Deriving Population Exposure Fatality Rate Estimates for Tornado Outbreaks Using Geographic Information Systems (GIS)

Amber R. Cannon[1, 2], K. Klockow[3], R. Peppler[4], H. Brooks[5]

[1]National Weather Center Research Experiences for Undergraduates Program
[2]East Central University, Ada OK

[3]Oklahoma Climatological Survey
University of Oklahoma Department of Geography & Environmental Sustainability, Norman, OK

[4]Cooperative Institute for Mesoscale Meteorological Studies
University of Oklahoma Department of Geography & Environmental Sustainability, Norman, OK

[5]National Oceanic and Atmospheric Administration
National Severe Storms Laboratory, Norman OK

ABSTRACT

In this study we looked at several issues regarding the derivation of population exposure and fatality rate estimates during widespread tornado outbreaks. The two events studied were the April 3, 1974 and April 27, 2011 tornado outbreaks in the state of Alabama. We attempted to determine if tornado warnings have become more effective over time at reducing the number of fatalities. We used GIS to perform an analysis on these outbreaks. We found that the effectiveness of tornado warnings did improve between the two outbreaks, and we can have reasonably high confidence in using county level population data to compare recent and historical outbreaks, although the higher resolution of the census track data is preferred for studying a single tornadic event. We also found that the accuracy of fatality rates is directly related to the accuracy of the path data. Finally, GIS can be used to innovatively evaluate tornado warning effectiveness.

1. INTRODUCTION

The use of GIS techniques to evaluate tornado-related damage and fatalities, while not yet common practice in the field of meteorology, has recently gained some attention. A theoretical study was performed by Rae and Stefkovich (2000) addressing the possible ramifications of a major tornado outbreak in the Dallas-Fort Worth Metro Area using modeled data from the May 3, 1999 tornado outbreak in Oklahoma, and one study, in the field of demography, has performed a similar analysis using GIS and actual storm data (Donner 2007). There were three major research questions in this study that were addressed. Firstly, "Has the effectiveness of tornado warnings improved?" This question was answered by comparing the deadliness of 1974 outbreak to the deadliness of 2011 tornado outbreak within the state of Alabama. Secondly, "How well can we compare historical outbreaks to present day outbreaks from a population density perspective?" This question was answered by using the geographic location and storm event as a control and the variables were the spatial resolution from two data sets. The two data sets compared were the higher spatial resolution census tract level population data with the coarser spatial resolution population data from a county level. Finally, "Can GIS be used to evaluate the effectiveness of tornado warnings?" This question was answered by using multiple maps as a visual representation of effectiveness. The use of geographical representations for storm data can be extremely beneficial as this study will show.

[1] *Amber Renee' Cannon*
1707 Pecan Grove Street
Tecumseh, OK 74873
soccernay03@yahoo.com

2. DATA AND METHODOLOGY

2.1 *DATA TYPES*
There were four main types of data that were used for this study, which are outlined in the sections below; these data are population data, geographic data, storm data and orthophotography.

2.2 *POPULATION DATA*
Two sets of population data were used in this study. The 2010 population data were acquired from the U.S. Census Bureau, American FactFinder (http://factfinder2.census.gov/faces/nav/jsf/pages/index.xhtml) and the 1974 population data were acquired from the National Bureau of Economic Research, Census U.S. Decennial County Population Data, 1900-1990 (http://www.nber.org/data/census-decennial- population.html). Each data set was downloaded in comma delimited (.csv) format and then imported into Microsoft Excel and converted to a single sheet workbook (.xls) format that could then be imported into the ESRI ArcGIS 9.3 software. There, it was converted to a database (.dbf) format that could later be joined with its respective geographic locations.

2.3 *GEOGRAPHIC DATA*
The geographic data used were acquired from the U.S. Census Bureau, Geography Division, Geographic Products Branch 2010 Census TIGER/Line® Shapefiles (http://www.census.gov/geo/www/tiger/tgrshp2010/tgrshp2010.html). These included the state and county boundaries of Alabama, all census tracts, all major hydrography (including lakes, rivers, and marsh land), and all major roads (including interstates and us highways) within the state.

2.4 *STORM DATA*
Tornado tracks data were acquired from the National Weather Service (NWS) Southern Region Headquarters (http://www.srh.noaa.gov/srh/ssd/mapping/) in KML format. This office provided a combined file of the tornado tracks reported by each individual forecast office within the region. The detailed storm data and preliminary fatality numbers were acquired from the NOAA / NWS National Centers for Environmental Prediction, Storm Prediction Center web page (http://www.spc.noaa.gov/climo/reports/110427_r pts.html).

2.5 *ORTHOPHOTOGRAPHY*
Orthophotography was used to compare the accuracy of the reported tornado tracks and was acquired from the National Geodetic Survey, April 2011 Tornado Response Imagery Viewer (http://ngs.woc.noaa.gov/storms/apr11_tornado/). The aerial photography missions were conducted after the 2011 outbreak by the NOAA Remote Sensing Division. The images were acquired from a nominal altitude of 5,000 - 15,000 feet, using a Trimble Digital Sensor System (DSS) aboard the NOAA King Air 350CER aircraft.

2.6 *CARTOGRAPHIC REPRESENTATIONS*
For the purpose of the most accurate areal calculations in the study area of interest within Alabama, the North American Datum 1983 was used in conjunction with the UTM zone 16 North Projection. All other shapefiles used were properly transformed from WGS 1984 and projected to the same. Using the UTM projection allowed the software to make the calculations of area in metric units.

2.7 *DERIVATION OF PATH WIDTH*
A rubric for path width was derived from the probability models in Figures 4 and 6 of Brooks (2004). This is shown in Table 1.

EF Scale	Buffer Dist.	Tot. Path Width
EF0	50 meters	100 meters
EF1	125 meters	250 meters
EF2	250 meters	500 meters
EF3	500 meters	1000 meters
EF4	1000 meters	2000 meters
EF5	1500 meters	3000 meters

Table1. *Tornado Width rubric based on EF Scale (from Brooks 2004).*

2.8 *DATA COMPILATION AND REFINEMENT*
The data in this study were combined using the ESRI ArcGIS version 9.3 software (http://www.esri.com/software/arcgis/index.html). Once all of the data were collected, ArcGIS project files (.mxd) were created for each topic of interest in this study as well as individual geodatabases for each case. The data then were imported into the project files to begin analysis. Each project file contains multiple feature classes. Most project files contain cities, major roadways, and major waterways within the boundaries of Alabama.

Individual case studies were built to suit the particular needs of the study; for example, the county level comparisons contain the county's shape file, population by the county's database table, and the tornado tracks, in addition to the previously listed feature classes.

Once all of the data were imported into the project file and the projections were set, the first step toward the analysis was to create an area feature representing the tornado paths. To achieve this, the linear features representing the tornado paths were buffered using the rubric shown in Table 1. This process derived a polygon shape file that was stored in the respective study's geodatabase. From this point, ESRI's ArcCatalog 9.3 software (http://www.esri.com/software/arcgis/index.html) was used to create topology for the new polygon shapefile of tornado path widths.

After the creation of topology, the topology edit toolbar in ArcGIS was used to split the tornado paths by the boundaries of other features in the project file. For example, in the study using the county level population data, the county boundaries were selected and the cut polygon feature in the editor toolbar was used to split the tornado paths by the county boundaries. This results in one tornado path being divided into multiple segments that occur within only one county. After the paths were divided, the area of each segment was calculated using the calculate geometry feature within the attribute table. The land area of each county was then calculated the same way. All calculations were performed in square meters and then converted to square kilometers.

The population density of each county was then derived using the field calculator command in the attribute table to divide the total population of the county by the total land area in kilometers of the county resulting in a population density of persons per square kilometer. The county shapefile containing population density of the county was then joined to the tornado path shapefile. Once the files were joined, the population density was then multiplied by the area of each path segment which resulted in the number of affected persons per segment, or population exposure. The number of affected persons column was then summed to achieve a total population exposure for that event. The total number of fatalities was then divided by the total population exposure to derive the fatality rate per affected residents. The path area column was also summed to achieve a total path area and a fatality rate per square kilometer of affected land was also calculated. Statistical computations were also calculated for the population densities resulting in minimum, maximum, and mean population densities.

2.9 *DATA CHALLENGES*

There were many challenges associated with the collection of data in this study. The most challenging situation arouse while attempting to acquire the 2010 census population data at the county and tract level. At the time of this study, the data from the 2010 decennial census was still being processed and was not fully released to the general public. Although the idea of using the older data from the 2000 census was discussed, the decision was made to remain persistent to acquire the newest data available for the greatest accuracy. After many frustrations and long hours of trial and error with the new census web site, the 2010 census data was acquired and used for this study.

3. RESULTS

Figures 1 and 2 provide context on the 1974 and 2011 April outbreaks that impacted Alabama. These two particular outbreaks are the two largest tornado outbreaks on record for a single 24-hour period. They are also the two deadliest tornado outbreaks in the last seventy-five years. The 1974 outbreak was confined to the northern third of the state and resulted in 77 fatalities. Long-track tornadoes impacted approximately 971 square kilometers and more than 27,000 people along the paths. The 2011 outbreak was a bit more widespread, covering the northern two-thirds of the state, and resulted in 246 fatalities. Tornadoes, several of which were long-track, impacted approximately 2,088 square kilometers and more than 94,000 people along the paths. Both outbreaks occurred during nearly the same time of day. During the 1974 outbreak, the first tornado in Alabama touched down around 6:30pm, while in the 2011 outbreak the first touchdown in Alabama happened at around 3pm and both outbreaks spanned throughout the evening and into the night.

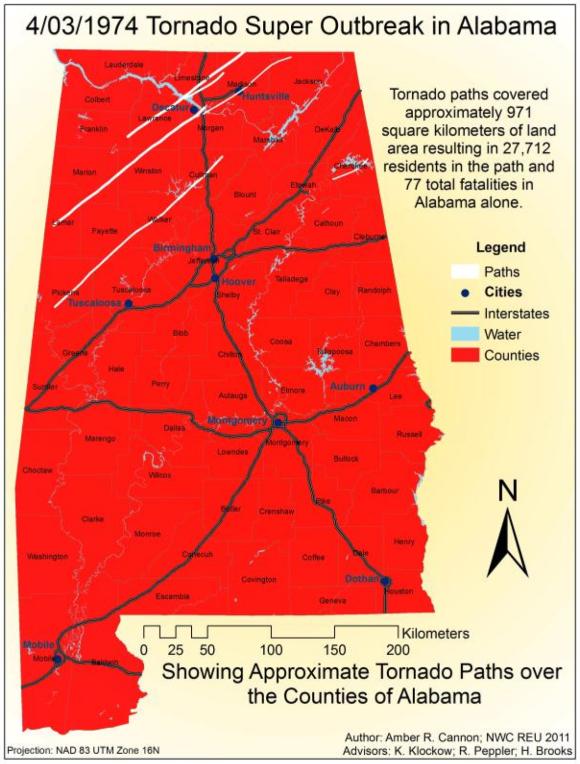

Figure 1: Paths of tornadoes (shown in white) in Alabama on April 3, 1974.

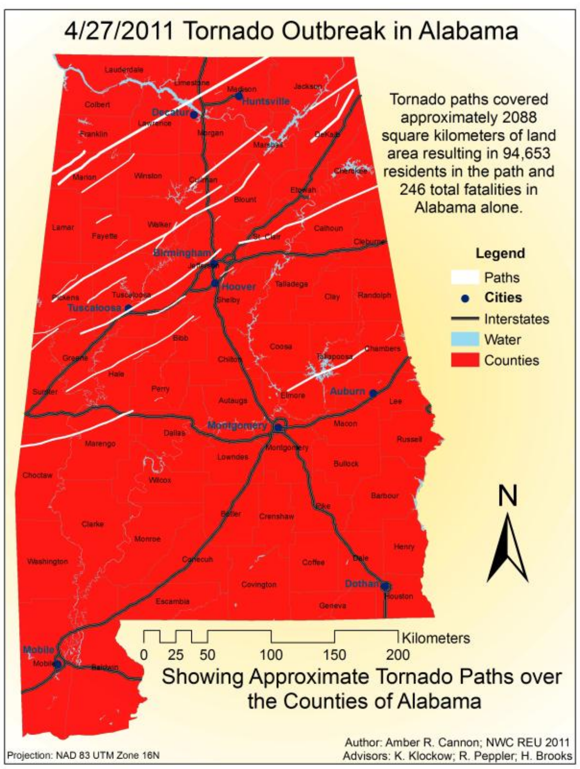

Figure 2: Paths of tornadoes in Alabama on April 27, 2011.

Figures 3 and 4 show the paths of the 1974 and 2011 tornadoes superimposed over the population densities based on the 1970 and 2010 U.S. Census, respectively, aggregated at the county level. As is shown, tornadoes impacted more urban areas in 2011 than in 1974, but not much detail can be provided beyond the county level.

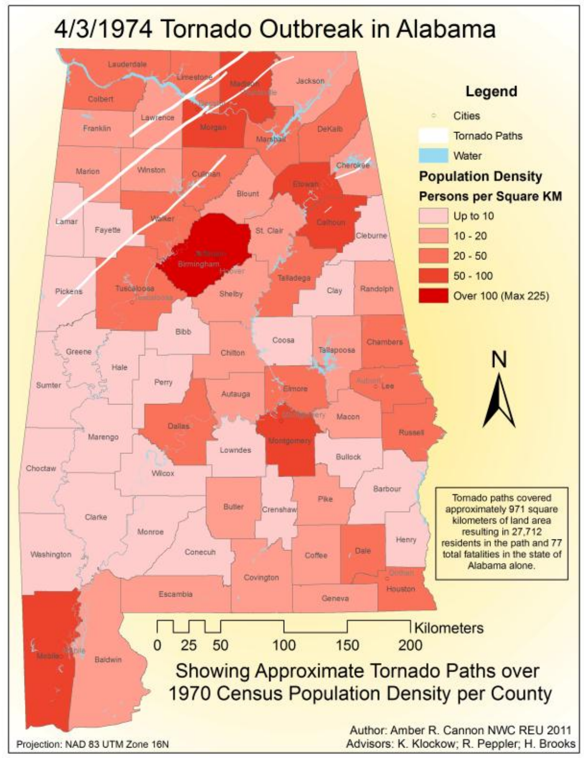

Figure 3: Paths of the 1974 tornadoes statewide overlaid on 1970 population density at the county level.

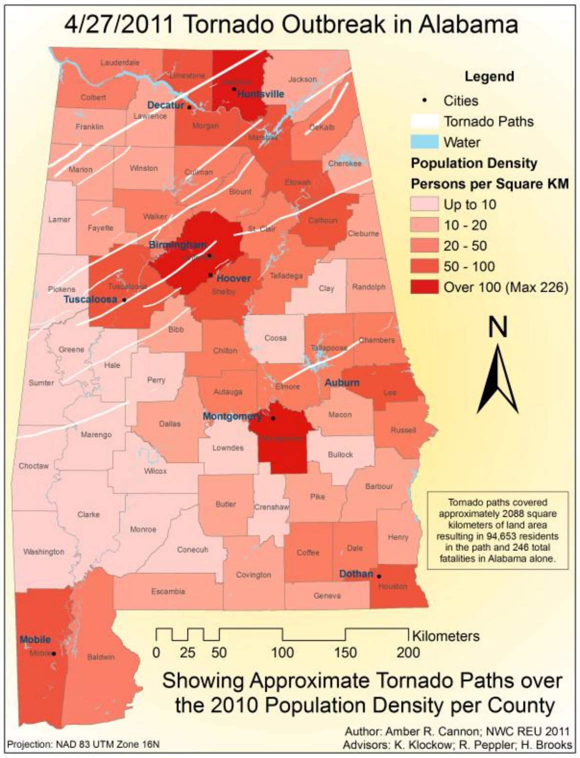

Figure 4: Paths of the 2011 tornadoes statewide overlaid on 2010 population density at the county level.

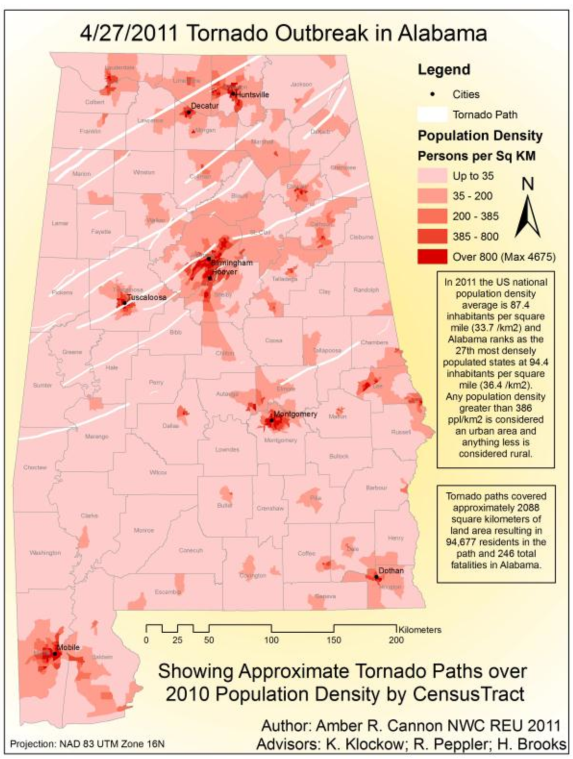

Figure 5: Paths of the 2011 tornadoes statewide overlaid on 2010 population density at the census tract level.

Figure 5 looks a little more closely at 2011 by superimposing the tornado paths over population densities aggregated at the census tract level. Here, population density is much more refined spatially and it can be seen how the Tuscaloosa, Birmingham, and Huntsville metropolitan areas were impacted by the tornadoes. The first three categories in population density represent rural areas while the fourth and fifth categories represent urban areas.

Figure 6 zooms in on the census tract analysis of Figure 5 for the Tuscaloosa and Birmingham areas. The main tornado made a direct hit on Tuscaloosa but skirted the north and west sides of Birmingham. The tract level data allows us to see this with fine detail. Figure 7 is the same but zooms in even more on the Birmingham metro area. Here, although many densely populated areas were affected by this storm, the most populous areas of Birmingham were spared. Figure 8 shows how Tuscaloosa was not as lucky, as the tornado traveled right through the heart of the city, barely sparing the University of Alabama.

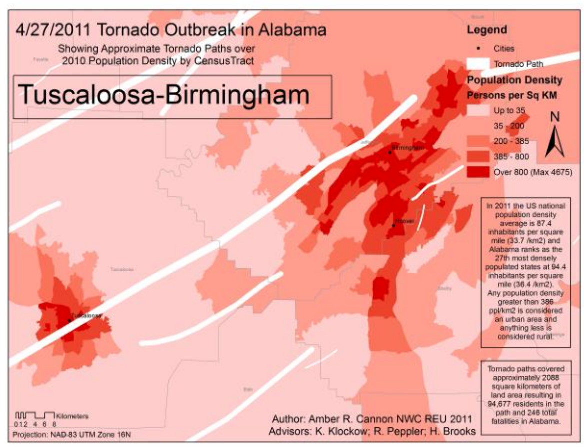

Figure 6: Same as Figure 5 but for the Tuscaloosa and Birmingham areas.

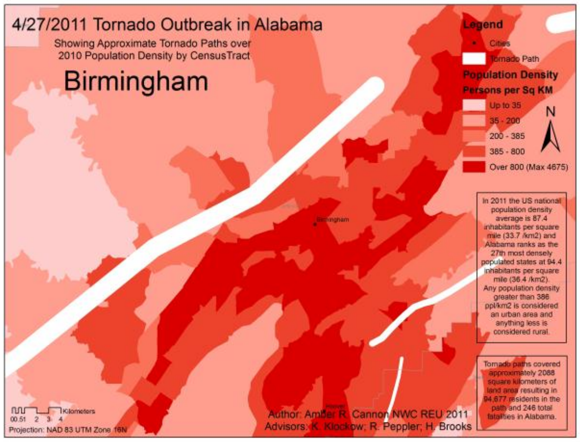

Figure 7: Same as Figure 6 but for the urban area of Birmingham.

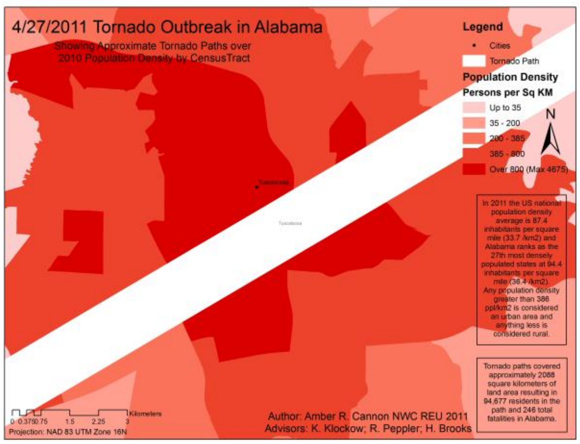

Figure 8: Same as Figure 7 but for the urban area of Tuscaloosa.

The sensitivity analysis of EF-scale rating is shown in Figure 9. Rating tornadoes with the EF-scale is a rigorous yet ultimately subjective exercise, especially when deciding between two ratings such as EF-4 and EF-5, when damage is severe in each. The Tuscaloosa/Birmingham long-track tornado was rated as high-end EF-4. We wanted to know how many more people would have been included in the path if the EF rating for this tornado was upgraded from EF4 to EF5 using the rubric of Brooks (2004).

It turns out that this rubric increased the width of the path from 2000 meters to 3000 meters, which added 65 square kilometers to the damage path. This increased the number of residents exposed from 94,677 to 105,761, which is an increase of 11,084 residents.

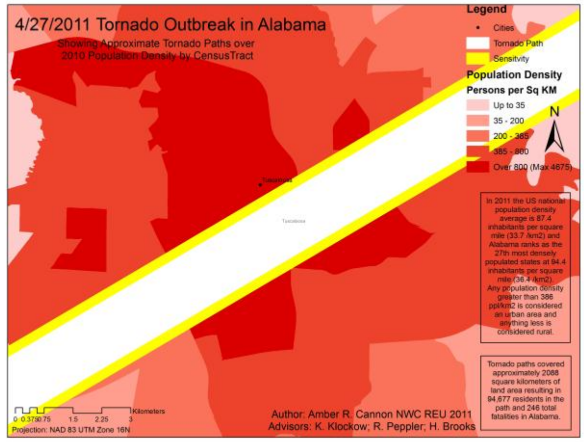

Figure 9: Sensitivity analysis of the intensity of the Tuscaloosa/Birmingham tornado. The increase in path width from EF-4 to EF-5, based on the rubric of Brooks (2004), is shown in yellow.

Figures 10 and 11 show how the accuracy of NWS storm track coordinates can be verified using orthophotography. Figure 10 shows how the designated storm path through Tuscaloosa proves extremely accurate compared to the photographic data, as the actual path falls directly in the center of the yellow lines representing the path shown using GIS coordinates. Figure 11, however, shows how the designated storm path near Haleyville in Winston County in northern Alabama (an EF-3 tornado) is inaccurate compared to aerial photographic evidence. Thus, care must be taken when using designated GIS tornado track coordinates with the more refined census tract population data, but the discrepancies found should not matter as much (or at all) with county level data.

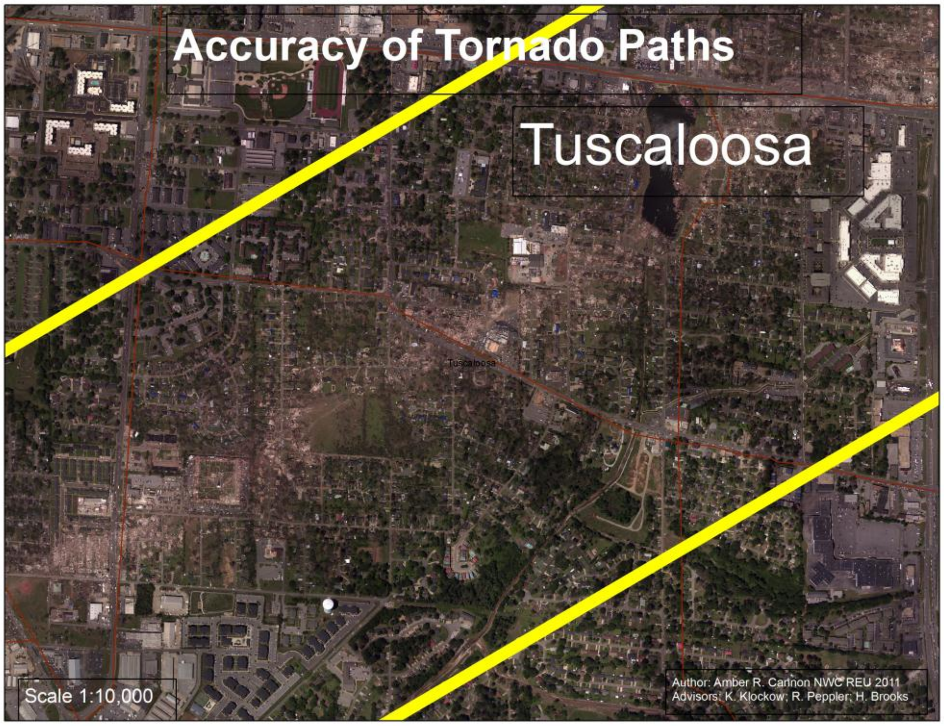

Figure 10: Test of the accuracy of NWS storm track coordinates for Tuscaloosa using orthophotography. Actual damage from photography can be seen on the image, and path boundaries as designated by GIS coordinates are shown in yellow.

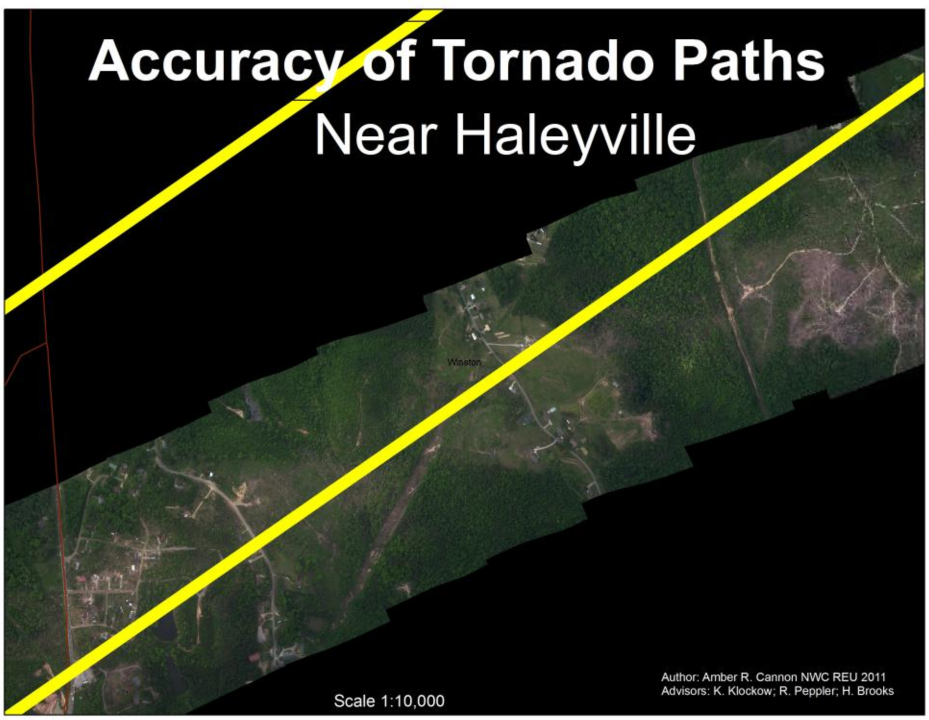

Figure 11: Test of the accuracy of NWS storm track coordinates for Haleyville using orthophotography. Actual damage from photography can be seen on the image, and path boundaries as designated by GIS are shown in yellow.

Finally, Table 2 shows numerical results derived from this study. As can be seen, the percentage of fatalities per affected residents decreased slightly from the 1974 outbreak (0.278%) to the 2011 outbreak (0.260%), with little difference when using the 2011 county or tract data (which means that county level data should be adequate for this type of analysis). However, the percentage of fatalities over the entire population increased slightly from 1974 to 2011 (0.002% to 0.005%) – nearly twice the number of counties was impacted in 2011 (32) than in 1974 (17) and population statewide has increased from 1970 to 2010 (3.4 million to 4.8 million).

Thus, accounting for absolute and mean density increases in population, the rate of death by exposure decreased between 1974 and 2011. This presumptively is attributable to improvements in the tornado warning system. Further research into warning information reception, interpretation and response may support this finding. It also is evident from the last column that increasing the EF scale rating of the Tuscaloosa storm path made a significant impact on the overall number of persons affected and as such reduced the percentage of fatalities of affected residents.

Tornado Outbreak Statistical Comparison Results

4/3/1974 vs. 4/27/2011

By Amber R. Cannon

6/30/2011

	4/3/1974 Outbreak by County	4/27/2011 Outbreak by County	4/27/2011 Outbreak by Tract	4/27/2011 Outbreak by Tract - Sensitivity Study
Total Population of Alabama	3,444,165	4,779,736	4,779,736	4,779,736
Total Number of Residents in the Path	27,712	94,653	94,677	105,761
Total Area of Path	971 km^2	2088 km^2	2088 km^2	2153 km^2
Number of Fatalities	77	246	246	246
Fatalities per Square Kilometer	.079	.118	.118	.114
Percentage of Fatalities of Affected Residents	**.278%**	**.260%**	**.260%**	**.233%**
Percentage of Fatalities of Total Population	.00224%	.00515%	.00515%	.00515%
Total Number of Counties Affected	17	32	32	32
Minimum Population Density	5.8 ppl/km^2	5.3 ppl/km^2	0 ppl/km^2	0 ppl/km^2
Maximum Population Density	224 ppl/km^2	226 ppl/km^2	459,642 ppl/km^2	459,642 ppl/km^2
Mean Population Density	24.8 ppl/km^2	48.6 ppl/km^2	389 ppl/km^2	389 ppl/km^2

In 2011 the US national population density average is 87.4 inhabitants per square mile (33.7 /km2) and Alabama ranks as the 27th most densely populated states at 94.4 inhabitants per square mile (36.4 /km2).

Any population density greater than 386 ppl/km^2 is considered an urban area and anything less is considered rural.

Table 2: Numerical results derived from this study. The first column shows 1974 county level results, the second columns shows 2011 county level results, the third column shows 2011 tract level results, and the last columns shows the results of the Tuscaloosa sensitivity analysis.

4. SUMMARY AND CONCLUSIONS

This study has shown that GIS can be used to innovatively evaluate tornado warning effectiveness. We showed that one can have reasonably high confidence in using just county level population data to compare recent and historical outbreaks, although the higher resolution of the census tract data would be preferred for studying a single tornadic event.

We also found that the accuracy of fatality rates is directly related to the accuracy of the path data, and large errors can result if the path data are inaccurate. Based on the experience of this study, it is of utmost importance to assess the overall quality of the path data before instilling a high level of confidence in the results that area derived.

5. FUTURE WORK

This project was undertaken as a pilot project to determine if we can go back in time to meaningfully analyze the population statistics of past tornado events. Since we believe we have done that, we would like to analyze other historical outbreaks. We also would like to compare the entire 1974 and 2011 outbreaks, both of which occurred over multi-state areas. In addition, we would like to analyze significant individual events at a tract level, such as the deadly Joplin, Missouri tornado of May 2011. And, we would like to develop an automated system for the GIS part of this analysis procedure for increased productivity.

Once the process is automated, we also believe that it can be used in conjunction with forecasting models to further increase the effectiveness of NWS warnings by pinpointing the geographic locations of severe damage. This will also allow for increased lead time for emergency managers, emergency responders, and government organizations like FEMA to better prepare for and dispatch response teams. This could be extremely useful in tornado outbreak situations that span multiple states or FEMA districts.

6. ACKNOWLEDGMENTS

The authors would like to thank Daphne LaDue, Kevin Kloesel, Somer Erickson, Stephanie Hoekstra and the Research Experiences for Undergraduates (REU) Program for making this paper possible. Additionally, they are indebted to the Oklahoma Climatological Survey, Norman, OK.

I would also like to thank Dr. Mark Micozzi, Dr. Greg Plumb and Dr. Robert Newcomer, East Central University Department of Cartography and Geography, Ada, OK for their advice, guidance, and support in helping to make this opportunity a reality for me. Lastly, I would like to thank Jason Redden, University of Oklahoma Geoinformatics Norman, OK and Joel Foster, East Central University Department of Cartography and Geography, Ada, OK for assistance with census data and GIS.

This work was prepared by the authors with funding provided by National Science Foundation Grant No. AGS-1062932, and NOAA/Office of Oceanic and Atmospheric Research under NOAA-University of Oklahoma Cooperative Agreement #NA08OAR4320904, U.S. Department of Commerce.

The statements, findings, conclusions, and recommendations are those of the author(s) and do not necessarily reflect the views of the National Science Foundation, NOAA, or the U.S. Department of Commerce.

7. REFERENCES

Brooks, H. E., 2004: On the relationship of tornado path length and width to intensity. *Weather and Forecasting*, **19(2)**, 310-319.

Donner, W. R., 2007: The political ecology of a disaster: An analysis of factors influencing U.S. tornado fatalities and injuries, 1998-2000. *Demography*, **44(3)**, 669-685.

Rae, S., and Stefkovich, J., 2000: The Tornado Damage Risk Assessment predicting the impact of a big outbreak in Dallas-Fort Worth, Texas. *20th Conference On Severe Local Storms*, American Meteorological Society, Paper 9.6.