HIGH SCHOOL EDITION
Journal of Student Research

Volume 13 Issue 1 (2024)

An Investigation of the Influence of the Fall Line on Tornado Occurrence in the Southeast United States

Adhav Ravikumar[1] and Varun Venkatagiri[1]

[1]Innovation Academy, USA

ABSTRACT

Tornadoes, highly destructive natural phenomena, cause substantial damage and loss of life annually. This study investigates a peculiar trend: the clustering of tornadoes near the geological fall line in the Southeastern US, focusing on Georgia, South Carolina, and North Carolina. Utilizing geospatial analysis and NOAA data from 2012 onwards, the research examines the correlation between tornado occurrence and the fall line's position. The analysis reveals a noteworthy pattern: around 32.657% of tornadoes occurred within the fall line region (FLR), a proportion significantly higher than expected by chance (p-value = 0.0732). While the correlation is modest, it holds statistical significance. The study underscores the need for heightened vigilance, protection, and public education in FLR areas, particularly in socio-economically challenged zones such as the "Black Belt." Additionally, the paper proposes future research avenues, including expanding analysis to varied terrains and historical periods. The findings underscore the importance of considering geographical factors for accurate tornado risk assessment and preparedness. By enhancing our understanding of tornado patterns, this study contributes to fortifying vulnerable communities and refining disaster management strategies.

Introduction

Tornadoes are a strong example of one of nature's most deadly and devastating natural disasters, resulting in over 400 million dollars in damage on average per year in the United States (National Geographic). Tornadoes form when warm, humid air collides with cold, dry air. The warm air rises through the cold air, which results in an updraft or in other words a current of rising air. The result is a deadly whirling storm that can have devastating effects on anything in and around its path. When studying the April 2022 tornado outbreak in Georgia, one may have noticed an interesting trend. Seemingly, many of the tornadoes were in and around the fall line (Fig. 1).

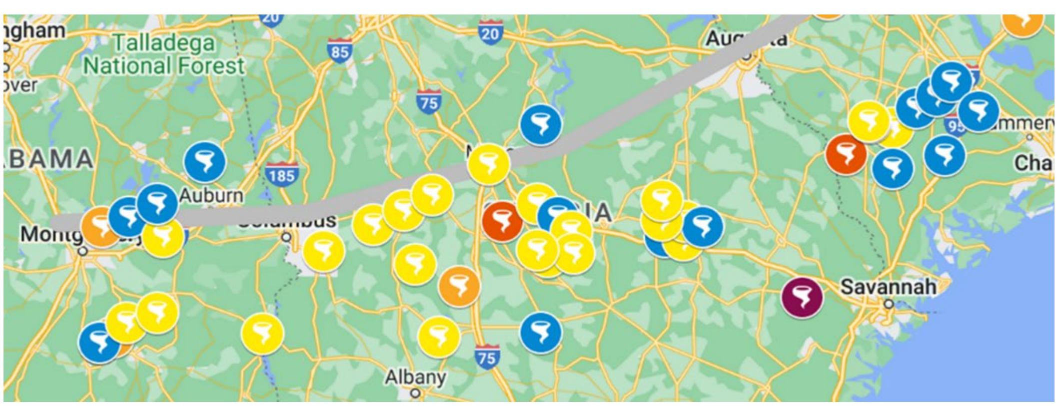

Figure 1. Dot Distribution map of the April 2022 tornado outbreak's locations and intensities.

The Fall Line

In order to understand the importance and the results of the study, an understanding on the fall line is necessary. The fall line is a geological line that stretches across the South, dividing the flat fertile Coastal Plains and hilly Piedmont regions of Georgia (Shankman & Hart, 2007). It runs through many cities in the south, such as Macon, Augusta, Montgomery, Columbus, and Columbia. It also marks the prehistoric shoreline of the Atlantic Ocean. As a result, the area south of the fall line, the Coastal Plains, is very fertile and home to large rural populations due to the decomposition of the high biodiversity that lived in the ocean through phenomena like marine snow, combined with clay soil that allowed limited percolation of those nutrients (Phillips, 1993).

Interestingly, studies have been conducted showing that “transition zones”, such as the region near the fall line, can have higher rates of tornadoes. As stated in a Purdue University study, “Areas where landscape shifts from urban to rural or forest to farmland may have a higher likelihood of severe weather and tornado touchdowns” (Kellner & Niyogi, 2014). Considering that the fall line marks the line that divides the flat, farmlands of the Coastal Plains to the hilly, forested Piedmont region, the trend found in the April 2022 tornado outbreak may not be a coincidence. Consequently, it was necessary for research to done utilizing available open-use data from national weather organizations in order to conduct a geospatial analysis of whether tornado patterns in the Southeast US really does follow the trend of occurring near the fall line.

Through our research process we reviewed other studies that compare the topography of a region to prevalence of tornadoes. For example, a study conducted by R. Panneer Selvem and Nawful Ahmed that found that tornadoes tend toward higher elevations and make more damage moving uphill. Another study, conducted by Northern Illinois University, has shown that “tornado frequency has begun shifting eastward from the Great Plains” (Gensini & Brooks, 2018). This shift in frequency could mean that the Southeast region, the region relevant to our study, could be facing an increase in frequency of tornadoes over the last couple of decades. However, we learnt that there has been no previous study done to determine the correlation between the fall line and the prevalence of tornadoes in the region.

For our study we will compare the frequency of tornados inside the fall line region (FLR) with the frequency of tornadoes outside the FLR in the tri-state region of Georgia, South Carolina, and North Carolina. For the purpose of the study, we will consider the fall line region to be 1-degree longitude or fifty-five miles south of the fall line. The reason why we chose Georgia, South Carolina, and North Carolina for our study is because these three states are within the same geographical region of the United States, the South-east, whereas other states along the fall line, from Virginia and northward, can classify as a part of the mid-Atlantic region. Hence in order to limit other factors, namely regional climate differences, from influencing our data, we chose to limit our study to the three states of Georgia, South Carolina, and North Carolina, which we will refer to as the tri-state region.

Methods

The goal of this research is to determine whether there is a correlation between the fall line region (FLR) and tornado occurrence. The methods for this research is based on previous geospatial analysis studies, such as that done by Newcastle University researcher Gainbi Park for hurricane damage and Clemson University researcher Joel Feltman’s team for wildfire damage. Both of these studies mapped natural disaster occurrences over a large area of land in the US and then found patterns, which they then tested in order to determine a prediction for future occurrence of these disasters. Data of all tornadoes since 2012 was obtained from the National Oceanic and Atmospheric Administration (NOAA) and filtered using Excel, removing tornados determined to be too small intensity or out of region. Then, the data was analyzed against the FLR through a multi-step approach involving region bounding, FLR land area calculation, data to FLR overlaying, and a single-tailed chi-square test.

First, the data is needed. In order to get the data, we source it from NOAA. The NOAA provides free and open access to data, meaning this method of data collection is legal and ethical. Then, utilizing Excel operations, we filtered out tornadoes occurred before 2012 or occurred outside Georgia, South Carolina, and North Carolina. After doing this, we had a sample size of 986.

Then, we must bound a region. To do this, we have to mathematically express the FLR as a function of longitude and latitude, which can be done using quadratic regression and algebraic manipulation. North and West bounds will be set by the Georgia-Alabama and North Carolina-Virginia borders.

Then, the area of the FLR and tri state area must be calculated and compared to each other. This can be done by evaluating the area between the curves of the mathematical model of the FLR using two double integrals, and converting square degrees into square miles. The area of the tri-state area will be calculated by adding the area of Georgia, South Carolina, and North Carolina. Finally, in order to compute the proportion of land in the tri-state area taken up by the FLR, we would divide the FLR's area by the tri-state area's area.

After this, we need to compute the proportion of tornadoes that lie within the FLR. To do this, we will utilize Python in order to write a script that will process the Excel data and see whether a given tornado falls into the mathematical representation of the FLR based on the latitude and longitude of the starting point of the tornado. This procedure will be repeated 986 times, once for each tornado. The total number of tornados within the FLR will be divided by 986 to find the proportion of tornadoes that occur inside the FLR.

Finally, we need to find whether this proportion is statistically significant. To do this, we decided to perform a one-tailed chi-squared test with $\alpha = .1$. We chose to perform a chi-squared test because the data was collected qualitatively - whether or not they were in the region - rather than a numerical quantity. Hence, the distribution of the tornadoes inside and outside the region should be compared to that of the expected based on the land area. It is also important to address the fact that we chose a non-traditional significance level of .1 due to the fact that a large amount of people would be affected by even a weak trend that is found by our research. So, even though there is an increased possibility of a Type I error with this level, a Type II error would have a much larger negative impact than Type I in this case, justifying the use of $\alpha = .1$.

Results

After defining methods, the study was conducted using the outlined four-step process: region bounding, FLR land area, data to FLR overlaying, and conducting a chi-squared test to analyze the data.

Region Bounding

Firstly, we had to determine the extent of which to define the FLR. Consequently, we had to mathematically express the fall line as a function of the longitude and latitude in order to create a objective way to determine whether a given tornado fell within the FLR. In order to do this, we utilized quadratic regression as the fall line follows a curve that roughly resembles that of a parabola (Fig. 2). This process output Equation 1.

Equation 1. Regression output for fall line, where x is the longitude value in degrees and y is the latitude value in degrees.

$$y = .062(x + 85.149)^2 + 32.037$$
$$R > .9$$

After this, we had to set a lower bound for the FLR, which we determined to be one degree of longitude, or approximately 55 miles. Through algebra and approximation of the scale factor to result in a curve that was one unit away from the original curve when measured normal to the original curve, the result was Equation 2.

Equation 2. Equation of the lower bound of the FLR.

$$y = .057(x + 85.149)^2 + 31.037$$

Finally, we had to set left and upper bounds to the region. Using online mapping tools, we found the latitude measure of the Alabama-Georgia border, or the western border of the three-state region considered for our purposes. That is 85.198°W, which can also be expressed as -85.198 for easy use mathematically here. This process was repeated, except for the North Carolina-Virginia border in order to determine the northern bound for our tri-state region. This was determined to be 38.22°N, or 38.22.

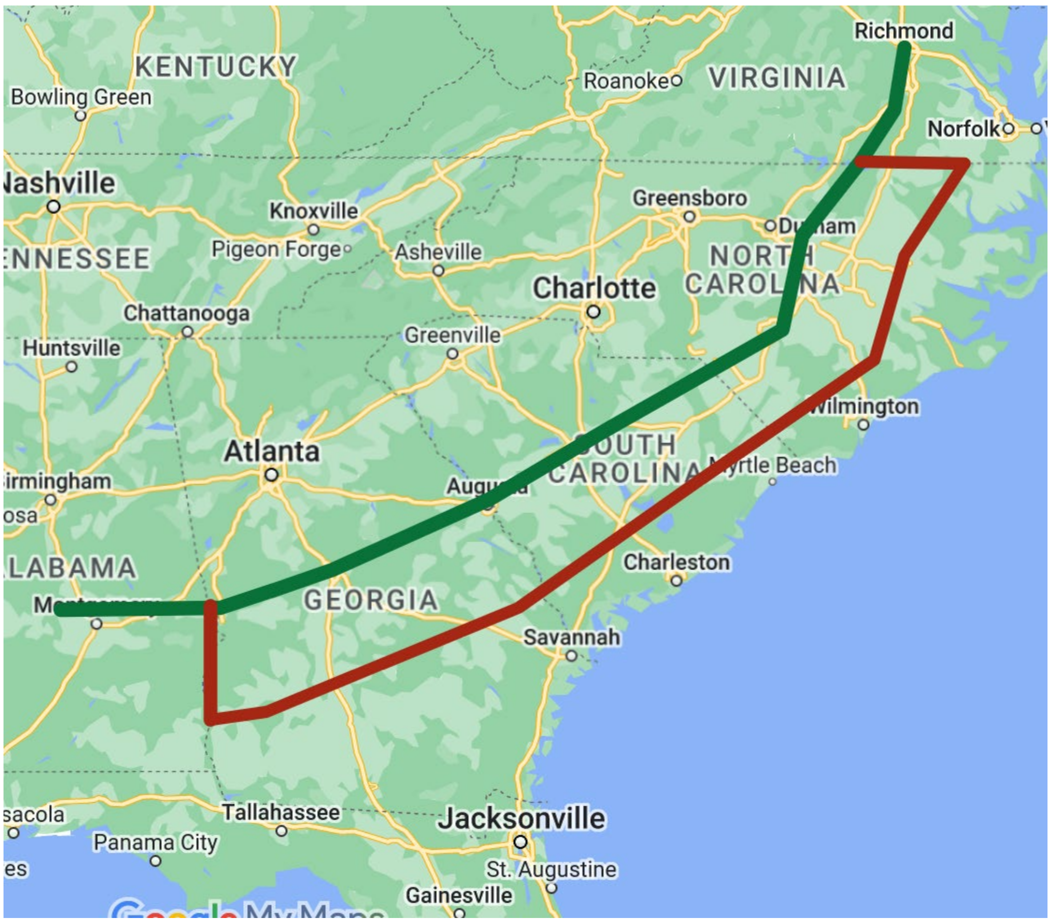

Figure 2. The FLR visualized as an enclosed region on Google MyMaps. The fall line is in green.

 Volume 13 Issue 1 (2024)

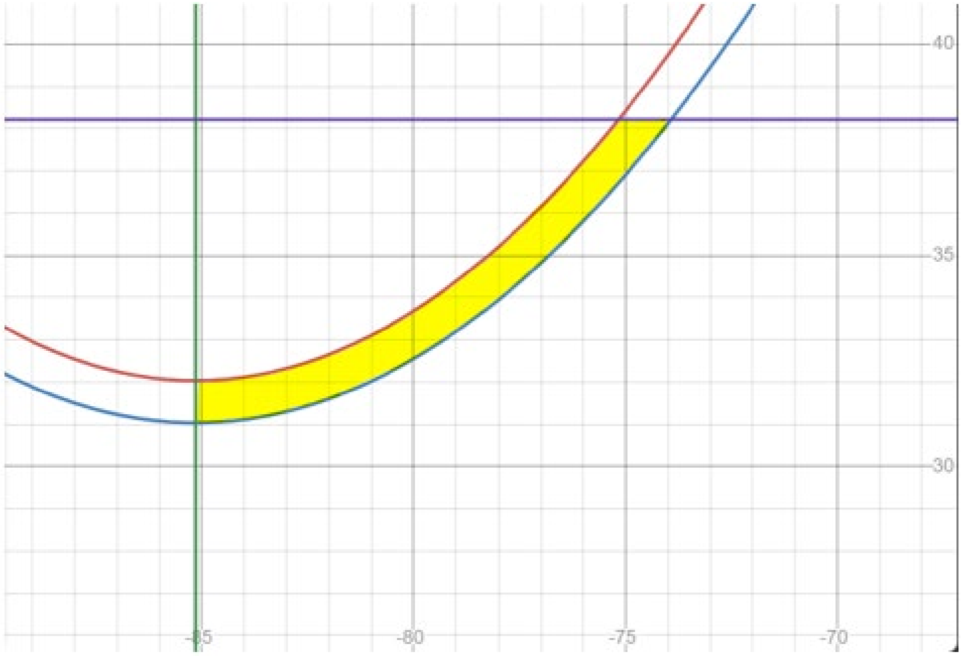

Figure 3. The functions expressed on a coordinate plane with the shaded region shown.

FLR Land Area

For the later statistical tests, the percentage of the three states covered by the FLR must be determined. This can be done mathematically using two double integrals (Equation 3).

Equation 3. Integral expression used to find the area of the FLR.

$$\int_{-85.198}^{-75.396}\int_{.057(x+85.149)^2+31.037}^{.062(x+85.149)^2+32.037} dydx + \int_{36.839}^{37.935}\int_{\sqrt{\frac{y}{.062}-32.027}-85.149}^{\sqrt{\frac{y}{.057}-31.037}-85.149} dxdy = 12.579$$

Hence, the area of the region is 12.579 square degrees, which equates to 44532.80 square miles. The total area of the tri-state area is found by adding Georgia, South Carolina, and North Carolina's area (Equation 4).

Equation 4. Expression used to find total area of the tri-state area.

$$59425\ (\text{GA}) + 32020\ (\text{SC}) + 53819\ (\text{NC}) = 145264\ \text{sq mi}$$

The proportion of the tri state area covered by the FLR is represented by Equation 5.

Equation 5. Expression used to find the overlap proportion of the FLR and tri-state area.

$$44532.80/145264 = .30525$$

Data to FLR Overlaying

Once this was complete, we had to check whether each tornado location was in the region. To do this, we put the NWS data (starting latitude and longitude of every tornado since 2012) into an Excel file and accessed it via the pandas Python library. Then, to determine whether it was within the region, it was checked whether it satisfied Equation 6.

Equation 6. Inequalities that define the FLR mathematically.

$$y < .062(x + 85.149)^2 + 32.037$$
$$y > .057(x + 85.149)^2 + 31.037$$
$$x > -85.1982$$
$$y < 38.22$$

This was accomplished through the script in Appx. 1. The code output is in Table 1.

Table 1. Output of the Python script that evaluates the number and proportion of tornadoes inside the FLR.

Number of tornadoes in FLR	322
Proportion of tornadoes in FLR	0.3265720081135903

Hence, 322 out of the 986 tornadoes, or 32.657%, in the excel sheet fell within the FLR. In order to analyze this, we must use a one-tailed chi-squared test.

Discussion

In order to determine whether this number was statistically significant, a one-tailed chi-squared test with one degree of freedom (based on the standard chi-squared degrees of freedom formula) was performed. The proportion of tornadoes that fell within the FLR was compared to the percentage of the tri-state area occupied by the FLR through the following hypotheses:

H_o: The proportion of tornadoes in the tri state area that occur inside the FLR will follow the proportion of land area in the tri state area that is inside the FLR.

H_a: The proportion of tornadoes in the tri state area that occur inside the FLR will be greater than the proportion of land area in the tri state area that is inside the FLR.

Table 2. Distribution of observed and expected tornadoes inside and outside the FLR.

	Category	Observed	Expected #	Expected %
1	In FLR	322	301	30.527%
2	Out of FLR	664	685	69.473%

The resulting p-value of the test is .0732. Based on a significance level of $\alpha = .1$, $p < \alpha$, so we can conclude that the proportion of tornadoes in the tri state area that occur inside the FLR will be greater than the proportion of land area in the tri state area that is inside the FLR. Hence, we have weak evidence of tornado occurrences following the trend of the FLR. However, it is important to note that with a significance level of $\alpha = .05$, the test would not be significant, meaning there is a statistically significant, albeit weak trend in this study.

Conclusion

Due to there being significant evidence for a majority of tornadoes occurring inside the FLR, it is important to act on these findings and give added surveillance and protection to the people living in the FLR due to tornadoes' devastating effects on individual people and communities as a whole. Our results are not surprising, nor do they go against previous research; the aforementioned Kellner & Niyogi study provides a scientific explanation for this result. Additionally, many of the residents in the FLR are low-income or minorities, leading to systemic bias against them. Additionally, many residents of these areas have significant misconceptions about the intensity and geographical location of these tornadoes.

The "Black Belt"

Stephen Strader, a professor at Villanova University who specializes in the intersection between tornadoes and geography states, "Turns out a lot of people don't understand the historical relationships between minority populations, rural areas, and poverty across the Southeast" (Strader, 2021). This is further elaborated by researcher Jack Sillin of Cornell University, who highlights a correlation between radar coverage, the "Black Belt", and the FLR we focused on. According to Gerald Webster of the University of Alabama, "In the nineteenth century settlement focused upon the region's rich dark soils for which it was originally named. The Black Belt became the site of the South's Antebellum plantation-cotton-slave complex. Today many of the counties in the region have large African American populations and are more noted for their lack of economic opportunities than the fertility of their soils" (Webster, 2008). Due to many reasons, including systemic discrimination, the residents of the "Black Belt" have been found to have the least surveillance, but based on our study, one of the highest tornado occurrence rates in the tri state area. Hence, using our results, it is vital for the protection and livelihood of these residents for radar and tornado surveillance instruments to be mobilized in the FLR area, and especially the "Black Belt."

Misconceptions

Additionally, according to a Carnegie Mellon University study by Stephen Broomell, "many members of the Southeastern public deviate from the expert sample on tornado likelihood" (Broomell et al, 2020). These deviations include the time of year for tornado and geographical location of tornado. The study further elaborates, saying many residents of the Southeast get their tornado-related information by extrapolating that about the Great Plains tornado. This leads to numerous misconceptions, as the weather patterns in the two regions are vastly different from each other. Hence, the results of this study and many other related ones about the occurrences of tornadoes is important to educate the general public of the American southeast about in order to make sure they are fully prepared for these deadly natural disasters.

Limitations

While the study and its methodology were constructed in a manner that allowed for as little bias or error as possible, there are some limitations to it. Firstly, as mentioned above, the significance level used of .1 can increase the odds of a type I error, which means that the null hypothesis was rejected when it was true. However, we chose to use a higher significance level in order to mitigate the chance of a type II error, which means that the alternate hypothesis was rejected when it was true. This is because a type II error would mean that a correlation that tornadoes adversely affected the FLR is there, but was not detected in the study, meaning that the people who are affected more in these areas by tornadoes would not get the added surveillance and protection they need.

Another limitation is that we did not use tornado data before 2012. This was because those tornadoes and outbreaks that did not have a set start or end coordinate. Instead of filtering out those tornadoes and causing convenience bias, we chose to focus the study on tornadoes that occurred after 2012. While this means that we had to limit our sample size, using more recent tornadoes also allows us to easier apply our results to the present day.

Lastly, tornado patterns have some randomness that we did not account for. Tornado density will depend on factors that was not analyzed, such as atmospheric wind, temperature and humidity conditions. These conditions are unlikely to be spatially uniform in the FLR, leading to confounding factors that were not accounted for. However, accounting for these factors in the long term is essentially impossible with present day meteorological technology and models. It is possible that future research could take these factors into account in the short term and produce more definite predictions.

Acknowledgments

I would like to thank my advisor for the valuable insight provided to me on this topic.

References

Broomell, S. B., Wong-Parodi, G., Morss, R. E., & Demuth, J. L. (2020). Do we know our own tornado season? A Psychological investigation of perceived tornado likelihood in the Southeast United States. *Weather, Climate, and Society*, *12*(4), 771–788. https://doi.org/10.1175/wcas-d-20-0030.1

Feltman, J. A., Straka, T. J., Post, C. J., & Sperry, S. (2012). Geospatial Analysis application to forecast wildfire occurrences in South Carolina. *Forests*, *3*(2), 265–282. https://doi.org/10.3390/f3020265

Gensini, V. A., & Brooks, H. E. (2018). Spatial trends in United States tornado frequency. *Npj Climate and Atmospheric Science*, *1*(1). https://doi.org/10.1038/s41612-018-0048-2

Kellner, O., & Niyogi, D. (2014). Land Surface Heterogeneity Signature in Tornado Climatology? An Illustrative Analysis over Indiana, 1950–2012*. *Earth Interactions*, *18*(10), 1–32. https://doi.org/10.1175/2013ei000548.1

National Oceanic and Atmospheric Administration. (1995, February). *Globdisttornado*. https://www.ncdc.noaa.gov/img/climate/research/tornado/globdist.jpg

NOAA National Centers for Environmental Information. (2023, January). *Monthly Tornadoes Report for Annual 2022*. https://www.ncei.noaa.gov/access/monitoring/monthly-report/tornadoes/202213

NOAA Storm Prediction Center. (2023). *Storm Prediction Center Severe Weather GIS (SVRGIS)* [Dataset; CSV file]. National Oceanic and Atmospheric Administration. https://www.spc.noaa.gov/gis/svrgis/

NWS Charleston. (2022, May 31). April 5-6, 2022 Tornado outbreak. *ArcGIS StoryMaps*. https://storymaps.arcgis.com/stories/65a1fd5ce0534809a5b0e75054ac0615

Park, G. (2021). A Comprehensive Analysis of Hurricane Damage across the U.S. Gulf and Atlantic Coasts Using Geospatial Big Data. *ISPRS International Journal of Geo-information*, *10*(11), 781. https://doi.org/10.3390/ijgi10110781

Phillips, J. D. (1993). CHAOTIC EVOLUTION OF SOME COASTAL PLAIN SOILS. *Physical Geography*, *14*(6), 566–580. https://doi.org/10.1080/02723646.1993.10642498

Shankman, D., & Hart, J. L. (2007). The Fall Line: a Physiographic-Forest Vegetation Boundary. *Geographical Review*, *97*(4), 502–519. https://doi.org/10.1111/j.1931-0846.2007.tb00409.x

Shepherd, M. (2021, March 20). Are Black And Rural Residents In The South More Vulnerable To Tornadoes Due To Radar Gaps? *Forbes*. https://www.forbes.com/sites/marshallshepherd/2021/03/20/are-black-and-rural-residents-in-the-south-more-vulnerable-to-tornadoes-due-to-radar-gaps/?sh=6ad028454988

Sillin, J. (2021). *Are Black Americans Underserved by the NWS Radar Network?* https://sillinweather.files.wordpress.com/2021/04/southeast-radar-coverage-inequality.png?strip=info&w=2000

Webster, G. R., & Bowman, J. (2008). Quantitatively delineating the Black Belt geographic region. *Southeastern Geographer*, *48*(1), 3–18. https://doi.org/10.1353/sgo.0.0007