| **Script Name** | **What Script Does** | **Inputs** | **Outputs** |
| --- | --- | --- | --- |
| PerfTables.Rmd | * Creates interactive tables of classification performance measures | * PerfMeasuresP\_Hats\_ALL.rda * PerfMeasuresP\_Hats\_ALL.csv | * Interactive webpage |
| PerfMeasuresP\_Hats\_ALL.R | * Four big for loops through all the predicted probability estimates by:   + meth=method:     - majority vote     - average   + algor=algorithm:     - ridge regression     - random forests     - generalized additive model with splines     - ensemble   + datasrc=Census data used:     - None     - ACS     - LODES     - Both   + gran=granularity (i.e., spatial resolution):     - 1.0 degree     - 0.25 degree   + measure=classification performance measure:     - AUC     - Recall     - Specificity | * Prediction\_Hail\_ALL\_1deg\_3pred.rda * Prediction\_Hail\_LODES\_1deg\_3pred.rda * Prediction\_Hail\_ACS\_1deg\_3pred.rda * Prediction\_Hail\_NONE\_1deg\_3pred.rda * Prediction\_Hail\_ALL\_qtrdeg\_3pred.rda * Prediction\_Hail\_LODES\_qtrdeg\_3pred.rda * Prediction\_Hail\_ACS\_qtrdeg\_3pred.rda * Prediction\_Hail\_NONE\_qtrdeg\_3pred.rda | * PerfMeasuresP\_Hats\_ALL.rda * PerfMeasuresP\_Hats\_ALL.csv |
| Prediction\_Hail\_ALL\_1deg\_3pred.R  Prediction\_Hail\_LODES\_1deg\_3pred.R  Prediction\_Hail\_ACS\_1deg\_3pred.R  Prediction\_Hail\_NONE\_1deg\_3pred.R | * Generates predicted probabilities at 1.0 degree resolution using the algorithms listed above and using different sets of explanatory variables as listed above. | * SWDI\_SE\_ACS\_LODES\_1deg.rda | * Prediction\_Hail\_ALL\_1deg\_3pred.rda * Prediction\_Hail\_LODES\_1deg\_3pred.rda * Prediction\_Hail\_ACS\_1deg\_3pred.rda * Prediction\_Hail\_NONE\_1deg\_3pred.rda |
| Prediction\_Hail\_ALL\_qtrdeg\_3pred.R  Prediction\_Hail\_LODES\_qtrdeg\_3pred.R  Prediction\_Hail\_ACS\_qtrdeg\_3pred.R  Prediction\_Hail\_NONE\_qtrdeg\_3pred.R | * Generates predicted probabilities at 0.25 degree resolution using the algorithms listed above and using different sets of explanatory variables as listed above. | * SWDI\_SE\_ACS\_LODES.rda | * Prediction\_Hail\_ALL\_qtrdeg\_3pred.rda * Prediction\_Hail\_LODES\_qtrdeg\_3pred.rda * Prediction\_Hail\_ACS\_qtrdeg\_3pred.rda * Prediction\_Hail\_NONE\_qtrdeg\_3pred.rda |
| merge\_ACS\_SWDI\_SE\_LODES.R | * Merges LODES data to the merged SWDI/SE/ACS data. * Sets employment data to zero for non-merges. * Constructions explanatory variables (jobs per square mile at tract and county level). | * LodesData2011.rda * SWDI\_SE\_ACS.rda | * SWDI\_SE\_ACS\_LODES.rda |
| LODES.R | * Downloads and unzips LEHD Origin-Destination Employment Statistics data from the Census Bureau website. * Sums employment data by tract and by county FIPS and merges the two together. | * ss\_wac\_S000\_JT00\_2011.csv.gz, where “ss” is lowercase state postal code for lower 48 states + DC downloaded from <http://lehd.ces.census.gov/data/lodes/LODES7/> | * LodesData2011.rda |
| merge\_ACS\_SWDI\_SE\_rev.R | * Merges ACS data to the merged SWDI/SE data (after getting rid of SWDI/SE records that could not be geocoded because they were in the ocean or Canada or Mexico). * Constructs explanatory variables. | * ACS\_2006\_2010\_counties\_Tracts.rda * SWDI\_SE\_geo.rda | * SWDI\_SE\_ACS.rda |
| prepare\_SWDI\_SE\_for\_ACS\_merge.R  prepare\_SWDI\_SE\_for\_ACS\_merge\_remainder.R | * Creates dataframe of unique coordinates from the SWDI/SE file * Uses FCC API to create a table of coordinates and their corresponding 2010 Census Block FIPS code * Takes a long time to run; had to fill in the remainder of table after a time out. * Note that some SWDI coordinates could not be matched because they were not located on US land territory (i.e., in an ocean or in Canada or Mexico). | * SWDI\_SE\_merge\_2011\_2015.rda | * SWDI\_SE\_geo.rda |
| merge\_ACS\_county\_tract.R | * Merges tract- and county-level ACS data | * ACS\_2006\_2010\_counties.rda * ACS\_2006\_2010.rda | * ACS\_2006\_2010\_counties\_Tracts.rda |
| ACS\_county\_level.R | * Downloads 2006-2010 ACS Summary File data by county * Merges it with Census geodata by county | * ACS Summary File data via API calls from acs package * Gaz\_counties\_national.txt | * ACS\_2006\_2010\_counties.rda |
| ACS.R | * Downloads 2006-2010 ACS Summary File data by census tract * Merges it with Census geodata by census tract | * ACS Summary File data via API calls from acs package * Gaz\_tracts\_national.txt | * ACS\_2006\_2010.rda |
| assess\_merge\_SWDI\_SE\_hail\_2011\_2015\_new.R | * Merges SWDI and SE data and assesses the quality of the merge | For years 2011-2015:   * DeDuplicateHailStormEventData\_[year].rda * DeDuplicateSWDIHailData\_[year].rda | * SWDI\_SE\_merge\_2011\_2015.rda |
| prepare\_SWDI\_for\_storm\_events\_data\_merge\_2011\_2015.R | * Reads in raw SWDI hail data * Limits data to lower 48 states * Rounds coordinates to the nearest ¼ degree * Deduplicates data by date/coordinates keeping max/min for SEVPROB and MAXSIZE | For years 2011-2015:   * Nexrad\_Hail\_[year].rda | For years 2011-2015:   * DeDuplicateSWDIHailData\_[year].rda |
| prepare\_storm\_events\_data\_for\_SWDI\_merge\_2011\_2015.R | * Extracts hail data from Storm Events data * Rounded coordinates to nearest ¼ degree * Converted times (hence dates) from local standard time to GMT * Used beginning and ending coordinates to develop records reflecting full path of hail storms * Deduplicated by date/coordinates keeping max/min values of hail size, economic and health damages | For years 2011-2015:   * merged\_SE\_[year].rda | For years 2011-2015:   * DeDuplicateHailStormEventData\_[year].rda |
| StormEvents\_2008\_2015\_merge\_location\_details.R | * Downloads and unzips Storm Event details and location files and merges them by Episode and Event ID | <http://www1.ncdc.noaa.gov/pub/data/swdi/stormevents/csvfiles/> | For years 2008-2015:   * merged\_SE\_[year].rda |
| NexradHail\_2007\_2015.R | * Downloads SWDI data on hail events | See CDS tutorial on hail data. | For years 2007-2015:   * Nexrad\_Hail\_[year].rda |

One Degree Resolution

|  |  |  |  |
| --- | --- | --- | --- |
| merge\_ACS\_SWDI\_SE\_LODES\_1deg.R | * Merges ACS data to the merged SWDI/SE data (after getting rid of SWDI/SE records that could not be geocoded because they were in the ocean or Canada or Mexico). * Constructs explanatory variables. | * LodesData2011.rda * SWDI\_SE\_ACS\_1deg.rda | * SWDI\_SE\_ACS\_LODES\_1deg.rda |
| merge\_ACS\_SWDI\_SE\_1deg.R | * Merges ACS data to the merged SWDI/SE data (after getting rid of SWDI/SE records that could not be geocoded because they were in the ocean or Canada or Mexico). * Constructs explanatory variables. | * ACS\_2006\_2010\_counties\_Tracts.rda * SWDI\_SE\_geo\_1deg.rda | * SWDI\_SE\_ACS\_1deg.rda |
| prepare\_SWDI\_SE\_for\_ACS\_merge\_1deg.R | * Creates dataframe of unique coordinates from the SWDI/SE file * Uses FCC API to create a table of coordinates and their corresponding 2010 Census Block FIPS code * Note that some SWDI coordinates could not be matched because they were not located on US land territory (i.e., in an ocean or in Canada or Mexico). | * SWDI\_SE\_merge\_2011\_2015\_1deg.rda | * SWDI\_SE\_geo\_1deg.rda |
| assess\_merge\_SWDI\_SE\_hail\_2011\_2015\_new\_1deg.R | * Merges SWDI and SE data and assesses the quality of the merge | For years 2011-2015:   * DeDuplicateHailStormEventData\_[year]\_1deg.rda * DeDuplicateSWDIHailData\_[year]\_1deg\_.rda | * SWDI\_SE\_merge\_2011\_2015\_1deg.rda |
| prepare\_SWDI\_for\_storm\_events\_data\_merge\_2011\_2015\_1deg.R | * Reads in raw SWDI hail data * Limits data to lower 48 states * Rounds coordinates to the nearest 1 degree * Deduplicates data by date/coordinates keeping max/min for SEVPROB and MAXSIZE | For years 2011-2015:   * Nexrad\_Hail\_[year].rda | For years 2011-2015:   * DeDuplicateSWDIHailData\_[year]\_1deg.rda |
| prepare\_storm\_events\_data\_for\_SWDI\_merge\_2011\_2015\_1deg.R | * Extracts hail data from Storm Events data * Rounded coordinates to nearest 1 degree * Converted times (hence dates) from local standard time to GMT * Used beginning and ending coordinates to develop records reflecting full path of hail storms * Deduplicated by date/coordinates keeping max/min values of hail size, economic and health damages | For years 2011-2015:   * merged\_SE\_[year].rda | For years 2011-2015:   * DeDuplicateHailStormEventData\_[year]\_1deg.rda |