Modern Visualization Techniques for Severe Weather Data Sets, Case Study: Correlation between Tornado and Hail

Faruk Uysa, *Member, IEEE*

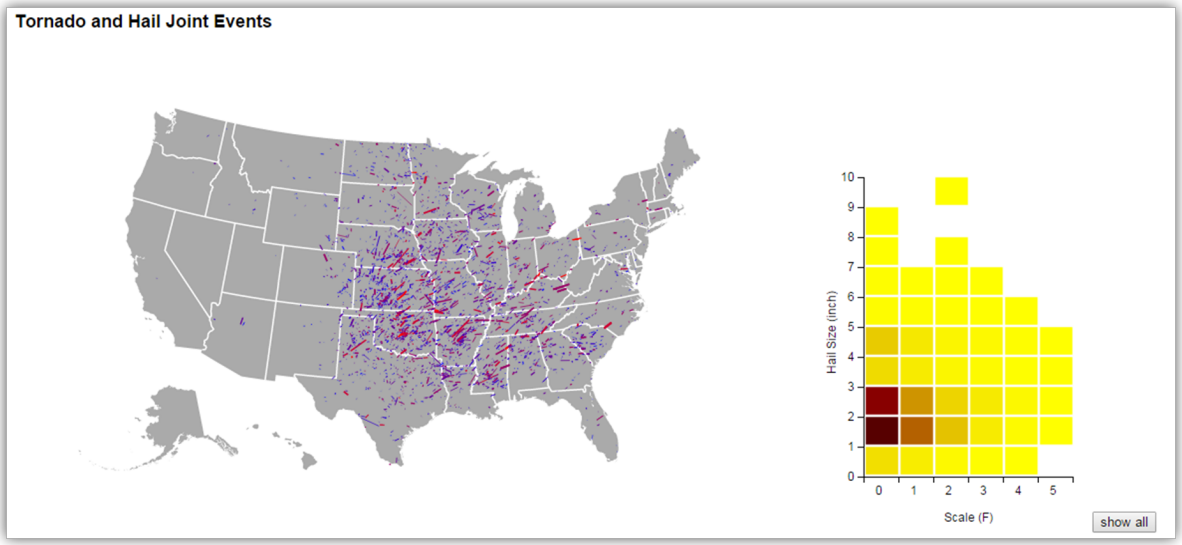

Fig. 1. Visulization Interface.

Abstract—It is difficult to make prediction or understand the results of the severe weather events by just looking the available data sets. Data mining and visualization is necessary to find the related data and emphasize its meaning. In this project, we will work on the modern visualization techniques for the severe weather data sets. We propose a visualization tool that give emphasis of hail and tornado correlation. This tool will have the ability to process and visualize the tornado and hail data sets that will help the meteorologist and researchers to understand the relationship between hails size and tornado scale in US.

Index Terms—Severe Weather, Tornado, Hail, Correlation

1 Introduction

National Oceanic and Atmospheric Administration's (NOAA's) National Weather Service's Storm Prediction Center makes the US severe weather data available for researchers and meteorologist to support the research and development in this field. Unfortunately, it is challinging to make prediction or understand the results of the severe weather events by just looking the available data sets. Thus, data mining and visualization is necessary to find the related data and emphasize its meaning. For example, the location of tornadoes given by latitude and longitude in raw data. By looking the raw data (latitude and longitude), it is not possible to quickly determine the tornadoes' spatial location. Visualize is necessary to make this data meaningful to the end user. To this end, some figures and graphs (such as; charts of the latest severe storm reports; annual summaries; etc.) are presented to explain the important and unique events in NOAA's Storm Prediction Center's Warning Coordination Meteorologist's web page (SPC-WCM) [2]. SPC-WCM's current web page is not designed for interactive user interaction and does not have any effective visualization tool. The graphs and figures are mostly created for single events (to answer a very specific task) and do not give any other information in most cases.

Achieving an information such as correlation of tornado and hail events almost impassible using the raw data files. An interactive tool is necessary to support the tornado and hail joint event information to meteorology's.

- *F. Uysal is with Dept. of Electrical and Computer Engineering, Polytechnic School of Engineering, New York University, 6 Metrotech Center, Brooklyn, NY 11201, USA. E-mail: faruk.uysal@nyu.edu*
- *Submitted as NYU CS-GY 6313 Information Visualization course project report.*

1.1 Related Work:

In literature, there are interactive visualization tools available for tornado data visualization. Some of them are - Twister [4], - Twister Dashboard: Exploring Three Decades of Violent Storms [3], - Interactive Tornado Tracks Map [1] and etc...

The first two visualization tool is freely available on the web. They cover up some of the data variables. Each of them have different unique features. The Twister lets user to visualize the tornadoes for US or selected state. The information of selected tornadoes are displayed using different techniques. Filtering is possible. Twister Dashboard lets user to select the year of interest and display the information about the tornadoes where filtering is also available. The third one Interactive Tornado Tracks Map is not freely available and developed by a professional visualization company. It has very advanced filtering options which lets end-user to select what to visualize for selected tornado data set. In addition to these tools, there are plenty number of graphs in SPC-WCM web page [2] to support the published data. Unfortunately, none of these tools give information about the hails. Thus none of this can visualize the hail and tornado correlation.

2 Data Analysis and Abstraction

Tornado, Wind and Damaging Hail historical data sets are available at SPC-WCM web page. The files are represent the data that is submitted to the Storm Data publication by National Weather Service field offices. Careful review of the data is conducted at the National Climate Data Center and the Storm Prediction Center. We are going to work on tornado and hail data sets which cover the tornadoes and hail events in all states from 1950 to 2015. These are spatio-temporal data sets which include the location of the past events together with occurrence time and some other meteorological information. The data type is table and given as Comma Separated Values (csv). The data sets include multiple attributes (quantitative, ordinal and categorical). Some of the attributes that used proposed project are given in Table 1

For the proposed visualization tool, data fields 2, 9, 11, 12, 13, 14 and 16 through 21 will be used to extract the necessary information. The data field number 2 (Year) is necessary to filter data for selected years and this attribute will be used to control the temporal behavior of data. Field number 11 (F-scale) is determine the F-scale of the tornado. Field number 11 in hail data set represents the hail size in inches. Fields 12, 13 and 14 will be used to extract important information about the selected tornado. Fields 16 through 19 determine the spatial location of the tornadoes and hails.

3 Task Analysis and Questions

In meteorology, it has been known that hail is created by updraft and tornado needs updraft to develop. Hail can also be a warning signal of a tornado. Usually large size hails are observed when strong tornadoes present. As a result of an interview with meteorologist, we decided to investigate the correlation between the tornadoes and hails by synthesizing two separate data sets (tornado and hail data sets) supported by National Weather Services. This tool will help meteorologist to exploit the correlation between the hail size and the tornado intensity (F-Scale) and understand the location of these joint severe weather event. The tasks can be summarized as follow:

Task1: Is there a correlation between hail size and tornado intensity?

Task2: If there is a correlation in which areas this special weather event occurred?

4 Visualization and Interaction Design

Two data sets tornado and hail are processed jointly to match the events created by same storm. The relation between the hail data and the tornado data sets are investigated by using Matlab. First, two data sets (csv files) are imported to the Matlab using import.m function (see in Github). A Matlab script (DataManipulation.m) is implemented to pair two data sets. Figure 2 depicts the basic flow diagram of the pairing algorithm. The hail and tornado events are assumed to be formed by same storm cell if they occurred in similar time and similar region. To determine the relationship for a tornado and hail we propose the following algorithm; the tornado and hail data will be searched together to find the similarities in data sets. For example, for a selected tornado, the hail data set will be searched for date. If hail data is available for same day with tornado then time information will be compared. We will assume that tornado and hail created by same storm cell, if the occurrence time between two event is less than 1 hour and the location of two event is close enough (roughly 25 km^2 or with in a 0.25° degree error in latitude or longitude). Then we will create the density estimation scatter plot by looking the hail size and the tornado intensity (F-Scale) for paired data.

As seen in Fig.1, each tornado track is codded with different color. Color code represent the F-scale (intensity) of the tornadoes. Color map span between blue and red which represents the range F-scale 0 to F-scale 5 respectively. In addition to the color code, opacity is used to emphasize the differences between strong and weak tornadoes. The weak tornadoes are codded more transparent comparing to the strong tornadoes. The length and width of the tornado track plotted on map scaled according to the original length and width of the tornadoes.

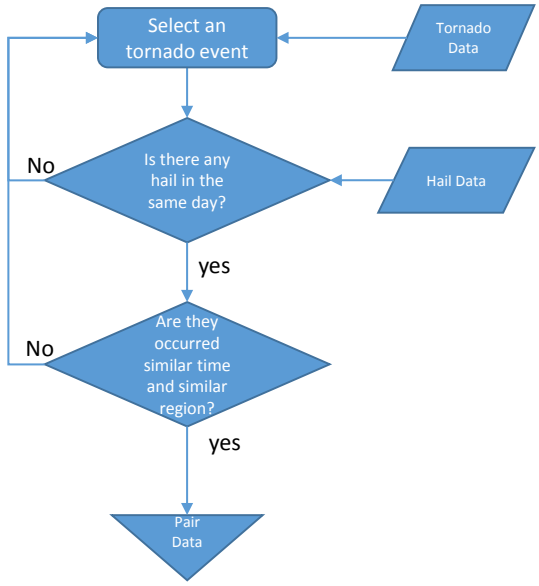

Fig. 2. Flow diagram of the pairing algorithm

Note that some of the small tornadoes which are insignificant and hard to see in visualization are omitted during the plotting for fast response.

Fig. 3 shows the user interaction with the map. When user select (click) a point inside the state borders then map automatically updated (zoom) to show the selected state. This feature allows user to understand the behaviour of selected weather events in selected state more closely. Similarly, user can interact the map again to see the all states again.

A density estimation scatter plot is created to show the correlation between the tornadoes and hail events. There are different hail size are reported between 0 to 10 inches. It is not practical to show each different hail size in density estimation plot so for simplicity hail size dimension is reduced to hail size groups. Any hail size between integer number fall in appropriate group. For instance, hail size 0.33 inches is assumed to be in the first group which represents any hail size between 0 to 1 inches. As a result, 10 different hail size group versus F-scale of the tornadoes are created and plotted as in Fig. 4.

The density scatter plot shown in Fig. 4 is color coded to identify the dense occurrence. In addition to that, a tooltip displays the number of occurrence when mouse hover on scatter plot. User can update the tornado map by clicking scatter plot. This feature helps user to

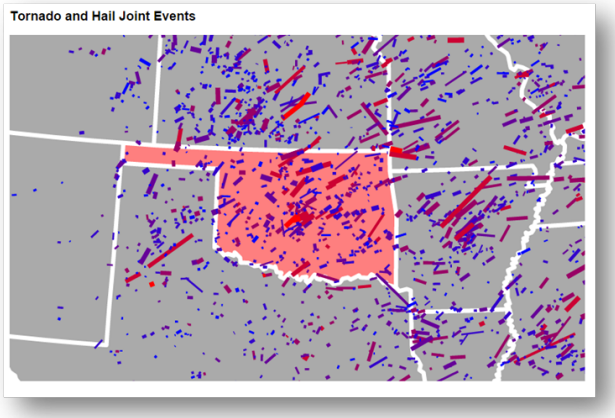

Fig. 3. Interaction with the map. Close look (zoom) to Oklahoma.

Table 1. Some of the fields of given tornado and hail data sets

Field #	Abbreviation	Explanation or Range
1	Tornado or hail number	A count of tornadoes or hails during the year: Prior to 2007
2	(yr)	year, between 1950-2009
3	(mo)	Month, 1-12
4	(dy)	Day, 1-31
5	(date)	Date in yyyy-mm-dd format
6	(time)	Time in HH:MM:SS
8	(st)	State , two-letter postal abbreviation
11	(f) or (s)	F-scale (EF-scale after Jan. 2007): values -9, 0, 1, 2, 3, 4, 5 (-9 = unknown). Or, hail size in inches
12	(inj)	Number of injuries
13	(fat)	Number of fatalities
14	(loss)	Estimated property loss (U.S $)
16	(slat)	Starting latitude in decimal degrees
17	(slon)	Staring longitude in decimal degrees
18	(elat)	Ending latitude in decimal degrees
19	(elon)	Ending longitude in decimal degrees
20	(len)	Length of the tornado in miles
21	(wid)	Width of the tornado yards

understand the spacial location of selected events. User may show all hail related tornado events by clicking the show all button.

5 Findings and Insights

Proposed visualization tool effective to exploit the information buried in two different data set by joint visualization and helps researcher to understand the joint events. Especially, expose the following information

Finding 1: Under the assumptions mentioned in Section 4, proposed method is able to pair the tornadoes and hail successfully.

Finding 2: Proposed method achieve the density estimation plot of the joint occurrence of hail and tornadoes as shown in Fig. 4. It is obvious that the F-scale 0 tornadoes produces hail between 1 to 2 inches (2763 paired events). Then it is followed by hails 2 to 3 inches (2042 paired events). As seen from Fig. 4, the extrema cases are observable but the occurrence of these events are more rare.

Finding 3: The joint hail and tornado events occur mostly in mid US. There are significant amount of distribution for east cost whereas there is rare events observed in west cost of US.

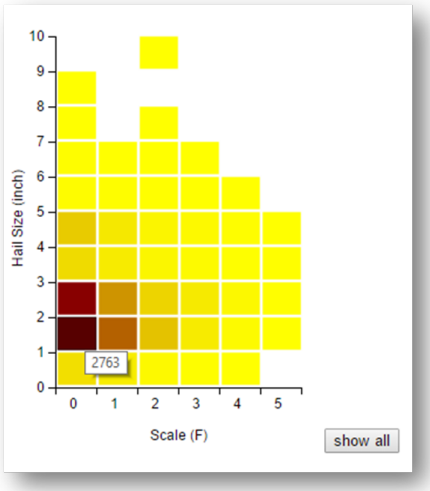

Fig. 4. Interaction with the map. Close look (zoom) to Oklahoma.

6 Limitations and Future Work

The proposed visualization tool works under the assumption of that the hail and tornado events are joint events and created by same storm cell if they occurred in similar time and similar region. The scientific reasons of this assumption is not covered in this paper and needs extensive research (which is an active research field of meteorology). However this tools may help meteorologist and researcher to understand the hail and tornado correlation.

It is known that the tornado events are observed more often during the spring season. Unfortunately, the temporal information is not provided with current visualization tool to exploit this information. This important feature may be included to the next releases. In addition to that, different sliders may be added to control the data filtering process as shown in Fig. 5. The sliders implementation work in progress during the reporting period.

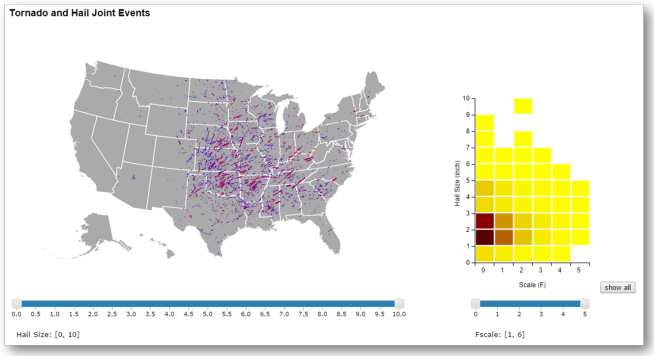

Fig. 5. Possible slider improvement to the proposed visualization tool for future releases.

7 Conclusion and Lessons Learned

Even though there are different visualization tools available on web, they dont give enough information about the correlation of hail and tornado events. We propose a visualization tool that give emphasis of hail and tornado correlation. This tool has the ability to process and visualize the tornado and hail information given in SPC-WCM web page. This tool aims to help meteorologist and researchers for understanding the correlation between hails and tornadoes in US.

During the course of design, we learn how to convert two complicated data set into a more meaningful visual information. It has been

also shown that visualization is an effective tool to extract hail and tornado correlation information.

ACKNOWLEDGMENTS

The authors wish to thank Dr Enrico Bertini and Mr. Josua Krause for their help and support during the course of work.

REFERENCES

[1] Interactive tornado tracks map. http://uxblog.idvsolutions.com/2012/07/interactive-tornado-tracks-map-is-live.html. Accessed: 2015-05-15.

[2] National Oceanic and Atmospheric Administration Storm Prediction Center Warning Coordination Meteorologist Web Page. http://www.spc.noaa.gov/wcm/. Accessed: 2015-05-15.

[3] Twister Dashboard: Exploring three decades of violent storms. http://storymaps.esri.com/stories/2014/tornadoes/#sthash.aTbs0WnG.dpuf. Accessed: 2015-05-15.

[4] Twister tornado visualization tool. http://roadtolarissa.com/twisters/. Accessed: 2015-05-15.

LINK

A working demo of the proposed visulization tool can be find online at http://nyu-cs6313-projects.github.io/Tornado-Hail/