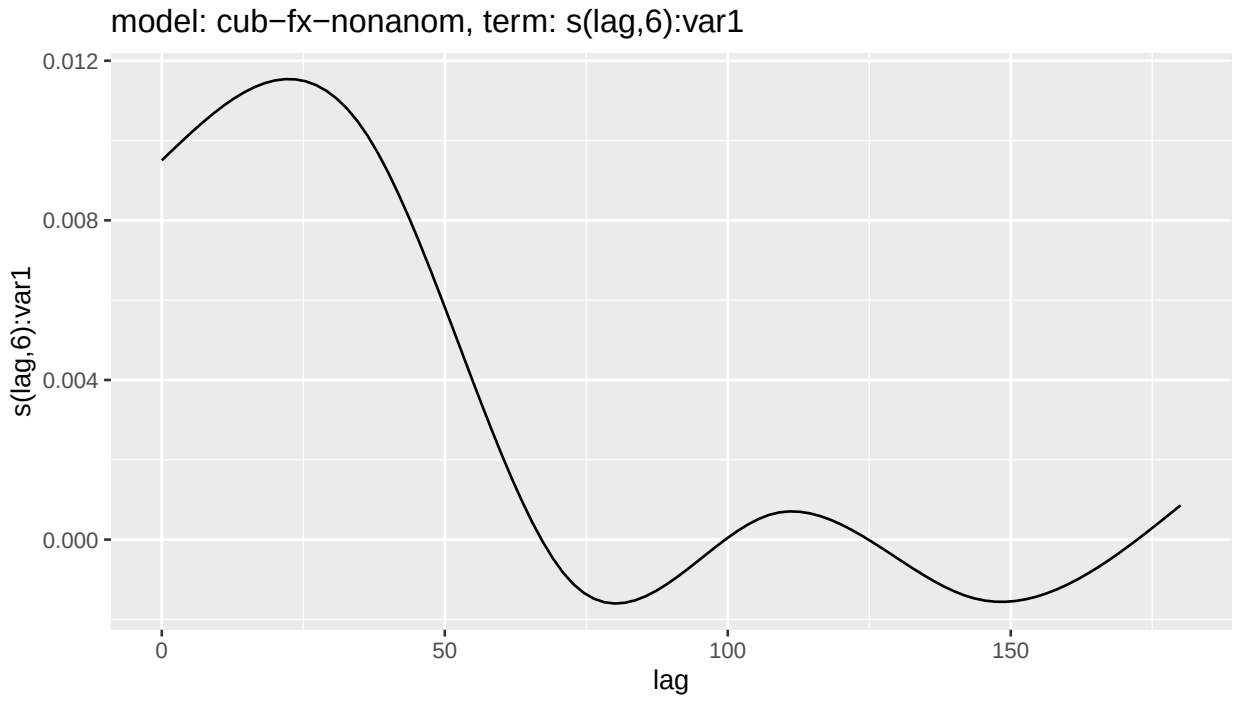
This map indicates whether probabilities reported in the previous map are higher (red) than average, lower (blue) than average, or right about normal (yellow) compared to the same week in previous years.

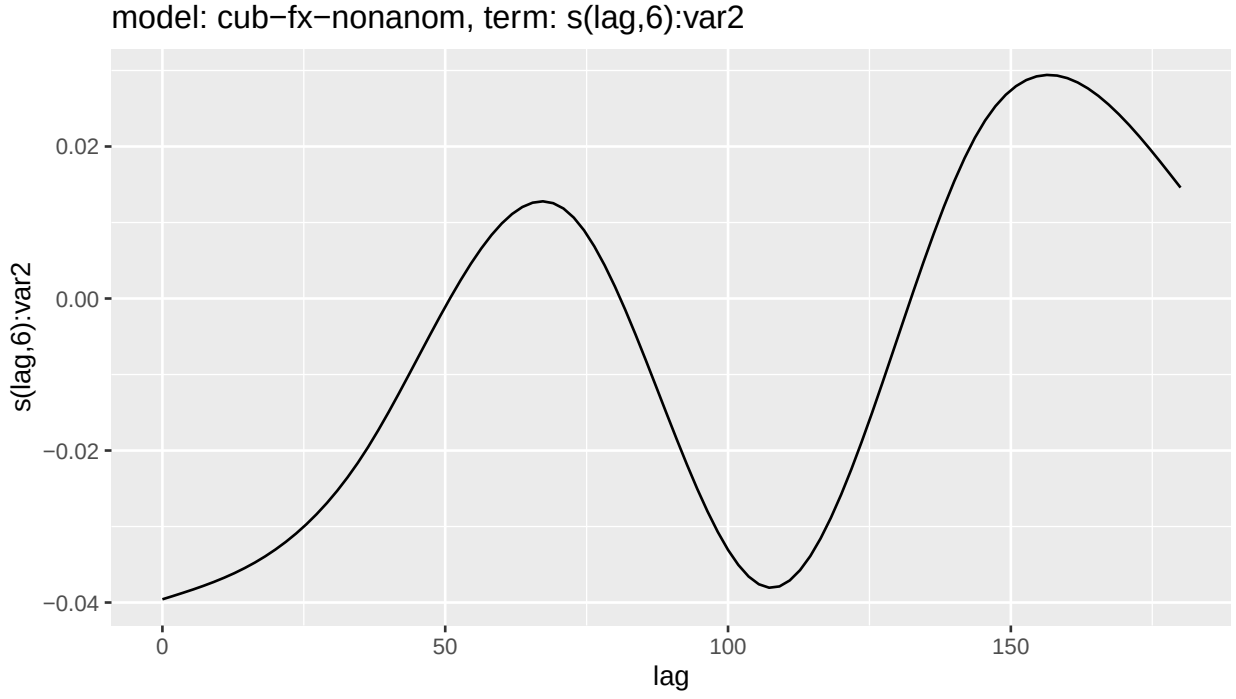


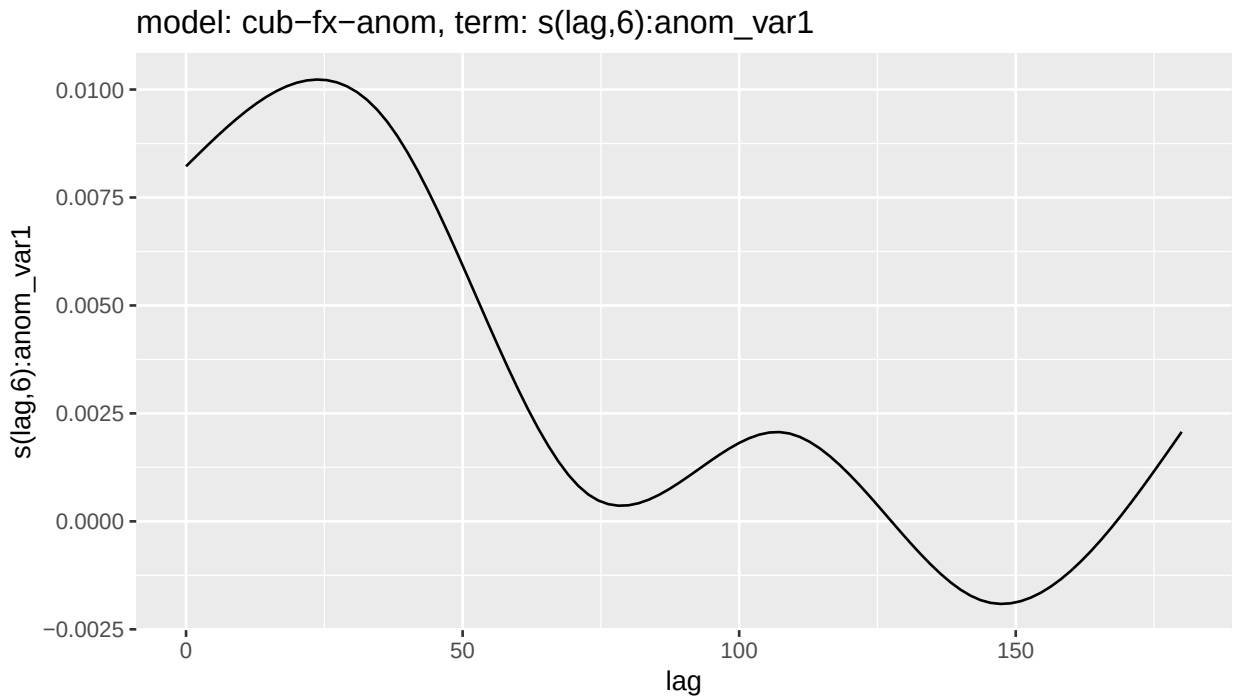
Risk for 2018−09−09 to 09−15

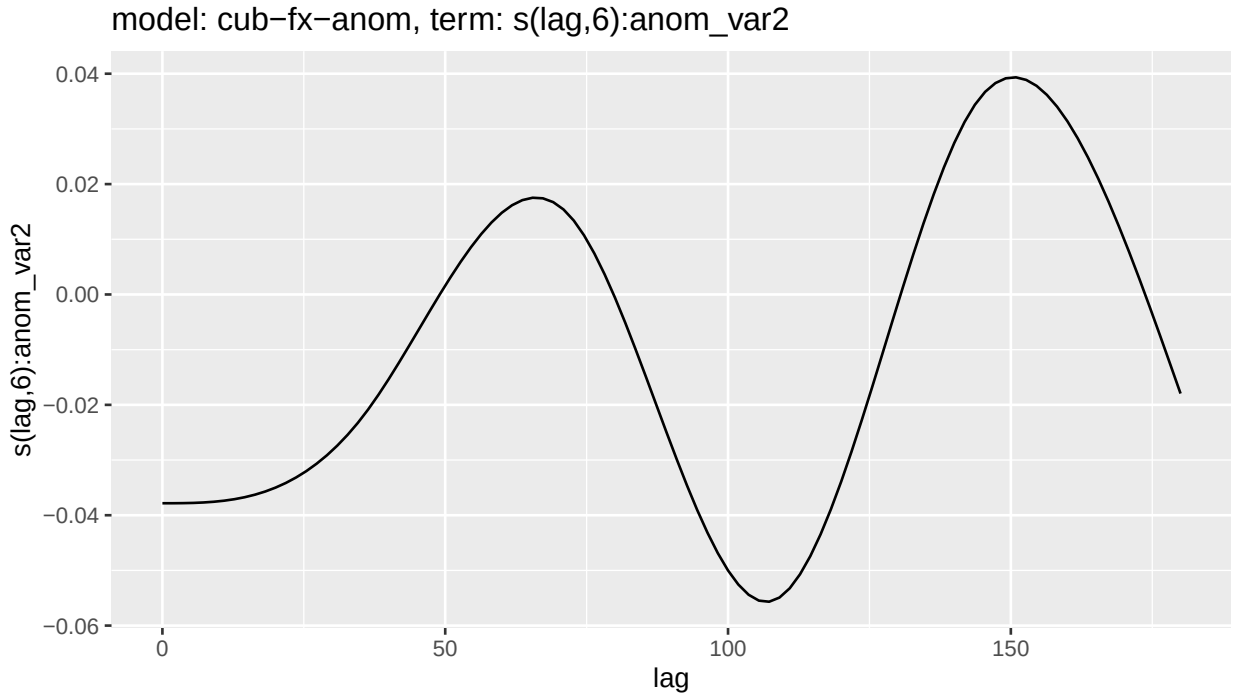
About average

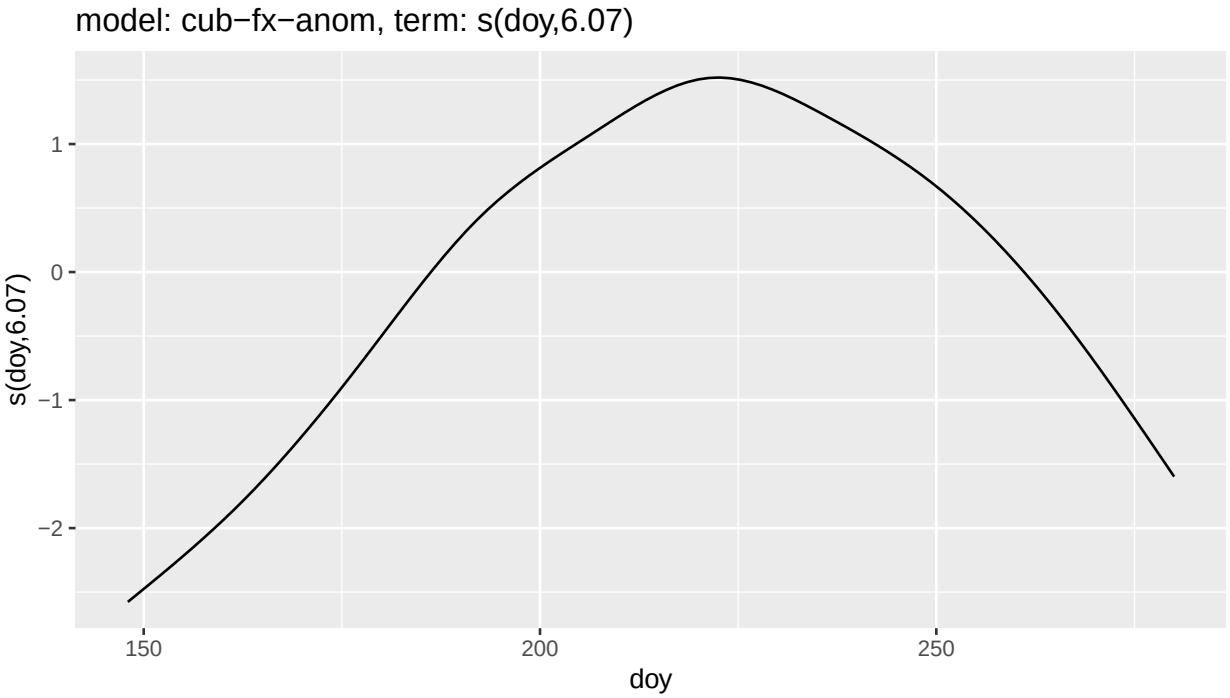
Estimated dependence functions

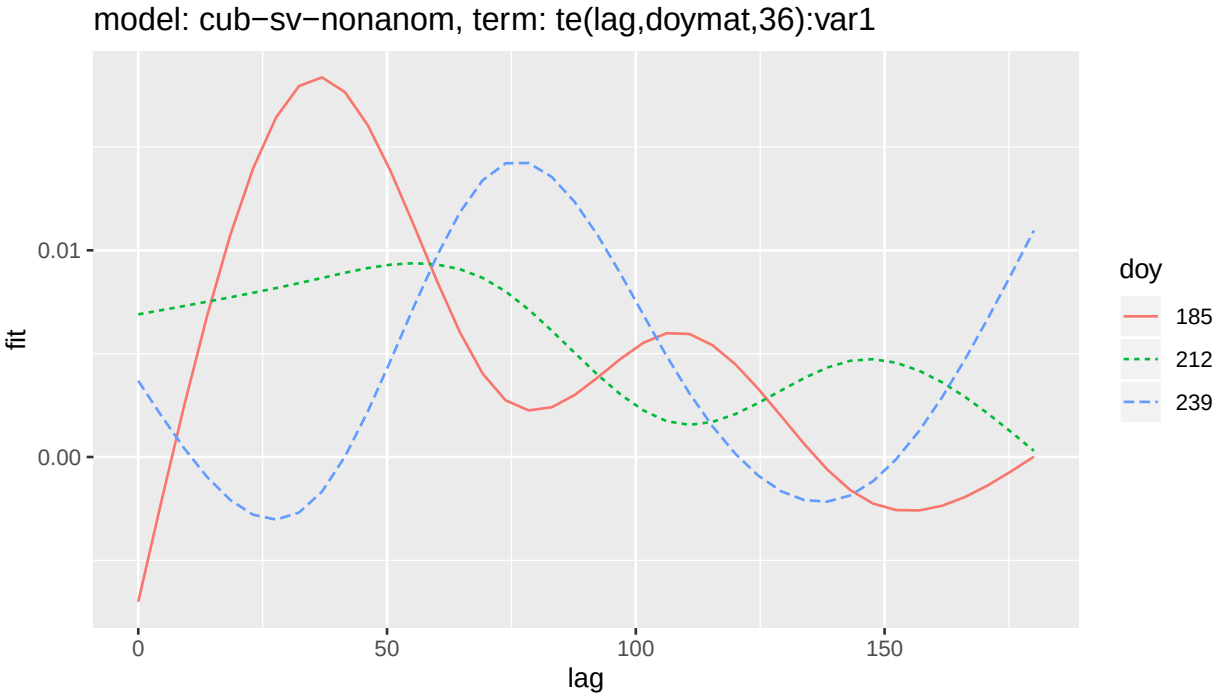


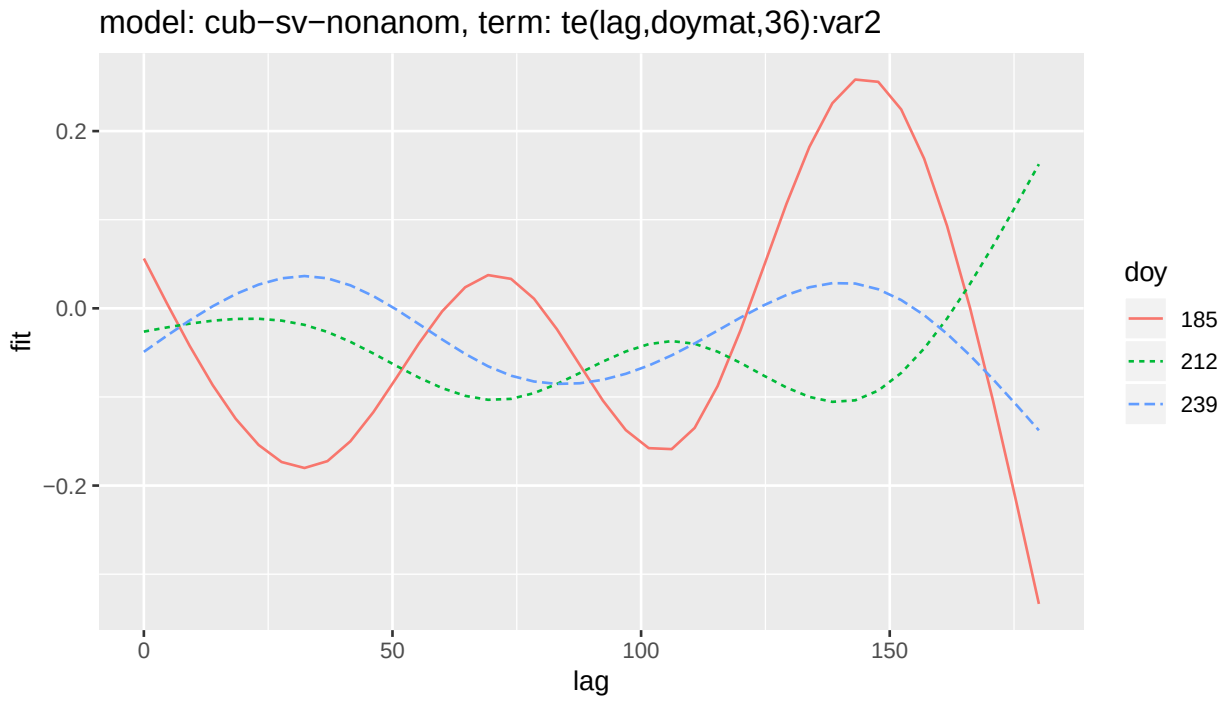


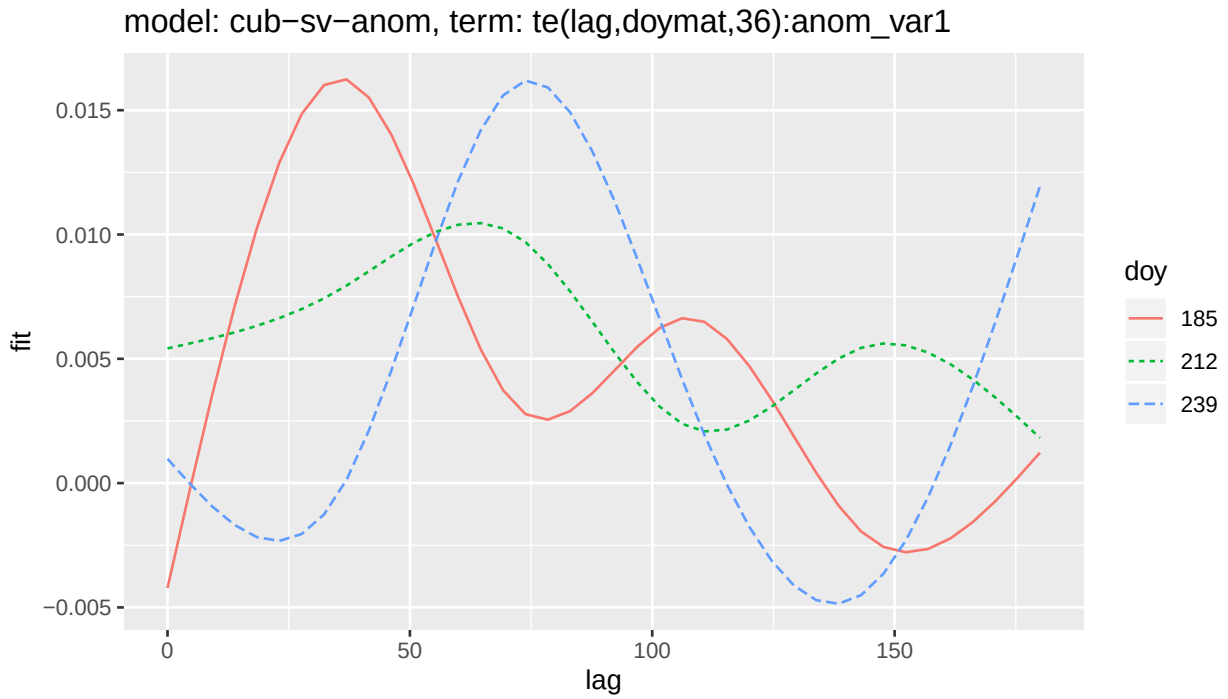

model: cub−fx−anom, term: s(lag,6):anom_var1
s(lag,6):anom_var1
0.0100
0.0075
0.0050
0.0025
0.0000
−0.0025
0
50
100
150
lag

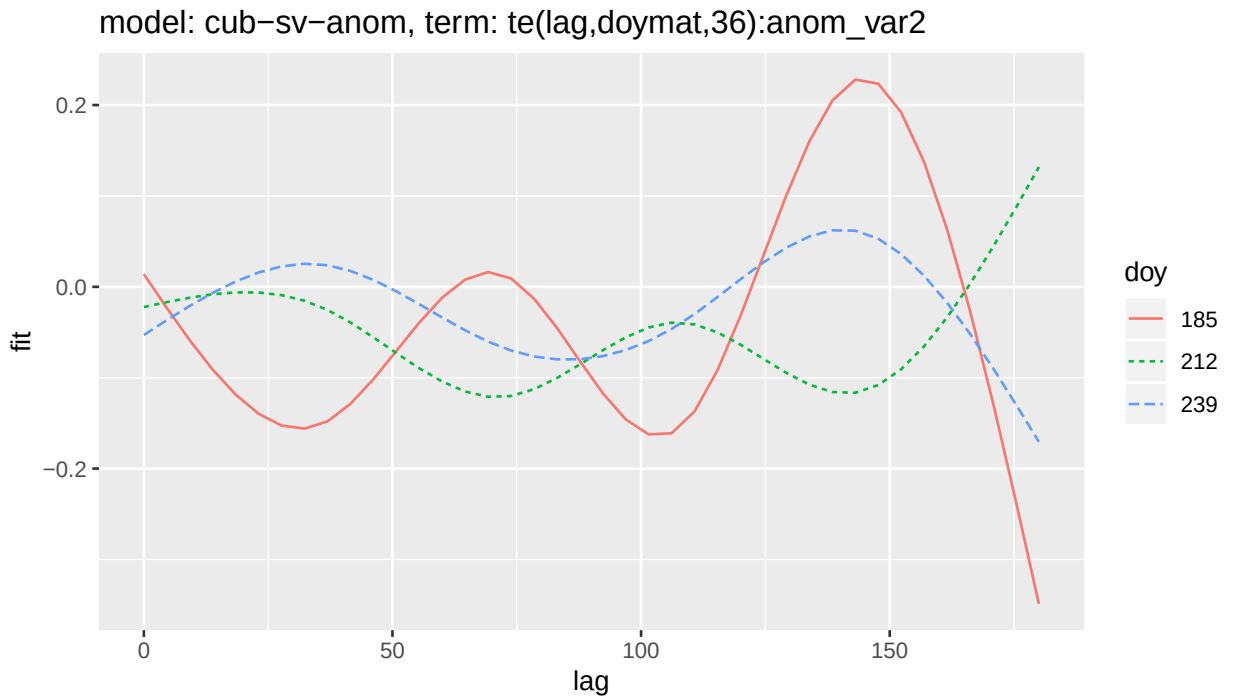

model: cub−fx−anom, term: s(lag,6):anom_var2
s(lag,6):anom_var2
0.04
0.02
0.00
−0.02
−0.04
−0.06
0
50
100
150
lag

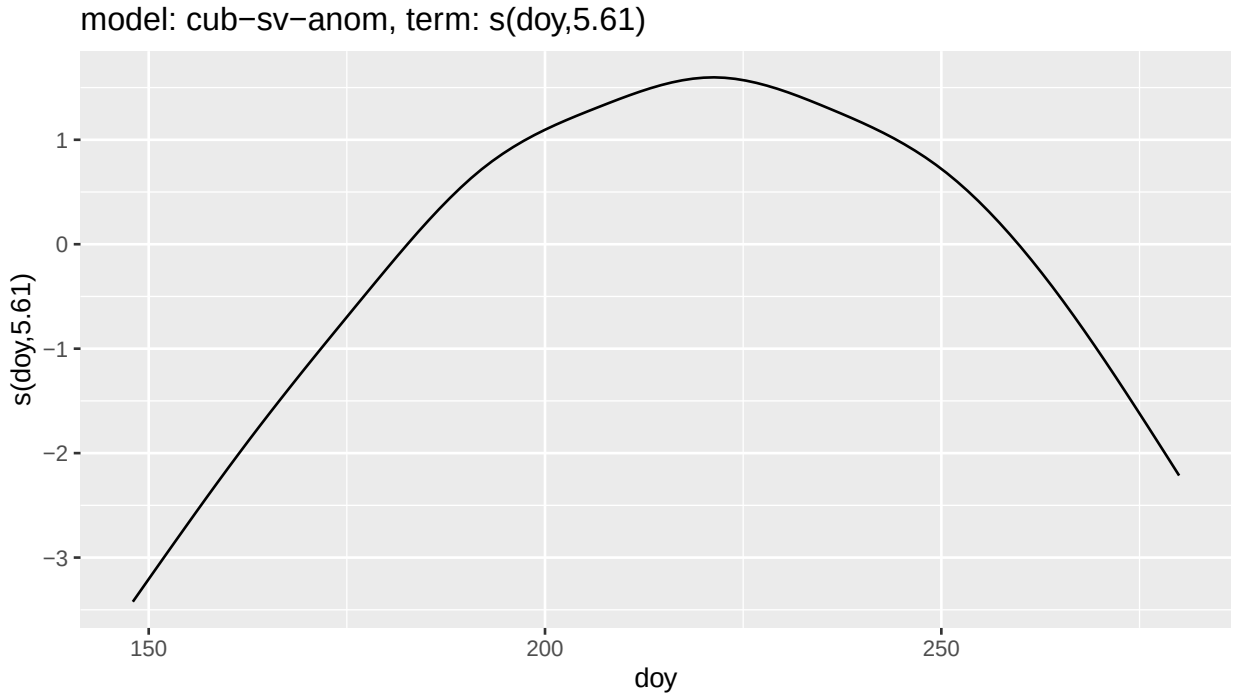

model: cub−fx−anom, term: s(doy,6.07)
s(doy,6.07)
1
0
−1
−2
150
200
250
doy

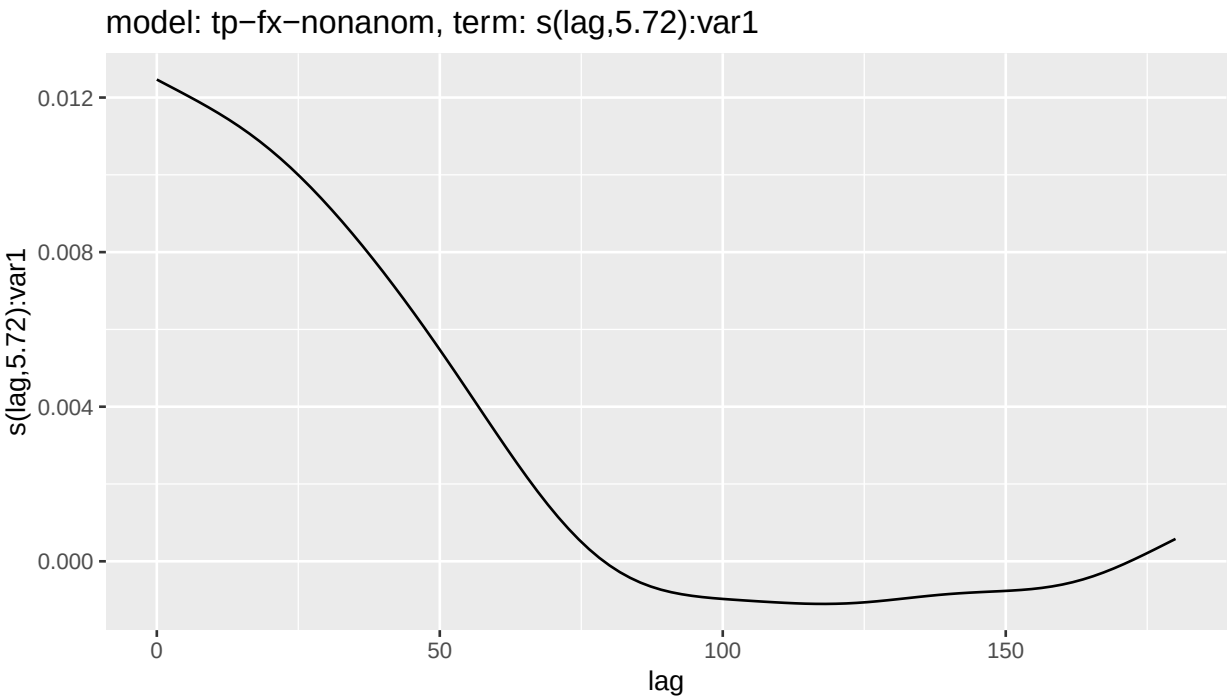

model: cub−sv−nonanom, term: te(lag,doymat,36):var1
fit
0.01
0.00
0
50
100
150
lag
doy
185
212
239

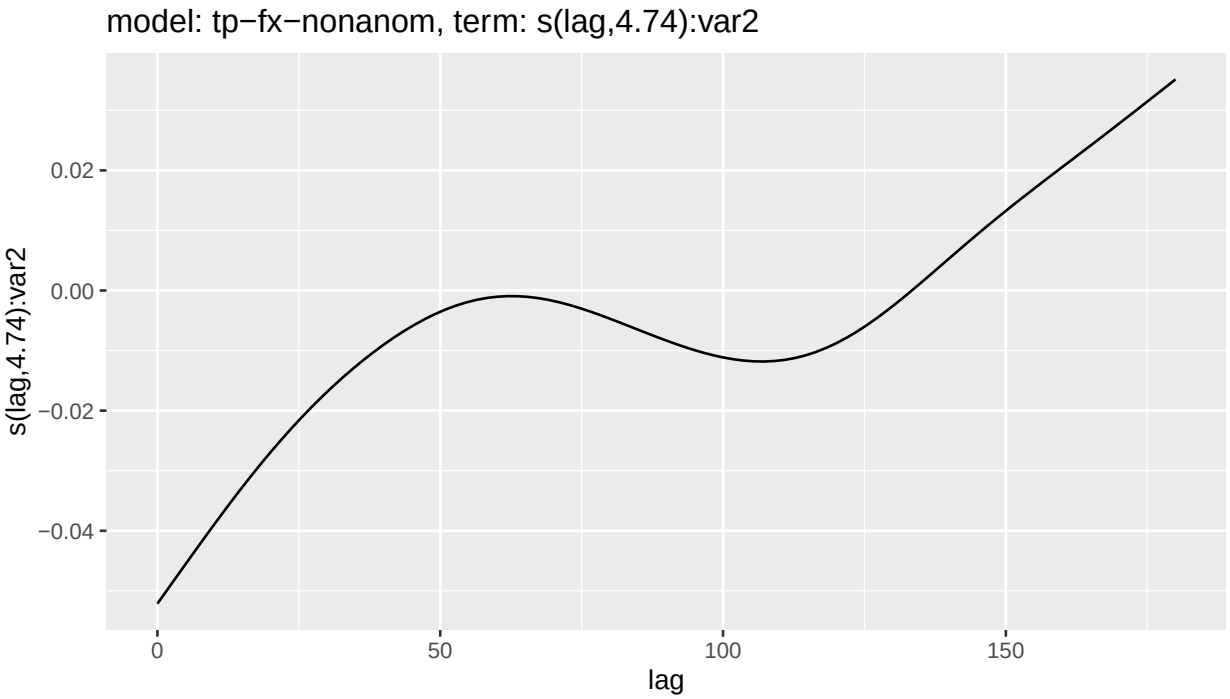

model: cub−sv−nonanom, term: te(lag,doymat,36):var2
0.2
0.0
−0.2
fit
0
50
100
150
lag
doy
185
212
239

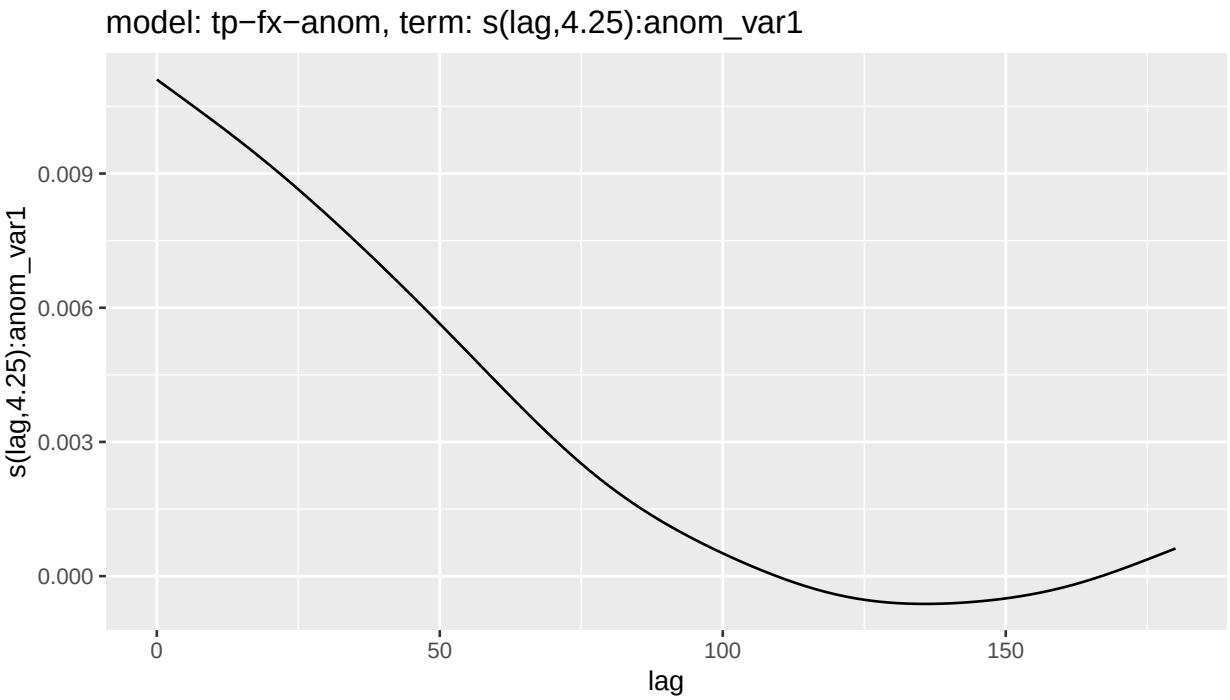

model: cub−sv−anom, term: te(lag,doymat,36):anom_var1
0.015
0.010
0.005
0.000
−0.005
fit
0
50
100
150
lag
doy
185
212
239

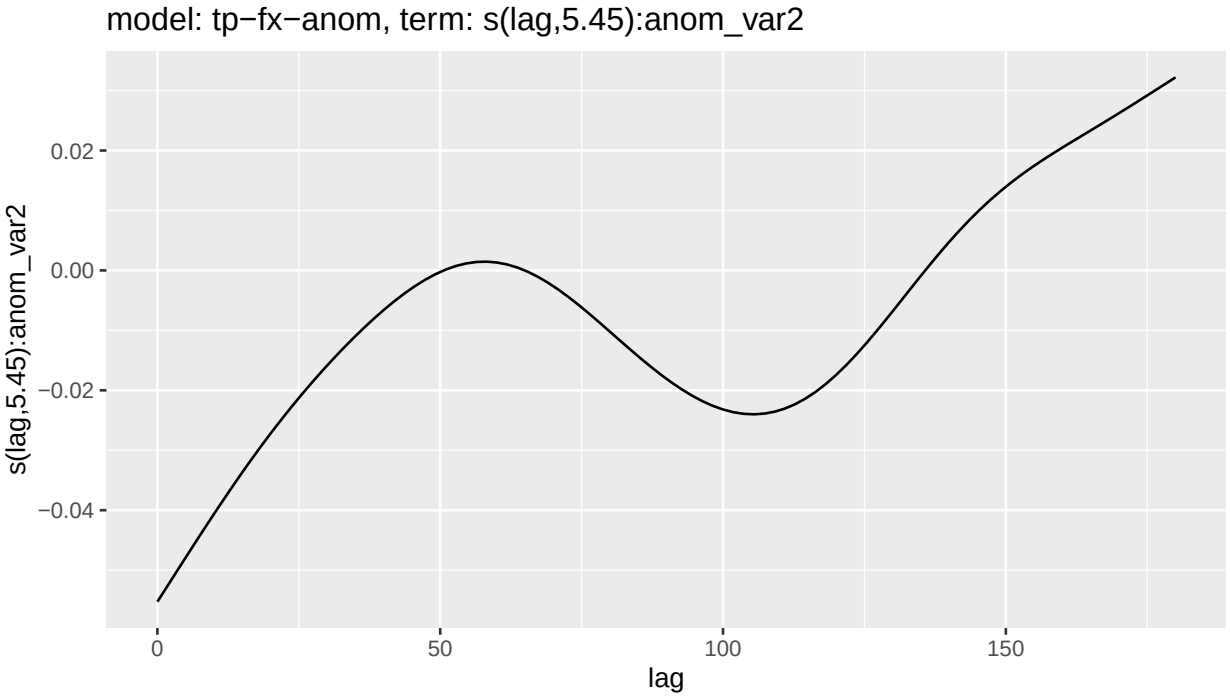

model: cub−sv−anom, term: te(lag,doymat,36):anom_var2
0.2
0.0
−0.2
fit
0
50
100
150
lag
doy
185
212
239

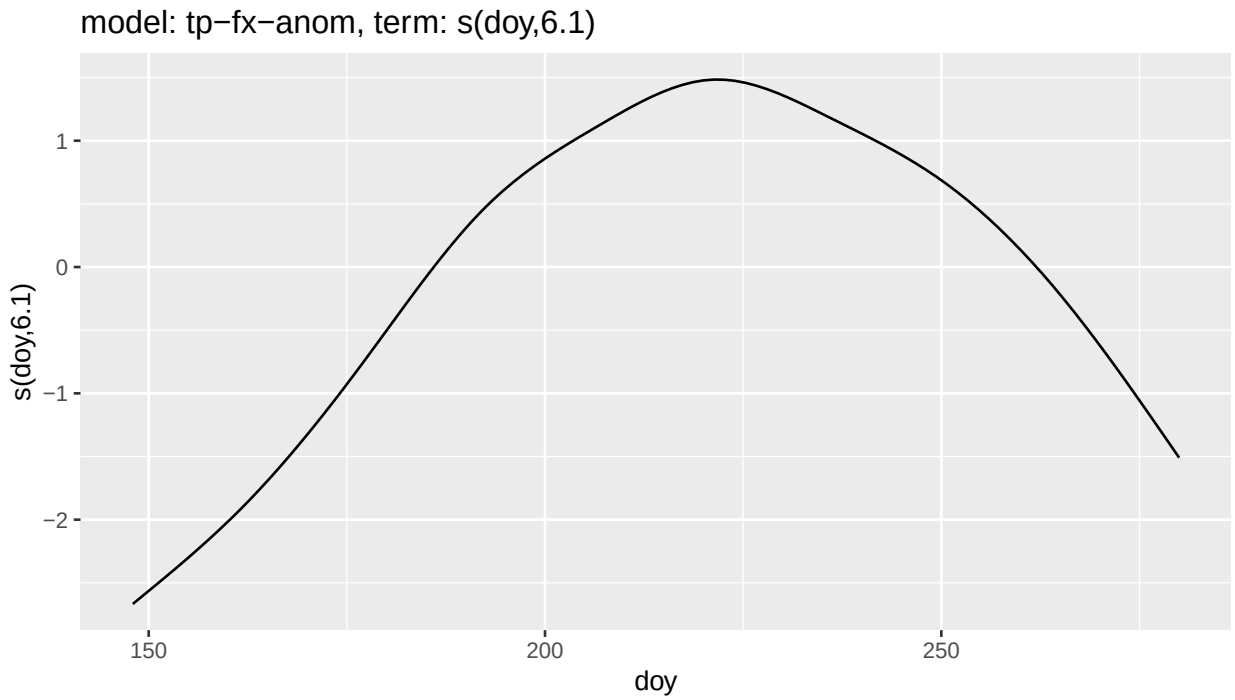

model: cub−sv−anom, term: s(doy,5.61)
1
0
−1
−2
−3
s(doy,5.61)
150
200
250
doy

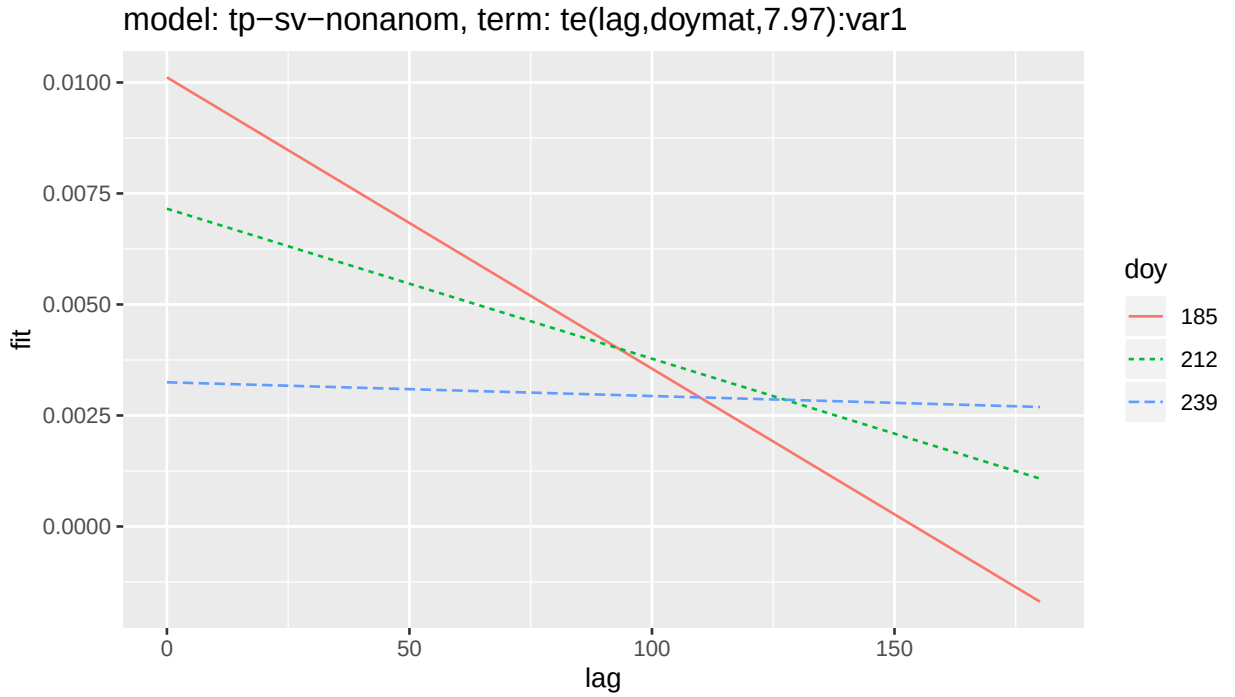

model: tp−fx−nonanom, term: s(lag,5.72):var1
0.012
0.008
0.004
0.000
s(lag,5.72):var1
0
50
100
150
lag

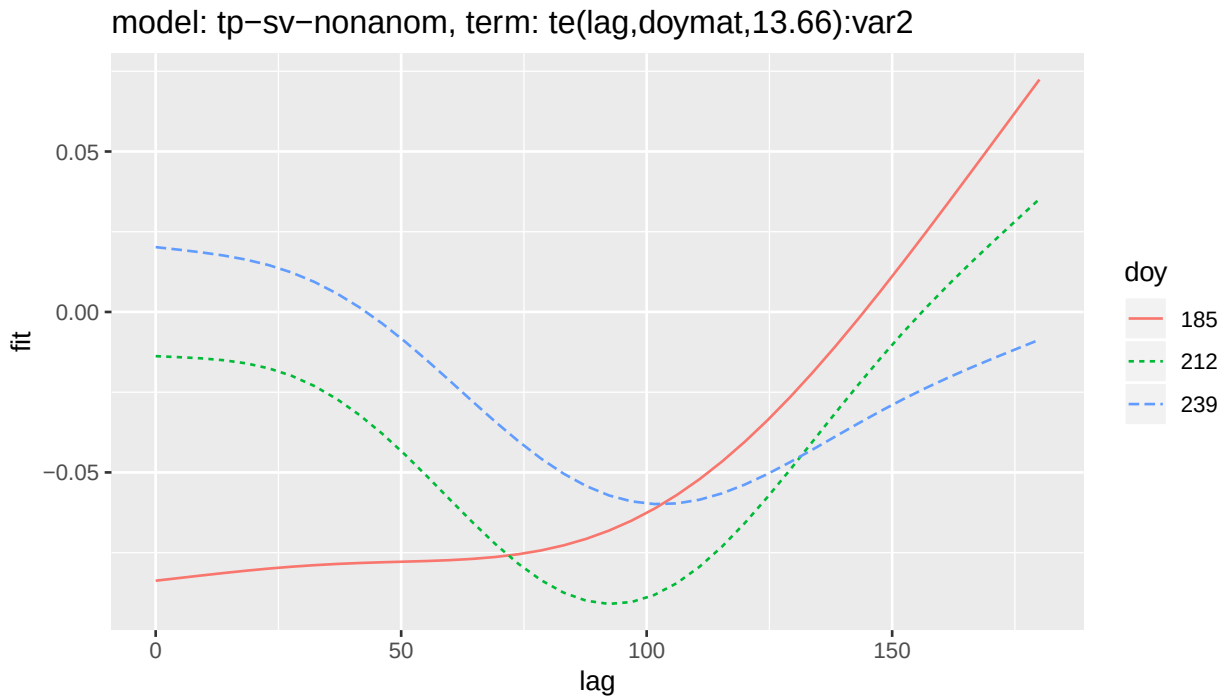

model: tp−fx−nonanom, term: s(lag,4.74):var2
0.02
0.00
−0.02
−0.04
s(lag,4.74):var2
0
50
100
150
lag

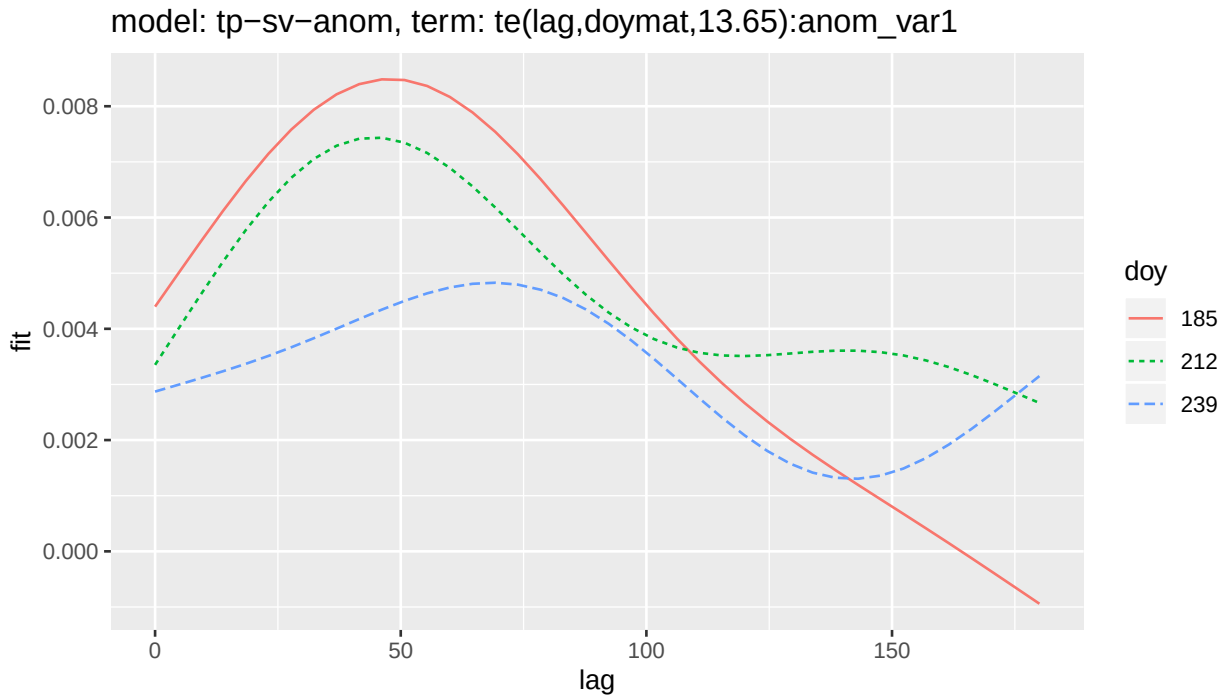

model: tp−fx−anom, term: s(lag,4.25):anom_var1
s(lag,4.25):anom_var1
0.009
0.006
0.003
0.000
0
50
100
150
lag

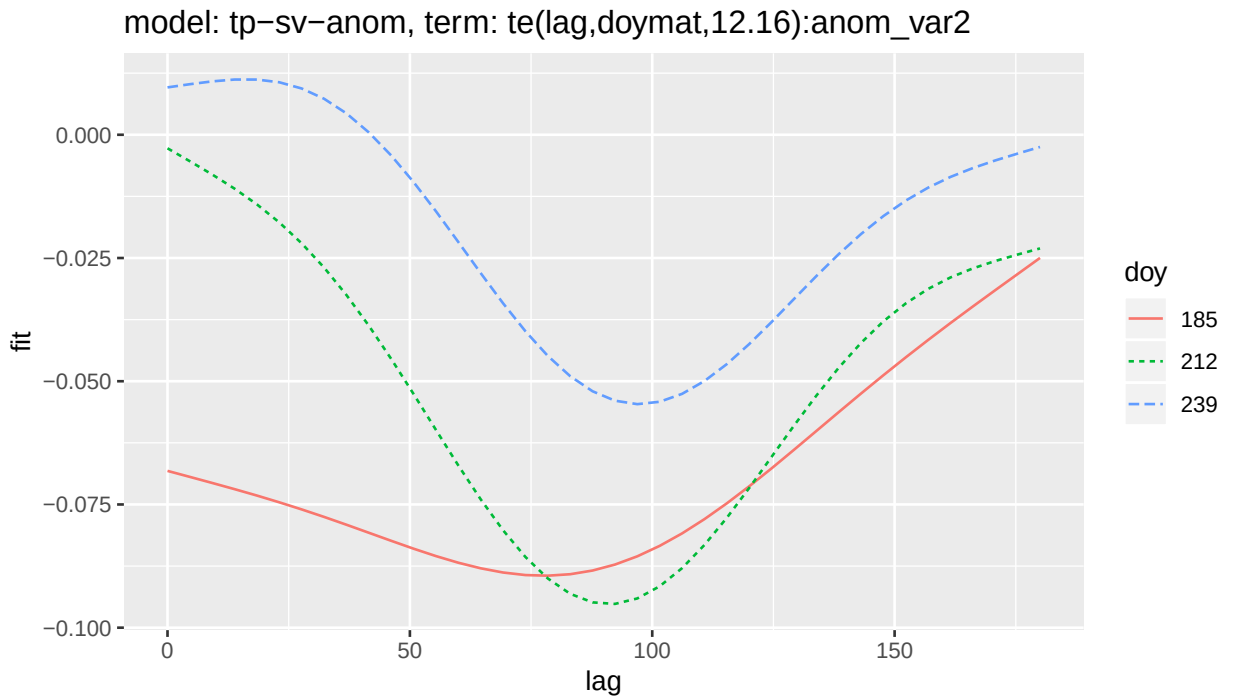

model: tp−fx−anom, term: s(lag,5.45):anom_var2
s(lag,5.45):anom_var2
0.02
0.00
−0.02
−0.04
0
50
100
150
lag

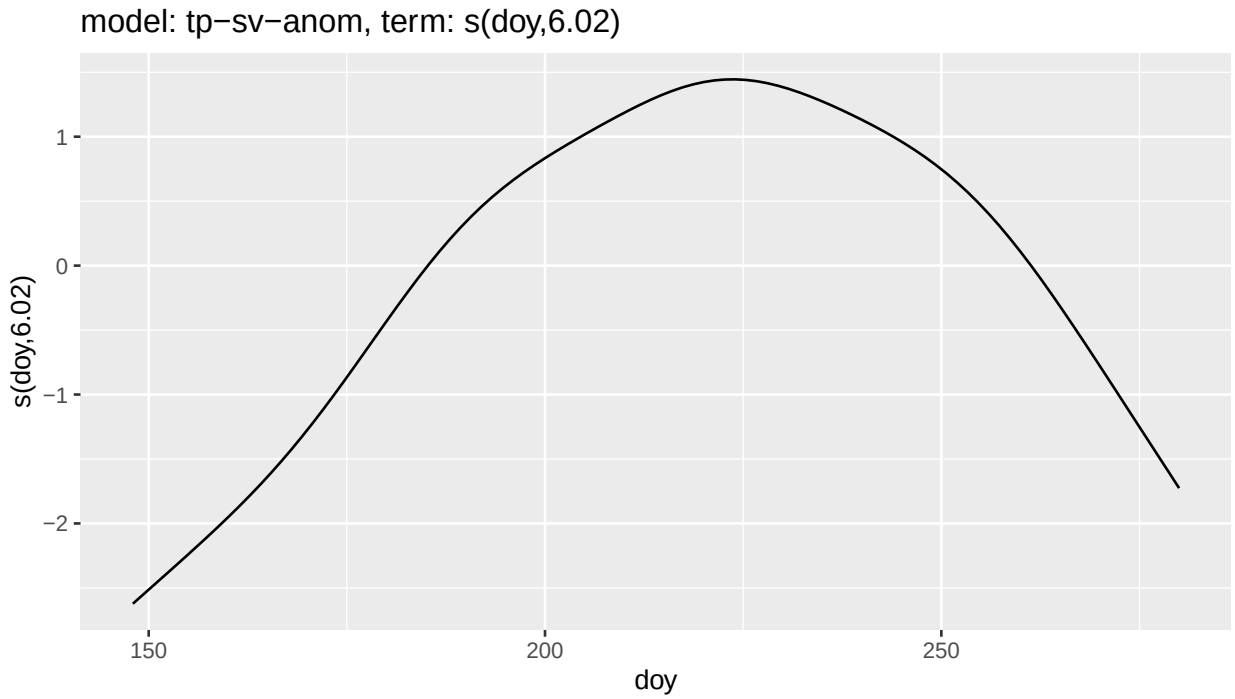

model: tp−fx−anom, term: s(doy,6.1)
s(doy,6.1)
1
0
−1
−2
150
200
250
doy


model: tp−sv−nonanom, term: te(lag,doymat,7.97):var1
fit
0.0100
0.0075
0.0050
0.0025
0.0000
0
50
100
150
lag
doy
185
212
239


model: tp−sv−nonanom, term: te(lag,doymat,13.66):var2
0.05
0.00
−0.05
fit
0
50
100
150
lag
doy
185
212
239


model: tp−sv−anom, term: te(lag,doymat,13.65):anom_var1
0.008
0.006
0.004
0.002
0.000
fit
0
50
100
150
lag
doy
185
212
239


model: tp−sv−anom, term: te(lag,doymat,12.16):anom_var2
0.000
−0.025
−0.050
−0.075
−0.100
fit
0
50
100
150
lag
doy
185
212
239


model: tp−sv−anom, term: s(doy,6.02)
1
0
−1
−2
s(doy,6.02)
150
200
250
doy

District names

It is possible that district names in the human, mosquito, or weather data might disagree with the names found in the district shapefile. We have chosen the shapefile to unify all of the various data sources, so the human, mosquito, and weather data files should be updated to match these. Typically, check the TIGER shapefiles or census for standardized names.

There are 66 districts in the shapefile. This is the maximum number of districts ArboMAP will recognize in the human, mosquito, or weather data.

There are 66 districts in the human WNV data. If some districts never reported cases, this number may be less than the number of districts in the shapefile. If there are more, there are likely misspellings in the human data file. Human districts not found in the district shapefile: none.

There are 73 districts in the mosquito WNV data. If some districts never reported mosquito data, this number may be less than the number of districts in the shapefile. Mosquito districts not found in the district shapefile: bon, washabaugh, unknown, , ogalalakota, doh, doh-withers, mitchell-carlson.

There are 66 districts in the mosquito WNV data. If these data were downloaded with the GEE app, they should match the shapefile districts exactly. Weather districts not found in the district shapefile: none.