Here are the first few rows of the data that we currently have for county-level resolution:

Table

Description automatically generated

The data includes nine rows per combination of county and storm—one for each day from 5 days before the date the storm came closest to that county to three days after. As a note, the lag 0 date (the date the storm came closest to the county) is county specific for each storm—it is *not* the date of landfall. The figure below shows an example for Hurricane Floyd of how this date is county-specific.



Only counties in the eastern half of the US are included. Each county is identified by its five-digit FIPs code. Each storm is identified by a storm ID (combo of storm name and year) as well as its Automated Tropical Cyclone Forecasting System ID. Here are the descriptions for each column:

Graphical user interface, text, application, email

Description automatically generated

Most file formats would work fine for the data. We store most of our original data in CSV files, which are very easy to read into R, but we should be able to work with most file formats.

For ZIP-code level data, we’ll want one row for each combination of ZIP code and storm. We are including all areas in the states of: Alabama, Arkansas, Connecticut, Delaware, District of Columbia, Florida, Georgia, Illinois, Indiana, Iowa, Kansas, Kentucky, Louisiana, Maine, Maryland, Massachusetts, Michigan, Mississippi, Missouri, New Hampshire, New Jersey, New York, North Carolina, Ohio, Oklahoma, Pennsylvania, Rhode Island, South Carolina, Tennessee, Texas, Vermont, Virginia, West Virginia, and Wisconsin.

I will create a datafile with the centroid of each ZIP code we want to get storm wind data for. I think the ZIP-code boundaries might change somewhat frequently, so I can try to get a separate list of latitudes and longitudes for each ZIP code center for each year.