REAL-TIME RECOMMENDATIONS
FOR TRAFFIC CONTROL IN AN
ITS SYSTEM DURING AN
EMERGENCY EVACUATION

FINAL REPORT
SEPTEMBER 2021

AUTHORS
Lauren Davis, Xiuli Qu, Younho Seong
North Carolina A&T State University (NCAT)

US DEPARTMENT OF TRANSPORTATION GRANT 69A3551747125
DISCLAIMER

The contents of this report reflect the views of the authors, who are responsible for the facts and the accuracy of the information presented herein. This document is disseminated under the sponsorship of the Department of Transportation, University Transportation Centers Program, in the interest of information exchange. The U.S. Government assumes no liability for the contents or use thereof.
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Subtitle: Phase I Studies</td>
<td></td>
</tr>
<tr>
<td>7. Author(s): Qu, Xiuli; Davis, Lauren; Seong, Younho; Meda, Harshitha; Mhatre, Sachin; Alabi, Miriam; Richmond, David</td>
<td></td>
</tr>
<tr>
<td>9. Performing Organization Name and Address</td>
<td>10. Work Unit No. (TRAIS)</td>
</tr>
<tr>
<td>Center for Advanced Transportation Mobility Transportation Institute 1601 E. Market Street Greensboro, NC 27411</td>
<td></td>
</tr>
<tr>
<td>12. Sponsoring Agency Name and Address</td>
<td>13. Type of Report and Period Covered</td>
</tr>
<tr>
<td>University Transportation Centers Program (RDT-30) Office of the Secretary of Transportation–Research U.S. Department of Transportation 1200 New Jersey Avenue, SE Washington, DC 20590-0001</td>
<td>Final Report: May 2019 – December 2020</td>
</tr>
<tr>
<td>16. Abstract</td>
<td>USDOT/OST-R/CATM</td>
</tr>
<tr>
<td>Recent hurricanes such as Irma (2017) and Florence (2018) caused mass evacuations and the issues occurring during the evacuations have been brought to the public attention. To address these issues, the project team has investigated significant cues for evacuation planners’ decisions using a Linear Lens Model and machine learning algorithms, and has analyzed traffic data during hurricane evacuations in North Carolina to discover spatial-temporal evacuation traffic patterns and create predictive models of hurricane evacuation traffic volumes. Our results show that only one of the seven cues tested (wind speed) contributes to evacuation planners’ decisions and the sensor locations at the same county and adjacent counties form the cluster for evacuation traffic prediction. These models and findings can support the deployment of an effective evacuation and improve the mobility of the people and evacuation resources during a hurricane. We have also proposed and tested the quantitative methods to quantify a hurricane disruption to the U.S. airport network and identify feasible airports to reroute disrupted flights. Our results show that the proposed methods can identify the airports to be disrupted by an approaching hurricane and feasible airports for flight rerouting, which can support airlines administrators to divert flights from the affected airports.</td>
<td></td>
</tr>
<tr>
<td>17. Key Words</td>
<td>18. Distribution Statement</td>
</tr>
<tr>
<td>Hurricane Evacuation, Air Transportation, Road Transportation Networks, Evacuation Traffic Prediction, Decision Making, Rerouting, Resilience, Judgment Model, Machine Learning</td>
<td>Unrestricted; Document is available to the public through the National Technical Information Service; Springfield, VT.</td>
</tr>
<tr>
<td>Unclassified</td>
<td>Unclassified</td>
</tr>
<tr>
<td>21. No. of Pages</td>
<td>22. Price</td>
</tr>
<tr>
<td>80</td>
<td>…</td>
</tr>
</tbody>
</table>
EXECUTIVE SUMMARY

In recent years natural disasters such as hurricanes, floods, and winter storms, have caused significant losses and disruptions to infrastructure, communities, and the economy. Effective preparation and quick response to natural disasters is very important for the mitigation of such losses and disruptions. As an important part of natural disaster preparation and response, evacuations often occur before or after a natural disaster. However, an effective evacuation involves complex planning, preparation and operations. Moreover, diverse human evacuation behavior, such as staying or evacuating from home, taking public transportation or driving private vehicles, and different times and routes to leave from the affected area, complicates the management of mass evacuations.

In the first phase of this CATM project, we investigated the components of North Carolina’s Intelligent Transportation System (ITS), which include smart traffic signals, cameras, meters, and dynamic message signs, and then identified potential actions for traffic control of an ITS during a hurricane evacuation. We also investigated the most affected airlines at the most affected airport during Hurricane Matthew, and proposed and tested the quantitative methods to estimate the number of affected passengers during the hurricane, quantify a hurricane disruption to the U.S. airport network, and identify feasible airports to reroute flights from a disrupted airport. Our results showed that the topology of an airline’s flight network may influence the number of flight passengers affected by a hurricane disruption, and that the proposed quantitative methods can identify the airports to be disrupted by an approaching hurricane and feasible airports for rerouting flights from the disrupted airports.

For hurricane evacuation, we requested and collected historical data related to hurricane evacuations from different sources. These data were used to discover spatial-temporal evacuation traffic patterns, predict the hourly traffic during the hurricane evacuation, and investigate significant cues for Evacuation Planners’ decisions during a hurricane evacuation. Our analysis results showed that the traffic volume at a sensor location could be predicted by the traffic volumes at the sensor locations in the same county and adjacent counties, and revealed that the sensor locations at the same county and adjacent counties form the cluster, and the 25 sensor locations can be grouped into the four clusters.
Our results also showed that only one (wind speed) of the seven cues tested contributes to Evacuation Planners’ decision.

The proposed graph theoretical and rerouting methods can support airports and airlines administrators to recognize the airports that might be affected by an approaching hurricane, and identify potential airports to divert flights from the affected airports. The predictive models created in our studies can be used to manage hurricane evacuations, including proper traffic control, estimating emergency responders needed, and preparing enough evacuation supplies such as gas, water, food, and hotel rooms. These models and findings will help emergency response agencies deploy effective evacuation operations and improve the mobility of the people and evacuation resources during a hurricane. Our results of the first phase of this CATM project have been published as three peer-reviewed conference papers and presented as posters and oral presentations at national professional conferences and regional transportation conferences and symposiums. One journal paper has been submitted. In addition, four graduate students (including one African American student and two female students) and two female undergraduate students (including one African American student) have been involved in this CATM project. The participation of these students can contribute to the diversity of the U.S. transportation workforce in the future.
DESCRIPTION OF PROBLEM

As part of natural disaster preparation and response, evacuations often occur before or after natural disasters such as hurricanes and earthquakes. For example, nearly 7 million Florida residents evacuated from the state during Hurricane Irma (2017), making it the largest hurricane evacuation in the U.S. and causing significant traffic congestion and fuel shortage in Florida. Hurricane Irma also caused significant disruption to air transportation in Florida. Nearly 4000 flights were canceled according to one report by Flight Aware. In addition to canceling and diverting flights, airlines also added flights to get passengers out of the storm’s path and moved their planes (some of which cost $100 million) to other safe cities. United Airlines, Delta, and American added flights in advance of the hurricane to help get stranded passengers out of the storm’s path. The storm’s ripple effects were felt in other cities like Atlanta, where Delta canceled nearly 1000 flights, resulting in many passengers stranded at the airport. During Hurricane Florence (2018), evacuation orders were issued to 25 counties in North Carolina and South Carolina, causing approximately 1 million Carolinians to evacuate from their homes. However, some residents chose to shelter in place despite mandatory evacuation issued in their counties. Traffic congestion and fuel shortages occurred in the eastern coastal areas in North Carolina even though the state government prepared for the evacuation by closing public schools and issuing evacuation orders 3 days before landfall, arranging evacuation paths, and providing evacuation guides. A worse issue was that some people could not evacuate due to the closure of a bridge when they chose to evacuate later. Therefore, it is obvious that effective and proper traffic control is crucial during a mass evacuation.

Recently, information and communication technologies (ICT) have been incorporated in the N.C. transportation infrastructure to build intelligent transportation systems (ITSs), which include smart actuated signals, dynamic message signs, the roadway weather information system, reversible lane systems, and the traveler information management system. These ITSs provide us with opportunities to improve the effectiveness and efficiency of emergency response. For example, smart traffic signals and the traffic coordination network enable emergency vehicles to respond to incidents rapidly. During natural disasters, ITSs can also play an important role in mass emergency evacuations. In this CATM project,
we aimed to develop and integrate ecological models for human evacuation behavior prediction and hurricane evacuation traffic control in intelligent transportation infrastructure. The ultimate goal is to create a human-centered intelligent traffic control recommendation system to support mass evacuations. The research questions of the first phase of our CATM project are:

- What factors and actions in the current NC ITSs should be considered when recommending an initial traffic control plan to prepare for a hurricane evacuation?
- What are significant factors or cues for evacuation planners to make hurricane evacuation decisions?
- How should effective policies be designed for airline carriers to minimize the number of passengers and planes stranded in the path of a hurricane, i.e., scheduling additional flights?
- How are evacuation traffic flows spatial-temporal associated, and how can the evacuation traffic volumes be predicted based on their spatial-temporal associations?
METHODOLOGY AND RESULTS

In this CATM project, we conducted five studies addressing hurricane evacuation and weather-related flights recovery. Before we conducted these studies, we investigated the components of North Carolina’s Intelligent Transportation System (ITS), and identified the potential actions that are supported by the smart components in the current N.C. road transportation system, and can be used to improve traffic control during a hurricane evacuation. We also requested and collected historical data related to hurricane evacuations from different sources to support our CATM studies. The methodology and results of our five studies in Phase 1 of this CATM project are described in detail in the following subsections.

Study 1 – Transportation Network Resilience in Charlotte, North Carolina for Day-to-Day Disruptions

1.1 Research Problem

Daily disruptions to transportation systems are almost inevitable. These disruptions such as weather conditions, human errors, and/or technical failures, that occur on a daily basis, often cause delays and congestion to the transportation system, which ultimately result in the waste of fuel and time for each automotive commuter. Cities with large populations and high utilization of their transportation systems are vulnerable to such disruptions. For example, Charlotte, North Carolina, is a major city, ranking the 16th most populous city in the United States, having a population that is over 872,000 [1]. Being such a large city, it is the hub of major businesses, travel, and events, which often means that there is high utilization of its road network. In other words, there is a high volume of vehicles that travel through Charlotte’s transportation system.

Transportation network disruptions can cause traffic congestion and delays. According to the 2019 Urban Mobility Report from Texas A&M Transportation Institute, it is estimated that an auto commuter in Charlotte was delayed 57 hours as well as consumed 22 gallons of excess fuel yearly due to congestion, resulting in a yearly congestion cost of $1,269 per auto commuter [2]. For this reason, this study aimed to analyze the network connectivity and topology as well as the traffic volumes of the road transportation network of...
Charlotte, NC, to quantify the resilience of the overall transportation network and to identify the road segments that are the most critical to the efficiency of the transportation network with respect to daily disruptions. In this study, two approaches, betweenness centrality and hierarchal cluster analysis, were used to identify critical road segments in Charlotte.

1.2 Methodology
To identify critical road segments and examine the resiliency of the transportation network of Charlotte, we first define the scope of the transportation network. In this study, we focused on the U.S. interstates and highways since they are the most driven upon and have the highest traffic volume on average. The transportation network we define consists of U.S. interstates 85, 485, 77, 277, and U.S. highways 21, 29, and 74. The links in our defined network are the road segments, while the nodes are the exits or the intersection of two road segments. We also defined a set of starting nodes and ending nodes. The starting and ending nodes represent the most populated cities surrounding Charlotte, including Concord City, Gastonia, Huntersville and Indian Trail in North Carolina, and Rock Hill in South Carolina. We then collected and processed the data to compute the weighted edge betweenness centrality of each road segment, as well as to conduct a hierarchal cluster analysis. The results from this study can be used to make recommendations for strategic route planning, which could ultimately aid in the reduction of excess fuel consumption, delay, and congestion costs per auto commuter.

1.2.1 Data Preparation and Processing
In this study, Google Maps was used to determine the latitude and longitude of each intersection, which were later used to determine the distance of each road segment in the transportation network. After data regarding the intersections and links had been collected, Python 3.4 and the package, Networkx were utilized to compute the betweenness centrality for the links.

The AADT for 2018 was collected from the NCDOT’s website, which was used as a variable in the hierarchal cluster analysis. Both variables, AADT and distance, were normalized to [0, 1] to reduce the influence of outliers and avoid that one variable is much
more dominant than the other. The variables were normalized using Equation (1). Note that the descriptive statistics and distributions of both variables can be seen in Section 1.3.

\[ x_{\text{norm}}^i = \frac{x^i - x_{\text{min}}}{x_{\text{max}} - x_{\text{min}}} \]  

(1)

1.2.2 Edge Betweenness Centrality Indicator

The weighted edge betweenness centrality algorithm is utilized to assess road resilience by identifying, topologically, vulnerable road segments and intersections in Charlotte’s highway transportation network. Let \( G(V, E) \) represent the transportation network of interest, where \( V = \{v_1, v_2, \ldots, v_i, s_1, s_2, \ldots, s_n, t_1, t_2, \ldots t_m\} \) denotes the set of locations and intersections (nodes) and \( E = \{e_1, e_2, \ldots, e_j\} \) represents the set of road segments (links) between each node. Each edge \( e_j \) is associated with a weight \( w_j > 0 \), which represents the traveling distance or time through the edge.

Given a transportation network \( G(V, E) \), the weighted edge betweenness centrality algorithm computes the betweenness centrality for each road segment within the network using

\[ C_B(e) = \sum_{s\in S, t\in T} \frac{\sigma(s, t|e)}{\sigma(s, t)} , \]  

(2)

where \( s \in \{s_1, s_2, \ldots, s_n\} \subset V, t \in \{t_1, t_2, \ldots t_m\} \subset V, \sigma(s, t) \) denotes the number of shortest paths from node \( s \) to node \( t \), and \( \sigma(s, t|e) \) represents the number of the shortest paths from \( s \) to \( t \) including edge \( e \). A shortest path from node \( s \) to node \( t \) is defined as a path from \( s \) to \( t \) with a minimum total weight of the edges in the path [3]. Using the edge betweenness centrality algorithm, a betweenness centrality value is computed for each road segment in the Charlotte transportation network. The road segments with the highest betweenness centrality values are the most critical to the overall topological structure of the network. Although there are other centrality indicators, such as degree centrality and closeness centrality, to consider, we chose edge betweenness centrality since it allows us to identify critical road segments such that the disruption (removal) of those road segments (links) to the transportation network could affect the fluidity of the traffic between many pairs of intersections through the shortest paths between them [4].
1.2.3 Hierarchal Cluster Analysis

A hierarchal cluster analysis was conducted using Python 3.4 to classify and identify critical road segments in the transportation network of Charlotte. The parameters used were the annual daily traffic for 2018 and the distance of each road segment (in miles). The dendrogram (Figure 1.1) was computed using the Euclidean distance formula. From the dendrogram, we conclude that the $k$ value (number of clusters) should be 3.

![Dendrogram of road segments using Euclidean distance.](image)

1.3 Results and Discussion

When exploring the data, we found that there were some extreme outliers in the average annual daily traffic (AADT) for 2018. Although the AADT and distance variables used contained extreme outliers, there were no erroneous values or inputs; the data occurred naturally given the dataset. Thus, none of the extreme outliers were dropped; however, both variables were normalized. The descriptive statistics of the AADT and the distance of each road segment are shown in Table 1.1. Figures 1.2 and 1.3 show the distribution of both variables, AADT, and distance.
Table 1.1: Descriptive statistics of the distances and average annual daily traffic for each road segment

<table>
<thead>
<tr>
<th>Descriptive Statistics</th>
<th>Distance (miles)</th>
<th>AADT 2018</th>
</tr>
</thead>
<tbody>
<tr>
<td>count</td>
<td>36</td>
<td>36</td>
</tr>
<tr>
<td>mean</td>
<td>4.386</td>
<td>100546</td>
</tr>
<tr>
<td>std</td>
<td>2.880</td>
<td>39155</td>
</tr>
<tr>
<td>min</td>
<td>0.4</td>
<td>15500</td>
</tr>
<tr>
<td>0.25</td>
<td>1.95</td>
<td>83604</td>
</tr>
<tr>
<td>0.5</td>
<td>3.9</td>
<td>99833</td>
</tr>
<tr>
<td>0.75</td>
<td>5.65</td>
<td>126786</td>
</tr>
<tr>
<td>max</td>
<td>12.1</td>
<td>174500</td>
</tr>
</tbody>
</table>

Figure 1.2: Histogram of the distances of each road segment.
After conducting the betweenness centrality analysis, we found that the road segments received a betweenness centrality value of either 0, 1, or 2. Figure 1.4 shows the transportation network with the betweenness centrality values colored as blue for 0, red for 1, and green for 2. The higher the betweenness centrality, the more critical the road segment to the entire transportation network.

Figure 1.4: Charlotte highway segments grouped based on betweenness centrality.
To identify critical road segments, a hierarchal cluster analysis was conducted. After computing and examining the dendrogram, we allocated the observations to 3 different clusters, i.e., green, blue, and red clusters in Figure 1.5. In the figure, we notice that road segments that belong to the green cluster have the highest AADT, while those grouped in the red cluster, on average, have a lower AADT but a longer distance. On the other hand, the blue cluster contains road segments that have a higher AADT, on average, but a shorter distance. Figure 1.6 shows the transportation network considering the cluster groupings.

Figure 1.5: Clustering of road segments.

Figure 1.6: Charlotte highway segments clustered based on 2018 AADT.
Our results show that road segments of U.S. Interstates 485 and 77 and U.S. Highway 74 have the highest betweenness centrality value. This reveals that those are the most central links of the topological structure in the Charlotte highway transportation network. The disruption (removal) of the most critical links could affect the fluidity of the traffic through the shortest paths between many pairs of origins and destinations. Based on the hierarchal cluster analysis results, we conclude that the road segments belonging to the green cluster are the most critical since these road segments have the highest AADTs compared to the other road segments although they are not the longest road segments. This means that these road segments are more likely to be congested with an occurrence of disruption compared to the other road segments due to high traffic volume. The links that belong to the green cluster are road segments of U.S. Interstate 85 and 77 and U.S. Highway 74.

**Study 2 – Data-Driven Approach to Estimate Airline Passengers Disruption**

**2.1 Research Problem**

The United States (U.S.) airline companies serve millions of passengers and contribute to the country’s economy in billions every year [5, 6]. To ensure passenger safety, convenience, and economic value, uninterrupted and efficient functioning of airline services is necessary. However, disruptions are not uncommon in an air transportation system, and disruption of any kind impacts planned flight schedules and can also lead to downstream effects. As per Bureau of Transportation Statistics (BTS) annual reports, more than 50% of disruptions in the aviation system are weather-related [5]. Hurricanes are one of the weather-related incidents that can significantly disrupt air travel due to their inclement nature. Moreover, some studies pointed out the fact that the network structure of air transportation or airlines might influence the general functionality of airlines operations [7]. During a disruption, different recovery actions could be implemented, and over the years, different planning and recovery tools or methods have been proposed [8-10]. Research on developing robust recovery methods and flight schedules may be essential for alleviating an airport system-related disruption. To achieve this, a prior step of appropriate data analysis of airport or airline big data might be influential in recommending certain actions during a disruption. In this study, we consider two-month flight data pertaining to four main airports in the southeast
region of the U.S. During the time frame considered in this dataset, Hurricane Matthew 2016 occurred. In this research, we set out to achieve the following research objectives:

1. To identify the most affected airlines and understand the influence of their network structure on the impact created during a hurricane disruption.
2. To estimate the possible number of impacted passengers during a hurricane disruption.

2.2 Methodology

The data used for this study was purchased from a leading provider of digital information called OAG [11]. The data set consists of 117,644 rows and 18 columns of information, for the period between September 1st, 2016 to October 31st, 2016. The scope of the data is limited to the departures and arrivals of four major airports in the Southeastern United States for all airline carriers. These airports are Orlando International Airport (MCO), Raleigh-Durham International Airport (RDU), Norfolk International Airport (ORF), and Ronald Reagan Washington National Airport (DCA). Hurricane Matthew was characterized as a post-tropical cyclone between September 28th, 2016 and October 9th, 2016. Hurricane Matthew in 2016 had its impact on states of Florida, Georgia, North Carolina, South Carolina, and certain parts of Virginia between October 7th, 2016 and October 9th, 2016 [12, 13]. Therefore, we selected this time period for our analysis. The data was filtered to remove rows with erroneous/empty information for further analysis (i.e., “NULL” values). We used the Networkx [14] package and other Python packages to conduct the data analysis.

Among the four airports present in our data set, we identified the most affected airport based on the change in their flight connections. The most affected airlines were recognized based on the summation of the count of flights for every airline for each selected day and affected airport. The network structure of the identified affected airlines was further analyzed to understand its impact on operations. To estimate the possible number of affected passengers, we have developed a set of mathematical equations. To conduct experiments with the developed equations, we have utilized the available number of seats in our dataset and load factors from the BTS website [15]. For the detailed version of our methods, please refer to our published work [16].
2.3 Results and Discussion

As per our analysis, during Hurricane Matthew, 2016, Orlando International Airport (MCO) was the most affected airport among the four airports present in the data set. MCO flight data was analyzed to understand different affected airlines. From Figure 2.1, we recognized the trend for every airline as the visualization shows the number of departure flights for each airline. As far as MCO departures are concerned, the most affected airlines in terms of the number of flights were American Airlines (AA), Delta Airlines (DL), Frontier Airlines (F9), JetBlue Airways (B6), Southwest Airlines (WN), Spirit Airlines (NK), and United Airlines (UA). Figures 2.2 – 2.8 show the network structures of these seven airlines constructed with the flight data present in the data set. Southwest was the most affected carrier with the highest number of flights canceled on October 6th, 7th, and 8th followed by JetBlue Airways. Spirit Airlines and Frontier Airlines were the least affected among the seven carriers. From examining the network topology and patterns of the seven affected airlines, most of them follow either a traditional hub-spoke model or focus city model except Southwest. Southwest Airlines follows a rolling hub model [17, 18]. However, further analysis of the network topology of Southwest along with other airlines is required for a stronger validation.

Figure 2.1: Departures at MCO before, during, and after the hurricane.
Figure 2.2: Network structure for American Airlines

Figure 2.3: Network structure for Delta Airlines
Figure 2.4: Network structure for Frontier Airlines

Figure 2.5: Network structure for JetBlue Airways
Figure 2.6: Network structure for Southwest Airlines

Figure 2.7: Network structure for Spirit Airlines
Further, the predicted number of passengers for the seven affected airlines at MCO was determined. A comparison was made between these seven airlines based on the computation of the percentage of the predicted number of passengers. Figure 2.9 shows the percentage of the predicted number of affected passengers for departures of these seven airlines before, during, after the hurricane. The negative percentage value signifies the possible percentage of affected passengers for an airline carrier and excessively planned passengers if it is positive. From Figure 2.3, unlike other airlines, Southwest Airlines took almost two days to recover after the disruption at MCO caused by Hurricane Matthew. As MCO was completely closed on October 7th, 2016, it is evident that all the seven other carriers were affected only on this day.
Figure 2.9: Percent of affected passengers for seven carrier airlines departures at MCO before, during, and after the hurricane.

From our analysis, we may relate the airline’s network topology with its possible number of affected passengers. However, extensive experimentation is required to develop and validate this relation. Moreover, the data-driven approach proposed in this study serves as a preliminary step to understand factors that influence passenger scheduling before, during, and after a disruption. In this study, we identified MCO to be the most affected airport and Southwest to be the most affected airline at this airport. Being aware of such information might aid in alleviating passenger disruption to a certain extent. As the data-driven approach does not discriminate the type of airport and kind of disruption, this methodology can also be extended to any other airport and disruption type.

Study 3 – A Graph Theoretical Approach Integrating Geospatial Information to Analyze Airport Network Disruptions

3.1 Research Problem
Disruptions in an air transportation system are not uncommon in occurrence, and especially weather-related ones are harder to predict and control due to their stochastic nature. Hurricane and tropical storms are some of the weather-related disruptions that can lead to
significant economic impact along with root and propagated disruptions [19, 20]. In the past, a significant number of flights were canceled and delayed during certain hurricanes [21, 22]. The weather disruptions can also lead to significant delay costs, loss of passengers’ goodwill, distorted operations and many more [23]. Due to the inclement and stochastic nature of certain weather-related events, it is not always possible to eliminate disruptions of this kind. However, locating airports in an airport network that might be impacted ahead of time may help in either planning or providing opportunities for recovery of a disruption. Additionally, there will be airports that will be affected due to these originally impacted airports. Therefore, capturing these two types of airports in an airport network might help in efficient planning, rescheduling, and recovery when a disruption occurs. Moreover, there might be a scenario where a flight path gets affected due to its proximity to a disrupted geographical zone. There are certain real-time situations where flights were diverted using different airports to avoid disruption or any adverse events, and it is essential to consider geography to include such scenarios [24, 25]. From our literature review, we have identified that geography consideration is absent when conducting airport network analysis related research [26-28]. In this study, we introduce this missing element as a geospatial scenario. Overall, we propose a graph theoretical method integrating geospatial elements to quantify airports from two different perspectives: disruptor and disruptee. A disruptor is an airport that gets impacted during a disruption along with flight routes connecting from or via this airport. A disruptee is an airport that is impacted by a disruptor. A pair of disruptor and disruptee is two airports where the latter is dependent on the former. In addition, we propose a rerouting method to discover feasible airports for rerouting from a disrupted airport.

3.2 Methodology

In this study, we propose theoretical methods to quantify airports in an airport network from two different perspectives. In addition, we propose a rerouting method to identify feasible airports to reroute from a disrupted airport. To quantify the extent to which an airport is a disruptor, we propose a mathematical equation with three components: the immediate, the intermediate, and the geospatial components. The immediate component captures all immediate connections from an airport, whereas the intermediate component captures flight
routes for which an airport is an intermediate connection. The geospatial element captures flight routes that lie within a disrupted airport zone. The higher value of the disruptor equation indicates a highly disrupted airport and vice versa if it is a lower value.

To estimate the extent to which an airport could be a disruptee, we propose the disruptee equation. The disruptee equation consists of three components: the destination component, the transition component, and the geographical component. The destination component captures the number of flight paths originating from an airport that are affected due to a disruptor airport being their final destination. The transition component captures flight routes from an airport for which disruptor airports act as an intermediate connection. The geographical component estimates the flight paths from an airport that fall under the disrupted zone of the airport. Once the airports in an airport network are classified as disruptors and disruptees, rerouting passengers to different airports might aid in either avoiding or alleviating passenger disruption.

As a part of the rerouting method, we propose the rerouting metric that helps in identifying feasible airports for rerouting from a disrupted airport. It consists of three components: the change in disruptee equation value, the network proportion, and the airline network proportion. The change in disruptee equation value helps in quantifying an airport as either less or more disrupted than a disruptee. The network proportion helps in understanding if a similar set of destination airports are being operated from disruptee and rerouting airports. The airline network proportion indicates if the same airline companies are operating to the destinations from disruptee and rerouting airports.

To implement the proposed theoretical methods, we have constructed an airport network from world airlines routes data extracted from an open online source called “openflights.org”. The U.S. airport network constructed based on this data consisted of 2465 routes between 354 airports on 60 airlines. The hurricane forecast data available on National Hurricane Center is extracted to know storm location and time. For the detailed version of the proposed methodology and the data extraction process, please refer to the dissertation [29]. All the equations, methods, and analyses are developed as algorithms using Python programming [30].
3.3 Results and Discussion

To test the proposed methods, we have conducted a series of experiments considering hypothetical hurricane impact distance. We considered Fort Lauderdale/Hollywood International Airport (FLL) as a disruptor and 62 miles as a hurricane impact distance. With FLL as a disruptor, we have computed the disruptee equation values for remaining airports in an airport network. The top 20 disruptees with their respective component values are shown in Table 3.1. Though the destination and the transition component values are small for certain airports, they still appear as a top disruptee due to the geographical component value. This shows that there is a good chance of airports to be geographically impacted. The rerouting method was used to discover feasible airports within a distance of 50 miles for these top 20 disruptees. As a result, we have identified feasible airports for rerouting flights from each disruptee. The component values of the rerouting metric for the identified airports are as shown in Figure 3.1. Each airport is represented by its IATA code.

Table 3.1: Top 20 disruptees when FLL is a disruptor at a hurricane impact distance of 62 miles.

<table>
<thead>
<tr>
<th>Airport (IATA code)</th>
<th>Destination Component</th>
<th>Transition Component</th>
<th>Geographical Component</th>
<th>Disruptee Equation Value</th>
<th>Airport Name</th>
</tr>
</thead>
<tbody>
<tr>
<td>MIA</td>
<td>0.00000</td>
<td>0.02912</td>
<td>1.00000</td>
<td>0.50925</td>
<td>Miami International Airport, FL</td>
</tr>
<tr>
<td>PBI</td>
<td>0.00000</td>
<td>0.03196</td>
<td>0.00000</td>
<td>0.50275</td>
<td>Palm Beach International Airport, FL</td>
</tr>
<tr>
<td>ORH</td>
<td>0.50000</td>
<td>0.49150</td>
<td>0.03437</td>
<td>0.41035</td>
<td>Worcester Regional Airport, MA</td>
</tr>
<tr>
<td>IAG</td>
<td>0.20000</td>
<td>0.60032</td>
<td>0.03789</td>
<td>0.38100</td>
<td>Niagara Falls International Airport, NY</td>
</tr>
<tr>
<td>LBE</td>
<td>0.33333</td>
<td>0.42184</td>
<td>0.03073</td>
<td>0.33681</td>
<td>Arnold Palmer Regional Airport, PA</td>
</tr>
<tr>
<td>PBG</td>
<td>0.20000</td>
<td>0.38758</td>
<td>0.03908</td>
<td>0.28485</td>
<td>Plattsburgh International Airport, NY</td>
</tr>
<tr>
<td>EYW</td>
<td>0.12500</td>
<td>0.14984</td>
<td>0.29092</td>
<td>0.26624</td>
<td>Key West International Airport, FL</td>
</tr>
<tr>
<td>SWF</td>
<td>0.20000</td>
<td>0.24047</td>
<td>0.03509</td>
<td>0.21616</td>
<td>New York Stewart International Airport, NY</td>
</tr>
<tr>
<td>AVL</td>
<td>0.12500</td>
<td>0.17682</td>
<td>0.03611</td>
<td>0.15903</td>
<td>Asheville Regional Airport, NC</td>
</tr>
<tr>
<td>ISP</td>
<td>0.09091</td>
<td>0.13293</td>
<td>0.09280</td>
<td>0.15144</td>
<td>Long Island MacArthur Airport, NY</td>
</tr>
<tr>
<td>ACY</td>
<td>0.11111</td>
<td>0.13678</td>
<td>0.03450</td>
<td>0.13376</td>
<td>Atlantic City International Airport, NJ</td>
</tr>
<tr>
<td>HPN</td>
<td>0.06667</td>
<td>0.10938</td>
<td>0.07984</td>
<td>0.12382</td>
<td>Westchester County Airport, NY</td>
</tr>
<tr>
<td>LEX</td>
<td>0.07692</td>
<td>0.10731</td>
<td>0.03632</td>
<td>0.10619</td>
<td>Blue Grass Airport, KY</td>
</tr>
<tr>
<td>RIC</td>
<td>0.05556</td>
<td>0.07678</td>
<td>0.08405</td>
<td>0.10527</td>
<td>Richmond International Airport, VA</td>
</tr>
<tr>
<td>TTN</td>
<td>0.07692</td>
<td>0.10428</td>
<td>0.03480</td>
<td>0.10400</td>
<td>Trenton-Mercer Airport, NJ</td>
</tr>
<tr>
<td>GNV</td>
<td>0.00000</td>
<td>0.02736</td>
<td>0.19887</td>
<td>0.10283</td>
<td>Gainesville Regional Airport, FL</td>
</tr>
<tr>
<td>BDL</td>
<td>0.04348</td>
<td>0.06736</td>
<td>0.09290</td>
<td>0.09970</td>
<td>Bradley International Airport, CT</td>
</tr>
<tr>
<td>MYR</td>
<td>0.05000</td>
<td>0.09627</td>
<td>0.03527</td>
<td>0.08855</td>
<td>Myrtle Beach International Airport, SC</td>
</tr>
<tr>
<td>TYS</td>
<td>0.05882</td>
<td>0.08556</td>
<td>0.03567</td>
<td>0.08745</td>
<td>McGhee Tyson Airport, TN</td>
</tr>
<tr>
<td>JAX</td>
<td>0.04167</td>
<td>0.06560</td>
<td>0.07049</td>
<td>0.08707</td>
<td>Jacksonville International Airport, FL</td>
</tr>
</tbody>
</table>
The airline network proportion value is zero for BUF airport when IAG is a disruptee. This shows that there are no common airline companies that operate to the destinations of BUF and IAG. The rerouting metric component values provide guidance in determining the most feasible airport when there are multiple choices for rerouting for certain disruptees. In Figure 3.1, ISP and HPN airports have multiple rerouting choices, whereas IAG and PBG airports have a single one.

Further experimentation revealed that ORD, DEN, and ATL are the top three disruptors in the U.S. airport network irrespective of the hurricane impact distance. Moreover, mostly large hub airports appeared as top 20 disruptors. Based on the adapted Hurricane Matthew, 2016 scenario, our results showed that the change in hurricane impact distance and forecast track do have an impact on identifying potential disruptors and their respective disruptees. This, in turn, impacted in discovering potential airports for rerouting. Please refer to the dissertation for detailed experimental results and findings [29].

Figure 3.1: Rerouting airport choices within 50 miles radius for disruptees identified when FLL is a disruptor.

4.1 Research Problem
Intelligent Transportation Systems play an important role in mass emergency evacuations. In dealing with all humanitarian aspects of emergencies, emergency preparedness and disaster management are essential. Evacuation Planners (EP) and individuals’ decisions during emergency preparation involve complex behavioral factors [31]. This study proposes to characterize EPs’ decision-making behavior during emergency evacuation using machine learning algorithms and statistical methods. In doing so, a decision-making tool, the Brunswik Lens model, is used to describe the correlation between the environment and the behavior of organisms in the environment. Additionally, identifying significant cues for residents to decide dynamic evacuation routes under an uncertain environment and having incomplete information is somewhat limited due to how stochastic decision-making can be.

4.2 Methodology and Results
In the modeling process, a judgment model is created with cues influencing the judgment. The researchers [32-34] considered the variables including wind speed, rainfall, number of households affected, flood level, median household income, and the poverty level. These variables (cues) were found to influence evacuation in North Carolina while preparing for an impending hurricane. The goal is to create a judgment model for an emergency to gain insight into the decision behavior of the various entities involved when the environment presents multiple cues.

The Judgment data about Hurricane Matthew was retrieved from multiple federal, state governments and other websites. All data collected are related to counties that were affected by Hurricane Matthew in the state of North Carolina. The weather-related data were collected from October 8th to 9th, 2016, when the hurricane approached. Socio-economic data such as poverty, median household income level, poverty level data, and disaster data were obtained for the year 2016 from the United States Census Bureau (USCB), United States Department of Agriculture Economic Research Service (USDA), and WebEOC [35], respectively. Weather data were mainly taken from National Oceanic and Atmospheric
Administration (NOAA) - and United States Geological Survey, respectively. The data collected consisted of seven variables and 42 observations (counties). All seven variables examined were numerical and were normally distributed.

4.2.1 Brunswik Lens model

Figure 4.1 shows the Brunswik Linear Lens Model representation of single-system design (ecological criterion unavailable or not of interest) [36]. The Lens Model framework and its related parameters can capture and quantify judgment policies. Figure 4.1(a) represents the model of the criterion or environment. This model describes the relationship between the ecological criterion value (e.g., individuals' decision to evacuate or not) and the cue values accessible at the time a judgment is made. Figure 4.1(b) represents the judge's policy or strategy. It describes the relationship between the cue values and the criterion value. In Figure 4.1(b), the judgments are related to each cue, known as cue utilization validity. The pattern of cue utilization demonstrated by a judge determines the judgment policy, \( Y_S \), represents the EPs' judgment and is modeled as a linear combination of a set of \( k \) cues \( (X_i, i = 1, \ldots, k) \).

\[
Y_S = \sum_{i=1}^{k} \beta_{si} X_i + e \tag{3}
\]

In Equation (3), the \( \beta_{si} \) denotes the weights of the cues that contribute to the judge's decision, and \( e \) represents the scope to which the cognitive model misses the actual value when trying to predict the judgment, \( Y_S \) [36]. Thus if \( \hat{Y}_S \) denotes the cognitive model, then

\[
Y_S = \hat{Y}_S + e \tag{4}
\]

The correlation between \( Y_S \) and \( \hat{Y}_S \) denoted by \( R_s \) measures the cognitive control with which a judgment strategy is executed. Consistency refers to the similarity between judgments made to repeated profiles of cue information [36].

\[
R_s = \text{corr}(Y_S, \hat{Y}_S) \tag{5}
\]

4.2.2 Machine learning algorithms

Supervised learning is a machine learning technique that uses computational learning theory, pattern recognition, and algorithm construction to map inputs to the output. Six (6) machine learning algorithms were also used to create the judgment model, namely linear discriminant
analysis (LDA), K-nearest neighbor (KNN), logistic regression (LR), classification and regression tree (CART), Naive-Bayes (NB), support vector machine (SVM). Python packages provide algorithms built in their libraries. Figure 4.2 shows the steps in Machine Learning.

Figure 4.1: Brunswik Linear Lens Model representation of single system design (ecological criterion unavailable or not of interest).
4.2.3 Evaluation metrics

The two performance metrics for the evaluation of models were Cross-Validation Accuracy (CVA) and Prediction/Model Accuracy (Accuracy). A 10-fold cross-validation is used in this study, and the accuracy is defined as:

\[ \text{Prediction Accuracy (PA)} = \frac{TP + TN}{TP + TN + FP + FN}, \]  

where \( TP, TN, FP \) and \( FN \) represent true positive, true negative, false positive, and false negative, respectively.

4.2.4 Lens model parameter, \( \bar{Y}_S \) and \( R_s \)

The supervised machine learning techniques; LR, LDA, KNN, CART, NB, SVM as shown in Figure 4.3, also represented interesting accuracies but did not perfectly estimate the judgment model since model accuracies were 0.66, 0.72, 0.69, 0.6, 0.65 and 0.76, respectively. Figure 4.3 displays the cross-validation accuracy for all the machine learning models and data types after 75% of the judgment data is used in training them, and the model accuracy after 25% of the data is used to test the model. In the Judgment model, SVM generated a high CVA of approximately 73%, and LR, LDA and NB generated a PA of 72.7%.

The Lens model parameters are computed using the logistic regression for the judgment model. This technique is used as an idiographic-statistical approach to understanding the characteristics and conditions of individuals' behavior. This approach also helps to capture judgment policies as well as aspects of the judgment process. Logistic regression provides a statistical model that, in its basic form, uses a logistic function to model a dichotomous dependent variable.
Dependent Variable
In regression analysis, logistic regression (or logit regression) estimates the parameters of a logistic model (a form of binary regression). Mathematically, logistic regression estimates a multiple linear regression function defined as:

$$\text{logit}(p) = \log \left( \frac{p(y=1)}{1-p(y=1)} \right) = \beta_0 + \beta_1 x_{12} + \beta_2 x_{12} + \cdots + \beta_p x_{in} \text{ for } i = 1, \ldots, n.$$  \hfill (7)

In the selection of cues, backward elimination regression was used. In the backward elimination method, all the cues are initially used to build the model. Subsequently, the cue with the highest p-value is eliminated if the p-value is greater than the significance level (0.05). This is repeated, and a single cue is eliminated for each iteration until the cue shows a p-value less than 0.05. However, in this application, all the cues were eliminated except the wind speed. This means that only one of the cue weights contributed well enough to the model. The pseudo-$R^2$ value validates the performance of the model.

Figure 4.4 shows the results of logistic regression for judgment when all cues were used. Figure 4.4 shows the $R^2$ for the judgment model ($\hat{Y}_S$) and a correlation ($r$) of 50.2%. For the judgment model ($\hat{Y}_S$), $R^2$ values were very low indicating that the cues were not able to produce the best model. Table 4.1 shows the weighting applied to each cue that
contributed to the EPs’ judgment. These weights explain how the judge's policy was made based on the cue utilization validities.

![Logit Regression Results](image)

Figure 4.4: Logistic regression output for judgment when all cues were used

Table 4.1: Cue weights

<table>
<thead>
<tr>
<th>Cues</th>
<th>Relative(cue) weight</th>
</tr>
</thead>
<tbody>
<tr>
<td>Shelters</td>
<td>0.144</td>
</tr>
<tr>
<td>Wind speed (mph)</td>
<td>0.358</td>
</tr>
<tr>
<td>Rainfall (in)</td>
<td>0.217</td>
</tr>
<tr>
<td>Households</td>
<td>0.141</td>
</tr>
<tr>
<td>Food level (ft)</td>
<td>0.016</td>
</tr>
<tr>
<td>Median Income ($)</td>
<td>0.031</td>
</tr>
<tr>
<td>Poverty Percent (%)</td>
<td>0.092</td>
</tr>
</tbody>
</table>

**Independent variables**

The total number of independent variables, also known as the cues in this context, is seven (7), as shown in Appendix A. Checking the correlation between the variables (cues) helps eliminate all redundant cues. The correlation between two variables is defined as:

\[
 r_{xy} = \frac{\Sigma(x_i-x)(y_i-y)}{\sqrt{\Sigma(x_i-x)^2 \Sigma(y_i-y)^2}}, \tag{8}
\]
where

- $r_{xy}$ the correlation coefficient of the linear relationship between the variables $x$ and $y$
- $X_i$ – the values of the $x$-variable in a sample
- $\bar{X}$ – the mean of the values of the $x$-variable
- $y_i$ – the values of the $y$-variable in a sample
- $\bar{y}$ – the mean of the values of the $y$-variable

There were no high positive correlations among cues, as shown in Figure 4.5, indicating that all cues are independent and can be used for the prediction.

Figure 4.5: Correlation matrix for the cues
Study 5 – Visualization and Prediction Models of Hurricane Evacuation Traffic in Eastern North Carolina

5.1 Research Problem

Natural disasters affect millions of people, communities, and critical infrastructures every year in the world. According to the Center for Research on Epidemiology of Disaster (CRED), in 2019, there were 396 climate-related and geophysical disaster events recorded in the International Disaster Database (EM-DAT) with 11,755 deaths and over 95 million people affected across the world [37]. Over the last decade, the number and intensity of natural disasters, especially hurricanes, concerning the U.S. east coast have increased notably. According to the statistics, on average, 20 Hurricanes and eight floods have affected the United States every year since 2007. Many of them caused significant disruptions and losses to the communities, infrastructure, and economy. For example, recent hurricanes such as Matthew (2016), Harvey (2017), Irma (2017), Florence (2018), Michael (2018), Dorian (2019) affected the U.S. These hurricanes caused many casualties, damaged thousands of houses in the United States, and significantly disrupted critical infrastructures in the East Coast and Gulf Coast areas. These hurricanes also caused mass evacuations, in which millions of people traveled from the affected areas to safe locations before the landfalls.

Over the years, government agencies have guided millions of people for hurricane evacuation through road transportation systems in the best possible way. Similarly, in response to an approaching hurricane in North Carolina (N.C.), the government agencies, specifically disaster management agency which activates during the forecast of any hurricane, need to plan and carry out many preparation activities. One of the critical preparation activities is to plan and manage an effective evacuation in the potentially affected regions. Proper traffic control, sufficient emergency responders, and enough evacuation supplies such as gas, water, food, and hotel rooms are crucial for an effective mass evacuation. To decide traffic control policies and predict resources needed during a hurricane evacuation, the first step is to estimate the evacuation traffic volume and pattern when a hurricane approaches. To meet this need, in this study, we analyzed traffic flow data during Hurricane Florence evacuation in N.C. to discover space-time evacuation traffic patterns. Hourly traffic flow data during the hurricane evacuations were collected and provided by the
Departments of Transportation (DOTs) of North Carolina. Besides, we investigated the association of sensor locations to predict the hourly traffic during a hurricane evacuation.

5.2 Methodology
The primary sources of the data for this study are traffic counts from the traffic survey group of the NC Department of Transportation (NCDOT). The counties in N.C. are divided into fourteen DOT divisions, each division covering a group of the counties [38]. These divisions are also referred to as districts. The districts are allotted numbers starting from the eastern coast to western mountain regions. We requested the six coastal districts’ traffic volume data for Hurricane Florence and Hurricane Dorian. The data sets cover traffic volume one week before hurricane landfall, one week during hurricane landfall, and one week after the hurricane landfall. In the current study, we analyzed the Hurricane Florence (2018) traffic volume data during a hurricane evacuation, i.e., five days before the hurricane landfall at the N.C. coastline. The traffic volume data of six coastal districts cover 42 counties in eastern North Carolina. The data was provided for thirty-eight sensor locations, consisting of one EXCEL file for each sensor location. Each EXCEL file contains the sensor location county, route number, lane direction, district number, and hourly traffic volume data with dates. Figure 5.1 illustrates the locations of some sensors on the different routes in North Carolina.

![Figure 5.1: Illustration of different sensor locations in North Carolina.](image-url)
Predictive modeling was used in this study to explore and analyze the traffic count data, infer the information about the traffic volume at each sensor location, and discover the sensor association. The collection and preparation of the input data required for the predictive analysis is a crucial step in the analysis. The input data for the predictive analysis was retrieved from the traffic count data received from NCDOT. Original data we obtained is separated by districts for each hurricane in the EXCEL sheets. For the predictive analysis, input data was prepared using different data exploration and data cleaning techniques in Python. After preparing the input data for the investigation, the data analysis is performed using the steps described in Figure 5.3. After analyzing the predictive analysis results, the sensor association was analyzed. Figure 5.2 depicts the overall methodology of the study.

The linear regression [39] was used for assessing the association of traffic among the sensors in the study. The dependent variable (DV) for the linear regression is hourly traffic volume at a sensor location. The independent variables (IVs) for the linear regression consist of traffic volumes at the sensor location before 1 hour, 2 hours, 3 hours, and 4 hours and at other adjacent sensor locations before 1 hour, 2 hours, 3 hours, and 4 hours. Within four hours, the coastal county population can reach the resource center located at the States capital Raleigh.

The procedure for the statistical analysis is presented in Figure 5.3. In the statistical analysis, after preparing the input data and cleaning the data by omitting the data with missing value, if there were more than five feature variables (IVs), then the feature selection step is performed. In feature selection, we used the “backward elimination” method to select the most significant features. As the name suggests, firstly we need to feed all the possible
features to the model. We check the model's performance then iteratively remove the least correlated feature one by one until the model's overall performance comes in an acceptable range. The metric used to evaluate the feature correlation is $p$-values. If the $p$-value of a feature is above 0.05, then the feature is removed; otherwise, it is kept in the model. After that, we fitted the linear regression model of the selected features to the data, and tested the normality of residuals using the Normal QQ-plot and three different normality tests. The three normality tests are the Shapiro-Wilk test, the D’Angostino and Pearson’s test, and the Anderson-Darling test. If the model passed the three tests, we stopped and accepted the model; otherwise, we did the outlier detection analysis.

![Figure 5.3: Procedure for statistical analysis.](image)

In the outlier detection analysis, we calculate Cook’s distance (CD) to identify the most influential observations in outlier detection. The observations with $CD > (4/n)$ were dropped, where $n$ is the total number of observations. After dropping the most influential observations, we fitted the model to the remaining data and tested the normality of residuals again. If the normality tests were passed, we stopped and accepted the model. Following the procedure, we fitted the regression models for the 25 selected sensor locations. Furthermore,
the regression models were used to study the traffic associations among the sensor locations, predicting the traffic volume at a sensor location during hurricane evacuation in the future.

5.3 Results and Discussion

Figure 5.4 graphically shows the traffic associations among the sensor locations based on the regression models for each location. In the graph, the node representing the sensor, and the arrow's direction represents the sensors' association. The incoming arrows to a sensor location show that the traffic at that sensor location is predicted from those incoming directed sensor locations. For example, the traffic volume at sensor A4201 is predicted by traffic volumes at sensor locations A2501, R9111, R9112, R9103, R9113, R9106, R9102, R9107, A5001D4, R5001, and previous traffic volumes at the same location A4201. Figure 4.1 reveals that the 25 locations can be grouped into the four clusters based on their traffic associations analytically and topologically.

Figure 5.4: Sensor association in North Carolina for prediction of traffic volume during a hurricane evacuation.
Table 5.1: Sensor clusters and the counties they are located in

<table>
<thead>
<tr>
<th>Sensor</th>
<th>County</th>
<th>Cluster</th>
</tr>
</thead>
<tbody>
<tr>
<td>R9101</td>
<td>Wake</td>
<td>1</td>
</tr>
<tr>
<td>R9102</td>
<td>Wake</td>
<td>1</td>
</tr>
<tr>
<td>R9103</td>
<td>Wake</td>
<td>1</td>
</tr>
<tr>
<td>R9106</td>
<td>Wake</td>
<td>1</td>
</tr>
<tr>
<td>R9107</td>
<td>Wake</td>
<td>1</td>
</tr>
<tr>
<td>R9111</td>
<td>Wake</td>
<td>1</td>
</tr>
<tr>
<td>R9112</td>
<td>Wake</td>
<td>1</td>
</tr>
<tr>
<td>R9113</td>
<td>Wake</td>
<td>1</td>
</tr>
<tr>
<td>R5001</td>
<td>Johnston</td>
<td>1</td>
</tr>
<tr>
<td>A5001D4</td>
<td>Johnston</td>
<td>1</td>
</tr>
<tr>
<td>R3103</td>
<td>Durham</td>
<td>1</td>
</tr>
<tr>
<td>A4201</td>
<td>Harnett</td>
<td>1</td>
</tr>
<tr>
<td>A6301</td>
<td>Nash</td>
<td>1</td>
</tr>
<tr>
<td>A6302</td>
<td>Nash</td>
<td>1</td>
</tr>
<tr>
<td>A9501</td>
<td>Wayne</td>
<td>2</td>
</tr>
<tr>
<td>A3003</td>
<td>Duplin</td>
<td>2</td>
</tr>
<tr>
<td>A7101</td>
<td>Pender</td>
<td>2</td>
</tr>
<tr>
<td>W7002</td>
<td>Pender</td>
<td>2</td>
</tr>
<tr>
<td>A6401</td>
<td>New Hanover</td>
<td>2</td>
</tr>
<tr>
<td>W2301</td>
<td>Columbus</td>
<td>2</td>
</tr>
<tr>
<td>A2702</td>
<td>Dare</td>
<td>3</td>
</tr>
<tr>
<td>W4701</td>
<td>Hyde</td>
<td>3</td>
</tr>
</tbody>
</table>

Table 5.1. summarizes the sensor clusters and the county they are located in. Figure 5.1 and Table 5.1 reveal that the sensor locations from the same county and adjacent counties form the cluster, and the traffic volume at a sensor location can be predicted by the traffic volumes in the same county and adjacent counties. For example, sensors R9106, R9113, R9103, R9111, R9112, R9101, R9102, and R9107 in Wake County, sensors A6302 and A6301 in Nash County, sensor A4201 in Harnett County, sensors A5001D4 and R5001 in Johnston County, and sensor R3103 in Durham County form one significant cluster, and all those sensors are located in the same county or adjacent counties. Similarly, sensors A9501,
A3003, A7101, W7002, W2301, and A6401 form the second cluster, located in the neighboring counties in southeast North Carolina. The sensors in Hyde County and Dare County form the third cluster, and the fourth cluster consists of the only sensor (A9001) located in Vance County.

The topological association of the sensors is depicted in Figure 5.5 [40], where Wake County and its neighboring counties form the first cluster. The adjacent counties in southeastern North Carolina shown in brown color on the map are in the second cluster, and the easternmost counties (Dare and Hyde) form the third cluster. This implies that to predict the traffic volume at a location, we may need the historical traffic volumes at the same location and the traffic volumes at the locations in the same cluster.

![Figure 5.5: Locations of sensor clusters on the North Carolina map](image)

Real-Time Traffic Control in an ITS System during an Emergency Evacuation
FINDINGS, CONCLUSIONS, RECOMMENDATIONS

In this CATM project, we (1) identified road segments that are critical to the resiliency of road transportation system around Charlotte using weighted betweenness centrality and 2018 average daily traffic (AADT); (2) investigated the most affected airlines at the most affected airport during Hurricane Matthew and estimated the number of affected passengers during the hurricane using network theory; (3) proposed and tested a graph theoretical approach to analyzing the U.S. airport network during a hurricane disruption, and a rerouting method to identify feasible airports to reroute from a disrupted airport; (4) investigated significant cues for Evacuation Planners’ decisions during hurricane evacuation using the Brunswik Linear Lens Model and the six machine learning algorithms; and (5) analyzed traffic flow data during Hurricane Florence evacuation in N.C. to discover space-time evacuation traffic patterns, and investigated the association of sensor locations to predict the hourly traffic during the hurricane evacuation.

For road transportation, our resiliency analysis showed that the central links with high AADT, such as road segments of U.S. Interstate 77 and U.S. Highway 74, are the most critical to the resiliency of the highway transportation network around Charlotte, and that road segments of U.S. Interstates 85 and 485 around Charlotte are also critical to the road network resiliency. For air transportation, our analysis of flight delays due to Hurricane Matthew revealed that seven airlines were most affected at the airport of MCO (which was the most affected airport), and that the topology of an airline’s flight network may influence the number of flight passengers affected by a weather disruption. Based on this finding, a graph theoretical method, including two mathematical equations for disrupter and disruptee, respectively, has been proposed to quantify airports in an airport network from the two perspectives (disrupter and disruptee). Our testing results showed that there is a high chance for an airport to be geographically impacted by a hurricane disruption and the rerouting method can identify feasible airports for rerouting flights from the disruptees. The proposed graph theoretical and rerouting methods can support airports and airlines administrators to recognize the airports that might be affected by an approaching hurricane, and identify potential airports to divert flights from the affected airports.
For hurricane evacuation, the results of our evacuation traffic analysis showed that the sensor locations from the same county and adjacent counties form the cluster, and the traffic volume at a sensor location can be predicted by the traffic volumes in the same county and adjacent counties. Our results revealed that the 25 sensor locations could be grouped into four clusters. Wake County and its neighboring counties are grouped as the largest cluster, adjacent counties in southeast N.C. as the second-largest cluster, the easternmost counties (Dare and Hyde) as a cluster, and Vance County as the last cluster. Predicting the traffic volume at a location needs to consider historical traffic volumes at the same location and the locations in the same cluster. Our results of decision making cues for hurricane evacuation showed that among the seven cues we tested, only one cue (wind speed) contributes to Evacuation Planners’ decision (i.e., whether to issue an evacuation order). The predictive models created in our studies can be used to manage hurricane evacuations, including proper traffic control, estimating emergency responders needed, and preparing enough evacuation supplies such as gas, water, food, and hotel rooms. These models and findings will help emergency response agencies deploy effective evacuation operations and improve the mobility of the people and evacuation resources during a hurricane.
REFERENCES


APPENDIX A: Data for Hurricane Evacuation in North Carolina

Table A.1: Data obtained from multiple sources for the judgment model

<table>
<thead>
<tr>
<th>Name of County</th>
<th>No of Shelters</th>
<th>Wind Speed</th>
<th>Rainfall</th>
<th>Households affected</th>
<th>Flood Level (ft)</th>
<th>Median Household Income</th>
<th>Poverty Percent</th>
<th>Evacuate</th>
</tr>
</thead>
<tbody>
<tr>
<td>Anson</td>
<td>1</td>
<td>45</td>
<td>4</td>
<td>74</td>
<td>24.39</td>
<td>33,228</td>
<td>25.1</td>
<td>0</td>
</tr>
<tr>
<td>Beaufort</td>
<td>2</td>
<td>50</td>
<td>7.5</td>
<td>850</td>
<td>12.50</td>
<td>45,860</td>
<td>19</td>
<td>1</td>
</tr>
<tr>
<td>Bertie</td>
<td>2</td>
<td>40</td>
<td>10.5</td>
<td>1025</td>
<td>16.67</td>
<td>34,127</td>
<td>24.4</td>
<td>0</td>
</tr>
<tr>
<td>Bladen</td>
<td>4</td>
<td>39</td>
<td>12</td>
<td>2817</td>
<td>36.34</td>
<td>34,422</td>
<td>26.4</td>
<td>0</td>
</tr>
<tr>
<td>Brunswick</td>
<td>3</td>
<td>67</td>
<td>7</td>
<td>784</td>
<td>18.98</td>
<td>47,000</td>
<td>13.8</td>
<td>0</td>
</tr>
<tr>
<td>Carteret</td>
<td>2</td>
<td>60</td>
<td>6.5</td>
<td>49</td>
<td>8.48</td>
<td>52,000</td>
<td>12.3</td>
<td>1</td>
</tr>
<tr>
<td>Columbus</td>
<td>5</td>
<td>59</td>
<td>11.5</td>
<td>5189</td>
<td>2.00</td>
<td>40,000</td>
<td>24.6</td>
<td>0</td>
</tr>
<tr>
<td>Cumberland</td>
<td>6</td>
<td>60</td>
<td>13</td>
<td>14803</td>
<td>58.70</td>
<td>42,107</td>
<td>18.8</td>
<td>0</td>
</tr>
<tr>
<td>Dare</td>
<td>0</td>
<td>75</td>
<td>6</td>
<td>1121</td>
<td>3.00</td>
<td>56,489</td>
<td>10.9</td>
<td>1</td>
</tr>
<tr>
<td>Duplin</td>
<td>3</td>
<td>49</td>
<td>7</td>
<td>1322</td>
<td>19.92</td>
<td>39,146</td>
<td>21.3</td>
<td>0</td>
</tr>
<tr>
<td>Durham</td>
<td>1</td>
<td>60</td>
<td>11.5</td>
<td>850</td>
<td>17.73</td>
<td>54,093</td>
<td>16.1</td>
<td>0</td>
</tr>
<tr>
<td>Edgecombe</td>
<td>4</td>
<td>60</td>
<td>11.5</td>
<td>3139</td>
<td>36.15</td>
<td>35,000</td>
<td>23.9</td>
<td>1</td>
</tr>
<tr>
<td>Franklin</td>
<td>0</td>
<td>45</td>
<td>8.41</td>
<td>13</td>
<td>23.18</td>
<td>50,000</td>
<td>15.3</td>
<td>0</td>
</tr>
<tr>
<td>Gates</td>
<td>0</td>
<td>64</td>
<td>9</td>
<td>158</td>
<td>16.19</td>
<td>49,258</td>
<td>15.2</td>
<td>0</td>
</tr>
<tr>
<td>Greene</td>
<td>1</td>
<td>50</td>
<td>9</td>
<td>579</td>
<td>24.18</td>
<td>38,010</td>
<td>23.7</td>
<td>0</td>
</tr>
<tr>
<td>Harnett</td>
<td>1</td>
<td>57</td>
<td>7</td>
<td>1683</td>
<td>19.31</td>
<td>51,682</td>
<td>16.1</td>
<td>0</td>
</tr>
<tr>
<td>Hyde</td>
<td>0</td>
<td>75</td>
<td>6</td>
<td>194</td>
<td>2.00</td>
<td>56,285</td>
<td>22.3</td>
<td>1</td>
</tr>
<tr>
<td>Hertford</td>
<td>1</td>
<td>40</td>
<td>9</td>
<td>453</td>
<td>15.40</td>
<td>37,000</td>
<td>26.1</td>
<td>0</td>
</tr>
<tr>
<td>Hoke</td>
<td>1</td>
<td>60</td>
<td>11.5</td>
<td>1786</td>
<td>12.84</td>
<td>45,820</td>
<td>19.5</td>
<td>0</td>
</tr>
<tr>
<td>Johnston</td>
<td>3</td>
<td>50</td>
<td>11.5</td>
<td>1683</td>
<td>28.90</td>
<td>57,151</td>
<td>13.2</td>
<td>0</td>
</tr>
<tr>
<td>Jones</td>
<td>1</td>
<td>50</td>
<td>6.5</td>
<td>226</td>
<td>18.46</td>
<td>34,000</td>
<td>21.5</td>
<td>0</td>
</tr>
<tr>
<td>Lee</td>
<td>1</td>
<td>45</td>
<td>11.5</td>
<td>190</td>
<td>11.44</td>
<td>50,547</td>
<td>16.9</td>
<td>0</td>
</tr>
<tr>
<td>Lenoir</td>
<td>1</td>
<td>51</td>
<td>9</td>
<td>3291</td>
<td>28.24</td>
<td>38,000</td>
<td>20.6</td>
<td>1</td>
</tr>
<tr>
<td>Martin</td>
<td>1</td>
<td>50</td>
<td>9</td>
<td>213</td>
<td>11.60</td>
<td>35,080</td>
<td>22.5</td>
<td>0</td>
</tr>
<tr>
<td>Montgomery</td>
<td>0</td>
<td>50</td>
<td>4</td>
<td>0</td>
<td>10.40</td>
<td>37,800</td>
<td>21.4</td>
<td>0</td>
</tr>
<tr>
<td>Moore</td>
<td>1</td>
<td>55</td>
<td>5</td>
<td>343</td>
<td>8.91</td>
<td>56,678</td>
<td>11.4</td>
<td>0</td>
</tr>
<tr>
<td>Nash</td>
<td>1</td>
<td>55</td>
<td>7</td>
<td>927</td>
<td>15.26</td>
<td>47,200</td>
<td>16.5</td>
<td>0</td>
</tr>
<tr>
<td>New Hanover</td>
<td>1</td>
<td>75</td>
<td>6</td>
<td>23</td>
<td>3.00</td>
<td>56,200</td>
<td>17.3</td>
<td>0</td>
</tr>
<tr>
<td>Onslow</td>
<td>6</td>
<td>50</td>
<td>6.5</td>
<td>442</td>
<td>20.55</td>
<td>38,000</td>
<td>13.7</td>
<td>0</td>
</tr>
<tr>
<td>Orange</td>
<td>1</td>
<td>45</td>
<td>6.5</td>
<td>850</td>
<td>16.25</td>
<td>61,130</td>
<td>12.8</td>
<td>0</td>
</tr>
<tr>
<td>Pamlico</td>
<td>1</td>
<td>50</td>
<td>5</td>
<td>7569</td>
<td>2.00</td>
<td>46,762</td>
<td>18.5</td>
<td>0</td>
</tr>
<tr>
<td>Pasquotank</td>
<td>1</td>
<td>64</td>
<td>8.5</td>
<td>476</td>
<td>16.19</td>
<td>45,400</td>
<td>17</td>
<td>0</td>
</tr>
<tr>
<td>Pender</td>
<td>3</td>
<td>68</td>
<td>6</td>
<td>957</td>
<td>17.79</td>
<td>50,000</td>
<td>15</td>
<td>1</td>
</tr>
<tr>
<td>Pitt</td>
<td>3</td>
<td>69</td>
<td>8</td>
<td>3303</td>
<td>24.46</td>
<td>50,000</td>
<td>21.5</td>
<td>1</td>
</tr>
<tr>
<td>Richmond</td>
<td>0</td>
<td>50</td>
<td>6</td>
<td>23</td>
<td>17.60</td>
<td>37,000</td>
<td>24.9</td>
<td>0</td>
</tr>
<tr>
<td>Robeson</td>
<td>5</td>
<td>67</td>
<td>11</td>
<td>18482</td>
<td>25.00</td>
<td>33,000</td>
<td>27.8</td>
<td>1</td>
</tr>
</tbody>
</table>
Table A.1: (cont’d)

<table>
<thead>
<tr>
<th>Name of County</th>
<th>No of Shelters</th>
<th>Wind Speed</th>
<th>Rainfall</th>
<th>Households affected</th>
<th>Flood Level (ft)</th>
<th>Median Household Income</th>
<th>Poverty Percent</th>
<th>Evacuate</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sampson</td>
<td>3</td>
<td>60</td>
<td>12</td>
<td>2236</td>
<td>27.92</td>
<td>38,835</td>
<td>19.6</td>
<td>0</td>
</tr>
<tr>
<td>Scotland</td>
<td>0</td>
<td>55</td>
<td>7</td>
<td>500</td>
<td>15.47</td>
<td>52,000</td>
<td>27.6</td>
<td>0</td>
</tr>
<tr>
<td>Vance</td>
<td>1</td>
<td>55</td>
<td>5</td>
<td>850</td>
<td>2.00</td>
<td>32,733</td>
<td>24.2</td>
<td>0</td>
</tr>
<tr>
<td>Wake</td>
<td>1</td>
<td>45</td>
<td>9</td>
<td>916</td>
<td>5.47</td>
<td>76,000</td>
<td>9.2</td>
<td>0</td>
</tr>
<tr>
<td>Wayne</td>
<td>3</td>
<td>50</td>
<td>6.5</td>
<td>6695</td>
<td>2.00</td>
<td>45,000</td>
<td>20.6</td>
<td>1</td>
</tr>
<tr>
<td>Wilson</td>
<td>1</td>
<td>52.5</td>
<td>10.5</td>
<td>721</td>
<td>2.00</td>
<td>43,456</td>
<td>22.3</td>
<td>0</td>
</tr>
</tbody>
</table>
Table A.2: Summary of evacuation traffic data in eastern North Carolina

<table>
<thead>
<tr>
<th>Sr.No</th>
<th>File name</th>
<th>District</th>
<th>County</th>
<th>Location</th>
<th>Location ID</th>
<th>Route</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>A2702</td>
<td>1</td>
<td>Dare</td>
<td>US 64 Hwy</td>
<td>280000002</td>
<td>64</td>
</tr>
<tr>
<td>2</td>
<td>A3002</td>
<td>3</td>
<td>Duplin</td>
<td>W NC 24 Hwy</td>
<td>0310000003</td>
<td>24</td>
</tr>
<tr>
<td>3</td>
<td>A3003</td>
<td>3</td>
<td>Duplin</td>
<td>E NC 24 Hwy</td>
<td>0310000002</td>
<td>24</td>
</tr>
<tr>
<td>4</td>
<td>A5001</td>
<td>4</td>
<td>Johnston</td>
<td>I-40</td>
<td>0510000001</td>
<td>40</td>
</tr>
<tr>
<td>5</td>
<td>A5301</td>
<td>2</td>
<td>Lenoir</td>
<td>Hwy 55 W</td>
<td>0540000001</td>
<td>55</td>
</tr>
<tr>
<td>6</td>
<td>A6401</td>
<td>3</td>
<td>New Hanover</td>
<td>Castle Hayne Rd</td>
<td>0650000001</td>
<td>117</td>
</tr>
<tr>
<td>7</td>
<td>A6403</td>
<td>3</td>
<td>New Hanover</td>
<td>Wrightsville Ave</td>
<td>0650000002</td>
<td>74</td>
</tr>
<tr>
<td>8</td>
<td>A7101</td>
<td>3</td>
<td>Pender</td>
<td>US Hwy 421</td>
<td>0710000002</td>
<td>421</td>
</tr>
<tr>
<td>9</td>
<td>R0901</td>
<td>3</td>
<td>Brunswick</td>
<td>Ocean Hwy W</td>
<td>0100000028</td>
<td>17</td>
</tr>
<tr>
<td>10</td>
<td>R1401</td>
<td>1</td>
<td>Camden</td>
<td>US 17</td>
<td>0150000003</td>
<td>17</td>
</tr>
<tr>
<td>11</td>
<td>R2301</td>
<td>6</td>
<td>Columbus</td>
<td>N US 701 Bypass</td>
<td>02400000676</td>
<td>701</td>
</tr>
<tr>
<td>12</td>
<td>R2701</td>
<td>1</td>
<td>Dare</td>
<td>NC 12 Hwy</td>
<td>2800000004</td>
<td>12</td>
</tr>
<tr>
<td>13</td>
<td>R5001</td>
<td>4</td>
<td>Johnston</td>
<td>I-95</td>
<td>0510000002</td>
<td>95</td>
</tr>
<tr>
<td>14</td>
<td>R5301</td>
<td>2</td>
<td>Lenoir</td>
<td>Hwy 70 W</td>
<td>0540000002</td>
<td>70</td>
</tr>
<tr>
<td>15</td>
<td>W2301</td>
<td>6</td>
<td>Columbus</td>
<td>Andrew Jackson Hwy E</td>
<td>0240000001</td>
<td>74</td>
</tr>
<tr>
<td>16</td>
<td>W4701</td>
<td>1</td>
<td>Hyde</td>
<td>US 264 Hwy</td>
<td>0480000001</td>
<td>264</td>
</tr>
<tr>
<td>17</td>
<td>W7002</td>
<td>3</td>
<td>Pender</td>
<td>US Hwy 421</td>
<td>0710000002</td>
<td>421</td>
</tr>
<tr>
<td>18</td>
<td>A2501</td>
<td>6</td>
<td>Cumberland</td>
<td>All American Exp</td>
<td>0260000001</td>
<td>1007</td>
</tr>
<tr>
<td>19</td>
<td>A4201</td>
<td>6</td>
<td>Harnett</td>
<td>E Cornelius Harnett Blvd</td>
<td>0430000001</td>
<td>421</td>
</tr>
<tr>
<td>20</td>
<td>A5001</td>
<td>4</td>
<td>Johnston</td>
<td>I-40</td>
<td>0510000001</td>
<td>40</td>
</tr>
<tr>
<td>21</td>
<td>A6301</td>
<td>4</td>
<td>Nash</td>
<td>US 264 E</td>
<td>0640000005</td>
<td>264</td>
</tr>
<tr>
<td>22</td>
<td>A6302</td>
<td>4</td>
<td>Nash</td>
<td>W NC 97</td>
<td>0640000003</td>
<td>97</td>
</tr>
<tr>
<td>23</td>
<td>A9001</td>
<td>5</td>
<td>Vance</td>
<td>S Carroll St</td>
<td>0910000001</td>
<td>1501</td>
</tr>
<tr>
<td>24</td>
<td>A9501</td>
<td>4</td>
<td>Wayne</td>
<td>US 117 Alternate</td>
<td>0960000001</td>
<td>117</td>
</tr>
<tr>
<td>25</td>
<td>R3103</td>
<td>5</td>
<td>Durham</td>
<td>I-40 E</td>
<td>0320000010</td>
<td>40</td>
</tr>
<tr>
<td>26</td>
<td>R3201</td>
<td>4</td>
<td>Edgecombe</td>
<td>US 64 East</td>
<td>0330000002</td>
<td>64</td>
</tr>
<tr>
<td>27</td>
<td>R6301</td>
<td>4</td>
<td>Nash</td>
<td>US 64 East</td>
<td>0640000006</td>
<td>64</td>
</tr>
<tr>
<td>28</td>
<td>R7701</td>
<td>6</td>
<td>Robeson</td>
<td>I-74 E</td>
<td>0780000002</td>
<td>74</td>
</tr>
<tr>
<td>29</td>
<td>R9101</td>
<td>5</td>
<td>Wake</td>
<td>US 70 Hwy E</td>
<td>0920000081</td>
<td>70</td>
</tr>
<tr>
<td>30</td>
<td>R9102</td>
<td>5</td>
<td>Wake</td>
<td>Capital Blvd</td>
<td>0920000082</td>
<td>401</td>
</tr>
<tr>
<td>31</td>
<td>R9103</td>
<td>5</td>
<td>Wake</td>
<td>I-40 EB</td>
<td>0920000013</td>
<td>40</td>
</tr>
<tr>
<td>32</td>
<td>R9104</td>
<td>5</td>
<td>Wake</td>
<td>I-40 EB</td>
<td>0920000014</td>
<td>40</td>
</tr>
<tr>
<td>33</td>
<td>R9106</td>
<td>5</td>
<td>Wake</td>
<td>I-40 EB</td>
<td>0920000020</td>
<td>440</td>
</tr>
<tr>
<td>34</td>
<td>R9107</td>
<td>5</td>
<td>Wake</td>
<td>I 540 EB</td>
<td>920000024</td>
<td>540</td>
</tr>
<tr>
<td>35</td>
<td>R9108</td>
<td>5</td>
<td>Wake</td>
<td>I 440 EB</td>
<td>0920000021</td>
<td>440</td>
</tr>
<tr>
<td>36</td>
<td>R9111</td>
<td>5</td>
<td>Wake</td>
<td>I 40 EB</td>
<td>0920000016_EB</td>
<td>40</td>
</tr>
<tr>
<td>37</td>
<td>R9112</td>
<td>5</td>
<td>Wake</td>
<td>I 40 EB</td>
<td>0920000017</td>
<td>40</td>
</tr>
<tr>
<td>38</td>
<td>R9113</td>
<td>5</td>
<td>Wake</td>
<td>I 40 EB</td>
<td>0920000018</td>
<td>40</td>
</tr>
<tr>
<td>39</td>
<td>R9114</td>
<td>5</td>
<td>Wake</td>
<td>I 540 EB</td>
<td>0920000025</td>
<td>540</td>
</tr>
</tbody>
</table>
APPENDIX B: Codes for Prediction Models of Hurricane Evacuation Traffic in Eastern North Carolina

## Libraries required for analysis

```python
import pandas as pd
import matplotlib.pyplot as plt
import statsmodels.api as sm
from scipy.stats import normaltest, shapiro, anderson
```

### Data for the analysis

```python
def DailyTraffic(TData, Date):
    df1 = TData[TData['DATE'] == Date]

    df2 = df1.drop(['DIST', 'DIR', 'DATE'], axis=1)
    df3 = df2.sort_index()
    df4 = df3[~df3.index.duplicated()]
    df5 = df4.T

    Hours = pd.Series(list(range(1, 25)), index=df5.index)
    df5['Hour'] = Hours
    return df5

# End of DailyTraffic
```

### Function to add shifted traffic data to the dataset

```python
def addShiftTraffic(DataSet, OneSiteTraffic, periods=4):
    for i in range(1, periods+1):
        ShiftedTraffic = OneSiteTraffic.shift(periods=i)
        ColumnSuffix = '_' + str(i) + 'HrBf'
        DataSet = DataSet.join(ShiftedTraffic, rsuffix=ColumnSuffix)

    return DataSet
```

```python
Folder = "C:/Users/shmhatre/Desktop/Spring_2020/CATM project_2020/TrafficEvacAnalysis/Sensor_Analysis/

Counties = pd.read_csv(Folder + "AdjacentCounty.csv", index_col="County")
Counties['Sr.No'] = Counties.index

SensorSite = pd.read_csv(Folder + "Consolidate Data.csv", index_col="SITE")

D1 = pd.read_csv(Folder + "D1.csv",index_col='SITE')
D2 = pd.read_csv(Folder + "D2.csv",index_col='SITE')
D3 = pd.read_csv(Folder + "D3.csv",index_col='SITE')
D4 = pd.read_csv(Folder + "D4.csv",index_col='SITE')
D5 = pd.read_csv(Folder + "D5.csv",index_col='SITE')
D6 = pd.read_csv(Folder + "D6.csv",index_col='SITE')
```
# Merge the traffic data from all files
# Then separate them into four datasets by directions
RawData = pd.concat([D1,D2,D3,D4,D5,D6], sort=False)
EBTraffic = RawData[RawData['DIR'] == 'EB']
WBTraffic = RawData[RawData['DIR'] == 'WB']
NBTraffic = RawData[RawData['DIR'] == 'NB']
SBTraffic = RawData[RawData['DIR'] == 'SB']

# Filter the evacuation data and re-organize the dataset

WBframes = [0] * len(Dates)
NBframes = [0] * len(Dates)

for i in range(len(Dates)):
    WBframes[i] = DailyTraffic(WBTraffic, Dates[i])
    NBframes[i] = DailyTraffic(NBTraffic, Dates[i])

WBTData = pd.concat(WBframes, keys=Dates, sort=False)
NBTData = pd.concat(NBframes, keys=Dates, sort=False)

EvacuationTraffic = WBTData.join(NBTData, how='outer', rsuffix='_NB')

# Prepare the dataset for the prediction modeling
# PredInput is the dataset for the prediction modeling
AvailableSensors = list(EvacuationTraffic.columns)
Sensor = "R9103"

PredInput = EvacuationTraffic.loc[:,[Sensor,'Hour']]# PredInput is the dataset for the prediction modeling
SiteTraffic = EvacuationTraffic[Sensor]
PredInput = addShiftTraffic(PredInput, SiteTraffic)

SensorCounty = SensorSite.loc[Sensor,'County']

for adjCounty in Counties.loc[SensorCounty]:
    adjSensors = SensorSite[SensorSite['County'] == adjCounty].index

    for adjSensor in adjSensors:
        if adjSensor == Sensor:
            continue

        print(adjSensor)
        available = AvailableSensors.count(adjSensor)
        if available == 0:
            print('Sensor ' + adjSensor + ' is missing!')
        elif available == 1:
            adjTraffic = EvacuationTraffic[adjSensor]
```python
PredInput = PredInput.join(adjTraffic)
PredInput = addShiftTraffic(PredInput, adjTraffic)
PredInput = PredInput.drop([adjSensor], axis=1)

Df1 = PredInput.iloc[3:,
Df1
Data = Df1.dropna(axis=1)
X = Data.drop(['R9103', 'Hour'], axis=1)
y = Data['R9103']

# Adding constant column of ones, mandatory for sm.OLS model
X_1 = sm.add_constant(X)

# Fitting sm.OLS model
model = sm.OLS(y, X_1).fit()

# Backward Elimination
cols = list(X.columns)
pmax = 1
while (len(cols)>0):
    p = []
    X_1 = X[cols]
    X_1 = sm.add_constant(X_1)
    model = sm.OLS(y, X_1).fit()
    p = pd.Series(model.pvalues.values[1:], index = cols)
    pmax = max(p)
    feature_with_p_max = p.idxmax()
    if(pmax>0.05):
        cols.remove(feature_with_p_max)
    else:
        break
    selected_features_BE = cols
print(selected_features_BE)

# Adding constant column of ones, mandatory for sm.OLS model
X_1 = sm.add_constant(X)

# Fitting sm.OLS model
model = sm.OLS(y, X_1).fit()

# model.pvalues
X2 = sm.add_constant(X)
model = sm.OLS(y, X2)
result = model.fit()
res = result.resid
fig = sm.qqplot(res, fit=True, line='45')
h = plt.title('Normal QQ Plot ' +Sensor)
```

*Real-Time Traffic Control in an ITS System during an Emergency Evacuation*
plt.savefig("C:/Users/shmhatre/Desktop/Spring_2020/CATM project_2020/TrafficEvacAnalysis/Informs_result/" + Sensor+'_qqplot.png',transparent = True, dpi= 1000, bbox_inches = 'tight')
plt.show()

# output residual
res
res.to_excel("C:/Users/shmhatre/Desktop/Spring_2020/CATM project_2020/TrafficEvacAnalysis/Sensor_Analysis/" + "res" + Sensor + ".xlsx")

# Shapiro-Wilk Test
data = res

# normality test
stat, p = shapiro(data)
print('Statistics=%.3f, p=%.3f % (stat, p))

# interpret
alpha = 0.05
if p > alpha:
    print('Sample looks Gaussian (fail to reject H0)')
else:
    print('Sample does not look Gaussian (reject H0)')

# D'Agostino and Pearson's Test
data = res

# normality test
stat, p = normaltest(data)
print('Statistics=%.3f, p=%.3f % (stat, p))

# interpret
alpha = 0.05
if p > alpha:
    print('Sample looks Gaussian (fail to reject H0)')
else:
    print('Sample does not look Gaussian (reject H0)')

# generate univariate observations
data = res

# normality test
result = anderson(data)
print('Statistic: %.3f % result.statistic)
p = 0
for i in range(len(result.critical_values)):
    sl, cv = result.significance_level[i], result.critical_values[i]
    if result.statistic < result.critical_values[i]:
print("%.3f: %.3f, data looks normal (fail to reject H0)" % (sl, cv))
else:
    print("%.3f: %.3f, data does not look normal (reject H0)" % (sl, cv))

# Visualisation for Cook's distance (CD) to identify the most influential observation in outlier detection

# Instantiate and fit the visualizer
from yellowbrick.regressor import CooksDistance
visualizer = CooksDistance()
visualizer.fit(X, y)
visualizer.show()

# Adding constant column of ones, mandatory for sm.OLS model
X_1 = sm.add_constant(X)
# Fitting sm.OLS model
model = sm.OLS(y, X_1).fit()
infl = model.get_influence()
sm_fr = infl.summary_frame()
sm_fr.to_excel("C:/Users/shmhatre/Desktop/Spring_2020/CATM
project_2020/TrafficEvacAnalysis/Sensor_Analysis/" + "CD" + Sensor + ".xlsx")
        (‘9/6/2018’, ‘HR 17’),

        (‘9/6/2018’, ‘HR 17’),

from yellowbrick.regressor import CooksDistance
visualizer = CooksDistance()
visualizer.fit(X, y)
visualizer.show()

#Adding constant column of ones, mandatory for sm.OLS model
X_1 = sm.add_constant(X)
#Fitting sm.OLS model
model = sm.OLS(y, X_1).fit()
#model.pvalues
sm_fr.to_excel("C:/Users/shmhatre/Desktop/Spring_2020/CATM
project_2020/TrafficEvacAnalysis/Sensor_Analysis/" + "7RS" + Sensor + ".xlsx")
X2 = sm.add_constant(X)
model = sm.OLS(y, X2)
result = model.fit()
res = result.resid
fig = sm.qqplot(res, fit=True, line='45')
# probplot = sm.ProbPlot(res)
# fig = probplot.qqplot()
h = plt.title('Normal QQ Plot 9RES ' + Sensor)
plt.savefig("C:/Users/shmhatre/Desktop/Spring_2020/CATM
project_2020/TrafficEvacAnalysis/Informs_result/" +
Sensor+'_qqplot.png', transparent=True, dpi=1000, bbox_inches='tight')
plt.show()

# Shapiro-Wilk Test
data = res
# normality test
stat, p = shapiro(data)
print('Statistics=%.3f, p=%.3f % (stat, p))
# interpret
alpha = 0.05
if p > alpha:
    print('Sample looks Gaussian (fail to reject H0)')
else:
    print('Sample does not look Gaussian (reject H0)')

# D'Agostino and Pearson's Test
from scipy.stats import normaltest
data = res
# normality test
stat, p = normaltest(data)
print('Statistics=%.3f, p=%.3f % (stat, p))
# interpret
alpha = 0.05
if p > alpha:
    print('Sample looks Gaussian (fail to reject H0)')
else:
    print('Sample does not look Gaussian (reject H0)')

# generate univariate observations
data = res
# normality test
result = anderson(data)
print('Statistic: %.3f % result.statistic)
p = 0
for i in range(len(result.critical_values)): 

sl, cv = result.significance_level[i], result.critical_values[i]
if result.statistic < result.critical_values[i]:
    print('%.3f: %.3f, data looks normal (fail to reject H0)' % (sl, cv))
else:
    print('%.3f: %.3f, data does not look normal (reject H0)' % (sl, cv))

# After passing the normality check we accept the data and do the analysis for other sensors
Real-Time Traffic Control in an ITS System during an Emergency Evacuation
Real-Time Traffic Control in an ITS System during an Emergency Evacuation
Real-Time Traffic Control in an ITS System during an Emergency Evacuation
Real-Time Traffic Control in an ITS System during an Emergency Evacuation

Traffic Analysis During Hurricane Florence Evacuation

Morgan Lipkin, Sachin Mhatre, Xiuli Qu, North Carolina A&T State University, Greensboro, NC 27411

Introduction

Due to its catastrophic damage, Hurricane Florence (2018) is said to be the worst tropical cyclone on record in the United States. The hurricane made landfall as a Category 1 hurricane in the east coast of the United States. The hurricane was accompanied by widespread flooding and strong winds that caused significant damage. The storm caused more than $1 billion in damage, with at least 100 fatalities reported. More than 1 million voluntary evacuations were reported in 39 states. The storm's impacts were felt across the southeastern United States, with a total of 3,000 people killed and 6,000 people injured. The cumulative damage from Hurricane Florence was estimated to be $24 billion, making it one of the most destructive hurricanes in U.S. history.

Traffic Data

- Time Period: September 15 to September 22
- Locations: Wake County (East and West), Johnston County (East and West), and Wayne County

Observations and Conclusions

- More westbound traffic volumes from 9/15 to 9/19, and 9/18 compared to the overall traffic volumes on the other days in Wake county (R59101).
- More traffic on 9/11 than on 9/10, 9/12, and 9/13.
- The change of traffic volumes over time during a hurricane evacuation depends on the location and the transportation network.

Acknowledgement

This research was supported by the Center for Advanced Transportation Modeling (CATM), UTC-CTP grant #46462-1471725.
ANALYSIS OF HURRICANE MATTHEW 2016 DATA TO IDENTIFY AIRLINE PASSENGERS DISRUPTION

Harshitha Meda (Ph.D. Candidate)
Advisor: Dr. Lauren B. Davis
Co-advisor: Dr. Chrysaftis Vogiatzis

Industrial and Systems Engineering

Outline
- Introduction
- Data
- Results
- Conclusions
- Future Work
Introduction

Research Objectives

- To identify the affected airports
  - To visualize the outgoing and incoming connections at these airports
- To identify the most impacted airlines
  - To understand the influence of their network structure
- To estimate the percentage of affected passengers
Objective 1

- To identify the affected airports
  - To visualize the outgoing and incoming connections at these airports
Networks with outgoing connections (before hurricane)

10/5/2016 (77)

10/6/2016 (74)

Networks with outgoing connections (during hurricane)

10/7/2016 (0)

10/8/2016 (73)

10/9/2016 (83)
Real-Time Traffic Control in an ITS System during an Emergency Evacuation
Networks with incoming connections (before hurricane)

10/5/2016 (78)

Networks with incoming connections (during hurricane)

10/7/2016 (0)

10/8/2016 (76)

10/9/2016 (84)
Objective 2

- To analyze the most impacted airlines
- To understand the influence of their network structure models
Objective 3

- To estimate the percentage of affected passengers

Methodology to estimate percent of affected passengers at the MCO

- Analysis is done utilizing the number of available seats present in the data set
- Load factors for September and October 2016 are retrieved from Bureau of Transportation Statistics

\[
\text{Load Factor} = \frac{\text{Revenue Passenger miles}}{\text{Available Seats miles}}
\]

- Number of miles are assumed to be constant for this study
Methodology cont.

\[ P_{ij} = L_{io} \times S_{ij} \]

- \( P_{ij} \) is predicted number of passengers for an airline \( i \) on day \( j \) during October month in 2016
- \( L_{io} \) is the load factor for an airline \( i \) during October month in the year 2016
- \( S_{ij} \) is number of available seats for an airline \( i \) on day \( j \)

\[ L_{avg} = \frac{(L_{is} + L_{io})}{2} \]

- \( L_{is} \) is the load factor for an airline \( i \) during September month in the year 2016
- \( L_{avg} \) is the average load factor for an airline during September and October months in the year 2016

\[ \overline{P}_i = \frac{\sum_{t=1}^{T} S_{it}}{N_T} \times L_{avg} \]

- \( \overline{P}_i \) is average number of passengers for an airline \( i \)
- \( \sum_{t=1}^{T} S_{it} \) is total number of available seats for an airline \( i \) for a total of \( T \) days i.e. 81
- \( N_T \) is the total number of days available in the data set

\[ PV_{ij} = \left( \frac{P_{ij} - \overline{P}_i}{\overline{P}_i} \right) \times 100 \]

- \( PV_{ij} \) is the percentage of the predicted number of affected passengers for an airline \( i \) on day \( j \)
Conclusions

- This study presents a data-driven approach
  - To identify the impacted airlines
  - To predict the affected number of passengers using network theory related concepts
- Southwest Airlines is the most affected carrier airline
- Orlando International Airport is the most affected airport
- This work could be relevant to apply to either a regional or international airports
- This approach could be applied to any kind of weather or human-made disruption at different airports

Storm distances on hurricane occurring days
Future Work

- Future work would be
  - To extend this analysis to other major airports present in the data
  - Based on the hurricane dates
- Other possible work would be to analyze
  - The role of different carrier airlines network topologies in hurricane affected areas.

Thank you
TRAFFIC PATTERN DISCOVERY DURING
RECENT HURRICANE EVACUATIONS
IN NORTH CAROLINA: A CASE STUDY

Sachin Mhatre
Graduate student at NC A&T
Guided by: Dr. Xiuli Qu
Associate Professor at NC A&T
7th UTC Conference, March 26-27, 2020

Outline

- Introduction
- Motivation
- Research Objectives
- Previous Studies
- Data
- Results
- Conclusions
Introduction

Overview

- Frequency and intensity of natural disasters increasing
- Effect of natural disasters to critical infrastructure
- Mass evacuation during hurricanes

Motivation

Hurricane Florence in North Carolina

- Evacuation orders in southeastern North Carolina
- Government agencies response to the incoming hurricane
- Critical preparation plan for mass movement of people
Research Objectives

- To analyze the hourly traffic flow data during hurricane evacuations.
- To discover space-time traffic patterns during hurricane evacuations.
- To investigate the relationship between traffic patterns and factors such as forecasted hurricane path and intensity, evacuation orders, etc.

Previous Studies

<table>
<thead>
<tr>
<th>Study/Methodology</th>
<th>Optimization</th>
<th>Simulation</th>
<th>Temporospaital</th>
<th>Empirical</th>
</tr>
</thead>
<tbody>
<tr>
<td>Liu et. al, 2005</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Wolshon, 2008</td>
<td></td>
<td></td>
<td></td>
<td>✓</td>
</tr>
<tr>
<td>Wolshon &amp; McArdie, 2009</td>
<td></td>
<td></td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>Archibald and McNeil, 2012</td>
<td></td>
<td></td>
<td></td>
<td>✓</td>
</tr>
<tr>
<td>Bish &amp; Sherali, 2013</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sun et al, 2017</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Data

Source of the Data

- North Carolina Department of Transportation
- Traffic Survey Group

Description of Data

- Traffic volume data in North Carolina during a hurricane evacuation
- Time period: 9/5/2018 - 9/20/2018
- Counties (Data): 42 in NC
- 1 or more sensors for each county in NC
- Each sensor: Sensor Location County, Route, lane direction
**Results continue**

- Eastern North Carolina
- Western North Carolina
- Eastern North Carolina

Total Westbound traffic volume between 12 to 4 PM
Total Westbound traffic volume between 5 to 9 PM
Total Westbound traffic volume between 9 to 1 AM

---

**Conclusions**

- Created spatial time traffic volume pattern visualization during hurricane evacuation for southeastern North Carolina.
- The traffic volume visualization indicates that the evacuation traffic volume increases during day time i.e. between 12PM to 9 PM.
- Also, the evacuation traffic volume is low during early morning hours and late night.