Visualization of Tornado Data

Anu Joy, Harshal Chheda, James Chy and Yuxing Sun

Drexel University

ABSTRACT— The aim of this project is to study the nature of tornadoes with the use of information visualization tools. The emphasis of this project is to study the tornadoes that change directions and their effects. For this report we used a subset of published storm data from the National Oceanographic and Atmospheric Administration (NOAA). The data set included many types of storms, but we focused upon just the storm data relating to tornadoes. Both Tableau and IBM Many Eyes were used as the information visualization tools to perform graphical analysis of the data. When visualizing the data, we focused on the BEGIN_AZIMUTH and the END_AZIMUTH parameters in the data, which reflect the change of tornado directions. We found that tornadoes do frequently change direction. Tornadoes that probably went straight seemed to have much less in injuries than ones, which probably changed course. We had difficulties using the visualization tools convey any aggregated information regarding directions in general because the direction was not presented as numeric values in NOAA data. During the preparation of visualizations, we found some skills to change and improve the presentation of data.

1 INTRODUCTION

1.1 Brief History of Tornado Research

Researchers continue to be fascinated by tornadoes and measure and observe the paths tornadoes follow, the months they are occur, and regions in which are tornado prone. Researchers exploring tornado mysteries examine tornadoes from different viewpoints such as the following: genesis, dynamics, detection, warning decision support, forecasting, preparedness and climatology**.** Tornado research targets ways to improve forecasts and warnings in order to help save lives. Officials documented the first observation of tornadic storms in 1953 in Illinois (9 April) and Massachusetts (9 June 1953) by radar at the Illinois State Water Survey and the Massachusetts Institute of Technology (MIT), respectively. Stuart Bigler made other similar observations at Texas A&M University. These early observations showed that hook echoes are related to tornadoes. From the late 1950s onward, and especially in the 1960s and 1970s, a Tornado researcher at the University of Chicago named, **Ted Fujita,** made painstaking analyses of tornado damage, tornado photographs. He integrated these analyses with "mesoanalysis" of surface data. Based on photographs of a Fargo, ND tornado in June 1957, he described tornadoes using terminology such as "wall cloud" and "tail cloud." Because of his work many of these terms persist today. He also categorized tornadoes on a scale based upon wind speed estimates and inflicted damage. Researchers later named this scale the "Fujita Scale" in his honor. The lowest tornado on the scale was an F0 and progressed in strength to F5. In 2007, the NOAA data officially switched to an enhanced Fujita scale to improve the precision of the scale. In 2007 and later, the tornadoes ranged from EF0 to EF5 where EF indicates the "Enhanced Fujita" scale. After viewing tornado damage from aircraft, he documented cycloidal ground marks and associated them with multiple-vortex tornadoes. On the basis of aircraft flights over a tornadic storm in 1977 he suggested that thunderstorm downdrafts generate tornadoes.

The history of tornado research consists of a series of both serendipitous and planned observations and measurements, numerical studies with idealized models, laboratory-vortex experiments, and numerical simulations. Observations have improved in step with advances in technology, most prominently with increasingly more sophisticated radar systems. Numerical simulations continue to improve with increased computer power, speed, and storage capabilities.

1.2 Brief History of Research into Tornado Path Research

Most research suggests that Tornado paths range from 100 yards to 2.6 miles wide and are rarely more than 15 miles long, although some strong tornadoes on record have crossed through multiple states (e.g. the Tri-State Tornado of 1925). Tornadoes can last from several seconds to more than an hour. However, most don't exceed 10 minutes. Most tornadoes travel from the southwest to northeast with an average speed of 30 mph, but the speed has been observed to range from almost no motion to 70 mph. Also the timeframe for tornadoes occur in the Deep South and in the broad, relatively flat basin between the Rockies and the Appalachians, but no state is immune. Peak months of tornado activity in the U.S. are April, May, and June. However, tornadoes have occurred in every month and at all times of the day or night. A typical time of occurrence is on an unseasonably warm and sultry spring afternoon between 3 p.m. and 9 p.m.

Emergency service personnel notify the public of a tornado in their area through many channels. In a survey of an Oklahoma City tornado on May 3rd, 1999, survivors indicated they took safety precautions when they heard the tornado warning on television, a telephone call, warning sirens, or from AM/FM radio. Weather band radios were another source, but few of those survivors indicated a weather band radio provided their first warning notice. Many survivors received the warning from multiple sources. Those with injuries stayed in their homes after receiving the warning. The survivors without injuries left their home to enter a personal or public storm shelter [9]. Other researchers of that same outbreak indicated that relatively inexpensive structural improvements could greatly improve safety to those taking shelter in their homes or in other buildings [4] We unable to find any documentation indicating that emergency services notified people based upon the direction the tornado was heading at the time it first touched down. The direction a tornado around Kellyville, Texas, on June 8th, 1995 had not only a distinct cusp in its path, but also followed a sine-like wave during a portion of its travel [15]. Researchers also identified unusual damage tornado paths as they alternate side-to-side within a mesocyclonic region, a large band encompassing the parent storm during a rare Massachusetts tornado on May 29rd, 1995 [14].

1.3 Brief History of Research into Predicting Tornado Path

Studies have found that tornadoes usually travel at the same speed and in the same direction as their parent thunderstorms. In a pioneering research paper published in 1987, Dr. Tetsuya Fujita (1920-1998), the severe weather researcher, categorized the known direction of 17,081 tornadoes in the U.S. for the period 1916-1985. He found that 59 percent of those tornadoes traveled from the southwest; 19 percent from the west; 11 percent from the northwest; 6 percent from the south; 2 percent from the southeast; and 1 percent each from the east, north and northeast.

1.4 Research into Predicting when a Tornado might change its path

Two studies about tornadoes, which changed directions, made some interesting points in predicting when a tornado might change its path.

One is about the tornado occurred on 5/15/2013 in Granbury, TX [23]. The tornado was a strong one, rated as EF-3. At first, it moved along southeaster but later turned about 110 degree left to due north. TX tornado was compared to a Greensburg, KS tornado, which also turned its path to, left and did a full 360-degree. The study indicated that both tornadoes experienced a "mesocyclone core evolution": When the cold air from the rear flank downdraft wrapped around the tornado completely, it stopped the warm air coming into the tornado. Then the tornado occluded and turned left, and a new mesocyclone would regenerate to its southeast. The study pointed out that it appears a tornado turns left only when it decays rapidly.

The second study is about a tornado in El Reno, Oklahoma on 5/31/2013, which had a weak speed wind in the middle level [8]. It also changed direction and took a left turn. The study compared El Reno tornado with other similar weak mid level wind tornadoes and found out that these tornadoes all had weird paths and inflicted more damage. So the study inferred that the weaker winds appear to leave tornadoes vulnerable to other influences closer to the surface, resulting in taking a left turn. Questions for future studies could focus on forecasting weak middle level tornadoes and observing their paths to test the inference.

These two studies happen to hold the similar view. Both viewpoints can be interpreted as follows: the change of tornado's direction will occur only when its wind speed is weak, no matter it is weak from the beginning or occluded by rear flank downdraft cold air. And it takes left turn in north Hemisphere in most cases. Thus in the further research area we should pay attention to those weak wind tornadoes to test this hypothesis.

1.5 Existing Visualizations of Tornado Data

The tornado history project [20] is a free, searchable database of the tornadoes, which are reported in the US between 1950 and 2012. Each tornado is mapped out on the map of US and has a forum associated with it, which allows users to comment on the tornado data. The map displays the intensity of each tornado along with its path. A table below the map shows a summary of the tornado injuries, fatalities, longest path and the widest path for each tornado. While the visualization provides a method to view each individual tornado, it does not provide any mechanism to compare data between various tornadoes.

An interesting visualization in Tableau created by Alex Kerin [22] showed a very detailed multiple visualizations of Tornado data from NOAA, superimposed on state area and population data. A number of tabs display various tornado data visualizations from different perspectives. Each section of the visualization has filters that can be applied to zero in on a specific piece of information.

A New York Times tornado map [6] from 2011 portrayed tornado path and showed tornado deaths by county, spread over the years 1950-2011. The number of deaths associated with the tornado was depicted with the help of circles. The visualization showed 2011 as one of the deadliest year for tornadoes.

Another New York Times visualization [5] mapped the number of tornadoes each hour between the time periods of April 21 to April 28 on the map of US. The map is motion animated and depicts the distribution of tornadoes over the given time period. The map uses color to depict the number of tornadoes.

1.6 Existing Visualization of Tornadoes changing direction

Data visualization expert John Nelson created two visualizations associated to paths and directions of tornadoes. The first one termed 'Tornado Travel Map' showed a breakout of the directions in which tornadoes traveled over the last 63 years. The idea behind this visualization was driven by the prospect of the average tornado bearing as input to divide the country into tornado warning zones. He used the compass to provide an overall reference of tornado travel in the visualization. The data was extracted from the archived tornado data available in the NOAA/NWS Storm Prediction Center sites. The most significant insight provided by this visualization was that most storms tend to travel in the northeastern direction along with the prevailing winds. Other observations included that tornadoes that move in the northwest direction mostly occur in the western Great Plains and along the Atlantic Coast. Also, the ones that travel southeast cluster heavily in the northern plains states than in the south.

The second visualization called 'Tornado Tracks' traces the paths of tornado for 56 years. The tracks in this visualization are categorized by their F-scale, with the brightest strokes representing the most violent storms. The data for this visualization came from NOAA via Data.gov website. Here he focused on the tracks of the tornadoes over the years, showing the most destructive tornadoes as measured by injuries, fatalities and property damage. The two most notable pieces of information that jump out from this visualization is the existence of the 'tornado alley' in southeastern part of US and the catastrophic 2011 tornado season. We also found this visualization interesting because we saw that most tornadoes seemed to follow a straight path. This led us to examine the data from NOAA to check for anomalies, which led us to explore tornadoes that change direction.

2 TOOLS

2.1 NOAA Storm Data

2.1.1 History of the Data

The NOAA (National Oceanic and Atmospheric Administration) receives the storm data from the National Weather Service. The National Weather service receives their information from a variety of sources. Theses sources include the following: county, state and federal emergency management officials, local law enforcement officials, sky warn spotters, NWS damage surveys, newspaper clipping services, the insurance industry as well as the general public.

Due to the amount of time it takes to collect, validate, and enter post-storm data, the National Climatic Data Center (NCDC) regularly receives Storm Data from the National Weather Service (NWS) approximately 60-90 days after the end of the data month. The NWS has 60 days to submit their data files to the NWS Headquarters in Silver Spring, MD. The NWS Headquarters (NWSHQ) then collects all of the data files from the 124 NWS Forecast Offices (NWSFO). The NWSHQ then uses several algorithms to consolidate the product into one database. The NCDC receives a copy of this database approximately 75-90 days after the end of the month, prepares a publication and an archive. The NCDC then updates the Storm Events Database within 90-120 days after the end of the month.

2.1.2 Content of the Data

NOAA Storm Database has information on various types of storms recorded in each county. Storm Data is an official publication of the National Oceanic and Atmospheric Administration (NOAA), which documents the occurrence of storms and other significant weather phenomena. Some of these storms had sufficient intensity to cause loss of life, injuries, property damage, or disruption to commerce. In addition, it is a partial record of other significant meteorological events, such as record maximum or minimum temperatures or precipitation that occurs in connection with another event. The media, law enforcement, government agencies, private companies, and individuals contribute information into the NOAA storm database beyond the National Weather Service (NWS) data. NOAA is unable to verify all of the data it receives through the NWS because of time and resource constraints. Therefore, when using information from Storm Data, customers should be cautious, as the NWS does not guarantee the accuracy or validity of the information. Further, when it is apparent information appearing in Storm Data originated from a source outside the NWS (frequently credit is provided); Storm Data customers requiring additional information should contact that source directly. In most cases, NWS employees will not have the knowledge to respond to such requests. In cases of legal proceedings, federal regulations generally prohibit NWS employees from appearing as witnesses in litigation not involving the United States.

The data extracted from the NOAA database contained information like the begin time and end time of the tornado. The fields captured the start and end time of the tornado. The NOAA data included a field called BEGIN_AZIMUTH, which reflected the direction the tornado was traveling when it first appeared and another called END_AZIMUTH, which reflected the direction the tornado was heading when it ended. We used these fields to determine if a field went straight or changed course. Also, for each tornado, the NOAA data captured numbers of deaths and injuries and further breaks the data down between whether the death or injury occurred during the storm (direct) or after the storm (indirect). This data is useful to capture the deaths and damage caused when tornadoes change the directions. Along with that the event narrative field is used to capture and provide detailed about the intermediate locations where the tornado turned sharply about the intermediate locations where the tornado turned. The NOAA data also captured the latitude and longitude of the tornado start and end locations as well as the name of the state in which it started and ended.

The NOAA data captured tornado length in miles tornado width in yards. The azimuth fields were in directions on a 16-point compass scale and the descriptions included this reference to provide relationships from a known landmark, usually a town or city, such as 4.5 miles ESE Atlanta. The NWS uses a database of over 106,000 cities and towns including their latitudes and longitudes. Using an algorithm, the location 4.5 miles ESE of Atlanta can be derived from the known latitude and longitude of Atlanta. These latitude and longitude pairs are generated by the NWS and populated into the database. The latitude and longitude are in decimal degrees format.

The data contains descriptive information about the times, the direction of tornados or locations, and severity of destruction of property, trees, crops, power lines, roads, bridges, etc. Additionally, a brief summary of fatalities and injuries are also recorded. Below are the few categories on which data is collected.

Direct Fatalities/Injuries

- Structures or trees were blown over and landed on someone, resulting in a fatality/injury.
- People became airborne and struck the ground or objects, resulting in a fatality/injury.
- High voltage power lines were blown onto a car, killing or injuring those inside.
- A high-profile vehicle was blown over, resulting in a fatality/injury.
- A vehicle was blown into structure or oncoming traffic, resulting in a fatality/injury.
- Objects became airborne (debris, missiles), resulting in a fatality/injury.
- A boat on an inland lake or river was blown over or capsized, resulting in a drowning.

Indirect Fatalities/Injuries

- A person was killed or injured after running into a tree downed by the tornado.
- Someone was electrocuted by touching downed power lines.
- Someone suffered a heart attack and died as a result of removing debris.

Property damage estimates are entered as actual dollar amounts, by using reasonably accurate estimate from an insurance company or other qualified individual based on availability. If this estimate is not available, then the data is categorized into two categories: either check the “no information available” box, or make an estimate.

2.1.3 Fujita scale

A Fujita Scale was implemented by the National Weather Service in 2007 to rate tornadoes in a more consistent and accurate manner. The EF-Scale takes into account more variables than the original Fujita Scale (F-Scale) when assigning a wind speed rating to a tornado, incorporating 28 damage indicators such as building type, structures and trees. For each damage indicator, there are 8 degrees of damage ranging from the beginning of visible damage to complete destruction of the damage indicator. The original F-scale did not take these details into account. The original F-Scale historical database will not change. An F5 tornado rated years ago is still an F5, but the wind speed associated with the tornado may have been somewhat less than previously estimated. A correlation between the original F-Scale and the EF-Scale has been developed. This makes it possible to express ratings in terms of one scale to the other, preserving the historical database.

2.2 Types of Records

The tornado records are entered based on the type of event. Since Verification of tornado is difficult its existence is based on available evidence and once established a Tornado event occurred, the data is entered as an event in Storm Data. An event will be characterized as a tornado if the type or intensity of the structural and vegetative damage and/or scarring of the ground could only have been tornadic, or if any two of the following guidelines are satisfied:

- Fairly well-defined lateral boundaries of the damage path;
- Evidence of cross-path wind component, e.g., trees lying 30 degrees or more to the left/right of the path axis (suggesting the presence of a circulation);
- Evidence of suction vortices, ground striations, and extreme missiles; or
- Evidence of surface wind convergence as suggested by debris-fall pattern and distribution. In fast-moving storms, the convergence pattern may not be present and debris pattern may appear to fall in the same direction.

Additionally, an event will be characterized as a tornado if:

- Eyewitness reports from credible sources, even with little or no structural or vegetative damage, and/or little or no scarring of the ground, indicate that a violent circulation extended from the convective cloud base to the ground; or
- Videotapes or photographs from credible sources, even with little or no structural or vegetative damage, and/or little or no scarring of the ground, indicate that a violent circulation extended from the convective cloud base to the ground.

2.3 Data format

The database currently contains data from January 1996 to November 2013, as entered by NOAA's National Weather Service (NWS). This data can be extracted in CSV format.

2.4 Tableau

Tableau is a desktop visualization application from Tableau Software. Tableau provides visualizations as text tables, heat maps, highlight tables, symbol maps, filled maps, pie charts, horizontal bars, stacked bars, side-by-side bars, tree maps, circle views, side-by-side circles, discrete and continuous lines, dual lines, discrete and continuous area charts, dual combination charts, scatter plots, histograms, Gantt views, bullet graphs, and packed bubbles.

2.5 IBM Many Eyes

Many Eyes is IBM's visualization tool that allows users to perform graphical analysis with a community. The concept behind the tool is to encourage users to generate visualizations and to share and discuss information. It creates a web community that connects visualization experts, citizen scientists, academics and others. The site has over 450,000 data sets to choose from and many of these data sets already have visualizations created. The user can upload and post data by creating a free account on the site, although one does not need to create an account to simply browse through visualizations or data. The software accepts all kinds of data— even the contents of a spreadsheet without having the user worry much about formatting. For most visualization, the basic formatting requirement is that the data must be in a tab-separated text file with column headers in the first row. After the data is uploaded and the Many Eyes software approves the data set, the user can click the 'visualize' button to create visualizations. Data can be visualized in many ways, such as scatter-plots, bar charts, tree maps, phrase nets, word clouds, plots, network diagrams and geographic maps. The Many eyes software is easy to use and requires minimal formatting of data. The site also provides very good documentation of all possible types of visualizations and suggestions of the types of data each visualization is suited for. The features and tasks available for each type of visualization are also well documented.

One drawback with the software is that once datasets are uploaded to Many Eyes, they are made available to all users and the visualizations produced become the property of IBM. The data and visualizations can be easily downloaded, shared and commented upon by other users. So, while it offers a lot of options in terms of visualizations, Many Eyes is not appropriate for use with confidential datasets. In addition, it does not allow any significant customization and also limits size of data sets to 5Mb.

3 METHODS

We downloaded the NOAA Storm Watch data from http://www.ncdc.noaa.gov/stormevents/. NOAA provided this data at plain text files. NOAA provided one storm file for each year. Each data file was in Comma Separated Values (CSV) format. The first line of each file had a comma-delimited list of labels, and then each line had the data for each storm event delimited by commas. There were 18 files for the years 1996 to 2013. The total combined volume of these files was 767 megabytes. We opened each file in Microsoft Excel 2013 and saved each file as an Excel format file (extension xlsx). In Excel, the CSV data could be treated like a database table. In each Excel file we sorted the rows by EVENT_TYPE, selected and deleted all rows prior to the rows where the EVENT_TYPE was "Tornado", and selected and deleted all rows after the rows where the EVENT_TYPE was "Tornado". We re-sorted the results dropping the blank rows to the bottom of each Excel file then resaved the files. We then loaded each Excel file into a single table called Tornadoes within a Microsoft Access 2013 database (extension accdb). We ran a delete query against the Tornadoes table to delete all blank records. This step gave us one table with a consolidated list of every tornado logged by NOAA between 1996 and 2013. We converted this database to Microsoft Access 2002-2003 format (extension mdb) to support Tableau. We still found Excel easier to use, so we exported the Tornadoes table as an Excel 2013 spreadsheet (back to xlsx). To support Many Eyes, we kept the first 15 years of data and dropped years 1996, 1997, and 1998 from the Excel spreadsheet. Many Eyes had a 5-megabyte limit; 18 years worth of NOAA tornado data came out to 5.7 megabytes; 15 years worth of data came out to about 4.7 megabytes of data. We went back and adjusted the queries in Tableau to only use the data after 1998 to keep visualizations between Many Eyes and Tableau synchronized.

When looking at this data with the idea of turning tornadoes, we wanted the data to provide some insight about tornadoes and identified the following questions: Do tornadoes tend to go straight or turn? Do tornadoes that turn cause more deaths and injuries? Has the frequency of turning tornadoes fluctuated over time? Are there any patterns in turning tornadoes? For example, do they tend to start off going North then turn East? Do stronger tornadoes tend to turn one direction, but weaker ones another direction? Is there any relationship between path length, path width, and likelihood of turning? Do turning tornadoes follow well-worn tracks?

Many Eyes provides an array of visualization options for the tornado data. The path of tornadoes that changed directions can be plotted on a map of USA using the BEGIN_AZIMUTH and the END_AZIMUTH fields in the data. This will provide an understanding of the specific regions over which such tornadoes occur. This data can then be superimposed on the topography map of the area to show any possible relationship between topography and change in direction of tornadoes. Line graphs or stack graphs could be used to conclude trends in tornado occurrences over the years or by state. Other interesting information about the tornado data such as comparisons of damage caused by tornadoes that change directions to ones that go straight, comparisons of fatalities or tornadoes intensities etc., also can be plotted using the various types of visualizations available in Many Eyes.

For Tableau, we wanted to choose visualizations which could show tornadoes which change direction in relation to time, location, starting direction and compare those to tornadoes which ended up going straight. We also expect that showing state-by-state statistics may be helpful to show impacts of tornadoes to a general audience.

4 RESULTS

4.1 Continental US Tornadoes With Some Injuries

The NOAA tornado data included latitude and longitude of the start of the tornado, but several errors required cleanup before they could be used in the visualization. Some of the longitudinal entries were indicated with a positive sign and others with a negative sign. The only portion of the United States with a positive longitude would be the tail of the Alaskan Aleutian Islands. We converted all positive longitudinal entries to be negative. Some entries had values that could not be converted to a number. We excluded these entries. Some latitude and longitude entries were miskeyed with the decimal point in the wrong place. If the latitude was greater than 900, we divided the number by 100, else if it was greater than 90, we divided the number by 10, otherwise we left the number alone. If the longitude was less than -170, then we divided by 10, else if the longitude was greater than -65, then we multiplied by 10, otherwise we left the number alone. Showing all tornadoes made the graphic too hard to read. Again, we only included data between 1999 and 2013. In *Figure 1*, showing just tornadoes with at least one death made the graph too sparse. A good balance appeared to be to include all tornadoes with at least one injury and then indicate the size by the number of deaths. The deaths were direct deaths; indirect deaths were much higher. Coloring by state made the state coverage apparent but not overwhelming. Tornadoes in Alaska and Hawaii were not without injury but it was too hard to place them in the same map without losing the detail over the data in the Continental states. The direct death toll from the 2011 Jasper County, Missouri tornado was 158; the number of direct injuries was much higher at 1158. There may be some confusion about whether the graph is showing injuries or deaths. This image also does not take into account tornado direction or tornadoes that changed direction while traveling.

4.2 Direct death associated with different Tornado_F_Scale and directions

The interactive visualization in *Figure 2* was created using the tool Matrix Chart in Many Eyes. Rows represent different Tornado_F_Scale and columns represent begin azimuth of tornadoes. Different colors denote end azimuth of tornadoes. Sizes of bars indicate direct death in tornadoes with number attached to the bar. When I move mouse over a bar, the exact value of the bar appears which contains direct death numbers and proportions of each end azimuth with the same begin azimuth. The most direct death number is 158 presented by the largest bar. It is one strong tornado with a Tornado_F_Scale of EF 5 and its begin azimuth is SE with an end azimuth of SSE. I can also highlight an item by clicking it.

4.3 Direct Deaths caused by Tornadoes

Figure 3 shows an interactive information visualization which is created using Bubble chart style by using ManyEyes Tool. This bubble chart takes a table with one text column, for labels, and one or more numeric value columns, to create the chart. The table above "State" is the label column and "Direct_Deaths" is the value column. This visualization gives the total direct death across years 1999-2013. The visualization gives the breakup of deaths per month across the years. It also uses different color codes to represent different months. The Visualization is useful as it helps in answering the question the damage done by tornadoes across the years. The use of different sizes of circles answers that question as it gives a clear picture that the maximum death till date has been in the state Alabama.

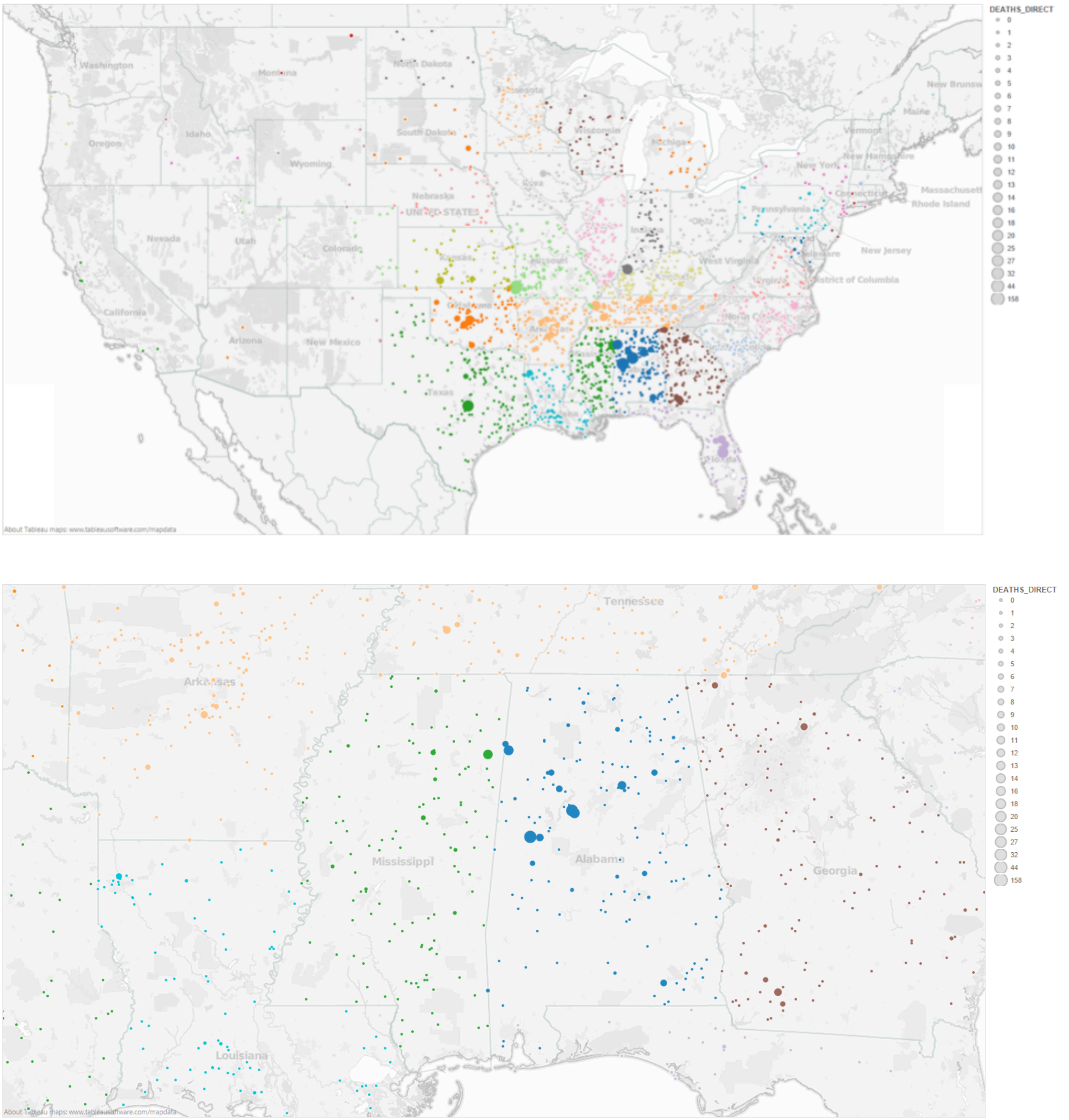
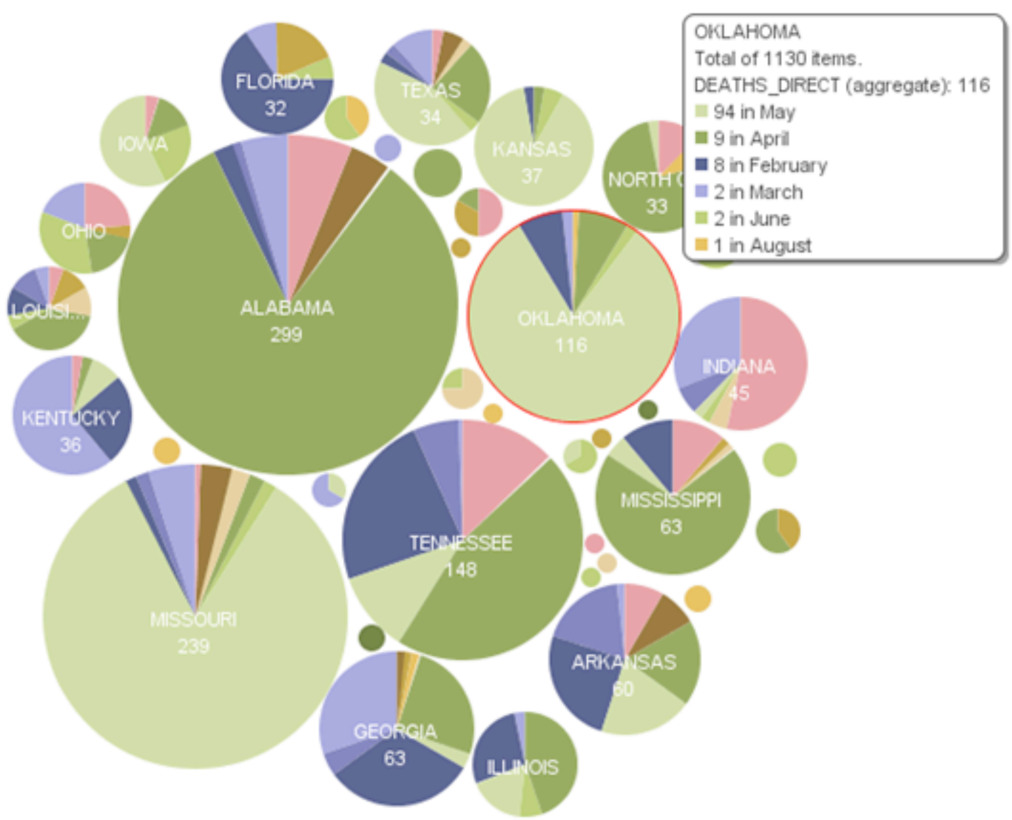


Figure 1: Continental US Tornadoes With Some Injuries - Circle Size Indicates Deaths

Figure 2: Direct death associated with different Tornado_F_Scale and directions



To

Figure 3: Direct Deaths caused from Tornadoes

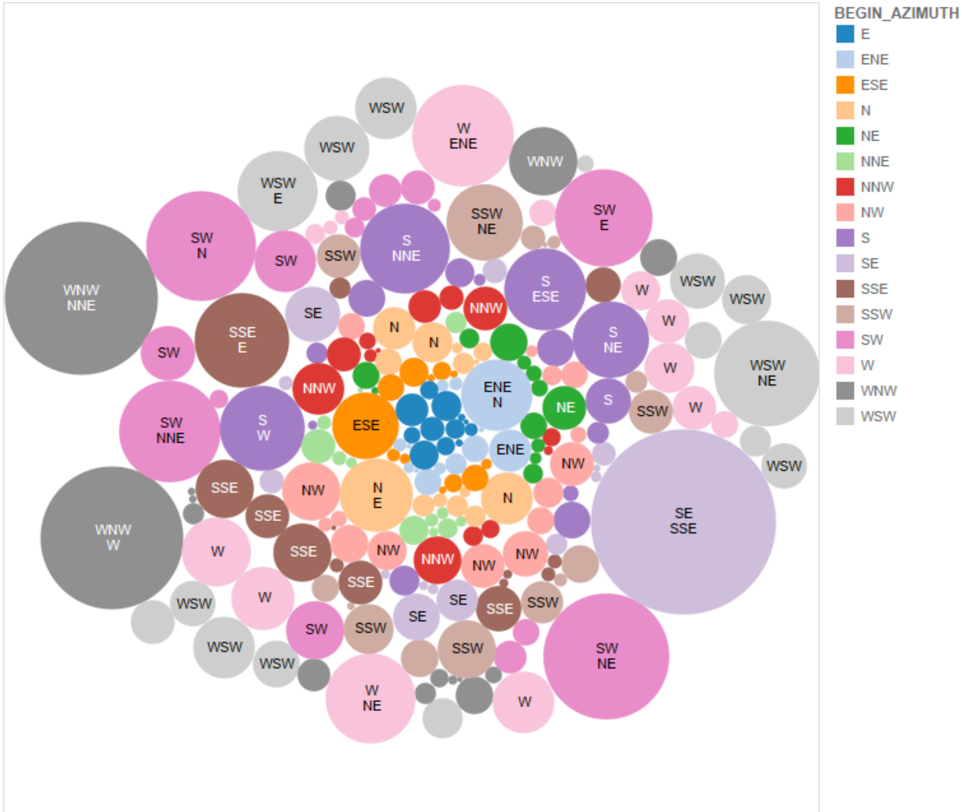


Figure 4: Direct Injuries from Tornadoes which Changed Direction

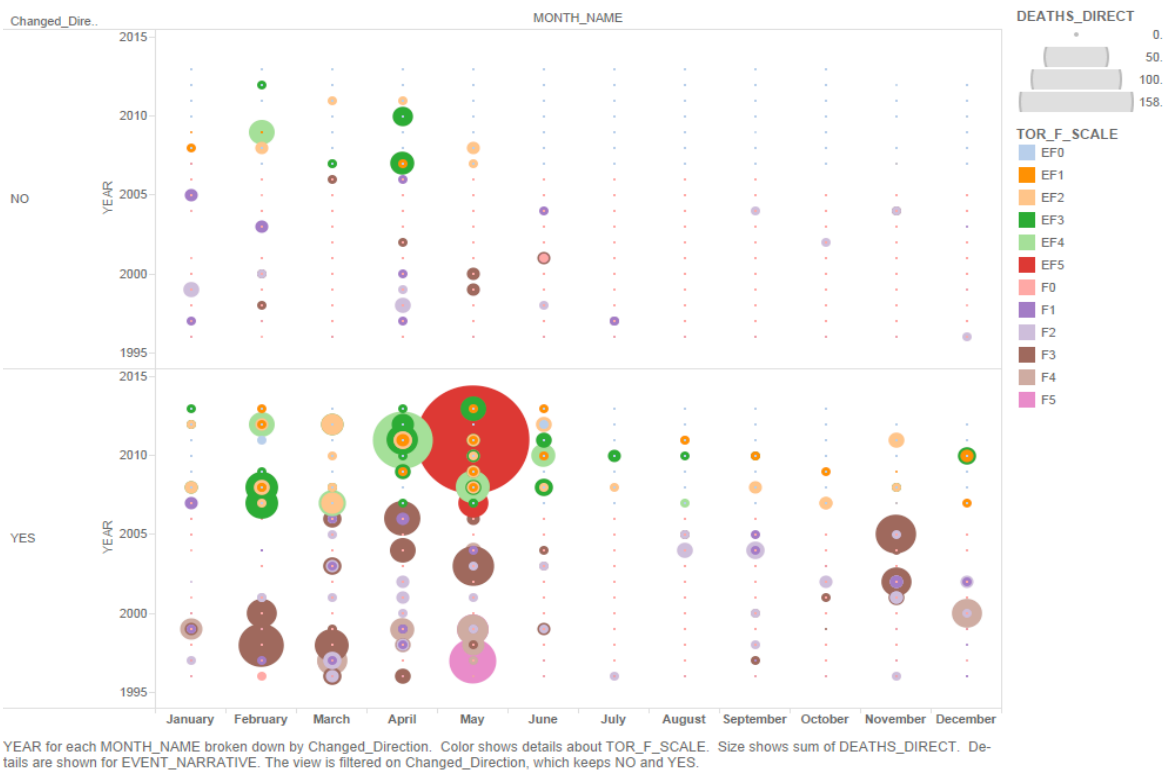


Figure 5: Direct Injuries from Tornadoes which Changed Direction

4.4 Direct Injuries from Tornadoes which Changed Direction

For this visualization in *Figure 4*, we excluded all records where the BEGIN_AZIMUTH was blank, the END_AZIMUTH was blank, and where BEGIN_AZIMUTH was different from the END_AZIMUTH. Because of file size limitations in ManyEyes (5 megabyte), our team also agreed to filter the data to years 1999 and 2013. Of the 20,944 tornadoes documented by NOAA between 1999 and 2013, 1,790 had no BEGIN_AZIMUTH or END_AZIMUTH listed, and 1,875 had one but not the other. The remaining 17,279 tornadoes had 7,637 where the BEGIN_AZIMUTH equaled the END_AZMIMUTH. The visualization covers the remaining 9,642 tornadoes. The top value is the BEGIN_AZIMUTH and the bottom value is the END_AZIMUTH. The bubble color is the BEGIN_AZIMUTH and the bubble size is the sum of the number of DIRECT_INJURIES.

4.5 Relationship between direct injuries and tornadoes that change direction

This interactive visualization in *Figure 5* was generated in Tableau. We first started out by plotting direct tornado injuries over years on the y-axis and then spread this out over months on the x-axis. The resulting graph showed trends in direct tornado injuries for different months of the year. The size of the circle shows the number of direct injuries. In order to understand whether tornadoes that changed direction had an impact on the number of injuries, we computed a new value called Change_Direction by comparing the BEGIN_AZIMUTH and the END_AZIMUTH values for tornadoes. The rows with null values in either of these columns were filtered out of the dataset. The new calculated column was created with the formula IIF([BEGIN_AZIMUTH]! =[END_AZIMUTH],'YES','NO'). Adding this to the visualization created two separate areas in the graph: one showing direct injuries for tornadoes that changed direction and the other for the tornadoes that did *not* change direction. In order to understand what types of tornadoes caused the direct injuries, we included Fujita scale for each of the tornado in the visualization. The colors in the graph represent the magnitude of the tornado.

5 INTERPRETATION

5.1 Direct Deaths from Tornadoes which Changed Direction

Once the filtering was completed by the selected year and by removing blank azimuth values, we found it striking that the number of tornadoes that changed direction actually exceeded (slightly) the number of tornadoes that went straight. Of the 17,279 tornadoes, 55.8% changed direction. The visualization indicates that when tornados changed direction from South-East to South-South-East, it caused the most number of deaths (159 deaths).

5.2 Continental US Tornadoes With Some Injuries - Circle Size Indicates Deaths

The overall clustering of tornadoes in the MidWest states was not surprising. What surprised us was that the density was not as prevalent around Kansas and Oklahoma. The center seemed much further East than expected and there were far more tornadoes in the Southern states like Louisiana, Mississippi and Alabama. Looking closely at the visualization, we also saw a path going from North-East Louisiana all the way to the Southern tip of the Appalachian Mountains. Tennessee had several tornadoes with a high death count. The highest immediate death count was in Jasper County, Missouri on the corners between Kansas, Oklahoma, Missouri and Arkansas. Other tornadoes of importance include the Southern tip of Indiana, Central Florida and Central Iowa. There is also a tight cluster around Central Oklahoma.

5.3 Deaths caused due to Tornadoes

These visualizations have been helpful to interpret the death that has been caused due to tornadoes. They show the increase or decrease in death rate across the years to take necessary actions. Since many parameters have to be considered by the researchers to come up with a prediction on the patterns such visualizations can play a key role to give multi-parameter view, which will help them in their predictions. Sometimes due to multiple parameter consideration can also lead to incorrect predictions that on the other hand lead to damage on a large scale.

5.4 Direct death associated with different Tornado_F_Scale and directions

It is quite evident that tornadoes with large F_Scale kill more people. The bars with F_Scale of EF3, EF4, EF5, and F3 are larger than other F_Scale tornadoes. As discussed in former section, most tornadoes travel from the southwest to northeast. Based on this conclusion, let's take a look at the visualization. On the x-axis lies different tornado begin azimuths. We can easily notice that the sizes of bars are quite large with begin azimuths of S, WSW, SW, SSW, and W, which shows that direct death numbers are relatively high compared to other begin azimuths. Also, when we click the end azimuths to highlight them, we can see that direct death numbers are very high with end azimuths of ENE, NE, NNE, and N. The result verifies the previous research conclusion. Those large pure color parts of bars indicate big numbers of direct death associated with the same end azimuths. When focusing on those large pure color parts, we can see that most of their end azimuths are different with their begin azimuths. So it demonstrates that tornadoes that change directions indeed have large destructive power. One thing that is easily confused is that each color does not represent a different tornado but represents tornadoes with the same end azimuths.

5.5 Relationship between direct injuries and tornadoes that change direction

The one thing that stood out in this visualization is the perceivable difference between the number of injuries caused by tornadoes that changed direction and those that did not. The finding was consistent with the results from the other visualizations: tornadoes that change directions cause more direct injuries than the ones that did not. The tornado that resulted in most number of injuries was an EF5 tornado in May 2011, shown by the biggest red circle on the graph.

6 DISCUSSION

6.1 Direct Deaths caused from Tornadoes

From a predictive point of view, we cannot expect future tornadoes that start going South-East then changing to South-South-East to cause significant damage in the future. That data point in the visualization is skewed by a single powerful tornado directly killing 158 people on May 22, 2011 in Jasper County, Missouri. This tornado caused the most direct deaths in the entire data set.

6.2 Continental US Tornadoes With Some Injuries - Circle Size Indicates Deaths

Using Tableau for the visualization still required a bit of a learning curve as it may be unclear what Tableau considers a Dimension and what it considers a Measure. Pre-processing the data was also important to get the data just right. We were really unable to find any visualizations that could really address tornado direction in general, let alone changes in direction. For example, it would be hard to determine if a tornado started going North but ended going South, whether it went toward the East or to the West to get there. Using the latitude and longitude to see the overall path might help but it would have been tricky to pre-process the data to work out those calculations. Those calculations would have worked out to be a guess on the actual tornado path. With some work, we probably could have come up with a circular estimate of the path, but with the research we found on the mesocyclone, a circular path might not have been very accurate anyway. A bigger question might be whether mesocyclonic regions change direction.

6.3 Deaths caused due to Tornadoes

Our approach was to come up with a problem statement that could be resolved with the help of visualizations. So the aim for this project was to come up injuries and death caused by tornadoes due to change in directions or any other parameters. These visualizations have been created to resolve this issue by giving multifaceted view by considering various fields like INJURIES_Direct, DIRECT_DEATH, BEGIN_AZIMUTH (which is the tornado direction prediction). This visualization was straightforward and is easy to understand. Since most of the data displayed is numeric with the use of colors for different months and circle sizes for easy visibility. Bigger the circle more is the death count for that state and also broader color code represents the months that lead to highest damage.

6.4 Direct death associated with different Tornado_F_Scale and directions

This visualization is not very straightforward actually. However, it is not hard to understand if we pick the right perspectives to view it. That is because the most important data begin azimuth and end azimuth, which are related to the tornado directions, are not numeric value data. So that many tools in Many Eyes can't be applied to visualize the data. But with this visualization, we can still draw some valuable conclusions. We found out that tornadoes with high F_Scale do more damages. By highlighting different begin and end azimuths, we verified the conclusion of tornadoes general moving direction. Also, by comparing azimuths and death numbers, we found tornadoes that change directions have large destructive power.

6.5 Relationship between direct injuries and tornadoes that change direction

Tableau provided much more flexibility in terms of data manipulation than Many Eyes. Many Eyes required a fixed data set to be uploaded to create visualization. Tableau on the other hand, allowed the data to be filtered and manipulated (e.g., change in direction field) during the creation of the visualization. Tableau required some getting used to and was confusing in terms of the difference between measures and dimensions.

7 CONCLUSION

Our primary goal was to investigate and provide visualizations to show how the path a tornado took after touchdown affected those in its path. We were able to find a variety of sources of research into tornado paths. Some research indicated that tornadoes frequently do not follow a straight path once they hit the ground. We expected to be able to take the directions identified in the NOAA source data to make guesses into the path followed. We also expected to find that emergency services would try to alert those in the direction the tornado was going at touchdown and assume it would go in that same direction. A tornado changing direction would go in an unexpected direction, leaving those in the new path less prepared leading to more injuries and casualties. What we found was that tornadoes do frequently change direction but that emergency services appear to broadcast broadly without taking the initial direction into account. We can only speculate why the visualizations comparing tornadoes with the same start and end azimuth (probably went straight) seemed to have much less in injuries than ones where these values were different (probably changed course). We had difficulties using the visualization tools convey any aggregated information regarding directions in general, let alone change in directions. Many of the visualizations expected linear input such as linear time, length, width, but the visualizations did not really lend themselves to an angular direction let alone directions indicated with North, South, East or West. ManyEyes is a great free tool. Tableau is excellent product for those creating professional visualizations. Future work and experience with these tools would greatly improve our abilities to generate more sophisticated visualizations. When preparing our final visualizations, our team identified modifications which would improve the foreground of the data from background, bring the interesting items pop out from the background distractors, improve consistency through the paper and to improve the flow of the story we were trying to portray. We found it very easy to change visualizations within the tools by applying zoom and filter, changing major characteristics, and by looking at details on demand. One major characteristic we wanted in consistency was whether we were showing deaths or injuries. The NOAA data had both and even included indirect forms of both deaths and injuries.

8 REFERENCES

[1] M. Yuan, M. Dickens-Micozzi, and M. A. Magsig, "Analysis of Tornado Damage Tracks from the 3 May Tornado Outbreak Using Multispectral Satellite Imagery," Weather and Forecasting, vol. 17, no. 3, pp. 382–398, Jun. 2002.

[2] M. S. Passe-Smith,, "Exploring Local Tornado Alleys for Predictive Environmental ...," esri.com. [Online]. Available: http://training.esri.com/bibliography/index.cfm?event=general.recorddetail&id=83263. [Accessed: 16-Mar-2014].

[3] "IBM Advanced visualization," 03-Jan-2014. [Online]. Available: http://www-01.ibm.com/software/analytics/many-eyes/. [Accessed: 16-Mar-2014].

[4] C. A. Doswell and H. E. Brooks, "Lessons Learned from the Damage Produced by the Tornadoes of 3 May 1999," Weather and Forecasting, vol. 17, no. 3, pp. 611–618, Jun. 2002.

[5] A. McLEAN and A. TSE, "Map of the Tornadoes Across the South - Interactive - NYTimes.com," 28-Apr-2011. [Online]. Available: http://www.nytimes.com/interactive/2011/04/28/us/map-of-the-tornadoes-across-the-south.html?_r=0. [Accessed: 16-Mar-2014].

[6] "Map of Tornado Deaths in the U.S. From 1950 to 2011." [Online]. Available: http://www.nytimes.com/interactive/2011/04/28/us/tornado-deaths.html?_r=0. [Accessed: 16-Mar-2014].

[7] J. Nelson, "Relative proportion of historic tornadoes by their direction of travel," *Tornado Travel Map*, 12-Jun-2013. [Online]. Available: http://uxblog.idvsolutions.com/2013/06/tornado-travel-map.html. [Accessed: 16-Mar-2014].

[8] S. Finger, "Researcher finds possible clue to some tornadoes' paths," Wichita Eagle. [Online]. Available: http://www.kansas.com/2014/03/03/3321809_researcher-finds-possible-clue.html. [Accessed: 16-Mar-2014].

[9] B. Hammer and T. W. Schmidlin, "Response to Warnings during the 3 May 1999 Oklahoma City Tornado: Reasons and Relative Injury Rates," Weather and Forecasting, vol. 17, no. 3, pp. 577–581, Jun. 2002.

[10] J. Nelson, "Sixty One Years of Tornado Tracks, by F-scale," *Tornado Tracks*, 23-May-2012. [Online]. Available: http://uxblog.idvsolutions.com/2012/05/tornado-tracks.html. [Accessed: 16-Mar-2014].

[11] P. W. Suckling and W. S. Ashley, "Spatial and Temporal Characteristics of Tornado Path Direction*," The Professional Geographer, vol. 58, no. 1, pp. 20–38, 2006.

[12] "Storm Events Database - FAQ | National Climatic Data Center." [Online]. Available: http://www.ncdc.noaa.gov/stormevents/faq.jsp. [Accessed: 16-Mar-2014].

[13] "Storm Events Database | National Climatic Data Center." [Online]. Available: http://www.ncdc.noaa.gov/stormevents/. [Accessed: 16-Mar-2014].

[14] L. F. Bosart, A. Seimon, K. D. LaPenta, and M. J. Dickinson, "Supercell Tornadogenesis over Complex Terrain: The Great Barrington, Massachusetts, Tornado on 29 May 1995," Weather and Forecasting, vol. 21, no. 6, pp. 897–922, Dec. 2006.

[15] R. M. Wakimoto, H. V. Murphey, D. C. Dowell, and H. B. Bluestein, "The Kellerville Tornado during VORTEX: Damage Survey and Doppler Radar Analyses," Monthly Weather Review, vol. 131, no. 10, pp. 2197–2221, Oct. 2003.

[16] W. R. Donner, "The political ecology of disaster: an analysis of factors influencing U.S. tornado fatalities and injuries, 1998-2000," Demography, vol. 44, no. 3, pp. 669–685, Aug. 2007.

[17] "The Science of Tornadoes | Weather Underground." [Online]. Available: http://www.wunderground.com/resources/severe/tornadoFAQ.asp. [Accessed: 16-Mar-2014].

[18] H. B. Bluestein, Tornado Alley: Monster Storms of the Great Plains. Oxford University Press, 2006.

[19] J. Jedlinski, "Tornado direction paths | Chicago Weather Center Blog," *Tornado direction paths*, 23-Jul-2010. [Online]. Available: http://blog.chicagoweathercenter.com/2010/07/23/tornado_direction_paths/. [Accessed: 16-Mar-2014].

[20] J. Lietz, "Tornado History Project: Maps and Statistics." [Online]. Available: http://www.tornadohistoryproject.com/. [Accessed: 16-Mar-2014].

[21] S. Brown, P. Archer, E. Kruger, and S. Mallonee, "Tornado-Related Deaths and Injuries in Oklahoma due to the 3 May 1999 Tornadoes," Weather and Forecasting, vol. 17, no. 3, pp. 343–353, Jun. 2002.

[22] A. Kerin, "US Tornadoes 1950 to 2011 « Data Driven: Data Analytics, Dashboard Design." [Online]. Available: http://www.datadrivenconsulting.com/2012/04/us-tornadoes-1950-to-2011/. [Accessed: 16-Mar-2014].

[23] M. Weinberg, "Texas Tornado Makes Very Rare Left Move & Completely Changes Direction! How Can This Happen? - WDRB Weather Blog," 16-May-2013. [Online]. Available: http://fox41blogs.typepad.com/wdrb_weather/2013/05/texas-tornado-makes-very-rare-left-move-completely-changes-direction-how-can-this-happen.html. [Accessed: 16-Mar-2014].