Observations on Pennsylvania Flood Data

Ajani Adovor

Introduction

Pennsylvania's topography makes many areas in the state prone to flooding. While there are measures in place to deal with flood events, their implementation varies, and areas face damages with differing levels of severity when major weather events occur. The goal of this research is to answer the question, "How Do Severe Flood Events Affect Certain Infrastructure and Populations in Pennsylvania?" This is a broad question that is not necessarliy qualifiable in all aspects. In order to answer this question specifically, a statistical analysis of Pennsylvania flood data will highlight key predictors that affect both bridges, critical pieces of Pennsylvania infrastructure, and populations within the state. An ideal answer to this question would highlight the association between certain variables and damages so that government officials can invest in safety measures in areas more prone to feel the damage of flooding. Given greater fluctuations in weather intensity due to climate change (include extra EDA here), identifying vulnerable infrastructure and populations is critical in order to prepare for catastrophic flooding events.

Data

The data set for this research comes from the United States Geological Survey and FEMA's flood data release for Pennsylvania. The data can be found here: Flood Data in Pennsylvania

The following are key variables being used from the data set. Please note that variables having "Var" in the name have been normalized (by dividing the initial variable by its maximum value) in the data set already so that they are easier to work with:

Control Variables:

YEARCON: The year in which a bridge was reconstructed (relabeled in this analysis as recon_yr)

DKSURFTYPE: The surface type of the bridge deck (relabeled in this analysis as dk_surf)

COUNTY_NAM: the name of the county in which the data point was recorded (relabeled in this analysis as county)

POPVar: The mean population count within a 1 km radius of the location

HURVar: Hurricane track density in kilometers per square kilometers

FLEVVar: Distance to the nearest point in the historical flood event data collection site data set in meters

DDVar: Number of disaster declarations per 1000 square kilometers

CLAVar: Number of flood-related insurance claims per square kilometer

INDAVar: Number of flood-related individual assistance applications per square kilometer

REPLVar: Number of flood-related repetitive loss records per square kilometer

Response Variables

ScoCritVar: Scour critical bridge indicator SCOURCRIT variable reclassified (1 = SCOURCRIT values of '0', '1', '2', or '3'; 0 = all other SCOURCRIT values)

SCBI Guide

The Scour Critical Bridge Indicator (SCBI) Code and Scour Assessment Rating (SAR) calculator developed by the Pennsylvania Department of Transportation (PennDOT) and the U.S. Geological Survey to identify Pennsylvania bridges with excessive scour conditions or a high potential for scour.

Scour: ***Bridge scour is the removal of sediment such as sand and gravel from around bridge abutments or piers. Hydrodynamic scour, caused by fast flowing water, can carve out scour holes, compromising the integrity of a structure. In the United States, bridge scour is one of the three main causes of bridge failure.*** **Scour Definition**

Use of the calculator will enable PennDOT bridge personnel to quickly calculate these scour indices if site conditions change, new bridges are constructed, or new information needs to be included.

SVIVar: Social Vulnerability Index measures the potential negative effects on communities caused by external stresses on human health. Such stresses include natural or human-caused disasters, or disease outbreaks. Reducing social vulnerability can decrease both human suffering and economic loss. SVI

Methodology

Bridge Infrastructure Analysis: The ideal way to determine infrastructural damage is by using an official government metric. Given the existence of the Scour Critical Bridge Indicator, this makes defining potential infrastructural damage straightforward. Thus, in terms of infrastructure, SCBI will be the response variable.

Because the

Population Analysis

Analysis Method Notes: Talk about pain here.

Results

```
Call:
glm(formula = ScoCritVar ~ recon_yr + dk_surf + POPVar + HURVar +
    FLEVVar + DDVar + (POPVar * DDVar), family = "binomial",
    data = fludd)

Coefficients:
               Estimate Std. Error z value Pr(>|z|)
(Intercept)    -1.74768    0.12546 -13.930  < 2e-16 ***
recon_yr1875   18.90266 2399.54472   0.008 0.993715
recon_yr1882   17.83980 2399.54472   0.007 0.994068
```

```
recon_yr1902  -15.46037 2399.54472  -0.006 0.994859
recon_yr1907   19.44434 2399.54472   0.008 0.993535
recon_yr1909  -14.21186 2399.54472  -0.006 0.995274
recon_yr1914   17.74848 2399.54472   0.007 0.994098
recon_yr1915   18.23810 1669.60644   0.011 0.991284
recon_yr1918    1.29946    1.30349   0.997 0.318812
recon_yr1919  -15.44950  845.59314  -0.018 0.985423
recon_yr1920    0.42738    1.16120   0.368 0.712838
recon_yr1921   17.73123 2399.54472   0.007 0.994104
recon_yr1923   19.01369 2399.54472   0.008 0.993678
recon_yr1926   17.84914 2399.54472   0.007 0.994065
recon_yr1927    0.44036    1.18123   0.373 0.709300
recon_yr1928  -14.65179 1163.59142  -0.013 0.989953
recon_yr1929    2.74053    1.24316   2.204 0.027490 *
recon_yr1930    2.09890    0.75003   2.798 0.005135 **
recon_yr1931   -0.05924    1.08356  -0.055 0.956397
recon_yr1932    2.61724    1.30299   2.009 0.044576 *
recon_yr1933    1.12826    0.90355   1.249 0.211774
recon_yr1934   -0.58479    1.07121  -0.546 0.585127
recon_yr1935    0.71655    1.18709   0.604 0.546097
recon_yr1936    1.95332    0.60723   3.217 0.001296 **
recon_yr1937    0.91219    0.71227   1.281 0.200309
recon_yr1938   -0.09195    1.11654  -0.082 0.934367
recon_yr1939    0.00224    1.11085   0.002 0.998391
recon_yr1940    0.59095    0.48870   1.209 0.226581
recon_yr1941    1.07278    0.44159   2.429 0.015125 *
recon_yr1942    1.10271    0.52097   2.117 0.034289 *
recon_yr1943    2.53000    1.27324   1.987 0.046915 *
recon_yr1944    0.12517    1.12862   0.111 0.911694
recon_yr1945    0.63066    0.60393   1.044 0.296369
recon_yr1946    1.00655    0.44109   2.282 0.022491 *
recon_yr1947    0.59000    0.37617   1.568 0.116776
recon_yr1948    0.45183    0.32451   1.392 0.163815
recon_yr1949    0.69921    0.24835   2.815 0.004872 **
recon_yr1950    0.90358    0.24180   3.737 0.000186 ***
recon_yr1951   -0.04181    0.42005  -0.100 0.920712
recon_yr1952    0.41540    0.36612   1.135 0.256534
recon_yr1953    0.13121    0.37227   0.352 0.724498
recon_yr1954   -0.02816    0.39507  -0.071 0.943182
recon_yr1955    0.93635    0.40396   2.318 0.020453 *
recon_yr1956    0.44741    0.38870   1.151 0.249724
recon_yr1957   -0.24448    0.54457  -0.449 0.653478
recon_yr1958    1.22311    0.46481   2.631 0.008503 **
```

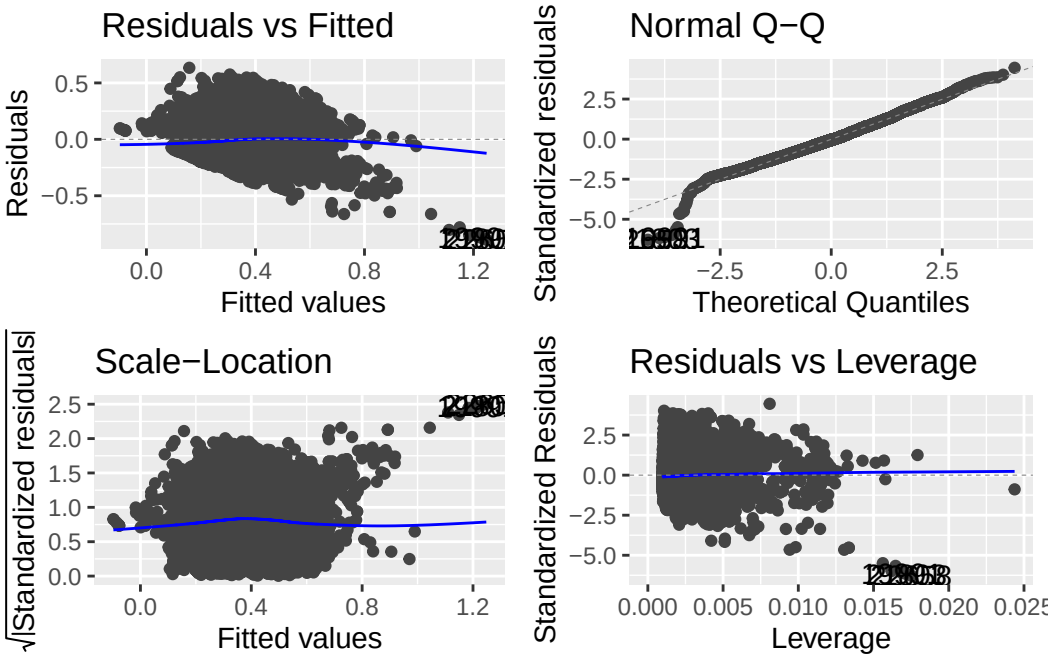
```
recon_yr1959   -0.28104    0.48755  -0.576 0.564328
recon_yr1960    0.41594    0.30464   1.365 0.172140
recon_yr1961    1.00970    0.34789   2.902 0.003704 **
recon_yr1962    0.17150    0.32993   0.520 0.603202
recon_yr1963    0.38742    0.36638   1.057 0.290315
recon_yr1964    0.65435    0.31652   2.067 0.038702 *
recon_yr1965    0.83691    0.37959   2.205 0.027468 *
recon_yr1966    0.18765    0.45995   0.408 0.683284
recon_yr1967    0.76474    0.43451   1.760 0.078408 .
recon_yr1968    0.53783    0.43683   1.231 0.218243
recon_yr1969    0.12659    0.55758   0.227 0.820402
recon_yr1970    1.21281    0.31078   3.902 9.52e-05 ***
recon_yr1971    1.14703    0.47171   2.432 0.015031 *
recon_yr1972    0.51319    0.41163   1.247 0.212491
recon_yr1973    0.92762    0.35110   2.642 0.008241 **
recon_yr1974    0.74866    0.37050   2.021 0.043313 *
recon_yr1975    0.21464    0.38594   0.556 0.578105
recon_yr1976    0.24426    0.40351   0.605 0.544954
recon_yr1977    0.44662    0.35997   1.241 0.214709
recon_yr1978    1.16396    0.26903   4.327 1.51e-05 ***
recon_yr1979    0.80404    0.30205   2.662 0.007768 **
recon_yr1980    0.81341    0.27453   2.963 0.003047 **
recon_yr1981   -0.28612    0.36428  -0.785 0.432197
recon_yr1982    0.80749    0.22081   3.657 0.000255 ***
recon_yr1983    1.01988    0.23159   4.404 1.06e-05 ***
recon_yr1984    0.70632    0.24685   2.861 0.004219 **
recon_yr1985    0.57298    0.22299   2.570 0.010184 *
recon_yr1986    0.94266    0.21348   4.416 1.01e-05 ***
recon_yr1987    0.54386    0.25070   2.169 0.030057 *
recon_yr1988   -0.02023    0.29632  -0.068 0.945562
recon_yr1989    0.33663    0.32720   1.029 0.303563
recon_yr1990    0.15008    0.28873   0.520 0.603211
recon_yr1991    0.71292    0.24099   2.958 0.003093 **
recon_yr1992    0.67356    0.24488   2.751 0.005950 **
recon_yr1993    0.41392    0.29611   1.398 0.162152
recon_yr1994   -0.50801    0.35569  -1.428 0.153220
recon_yr1995    0.24615    0.27683   0.889 0.373921
recon_yr1996    0.74140    0.27317   2.714 0.006646 **
recon_yr1997    0.23986    0.30872   0.777 0.437185
recon_yr1998    0.90183    0.24516   3.679 0.000235 ***
recon_yr1999    0.58560    0.26920   2.175 0.029603 *
recon_yr2000    0.14524    0.33357   0.435 0.663274
recon_yr2001    0.45760    0.25517   1.793 0.072927 .
```

```
recon_yr2002    0.13175    0.31180   0.423 0.672628
recon_yr2003   -0.08135    0.35078  -0.232 0.816617
recon_yr2004   -0.10888    0.39003  -0.279 0.780127
recon_yr2005    0.60137    0.29332   2.050 0.040344 *
recon_yr2006    0.43962    0.28390   1.549 0.121500
recon_yr2007   -0.25560    0.32614  -0.784 0.433207
recon_yr2008    0.52727    0.22617   2.331 0.019737 *
recon_yr2009    0.42721    0.20174   2.118 0.034203 *
recon_yr2010    0.30610    0.19399   1.578 0.114576
recon_yr2011    0.27931    0.21362   1.308 0.191041
recon_yr2012    0.10096    0.24734   0.408 0.683153
recon_yr2013    0.44682    0.23815   1.876 0.060632 .
recon_yr2014    0.10017    0.26663   0.376 0.707150
recon_yr2015    0.24798    0.27760   0.893 0.371687
recon_yr2016   -0.57616    0.37339  -1.543 0.122820
recon_yr2017   -0.74301    0.39543  -1.879 0.060244 .
recon_yr2018   -0.32910    0.33896  -0.971 0.331597
recon_yr2019   -0.39971    0.32273  -1.239 0.215524
recon_yr2020    0.09662    0.32997   0.293 0.769671
recon_yr2021   -0.04953    0.40842  -0.121 0.903469
recon_yr2022  -15.11049 2399.54473  -0.006 0.994976
dk_surf1       -1.55377    0.09635 -16.127  < 2e-16 ***
dk_surf2       -0.19502    0.28138  -0.693 0.488263
dk_surf3       -2.10413    0.22078  -9.531  < 2e-16 ***
dk_surf4       -0.86998    1.13183  -0.769 0.442105
dk_surf5       -2.09859    0.18963 -11.067  < 2e-16 ***
dk_surf6       -0.23677    0.08733  -2.711 0.006705 **
dk_surf7       -0.13895    0.17340  -0.801 0.422939
dk_surf8        0.18876    0.30008   0.629 0.529322
dk_surf9       -0.65242    0.42793  -1.525 0.127358
dk_surfN       -1.38577    0.09372 -14.787  < 2e-16 ***
dk_surfP      -15.41523  125.07627  -0.123 0.901912
POPVar         -0.27744    0.67172  -0.413 0.679591
HURVar         -0.40927    0.11591  -3.531 0.000414 ***
FLEVVar         0.44775    0.10578   4.233 2.31e-05 ***
DDVar           2.32495    0.18555  12.530  < 2e-16 ***
POPVar:DDVar   -4.79305    1.27261  -3.766 0.000166 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 21893  on 27428  degrees of freedom
```

```
Residual deviance: 19954  on 27303  degrees of freedom
AIC: 20206

Number of Fisher Scoring iterations: 15



Call:
lm(formula = SVIVar ~ county + HURVar + POPVar + FLEVVar + CLAVar +
    INDAVar + REPLVar, data = fludd)

Residuals:
     Min       1Q   Median       3Q      Max
-0.89020 -0.09782 -0.00751  0.09479  0.63397

Coefficients:
                     Estimate Std. Error t value Pr(>|t|)
(Intercept)          0.400623   0.008856  45.238  < 2e-16 ***
countyALLEGHENY     -0.115296   0.009914 -11.629  < 2e-16 ***
countyARMSTRONG      0.148413   0.011545  12.855  < 2e-16 ***
countyBEAVER        -0.025948   0.011685  -2.221 0.026392 *
countyBEDFORD        0.148984   0.010051  14.824  < 2e-16 ***
countyBERKS         -0.063602   0.009838  -6.465 1.03e-10 ***
countyBLAIR          0.094747   0.011028   8.592  < 2e-16 ***
countyBRADFORD       0.127304   0.009538  13.347  < 2e-16 ***
countyBUCKS         -0.187243   0.010073 -18.588  < 2e-16 ***
countyBUTLER        -0.017001   0.011103  -1.531 0.125724
countyCAMBRIA        0.150554   0.011446  13.153  < 2e-16 ***
countyCAMERON        0.027164   0.017557   1.547 0.121829
countyCARBON         0.086706   0.014066   6.164 7.19e-10 ***
countyCENTRE        -0.135664   0.010422 -13.017  < 2e-16 ***
countyCHESTER       -0.112494   0.009822 -11.454  < 2e-16 ***
countyCLARION        0.249235   0.013063  19.080  < 2e-16 ***
countyCLEARFIELD     0.210403   0.010346  20.336  < 2e-16 ***
countyCLINTON        0.179951   0.012198  14.752  < 2e-16 ***
countyCOLUMBIA       0.009925   0.011405   0.870 0.384177
countyCRAWFORD       0.169671   0.010537  16.102  < 2e-16 ***
countyCUMBERLAND    -0.037049   0.010700  -3.462 0.000536 ***
countyDAUPHIN       -0.009312   0.011063  -0.842 0.399934
countyDELAWARE      -0.077081   0.012189  -6.324 2.59e-10 ***
countyELK           -0.061823   0.014532  -4.254 2.10e-05 ***
countyERIE           0.000368   0.010248   0.036 0.971354
countyFAYETTE        0.207101   0.010990  18.845  < 2e-16 ***
```

```
countyFOREST          0.199183   0.017473  11.399  < 2e-16 ***
countyFRANKLIN        0.025369   0.010411   2.437 0.014825 *
countyFULTON          0.065634   0.012847   5.109 3.26e-07 ***
countyGREENE          0.168184   0.010690  15.732  < 2e-16 ***
countyHUNTINGDON      0.092977   0.010825   8.589  < 2e-16 ***
countyINDIANA         0.183778   0.011040  16.646  < 2e-16 ***
countyJEFFERSON       0.133492   0.012465  10.709  < 2e-16 ***
countyJUNIATA         0.265484   0.011382  23.325  < 2e-16 ***
countyLACKAWANNA     -0.029199   0.011237  -2.599 0.009368 **
countyLANCASTER       0.051095   0.009182   5.565 2.65e-08 ***
countyLAWRENCE        0.081257   0.012027   6.756 1.44e-11 ***
countyLEBANON         0.065037   0.012153   5.351 8.80e-08 ***
countyLEHIGH         -0.139323   0.010939 -12.736  < 2e-16 ***
countyLUZERNE        -0.019275   0.010366  -1.859 0.062984 .
countyLYCOMING        0.031464   0.010025   3.139 0.001699 **
countyMCKEAN          0.165203   0.011974  13.797  < 2e-16 ***
countyMERCER          0.169993   0.010667  15.936  < 2e-16 ***
countyMIFFLIN         0.097118   0.012548   7.740 1.03e-14 ***
countyMONROE          0.092431   0.011106   8.323  < 2e-16 ***
countyMONTGOMERY     -0.133826   0.010165 -13.166  < 2e-16 ***
countyMONTOUR         0.024355   0.014547   1.674 0.094090 .
countyNORTHAMPTON    -0.111308   0.010965 -10.151  < 2e-16 ***
countyNORTHUMBERLAND  0.007753   0.010776   0.719 0.471845
countyPERRY           0.100410   0.011205   8.961  < 2e-16 ***
countyPHILADELPHIA   -0.195791   0.017789 -11.006  < 2e-16 ***
countyPIKE            0.062014   0.013655   4.542 5.61e-06 ***
countyPOTTER          0.273556   0.011840  23.105  < 2e-16 ***
countySCHUYLKILL      0.034526   0.010712   3.223 0.001270 **
countySNYDER          0.075215   0.011474   6.555 5.64e-11 ***
countySOMERSET        0.184265   0.010947  16.833  < 2e-16 ***
countySULLIVAN        0.178125   0.013154  13.542  < 2e-16 ***
countySUSQUEHANNA     0.063841   0.010098   6.322 2.61e-10 ***
countyTIOGA           0.142225   0.009399  15.133  < 2e-16 ***
countyUNION           0.181137   0.012447  14.553  < 2e-16 ***
countyVENANGO         0.155240   0.012726  12.199  < 2e-16 ***
countyWARREN          0.118193   0.011813  10.005  < 2e-16 ***
countyWASHINGTON     -0.022962   0.009829  -2.336 0.019487 *
countyWAYNE           0.026198   0.010656   2.459 0.013955 *
countyWESTMORELAND   -0.031018   0.010158  -3.054 0.002263 **
countyWYOMING         0.049509   0.012192   4.061 4.90e-05 ***
countyYORK           -0.070750   0.009225  -7.669 1.78e-14 ***
HURVar                0.007419   0.006927   1.071 0.284230
POPVar                1.444860   0.023181  62.328  < 2e-16 ***
```

```
FLEVVar              -0.124188   0.009597 -12.941  < 2e-16 ***
CLAVar               -0.223889   0.037905  -5.907 3.53e-09 ***
INDAVar              -0.259271   0.022491 -11.528  < 2e-16 ***
REPLVar               0.170122   0.041972   4.053 5.07e-05 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.143 on 27356 degrees of freedom
Multiple R-squared:  0.4017,    Adjusted R-squared:  0.4001
F-statistic: 255.1 on 72 and 27356 DF,  p-value: < 2.2e-16
```



Discussion