Update 4/19

2022-04-19

Absense-Presence Model

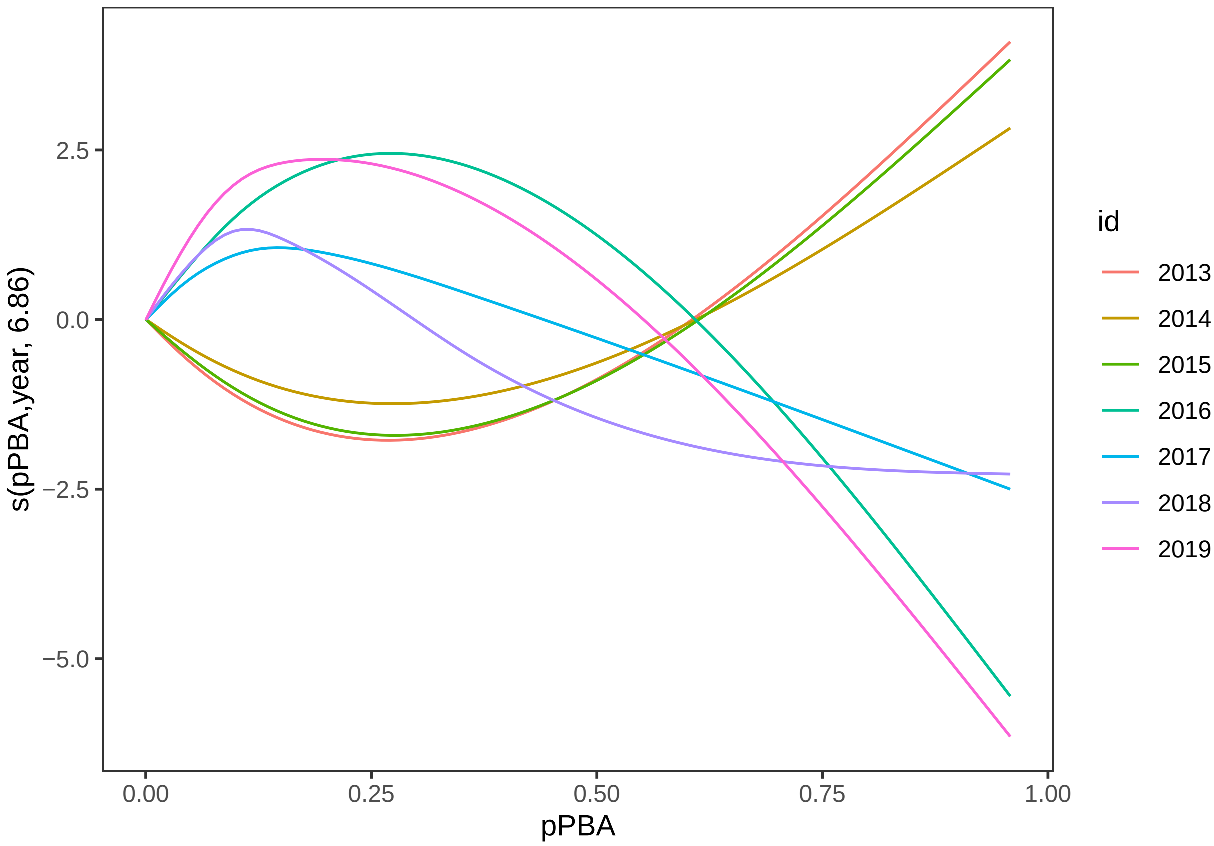
For my binary model, after model selections, the best-performing model is one that fits a random smooth of time-dependent variables (e.g. temperature, precipitation) for each year, but lets the random smooths have a shared panelty term (same wiggliness). This is achieved using the `bs = "fs"` option.

```
##
## Family: binomial
## Link function: logit
##
## Formula:
## wf ~ s(ppt, year, bs = "fs") + s(vpma, year, bs = "fs") + s(perc_cloud) +
##     s(wind_speed) + s(pop_den) + s(pFederal) + s(pState) + s(dRoad) +
##     s(dTrail) + s(dTemp, year, bs = "fs") + s(dPpt, year, bs = "fs") +
##     s(dVPM, year, bs = "fs") + s(dPowerLine) + s(tree) + s(herb) +
##     s(slope) + s(aspect) + s(CENTROID_X, CENTROID_Y, year, bs = "fs") +
##     s(temp, year, bs = "fs") + s(pPBA, year, bs = "fs") + s(pb1Prev,
##     year, bs = "fs") + s(pb2Prev, year, bs = "fs") + s(pb3Prev,
##     year, bs = "fs") + s(ppt) + s(vpma) + s(temp) + s(dTemp) +
##     s(dPpt) + s(dVPM)
##
## Parametric coefficients:
##             Estimate Std. Error z value Pr(>|z|)
## (Intercept)  -7.9871     0.5151  -15.51   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Approximate significance of smooth terms:
##                                   edf  Ref.df   Chi.sq  p-value
## s(ppt,year)                    29.556  68.000  220.920  < 2e-16 ***
## s(vpma,year)                   28.689  68.000  175.312  < 2e-16 ***
## s(perc_cloud)                   5.895   6.951   23.152 0.002352 **
## s(wind_speed)                   2.981   3.832    9.711 0.040515 *
## s(pop_den)                      1.365   1.647    0.561 0.555875
## s(pFederal)                     3.423   4.205  114.261  < 2e-16 ***
## s(pState)                       1.883   2.339   20.889 7.19e-05 ***
## s(dRoad)                        1.679   2.099    1.425 0.553553
## s(dTrail)                       4.870   5.956   19.712 0.003185 **
## s(dTemp,year)                  18.020  68.000   78.993  < 2e-16 ***
## s(dPpt,year)                   30.096  67.000  300.206  < 2e-16 ***
## s(dVPM,year)                    5.847  68.000   12.485 0.006749 **
## s(dPowerLine)                   2.514   3.212   12.798 0.006536 **
## s(tree)                         2.539   3.182   11.409 0.012306 *
## s(herb)                         6.101   7.254  268.711  < 2e-16 ***
```

```
## s(slope)                         7.437   8.383   57.243  < 2e-16 ***
## s(aspect)                        1.697   2.129    3.673 0.177381
## s(CENTROID_X,CENTROID_Y,year) 167.181 209.000 1955.565  < 2e-16 ***
## s(temp,year)                    28.895  68.000  201.628  < 2e-16 ***
## s(pPBA,year)                     6.857  66.000   35.618 3.62e-06 ***
## s(pb1Prev,year)                  2.391  67.000    3.329 0.178748
## s(pb2Prev,year)                  4.027  67.000    8.572 0.027721 *
## s(pb3Prev,year)                  4.536  67.000   14.906 0.001304 **
## s(ppt)                           4.291   4.825    4.915 0.322947
## s(vpma)                          4.800   5.301    4.708 0.410015
## s(temp)                          2.274   2.594    1.141 0.732734
## s(dTemp)                         1.000   1.001    0.911 0.340128
## s(dPpt)                          1.661   1.736    0.478 0.645066
## s(dVPM)                          7.157   8.101   26.944 0.000656 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## R-sq.(adj) =  0.253   Deviance explained = 41.8%
## fREML = 1.4146e+05  Scale est. = 1         n = 99141
```

This model has a deviance explained of over 40%, which is doulbe the value from the model we discussed last time. Some notes on the variable names: `dTemp`, `dPpt`, and `dVPM` refer to the deviation of, say, temperature of a particular year, from the average temperature of that grid cell across all years. `pbiPrev` is the percent prescribe burned from `i` years prior. By the way, I tried transforming aspect using the folded aspect structure you showed me, but the result is quite similar with no significance. I also tried using a cyclic basis on the untransformed aspect, and the result was still no significance.

The main concern is how to intepret the model output. The most important variable, the proportion of a gridcell burned, looks like this:



The smooths seem to form mirror images of each other. I think this could still be interesting because the ones that showed a decrease in ignition probability are the ones that come from the most recent years. I think we could say something along the lines of: for most recent years, you need to burn a suffcient amount of the area in order to effectively prevent a wildfire. But I'm not sure how to intepret the results from less recent years.

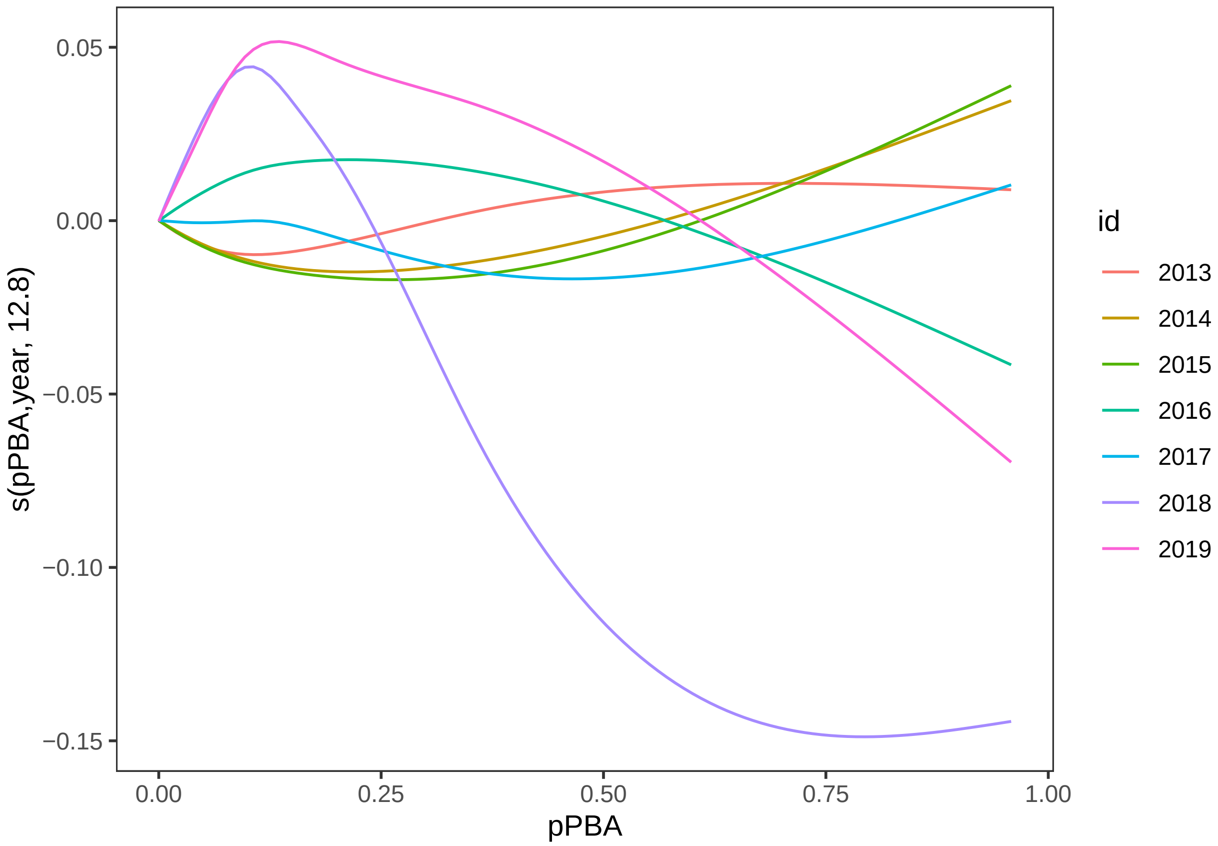
Continuous Model

Using the same model, I fitted the percentage of a grid cell burned by wildfires instead of just the absence-presence of a wildfire. The result is more chaotic.

```
##
## Family: gaussian
## Link function: identity
##
## Formula:
## pWFA ~ s(ppt, year, bs = "fs") + s(vpma, year, bs = "fs") + s(perc_cloud) +
##     s(wind_speed) + s(pop_den) + s(pFederal) + s(pState) + s(dRoad) +
##     s(dTrail) + s(dTemp, year, bs = "fs") + s(dPpt, year, bs = "fs") +
##     s(dVPM, year, bs = "fs") + s(dPowerLine) + s(tree) + s(herb) +
##     s(slope) + s(aspect) + s(CENTROID_X, CENTROID_Y, year, bs = "fs") +
##     s(temp, year, bs = "fs") + s(pPBA, year, bs = "fs") + s(pb1Prev,
##     year, bs = "fs") + s(pb2Prev, year, bs = "fs") + s(pb3Prev,
##     year, bs = "fs") + s(ppt) + s(vpma) + s(temp) + s(dTemp) +
##     s(dPpt) + s(dVPM)
##
```

```
## Parametric coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.012380   0.006363   1.946   0.0517 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Approximate significance of smooth terms:
##                                    edf  Ref.df      F  p-value
## s(ppt,year)                   51.83651  68.000  9.941  < 2e-16 ***
## s(vpma,year)                  43.96998  68.000  4.070  < 2e-16 ***
## s(perc_cloud)                  6.94560   8.116  7.069  < 2e-16 ***
## s(wind_speed)                  2.59799   3.353  3.354 0.014912 *
## s(pop_den)                     4.74877   5.821  4.055 0.000501 ***
## s(pFederal)                    2.84078   3.492 49.173  < 2e-16 ***
## s(pState)                      1.00032   1.001 18.732 1.50e-05 ***
## s(dRoad)                       1.00031   1.001  0.036 0.849242
## s(dTrail)                      4.49956   5.543  4.782 0.000129 ***
## s(dTemp,year)                  4.48664  68.000  0.087 0.124336
## s(dPpt,year)                  35.42783  68.000 17.945  < 2e-16 ***
## s(dVPM,year)                  25.90046  68.000  1.563  < 2e-16 ***
## s(dPowerLine)                  3.86618   4.824  6.594 9.77e-06 ***
## s(tree)                        3.42242   4.258 15.594  < 2e-16 ***
## s(herb)                        4.68204   5.746 27.444  < 2e-16 ***
## s(slope)                       5.86913   7.081 14.609  < 2e-16 ***
## s(aspect)                      1.00056   1.001  0.106 0.745525
## s(CENTROID_X,CENTROID_Y,year) 193.68073 209.000 25.632  < 2e-16 ***
## s(temp,year)                  54.21714  68.000 10.612  < 2e-16 ***
## s(pPBA,year)                  12.80351  69.000  1.178  < 2e-16 ***
## s(pb1Prev,year)               10.09840  67.000  0.602 2.78e-06 ***
## s(pb2Prev,year)                2.87680  67.000  0.063 0.158069
## s(pb3Prev,year)                0.00369  67.000  0.000 0.066724 .
## s(ppt)                         2.03557   2.155  0.162 0.870761
## s(vpma)                        1.00057   1.001  0.337 0.561583
## s(temp)                        1.00051   1.001  1.337 0.247505
## s(dTemp)                       3.71171   4.758  3.401 0.006704 **
## s(dPpt)                        1.00009   1.000  0.127 0.721895
## s(dVPM)                        6.64977   7.080  0.824 0.566945
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## R-sq.(adj) =  0.122   Deviance explained = 12.7%
## fREML = -1.1105e+05  Scale est. = 0.0061067  n = 99141
```

This model has a really low deviance explained, which prompts me to think that there are some covariates that are missing from the data. I think I can spin this into a suggesting for future studies. But, again, the issue lies in the interpretability of the model. For percent prescribed burned, the smooths look like this:



For 2018, the effect is much larger than other years. And overall, the effect is quite inconsistent across years.