Computers, Environment and Urban Systems 72 (2018) 212–220

Contents lists available at ScienceDirect

Computers, Environment and Urban Systems

journal homepage: www.elsevier.com/locate/ceus

Data-driven analysis on the effects of extreme weather elements on traffic volume in Atlanta, GA, USA

David Sathiaraj[a,*], Thana-on Punkasem[b], Fahui Wang[c], Dan P.K. Seedah[d]

[a] Dept of Geography and Anthropology and NOAA Southern Regional Climate Center, Louisiana State University, United States
[b] Dept. of Human Centered Design & Engineering, University of Washington, United States
[c] Dept. of Geography and Anthropology, Louisiana State University, United States
[d] Texas Transportation Institute, Texas A&M University, United States

ARTICLE INFO

Keywords:
Weather hazards
Transportation
Urban infrastructure
Climate
Big data
Data analytics
Machine learning
Data visualization

ABSTRACT

Severe weather events pose a significant threat to transportation networks. This research analyzes and discusses the impact of precipitation, temperature, visibility and wind speed on hourly weekday traffic flow volume in Atlanta, Georgia. The study involves the following: determine weather variables that affect traffic volume, develop a machine learning technique to derive decision rules based on weather and traffic volume, and create a web-based decision support visualization tool using the analyzed results. The relationship between extreme weather events and traffic volume was investigated by comparing traffic volume between a base case scenario and an extreme weather scenario. Data from 48 Automatic Traffic Recorder (ATR) sites around Atlanta, GA, USA and hourly precipitation data from 4 climate measurement stations were used to conduct this study. The spatiotemporal relationships between traffic volume and weather variables were analyzed individually and evaluated using a non-parametric statistical test. A machine learning technique is applied to derive decision rules that result in reduction in traffic volume. Results show significant impacts on traffic volume from visibility, precipitation and temperature and helps in isolating hours in a typical weekday when such impacts are felt. A decision support tool was also developed to visualize traffic volume and weather interactions. The data-driven insights from this analysis is applicable to transportation planners, centralized traffic control rooms and urban infrastructure decision makers.

1. Introduction

The transportation sector is an important component for the economic development of an area. An efficient transportation network increases economic and social opportunities by improving accessibility to employment, markets, and additional investments. As population and the number of vehicles increases, congestion on road networks becomes more common. Traffic congestion accounts for over three billion hours of traffic delay annually in the US (*Press release - urban mobility information*, 2015).

Extreme weather events can adversely impact traffic volume (Cools, Moons, & Wets, 2010, Calvert & Snelder, n.d.). These weather events have negative effects on transportation network performance, travel speed, time, capacity, and volume (Bartlett, Lao, Zhao, & Sadek, 2013). The Federal Highway Administration (FHWA) provided a statistical report showing that severe weather events cause up to 22% of vehicle accidents. According to a 2015 Federal report (DoT, Nitch, Safety, & FLAP, 2015), more than 5897 deaths and 445,303 injuries occurred in the U.S. during a ten-year period because of extreme weather events. A better understanding of weather factors affecting traffic flow can eventually help both government and other infrastructure agencies to properly plan, design, and maintain road networks and thereby reduce the risk of fatalities.

This research investigates and analyzes the impact of four weather variables - precipitation, visibility, temperature and wind speed - on hourly traffic volume. The focus of the study was on the city of Atlanta, Georgia, USA. One of the main motivations for choosing Atlanta was it being the large urban city with a dense transportation network and its propensity to extreme weather hazards. NOAA climate normal reports show that Atlanta receives an average annual rainfall of 1262.63 mm (49.71 in.) which is 27% more than the average in other similarly-sized US cities (Arguez et al., 2012).

Two major data sets were used in this research. Traffic data for the state of Georgia, USA was provided by Transmetric LLC, a transportation data management firm located in Austin, Texas. The weather data was retrieved from the Integrated Surface Hourly (ISH) archives at the

* Corresponding author at: E328 Howe Russell Bldg, NOAA Southern Regional Climate Center, Department of Geography, Louisiana State University, LA 70803, United States.
E-mail address: davids@srcc.lsu.edu (D. Sathiaraj).

https://doi.org/10.1016/j.compenvurbsys.2018.06.012
Received 14 February 2018; Received in revised form 26 June 2018; Accepted 29 June 2018
Available online 09 July 2018
0198-9715/

National Centers for Environmental Information and housed locally.

The impact of weather hazards on traffic volume was then studied by applying rigorous statistical and machine learning tests. A decision support tool was built to visualize the impacts of inclement weather events on transportation volume in Atlanta.

1.1. Problem description

This work focused on analyzing and examining the following research questions:

- Do precipitation, temperature, visibility and wind speed have a significant impact on hourly traffic volume in Atlanta?
- If so, does the impact have a specific pattern? Is the impact different based on different times of the day and volumes of precipitation, temperature, visibility and wind speed?
- Is it possible to statistically correlate these impacts?
- Is it possible to develop a machine learning model that can account for interdependent weather variables and predict impacts on hourly traffic volume?

1.2. Background and similar work

Previous studies examined relationships between weather events and traffic characteristics. A recent study by (Yuan-Qing & Jing, 2017) looked at the effect of rainfall on traffic flow on a freeway in Hainan province in China. Studies have shown that extreme weather conditions could lead to reduction in speed, travel time, road capacity, and volume. For instance, research from the Federal Highway Administration showed that light rainfall could reduce driving speed by 6–9% in several cities (Hranac et al., n.d.). This decrease in speed during extreme weather caused 3.5% more travel time compared to days (Stern, Shah, Goodwin, & Pisano, 2003) with normal weather conditions. Heavy rain has been reported to produce a 14–15% reduction in traffic speed and capacity (according to Agarwal, Maze, & Souleyrette (2005)), and heavy snow events have been reported to reduce the capacity of road networks by up to 28% (Agarwal et al., 2005).

Some studies investigating traffic behavior using weather variables have applied machine learning models to predict traffic characteristics in different weather conditions. These machine learning models could accurately classify which variables contributed to change in traffic behavior. For example, a study of the effect of traffic parameters on road hazards using a classification tree model identified hazardous situations on the freeways (Hasan, 2012). Results of this research showed that traffic flow and vehicle speed were the most important factors that influence traffic volume. Another study of the mixed effects of precipitation on traffic crashes used a machine learning technique to build a crash risk prediction model based on precipitation and snowfall. The study discovered that if precipitation increased by 10 mm, the fatal crash rate would increase about 3% (Eisenberg, 2004).

Previous research found decreases in traffic volume during different severe weather events. For example, an average winter storm event can reduce traffic volume by 29% (Knapp, Smithson, & Khattak, n.d.). Another urban study (Kwon, Fu, & Jiang, 2013) looked only at the impact of winter weather conditions on traffic volume and the study monitored just 2 locations over a 2 year period. A study used a probabilistic approach to find reductions in traffic volume during intense snow and rain events (Samba & Park, 2010). A study of hourly traffic volume and precipitation in Buffalo, NY, USA investigated relationships between traffic volume and precipitation by dividing traffic volume into two subgroups including 1) traffic volume during non-inclement weather (base case) and 2) inclement weather (inclement case) (Bartlett et al., 2013). This research applied a regression model to the dataset to create the predictive model under specific weather conditions. Results showed a significant correlation between hourly rainfall and traffic volume (Bartlett et al., 2013). Similarities between this study and that of (Bartlett et al., 2013) were the sub-grouping of scenarios between extreme and normal weather conditions. The differentiating factors were different cities with differing extreme weather thresholds, analyzing a more extensive network of traffic counters, differing statistical techniques and use of a machine learning model to derive rules. Similar to our work, Angel, Sando, Chimba, & Kwigizile (2014) explored the effect of rain on the traffic volume at 2 freeway sites in Florida and Dehman & Drakopoulos (2017) studied the effect of inclement weather on 15 freeway traffic counters in Milwaukee, Wisconsin. Factors that differentiated our work were the extensive statistical analysis, the deriving of thresholds and specific hours of affected traffic volume using predictive models and the decision support visualization tool.

Wind speed, temperature and visibility could also create hazardous driving environments and cause a decrease in traffic volume. A study of the effect of wind speed, temperature, precipitation and visibility on traffic capacity in Minnesota found that precipitation had the most severe impact on traffic capacity, reducing it by 19–28% (Agarwal et al., 2005). Cold temperatures and low visibilities had moderate impacts on the capacity, leading to a 10–12% decrease. However, wind speed did not have any noticeable effect on capacity reduction in Minnesota (Agarwal et al., 2005). Work such as Saha, Schramm, Nolan, & Hess (1994–2012) seek to provide a link between adverse weather conditions with fatal vehicle crashes in the United States by looking at the Fatality Analysis Reporting System (FARS) data set.

1.3. Outline

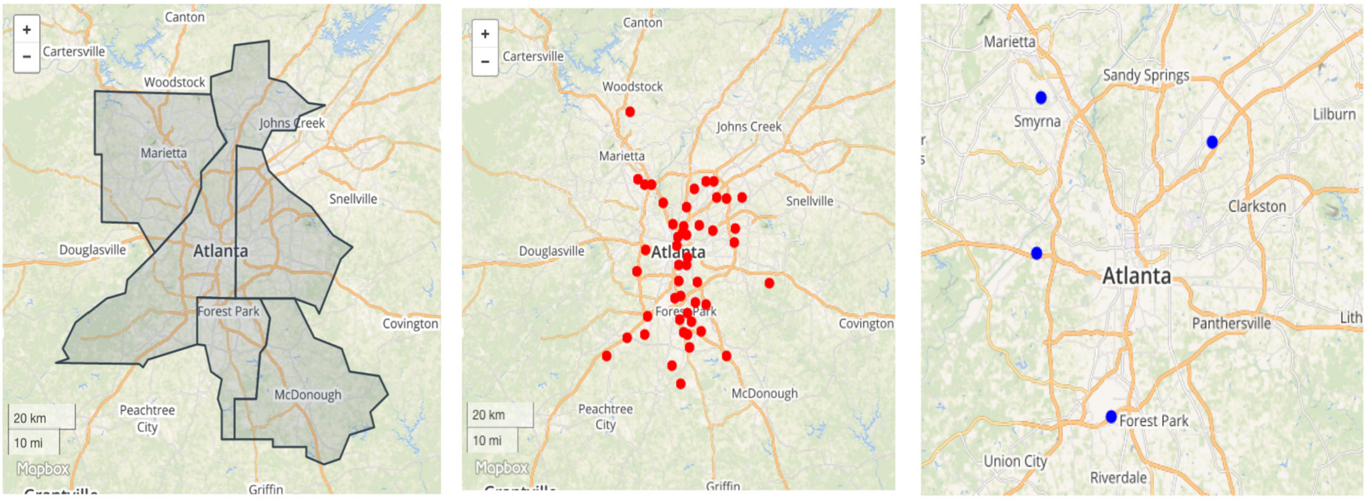
This work consists of three components. First, a statistical, data-intensive study is provided to analyze correlations between individual weather elements and traffic volume. Then, a machine learning model is discussed to predict changes in traffic volume based on weather conditions. Finally, a decision support tool is presented to visualize hourly traffic volume and its interaction with studied weather elements.

The paper is organized as follows: First, a background of prior research on this topic is provided. Also provided is a detailed description of the datasets, along with methodologies pertaining to data collection, storage and data processing. Secondly, the statistical experiments and computational results are provided. Thirdly, a discussion of the machine learning model and the development of the geo-visualization decision support tool (this includes programming methodologies used) is detailed. The last section provides conclusions and future work.

2. Study area and data processing

The focus area was the Atlanta Metropolitan Statistical Area (MSA) which has a dense network of traffic counting stations. Atlanta also happens to be an important business hub and one of the largest population centers in the southern U.S. According to the Georgia Department of Revenue, there are almost 3,000,000 personal passenger cars in the area with nearly 500,000 daily riders. Weather hazards can play a significant role in such a city.

The specific study area includes Cobb, Fulton, Dekalb, Henry and Clayton counties (see Fig. 1). The study area covered approximately 135 mile^2, containing latitudes between 33.4623° N and 34.0683° N and longitudes between 84.1462° W and 84.6059° W. Based on the NOAA 1981–2010 US Climate Normals dataset (Arguez et al., 2012), this area averages 113 days with rain and 1263 mm (49.7 in.) rainfall annually. The annual rainfall total is 27% larger than the average rainfall in other cities with similar size across the United States. Precipitation, visibility, wind speed and temperature data used was from the Integrated Surface Hourly (ISH) dataset (Smith et al. (2011)) of the National Centers of Environmental Information (NCEI). ISH is a large hourly data set spanning more than 2000 stations worldwide. Four weather measurement stations were available within the study area and reported regularly between 2010 and 2015 (see Fig. 1). The four weather stations included: KATL (Hartsfield-Jackson Atlanta



(a) Map of counties (b) Map of Traffic sites (c) Map of climate sites

Fig. 1. Map of counties. Map of Traffic sites. Map of climate sites.

International Airport), KFTY (Fulton County Airport-Brown Field), KPDK (Dekalb-Peachtree Airport) and KMGE (Dobbins Air Reserve Base). Spatial interpolation techniques such as Kriging, Inverse Distance Weighting (IDW), and Radial Basis Functions (RBF) were then used to interpolate weather observation values at each traffic counter location (see Fig. 1).

The study uses two main data sources: hourly traffic volume and weather variables. Hourly traffic volume data were collected from 48 Automatic Traffic Recorder (ATR) sites spanning the 5-year period, 2011–2015. The ATR stations are permanently installed sensors that count the number of vehicles passing through each counter location on state highways, major county roads, and city streets on a 24/7/365 day basis. The traffic dataset also consisted of information such as location (latitude and longitude) of the sites, names of the roads where each counter was located, amount of hourly traffic volume, and an hourly date stamp that the counter recorded the traffic volume.

4 hourly weather elements/variables were examined: precipitation, visibility, wind speed and temperature. Weather data from the Integrated Surface Hourly (ISH) data set of the National Centers of Environmental Information (NCEI) was used. This analysis retrieved nearly 225,000 records of hourly weather observations spanning the time period, 2011–2015. Data was reported according to the UTC timezone and were converted to US Eastern Time Zone (Atlanta's time zone).

2.1. Setting of parameters

This work analyzed only weekday traffic volume as weekday traffic has less variability than weekend traffic. Another aspect of our study with regards to temperature was to only use data from the winter months of December to February. This was to isolate the study to a season when such extreme temperatures are most likely to occur. To determine thresholds for the base and extreme cases, we use threshold guidelines for climate extremes as suggested by the World Meteorological Organization (Data, 2009). As per the guidelines, a typical threshold used is values above a certain percentile, 95th or 99th. In our study, we use values that are at the 99th percentile as a threshold for the extreme case. Since climate extremes vary from city to city, it was appropriate to use a percentile-based threshold. Climate data from the 4 weather stations in the Atlanta area and for the 4 weather elements were extracted. The 99th percentile values were then calculated and used as a threshold to distinguish between the base and extreme cases (and outlined in Table 1).

Table 1
Weather variable criteria used in the Atlanta transportation study.

Climate elements	Base case	Extreme case
Precipitation	0 mm/h	> 7.62 mm/h
Temperature	> 32 F	≤ 32 F
Wind Speed	< 10 m/s	≥ 10 m/s
Visibility	> 5 km	≤ 5 km

Table 2
Comparing of number of hours under base and extreme weather conditions in 2011–2015 for the 48 Traffic Counters in Atlanta, GA.

	Avg. base hours	Avg. extreme hours	Percent of extreme to base
Precipitation	33,467	2138	6.5
Temperature	34,273	1047	3.5
Wind Speed	34,461	222	0.7
Visibility	33,216	1880	5.9

Table 2 provides a summary of average number of hours that fit under the base condition category, the average number of hours that fit under the extreme case category and the average percentage of extreme to base condition hours.

2.2. Spatial interpolation analysis

Since the data were collected at 4 unique climate measurement sites, spatial interpolation techniques such as Inverse Distance Weighting (IDW), Kriging and radial basis functions (RBF) were used to estimate the 4 weather values at each of the 48 traffic counters. IDW (Mitas & Mitasova, 1999), Kriging (Hartkamp, De Beurs, Stein, & White, n.d.) and RBF (Rusu & Rusu, 2006) are commonly used spatial interpolation techniques.

In order to find the most accurate spatial interpolation technique, data from the 4 weather stations in the study were spatially interpolated. One of the weather stations was used as a reference station. The spatially interpolated values were then compared to the actual observed values of the reference station and root-mean-square error (RMSE) values were calculated. The technique with the smallest RMSE value was used in the study.

Based on the RMSE results as reported in Table 3, IDW produced the least error for precipitation, temperature, and visibility. However, for

Table 3
Root mean squared error (RMSE) values from each spatial interpolation techniques.

	IDW	Kriging	RBF
Precipitation	F8FF000.07596666	0.08476519	0.2435945
Temperature	F8FF001.778402	4.97	6.026029
Visibility	F8FF001.273612	1.49855	1.292256
Wind speed	3.191654	3.201029	FFCCC9 3.124553

wind speed, RBF produced the smallest error compared to IDW and Kriging. So IDW was used for interpolating temperature, rainfall and visibility while RBF was used for spatial interpolation of wind speed.

3. Correlation analysis

After applying spatial interpolation techniques to get estimated weather values at every traffic counter site, the next step was to find correlations between hourly traffic volume and hourly weather events. This was achieved for each traffic counter, by separating the traffic volume data set on an hourly basis into base and extreme weather categories (using the thresholds outlined in Table 1.

A computational result of such a comparison is provided on an hourly basis for 2 traffic counters (with traffic counters # 36764 and #25065) in Figs. 2 and 3. The percentage drops in traffic volume caused by an extreme weather hazard is computed for each hour of a given day (see example computation in Table 4). It is also evident from Fig. 2 that on certain hours of a day, there are substantial drops in traffic volume caused by precipitation and temperature. For example, for precipitation, during the peak evening hours of 16 and 17 h, there is nearly a 10% drop in traffic volume in Figs. 2 and 3.

To examine the influence of the 4 hazardous weather events on traffic volume, the traffic dataset was divided into two subgroups - one corresponding to hours that formed the base case (hours where there were normal driving conditions with no extreme weather hazards) and another corresponding to hours that experienced any or all of the 4 extreme weather hazards. The base and extreme cases are categorized in Table 1.

The 2 subgroups or time-series were created for each station. For the statistical analysis, the hypothesis was that there was a difference between the mean of hourly traffic volume in normal driving conditions and the mean hourly traffic volume due to an extreme weather hazard. The non-parametric statistical test, Wilcoxon Test (Wilcoxon, 1945) was conducted on each of the 4 weather elements used in the study.

In many cases, due to the rarity of an extreme weather hazard numbers of base cases outnumbered extreme cases more than five times. To avoid issues from different sample sizes, a bootstrapping technique was used to randomly pick equal numbers of extreme and base condition records and the statistical tests were repeated 100 times. The *p*-values were sorted in ascending order. After empirical analysis of the results and to ensure confidence in the repeatability of the results, *p*-values at the 80 percentile were chosen as representative of the group. If more than 80% of the *p*-value results fell under 0.05, it was interpreted that there was a significant difference between the means of the extreme and base case scenarios.

4. Analysis of collective effect of weather variables on traffic volume

To consider together all 4 weather elements and their collective effect on traffic volume, we explore a predictive modeling and machine learning technique, Random Forest (Breiman, 2001). In machine learning, decision tree based learning involves the construction of a tree like structure to depict the features (or independent variables) in a dataset. The constructed tree can help derive decision rules and for predicting a target value. Decision trees can predict a categorical value for classification tasks (similar to this problem) or predict continuous values for regression. Decision trees use all the independent variables as input and help provide decision rules with thresholds as an output. The Random Forest technique is an extension of the decision tree based methodology and involves the randomized construction of a number of decision trees and picking out the most commonly occurring tree structure for deriving decision rules. This technique finds applicability in areas such as behavioral science (Sathiaraj, Cassidy Jr., & Rohli, 2017) and in transportation (Ghasri, Rashidi, & Waller, 2017). Machine learning techniques such as Random Forest are also agnostic to the underlying statistical distribution of the data set being analyzed. The derived decision rules can help traffic planners and signal operators to know thresholds and conditions when traffic volume will be below normal and when it will be normal under abnormal weather conditions.

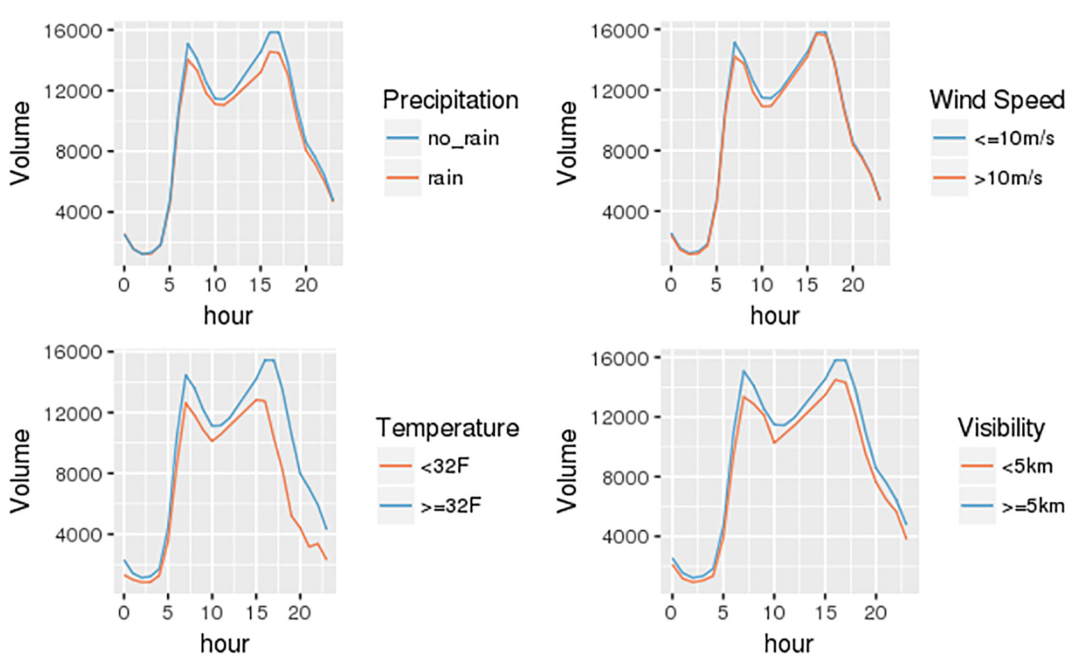


Fig. 2. Effect of different weather hazards on traffic volume by hour of day for station id 25065.

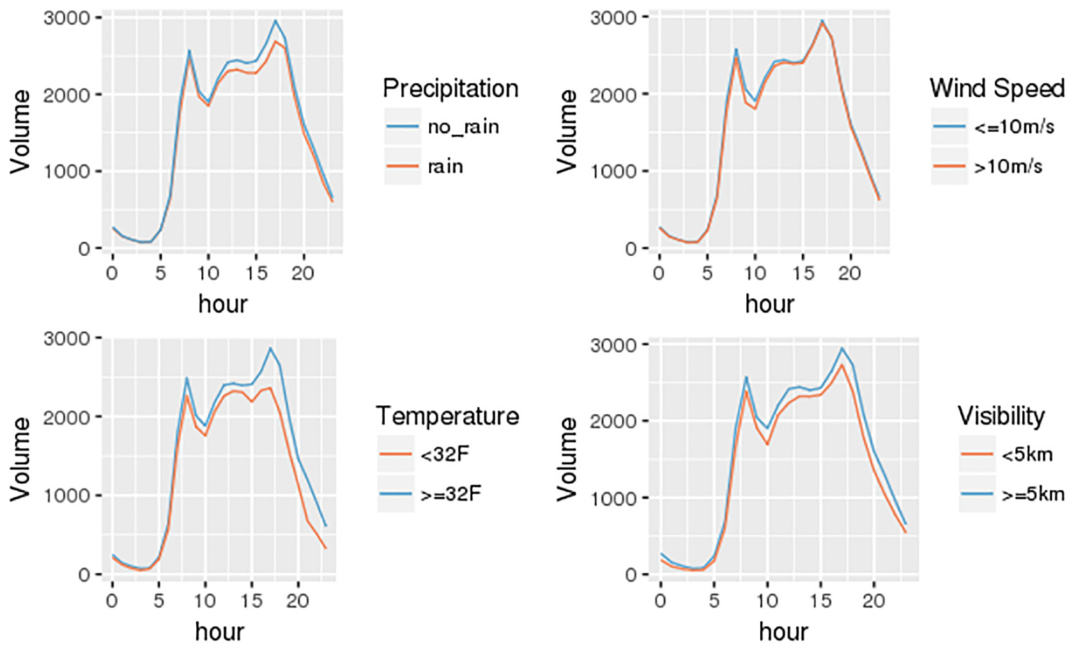


Fig. 3. Effect of different weather hazards on traffic volume by hour of day for station id 36764.

Table 4
Percentage drops in traffic volume ((base case-extreme case)/base case) under 4 weather hazards for traffic counter 36764.

Hour	Precipitation	Temperature	Wind	Visibility
0	7.69	14.86	4.74	31.87
1	5.13	16.89	7.01	34.39
2	1.85	24.51	2.78	35.78
3	7.89	31.08	6.58	30.26
4	1.22	10.26	3.61	28.92
5	2.5	13.57	6.67	27.8
6	7.4	12.38	6.52	13.66
7	8.14	10.55	7.08	12.99
8	3.07	8.97	4.46	7.35
9	3.57	7.02	8.46	6.66
10	2.74	6.54	5.46	10.99
11	2.86	5.54	2.9	5.85
12	4.88	5.83	2.32	7.53
13	5.07	4.01	1.39	4.96
14	5.24	3.76	0.58	3.46
15	6.49	9.34	1.19	3.79
16	8.7	9.41	0.45	5.89
17	9.16	17.56	1.02	7.27
18	4.69	22.46	−0.59	12.68
19	6.48	21.91	1.71	14.65
20	7.66	22.95	2.74	15.92
21	7.69	43.65	2.69	19.68
22	12.33	43.56	3.23	19.42
23	8.45	46.68	5.96	16.82

We derive predictive models for each of the traffic counters considered in the study.

Previous research has applied entropy based decision trees to predict climate events and traffic behavior. For example, one study used a decision tree classifier to predict precipitation in India. It showed that the decision tree learning approach provided an average accuracy of 79.42% (Prasad, Reddy, & Naidu, n.d.). Another study used a decision tree approach to analyze air traffic delay by comparing the accuracy of results from three machine learning models (Kulkarni, Wang, & Sridhar, 2014). For this research, a decision tree based predictive technique is used to classify and predict as to which weather variables and their thresholds, significantly contribute to the decrease of traffic volume. The predictive model splits the traffic dataset into a tree comprising of nodes and branches (as depicted in Fig. 5). Training sets were generated by classifying traffic volume into two categorical classes – normal and abnormal. For each traffic counter, the traffic volume that fell within one standard deviation (spanning the time period 2011–2015), was considered as the normal class, otherwise, it was considered as the abnormal class (or low traffic volume). The training data used were the hour of day, temperature, precipitation, visibility and wind speed. For each traffic counter, the data was normalized before a training set was generated. A random forest decision tree model was then developed to predict the traffic volume case scenarios (normal vs abnormal). The models at each traffic counter averaged about 80–85% in predictive accuracy after a 10-fold cross-validation process.

An analysis of results is now described. First, the statistical analysis examined the correlation between individual weather elements and traffic volume. A machine learning model was built to analyze all weather elements simultaneously and to predict under which conditions traffic volume drops or stays normal. The decision support tool visualized the interaction between weather elements and traffic volume on an hourly and daily basis. Data from each traffic counter was split into base and extreme cases and then subjected to the Wilcoxon test. The statistical analysis was subjected to cross validation, the *p*-values from each dataset were then collected and analyzed graphically. The Wilcoxon test was chosen as it is a non-parametric test, suited for randomly selected, paired samples and unlike the *t*-test, it does not assume that the underlying distribution is normally distributed.

Based on results from the Wilcoxon test, the effect of each weather variable across all the traffic counters is plotted and depicted as Fig. 4. The graphic provides information on specific hours of a typical day when a weather variable can impact traffic volume and the number of traffic counters where the decrease in traffic volume was statistically significant. Effects of precipitation, temperature and visibility on traffic volume were felt at a majority of the traffic counters during certain hours of the day. Precipitation events accounted for large decrease in traffic volume during afternoon to late night hours. During this period of a typical day, more than 70% of traffic counters showed a significant drop of traffic volume. Precipitation had a significant impact on the decline of traffic volume for more than 95% of the traffic stations during hours 19–21.

Extreme temperature had a significant impact on almost all hours of the day during the winter months. Temperature accounted for a decline

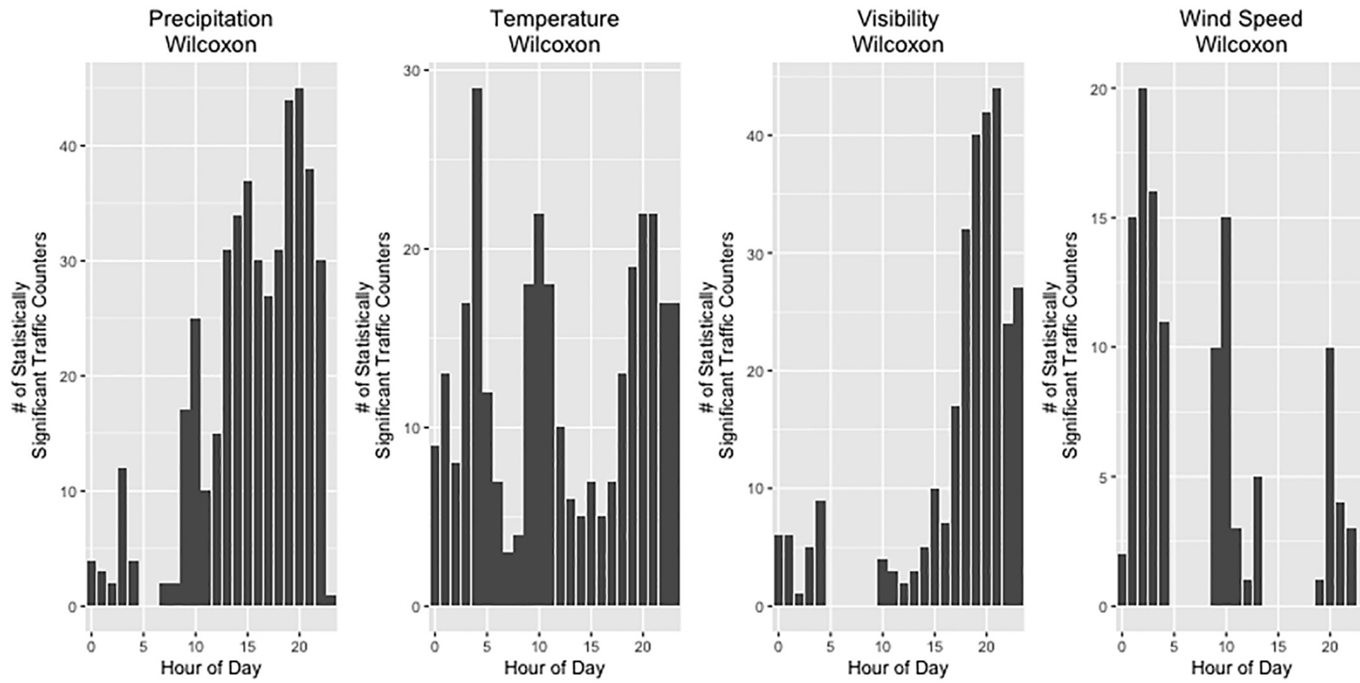


Fig. 4. Results of statistical analysis.

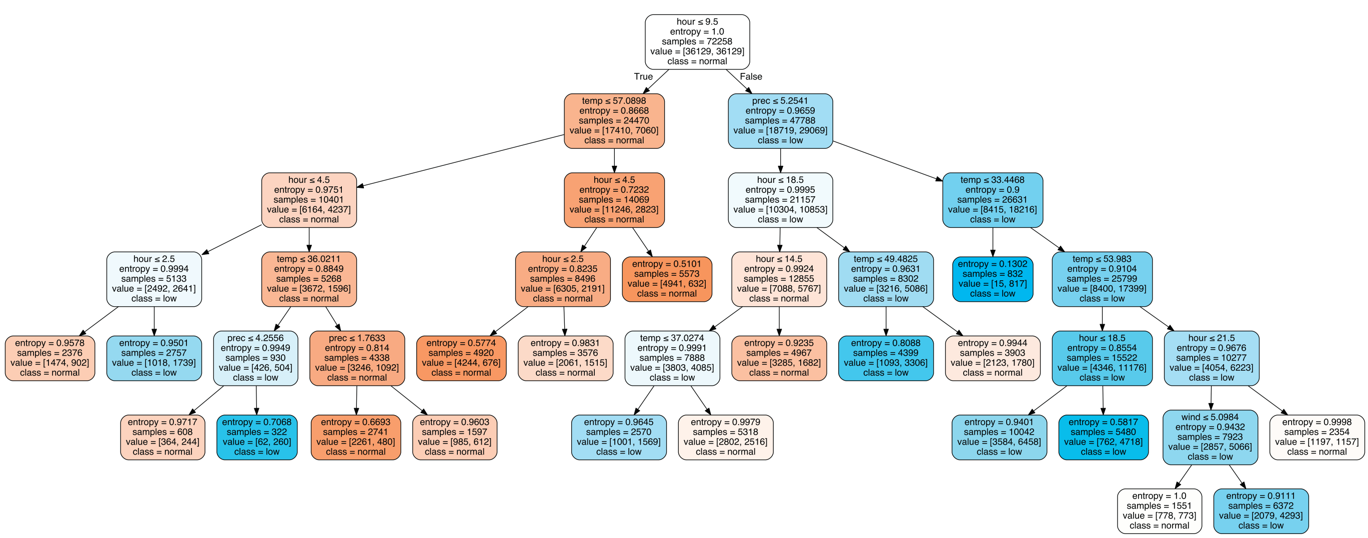


Fig. 5. Decision tree analysis for traffic station id 15287 for traffic patterns in Atlanta, GA.

in traffic volume for more than 60% of traffic counters and for nearly 11 out of 24 h in a day. Significant impacts were felt between 19 and midnight hours. Extreme low temperatures during the winter months also played a role in the decline in traffic volume. This was evidenced in more than 75% of traffic stations between the hours of 10 pm–12 am.

During poor visibility hours, more than 95% of traffic counters indicated a significant decline in traffic volume. The main hours in a day that showed a decline were between 17 and 23 h. During evening rush hours and early evening hours, more than 95% of traffic counters showed a dramatic drop in traffic volume. There were also effects on traffic volume in counters in certain areas of Atlanta, mainly between 1 and 4 am hours and 15–16 afternoon hours.

For a majority of traffic counters, wind speed did not factor in as having a significant impact on traffic volume.

In summary, based on the results of the Wilcoxon test, precipitation, visibility and temperature emerged as important factors that accounted for a decrease in traffic volume in Atlanta. These variables had significant impacts on traffic volume during the evening hours of 19–21. Temperature also had a significant impact on traffic volume between 19 - midnight hours.

5. Deriving decision rules using machine learning models

This machine learning model seeks to understand conditions under which the weather variables can reduce the traffic volume. Since traffic volume can vary across different traffic counters and under different extreme weather conditions, models were derived for each traffic counter in the study. Five variables - temperature, precipitation, visibility, wind speed, and hour of day - were used as attributes of the training set. The objective is not to predict the traffic volume (as typical supervised classification systems do) but instead predict combination(s) of weather hazards and thresholds that can result in a drop in traffic volume. Decision tree based models are suited for such a task as they provide tree-based rules and conditions that are concise for traffic and

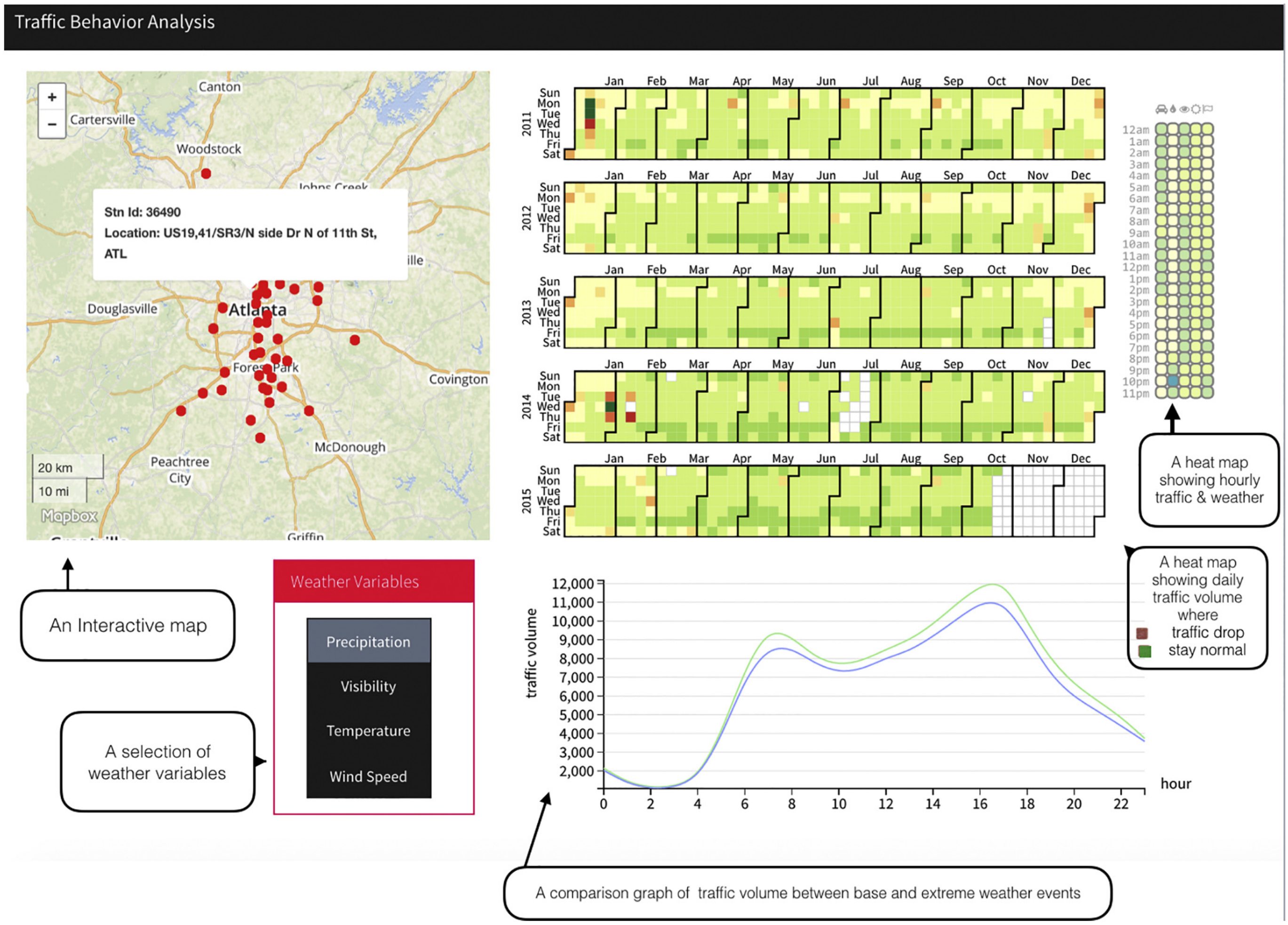


Fig. 6. A screenshot of the geo-visualization decision support tool.

infrastructure planners to use and implement. Based on the decision tree results from the model, hour of the day was a critical decision marker for deriving the decision rules, followed by temperature and precipitation. An example of one such decision tree is provided for traffic station 15287 as Fig. 5. The patterns of traffic volume were different before and after the 10th hour. For hours prior to 10 am, temperature was the major variable that caused a drop in traffic volume. Then, precipitation was the second leading cause for the decline in traffic volume. From Fig. 5, it is also evident, that traffic volume drops when temperatures fall below 36 °F and precipitation is more than 4 mm per hour. However, for hours after 10 am, precipitation emerged as a significant variable. Temperature was the second most important variable that led to the significant decreases in traffic volume. During the hours after 10 am, traffic volume dropped drastically when precipitation was more than 5 mm per hour and temperature was around 33–35 °F.

In conclusion, after analyzing all variables together, precipitation and temperature were variables that greatly affected traffic volume in Atlanta. Wind speed did not emerge as a significant factor to cause a drop in traffic volume. Notably, the combination of temperature when it was below 35 °F and precipitation that was more than 5 mm/h created the most significant impact on traffic volume throughout the day. The predictive models also provided support to the notion (similar to the results from the statistical analysis) that low temperatures during the early morning hours caused a decline in traffic volume (Fig. 5).

Based on the results from the statistical tests and the predictive machine learning model, one can conclude that extreme precipitation and temperature (during the winter months) during certain hours of a typical work day can negatively influence traffic volume. There were different impact patterns among the weather elements. During midnight to late morning hours in the winter months, extreme minimum temperature in combination with poor visibility influences traffic volume. Precipitation also contributed to a drop in traffic volume. Between noon and midnight, extreme precipitation was a significant factor in decreasing traffic volume. For each traffic counter, the statistical analysis yields specific hours when extreme weather hazards can influence a drop in traffic volume and the machine learning model provides combination(s) of weather conditions and their respective thresholds under which traffic volume drops.

6. Technology stack and geo-visualization decision support tool

The technology behind the geo-visualization comprised of 3 components: a database to store the data, a front-end, interactive user interface for visualization of analytics, and a middle-layer web framework (Django, http://www.djangoproject.com) as a communication layer between the data base and the user-interface. The database used was Postgresql http://www.postgresql.org and the front-end was written in Javascript and included the mapping library mapbox.js (http://mapbox.com) and the visualization library d3.js (http://d3js.org).

Geovisualization techniques were utilized to build a web-based geographic decision support tool. This tool was built to visually support and provides a better understanding of the impact of hazardous weather events on traffic volume. The 5 main components of the geo-visualization tool includes a map-based component, interactive displays for daily and hourly traffic volume and comparative display of traffic volume under normal and extreme weather conditions.

The map component displays the traffic counters used in the study and allows a user to select a counter. Upon clicking a traffic counter, a calendar-based visualization is loaded that displays daily traffic volume as a heat map. The calendar spans the 5 year study period. The heat map is color-coded from red to green, red representing a daily traffic volume that is 70% below the average traffic volume for that day from 2011 to 2015. Dark green color is used when the daily traffic volume is greater than or equal to the average traffic volume for that day across 5 years. Yellow and light green are used for intermediary ranges. When a user clicks a single day in the calendar, an hourly breakdown of traffic volume is represented as a heat map with similar color codes as the daily heat map.

The visualization tool also includes a chart-based representation of the analysis between the extreme and base cases of weather elements. When a user clicks on a weather element used in this study, 2 line charts are loaded for a traffic counter. Each line chart spans a 24 hour window in a day. The line charts depict traffic volume under extreme and base weather conditions. The visualization enables a graphical depiction of hours in a day when traffic volume is impacted by extreme weather condition. The geographic information tool is available at http://wxtr.srcc.lsu.edu (screenshots of the tool have been included as Figs. 6 and 7).

7. Conclusion

The impact of extreme weather elements - precipitation, minimum temperature, visibility and wind speed - on hourly traffic volume was analyzed across 48 traffic counters in the city of Atlanta, Georgia, USA. This comprehensive, data-driven study comprised of analyzing the impact of individual weather elements on traffic volume, developing a predictive machine learning model to predict conditions that caused the change in traffic volume, and deriving a decision support tool to visualize traffic volume and its interaction with the weather elements.

The influence of individual weather events on traffic volume was studied using the Wilcoxon test. Based on the statistical analysis, precipitation, temperature and visibility emerged as factors that can cause a significant decline in traffic volume. These weather elements had a significant impact on traffic volume during the evening rush hour times of 6–9 pm. Each weather variable indicated a decline in traffic volume during different hours of a typical week day. Temperature had the greatest impact on traffic volume between 19 and 21 h in the winter months; precipitation indicated the most significant impact on traffic volume between 13 and 22 h; and visibility had an impact on traffic volume between 20 and 22 h. Overall, this approach of individually analyzing historical weather data, enabled the isolation of hours in a week day that are likely to be impacted in terms of traffic volume. These are depicted in the geo-visualization tool and enable traffic planners and transportation officials in Atlanta to make appropriate preparations in the event of extreme weather.

Predictive machine learning models were developed for each of the traffic counters by considering all the weather hazard elements together. This helping in deriving rules or conditions that included combination of weather hazard elements and under different thresholds. Temperature during the winter months and precipitation again emerged as influencing factors that reduced traffic volume. Winter precipitation and extreme minimum temperature combined together to cause a decline in traffic volume. The predictive models were also useful in deriving decision rules that are useful for traffic managers and transportation planners. The statistical analysis and the predictive models have helped derive specific hours in a typical week day during which extreme weather conditions can severely affect traffic volume. The identification of weather element thresholds and affected hours has applications for efficient decision support and transportation planning across major arteries in Atlanta's transportation network.

A decision support tool was created that provides easy access to traffic and weather data. The visualizations are useful for transportation planners, traffic monitoring control rooms and urban planners to use this information for facilitating smooth traffic flows and mitigating bottlenecks caused by weather hazards. The geographic information tool consists of components that help users understand the effect of extreme weather elements on hourly and daily traffic volume. This tool also depicted the comparison of traffic volume between normal and extreme weather conditions to show the hourly impact of each weather variable on traffic volume. Results from this research could ultimately help transportation officials, stakeholders and planners in the overall

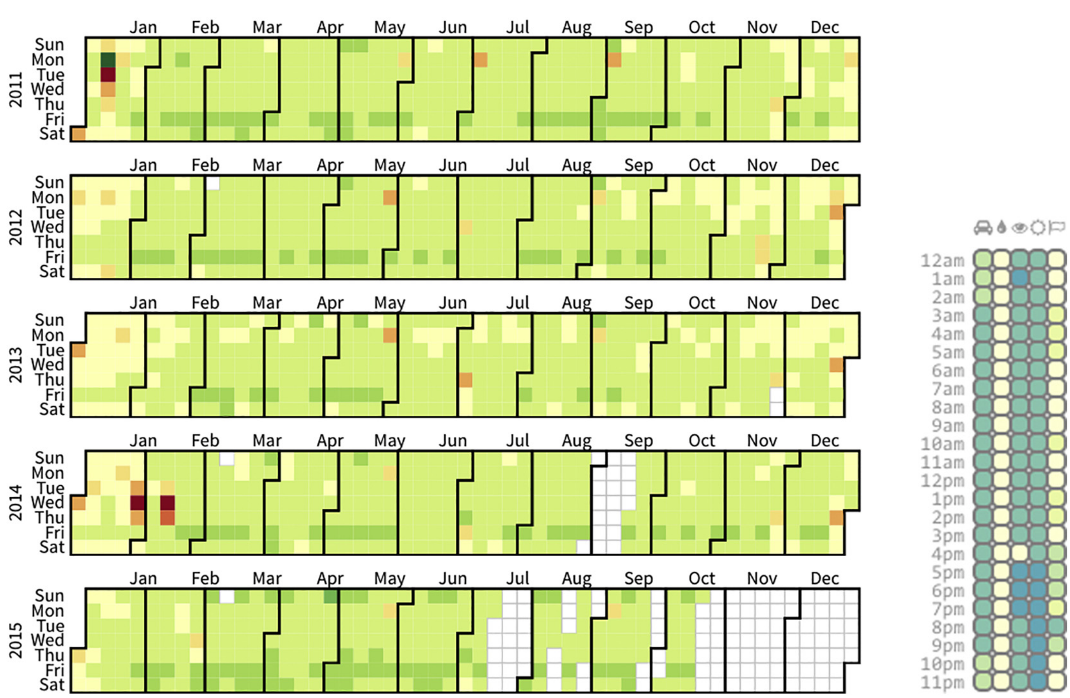


Fig. 7. Daily and hourly summary of traffic volume spanning 2010–2015 for a single traffic station in a calendar like view.

hourly assessment of extreme weather impacts on traffic volume in Atlanta, GA. Future work involves expanding the scope of the weather elements and including winter elements such as snow and ice. Similar studies such as this can also be undertaken for other large metropolitan areas and cities in the US.

References

Agarwal, M., Maze, T. H., & Souleyrette, R. (2005). *Impacts of weather on urban freeway traffic flow characteristics and facility capacity.* Proceedings of the 2005 mid-continent transportation research symposium18–19.

Angel, M. L., Sando, T., Chimba, D., & Kwigizile, V. (2014). Effects of rain on traffic operations on Florida freeways. *Transportation Research Record, 2440*, 51–59. http://dx.doi.org/10.3141/2440-07.

Arguez, A., Durre, I., Applequist, S., Vose, R. S., Squires, M. F., Yin, X., ... Owen, T. W. (2012a). Noaa's 1981-2010 us climate normals: An overview. *Bulletin of the American Meteorological Society, 93*(11), 1687–1697. http://dx.doi.org/10.1175/BAMS-D-11-00197.1.

Bartlett, A., Lao, W., Zhao, Y., & Sadek, A. W. (2013). Impact of inclement weather on hourly traffic volumes in buffalo, New York. *Transportation Research Board 92nd annual meeting, no. 13-3240*.

Breiman, L. (2001). Random forests. *Machine Learning, 45*(1), 5–32. http://dx.doi.org/10.1023/A:1010933404324.

S. C. Calvert, M. Snelder, Influence of weather on traffic flow: An extensive stochastic multi-effect capacity and demand analysis, EUROPEAN TRANSPORT-TRASPORTI EUROPEI (Vol. 60).

Cools, M., Moons, E., & Wets, G. (2010). Assessing the impact of weather on traffic intensity. *Weather Climate and Society, 2*(1), 60–68. http://dx.doi.org/10.1175/2009WCAS1014.1.

Data, C. (2009). Guidelines on analysis of extremes in a changing climate in support of informed decisions for adaptation. http://www.wmo.int/pages/prog/wcp/wcdmp/wcdmp_series/, Accessed date: 16 May 2018.

Dehman, A., & Drakopoulos, A. (2017). How weather events affect freeway demand patterns. *Transportation Research Record, 2615*, 113–122. http://dx.doi.org/10.3141/2615-13.

DoT, U., Nitch, D., Safety, M. R. A., & FLAP, A. I. (2015). Federal highway administration. *Money, 103*, 85–925.

Eisenberg, D. (2004). The mixed effects of precipitation on traffic crashes. *Accident Analysis & Prevention, 36*(4), 637–647.

Ghasri, M., Rashidi, T. H., & Waller, S. T. (2017). Developing a disaggregate travel demand system of models using data mining techniques. *Transportation Research Part A-Policy and Practice, 105*, 138–153. http://dx.doi.org/10.1016/j.tra.2017.08.020.

A. D. Hartkamp, K. De Beurs, A. Stein, J. W. White, Interpolation techniques for climate variables.

Hasan, M. M. (2012). Investigation of the effect of traffic parameters on road hazard using classification tree model. *International Journal for Traffic and Transport Engineering, 2*(3), 271–285.

R. Hranac, E. Sterzin, D. Krechmer, H. A. Rakha, M. Farzaneh, M. Arafeh, et al., Empirical studies on traffic flow in inclement weather.

K. Knapp, L. Smithson, A. Khattak, The mobility and safety impacts of winter storm events in a freeway environment, mid-continent transportation symposium 2000 proceedings, Iowa, USA.

Kulkarni, D., Wang, Y., & Sridhar, B. (2014). Analysis of airport ground delay program decisions using data mining techniques. *14th AIAA aviation technology, integration, and operations conference* (pp. 2025). .

Kwon, T. J., Fu, L., & Jiang, C. (2013). Effect of winter weather and road surface conditions on macroscopic traffic parameters. *Transportation Research Record, 2329*, 54–62. http://dx.doi.org/10.3141/2329-07.

Mitas, L., & Mitasova, H. (1999). *Spatial interpolation, geographical information systems: Principles, techniques, management and applications. 1*, 481–492.

N. Prasad, P. K. Reddy, M. M. Naidu, A novel decision tree approach for the prediction of precipitation using entropy in sliq, in: Computer modelling and simulation (UKSim), 2013 UKSim 15th international conference on, IEEE, 2013, pp. 209–217.

Press release - urban mobility information. https://mobility.tamu.edu/ums/media-information/press-release/, Accessed date: 5 February 2018.

Rusu, C., & Rusu, V. (2006). Radial basis functions versus geostatistics in spatial interpolations. *Artificial intelligence in theory and practice* (pp. 119–128). Springer.

Saha, S., Schramm, P., Nolan, A., & Hess, J. (1994-2012). Adverse weather conditions and fatal motor vehicle crashes in the United States. *Environmental Health, 15*. http://dx.doi.org/10.1186/s12940-016-0189-x.

Samba, D., & Park, B. B. (2010). Probabilistic modeling of inclement weather impacts on traffic volume. *Transportation Research Board 89th annual meeting, no. 10-0646*.

Sathiaraj, D., Cassidy, W. M., Jr., & Rohli, E. (2017). Improving predictive accuracy in elections. *Big Data, 5*(4, SI), 325–336. http://dx.doi.org/10.1089/big.2017.0047.

Smith, A., Lott, N., & Vose, R. (2011). The integrated surface database: Recent developments and partnerships. *Bulletin of the American Meteorological Society, 92*(6), 704–708. http://dx.doi.org/10.1175/2011BAMS3015.1.

Stern, A. D., Shah, V., Goodwin, L., & Pisano, P. (2003). Analysis of weather impacts on traffic flow in metropolitan, Washington DC. *Proceedings of the 19th international conference on interactive information and processing systems (IIPS) for meteorology, oceanography, and hydrology*.

Wilcoxon, F. (1945). Individual comparisons by ranking methods. *Biometrics Bulletin, 1*(6), 80–83.

Yuan-Qing, W., & Jing, L. (2017). Study of rainfall impacts on freeway traffic flow characteristics. *Transportation Research Procedia, 25*, 1533–1543.