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

Summary of Model Outputs (v2.0)

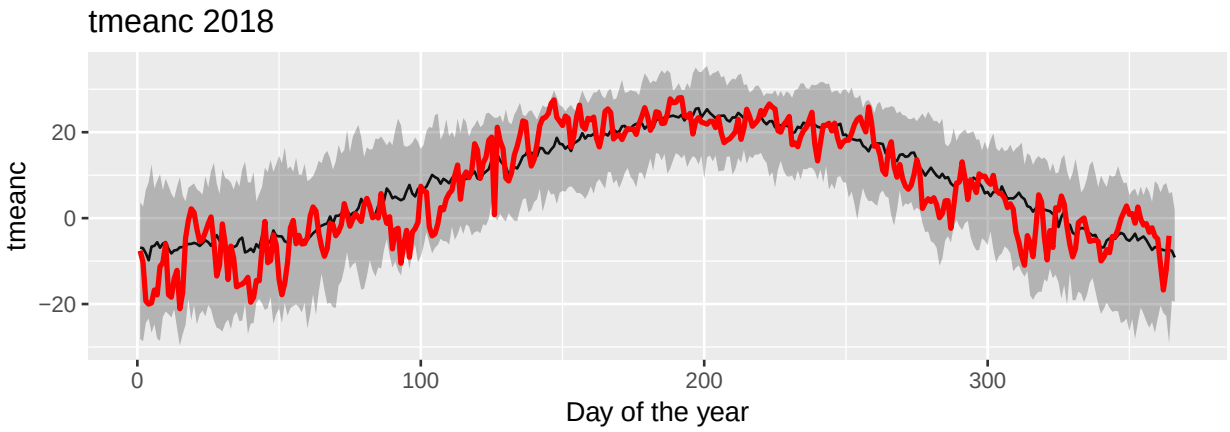
Justin K. Davis and Michael C. Wimberly
(justinkdavis@ou.edu, mcwimberly@ou.edu)
Geography and Environmental Sustainability, University of Oklahoma

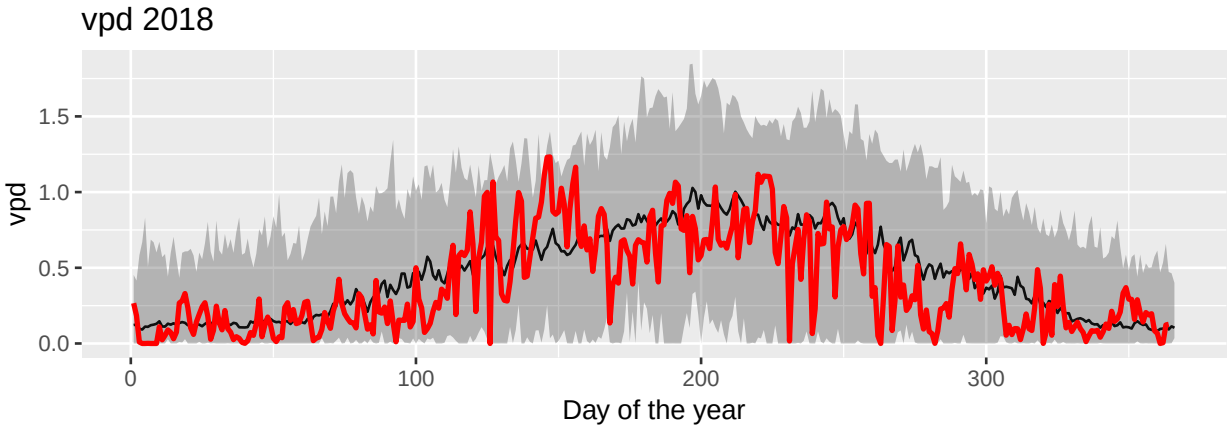
Updated October 25, 2019

Data used for predictions

Weather data

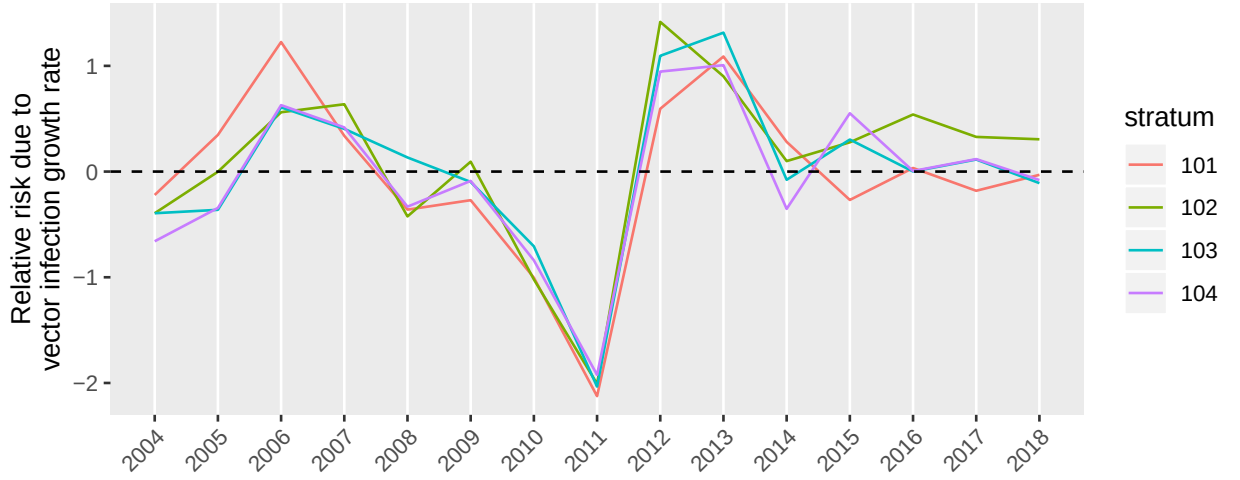
Weather data from the gridMET data set range from 2000-01-01 to 2018-12-30. Below are graphs of statewide daily averages of tmeanc and vpd. Observations for the current year are in red. Black is the medium from all other years, and the grey band indicates the max/min ever observed.



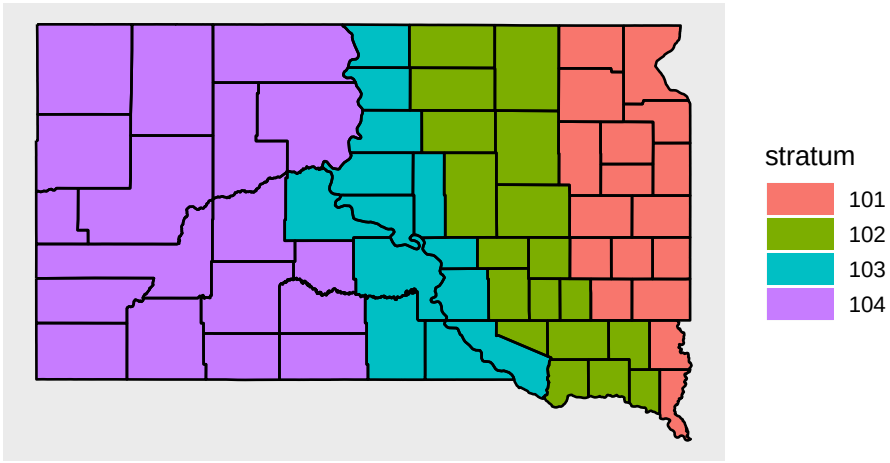


Vector infection data

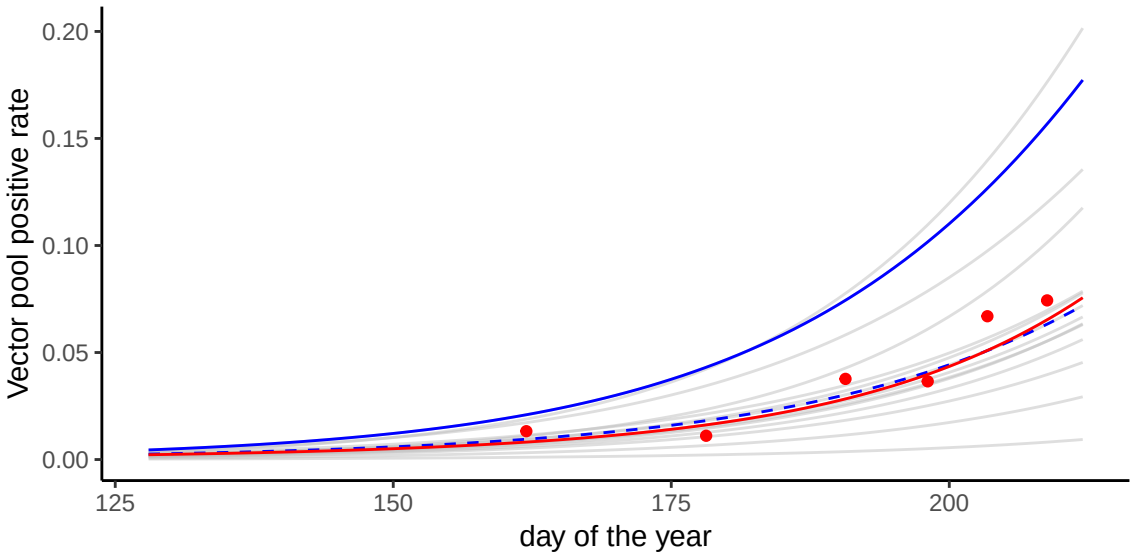
There are 16274 samples in the vector testing database. For 2018, there are 1697 tested samples, with 66 (3.9%) positive. The estimated risk of human infection due to the early-season vector infection growth rate is shown below. Higher means that the pathogen is spreading more rapidly among vectors, 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.



State stratification map



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.



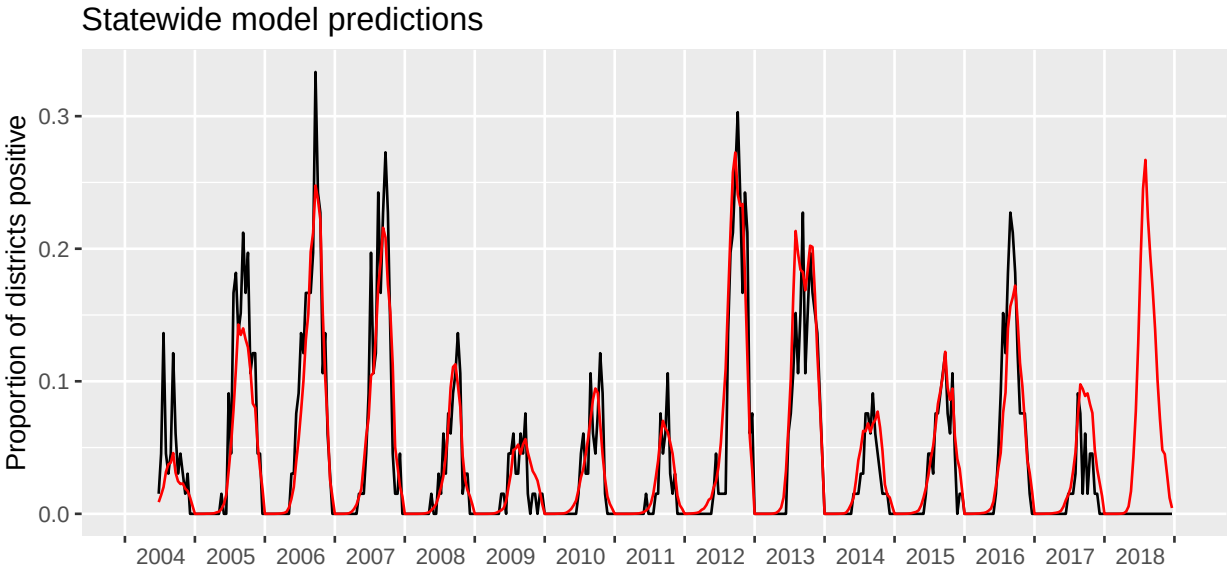
Human data

The predictive model of human cases was calibrated using 1189 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, 62% of a year's cases occur before the end of this week in any given year.

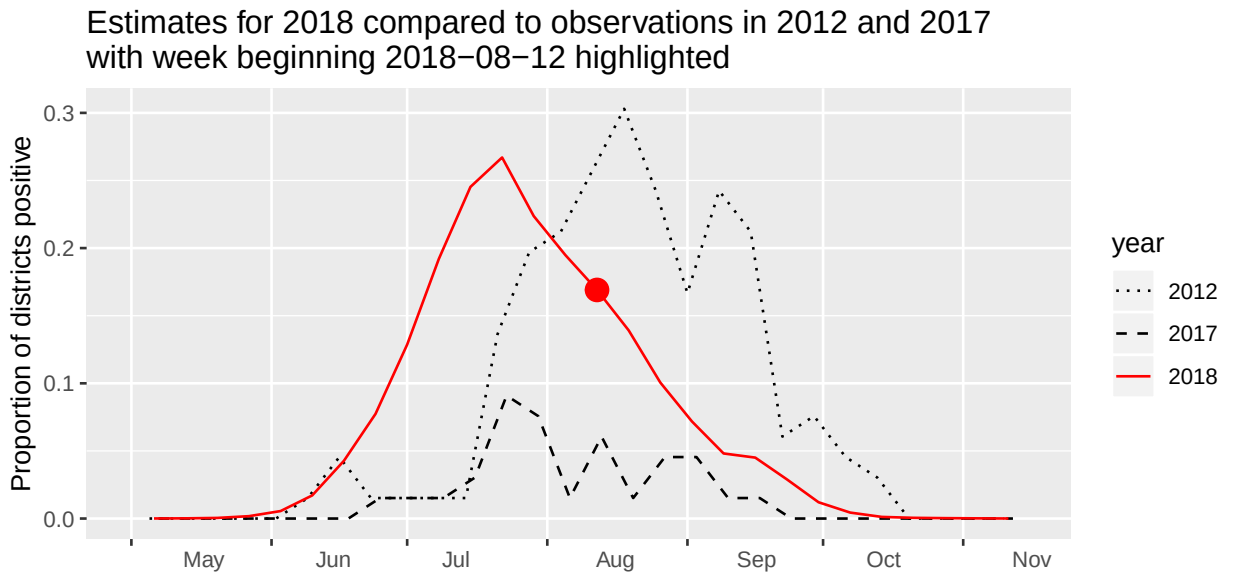
Model results

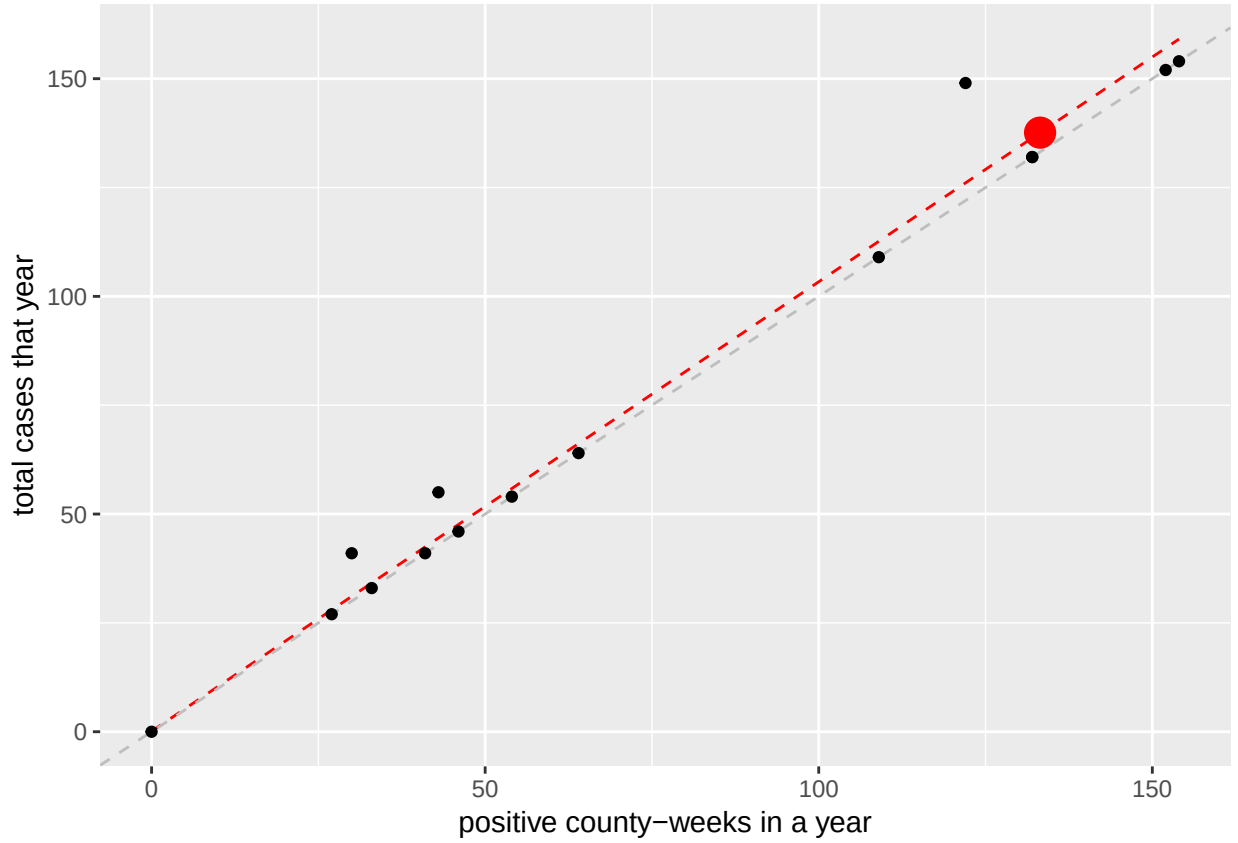
Statewide trends

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.



Predictions for 2018 are compared to observations in 2012 and 2017. We expect 11.2 districts to have at least one human case between Sunday 2018-08-12 and Saturday 2018-08-18.



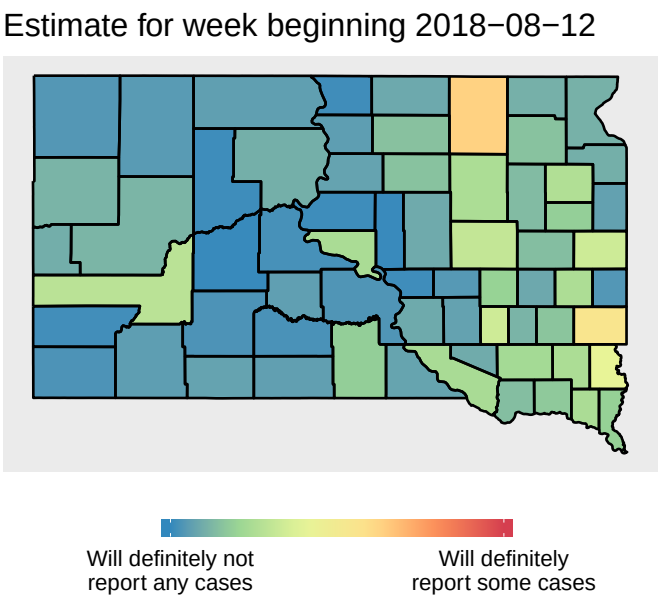


We expect that there will be 133.2 positive district-weeks this year. Since ArboMAP models positive district-weeks rather than total cases, we can only make a rough estimate based on the relationship of cases to district-weeks. If historical patterns hold (dashed red line), we expect there to be 137.6 cases this year (red

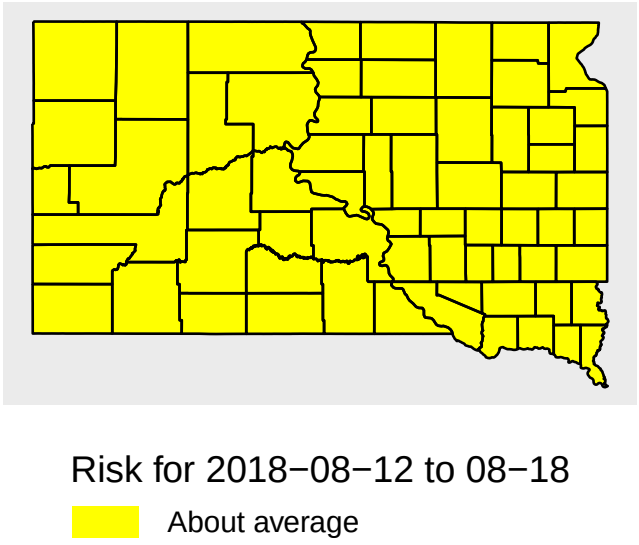
dot). The grey line indicates the case count if there were only one case per positive district-week. Especially in the height of the season, some districts that have one case that week will have more than one case.

Results for 2018-08-12 to 08-18

We visualize the raw estimated risk for 2018-08-15 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.

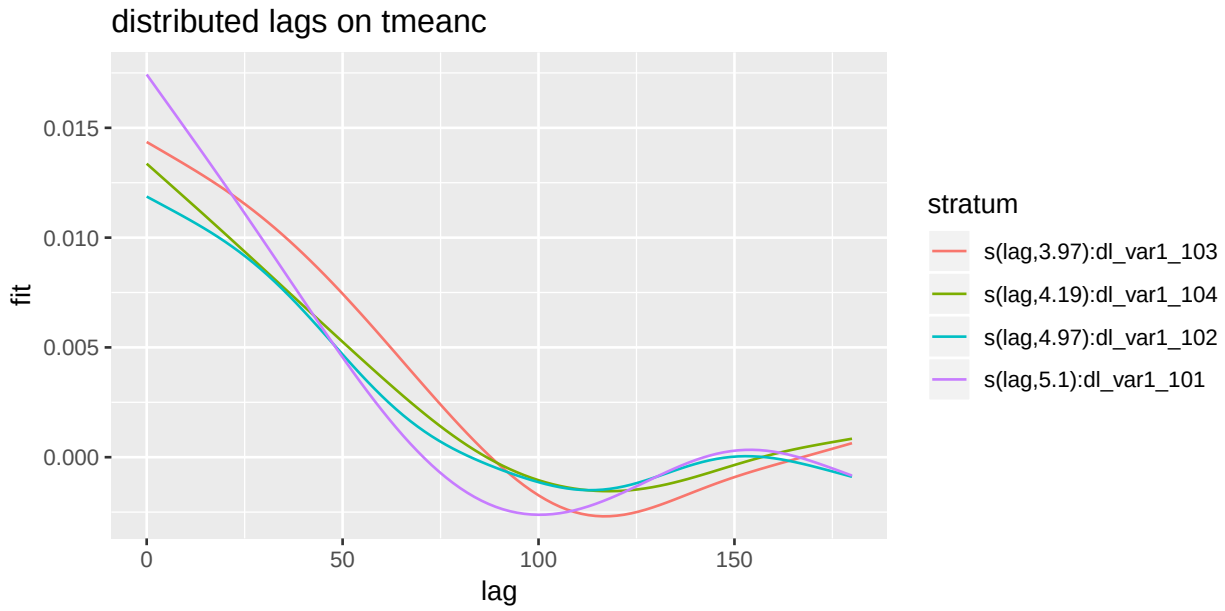


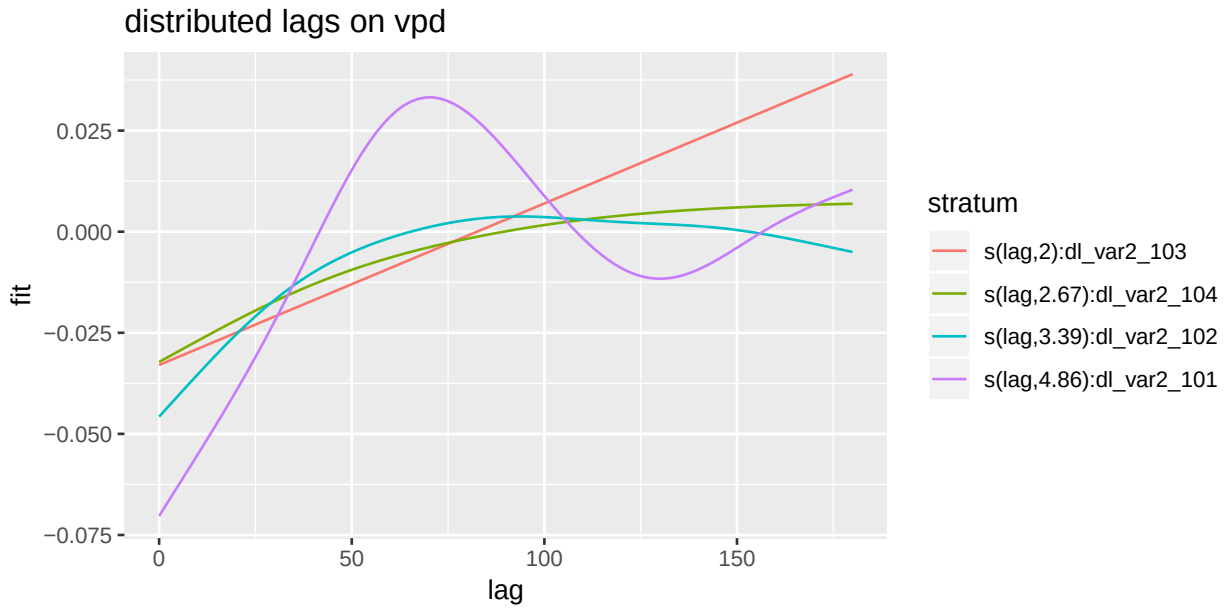
This map indicates whether probabilities reported in the previous map are higher (red) than average, lower (blue) than average, or right about normal (yellow) compared to the same week in previous years.



Dependence on environmental data.

Below, we visualize the distributed lags by stratum. If a line is above (below) zero at some lag, this means risk today depends positively (negatively) on the environmental data at that date; e.g. precipitation a week ago (lag = 7) might raise risk today.





The following two graphs show the influences of the two environmental covariates on estimated risk. The current year (red) is compared to all others (black). If the line is higher, this indicates the variable in the past has a positive influence on risk estimated for that date.