**Summary of Preliminary Findings for 2015 SWDI/SE Hail Merge**

**Summary of 2015 Storm Events Hail Data**

dim(original Storm Events data):

[1] 62942 61

dim(Hail Storm Events data:

[1] 8744 61

dim(deduplicated Hail Storm Events data):

[1] 5674 13

Summary of deduplicated Hail Events Storm data

DATE LON LAT mininjuries maxinjuries

Min. :2015-02-01 Min. :-149.75 Min. :18.00 Min. :0 Min. :0.0000000

1st Qu.:2015-04-28 1st Qu.:-101.50 1st Qu.:35.50 1st Qu.:0 1st Qu.:0.0000000

Median :2015-06-12 Median : -96.50 Median :39.25 Median :0 Median :0.0000000

Mean :2015-06-15 Mean : -94.88 Mean :38.98 Mean :0 Mean :0.0003525

3rd Qu.:2015-07-18 3rd Qu.: -88.75 3rd Qu.:42.50 3rd Qu.:0 3rd Qu.:0.0000000

Max. :2015-12-30 Max. : -65.75 Max. :61.50 Max. :0 Max. :1.0000000

mindeaths maxdeaths minproperty maxproperty mincrops

Min. :0 Min. :0 Min. :0.00000 Min. :0.00000 Min. :0.000000

1st Qu.:0 1st Qu.:0 1st Qu.:0.00000 1st Qu.:0.00000 1st Qu.:0.000000

Median :0 Median :0 Median :0.00000 Median :0.00000 Median :0.000000

Mean :0 Mean :0 Mean :0.05058 Mean :0.07931 Mean :0.006697

3rd Qu.:0 3rd Qu.:0 3rd Qu.:0.00000 3rd Qu.:0.00000 3rd Qu.:0.000000

Max. :0 Max. :0 Max. :1.00000 Max. :1.00000 Max. :1.000000

maxcrops minanydamage maxanydamage

Min. :0.00000 Min. :0.00000 Min. :0.00000

1st Qu.:0.00000 1st Qu.:0.00000 1st Qu.:0.00000

Median :0.00000 Median :0.00000 Median :0.00000

Mean :0.01163 Mean :0.05146 Mean :0.08037

3rd Qu.:0.00000 3rd Qu.:0.00000 3rd Qu.:0.00000

Max. :1.00000 Max. :1.00000 Max. :1.00000

Table of minanydamage (rows) by maxanydamage (columns):

0 1

0 5218 164

1 0 292

**Summary of 2015 SWDI Hail Data**

dim(SWDI excluding obs outside of continental US “box”):   
[1] 786202 10

dim(SWDI deduped by DATE/LAT/LON):   
[1] 91984 8

summary(dedupedSWDI$minSEVPROB):  
 Min. 1st Qu. Median Mean 3rd Qu. Max.

0.00 0.00 10.00 14.32 20.00 100.00

summary(dedupedSWDI$maxSEVPROB):

Min. 1st Qu. Median Mean 3rd Qu. Max.

0.00 20.00 40.00 42.85 60.00 100.00

summary(dedupedSWDI$minMAXSIZE):

Min. 1st Qu. Median Mean 3rd Qu. Max.

0.5000 0.5000 0.5000 0.6793 0.7500 4.0000

summary(dedupedSWDI$mAXMAXSIZE):

Min. 1st Qu. Median Mean 3rd Qu. Max.

0.500 0.750 1.000 1.179 1.500 4.000

**Merge and Logit Results**

Table of inSE (dummy variable if in Storm Events data)

0 1

88281 3703

Table of maxanydamage (max in deduped SE data of the anydamage flag)

0 1

91671 313

Call:

glm(formula = inSE ~ I(maxSEVPROB - mean(maxSEVPROB)) + I(maxMAXSIZE -

mean(maxMAXSIZE)), family = "binomial", data = merge\_SWDI\_SE)

Deviance Residuals:

Min 1Q Median 3Q Max

-0.8484 -0.2782 -0.1766 -0.1118 3.4629

Coefficients:

Estimate Std. Error z value Pr(>|z|)

(Intercept) -3.985940 0.029639 -134.482 < 2e-16 \*\*\*

I(maxSEVPROB - mean(maxSEVPROB)) 0.042932 0.001226 35.013 < 2e-16 \*\*\*

I(maxMAXSIZE - mean(maxMAXSIZE)) 0.246692 0.034678 7.114 1.13e-12 \*\*\*

---

Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

(Dispersion parameter for binomial family taken to be 1)

Null deviance: 31046 on 91983 degrees of freedom

Residual deviance: 25252 on 91981 degrees of freedom

AIC: 25258

Number of Fisher Scoring iterations: 7

Waiting for profiling to be done...

2.5 % 97.5 %

(Intercept) 0.01751606 0.01967427

I(maxSEVPROB - mean(maxSEVPROB)) 1.04136601 1.04638353

I(maxMAXSIZE - mean(maxMAXSIZE)) 1.19566522 1.36977391

Call:

glm(formula = maxanydamage ~ I(maxSEVPROB - mean(maxSEVPROB)) +

I(maxMAXSIZE - mean(maxMAXSIZE)), family = "binomial", data = merge\_SWDI\_SE)

Deviance Residuals:

Min 1Q Median 3Q Max

-0.3639 -0.0706 -0.0410 -0.0262 4.2132

Coefficients:

Estimate Std. Error z value Pr(>|z|)

(Intercept) -6.844582 0.121014 -56.560 < 2e-16 \*\*\*

I(maxSEVPROB - mean(maxSEVPROB)) 0.035364 0.004314 8.198 2.44e-16 \*\*\*

I(maxMAXSIZE - mean(maxMAXSIZE)) 0.759331 0.097750 7.768 7.97e-15 \*\*\*

---

Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

(Dispersion parameter for binomial family taken to be 1)

Null deviance: 4182.6 on 91983 degrees of freedom

Residual deviance: 3442.3 on 91981 degrees of freedom

AIC: 3448.3

Number of Fisher Scoring iterations: 10

Waiting for profiling to be done...

2.5 % 97.5 %

(Intercept) 0.0008318207 0.001337481

I(maxSEVPROB - mean(maxSEVPROB)) 1.0273473296 1.044871936

I(maxMAXSIZE - mean(maxMAXSIZE)) 1.7638085739 2.587704470

**Discussion**

While the number of deduped SWDI records is 91,984, which is a lot more than the number of deduped Storm Event records (5,674), it may make sense to drop the SWDI records with low probability of storm severity.  At any rate, after merging 3,700 out of 91,984 were in both SWDI and Storm events data; only 313 out of the 3,700 had any damage reported.

I estimated two logits – one for whether or not the record was in the Storm Events data and the other for whether there was any damage, using maxSEVPROB and maxMAXSIZE as predictors. Signs of the results seem to make sense. We might want to whittle down the SWDI data based on the maxSEVPROB variable.