USING CROWDSOURCED DATA TO ASSESS AND MITIGATE IMPACTS OF RECURRENT FLOODING ON THE ROADWAY NETWORK

A Dissertation Submitted to

Department of Engineering, Systems and Environment

University of Virginia

In Partial Fulfillment of the Requirements for the

Degree of Doctor of Philosophy

By

Shraddha Praharaj

June 2021

Advisor: T. Donna Chen
Committee Chair: Jonathan L. Goodall
Committee Members: Brian L. Smith
Madhur Behl (CSE & ESE)
John S. Miller (External)

ACKNOWLEDGEMENT

Starting PhD in Fall 2016 was an overwhelming experience for me since I had to start my life from scratch in the US for the second time. The five years I spent pursuing my PhD made me more knowledgeable and helped me grow both personally and professionally. As I end this journey, I want to take a moment expressing my sincere gratitude towards the people who enabled it.

I would like to thanks my dissertation committee during this time: Dr. Brian Smith, Dr. John Miller, Dr. Jon Goodall, Dr. Madhur Behl for agreeing to be a part of my committee and providing their feedback on my work. Their knowledge in their fields have always helped me get perspective on various parts of my dissertation. I'm also grateful for the constant support and encouragement from Dr. Michael Fontaine despite not being a part of my committee.

This dissertation would have not been possible without the dMIST research group which included professors as well students from different areas of research outside of the transportation realm. The diverse set of expertise allowed immense collaborative work, and was very helpful in leading the project forward. These researchers became friends along the way, and enabled discussion of various ideas and collaborative work in a casual setting. I'm really grateful to all the team members. In addition to this, Chen research group members are invaluable friends that I have gained along the way, who have enabled many research discussions. I also gained a best friend who I now can't imagine my life without.

The biggest contributor in the academic environment that led to the degree completion is Dr. T. Donna Chen. She created a safe environment in our research group, is very accommodating, let us explore our ideas, share opinions, enable learning from every failure, and was present to give us advice on any matter, personal or professional. She has always been very supportive in the current political climate and all the hindrances that COVID had to bring. I'm grateful to have a mentor like her present in my life.

I also want to express my profound gratitude to my family. My brother, Seemit Praharaj, who has always been the Yoda to my Luke, providing guidance in every step of the way, and being my motivation in finishing this degree. I'm also indebted to my parents, who were a constant mental support in this entire journey, patiently guiding me towards completion.

Lastly, the experience of PhD life was made exponentially better due to my husband, Narendra Sadhwani. He took a decision that my PhD was more important than us living together in the same city. He stood through various ups and down in our relationship with long-distance, always understood me, let me grow as a person, and did everything in his power to keep us strong. Going from being in a long-distance relationship to long-distance marriage, and finally living together due to COVID, he has been the most accommodating person I know. I could not have asked for a better partner to take this journey with.

ABSTRACT

Recurrent flooding is becoming an increasingly common phenomenon in coastal cities all over the world. In the US, many cities on the east coast are facing the problem of recurrent flooding. This is a phenomenon caused by climate change and sea-level rise, yielding higher high tide events, rise in river water levels, and more severe rainfall events. In recent decades, recurrent flooding events are becoming more intense and frequent. These events typically last for a few hours, and while they do not necessitate evacuation plans, can cause significant impacts on traffic operations with roadway inundation. There have been various studies in the hydrology and transportation domains that cover the impacts of flooding on the transportation network. However, past research uses either static traffic volumes while altering the roadway network, or use simulated traffic flow under altered roadway network conditions. Previous research has mainly been conducted this way due to the lack of reliable, comprehensive empirical data that can provide information about the location of flood incidents, water depth on the streets, and traffic volumes at the locations of the flood incidents. There are very few cities where agencies collect any hydrology or transportation data related to recurrent flooding, and the few datasets that do exist do not provide sufficient spatial coverage or temporal resolution. Apart from agency datasets, there are also crowdsourced datasets providing information on roadway incidents such as flooding. However, the crowdsourced datasets are unregulated and can be prone to erroneous reporting. While there exists research exploring the potential of crowdsourced datasets, there are far fewer studies which establish the reliability of crowdsourced data.

This dissertation aims to address the knowledge gap between the incidence of recurrent flooding and guidance provided to local agencies to mitigate subsequent transportation impacts by establishing four research objectives aimed at closing the research gaps mentioned above (in three papers). First, the research addresses the lack of spatially comprehensive and disaggregate agency-collected traffic volume data. This dissertation proposes a machine learning model which uses unregulated crowdsourced traffic count data to build a traffic volume estimation model. Using these estimated traffic volumes, recurrent flooding impact are calculated on a citywide level using agency-provided flood incident data and on a local level using crowdsourced flood incident data. The second part of the research addresses the flaws of using an unregulated crowdsourced dataset by creating a model to estimate the trustworthiness of crowdsourced flood incident reports. With the ability to distinguish high-confidence crowdsourced reports from low-confidence reports, the filtered dataset can be applied towards the goal of real-time traffic

management under recurrent flooding conditions. The last part of this dissertation assesses dynamic vulnerability of different locations throughout the roadway network under predicted recurrent flooding, through a traffic impact index. This research establishes a relationship between the hydrological, roadway, and environmental characteristics of the locations of the high-confidence crowdsourced flood reports and the observed traffic impact index. This relationship can then be used to predict near-real time traffic impact index at locations with insufficient data, such that emergency services can prioritize the locations predicted to experience high impact for faster mitigation of recurrent flooding disruption.

The results of the research show that unlike major disasters, recurrent flooding impacts are not well-observed when generalized on a citywide scale. While citywide impacts in the Norfolk case study showed a 3% overall decrease in vehicle-hours of travel (VHT) for agency-collected flood incidents, the localized impacts analysis showed a 7% decrease in VHT and a 12% reduction in VMT in the sub-areas of flood incidents for crowdsourced flood incidents. These impacts of recurrent flooding are also not uniform and vary across different roadway types, time of day, and location. Since these crowdsourced flood incident reports are unregulated, a trustworthiness model is created using contextual data to separate trustworthy and untrustworthy reports. It was found that the model's prediction accuracy was about 91%. When applied to the crowdsourced flood incident data, about 72% of the data was deemed trustworthy. Using the trustworthy crowdsourced flood report data and traffic volumes data, a machine learning model is created to estimate the traffic impacts at locations where near-real time traffic volume information cannot be estimated. The preferred prediction model showed a normalized root-mean-square-error (n-RMSE) of 14% in the dataset. About 67% of the data was predicted within one standard deviation of the observed values. These models are limited in interpreting the results due to a small sample size. Nevertheless, this dissertation demonstrates a framework for using crowdsourced data that has a potential to quickly identify and predict likely high impact flood event locations so that appropriate emergency response measures can be taken to reduce the traffic impacts due to flooding.

This research shows the potential of crowdsourced data as more than just information and as a significant reinforcement to the spatially and temporally limited agency datasets. With the use of various crowdsourced datasets, spatially and temporally disaggregate analyses on the roadway network are now possible, with potential applications in the decision-making framework for the cities experiencing deteriorating impacts of recurrent flooding. The various trustworthiness, volume estimation, and traffic impact prediction models created through this dissertation have a broad spectrum of applications within and outside of transportation networks. The models developed in this dissertation will help to close the

knowledge gap in existing practice and will also ensure the best use of available resources in applications that examine the use of crowdsourced data in disruption mitigation efforts.

TABLE OF CONTENTS

CHAPTER 1: INTRODUCTION

1.1 **BACKGROUND**

Coastal cities are increasingly facing the impacts of climate change and sea-level rise, with increasing instances of flooding in recent years (Jacobs et al., 2018). Recurrent flooding is a low severity and low intensity event that is caused due to high tides, storm surges, heavy rainfall, or any combination of these factors (Sadler, 2017). Increased frequency and intensity of these recurrent events in the recent decades have increased instances of roadway flooding (Karager, 2017). The US Atlantic coast has become particularly vulnerable due to its low elevation, large population concentration, and economic importance (Wdowinski, 2016). The recurrent roadway flooding incidents disrupt road users' daily activities, leaving the transportation network vulnerable and its performance hampered.

Historically, the planning, design, and management of transportation and stormwater infrastructure systems have been independent of each other. However, with climate change and growing prevalence of urban coastal flooding, these systems are becoming increasingly interdependent. Currently, there is a lack of decision-making frameworks for long-term planning and short-term operational strategies to mitigate urban coastal flooding that simultaneously considers both interdependent systems. This is largely due to the absence of sufficient records of flood locations, inundation levels, and traffic movement. In recent years, the use of crowdsourced data has emerged as a low-cost, widely available source for data collection in various fields. Insufficient or non-existent agency-provided data can now be enhanced with this readily available abundant data. The use of reliable crowdsourced data can enable near real-time decision-making for coastal cities to minimize impacts due to recurrent roadway flooding. This dissertation aims to utilize different agency and crowdsourced datasets to develop a vulnerability indicator that identifies locations highly susceptible to large traffic impacts as a result of roadway flooding. A data-predictive model is then developed which can identify other locations in the roadway network that have similar vulnerabilities but not sufficient data. To achieve this goal, the dissertation research is divided into three stages. The motivation, research gap, and research objectives in each phase of the study are explained in the next paragraphs.

The first study enables the estimation and understanding of the impacts of recurrent flooding on traffic flow. The motivation of this study stems from the gap in existing literature regarding the analysis of transportation impacts due to small scale-disruptions like recurrent flooding, the use of empirical data

to quantify these impacts and providing near real-time guidelines for cities to mitigate small-scale disruption related impacts. Most existing disruption-related studies have focused on large-scale disruptions for evacuation or emergency routing. Empirical studies for large-scale or small-scale disruptions are very limited in this domain, relying heavily on public agency datasets, which are only available in specific regions. New York City subway and taxi datasets were some of the very few public agency datasets that the researchers have used for analyzing system disruptions due to hurricanes (Donovan and Work, 2017, Zhu et al., 2016, Mudigonda et al. 2019). However, these cannot be translated to other regions due to the data-location specificity. The first part of the dissertation aims to establish the extent of recurrent flooding impacts, and subsequently quantify the impacts on traffic flow. The methodology addresses how to combine crowdsourced and agency-based data sources to estimate these impacts.

The second study is motivated by the lack of spatially comprehensive and temporally disaggregate agency datasets to identify flooding locations. In the past decade, crowdsourced datasets have become increasingly popular in transportation applications where traditional data collection methods are cost-prohibitive. However, crowdsourced data is not regulated, and there could be inaccurate reporting due to misunderstanding, confusion, carelessness, incompetence, or even reporting with an intent to deceive (Amin-Naseri et al., 2018). For efficient use of crowdsourced incident reporting data source in vulnerability analysis, a filtered dataset of high confidence reports is required. In this portion of the study, we build a trustworthiness model for crowdsourced incident reporting data, which can identify suspect reports and increase the reliability of the crowdsourced data.

The third study in the dissertation defines a vulnerability indicator that represents the traffic impacts at trustworthy flood event locations. Several studies have used network properties to define the vulnerability of the roadway network caused by various disruptions. However, there are very few studies that use near real-time data to define vulnerability in a truly dynamic manner, which can estimate the degree of traffic impacts at various locations, given the contextual information and traffic conditions. In the last part of the dissertation, we build upon the models developed in the previous two sections (the data-predictive traffic volumes estimation model and flood incident reporting trustworthiness model) to define a traffic impact index for locations with known flood-related disruptions. This part of the study provides a framework that can enable emergency management services to identify locations with similar characteristics in the network, which may not all have near- real-time traffic and flood incident data. This framework has a potential to provide decision-making support to the city agencies to prioritize areas likely

to experience high traffic impacts due to flooding, and enable faster mitigation responses that can reduce traffic disruptions due to flooding.

The research gap, motivation, and development of research objectives described in the previous paragraphs have emerged from the limitations of data in various studies, which can be filled by the use of crowdsourced data. We use crowdsourced data to fill the knowledge gap in the impact analysis of flooding and work towards providing real-time decision-making support for reducing the disruption impacts. Based on these, the research objectives that this dissertation seeks to address are listed below:

1. Develop a data-predictive model that utilizes crowdsourced traffic count data to estimate traffic volumes at any location on the roadway network in a computationally efficient manner.
2. Develop a framework to quantify the spatial and temporal impacts of recurrent flooding on personal vehicle travel.
3. Build a trustworthiness model using contextual factors for crowdsourced flood-incident reporting data which identifies suspect reports and increases the reliability of the crowdsourced data.
4. Create a data-predictive dynamic vulnerability model framework, which, with superior data, can be used to identify locations vulnerable to recurrent flooding in near-real-time, even in the absence of complete incident reports or traffic data.

These objectives can be applicable for other cities also experiencing the deteriorating impacts of recurrent flooding. The framework can be instrumental in using crowdsourced data to obtain near-real-time information for directing road users from more vulnerable areas to reduce travel delays and trip abandonment.

1.2 **ORGANIZATION OF REPORT**

The remainder of this dissertation is organized as follows:

- Chapter 2: Estimating Impacts of Recurring Flooding on Roadway Networks: A Norfolk, Virginia Case Study.
    - This paper uses crowdsourced traffic volume data to assess the impacts of recurrent flooding on traffic flow.
    - The impacts of recurrent flooding are estimated using a citywide and localized area of analysis, using flood incident data from agency and crowdsourced datasets.

    - This paper addresses research objective 1 and 2.
- Chapter 3: Assessing Trustworthiness of Crowdsourced Flood Incident Reports using Waze Data: A Norfolk, Virginia Case Study. Transportation Research Record (Under Review). Presented at Transportation Research Board 100th Annual Meeting, January 2020.
    - This paper uses contextual variables to build a model to assert the trustworthiness of Waze-reported flood incidents.
    - This paper addresses research objective 3.
- Chapter 4: Predicting traffic impacts due to recurrent flooding using crowdsourced data: A Norfolk, Virginia Case Study (To be submitted to the Transportation Research Board 101st Annual Meeting)
    - This paper defines a traffic impact index which is a measure of the impact of traffic flow due to recurrent flooding.
    - This index is then used to develop a data-predictive framework, which with superior data can identify the impact experienced by other locations with no traffic flow data.
    - This paper addresses research objective 4.
- Conclusion
    - Summarizes the findings and limitations of each of the three papers listed.
    - Discusses the overall contribution of this dissertation.
    - Discusses potential applications of this dissertation.
    - Discusses ideas on future work.

1.3 REFERENCES

1. US EPA. *Climate Change Indicators in the United States: Coastal Flooding*, United States Environmental Protection Agency, 2016. www.epa.gov/climate-indicators. Accessed May 10, 2018.
2. Jacobs, J. M., L. R. Cattaneo, W. Sweet, and T. Mansfield. Recent and Future Outlooks for Nuisance Flooding Impacts on Roadways on the US East Coast. *Transportation Research Record: Journal of Transportation Research Board*, 2018. 2672(2): 1-10. DOI 10.1177/0361198118756366
3. Holling C. S. (1973). Resilience and stability of ecological systems. *Ann. Rev. of Ecol. Systems, 4*(1), 1-23.
4. Berdica, K. (2002). An introduction to road vulnerability: what has been done, is done and should be done. *Transport Policy, 9*, 117–127.

5. Johansson J., & Hassel H. (2010). An approach for modeling interdependent infrastructures in the context of vulnerability analysis. *Reliability Engineering and System Safety, 95*(12), 1335–1344.
6. Duan Y., & Lu F. (2013). Structural Robustness of City Road Networks based on Community. *Computers, Environment, and Urban Systems, 41*, 75-87.
7. Grenzeback, L. R., Lukman, A. T., & Systematics, C. (2008). *Case study of the transportation sector's response to and recovery from hurricane's Katrina and Rita*. Transportation Research Board.
8. Hoseinzadeh, N., Y. Liu, L. D. Han, C. Brakewood, and A. Mohammadnazar. Quality of location-based crowdsourced speed data on surface streets: A case study of Waze and Bluetooth speed data in Sevierville, TN. *Computers, Environment and Urban Systems*, 2020. Volume 83.
9. Amin-Naseri, M., P. Chakraborty, A. Sharma, S. B. Gilbert, and M. Hong. Evaluating the Reliability, Coverage, and Added Value of Crowdsourced Traffic Incident Reports from Waze. *Transportation Research Record*, 2018. Volume: 2672 (43), pp. 34–43.
10. Goodall, N., and E. Lee. Comparison of Waze crash and disabled vehicle records with video ground truth. *Transportation Research Interdisciplinary Perspectives*, 2019. Volume: 1.
11. Ouyang, R. W., M. Srivastava, A. Toniolo, and T. J. Norman. Truth Discovery in Crowdsourced Detection of Spatial Events. *IEEE Transactions on Knowledge and Data Engineering*, 2016. Volume: 28 (4), pp. 1047–1060.
12. Dong, X. L., L. Berti-Equille, and D. Srivastava. Truth Discovery and Copying Detection in a Dynamic World. *Proc. VLDB Endow*, 2009. Volume: 2 (1), pp. 562–573.
13. Prandi, C., S. Ferretti, S. Mirri, and P. Salomoni. Trustworthiness in crowd- sensed and sourced georeferenced data. *IEEE International Conference on Pervasive Computing and Communication Workshops (PerCom Workshops)*, 2015. pp. 402-407, doi: 10.1109/PERCOMW.2015.7134071.
14. Mattsson L. G., & Jenelius E. (2015). Vulnerability and Resilience of Transportation Systems- A discussion of recent research. *Transportation Research Part A, 81*, 16-34.
15. Derrible S., & Kennedy C. (2010). The complexity and robustness of metro networks. *Physica A, 389*, 3678–3691.
16. Leu L., Abbass H., & Curtis N. (2010). Resilience of ground transportation networks: A case study on Melbourne. Presented at 33rd Australasian Transport Research Forum Conference, Canberra, Australia.
17. Dorbritz, R. (2011, April). *Assessing the Resilience of Transportation Systems in case of Large Scale Disastrous Events.* 8th International Environmental Engineering Conference, Vilnius, Lithuania, 1070-1076.

18. Osei- Asamoah A., & Lownes N. E. (2014). A complex network method of evaluating resilience in surface transportation networks. *Transportation Research Record: Journal of the Transportation Research Board*, *2467*, 120- 128.
19. Testa A. C., Furtado M. N., & Alipour A. (2015). Resilience of Coastal Transportation Networks in the face of Climatic Extreme Events. *Transportation Research Record: Journal of the Transportation Research Board*, *2532*, 29- 36.
20. Donovan B., & Work D. B. (2015). Using coarse GPS data to quantify city-scale transportation system resilience to extreme events. arXiv preprint arXiv:1507.06011.
21. Zhu, Y., Xie, K., Ozbay, K., Zuo, F., & Yang, H. (2017). Data-driven spatial modeling for quantifying networkwide resilience in the aftermath of hurricanes Irene and Sandy. Transportation research record, 2604(1), 9-18.
22. Mudigonda, S., Ozbay, K., & Bekir Bartin (2019) Evaluating the resilience and recovery of public transit system using big data: Case study from New Jersey, *Journal of Transportation Safety & Security*, 11:5, 491-519, DOI: 10.1080/19439962.2018.1436105

CHAPTER 2

ESTIMATING IMPACTS OF RECURRING FLOODING ON ROADWAY NETWORKS: A NORFOLK, VIRGINIA CASE STUDY

Shraddha Praharaj, T. Donna Chen, Faria T. Zahura, Madhur Behl, Jonathan L. Goodall

Natural Hazards - January 2021

Presented *at Bridging Transportation Researchers*, August 2020

Presented *at Transportation Research Board 99th Annual Meeting*, January 2020

Presented at 2nd *International Conference on Transportation System Resilience to Natural Hazards and Extreme Weather*, November 2019

Presented at 3rd World Conference for Transport Research, Mumbai, India, May 2019

Presented at Engineering Sustainability 2019, April 2019.

Abstract

Climate change and sea level rise have increased the frequency and severity of flooding events in coastal communities. This study quantifies transportation impacts of recurring flooding using crowdsourced traffic and flood incident data. Agency-provided continuous count station traffic volume data at 12 locations is supplemented by crowd-sourced traffic data from location-based apps in Norfolk, Virginia to assess the impacts of recurrent flooding on traffic flow. A random forest data predictive model utilizing roadway features, traffic flow characteristics, and hydrological data as inputs scales the spatial extent of traffic volume data from 12 to 7,736 roadway segments. Modeling results suggest that between January 2017 and August 2018, City of Norfolk reported flood events reduced 24-hour citywide vehicle-hours of travel (VHT) by 3%, on average. To examine the temporal and spatial variation of impacts, crowdsourced flood incident reports collected by navigation app Waze between August 2017 and August 2018 were also analyzed. Modeling results at the local scale show that on weekday afternoon and evening periods, flood-impacted areas experience a statistically significant 7% reduction in VHT and 12% reduction in vehicle-miles traveled, on average. These impacts vary across roadway types, with substantial decline in traffic volumes on freeways while principal arterials experience increased traffic volumes during flood periods. Results suggest that analyzing recurring flooding at the local scale is more prudent as the impact is temporally and spatially heterogeneous. Furthermore, countermeasures to mitigate impacts require a dynamic strategy that can adapt to conditions across various time periods and at specific locations.

Key Words: recurring flooding, crowdsourced data, data-predictive model, impact analysis

1. Introduction

Recurring flooding is a type of disruption commonly observed in coastal cities due to heavy rainfall, high tides, or both. Historically, recurring flooding has been a low-frequency, low spatial- and temporal-scale disruption to the transportation system, assumed to have minor impacts. However, in recent years, rising sea levels and coastal flooding are increasingly affecting coastal communities across the US, with almost 30 coastal cities witnessing more than double the number of annual flood days in the 2010s as compared to the 1950s (US EPA, 2016, p. 36). NOAA (Sweet et al., 2018, p. 17) projects tide-related flooding in east coast cities in the US to increase three-fold by 2030 and ten-fold by 2050, relative to 2019 estimates. Tidal flooding, combined with rainfall-induced flooding, is expected to increase the number of flood events in these cities even further. With continued relative sea level rise, recurring flooding is expected to occur more frequently, and propagate to more inland locations.

As recurrent flooding frequency and intensity increases, there is a growing need to understand the subsequent impacts of these flood events on people and civil infrastructure (traveler response, frequency and duration of roadway closures, reduction of infrastructure life, impact on stormwater drainage capacity, etc.). Coastal recurring flooding is considered a minor disruption compared to consequences of catastrophic storms. However, inundated areas in coastal cities greatly deteriorate the mobility of road users, by increasing travel delay and by disrupting their ability to complete trips. The existing literature on transportation disruptions due to flooding are largely focused on major storms, with much of the research oriented towards evacuation and rehabilitation efforts, and not the recovery of daily transportation activities.

Only a few studies have examined the transportation impacts of recurring flooding through projected data and scenario analysis (e.g., Suarez et al, 2005 and Chang et al, 2010). Due to lack of availability of real time crowdsourced datasets, none of the previous studies have used empirical data to observe the impacts of recurrent flooding on the roadway network. This study is the first to utilize empirical data to examine the impact of recurring flooding on roadway users, using a combination of agency-provided and crowdsourced data sets in Norfolk, Virginia. The analysis in this study is two-fold: first analyzing the daily (24-hour) impacts on a citywide scale using agency-provided flood report data, then analyzing the time-of-day (TOD) impacts on a localized scale using crowdsourced flood report data.

2. Literature Review

There are few studies focused on recurrent flooding in the transportation infrastructure resilience literature due to its historical categorization as low severity and low frequency. Among studies examining recurring flooding events, most use projected transportation and hydrological data to create disruption scenarios for predicting roadway impacts. Lu and Peng (2011) developed an accessibility-based analysis to quantify roadway network vulnerability to sea level rise (SLR). They considered land use and population variables in defining an accessibility index in the Miami, Florida network. Their model assessed portions of the roadway network and traffic analysis zones that would be inundated at different SLR scenarios. Jacobs et al. (2018) combined flood projection maps with annual average daily traffic data (AADT) from the US Federal Highway Administration's Highway Performance Monitoring System along east coast highways. They estimated the current total vehicle hours of delay due to recurring flooding at over 100 million hours annually, and projected this delay will increase to 160 million vehicle-hours by 2020 and 1.2 billion vehicle-hours by 2060. On a citywide scale, Suarez et al. (2005) estimated the indirect costs of increased flooding in Boston by examining the effects of coastal flooding due to SLR and riverine flooding due to heavy rainfall events. The study simulated these effects in an urban transportation model and projected an increase in delay and lost trips of around 80% in 2100 compared to 2000, with an assumed SLR of 0.3 cm per year and an increase in intensity of heavy rainfall events of 0.31% per year. Sadler, et al. (2018) estimated the impact of SLR on flooding of roadways, by running different high tide scenarios for the cities of Norfolk and Virginia Beach. Critical roadways vulnerable to flooding were identified based on the annual average weekday daily traffic, elevation of roadways, and different high tide and storm surge scenarios. The study yielded an annual generalized estimate that nearly 10% of major roadways would be affected for every high tide event by the year 2100. As a part of a larger study in Portland, Oregon, Chang et al. (2010) used predicted flooding frequency and locations based on hydrological models to determine the impacts of coastal flooding on the roadway network in 2035, using the four-step regional travel demand model. The study found an inconsistent relationship between precipitation and travel disruption impacts and estimated a negligible change in vehicle miles traveled (VMT). However, vehicle-hours of delay increased by up to 10% in one of the sub-areas analyzed.

None of the studies discussed thus far use empirical data for analysis. Only a few studies have characterized the impact of flood events on transportation systems using empirical data, and these studies focus on large-scale disruptions. For example, New York City taxi and subway ridership datasets were made publicly available for 2010 through 2013, during which hurricanes Irene and Sandy significantly disrupted the transportation and power networks in the area. Zhu et al. (2016) and Donovan and Work

(2016) used these datasets to propose new methodologies to quantify city-scale transportation system resilience to extreme events. Zhu et al. presented resilience curves, which showed that Hurricane Sandy had a slower transportation recovery rate than Hurricane Irene. Resilience of the roadway network was found to be higher in both disruptions compared to the subway network. In the post disruption period of Hurricane Sandy, Donovan and Work found an increase in delay of over two minutes per mile about two days after the hurricane had struck, although a faster traffic flow was observed during most of the post-disruption period.

A significant challenge to using real-time data for estimating the impacts of disruption incidents is simply the lack of availability of such data through traditional agency sources. Installing sensors on the roadway network to obtain comprehensive real-time information is cost-prohibitive, which makes passively generated crowdsourced data an attractive source for transportation analysis. Crowdsourced data is not regulated, and may contain erroneous reporting due to misunderstanding, confusion, carelessness, incompetence, or even intent to deceive (Ouyang et al, 2016). This data, however, may still contain useful information to improve the understanding of any situation. Various studies (e.g. Amin-Naseri et al., 2018, Lenkei, 2018, Goodall and Lee, 2019) have conducted analyses to quantify traffic incidents through the crowdsourced navigation app, Waze (owned by Google). This study is the first to apply crowdsourced real-time data to assess the impacts of recurring flooding, using citizen-reported flood incident data from Waze and crowdsourced traffic data from location-based service (LBS) app data aggregator Streetlight (founded, 2011). The goal of this study consists of two parts to contrast the extent of analysis possible with traditional vs. crowdsourced data: analyzing citywide impacts of recurring flooding using agency-obtained flood incident data, and localized impacts using crowdsourced flooding incident data.

3. Data Sources

For this study, a combination of agency-provided traffic volume data in limited locations and crowdsourced LBS data is used to build a predictive model which estimates the traffic volumes across the entire Norfolk roadway network. Transportation datasets include agency-provided roadway geometry data along with agency-provided and crowdsourced traffic volumes. Hydrology datasets include agency-provided tide and rainfall data, along with agency-provided and crowdsourced flood incident data. The following subsections discuss each dataset in detail, and Table 1 shows basic summary statistics for the various data sets.

3.1. Roadway characteristics data

The roadway characteristics considered in the data predictive model consist of geometric features obtained from Hampton Roads Regional Travel Demand Model (HRRTDM), provided by Virginia Department of Transportation (VDOT), and include number of lanes, posted speed limit, and per lane capacity for each of 7737 unique links in the city of Norfolk. Thus, the roadway network analyzed in this study was limited to the links in HRRTDM (shown in blue in Figure 1), which includes interstates, freeways, arterials and collectors in Norfolk. Minor streets (shown in grey in Figure 1) are represented by aggregate centroid connectors (blue links in Figure 1 that end at a cluster of grey links), but are not individually analyzed.

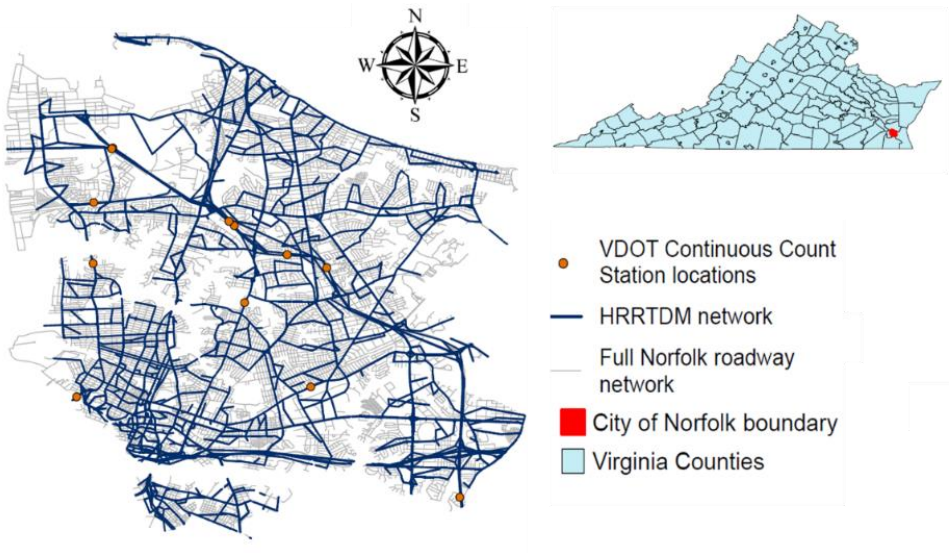


Figure 1 City of Norfolk, Virginia roadway network coverage with data sources

3.3. Flood incident data

3.3.1. Agency-provided flood incident data

Flood incident data from the City of Norfolk was collected from city employees' reported flood locations in a mobile phone application (System to Track, Organize, Record, and Map [STORM]). Due to the lack of a timestamp associated with the flood reports (only dates were included), flood report data is coded as a binary variable, with any day with one or more flood reports considered a flood day (FD) and any day without flood reports considered as a non-flood day (NFD). The spatial distribution of flood locations could not be considered while using the city's flood incident data set due to a lack of citywide spatial representation of the small sample of reports for each FD. The incident data collected spanned from January 2017 to August 2018 with floods reported on 10 unique days. This data is used to analyze the citywide flood impacts.

3.3.2. Crowdsourced flood incident data

The mobile navigation application Waze also collects flood incident reports (alongside other incident data like road closures and congestion) via their real-time information reporting tool. The application provides aggregated user-reported incident data via its Waze for Cities data sharing program, open to public entities worldwide. In this study, Waze timestamped and location-specific incident reports related to flooding in Norfolk (106 unique days with flooding between August 2017 and August 2018) are analyzed. While the Waze flood report data is not comprehensive of all instances of roadway flooding in Norfolk, its spatial coverage is significantly greater than the agency data available through City of Norfolk.

3.4. Hydrological data

The hydrological characteristics considered in this study include rain and tidal gauge data. The rainfall data, collected at 15-minute intervals, is from the Hampton Roads Sanitation Department (HRSD), which has seven rain gauge stations in the city. Tide level data is available through the sole tidal gauge in the city at Sewell's Point, and data collected every six minutes is archived and obtained by NOAA Tides and Currents. These two datasets were aggregated to match the time periods specified in the traffic volumes data description.

Table 1 Summary statistics for datasets utilized

Dataset (data source)	**# locations**	**Frequency of data collection**	**Minimum Value**	**Maximum Value**
Traffic volumes (VDOT CCS)	12	15 mins	0 vehicles	2080 vehicles
Traffic volumes (trip counts from Streetlight Data)	7737	Time period (3 hours or 6 hours)	0 pings	863 pings (in period 3: 9a – 3p)
Flood incident data (City of Norfolk)	Theoretically all	Daily	1 location in a single day	40 locations in a single day
Flood incident data (WAZE)	Theoretically all	1 minute	1 location in a single day	107 locations in a single day
Rainfall data	7	1 hour	0 inches	2.26 inches

(HRSD)				
Tide level data (NOAA)	1	6 minutes	-10.59 m	13.61 m

4. METHODS

The overall framework to estimate citywide impacts of recurring flooding is shown in Figure 2(a). For local impact analysis (Figure 2[b]), two components of the framework are changed. First, a localized boundary around flood incident reports is selected based on the roadway network structure and AADT of adjacent roadway links. Second, Waze flood incident report data is used in place of the City of Norfolk data.

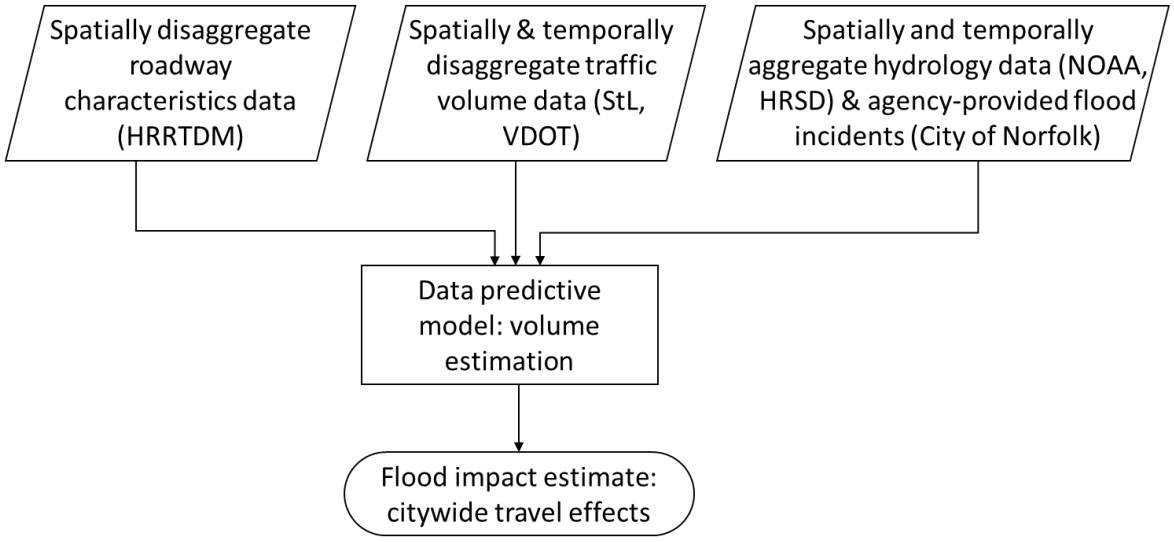


Figure 2(a) Citywide flood impact analysis framework

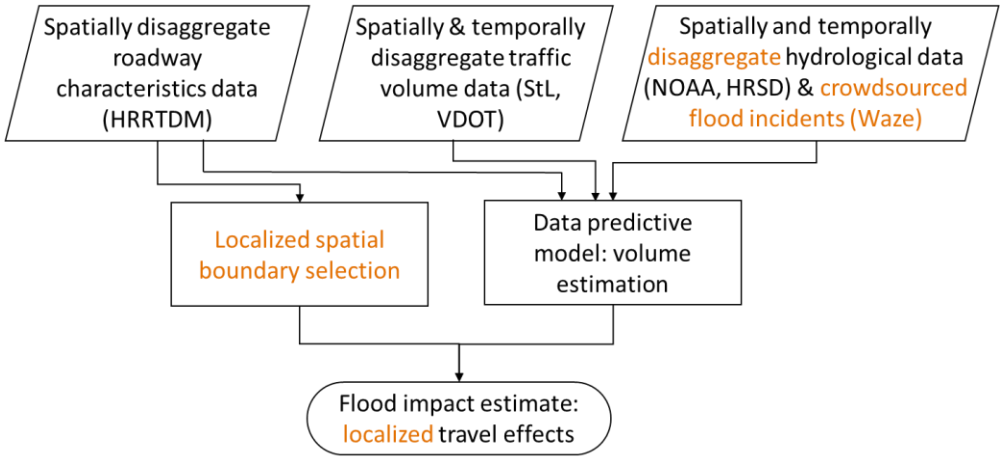


Figure 2(b) Localized flood impact analysis framework

4.1. Data Predictive Model: Volume Estimation

For the volume estimation step of the flood impact analysis, different roadway, traffic flow, and hydrological variables were used to create a data predictive model, which uses a set of input variables to provide traffic volume estimates on each roadway link for each time period. For the citywide flood impact analysis, the number of lanes, speed limit, and capacity per lane data were collected from the HRRTDM; trip counts and speeds were collected from Streetlight Data; tide levels from NOAA; rainfall values from HRSD (averaged over the city); and flood incident reports from City of Norfolk. In the localized impacts study, the same datasets are used, except the City of Norfolk flood incident report data is replaced by Waze flood incident reports, and the rainfall values are interpolated at the point of the flood incident report (from 7 rain gauges across the city, using the inverse distance weighting [IDW] technique in ArcGIS). As seen in Figure 3, several predictive models were tested in this study to predict the link volumes in order to determine a preferred model with the best prediction accuracy without overfitting the data.

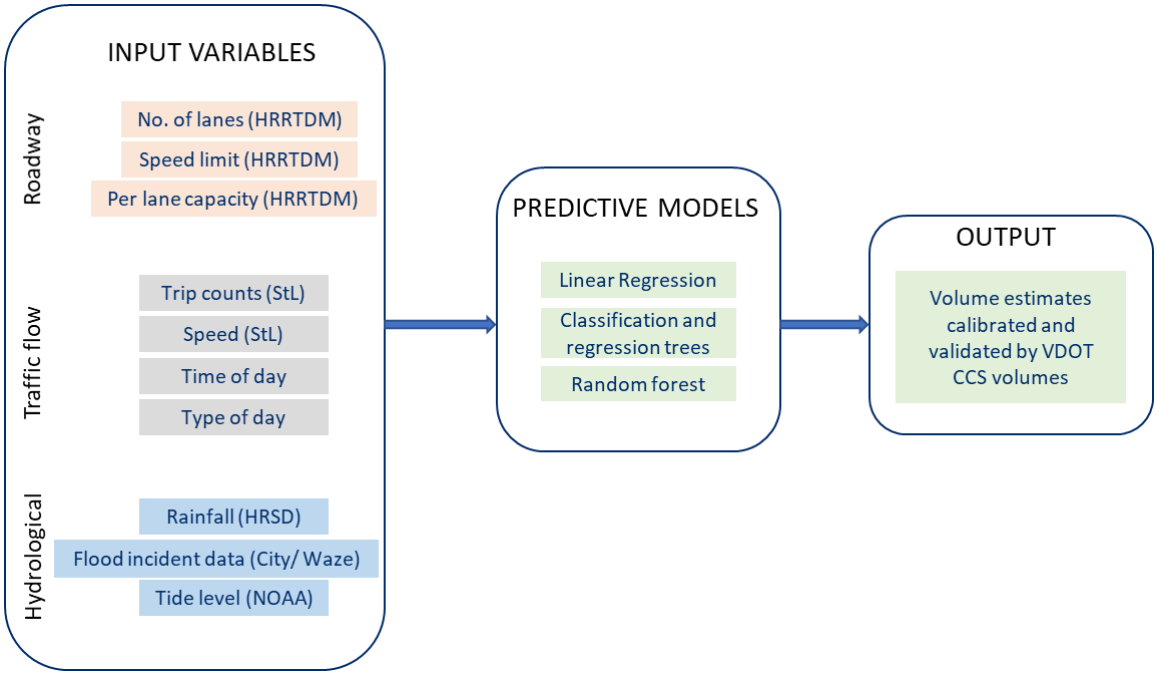


Figure 3 Traffic volume prediction process

A linear regression model was first developed as a baseline model for comparison. Classification and regression trees (CRT) and Random forest (RF) models, which group data points with similar dependent variable values together based on their independent variables, were also developed.

In CRT models, a parent node in the CRT is divided based on any independent variable into two child nodes, such that each child node is more homogenous (or less impure) than the parent node. Homogeneity is measured by the least squared deviation measure of impurity (within-node variance). The process continues until constraints, such as a minimum number of cases per node, maximum tree depth, node homogeneity, or a minimum change in improvement, are satisfied. In this study, 70% of the observations were reserved for training the dataset, and 30% were reserved for validation. Through trial and error, a 50-20 split of data in parent and child nodes was used (a minimum of 50 observations from the dataset in the parent node, and a minimum of 20 observations in the child node), which was pruned to avoid overfitting. Pruning reduces the size of decision trees in an attempt to prevent the nodes from being too specific (thereby keeping the model more generalized).

In random forests, similar to the CRT models, a 70-30 split of observations are used for training and testing the dataset, respectively. Random sampling of data subsets is performed on the training dataset to fit the samples into a model prediction, while reducing the total error in the model. The response variables are divided into groups until the resulting predictions reach a minimum amount of node impurity

(sum of the squared deviations between the predicted and actual value, a measure of error). Random forests are a strong modeling technique and much more robust than a single decision tree. They aggregate many decision trees to limit overfitting as well as error due to bias and therefore yield useful results. CRT models are also prone to overfitting the data, with random forest addressing the issue by creating various groups of randomly-selected regression trees while running the model.

Once the model is developed, errors are calculated for training and testing the data, which are used as criteria for selecting the appropriate model. Errors calculated for these models are the root mean squared error (RMSE) and normalized root mean squared error (NRMSE), given by Equations 1 and 2.

$$\boldsymbol{RMSE} = \sqrt{\frac{\sum_{i=1}^{I}(v_{obs,i} - v_{model,i})^2}{n}} \tag{1}$$

$$\boldsymbol{NRMSE} = \frac{\boldsymbol{RMSE}}{v_{obs,max} - v_{obs,min}} \tag{2}$$

where i= roadway link

$v_{obs,i}$=observed VDOT CCS volumes

$v_{model,i}$=predictive model's estimated volumes

4.2. Flood Impact Estimation (Citywide and Localized Analysis)

The citywide impacts of roadway flooding borne by travelers on the Norfolk roadway network are accounted for by comparing the 24-hour VHT across the entire city on a day with a recorded flood incident versus days without a flood incident. To assess the citywide VHT on a FD, the products of the estimated link volume and the average travel time for each link are aggregated across all TOD periods (Equation 3). FD traffic volumes and VHT were compared with an average NFD to estimate the network-wide impacts of recurring flooding. For each FD, four NFDs were selected to obtain an average NFD (and its associated link volume and travel speeds) (Equation 4). NFDs are selected from comparable days (workdays measured against other workdays, non-work days—weekends and holidays—measured against other non-work days) within three weeks prior to and after the FD to minimize effects of seasonal and weekday versus weekend traffic variation. However, the three days immediately before and after the FD were excluded to minimize potential anticipatory and residual traffic effects of the flood incident. Conceptually, this is similar to the approach taken by Zhu et al (2016, p. 2599), where the data was compared to the

same day in the prior year to observe differences in traffic flow while accounting for seasonal traffic variation.

In the localized analysis, Waze flood incident reports are assigned to a TOD flood period (FP) based on the timestamp of the report, and the comparable non-flood periods (NFPs) are defined as the same TOD periods during the three weeks (of the same day type, e.g., workday or non-work day) prior to and after the FP, when no flood incident was reported within a one mile radius of the location of the flood incident report. The incidents reported in Waze have an associated time duration (time between when a flood incident is first reported to 30 minutes after the last "thumbs up", indicating that report is still true; or until someone reports a "thumbs down", indicator that the report no longer holds true). The maximum duration of an incident reported in Waze is under 3 hours, which is shorter than any of the TOD periods considered for classification. Thus, the analysis here only accounts for impacts within the time period that contains the flood report (and not subsequent time periods). For each FP, all candidate NFPs within three weeks prior to and after the FP are considered to obtain an average NFP for link volume and travel speed comparisons.

Previous studies have shown that precipitation affects travel decisions and choices differently for peak and off-peak periods, weekday and weekend traffic, and in different seasons (Böcker, Dijst and Prillwitz, 2013, pp. 79-80). Thus, this study distinguishes the FDs (or FPs) with rainfall from those without, when considering candidate NFDs (or NFPs). Flood events in the study area are assumed to occur due to two environmental conditions: high tide, rainfall, or both. The candidate NFDs (or NFPs) for a high-tide only FD (or FP) were picked from days (or periods) with no rainfall within the comparison window. For FDs (or FPs) with rainfall, the NFDs (or NFPs) were chosen from days (or periods) that experienced rainfall, but recorded no flooding. Traffic impacts due to flooding are then evaluated using Equation 5, where change in VHT is assessed across links on a FD (or FP) compared to a NFD (or NFP).

$$VHT_{i\,F} = \sum_j (v_{i,j} * tt_{i,j})_F \tag{3}$$

$$VHT_{i\,NF} = 1/K(\sum_{k=1}^{K} \sum_j (v_{i,j} * tt_{i,j})_{NFk}) \tag{4}$$

$$\Delta\, Travel = \sum_i [(VHT_i)_F - (VHT_i)_{\,NF}] \tag{5}$$

where i= roadway link

j= time-of-day (TOD) period

K= maximum number of comparable days/periods considered for average NFD/ NFP calculations (4 for citywide impacts and all possible NFPs during 3 weeks prior to and after flood periods for localized impacts)

$\Delta\ Travel$ = change in VHT due to flooding, measured in veh-hrs

$tt_{i,j}$= travel time on segment i during TOD j on a FD/FP or NFD/NFP

$v_{i,j}$= traffic volume on segment i during TOD j on a FD/FP or NFD/NFP

$(VHT_{i,j})_F$= vehicle-hours of travel on segment i during TOD j on a FD/ FP

$(VHT_{i,j})_{NF}$= vehicle-hours of travel on segment i during TOD j on a NFD /FP

4.3. Localized Spatial Boundary Selection (Localized Analysis Only)

Since crowdsourced flood incident reports from Waze are spatially and temporally disaggregate, it provides an opportunity to analyze flood impacts on a more localized scale, as flooding on one roadway link is unlikely to impact traffic throughout City of Norfolk in a homogeneous way. Thus, in the localized flood analysis, a spatial boundary can be used to define the affected area around the location of the flood incident report. The roadway links within the localized boundary were selected as those most likely to be affected by the flood incident, based on the network theory measure hub dependence (Rodrigue, Comtis and Slack, 2017). Hub dependence, or *Hvalue,* is a measure of a node's vulnerability and represents the share of traffic borne by the highest volume traffic link among all links connected to a node, and is calculated as:

$$Hvalue = \frac{(AADT_{ij})_{max}}{\sum_{j=1}^{J}(AADT_{ij})} \quad (7)$$

where,

i = current node

j = adjacent node

J = maximum number of nodes that are connected to node j

$AADT_{ij}$ =annual average daily traffic of the link between node i and node j

Weak nodes, exhibiting higher hub dependence values, are heavily dependent on the conditions of the connected links for movement of traffic, and disruptions on any link connecting to the weak node

would greatly affect operations at the node. On the other hand, stronger nodes with lower hub dependence values have a more even distribution of traffic among the links they are connected to. Thus, their operations are less likely to be affected by disruptions of travel on any single connecting link. In a sense, nodes with lower hub dependence values may be more resilient when facing incidents and disruptions (Ducruet, 2008).

Hub dependence values for all nodes in the network based on 2009 AADT (collected from the HRRTDM) were calculated. To define the spatial boundary of roadway links impacted by a flood incident report, the node nearest to the flood report observation is assigned as the affected node and all the links connected to that node as affected links, by default. Then, the *Hvalue* for nodes connecting to these affected links are compared. If the adjacent node has a lower *Hvalue* than the affected node, it implies that the affected node is relatively stronger, and does not rely as much on the adjacent node for movement of traffic. In this case, the spatial boundary was cut off at the link leading to the adjacent weaker node. On the other hand, if the adjacent node has a higher hub dependence value than the affected node, it implies that the adjacent node's operations are highly affected by operations of the affected node. In this case, the adjacent weak node becomes an affected node, and the spatial boundary is expanded to include the set of links connected to the new affected node(s). A 10% threshold (see Figure 4) was used to ensure a sufficient difference in hub dependence values of adjacent nodes is observed before a boundary is set, and to prevent nodes with similar hub dependence values from being excluded from the boundary. The process of increasing the spatial boundary for each Waze-reported flood incident is repeated until there is a node that has a minimum 10% smaller hub dependence value than its previous adjacent node, or until the boundary of the study network (City of Norfolk border) was reached. This process of spatial boundary selection by hub dependence metric is illustrated in Figure 4.

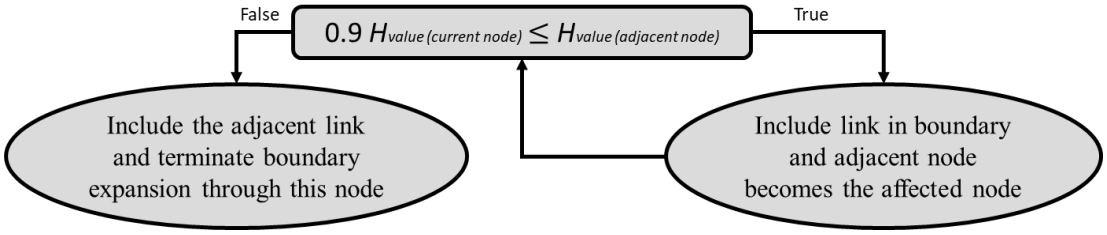


Figure 4 Framework for estimating localized spatial boundary of flooding impacts

Figure 5 shows an example of the spatial boundary selection starting with a flood report assigned to node A with $H_{value} = 0.427$. Following the procedure in Figure 4, the links within the localized

boundary are shown in yellow. The links connected to node *A* are included in the spatial boundary by default. Next, adjacent nodes *B*, *C*, *D*, and *E* were evaluated. Since the *Hvalue* of *B* was greater than that of *A*, *B* (and its connecting links) were included in the boundary. Moving on to the next adjacent node, *H*, which has an *Hvalue* less than 90% of node B, the boundary is terminated at node *H*. This propagation of spatial boundary takes place through each link that is connected to node *A* and terminates when the hub dependence value falls below 90% of the previous node's *Hvalue*.

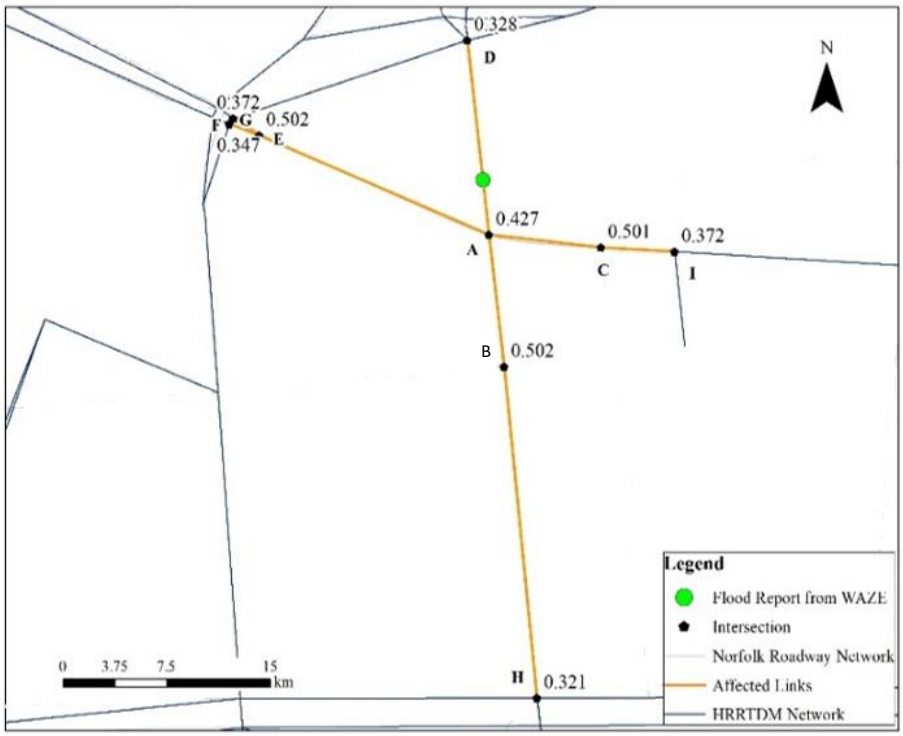


Figure 5 Localized spatial boundary assignment with flood incident assigned to Node A

5. Results and Discussion

5.1. Flood Impact Estimation at CCS Locations

Variation of traffic across all VDOT CCS locations with available data were first compared to understand the baseline roadway network impacts due to flood incidents. The CCS are strategically placed on major arterials and freeways where there are no historic congestion or bottlenecking issues to ensure accurate volume estimates. Due to the specific location selection criteria and sparse spatial representation of CCS across Norfolk, an accurate estimation of the flooding impacts throughout the network cannot be made,

but general trends can be observed, as shown in Table 2. The CCS data was collected on the same FDs and NFDs (January 2017 – August 2018) considered in the citywide flood impact analysis, at 15 minute intervals throughout the day. The table shows average speeds and volumes at across 15-minute intervals.

Table 2 CCS volume and speed trends, categorized by roadway type

Count Station ID	Facility Type	FD average volume (veh/15-mins)	NFD average volume (veh/15-mins)	Δ volume (%)	Δ volume p-value	FD average speed (mph)	NFD average speed (mph)	Δ speed (%)	Δ speed p-value
150114_1	Principal Arterial	188.35	210.80	-10.6%	<0.001	37.64	38.77	-2.9%	<0.001
050169_2	Freeway	709.50	789.70	-10.2%	<0.001	59.94	62.10	-3.5%	<0.001
050306_2	Freeway	594.59	655.47	-9.3%	<0.001	60.00	62.44	-3.9%	<0.001
150065_4	Freeway	553.78	627.52	-11.8%	0.001	71.91	77.76	-7.5%	<0.001
150110_2	Principal Arterial	88.76	106.74	-16.8%	<0.001	39.95	46.94	-14.9%	<0.001
150110_4	Principal Arterial	91.77	111.75	-17.9%	<0.001	35.52	38.63	-8.1%	<0.001
150119_2	Principal Arterial	174.11	188.87	-7.8%	<0.001	36.84	39.01	-5.6%	<0.001
150119_4	Principal Arterial	144.77	160.87	-10.0%	<0.001	36.19	37.14	-2.6%	<0.001
150120_4	Principal Arterial	98.32	109.86	-10.5%	<0.001	43.84	48.55	-9.7%	<0.001
Total	All	293.77	329.06	**-10.7%**	-	46.87	50.14	**-6.5%**	-

The CCS volumes and speeds on the 10 FDs (as reported by the City of Norfolk) were compared with their respective NFD counterparts. The NFDs used in this study come from the three weeks within (before and after) the flood incident report day, excluding the week of the flood incident. The data was compared at 15-minute intervals for the 24-hour day and then aggregated over the 10 FDs in the 20-month study period. A two-sample, one-tailed paired Student's t-test was conducted, and revealed link volumes and speeds at CCS locations were statistically significantly lower on FDs than on NFDs (with all p values < 0.01). An average 11% decrease in traffic volumes and 7% decrease in travel speeds were observed across the CCS locations on FDs. This result suggests that FDs consistently experience decreased traffic demand.

At the same time traffic volumes are decreased, those who are traveling on FDs also experienced slower speeds, which is indicative of increased travel times.

5.2. Traffic volume estimation model training and validation

While general trends of the traffic impacts of recurring flooding can be observed with the spatially limited CCS data, a network-wide impact assessment requires more spatial coverage. Here, the proposed data predictive model (using agency-provided roadway characteristics, hydrology data, and flood reports along with crowdsourced traffic flow data) estimates volumes across all freeway and arterial links in Norfolk. To create the ground truth dataset for model calibration (training) and validation (testing), all the days in the 20-month period were divided into categories based on environmental conditions. The days were categorized as combinations of three levels of rainfall (rainfall = 0 in., 0 < rainfall ≤ 0.5 in., and rainfall > 0.5 in.) and three levels of tide (tide level < 1 ft, 1ft < tide level ≤ 2 ft, and tide level > 2ft), thus creating 9 combinations of environmental conditions based on rainfall and tide levels. Twenty percent of the days in each category were randomly selected to create the ground truth dataset, ensuring representation of all combinations of rainfall and tide conditions.

Linear regression, CRT, and random forest models were developed with all the variables previously mentioned in three categories: hydrological, roadway, and traffic flow characteristics. The model fits (measured by RMSE and NRMSE values), along with statistically significant variables, are shown in Table 3 for comparison across models. The two random forest model specifications outperformed linear regression and CRT in terms of model fit. For random forest models, the first model (RF1) used only the roadway and traffic flow characteristics as input variables. In this model, the StL dynamic crowdsourced trip counts had less importance than other static variables such as number of lanes and type of day, which is counterintuitive. When the hydrological variables are introduced into the random forest mode specification (RF2), tide level and rainfall were found to be the least important variables, but the StL trip count became the highest significance variable, which is intuitive. In the RF2 model, other relatively high importance variables described patterns associated with traffic flow in specific environments, such as TOD, per lane capacity, posted speed limit, and link speed. This RF2 model specification also proved to be the best performing (with the lowest RMSE and NRMSE).

Table 3 Comparison of data predictive models

Model Type	RMSE	NRMSE	Significant/ high importance variables*	Insignificant/low importance variables**
Linear Regression	2384.43	0.085	Rainfall Tide level Flooding Number of lanes Posted speed limit TOD StL trip counts	Per lane capacity Segment speed Type of day
CRT	2512.22	0.157	StL trip count Posted speed limit TOD	Rainfall Tide level Flooding Number of lanes Per lane capacity Segment speed Type of day
RF1 with roadway and traffic characteristics	Train: 1573.97 Test: 3399.23	Train: 0.026 Test: 0.058	TOD Per lane capacity StL trip counts Posted speed limit	Number of lanes Type of day
RF2 with roadway, traffic, and hydrologic variables	Train: 1341.98 Test: 2865.60	Train: 0.022 Test: 0.048	StL trip counts TOD Per lane capacity Posted speed limit Segment Speed	Type of day Tide level Rainfall

Variable Importance: predictive power of the variables in the random forest model

*High importance variables: normalized variable importance over 0.5

**Low importance variables: normalized variable importance under 0.1

5.3. Citywide Roadway Network Impacts

The RF2 model was propagated to the HRRTDM roadway network in Norfolk to predict the volumes on each roadway segment across all TODs. The HRRTDM roadway network consists of 7736 segments, which were fed into Streetlight Data to retrieve the associated StL trip counts, segment speed, and travel time on each segment. StL trip counts and segment speed, along with other roadway and hydrological variables, were used as inputs into the random forest model (RF2) to obtain volume estimates on FDs and NFDs. Total VHT on FDs and NFDs were calculated per Equation 2. There were 11 FDs recorded in the 20-month analysis period by City of Norfolk employees. One of the FDs was discarded due to insufficient

comparable NFDs within the six-week window. Table 4 shows the total VHT on each FD compared to the corresponding NFDs.

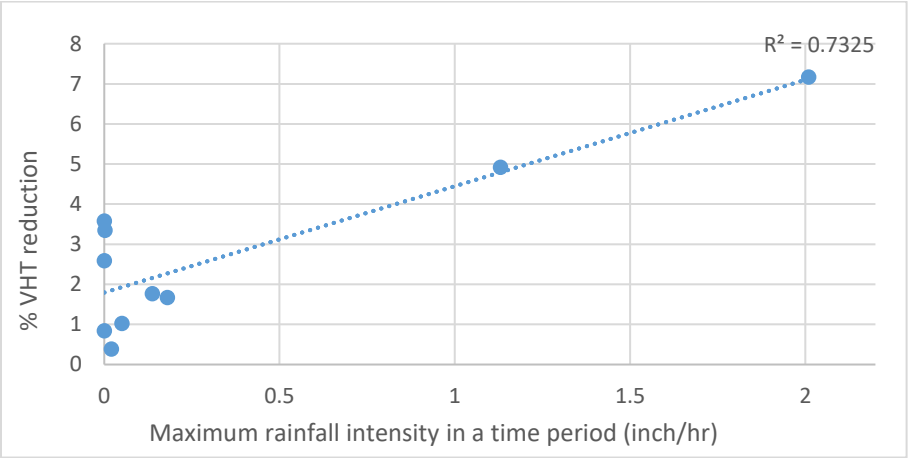
Table 4 Citywide network impact summary

Flood report date	Max tide level (ft)	Max rainfall intensity (in/hr)	FD VHT (veh-hrs)	NFD VHT (veh-hrs)	FD-NFD VHT (veh-hrs)	Δ VHT (%)
3/6/2017	1.000	0.002	6,261,710	6,478,423	-216,713	-3.35%
4/5/2017	2.208	0.000	6,594,179	6,649,965	-55,786	-0.84%
4/12/2017	1.148	0.000	6,311,667	6,546,158	-234,491	-3.58%
8/29/2017	2.854	1.130	6,005,899	6,316,644	-310,745	-4.92%
9/5/2017	1.877	0.050	6,416,229	6,482,857	-66,627	-1.03%
3/6/2018	3.051	0.137	6,133,580	6,244,388	-110,808	-1.77%
5/18/2018	2.430	0.180	6,447,560	6,557,150	-109,590	-1.67%
5/22/2018	1.791	0.000	6,462,511	6,634,472	-171,960	-2.59%
8/11/2018	2.493	2.010	4,503,163	4,851,056	-347,898	-7.17%
8/13/2018	2.572	0.020	6,526,115	6,551,927	-25,813	-0.39%

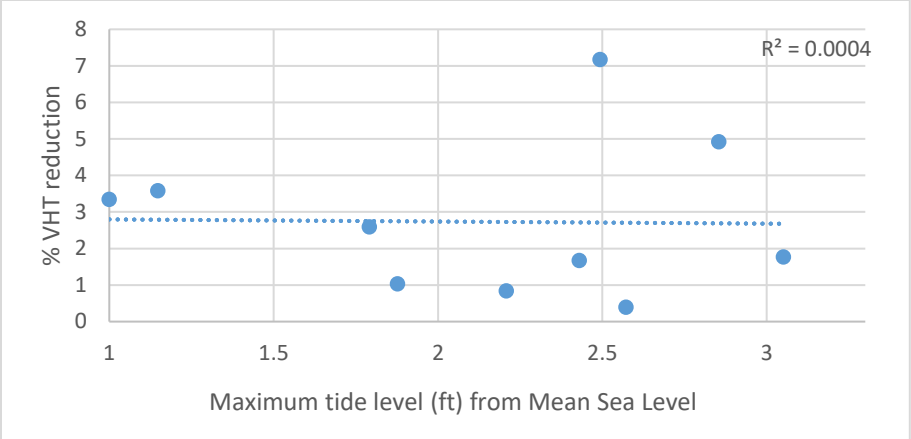
Table 4 shows that, based on the predicted vehicle volumes, network-wide VHT was consistently reduced on FDs compared to NFDs in the citywide analysis, on average by 3%. This decrease is consistent with trends in the CCS analysis, though not substantial. The result may be attributed to two factors: (1) cumulative change in VHT may not be a sufficient metric for quantifying the effect of flooding, and (2) the spatial aggregation at the city-level may be too large for assessing the impacts of local recurring flooding. Since the VDOT CCS data also showed a reduction in travel speeds and personal vehicle volumes on FDs, it is likely that the individual effects of increased travel time and reduced volumes were somewhat nullified when multiplying the two for the cumulative effect measured in VHT. Decreased network VHT may imply higher rates of abandoned trips, which would signify an economic impact of recurring flooding (due to decreased business transactions, work productivity loss, etc.). Considering a net difference in VHT over the roadway network of the entire city may also temper the more significant local

effects of flooding experienced by specific areas within the city. Thus, a more spatially disaggregate analysis of flooding impacts on the roadway network is necessary to fully understand the effects of recurring flooding.

While the sample of flood days is small (N=10), relationships between hydrological variables and traffic impacts on the roadway network still appear to exist. Figures 6(a) and 6(b) show the relationship between rainfall intensity, tide level, and VHT reduction. Increasing rainfall intensity had a relatively consistent positive correlation with reduction of VHT (Figure 6[a]); however, no relationship appears to exist between tide level and reduction in VHT (Figure 6[b]). It is possible that effects of tide-induced flooding are more local than that of rain-induced flooding. In other words, tide-induced flooding impact roadway segments near the shoreline in a spatially and temporally consistent manner. However, this temporal and spatial granularity cannot be analyzed with the City of Norfolk flood report data, which lacks timestamps and representative spatial coverage.



6(a) Estimated %VHT reduction vs. rainfall intensity



6(b) Estimated %VHT reduction vs. maximum tide level

Figure 6 Comparison of roadway network impacts by hydrological variables

5.4. Localized Roadway Network Impacts

Using City of Norfolk flood incident reports which are neither spatially nor temporally disaggregated, only a citywide analysis of recurring flooding impacts is feasible. However, Waze flood incident report data contains both timestamp and location (latitude and longitude) data, allowing for local analysis of recurring flooding impacts. In this section, the analysis of recurring flooding impacts is defined by five TOD periods (consistent with HRRTDM definitions) and a local geographic boundary around the location of the flood incident report (see Section 4.3 for methodology for localized spatial boundary selection). StL trip count data is too sparse to be aggregated at an hourly interval, particularly for off-peak travel periods, hence the selection of TOD periods for analysis. Total VHT during a flood period (FP) and an average non-flood period (NFP) is estimated using Equations 3 and 4, after obtaining the link volumes from the RF2 model (normalized RMSE for the training data: 0.03, testing data: 0.07). The candidate NFPs used in this study come from three weeks before and after the flood incident report, excluding any TOD with another reported flood incident within one mile radius of the original report. After removing the FPs that did not have any candidate NFPs, fell outside the roadway network being analyzed (e.g., on local streets, ramps, or centroid connectors), or had insufficient Streetlight Data trip counts to predict link volumes inside the localized spatial boundary, 340 flood report observations remained (representing 51% of the original Waze flood reported incidents between August 2017 and August 2018).

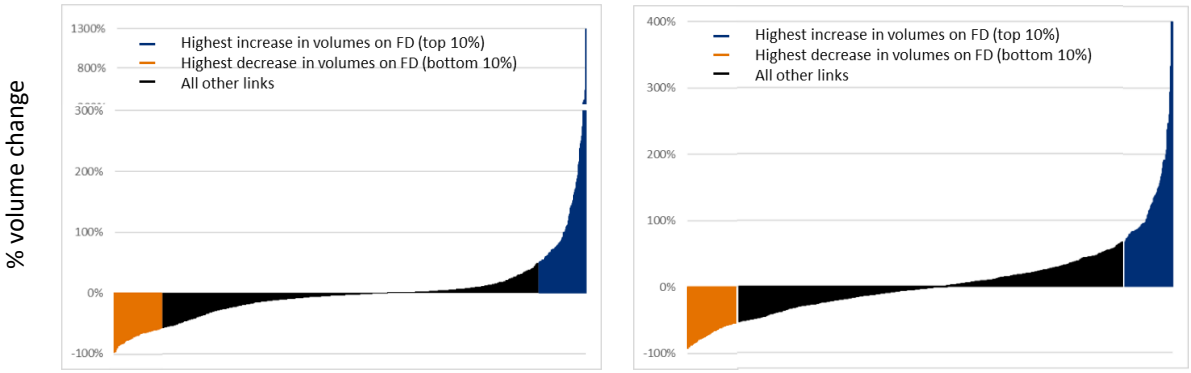
A link-by-link impact analysis for all roadway links within the localized boundaries was conducted to compare travel during FPs compared to NFPs. The localized network impact summary of the link-by-link analysis (categorized by time period) is shown in Table 5.

Table 5 Localized network impact summary

Day Type	Period	# of Waze flood reports	Δ Speed (FP-NFP)			Δ Volume (FP-NFP)			Δ VHT (FP-NFP)		Δ VMT (FP-NFP)	
			Mean % diff	Medi-an % diff	p-value	Mean % diff	Medi-an % diff	p-value	Aggr. % diff	p-value	Aggr. % diff	p-value
Wkday	12a - 6a	10	8%	-5%	0.41	2%	-19%	0.01*	-25%	0.04*	-16%	0.05*
Wkday	6a - 9a	26	1%	-4%	0.13	-3%	-11%	0.06	-1%	0.89	-9%	0.00*
Wkday	9a - 3p	59	19%	0%	0.42	21%	0%	0.00*	4%	0.08	3%	0.11
Wkday	3p - 6p	126	23%	-4%	0.15	1%	-3%	0.00*	-9%	0.00*	-16%	0.00*
Wkday	6p - 12a	65	23%	0%	0.00*	6%	-3%	0.02*	-6%	0.01*	-5%	0.04*

*Statistically significant at 95% confidence level

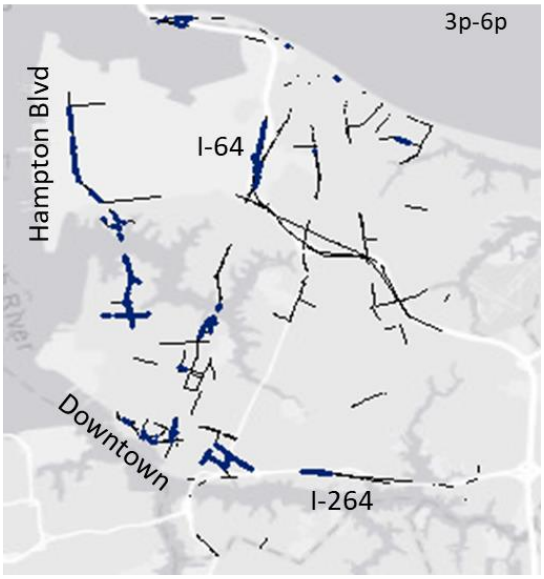
The discussion about localized impacts of flood incidents here is focused on workdays, because the number of Waze flood reports on non-workdays (weekends and holidays) was too small per each TOD period for analysis (N for each non-workday TOD ranged from 2 to 24). On workdays, across almost all time periods, the majority of affected links experience a decline in both speed and traffic volume during FPs compared to NFPs (example for 3p – 6p and 6p – 12a time periods shown in Figure 7), as indicated by the median percent differences in Table 5. However, the average link speed and volume changes during FPs compared to NFPs are predominantly positive, indicating a few links experiencing significantly higher speeds or volumes during FPs. As seen in Table 5, workday evening (3p – 6p) and overnight (6p – 12a) periods experience the most statistically significant reductions in travel speed, traffic volumes, and change in vehicle hours and miles of travel during flood periods compared to non-flood periods, and are examined in more detail here. As seen in Figure 7, for the workday 6p to 12a time period, the affected links observe a maximum 100% decrease in volume, while select links experience volume increases in excess of 300%. The significant change in traffic volumes in the workday evening and night periods (3p – 6p and 6p – 12a) are also reflected in the VMT measures, which show a reduction in travel during FPs compared to NFPs, with an aggregate VMT decline of 12% across all affected links. There are similar statistically significant reductions in VHT observed in the evening and night periods, suggesting that road users are either avoiding travel, or changing their destinations.



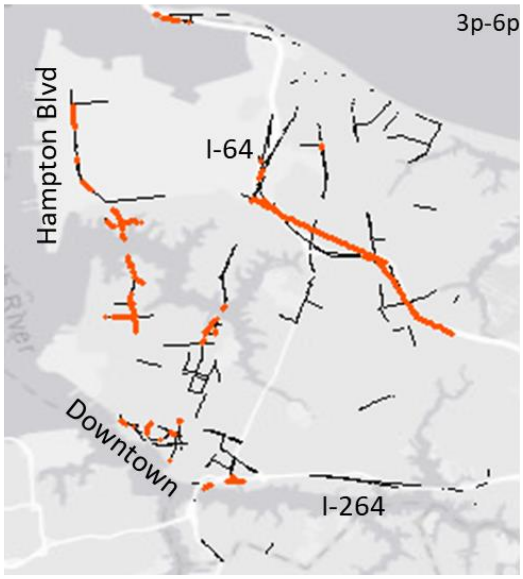
7(a) 3p – 6p **7(b)** 6p – 12a

Figure 7 Distribution of change in traffic volume between flood periods & non-flood periods

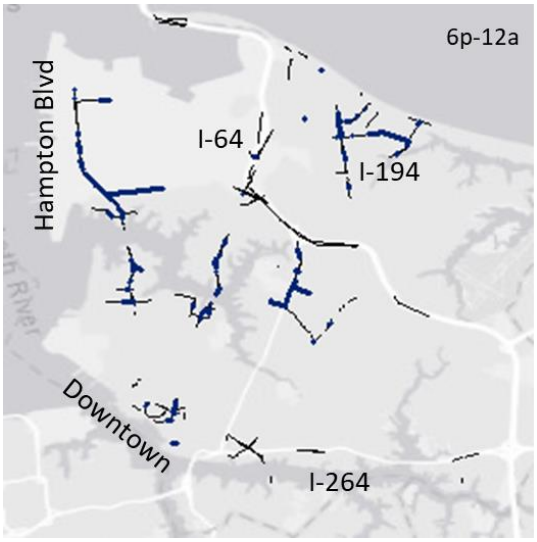
Figure 8 shows the spatial distribution of links experiencing the greatest increase (top tenth percentile, marked in blue) and greatest decrease (bottom tenth percentile, marked in orange) in traffic volumes during 3p – 6p (Fig 8[a] and 8[b]) and 6p – 12a (Fig 8[c] and 8[d]). The links in black represent all the other links that were considered in the localized spatial boundary analysis for the TOD period (middle 80th percentile). 2167 unique links are considered in the localized boundary analysis across all time periods, with the 3p to 6p time period showing the highest number of impacted links (985 unique links, with several links affected multiple times across multiple flood reports between August 2017 and August 2018). As shown in Figure 8, the Hampton Boulevard corridor is highly impacted during both TOD periods, with some flood incidents causing a large increase in traffic volumes, and some flood incidents causing a large decrease in volumes, compared to the NFPs. In the downtown area, the majority of impacted links experience increased traffic volumes. On the other hand, interstate corridors (I-64 and I-264) generally see decreased volumes during flood periods. A few of the highly impacted links (in orange or blue) appear in both maps, implying that these links (25% of highly impacted links during 3p-6p period and 14% of highly impacted links in the 6p-12a period) experience both increasing and decreasing traffic volumes during different flood events.



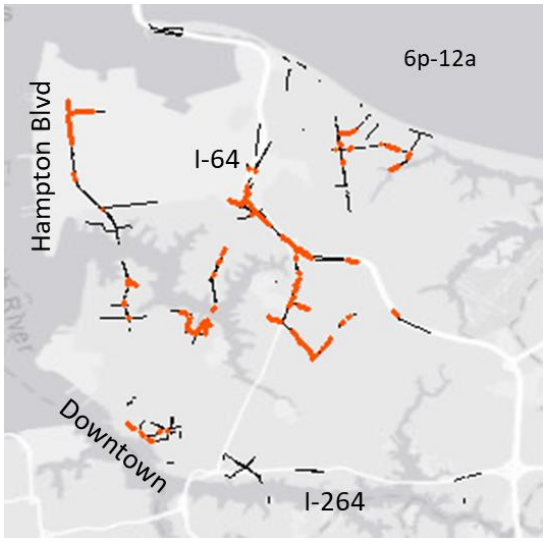
8(a) Links with greatest increase in volumes



8(b) Links with greatest decrease in volumes



8(c) Links with greatest increase in volumes



8(d) Links with greatest decrease in volumes

Figure 8 Highly affected links during flood periods

Impacts of flood incidents on the roadway network are also analyzed by roadway functional classification. The functional classification of affected links in the analysis include interstates, minor freeways, principal arterials, major arterials, minor arterials, minor collectors, and local collectors. Figure 9 shows the distribution of links experiencing greatest increase and decrease in traffic volume during FPs by roadway functional classification. The most impacted links within the localized boundary analysis were found to be interstates, principal arterials, minor arterials, and minor collectors (other functional classifications had less than 5 links impacted in the top and bottom 10th percentiles).

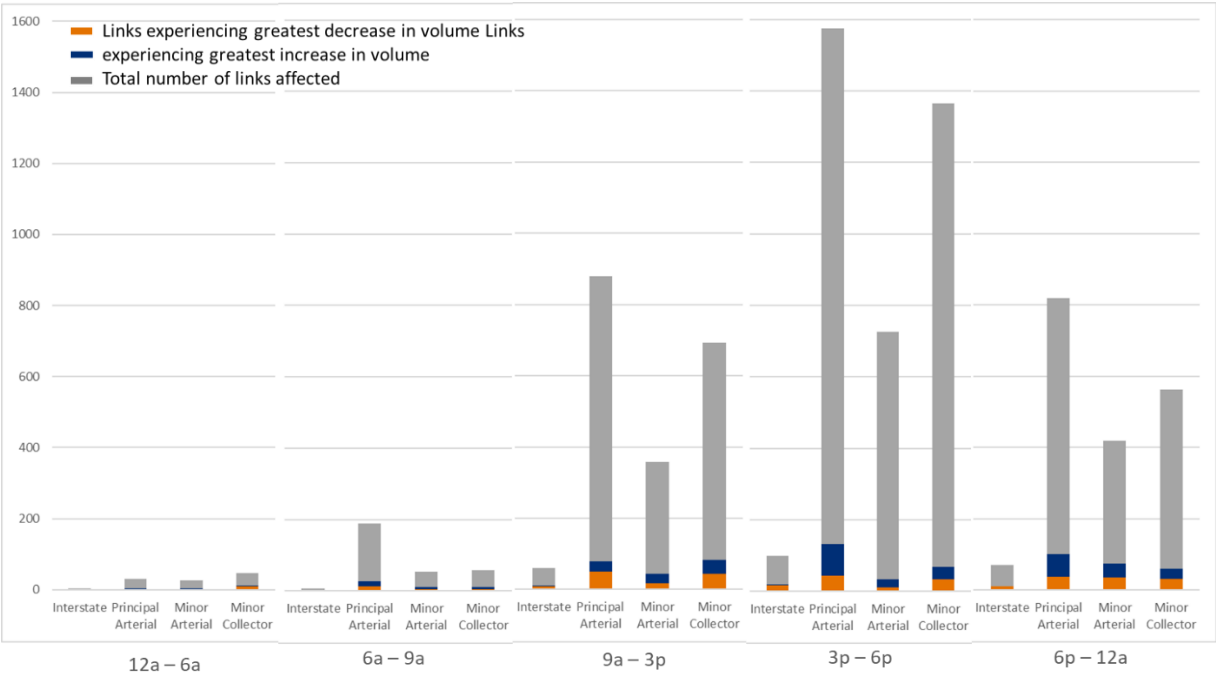


Figure 9 Distribution of most impacted links by roadway facility type

As seen in Figure 9, there are relatively few links impacted by flooding in the morning periods (12a – 6a and 6a – 9a), which is a result of lower traffic volumes and fewer flood incident reports for the overnight and early morning periods. In the remainder three time periods, interstate links within the localized boundary generally observe a decline in traffic volumes when a flood incident is reported. Principal arterials experience a decline in traffic volumes during the midday period (9am to 3pm) during FPs compared to NFPs, but see traffic volumes rise in the evening and night periods (3pm to 12am). This implies that during the evening peak and night time periods, when high numbers of flood incidents are reported, more road users opt to use the principal arterials (likely switching routes from interstates). It is possible that travelers perceive the access-controlled nature of interstate corridors as a disadvantage during flood periods, as their ability to switch routes dynamically is limited by access to ramps. Minor arterial links also appear to experience increased traffic volumes during FPs. Minor collector roads, though affected, do not show a consistent trend in direction of volume change by time period.

6. Conclusions and Limitations

Prior studies examining recurring flooding and subsequent impacts on the transportation network have used projected or simulated data. This study is the first to use empirical data to assess such impacts,

leveraging crowdsourced data to expand the spatial and temporal coverage of agency datasets in understanding the dynamic effects of recurring flood disruptions on roadway users. With recurring flooding becoming an increasing concern for coastal cities, this type of analysis demonstrates a framework for combining limited agency flood incident report data with crowdsourced flood reports, to understand the subsequent impacts on road users (in response to recurring flooding).

The study first estimates the citywide impact of recurring flooding on the Norfolk, Virginia roadway network. Due to lack of comprehensive traffic volume data, a framework to estimate traffic volumes using agency-provided and crowdsourced data was established, expanding the spatial coverage of volume data. These volume estimates show a 3% decline in VHT on FDs (as compared to NFDs) during the 20-month study period. With VHT being a cumulative measure of travel speeds and vehicle volumes, a simultaneous decrease in volumes and increase in travel times would not be sufficiently described by a single measure like VHT, especially when aggregated across the entire city which contains many links unaffected by flood events. Thus, the second part of the study examined the localized impacts of recurring flooding near the location of the crowdsourced flood incident report. Results suggest that majority of links within impacted areas show a decline in speeds and volumes during flood periods. Volume estimates show a significant change in traffic volumes in the workday evening and night periods during FPs, with an aggregate VMT decline of 12% across all affected links. However, select links experience sizable increases in speeds and volumes. There are similar statistically significant reductions in VHT observed in the evening and night periods, suggesting that travelers are either avoiding travel, or changing their destinations. Particularly in the evening and night periods, localized analysis results point to reductions in travel during FPs, with decreased volumes on interstate corridors and increased volumes on principal and minor arterials, compared to NFPs (suggesting route shifts as a result of flooding).

Results of this study strongly suggest that the impact of recurring flooding events on transportation networks is local, thus a citywide or regional analysis is not recommended due to the heterogeneous effects of flooding across various links. Analysis across a city or region may underestimate the impact of recurring flooding on travelers, as they abandon trips and shift routes in specific subareas. Spatial and temporal disparities in travel impacts are better explained through the localized impact assessment. Since recurring flooding is dynamic event, the framework provided in this study can serve as a precursor to identify recurring problem areas and periods for agency mitigation. In the case of Norfolk, since evening periods are more impacted than morning periods, mitigation efforts could be concentrated in key areas during those evening periods.

This study has certain limitations. First, like all studies that use crowdsourced data, the fidelity and accuracy of the data is an issue. Waze incident data is reported by Waze users, and there is no ground truth roadway flooding data to enable assertion of trustworthiness measures on the crowdsourced flood incident data. Regardless of trustworthiness, all Waze incident reports are included in the analysis. Second, the ground truth data used to train the data-predictive model is obtained from VDOT CCS locations which are located mostly on principal arterials or freeways. Thus the model's ability to predict traffic volumes on roads of lower functional classification is limited. In this study, only links contained in the regional travel demand model network are included (thereby excluding roads in the lowest functional classifications). Furthermore, there is some spatial mismatch between links in the Streetlight Data OSM layer and the HRRTDM roadway network, and these mismatches occur mostly on minor roads, further exacerbating the accuracy of volume predictions on these links. It is important to note that the results and conclusions in this study are focused on major roads, due to these limitations. Lastly, the Streetlight dataset analyzed here only considers light duty vehicle travel, thus heavy vehicle movement is not captured. Diversion from freight schedules incur significant economic costs, which has not been quantified with this study.

Nonetheless, the data-predictive framework presented in this study can be generalized to be applied to other crowdsourced datasets, which are highly valuable when and where agency data is limited. This framework is applicable to a wide range of incident analyses such as congestion or accident analysis, post-disruption analysis, etc. Additionally, with the emergence of smarter cities and increasing availability of crowdsourced data capturing real-time traffic flow and hydrological variables, this framework can be a key component in strategic traffic rerouting and dynamic stormwater management to mitigate the impacts of recurring flooding, to ultimately increase resilience of coastal cities against such events.

References

Amin-Naseri M, Chakraborty P, Sharma A, Gilbert S B, and Hong M (2018) Evaluating the Reliability, Coverage, and Added Value of Crowdsourced Traffic Incident Reports from Waze. Transportation Research Record: Journal of the Transportation Research Board. 2672 (43):34-43

Böcker L, Dijst M, Prillwitz J (2013) Impact of Everyday Weather on Individual Daily Travel Behaviours in Perspective: A Literature Review. Transport Reviews. 33 (1):71-91

Chang H, Lafrenz M, Jung I W, Figliozzi M, Platman D, Pederson C (2010) Potential Impacts of Climate Change on Flood-Induced Travel Disruptions: A Case Study of Portland, Oregon, USA. Annals Association of American Geographers. 100(4):938-952

Donovan B, Work D B (2017) Empirically quantifying city-scale transportation system resilience to extreme events. Transportation Research Part C: Emerging Technologies, 79:333-346

Ducruet C. (2008) Hub dependence in constrained economies: the case of North Korea. Maritime Policy & Management. 35(4):377-394.

Frequently Asked Questions about StreetLight Insight. www.streetlightdata.com/support/faq/general. Accessed on April 4, 2018.

Goodall N, Lee E (2019) Comparison of Waze crash and disabled vehicle records with video ground truth. In Transportation Research Interdisciplinary Perspectives. Volume 1

Jacobs J M, Cattaneo L R, Sweet W, Mansfield T (2018) Recent and Future Outlooks for Recurring Flooding Impacts on Roadways on the US East Coast. Transportation Research Record: Journal of Transportation Research Board. 2672(2):1-10.

Lenkei Z (2018). Crowdsourced traffic information in traffic management: Evaluation of traffic information from Waze, Dissertation, KTH School of Architecture and Built Environment.

Lu Q C, Peng Z R (2011) Vulnerability Analysis of Transportation Network under Scenarios of Sea Level Rise. Transportation Research Record: Journal of the Transportation Research Board. 2263:174-181.

Ouyang, R W, Kaplan L M, Toniolo A, Srivastava M and Norman T J (2016) Aggregating crowdsourced quantitative claims: Additive and multiplicative models. IEEE Transactions on Knowledge and Data Engineering, 28(7):1621-1634.

Rodrigue J P, Comtis C, Slack B (2017) The Geography of Transport Systems. 4th Edition. Routledge.

Sadler, J M, Goodall J L, Morsy M M, Spencer K (2018) Modeling Urban Coastal Flooding Severity from Crowd-sourced Flood Reports using Poisson regression and Random Forest. Journal of Hydrology, 59:43-55.

Suarez P, Anderson W, Mahal V, Lakshmanan T R (2005). Impacts of Flooding and Climate Change on Urban Transportation. Transportation Research Part D: Transport and Environment. 10 (3):231-244.

Sweet W, Park J, Marra J, Zervas C, Gill S (2016) Sea Level Rise and Recurring Flood Frequency Changes Around the United States. Publication NOAA Technical Report NOS CO-OPS 073. National Oceanic and Atmospheric Administration, U.S. Department of Commerce.

Sweet W, Dusek G, Marcy D, Carbin G, Marra J (2019) 2018 State of U.S. High Tide Flooding with a 2019 Outlook. Publication NOAA Technical Report NOS CO-OPS 090. National Oceanic and Atmospheric Administration, U.S. Department of Commerce.

System to Track, Organize, Record, and Map (STORM), accessible at http://gisapp1.norfolk.gov/storm-map

US EPA. Climate Change Indicators in the United States: Coastal Flooding, United States Environmental Protection Agency, 2016. Fourth Edition. EPA 430-R-16-004. www.epa.gov/climate-indicators. Accessed May 10, 2018.

Waze, via the Waze for Cities data-sharing program, https://www.waze.com/wazeforcities (data used covers August, 2017 to August, 2018).

Zhu Y, Ozbay K, Xie K, Yang H (2016). Using Big Data to study resilience of taxi and subway trips for hurricanes Sandy and Irene. Transportation Research Record: Journal of the Transportation Research Board. 2599:70 - 80.

CHAPTER 3

ASSESSING TRUSTWORTHINESS OF CROWDSOURCED FLOOD INCIDENT REPORTS USING WAZE DATA: A NORFOLK, VIRGINIA CASE STUDY

Shraddha Praharaj, Faria T. Zahura, T. Donna Chen (corresponding author), Madhur Behl, Jonathan L. Goodall

Accepted for publication in Transportation Research Record - May 2021

Presented *at Transportation Research Board 100th Annual Meeting*, January 2021

To be presented at *ASCE Transportation & Development Institute*, June 2021

ABSTRACT

In recent years, climate change and sea-level rise have caused higher and prolonged high tides which, in combination with rainfall, storm surges, and insufficient drainage infrastructure, result in recurrent flooding in coastal cities. There is a pressing need to understand the occurrence of roadway flood incidents in order to enact appropriate mitigation measures. Agency data for roadway flooding events are scarce and resource-intensive to collect. Crowdsourced data can provide a low-cost alternative for mapping roadway flood incidents in real time; however, the reliability is questionable. This research demonstrates a framework for asserting trustworthiness on crowdsourced flood incident data in a case study of Norfolk, Virginia. Publicly available (but spatially limited) flood incident data from the city in combination with different environmental and topographical factors are used to create a logistic regression model to predict the probability of roadway flooding at any location on the roadway network. The prediction accuracy of the model was found to be 90.48%. When applying this model to crowdsourced Waze flood incident data, 71.7% of the reports were predicted to be trustworthy. This study demonstrates the potential for using Waze incident report data for roadway flooding detection, providing a framework for cities to identify trustworthy reports in real-time to enable rapid situation assessment and mitigation to reduce incident impact.

Keywords: crowdsourced data, flooding, trustworthiness, logistic regression

INTRODUCTION

In recent years, crowdsourced data has emerged as a low-cost method for data collection in various fields. In the transportation domain, there are many areas of research which have insufficient or non-existent agency-provided data, where crowdsourced data shows promise to be a useful alternative resource for research and analysis in these domains, such as bicycle ridership ***(1)***, traffic analysis**,** ***(2***), and accident reporting ***(3)***.

One of the domains *with very limited* agency data is incidences of roadway flooding. Recurrent flooding as a result of rainfall, high tides, or both is becoming more prevalent in coastal cities. These flood incidents cause inundation of roadways for up to several hours, which deteriorates mobility and accessibility of travelers. Over the past six decades, almost thirty coastal cities in the US have witnessed a spike in the number of annual flood days, with some cities witnessing as many as 50 extra flood days every year ***(4)***. NOAA estimated a 125% increase in the number of annual flood days in the southeast coast of the US and 75% in the northeast coast between the years 2010 and 2015 ***(5)***. Some city or state agencies may collect flood incident data as a part of providing emergency management services. Preemptive, reliable knowledge of flood locations could significantly reduce response time for flood mitigation, thereby reducing delays and impacts to travelers. However, most cities do not collect standardized comprehensive data of roadway flooding incidents, making it difficult to make data-driven decisions for traffic rerouting and flood mitigation measures. One promising source for crowdsourced flood incident data is Google Waze, a GPS navigation app that allows users to report on a multitude of traffic-related incidents, including roadway flooding. While crowdsourced data can be a powerful alternative to revolutionize data collection where agency data is lacking, such data comes with its own set of limitations. Since crowdsourced data is not regulated, there can be human error, technical error, wrongful reporting, among other issues ***(6).*** This study explores the trustworthiness of crowdsourced Waze flood incident data in a case study of Norfolk, Virginia. The City of Norfolk has been collecting limited roadway flooding data due to the increasing frequency of recurrent flooding. This study combines the limited city flood report (ground truth) data with publicly available topographical and environmental data to build a model to assess trustworthiness of the crowdsourced Waze flood incident reports.

BACKGROUND AND LITERATURE REVIEW

In the past decade, crowdsourced datasets have become increasingly popular in transportation applications where traditional data collection methods are cost prohibitive. Waze has emerged as a popular crowdsourced dataset for transportation research, as the app allows users to enter location-specific and time-stamped reports of all sorts of traffic-related incidents: road closures, hazards (including roadway flooding), traffic jams, police presence, crashes, and more. Several studies have examined the usability of Waze incident data in transportation applications, though none have examined flood incident data. For example, Hoseinzadeh et al. ***(7)*** used independently collected Bluetooth speed data as ground truth and assessed the quality of Waze speed data. Their models concluded that the Waze data was more accurate in peak traffic hours, and achieved a prediction accuracy of almost 85%. Amin-Naseri et al. (***8***) conducted an analysis to quantify the potential added coverage of traffic incidents through Waze data, in addition to police reported data. It was estimated that 34.1% of Waze-reported incidents provide additional information not covered by any other dataset. Flynn et al. ***(3)*** developed machine learning models with spatial and temporal data which estimated police crash reports with high accuracy, and concluded that

Waze could be a potential data source to quickly identify crashes in real time, enabling faster police response. Goodall and Lee ***(9***) compared Waze-reported crashes and disabled vehicle information with video footage on a roadway segment. They found that 80% of crashes and 50% of the disabled vehicles were captured by Waze data, implying the potential to leverage Waze incident data to enable faster response times, shorter incident durations, and better incident information dissipation to the public. Eriksson ***(10)*** proposed a methodology to integrate crowdsourced Waze incident and congestion data with official traffic data to reduce redundancy, improve reliability, and measure severity of incidents. As evidenced by these studies, crowdsourced Waze data has potential to greatly expand and improve existing agency provided transportation data. However, thus far, no study has examined the value or trustworthiness of Waze flood incident report data.

Using real-time crowdsourced data to analyze impacts of roadway flooding is particularly appealing since the alternative (installing sensor technology to gain accurate spatially disaggregate data on rainfall, tide, and flood levels for every roadway link in a city) is cost-prohibitive. However, as crowdsourced data is not regulated, there could be erroneous reporting due to misunderstanding, confusion, carelessness, incompetence, or even intent to deceive ***(11)***. Emerging studies in the computer science and electrical engineering domains have considered asserting trustworthiness to various crowdsourced datasets dealing with environmental conditions ***(12, 13)***. For example, Flanagin and Metzger ***(14)*** suggest collaborating geospatial and environmental knowledge with crowdsourced data to assert credibility of the data. Similarly, a few studies in the geography domain incorporate topological and environmental characteristics to assert credibility of crowdsourced incident detection datasets. For example, Ostermann and Spinsanti ***(15)*** used crowdsourced Twitter and Flickr data to conduct a context analysis to identify hotspots of forest fires in Spain, using forest cover, distance to nearest hotspot, and inhabitant density in the area as contextual variables. The study concluded that geographic features of crowdsourced location information (also known as Volunteered Geographic Information, or VGI) provides a useful approach to filter crowdsourced data. Hung et al. ***(16)*** assessed the credibility of crowdsourced flood incident data by using contextual topological data. A binary logistic regression model with variables such as elevation and distance-to-flood-risk-zones showed prediction accuracy of 90% and 80% for training and testing datasets, respectively.

The studies conducted by Hung et al ***(16)*** and Ostermann and Spinsanti ***(15)*** combined VGI datasets with other relevant event-specific topographical variables to assert credibility on crowdsourced datasets. Most of the VGI datasets used in previous studies are very location specific, and do not have a presence globally. This study uses a similar approach, by using potential flood-related explanatory variables (environmental and topographic) to assert trustworthiness on crowdsourced Waze flood report data. Waze data has a much higher global footprint, and with that, this methodology can be applied to any city with enough Waze users. Also, unlike the previous studies, the ground truth data set is built from credible but spatially and temporally limited agency data.

DATA SOURCES AND PRE-PROCESSING

To build the trustworthiness model, this study uses contextual parameters such as environmental, topographic, and roadway infrastructure variables to explain the occurrence of a flood incident. This

section explains the different datasets used to build the model. The study period is restricted by the availability of flood incident data, which ranges from August 2017 to December 2018, with three weeks excluded due to loss of data.

Environmental data

The environmental data is composed of rainfall and tide level observations. Hourly tide levels referenced to the North American Vertical Datum (NAVD88) were obtained from National Oceanic and Atmospheric Administration's (NOAA) Sewell's Point station ***(17)***. Hourly rainfall data was collected from seven Hampton Roads Sanitation District (HRSD) observation sites. Both of these datasets are publicly available.

Topographic data

Three topographic features were used as model inputs: elevation, topographic wetness index (TWI) ***(18)***, and depth-to-water (DTW) ***(19)***. Elevation information at locations of the Waze flood incident reports is extracted from the United States Geological Survey (USGS) Digital Elevation Model (DEM), which has 1-meter horizontal resolution. The most recently published absolute vertical accuracy of the 3D Elevation Program (3DEP) DEMs within the United States is 3.04 meters at the 95% confidence level ***(27)***.

TWI and DTW were also derived using the DEM. TWI accounts for the tendency of any pixel (smallest grid in a raster file) in the topography to receive water from upstream and its tendency to drain that water. A high TWI value implies a high potential for accumulation of surface water runoff. TWI, defined by Beven and Kirkby ***(18)***, is a function of α, which is the upstream contributing area per unit contour length at a given pixel and $tan\ \beta$, which is the local slope at that pixel in the catchment, as shown in **Equation 1**:

$$TWI = ln\left(\frac{\alpha}{tan\ \beta}\right) \qquad (1)$$

DTW, defined by Murphy et al. ***(19)***, is a relative measure of soil moisture conditions, which approximates the elevation difference between a pixel in the topography and the nearest surface water pixel along the least slope path. DTW is a function of $\frac{dz_i}{dx_i}$, which is the slope of a pixel i in the topography along the least-cost path to the nearest surface water pixel, a, which is either 1 or $2^{0.5}$ depending on whether the path crosses the pixel parallel to the pixel boundary or diagonally, and x_c, which is pixel size, as shown in **Equation 2**:

$$DTW\ (m) = \left[\sum \frac{dz_i}{dx_i} a_i\right] x_c \qquad (2)$$

Topography pixels closer to surface water, in terms of both distance and elevation, tend to have smaller values of DTW, indicating wetter soil.

Predicted Surface Water Depth

Predicted street-level surface water depth was simulated using a physics-based hydrodynamic model (TUFLOW: Two-dimensional Unsteady Flow) model. The flood model solves 2D equations for shallow water and free surface flow to simulate overland flow, and it is coupled with 1D hydrodynamic network

software ESTRY (***20***) to simulate pipe flow. The 1D pipe/2D overland hydrodynamic flood model used in this study is described by Shen et al. ***(21, 22, 26)***, which provides details on the model construction, calibration, and evaluation process. The model used in this analysis covers roughly half the area of the city of Norfolk, VA (56.4 km^2). Surface flooding is simulated at one-hour time steps and at a spatial resolution of 2.5 m, which was then used to estimate water depth on street segments. TUFLOW is a high-fidelity model which simulates realistic flood depth, but is computationally intensive and not suitable for real-time trustworthiness assessment.

Roadway characteristics data

Roadway characteristics consist of geometric design features like number of lanes, per lane capacity, intersection (binary variable, based on if the report falls at an intersection or not), and freeway (binary variable, based on if the report falls on a freeway link or lower functional classification link). These roadway properties are obtained from the Hampton Roads Regional Travel Demand Model (HRRTDM). The roadway functional classification categorizes major and minor freeways as freeways (binary variable = 1), and all other roadway links as non-freeways (binary variable = 0). For capacity of each roadway link, per lane capacity is multiplied with number of lanes, both obtained from the HRRTDM. For roads that are not covered in the HRRTDM network, the number of lanes is obtained from the City of Norfolk's streets shapefile, and multiplied with a default per-lane capacity of 650 vehicles per hour per lane (minimum per lane capacity recorded in the HRRTDM).

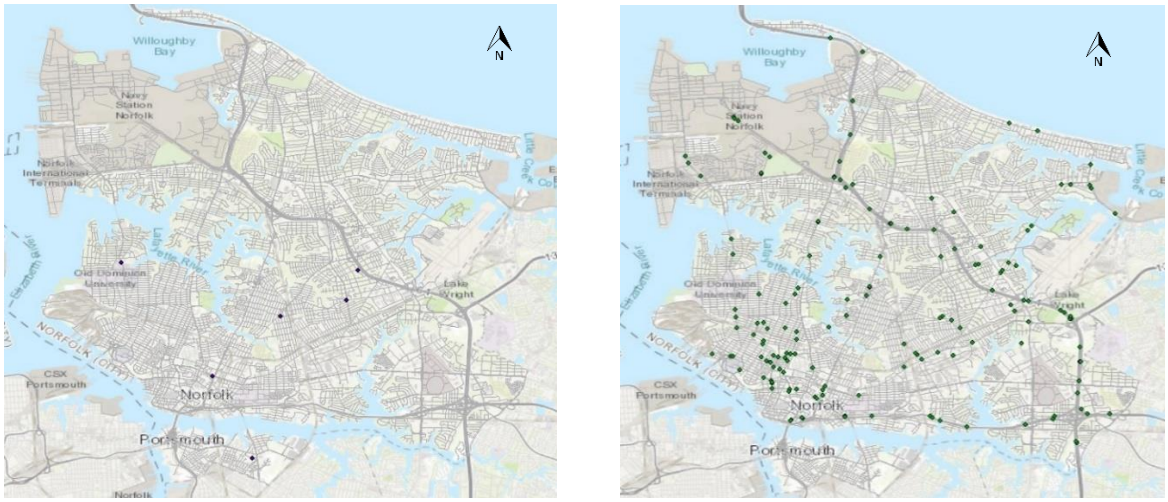
Agency-provided flood incident data

The flood incident data collected by City of Norfolk spanned from January 2017 to December 2018, using city employees' reported flood locations in a mobile phone application (System to Track, Organize, Record, and Map [STORM]). The app records the date and location of flooding. Furthermore, the app user can specify the flood location as an intersection, address, or block. Duplicate reports (reports occurring on the same day, within 50ft of each other) are eliminated. Then, several steps are followed in order to geo-tag the flood incident report to a specific location on the roadway network. In ArcGIS, the intersections named in the dataset are manually matched to the corresponding intersection, the addresses are relocated to the closest point on the roadway network, and the block locations are relocated to the lowest elevation point between the upstream and downstream intersections. These location-corrected reports are henceforth referred to as city reports. Due to the lack of a timestamp associated with the flood reports (only dates are recorded), the entire day is initially assumed to be flooded in this analysis. Then, these city reports are checked against TUFLOW model output to identify the flooded time periods with positive water depth, as explained in detail in the Methodology section.

Crowdsourced flood incident data

The mobile navigation application Waze contains a real-time information reporting tool, from which the crowdsourced flood incident reports are obtained. Waze provides user-reported incident data via its data sharing program (Waze for Cities), which is available to public entities worldwide. In this study, Waze incident reports (time-stamped and location-identified) related to roadway flooding in Norfolk (between August 2017 and December 2018, with three weeks excluded due to loss of data) are analyzed. Waze flood report data is only as comprehensive as the locations of road users reporting flooding; however, the spatial coverage is considerably greater than the agency data available through the City of Norfolk. Figure

1 shows an example of flood reports from both data sets in August 2018, to represent the coverage disparity between the City and Waze reporting of floods.



(a) (b)

Figure 1 Flooding reported by (a) City of Norfolk and (b) Waze in August 2018

Drainage characteristics data

A record of all the storm water structures (such as bridge drain, gutter basin, floor drain, manhole, etc.) is provided by the City of Norfolk in a GIS shapefile. This variable is regrouped from 18 structure types to 12 structure categories to combine similar drainage structure types on the roadway. Each flood report is characterized by the drainage infrastructure that is the closest feature by distance (in GIS) from the report location.

In addition to all the datasets described above, an additional parameter called "Proximity" is calculated. Proximity is a measure of closeness to other flood incident reports: the closer the other flood incident reports during the same time of day, the higher the proximity value. This score is assigned to each ground truth city flood report, and is calculated based on Waze- and city-reported flood incidents during the same time period (on the same date), as shown in **Equation 3**. Because Proximity can only be calculated for city reports when Waze flood report data is also available, the study period is defined as August 2017 to December 2018 (when both data sets overlap). Table 1 summarizes all the datasets used to build the ground truth model.

TABLE 1 Data inputs in predictive model

Variable (unit)	***Explanation***	***Data Dimension***	***Data Source or Method of Derivation***
Date	*Date of flood report*	*Temporal*	*City of Norfolk or Waze.*

Time period	*Time period of flood report* *(1: 12:00 to 6:00 am,* *2: 6:00 to 9:00 am,* *3: 9:00 am to 3:00 pm,* *4: 3:00 to 6:00 pm,* *5: 6:00 pm to 12:00 am)*	*Temporal*	*Timestamp obtained from Waze, and aggregated to corresponding time period.* *No timestamp for City reports, thus they are initially assumed to apply for all 5 time periods on the report day.*
Latitude and longitude	*Location of flood report.*	*Spatial*	*Obtained from geo-located City of Norfolk data or directly from Waze.*
Proximity score	*Measure of closeness to other flood reports on the same date and during the same time period. Calculated as the sum of the squares of inverse distances between the current flood report and all the other flood reports.*	*Spatial and temporal*	$Proximity_i = \sum_{j=1, j\neq i}^{J} \frac{1}{{d_{ij}}^2}$ *(3)* *Where:* *i: flood report for which proximity score is being assigned* *j: other flood reports (city and Waze) on the same date and during the same time period as i* d_{ij}*: bird's eye distance between coordinates of i and j, in miles*
Rainfall intensity (in/hr)	*Collected across seven rain gauges in the city, and interpolated for each flood report location.*	*Spatial and temporal*	*Data obtained from HRSD;* *Interpolation done by Inverse Distance Weighting (IDW), a spatial analysis tool in ArcGIS.*
Elevation (ft)	*Elevation of each flood report location.*	*Spatial*	*Obtained from DEM of Norfolk (from USGS), and extracted for each flood report point.*
Tide level (ft)	*Maximum tide level recorded during the given time period at gauge at Sewell's Point.*	*Temporal*	*Obtained from NOAA Tides and Currents*
Depth to water index (DTW)(m)	*Elevation difference between the flood report location and closest water body based on least slope path*	*Spatial*	*Created from rasters of DEM and waterbodies* ***(19)****, and extracted for each flood point*

Topographic Wetness Index (TWI)	*Measure of the tendency of an area to accumulate runoff: high TWI values imply a high potential for runoff accumulation*	*Spatial*	*Created from rasters of DEM and waterbodies **(18)**, and extracted for each flood point*
Intersection	*Binary variable to identify if the flood report occurs at an intersection*	*Spatial*	*Manually obtained from ArcGIS*
Total Capacity	*Total capacity of the roadway segment*	*Spatial*	*Obtained from HRRTDM* $Total\ Capacity = number\ of\ lanes \times per\ lane\ capacity$
Freeway	*Binary variable to identify if the flood report occurs on a freeway segment*	*Spatial*	*Functional classification obtained from HRRTDM*
Drainage	*Different types of drainage structures found closest to the report location*	*Spatial*	*Obtained from the City of Norfolk*

DATA PREPARATION & METHODOLOGY

Refining ground truth dataset

Due to the lack of a timestamp on city-reported flood events (and the initial assumption that the location is flooded for all five time periods of the day), the data undergoes another level of pre-processing before being included as a positive flood report in the ground truth dataset. The physics-based TUFLOW model is simulated on all the reported flood days to provide estimated water depth for each hour within the model boundary. The city report locations that are present within the TUFLOW model boundary are checked for maximum water depth within a 24ft buffer (width of the traveled way on a typical two-lane roadway) to ensure that time periods considered flooded have a predicted water depth greater than 0.1m ***(21)***. The TUFLOW model is not directly used to find the trustworthiness of Waze flood incident due to its computationally intensive nature, making such an approach infeasible for the end goal of using crowdsourced flood reports for real-time traffic management and flood mitigation. The TUFLOW model is used in this study purely as a filtering method to disaggregate city flood reports which only come with date information and no time stamp. In the 16-month study period, 19 days incurred city flood reporting, which translated into 95 initial positive flood observations across five time-of-day-periods. When verified against TUFLOW water depth models, 70 ground truth positive flood observations remained.

The ground truth dataset also requires negative (non-flood) observations. Since any location on the roadway network during any time period that is not reported as flooded could be a potential non-flood

observation, a procedure (outlined in Figure 2) was followed for random selection of negative observations at spatial and temporal scales.

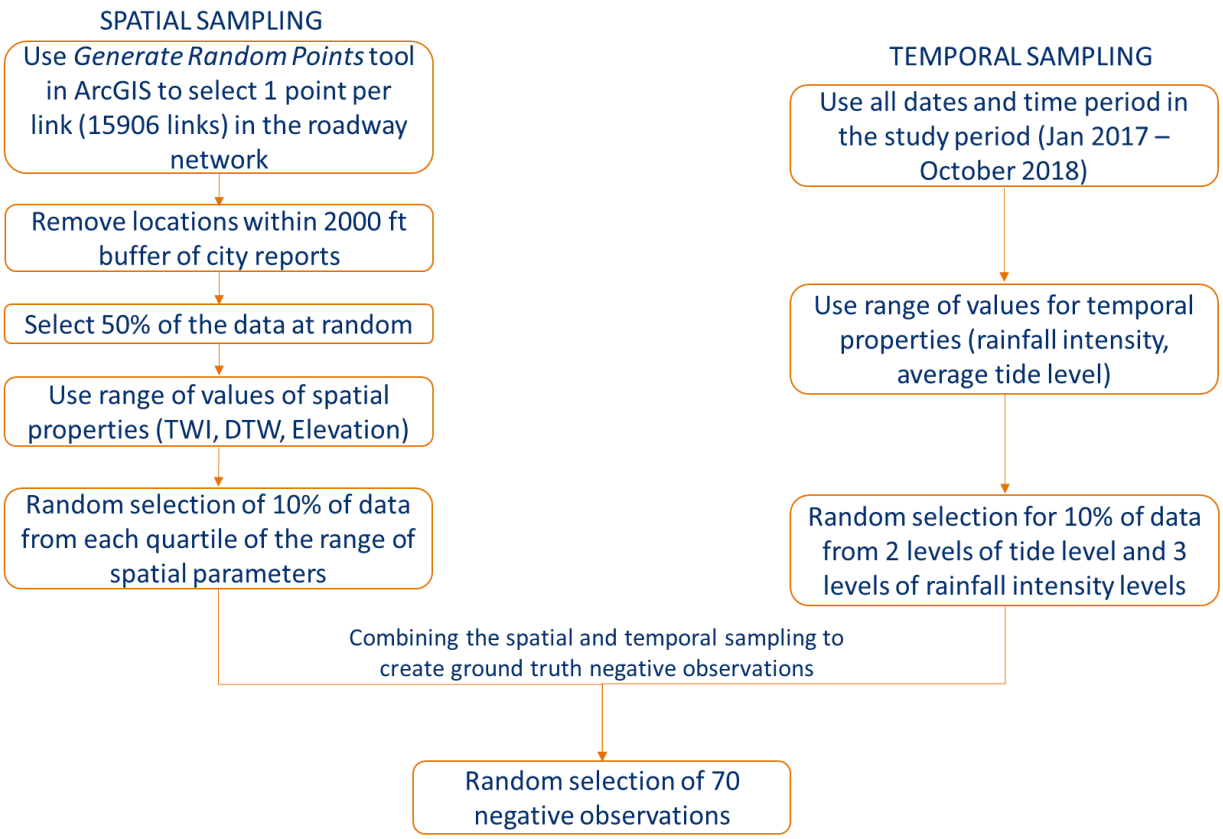


Figure 2 Combining spatial and temporal dimensions of negative ground truth observations

For the random sampling of negative observations, locations within a 2000ft buffer of the positive ground truth city flood reports were removed. Variable thresholds are also considered in the random selection process of negative ground truth data to ensure sufficient variation in random sampling. The spatial parameters (TWI, DTW, elevation) were divided into quartiles, based on their range of values, shown in Table 2. 10% of the data is randomly selected from each quartile to have some representation of each level of spatial characteristics. For temporal selection, two levels of tide (≤ 0 ft, 0 ft<) and three levels of rainfall (0, 0< and ≤0.1, 0.1< in/hr) were chosen to have a balanced sample of tide and rainfall levels. The non-flood observations were then sampled to fill a 1:1 of true-to-false flood report ratio in the ground truth dataset for the model. The total number of negative ground truth observations per variable threshold is shown in Table 2 for the balanced 1:1 dataset.

TABLE 2 Negative ground truth observations by environmental and topographic variable thresholds

DTW	Range (index)	DTW <1089	1089≤ DTW <2171	2171≤ DTW < 3260	DTW≥3260
	# of Observations (1:1)	44	21	4	1

TWI	Range (index)	TWI < 3.8	3.8 ≤ TWI <8.9	8.9 ≤ TWI <14	TWI ≥ 14
	# of Observations (1:1)	16	42	10	2
Elev-ation	Range (m)	Elev <-1.5m	-1.5m≤ Elev <3.3m	3.3m≤ Elev <8.2m	Elev≥8.2m
	# of Observations (1:1)	0	45	25	0
Rainfa ll	Range (in/hour)	Rain = 0	0< Rain ≤0.1 in/hr	Rain > 0.1 in/hr	
	# of Observations (1:1)	60	3	7	
Tide Level	Range (m)	Tide ≤ 0 m	Tide > 0 m		
	# of Observations (1:1)	30	40		

Trustworthiness modeling

Roadway flooding events are closely related to environmental, topographic, and infrastructure conditions (as listed in Table 1). Given the same probability of roadway flooding, the probability of the reporting of a roadway flood event is closely tied to the traffic volume on that roadway segment during the time period of the flood event (approximated here by the time of day and total capacity variables, as explained in Table 1). Then, all of these variables can be used to predict the trustworthiness of a crowdsourced flood report, estimated as the probability of a flood incident report at a location given the environmental, topographic, roadway, drainage, and time-of-day characteristics, via a binary logistic regression model. Logistic regression is widely used to study the effects of explanatory variables on binary outcomes, and the probability of the event occurring is calculated as shown in **Equation 4**:

$$P\ (1|X_1, X_2, \ldots\ldots X_n) = \frac{1}{1 + e^{-(\alpha + \sum \beta_i X_i)}} \qquad (4)$$

where,

P Probability of occurrence of the event (flood report)

i Observation

α Constant

X_i Independent variables (as listed in Table 1)

β_i Corresponding coefficient

A random 70-30 split of the 140 ground truth observations are used, where 70% of the entire dataset is reserved for training, and remainder 30% for testing the dataset.. Random sampling of data subsets is

performed on the training dataset to fit the samples into a prediction model, while reducing the total error in the model. Then, the regression model is used to calculate the probability of a Waze flood event explained by the independent variables.

Note that a classification tree model (which creates partitions in the dataset based on discrete characteristics) was also tested for this dataset. In the classification tree model, a parent node in the tree is divided based on an independent variable into two child nodes, such that each child node is more homogenous (or less impure) than the parent node. However, due to the large number of independent variables in the dataset, the classification tree model proved to be unstable (prediction accuracy and important variables varied widely when changing the observations in the training set). Hence, the logistic regression model is preferred for assessing trustworthiness.

Model Selection

Several different criteria are used to evaluate the fit of the various logistic regression models. These criteria and their definitions are listed in Table 3, including a confusion matrix and its associated criteria (true positive rate [TPR] (**Equation 5**), true negative rate [TNR] (**Equation 6**), false positive rate [FPR] (**Equation 7**), and false negative rate [FNR] (**Equation 8**)), Akaike Information Criterion (AIC), distance to corner (**Equation 9**), receiver-operating-characteristic (ROC) curves, and accuracy (**Equation 10**).

TABLE 3 Performance measures used for model selection

Performance Measure	***Definition***	***Preferred value directionality***	***Equation (if applicable)***
Akaike Information Criterion	*Estimator of out-of-sample prediction error, or the relative amount of information lost by a given model*	*Smaller*	
True Positive Rate (Sensitivity)	*Ratio of true positives identified to all positive ground truth reports*	*Higher*	$\frac{tp}{tp + fn}$ (5)
True Negative Rate (Specificity)	*Ratio of true negatives identified to all ground truth negatives reports*	*Higher*	$\frac{tn}{fp + tn}$ (6)

False Positive Rate	*Ratio of false positives identified by the model to all truly negative reports in the ground truth*	*Lower*	$\frac{fp}{fp + tn}$ (7)
False Negative Rate	*Ratio of false negatives identified by the model to all truly positive reports in the ground truth*	*Lower*	$\frac{fn}{tp + fn}$ (8)
Receiver Operating Characteristic (ROC) Curves ***(23)***	*Plot of Sensitivity vs 1-Specificity, helps to determine the diagnostic ability of the binary classifiers*	*Closer to the top left corner*	
Distance to Corner ***(24, 25)***	*Optimal threshold to minimize false positive and false negative rates*	*Lower*	$Distance\ to\ Corner = \sqrt{(1 - Sensitivity)^2 + (1 - Specificity}$
Accuracy	*Shows the ratio of correctly identified reports to all reports.*	*Higher*	$\frac{tp + tn}{all\ reports}$ (10)

RESULTS

Trustworthiness Model

A prediction model using the binary logistic regression structure was developed to analyze how well a ground truth report can be explained by the independent environmental, topographic, infrastructure, and temporal variables. The ground truth dataset was randomly split into 70% for model training, and the remainder 30% for model testing. First, all the continuous variables considered (as listed in Table 1) were tested for correlation. The highest correlation was found between elevation and TWI (-0.52), with all other correlation values less than 0.3. The correlation between elevation and TWI suggests that areas with lower elevation have a higher tendency to accumulate runoff (TWI), which is related to topography. Then, different binary logistic regression models with and without TWI were tested to examine the effect of inclusion of TWI on the tide level parameter estimate and overall model fit. Due to the relatively small

ground truth report sample size, explanatory variables with p-values less than 0.3 (confidence level 70%) are retained. Among the different model specifications tested, elevation, TWI, and DTW are always found to be statistically significant at 95% confidence level. When roadway characteristics are considered, total capacity was the only variable to emerge as statistically significant ($p<0.1$ in all three different model specifications). However, inclusion of the total capacity variable reduced the accuracy of the model (from 90.5% to 88%) with a higher false positive rate in the model prediction. Thus, roadway capacity is excluded from the final preferred model. Tide level and rainfall are marginally significant ($0.25 <$ p-value < 0.3) in the preferred model specification. These variables are retained due to their effect on the discrete time period variables (which emerge as non-significant if tide level and rainfall are excluded). Despite the proximity variable being a byproduct of the number and closeness of other reports in the same time period, it shows a low correlation with tide level and rainfall. Proximity variable is retained in the preferred model due to its high statistical significance. The final preferred model is shown in Table 4.

TABLE 4 Preferred binary logistic regression model results

Variables	*Estimate*	*Std. Error*	*z value*	*Pr(>|z|)*
(Intercept)	*13.54*	*5.26*	*2.58*	*0.01(*)*
Elevation	*-4.99*	*1.60*	*-3.12*	*0.00(**)*
Tide level	*3.35*	*3.05*	*1.10*	*0.27*
Rainfall	*6.88*	*6.54*	*1.05*	*0.29*
Period = 2 (6a – 9a)	*-5.01*	*2.61*	*-1.92*	*0.05 (.)*
Period = 3 (9a – 3p)	*-5.45*	*3.18*	*-1.71*	*0.09 (.)*
Period = 4 (3p – 6p)	*-3.47*	*2.58*	*-1.34*	*0.18 (.)*
Period = 5 (6p-12a)	*-6.24*	*2.95*	*-2.11*	*0.03(*)*
Proximity	*1.30*	*0.59*	*2.22*	*0.03(*)*
TWI	*-0.52*	*0.24*	*-2.13*	*0.03(*)*
DTW	*0.00*	*0.00*	*2.31*	*0.02(*)*

**: significant at 99% confidence level

* : significant at 95% confidence level

. : significant at 90% confidence level

In the final preferred model, all time periods are statistically significant at the 90% confidence level. In the absence of roadway characteristics in the final model, time period can be considered as a proxy exposure variable (as traffic volumes, and thereby active Waze users, are highly correlated to time-of-day). Proximity is also highly statistically significant in the preferred model. This indicates that the presence of other flood reports during the same time period (and physical proximity to those peer reports)

is an important predictor variable. The Root Mean Square Error (RMSE) values for testing and training subsets of the ground truth datasets were found to be 0.27 and 0.29 respectively, with AIC value of 46.15.

The preferred binary logistic model is then used to find an optimum threshold to separate trustworthy and untrustworthy reports. To do this, TPR and TNR are calculated at different thresholds, as shown in Table 5. In Figure 2, additional performance measures are considered for these thresholds as *1 - sensitivity* and *1 – specificity* are plotted to display Receiver Operating Characteristic (ROC) curves ***(23***). ROC curves help to determine the diagnostic ability of binary classifiers. An optimal threshold is defined as a point on the ROC curve with the value of *1 - sensitivity* and *1 - specificity* closest to 0 (i.e. when the FPR and the FNR are the lowest). Thus, the point on the curve with the shortest distance to the top-left corner of the plot, also known as distance-to-corner, corresponding to a threshold value of 0.8, is chosen as the threshold to differentiate trustworthy from untrustworthy crowdsourced reports. The corresponding confusion matrix for a trustworthiness threshold of 0.8 is shown in Table 6, which yields a model accuracy of 90.48% on the testing dataset.

TABLE 5 Performance measures of varying thresholds

Threshold Value	***0.60***	***0.70***	***0.80***	***0.85***	***0.90***	***0.95***
TPR (Sensitivity)	*0.80*	*0.80*	*0.80*	*0.80*	*0.80*	*0.75*
TNR (Specificity)	*0.91*	*0.95*	*1.00*	*1.00*	*1.00*	*1.00*
Distance to corner squared ($fpr^2 + fnr^2$)	*0.0206*	*0.0055*	*0.0000*	*0.0000*	*0.0000*	*0.0000*
Accuracy	*85.71%*	*88.10%*	*90.48%*	*90.48%*	*90.48%*	*88.10%*

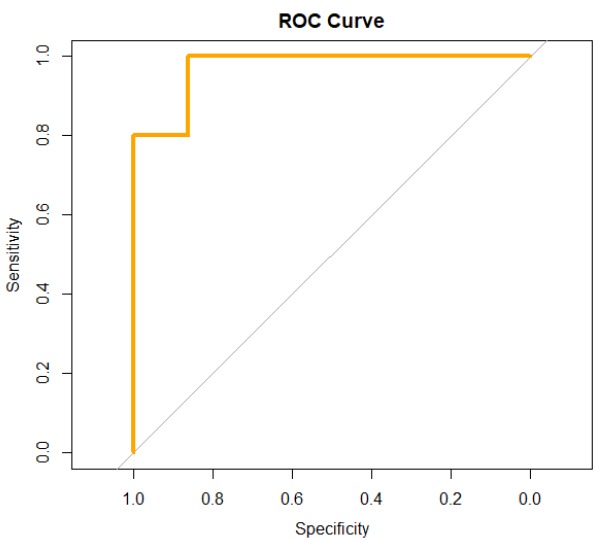


Figure 2 ROC Curve

TABLE 6 Confusion matrix (threshold = 0.8)

N = 42	*Predicted True*	*Predicted False*
True Report	*16 (tp)*	*4 (fn)*
False Report	*0 (fp)*	*22 (tn)*

Trustworthiness of Waze data

Figure 3 shows the predicted probability of occurrence of all crowdsourced Waze flood reports between August 2017 and December 2018 in Norfolk, when the preferred trustworthiness model is applied. 502 of the 697 reports exceed the threshold value of 0.8, implying 71.7% of the Waze flood reports can be considered trustworthy based on their topographic, environmental, temporal, and peer reporting characteristics.

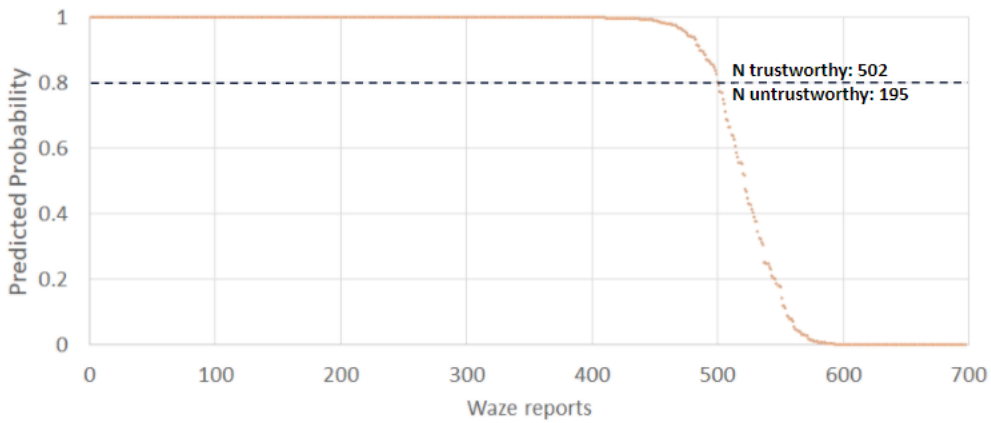


Figure 3 Waze report incident occurrence probabilities

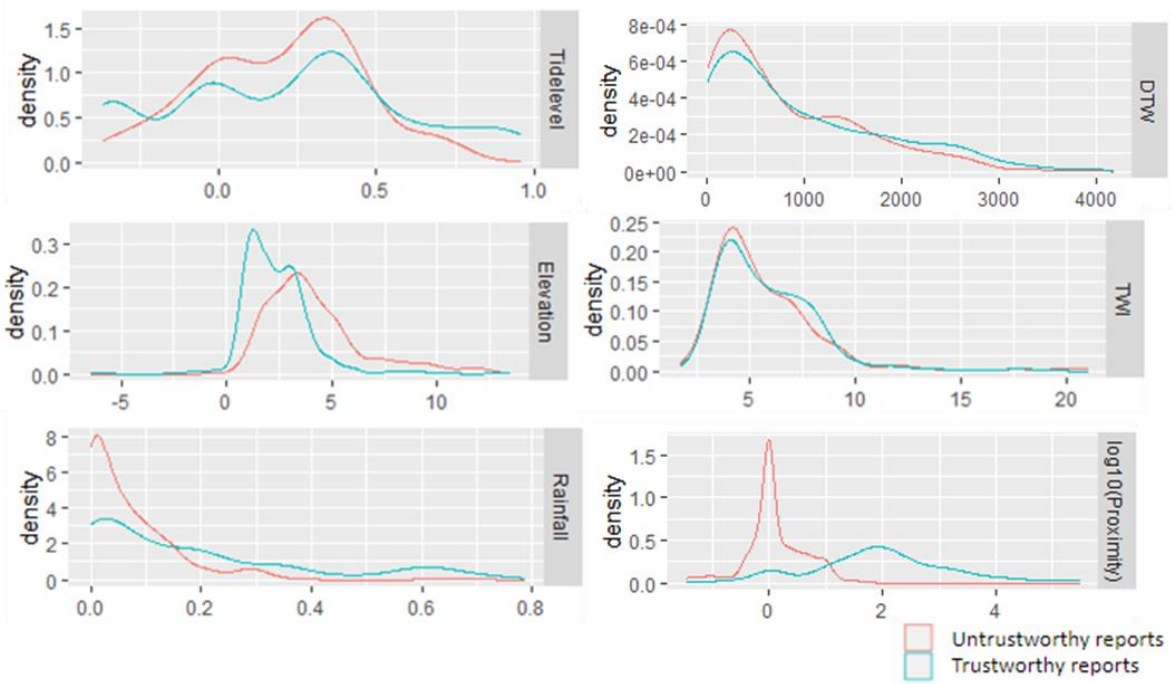
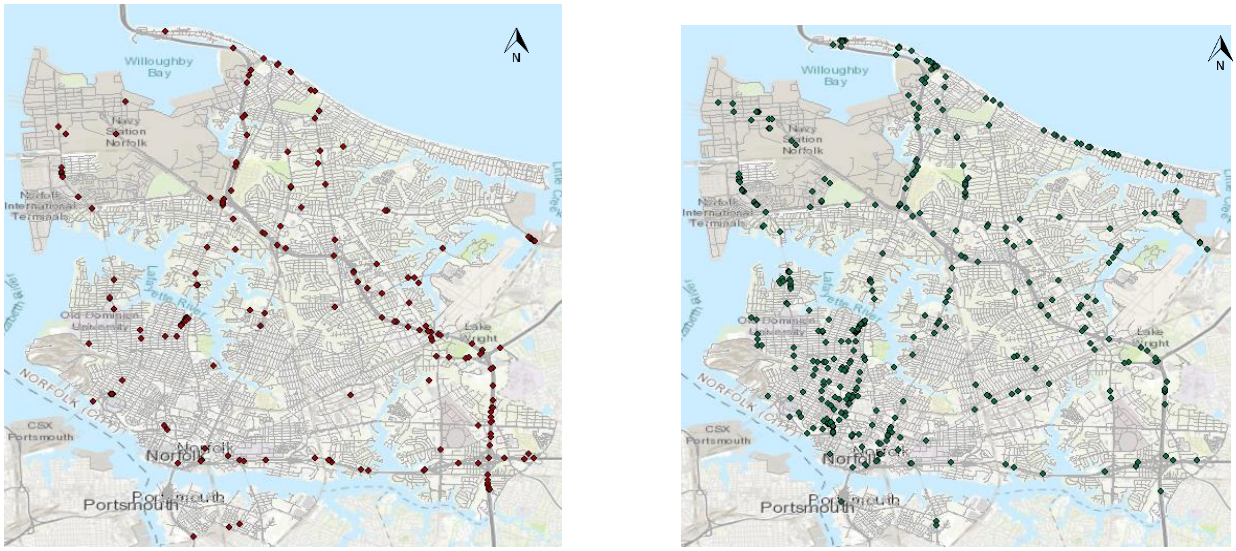


Figure 4 Characteristics of trustworthy and untrustworthy reports

Density plots in Figure 4 show the distribution of independent variable values of trustworthy (cyan) and untrustworthy (red) reports. In terms of environmental conditions, untrustworthy flood incident reports tend to cluster around 0 rainfall. Furthermore, flood incident reports are more likely to be considered trustworthy in higher intensity rainfall periods. For tide level, untrustworthy reports generally follow the same density plot as trustworthy reports, with two high density spikes at 0 ft and 0.4 ft. Tide level is not a spatially disaggregate variable as it is only collected at one point in the city. However, plotting these report locations on a map, there is a cluster of trustworthy reports (Figure 5[a]) near the Chesapeake Bay coast line in the northern part of the city, and another cluster of reports near downtown Norfolk (a low elevation area) in the southwest part of the city. Untrustworthy reports (Figure 5[b]), on the other hand, show far fewer reports in these locations. Similar to rainfall, untrustworthy reports show a low density of occurrence at higher tide levels (greater than 0.5 ft). In terms of the topographical variables, trustworthy reports generally occur at lower elevation compared to untrustworthy reports. When considering TWI (tendency of a location to accumulate runoff), the distributions are very similar for trustworthy and untrustworthy reports. On the other hand, trustworthy reports are more likely to report higher DTW (proximity to the closest water body) values, with the reports at the highest DTW values related to intense rainfall events. This implies that despite being at locations with high DTW, flooding does occur in instances of heavy rain. On the other hand, when examining untrustworthy reports with high DTW values, they occur during time periods with no rainfall. The proximity score of reports, which is based on the quantity and proximity of peer flood incident reports, has a large range. Trustworthy reports' mean proximity value is 6077.32 (log value 2.08) and untrustworthy reports' mean proximity value is 2.15 (log value ~ 0), implying that the majority of untrustworthy reports are either sole reports during the time-of-day period, or had peer reports very far away, resulting in a very low proximity score.



(a) Untrustworthy (b) Trustworthy

Figure 5 Spatial distribution of untrustworthy and trustworthy reports

Table 7 examines the characteristics of peer flood reports for trustworthy and trustworthy flood incident reports. Untrustworthy reports are more temporally dispersed (195 reports over 117 unique time-of-day periods) than trustworthy reports (502 reports over 100 unique time-of-day periods). Additionally, the maximum number of peer flood reports in the same time-of-day period was found to be 6 for untrustworthy reports and 66 in trustworthy reports, implying the importance of peer reports in asserting trustworthiness in crowdsourced data.

TABLE 7 Observations on trustworthy and untrustworthy time periods from Waze

# of reports	*Trustworthy Reports*	*Untrustworthy Reports*
Total reports	*502*	*195*
Total affected time periods	*100*	*117*
Sole flood report (in the time period)	*8.0%*	*39.4%*
1 peer flood report (in same time period)	*8.8%*	*20.5%*
>1 peer flood report (in same time period)	*83.2%*	*40.1%*
Maximum number of peer reports in a single time period	*66*	*6*

The distinguishable characteristics in trustworthy reports were found to be high rates of peer reporting, lower elevations, higher rainfall intensity, and proximity to the coast. These characteristics are intuitive for true flood events, but this study provides a trustworthiness cutoff that can be useful in helping decision makers quickly filter high-confidence crowdsourced reports from low-confidence ones, without requiring A the deployment of extensive resources for incident reporting. Accessing trustworthy information in near-real time can significantly improve response times for local agencies to delegate emergency management services, thereby mitigating flooding-related disruptions efficiently.

CONCLUSIONS AND LIMITATIONS

Crowdsourced data has the potential to provide real-time transportation information without the cost of additional sensors, cameras, or other cost-prohibitive measures. However, since crowdsourced data is usually unchecked, verification of the data becomes challenging. This study presents a framework to assess the quality of a subset of Waze incident reporting data related to roadway flooding in Norfolk, Virginia. Roadway flooding occurs as a combination of environmental conditions and insufficient drainage infrastructure. Environmental, topographic, and infrastructure variables which potentially contribute to flooding and flood reporting are used in this study to build a logistic regression model to estimate the probability of occurrence of a flood incident report. While this methodology does not directly identify misreports or false reports, it provides a conservative approach to distinguish crowdsourced flood incident reports with a high level of trustworthiness. The preferred model developed in this study shows a prediction accuracy of 90.48% when applied to a subset of ground truth data, implying a high rate of correct identification of reports. When applying the model to crowdsourced Waze data over a 16-month study period, 71.7% of the user reported flood incidents were predicted to be trustworthy. Among the untrustworthy reports, the most notable characteristics included low occurrence of peer reporting, inland locations with lower tide levels, higher elevation, and lower rainfall intensity in the reported periods.

This study has limitations which should be addressed in future research. To start, the positive ground truth data set utilized in model development has a small sample size, and is biased towards higher intensity flood events. Within these events, the ground truth data is biased towards tidal flooding compared to rainfall-induced flooding, due to the nature of city employees' data collection. In addition, the proximity variable defined in this study to account for peer reporting requires that crowdsourced datasets have high levels of user activity (and incident reporting), in order to achieve a wide range of proximity values. Furthermore, one of the assumptions in the study was a default minimum capacity on smaller roads. While the capacity variable was found insignificant in the regression model (and thus removed from the finalized model), it remains an important feature of the roadway network, and future studies should consider acquiring more detailed data on roadway capacity. Lastly, ground truth data (however limited) is still necessary to build a trustworthiness assessment model under this framework. A search for agency provided flood incident data used in recent research only yielded a handful of cities in the US (including New York City, NY; Norfolk, VA; Charleston, SC; Miami, FL; Houston, TX; San Francisco, CA; and Tacoma, WA). This becomes a challenge in transferring the model framework to other coastal cities without agency flood incident data. Given the increasing prevalence of crowdsourced incident data, coastal cities should consider investing in local ground truth data collection in limited areas during the periods in the year with heavy floods. Alternately, cities could also invest in strategically placing sensors

throughout the city for a short period of time to collect ground truth data. For cities that lack resources to collect locally relevant ground truth flood incident data, this research suggests that static variables such as topographic and roadway variables may also be used for initial screening of flood hotspot locations. The relative amount of peer reporting in crowdsourced datasets can then be used for assigning priority to potentially high-confidence locations.

Nonetheless, this study demonstrates the ability to assess trustworthiness of Waze flood incident reports with limited ground truth availability. This framework could eventually lead to identification of flooding hotspots in near-real time, allowing cities to deploy dynamic flood mitigation actions and ensure a faster recovery to normal conditions. The flooding hotspots identified through this methodology can be used to provide early improvements in addressing the long-term impacts of sea-level rise. Furthermore, the general methodology utilized in this study is not limited to assertion of trustworthiness of crowdsourced flood incidents, and can be used in other applications with available contextual and ground truth data sets.

ACKNOWLEDGEMENTS

The authors would also like to thank City of Norfolk and Waze for facilitating data acquisition. The authors would also like to thank Erin Robartes for giving valuable input in reviewing the manuscript. This work is supported by the National Science Foundation's Critical Resilient Interdependent Infrastructure Systems and Processes program (Award 1735587).

AUTHOR CONTRIBUTIONS

The authors confirm contribution to the paper as follows: study conception and design: S. Praharaj, T.D. Chen, J. L. Goodall; data collection and processing: Y. Shen, S. Praharaj, F. Zahura; analysis and interpretation of results: S. Praharaj, T.D. Chen, L. Zeng; draft manuscript preparation: S. Praharaj, F. Zahura, T.D. Chen. All authors reviewed the results and approved the final version of the manuscript.

REFERENCES

1. Medury, A., O. Grembek, A. Loukaitou-Sideris, and K. Shafizadeh. Investigating the underreporting of pedestrian and bicycle crashes in and around university campuses a crowdsourcing approach. *Accident Analysis & Prevention*, Volume: 130, 2019, pp. 99 – 107.

2. Nair, D.J., F. Gilles, S. Chand, N. Saxena, and V. Dixit. Characterizing multicity urban traffic conditions using crowdsourced data. PLoS ONE, 2019. Volume 14(3). https://doi.org/10.1371/journal.pone.0212845.

3. Flynn, D. F., M. M. Gilmore, and E. A. Sudderth. *Estimating Traffic Crash Counts Using Crowdsourced Data: Pilot analysis of 2017 Waze data and Police Accident Reports in Maryland*, 2018, Tech Report.

4. US EPA. *Climate Change Indicators in the United States: Coastal Flooding*, United States Environmental Protection Agency, 2016. www.epa.gov/climate-indicators. Accessed May 10, 2018.

5. Sweet, W., J. Park, J. Marra, C. Zervas, and S. Gill. *Sea Level Rise and Nuisance Flood Frequency Changes Around the United States.* Publication NOAA Technical Report NOS CO-OPS 073. National Oceanic and Atmospheric Administration, U.S. Department of Commerce, 2016.

6. Sanchez, L., E. Rosas, and N. Hidalgo. Crowdsourcing Under Attack: Detecting Malicious Behaviors in Waze. In *Trust Management XII* (N. Gal-Oz and P. R. Lewis, eds.), Springer International Publishing, Cham, 2018, pp. 91–106.

7. Hoseinzadeh, N., Y. Liu, L. D. Han, C. Brakewood, and A. Mohammadnazar. Quality of location-based crowdsourced speed data on surface streets: A case study of Waze and Bluetooth speed data in Sevierville, TN. *Computers, Environment and Urban Systems*, 2020. Volume 83.

8. Amin-Naseri, M., P. Chakraborty, A. Sharma, S. B. Gilbert, and M. Hong. Evaluating the Reliability, Coverage, and Added Value of Crowdsourced Traffic Incident Reports from Waze. *Transportation Research Record*, 2018. Volume: 2672 (43), pp. 34–43.

9. Goodall, N., and E. Lee. Comparison of Waze crash and disabled vehicle records with video ground truth. *Transportation Research Interdisciplinary Perspectives*, 2019. Volume: 1.

10. Eriksson, I. Towards Integrating Crowdsourced and Official Traffic Data: A study on the integration of data from Waze in traffic management in Stockholm, Sweden, 2019.

11. Ouyang, R. W., M. Srivastava, A. Toniolo, and T. J. Norman. Truth Discovery in Crowdsourced Detection of Spatial Events. *IEEE Transactions on Knowledge and Data Engineering*, 2016. Volume: 28 (4), pp. 1047–1060.

12. Dong, X. L., L. Berti-Equille, and D. Srivastava. Truth Discovery and Copying Detection in a Dynamic World. *Proc. VLDB Endow*, 2009. Volume: 2 (1), pp. 562–573.

13. Prandi, C., S. Ferretti, S. Mirri, and P. Salomoni. Trustworthiness in crowd- sensed and sourced georeferenced data. *IEEE International Conference on Pervasive Computing and Communication Workshops (PerCom Workshops)*, 2015. pp. 402-407, doi: 10.1109/PERCOMW.2015.7134071.

14. Flanagin, A. J. and M. J. Metzger. The credibility of volunteered geographic information. *GeoJournal*, 2008. Volume: 72 (3), pp. 137–148.

15. Ostermann, F. and L. Spinsanti. Context analysis of volunteered geographic information from social media networks to support disaster management: a case study on forest fires. *International journal of information systems for crisis response and management*, 2012. Volume: 4 (4), pp. 16–37.

16. Hung, K.-C., M. Kalantari, and A. Rajabifard. Methods for assessing the credibility of volunteered geographic information in flood response: A case study in Brisbane, Australia. *Applied Geography*, 2016. Volume 68, pp. 37 – 47.

17. NOAA, 2018b. Sewells Point - Station Home Page - NOAA Tides & Currents [WWW Document]. URL https://tidesandcurrents.noaa.gov/waterlevels.html?id=8638610 (accessed 10.1.19).

18. Beven, K. J. and M. J. Kirkby, A physically based, variable contributing area model of basin hydrology. *Hydrological Sciences Bulletin*, 1979. Volume: 24 (1), pp. 43–69.

19. Murphy, P. N. C., J. Ogilvie, K. Connor, and P. A. Arp. Mapping wetlands: A comparison of two different approaches for New Brunswick, Canada. *Wetlands*, 2007. Volume: 27 (4), pp. 846–854.

20. Syme, W. J. TUFLOW - Two one-dimensional Unsteady FLOW Software for Rivers, Estuaries and Coastal Waters, 2001.

21. Shen, Y., M. M. Morsy, C. Huxley, N. Tahvildari, and J. L. Goodall. Flood risk assessment and increased resilience for coastal urban watersheds under the combined impact of storm tide and heavy rainfall. *Journal of Hydrology*, 2019. Volume: 579.

22. Shen, Y. Flood Risk Assessment and Increased Flood Resilience for Civil Infrastructure in Coastal Regions Under Changing Climate. University of Virginia, 2019.

23. Zou, K. H., A. J. O'Malley, and L. Mauri. Receiver-Operating Characteristic Analysis for Evaluating Diagnostic Tests and Predictive Models. *Circulation*, 2007. Volume 115 (5), pp. 654–657.

24. Berrar, D. Performance Measures for Binary Classification. *Academic Press*, Oxford, 2019, pp. 546 – 560.

25. Krzanowski, W. J., and D. J. Hand. *ROC Curves for Continuous Data.* CRC Press, 2009.

26. Shen, Y. *Flood Risk Assessment and Increased Flood Resilience for Civil Infrastructure in Coastal Regions Under Changing Climate.* Doctoral Dissertation. University of Virginia, Charlottesville, VA, 2020.

27. United State Geological Survey (USGS). What is the vertical accuracy of the 3D Elevation Program (3DEP) DEMs? *United States Geological Survey (USGS)*. https://www.usgs.gov/faqs/what-vertical-accuracy-3d-elevation-program-3dep-dems?qt-news_science_products=0#qt-news_science_products. Accessed May 27, 2021.

28. Zahura, F. T., Goodall, J. L., Sadler, J. M., Shen, Y., Morsy, M. M., & Behl, M. (2020). Training Machine Learning Surrogate Models from a High-Fidelity Physics-Based Model: Application for Real-Time Street-Scale Flood Prediction in an Urban Coastal Community. *Water Resources Research*, 2020. Volume 56(10).

CHAPTER 4

Predicting traffic impacts due to recurrent flooding using crowdsourced data: A Norfolk, Virginia Case Study

Shraddha Praharaj, T. Donna Chen*, Luwei Zeng, Faria T. Zahura, Jonathan L. Goodall

Journal Name - August 2021

To be submitted at *Transportation Research Board 101st Annual* Meeting, January 2022.

ABSTRACT

Recurrent flooding is becoming an increasing cause of concern in coastal cities as climate change and sea level rise exacerbates its frequency and intensity. Recurrent flooding is a dynamic disruption, with varying impacts across a city based on the location and time of day. Since recurrent flooding typically lasts for a few hours, obtaining traffic impact information in near real-time is crucial for emergency management services. However, flood incident data is sparsely collected by state and local agencies (if at all), thus motivating the use of crowdsourced data. In this study, we explore the use of various types of reliable crowdsourced data that can provide us this information. With reliable traffic volume and high confidence flood incident data available, we train various machine learning models that use environmental, hydrological, and roadway characteristics to predict traffic impacts at trustworthy flood incident report locations. We compare different machine-learning models for this task, and the extreme gradient boosting model was found to be the preferred model with the least normalized root-mean-square-error. The model predicts 67% of the reports within the one-standard deviation of the observed values. The most prominent features from this model are annual average daily traffic (AADT), tide level and water depth. This study demonstrates a framework, which with more data has the potential of using crowdsourced data to classify locations that are likely to be highly impacted by recurrent flooding on the roadway network, a critical tool for emergency management, flood mitigation, and dynamic traffic management.

Keywords: crowdsourced data, flooding, machine learning models, emergency management

INTRODUCTION

Climate change and sea-level rise are causing coastal cities to confront exponentially increasing recurrent flooding in the past few decades *(Koetse & Rietveld 2009)*. This type of flooding is typical to coastal cities with low elevation and is traditionally considered a minor disruption that affects the traffic operations for short periods *(Chen et al., 2012)*. However, these recurrent flooding events have increased in frequency and severity. Recurring flooding can cause significant travel delays, even if they are short-term events *(Jenelius & Mattsson, 2012)*, depending on the location and time of occurrence. These disruption events obstruct the traffic flow and reduce accessibility to sub-networks, rendering the roadway network vulnerable. Under a major disruption event (such as a hurricane), city and region-wide impacts could last for several days, thus the traditionally defined static vulnerability metrics for the overall stability of the roadway network can be utilized *(Sohn et al., 2006, Yang et al. 2009)*. With recurrent flooding, the vulnerability of a location on the roadway network is not static, and can vary based on spatial, temporal, contextual, or geographical factors. Each of these categories has multiple independent factors, thereby making vulnerability a dynamic component of the system. Furthermore, with traffic flow also being dynamic spatially and temporally, mitigating the effects of recurrent flooding necessitates a space and time-dependent vulnerability metric which can cater to near-real-time decision-making efforts.

However, despite recurrent flooding's growing impact on coastal cities, these events often go unreported by local government agencies as data collection can be time and labor-intensive, and may require significant financial effort. Crowdsourced data can be a powerful alternative to revolutionize data collection where agency data is not available, with road users able to report a wide variety of incidents as they travel throughout the roadway network. In this study, we use such empirical roadway and topographic data to characterize the change in traffic volumes due to recurrent flooding, which relates to the vulnerability of the network. Subsequently, using topographical, roadway, and environmental characteristics, we build a machine learning model to estimate the traffic impact vulnerability of other locations which do not have sufficient empirical data. This study illustrates a framework for assessing dynamic vulnerability using empirical crowdsourced traffic and flood incident data, creating a more flexible measure for response to recurrent flooding compared to the traditional static vulnerability measures that come from simulation-based studies.

LITERATURE REVIEW

The vulnerability of a network is defined by Berdica et al. *(2002)* as the susceptibility of a transportation system to react to different changes. In the transportation system, the concept of vulnerability is closely attached to accessibility, as accessibility depends on the extent to which the transportation system is functioning *(Berdica, 2002)*. Vulnerability has also been defined differently in specific research areas. For example, Berdica *(2002)*, Taylor *(2017)* and Chen et al. *(2015)* have characterized vulnerability as the reduction in performance of transportation system infrastructure, considering accessibility and serviceability. Chen et al. *(2015)* considered vulnerability as a function of the probability of disruption events happening and the subsequent traffic impacts on the network. Jun-qiang et al. *(2017)* defined vulnerability as the reflection of sensitivity of traffic operation performance to disturbance.

Among the vulnerability literature, two categories emerge: static and dynamic vulnerability. Studies in the static vulnerability subset assess vulnerability of the entire roadway network using network topology parameters and network theory-based measures. For example, Jun-qiang et al. *(2017)* developed a traffic utility index, where vulnerability was calculated based on an even reduction of capacity on the roadway segments across the network. The study used sensitivity analysis to calculate the change in the traffic utility index by degradation of capacity. Similarly, Yang et al. *(2013)* calculated the vulnerability

of the roadway network by using a combination of several evaluation indices such as their topological properties and an emergency plan as a qualitative parameter. There are several other similar studies in the static vulnerability area of research *(Sohn et al., 2006, Yang et al. 2009, Lopez et al. 2017, Haghighi et al., 2018 etc.)*. On the other hand, studies in dynamic vulnerability subset assess vulnerability based on traffic flow which vary within the network. These assessments can be done by either using simulated data or by using crowdsourced empirical data. A large chunk of dynamic vulnerability studies have used simulated data for assessment. Appert & Laurent *(2007)* developed a vulnerability index from network theory measures by removing links and vertices sequentially, and simulating the traffic flow on a roadway network. Distinct hourly flows were assigned to the links and the vulnerability of the nodes were measured. Kim and Yeo *(2016)* developed a flow-based vulnerability measure, which can show the traveler response to disruptions by simulating different disabled links. This measure uses different simulated scenarios to rank vulnerability of different links in the roadway network. Shekar et al. *(2017)* uses traffic simulation methods to assess dynamic vulnerability of the road network, which can prioritize the critical links over time and is generalizable to the case where both link and node disruptions are of concern. Some other similar studies in the dynamic vulnerability domain include *(Lu & Peng, 2011, Jenelius & Mattsson, 2012, Rodriguez- Nunez & Palomares, 2014* etc.).

In the domain of flood-related disruptions, studies have used a combination of static and dynamic vulnerability with simulated data to assess the system performance. These studies usually also include a hydrological modeling component. Chen et al. *(2015)* developed an accessibility-based measure to evaluate transportation network vulnerability under flooding impacts for different modes of travel. The vulnerability calculation is based on the change in traffic flow with completely inundated segments in the roadway network. A study by Singh et al. *(2018)* combined the result of flood modeling simulation scenarios with a function that relates safety speed to water depth to generate a vulnerability index. They found that there was a significant decrease in average maximum speed in each roadway functional class during disruption conditions compared to normal conditions. Similarly, Pyatkova et al *(2015)* combines a flood model and a traffic model to simulate traffic flow on inundated roads. The vulnerability is calculated as loss of business hours, additional fuel consumption, and additional CO_2 emission. Kermanshah et al. *(2017)* developed a vulnerability assessment framework with a combination of climate models, network theory, and stochastic modeling. Stochastic modeling was used for extreme rainfall induced flash floods to generate the probability of failure of links in the network. Vulnerability of the network was estimated based on the change in traffic flow.

All the studies mentioned above either use static or dynamic vulnerability with simulated models for estimating impacts on the transportation network. This is a major drawback in recurrent flooding research due to the nature of disruption. Since recurrent flooding is a minor disruption that occurs for a few hours, there is a need for near-real time detection to reduce significant disruptions in traffic flow. This paper considers a combination of static characteristics with dynamic empirical data (instead of simulated data) to assess vulnerability, which can provide insight about the observed vulnerability of the sub-areas in the roadway network. This is similar to the vulnerability concept explained by Berdica *(2002)*, but focuses on the susceptibility to weather disruptions that can result in reductions in road network performance. We assess vulnerability by using empirical data in the form of observed traffic flow and observed flood incidents and express it as traffic impact index. We then build a prediction model with a combination of empirical, static and estimated data which can estimate traffic impact for locations in the roadway network that may not have sufficient empirical data. This study provides a framework to predict traffic impact index theoretically in all locations of the roadway network. This can enable identification of potential high impact locations under forecasted hydrological and transportation network conditions, that has the potential to be serving as a dynamic planning tool for city agencies.

DATA

In this paper, vulnerability is defined as the impact of recurrent flooding on traffic flow, coined as *traffic impact index*. A relationship between the traffic impact index and the explanatory variables is established, which include roadway, environmental and hydrological characteristics, using data from the city of Norfolk, Virginia. With this predictive model, traffic impact indices can be estimated for locations with spatial or temporal gaps in data, or for understanding potential traffic impacts in future scenarios. Figure 1 shows the analysis area and the flood incident locations that were assessed in this study.

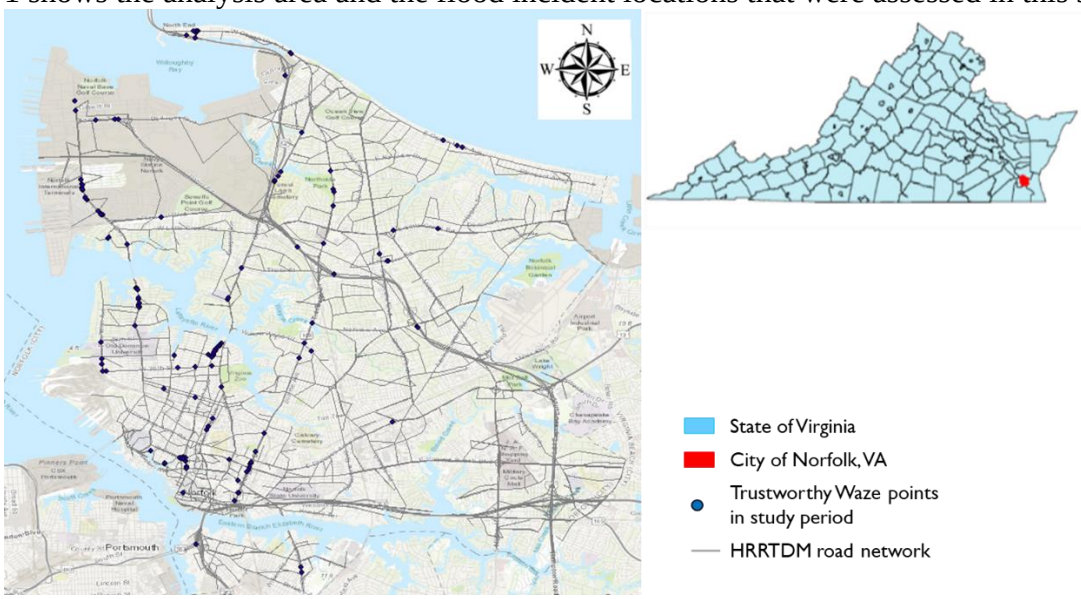


Fig 1 Study area and flood incident data

Crowdsourced flood incident data

The mobile navigation application Waze contains a real-time information reporting tool, from which the crowdsourced flood incident reports are obtained. Waze provides user-reported incident data via its data-sharing program (Waze for Cities), which is available to public entities worldwide. The mobile navigation application Waze collects flood incident reports via its real-time information reporting tool. The application provides user-reported incident data via its Waze for Cities data-sharing program, open to public entities worldwide.

Since the incidents can be reported by any user and are not regulated, it is essential to filter out potentially erroneous data. A methodology to assess the trustworthiness of Waze data for flood-related events is established by Praharaj et al. *(2021b)*. The dataset used in this study was obtained between August 2017 to December 2018 and contains 188 observations, which are the number of high-trust flood incident reports *(Praharaj et al., 2021b)*.

Network characteristics

Several road network characteristics are also considered as explanatory variables in this model. Some of these characteristics are obtained from the Hampton Roads Region Travel Demand Model (HRRTDM), which includes average annual daily traffic (AADT) of the impacted link, length of the link, nodal degree, and total capacity. In addition to the road characteristics, two categorical variables are also used to differentiate traffic flow characteristics within the study period. Type of day is a binary variable that distinguishes between the type of day being a workday, which includes all weekdays (Monday through

Friday) except holidays, and non-workday, which includes weekends and holidays. Another variable used is the time period, which divides the day into five categories in accordance with HRRTDM, as mentioned in Table 1. The degree of a node is defined as the number of road links connected to the intersection of the impacted link. For this study, the nodal degree is assumed to be a sum of degrees on both nodes of the impacted link. The total capacity of the impacted link is calculated by multiplying the per lane capacity obtained from HRRTDM with the number of lanes and number of hours in the time period. This is done to reflect the difference in the number of hours in each time period. For the roads that are not covered in the HRRTDM, the number of lanes is obtained from the City of Norfolk's streets shapefile and multiplied with a default per lane capacity of 650 vehicles per hour per lane (minimum per lane capacity recorded in the HRRTDM).

Crowdsourced traffic volume data

The traffic impact index is a measure of assessing the impact on traffic flow due to nuisance flooding. Praharaj et al. *(2021a)* defined change in travel as a difference between vehicle hours of travel during a flooding event and a comparable non-flood period. The traffic impact index in this study is adapted from the concept of change in travel and is defined as a percent change in volumes on the impacted link during the flood period as compared to a non-flood period, as shown in Equation 1.

$$Traffic\ Impact\ Index = \frac{Volume_F - Volume_{NF}}{Volume_{NF}} * 100 \quad (1)$$

where,

$Volume_F$ = traffic volume on trustworthy report's impacted link during flood period

$Volume_{NF}$= traffic volume on trustworthy report's impacted link during non-flood period

The traffic impact index is a measure of assessing the impact on traffic flow due to nuisance flooding. Praharaj et al. *(2021a)* defined change in travel as a difference between vehicle hours of travel during a flooding event and a comparable non-flood period. The traffic impact index in this study is adapted from the concept of change in travel and is defined as a percent change in volumes on the impacted link during the flood period as compared to a non-flood period, as shown in Equation 1.

Hydrological characteristics

The environmental data is composed of rainfall and tide level observations. Hourly tide levels referenced to the North American Vertical Datum (NAVD88) were obtained from National Oceanic and Atmospheric Administration's (NOAA) Sewell's Point station *(NOAA, 2018)*. Hourly rainfall data were collected from seven Hampton Roads Sanitation District (HRSD) observation sites. Both of these datasets are publicly available. That data obtained for rainfall and tide levels are matched with the study period of available flood incident datasets.

Predicted street-level surface water depth can be simulated using a physics-based hydrodynamic model (TUFLOW: Two-dimensional Unsteady Flow) model. The flood model solves 2D equations for shallow water and free surface flow to simulate overland flow, and it is coupled with 1D hydrodynamic network software ESTRY *(Syme, 2001)* to simulate pipe flow. A Random Forest surrogate model was developed by *(Zahura et al., 2020)* to approximate TUFLOW-simulated street flood depth while making predictions faster and is used in this study. Although the high-fidelity TUFLOW model simulates realistic flood depth, it requires 4 to6 hours to simulate each flood event. As there were 188 flood events in this study, computation using TUFLOW would be tedious. The random forest surrogate model *(Zahura et al., 2020)* can approximate TUFLOW model outputs (water depth) using environmental and topographic data. The ground truth flood depth used to calibrate the model was obtained from the TUFLOW model. This

model can predict flood depths at a one-hour time step with high accuracy. The predictions made from this model significantly reduce computational time, by running 3,000 times faster compared to the TUFLOW model.

Table 1 summarizes all the datasets and models used to build the ground truth data set for the predictive model.

TABLE 1 Data inputs in predictive model

Data Category	Variable (unit)	Explanation	Data Dimension	Data Source or Method of Derivation
Crowdsourced flood incident data	*Date*	Date of flood report.	Temporal	Trustworthy Waze flood reports *(Praharaj et al. 2021b)*
	Time period	Time period of flood report 1: 12:00 to 6:00 am 2: 6:00 to 9:00 am 3: 9:00 am to 3:00 pm 4: 3:00 to 6:00 pm 5: 6:00 pm to 12:00 am	Temporal	Timestamp obtained from Waze, and aggregated to corresponding time period.
	Latitude and longitude	Location of flood report.	Spatial	Obtained from Waze.
Network Characteristics	*AADT (veh/day)*	Annual average daily traffic on the roadway link	Spatial	Obtained from HRRTDM.
	Link Length	Length of the affected link	Spatial	Obtained from HRRTDM.
	Node Degree	Sum of number of links connected to both intersections (nodes) of the the affected link	Spatial	Calculated in ArcGIS, based on the HRRTDM road network shapefile.
	Total Capacity (veh/period)	Total capacity of the roadway segment	Spatial	Obtained from HRRTDM $Total\ Capacity$ $= number\ of\ lanes\ (ln)$ $\times\ per\ lane\ capacity\ (veh/hr/ln\ )$ $\times\ \#hours\ in\ the\ period(hr/period)$
Crowdsourced Traffic Volume Data	*Traffic Impact Index*	Percent change in estimated traffic volumes for the time period during flood periods and comparable non flood periods	Spatial and Temporal	Data obtained from Praharaj et al. 2021a model using Streetlight Data. $Traffic\ Impact\ Index = \frac{Volume_F - Volume_{NF}}{Volume_{NF}} * 100$
Hydrological Characteristics	*Rainfall intensity (in/hr)*	Collected across seven rain gauges in the city, and interpolated for each flood report location.	Spatial and temporal	Data obtained from HRSD; Interpolation done by Inverse Distance Weighting (IDW), a spatial analysis tool in ArcGIS.
	Tide level (ft)	Maximum tide level recorded during the given time period at gauge at Sewell's Point.	Temporal	Obtained from NOAA Tides and Currents
	Water Depth (m)	Maximum water depth estimated for an hour in the time period	Spatial and Temporal	Obtained from Zahura et al. *(2020)* model

METHODS

In this study, a modeling technique is proposed that can work with the existing data to estimate traffic impacts due to flooding for locations with similar characteristics, but not enough traffic impact information. In this modeling process, the traffic impact index is the output or dependent variable. The overall framework of estimating traffic impact index using various independent variables discussed above is shown in Figure 2.

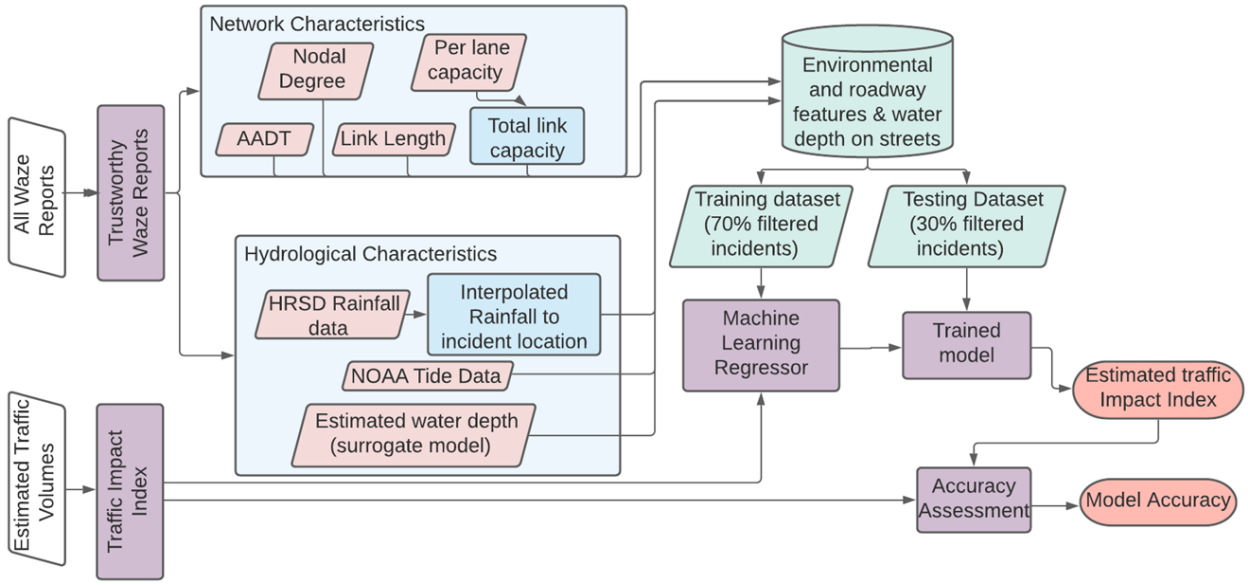


Fig 2 Framework for modeling traffic impact index

Several different modeling techniques were used to build the model. The model was tried out as a classification problem, with output as a categorical variable, and as a regression problem, with output as a continuous variable. The regression models were found superior to the classification models, therefore only regression models are discussed in this paper.

Classification and Regression Trees Model

Classification and regression trees (CRT) are a machine learning algorithm that group data points with similar dependent variable values together based on their independent variables. A parent node in the CRT is divided based on an independent variable into two child nodes, such that each child node is more homogenous (or less impure) than the parent node. Homogeneity is measured by the least squared deviation measure of impurity (within-node variance). The process stops until constraints such as minimum number of cases per node, maximum tree depth, node homogeneity, minimum change in improvement are not satisfied. In the transportation literature, CRT have been used for safety studies *(Iragavapu, et al, 2015, Harb et al., 2009, Pande et al., 2009)*, transit service quality determination *(de Ona et ak, 2012)*, and work zone applications *(Meng et al., 2012)*.

Random Forest Model

A random forest model *(Breiman, 2001)* is a machine learning technique that performs classification and regression tasks. The predictions are made by an ensemble of non-correlated decision trees, where the output is a mean prediction setting for regression trees or mode of classes for classification settings. The inherent randomness in the trees reduces the risk of overfitting, which is a drawback of CRT models. Random forest algorithms are optimized by tuning hyper-parameters such as number of trees, number of features at each split, maximum depth of a decision tree, and minimum number of samples in a node before splitting *(Probst et al, 2018, Pang et al., 2006)*. One of the outputs of the random forest model is the feature importance, which shows the importance of independent variables in predicting the values of

the output variable, which can help identify the unimportant features in the model. Random forests have been used for forecasting bicycle rentals *(Feng & Wang., 2017, Guidon et al., 2020)*, landslide susceptibility assessment *(Trigila et al., 2015)*, or roughness in asphalt pavements *(Gong et al., 2018)*, among others.

Gradient Boosting & Extreme Gradient Boosting Models

The extreme gradient boosting method is another machine learning developed by Friedman *(2001)* which can perform classification and regression tasks. This algorithm is similar to CRT and random forest models in a way that they also create trees for model prediction. However, the random forest model selects subset predictions from numerous randomly created regression trees based on the hyper-parameters. In contrast, the extreme gradient boosting algorithm works sequentially by generating tree models that sequentially work towards reducing model errors in each step from the previous step's models *(Anwar, 2021)*. This model also learns from creating trees with a combination of weak learner variables to create a stronger model. Thus, gradient boosting models have the potential to provide more accurate predictions *(Wang & Ross, 2018)*. Extreme gradient boosting model is optimized by reducing the second partial derivative of the loss function and uses hyper-parameters such as learning rate, number of trees, number of features at each split, maximum depth of a decision tree, and minimum number of samples in a node before splitting for tuning the model.

One-hot encoding

XGBoost has a limitation of not being able to work with categorical input data. Thus, during the dataset preparation, categorical datasets undergo one-hot encoding to ensure the efficient working of the algorithm. One-hot encoding is a method of data transformation, which converts categorical data into numerical data. Each category in the categorical data is transformed into binary variables, with 0 and 1 as the values. For example, the time period consists of 5 categories, and with on-hot encoding, now the model has 5 independent variables with 0 or 1 entries, depending on the observation's actual time period. The same method is used for the type of day variable.

Random search-cross validation

Random Search Cross-Validation (RS-CV) function uses a parameter called n-folds, where n is the number of groups that the given dataset is divided into. In a machine learning model, RS-CV randomly splits the training dataset into n-folds, and reserves one-fold for testing purposes. This technique protects the models against overfitting the data, particularly in small datasets. XGBoost model goes through cross-validation to set up the optimal hyper-parameters for the least error model.

Model Evaluation

Once the models are developed, model performances of various machine learning techniques are evaluated and compared using the root-mean-square-error (RMSE). These values are checked for training and testing data, and the model with least errors are selected. The RMSE and normalized-RMSE are calculated by equations 2 and 3 respectively.

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n} TI_{obs,i} - TI_{pred,i}}{n}} \quad (2)$$

$$n - RMSE = \frac{RMSE}{TI_{obs,max} - TI_{obs,min}} \quad (3)$$

Where i = observations in the dataset,

n = total observations,

TI = traffic impact index,

obs = observed value,
pred = predicted value,

RESULTS & DISCUSSION

Summary Statistics

When building the machine learning models, a 70-30 random split of observations were used for training and testing, respectively. The training set is randomly sampled until at least one observation of each of the categories of the categorical variables is present in the training set. The distribution of categorical variables is shown in Table 3.

Table 3: Summary of sampling of categorical variables

Categorical variables	Number of observations (train | test)
Total	188 (133 | 55)
Workday	155 (114 | 41)
Non- workday	33 (23 | 10)
Period 1 (12:00 am to 6:00 am)	2 (2 | 0)
Period 2 (6:00 am to 9:00 am)	11 (6 | 5)
Period 3 (9:00 am to 3:00 pm)	40 (28 | 12)
Period 4 (3:00 pm to 6:00 pm)	88 (65 | 23)
Period 5 (6:00 pm to 12:00 am)	47 (32 | 15)

Traffic Impact Index for the dataset is calculated as the percent change in traffic volumes during the flood period. This index value ranges from -100% to +180%, with an average of -1.3% and standard deviation of 36%. The distribution follows a near-normal bell curve, with a longer tail on the right, as shown in Figure 3.

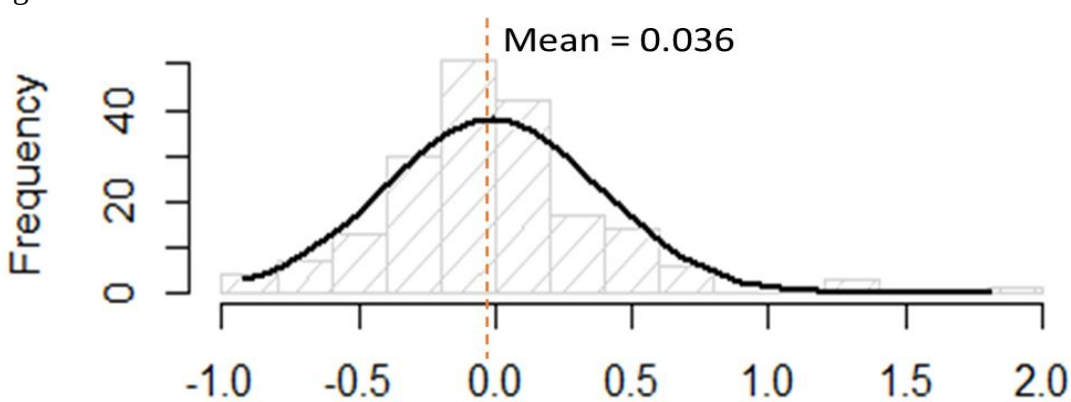


Fig 3 Frequency distribution of traffic impact index

Hyper-Parameter Selection

CRT, random forest model and extreme gradient boosting models are checked for least errors. This is done by setting some hyper-parameter values for least error model.

Classification and Regression Trees Model

Classification and regression trees use two parameters: maximum depth of the tree and minimum samples in each split. The tree is also pruned to avoid overfitting. Expansion of the tree is based on a complexity

parameter (cp), which is the minimum improvement needed in the model to split and propagate. The cp value was optimized at 0.008.

Random Forest Model

For random forest model, the number of trees (n_{tree}) and the number of features considered at each split(m_{try}) are selected based on the stabilization of out-of-bag error rates. Bagging is a statistical resampling technique used within the training dataset to quantify uncertainty in randomization. The training dataset is divided into in-bag and out-of-bag sample sets, where multiple trees are created from the in-bag sample set, and errors are calculated based on the predictions made on the out-of-bag sample set. Out-of-bag error is the average error of each observation in the out-of-bag sample set between predicted and observed values in their respective samples. Out-of-bag error rates are essential in deciding the hyper-parameters for the model. The out-of-bag error rate graphs for RF model are shown in Figure 4 (a) and 4 (b) respectively. n_{tree} value as 60, and m_{try} value as 12 were selected to ensure the least RMSE values.

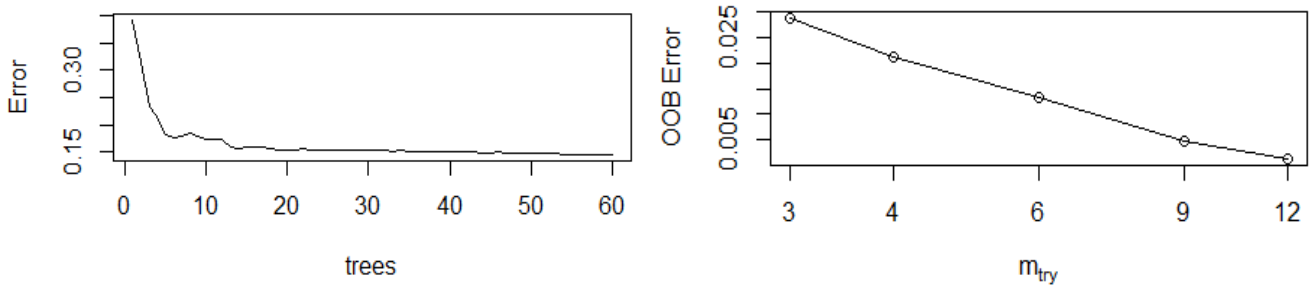


Fig 4 Selecting hyper-parameter in RF model for (a) number of trees and (b) m_{try}

XGBoost Model

There are various hyper-parameters used in the XGBoost model by using the boosting technique. These hyper-parameters are tuned using random search and cross validation techniques *(Wang & Ross, 2018)*. The hyper-parameters, their definition, default values, range and optimal values of the final model are listed in Table 4. These values are selected internally by iterative running of models until the model gives least errors.

Table 4: Selected hyper-parameter values for XGBoost model

Hyper-parameter	Definition	Default	Range	Optimal Value
n rounds	maximum number of iterations	100		
Eta	learning rate	0.3	(0,1)	0.292
gamma	prevents overfitting	0	(0, inf)	0
max_depth	depth of the tree	6	(0, inf)	8
min_child_weight	minimum number of observations required in a child node	1	(0, inf)	1
sub_sample	number of samples supplied to a tree	1	(0, 1)	0.998
colsample_bytree	number of features supplied to a tree	1	(0,1)	0.687

Feature Importance

Feature importance are in-built functions in the RF and XGBoost models in R. Feature importance helps us understand which features (or variables) contribute towards improving the model. Feature importance for RF model was calculated from *increase in node purity* with the introduction of each variable for

making the trees *(Menze et al., 2009)*. Feature importance for XGBoost models is calculated from *gain* of each feature. *Gain* is a relative contribution of the feature to the model, which is calculated by taking each feature's contribution for each tree in the model *(Rogers & Gunn, 2005)*.

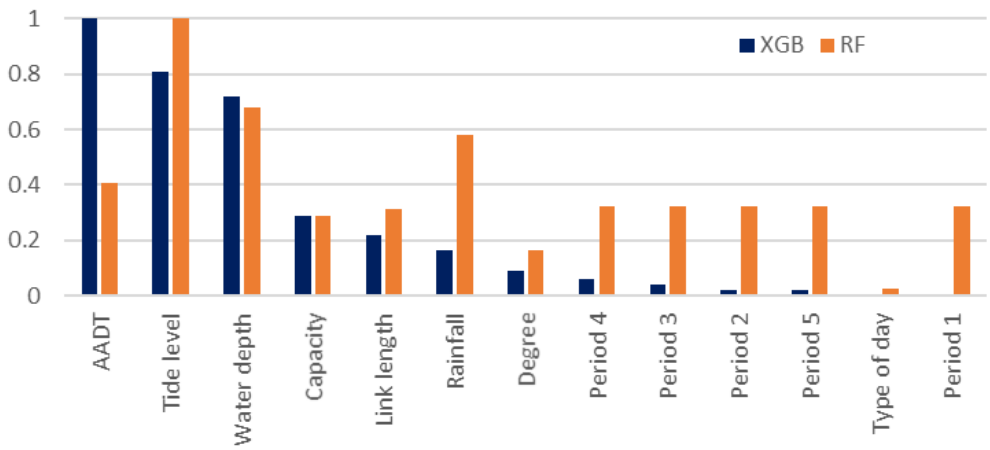


Fig 5 Relative importance of variables in RF model and XGBoost model

We find that AADT, tide level, water depth and rainfall are important explanatory variables in both models. These are intuitive indicators as they are the reasons of flooding, along with the amount of traffic on the road that gets impacted. Capacity and link length are next in the importance levels, which are roadway related features and influence the amount of traffic flow on the road. Some of the features such as period, type of day, degree have varying feature importance. The feature importance of different variables are different in both models due to the inherent difference in ensemble models. Thus, we go through a procedure of permuting each variable, as shown in the *Model Selection* section to find the model with least errors.

Model Selection

Once the hyper-parameters are set, the optimized models are compared to find the machine learning model that has least errors, and has most accurate predictions. We permute one variable in each iteration of both RF and XGBoost models to find which variable omission provides a model with least errors. This graph is shown in Figure 6.

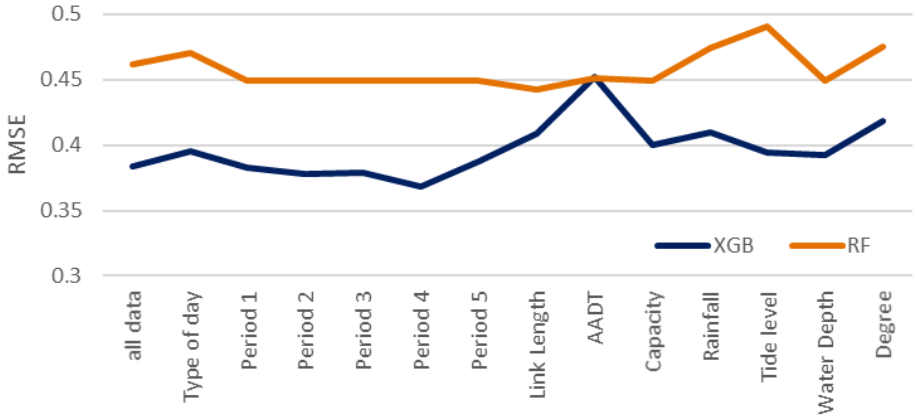


Figure 6 Selection between model algorithms and variable omission

From Figure 6 we find that XGBoost models have consistently lower RMSE values when compared to RF models. In addition to that, among the XGBoost models, removal of the period 4 variable shows the minimum RMSE value. However, since time period is a categorical variable, and is one-hot encoded for XGBoost modeling only, we do not omit this variable. Thus we select a preferred model with all features included which gives the lowest RMSE values. The normalized RMSE values with all features for CART, RF and XGBoost models are summarized in Table 5.

Table 5 Model selection based on testing errors

Model	Normalized RMSE (test)	Normalized RMSE (train)
CRT	0.172	0.134
RF	0.168	0.062
XGBoost	0.140	0.095

The training RMSE error for XGBoost model is very low since the model learns from each step in model training. We compare the testing errors of the 3 models to select the finalized model. XGBoost model has the least error of 14% and is thus selected as the preferred model. This framework enables a better understanding of how to compare these models for better prediction accuracy.

Testing Results

We plot the predicted results of the XGBoost model (orange) against the observed Traffic Impact Index (dark blue), as shown in Figure 7. The dashed grey lines show the limits for one standard deviation away from the observed Traffic Impact Index.

Fig 7 Comparison of observed & predicted traffic impact index from XGBoost model

In Figure 7, we find that 37 of the 55 observations (67.2%) in the test set fall within this buffer area. These observations that are predicted within one standard deviation from the observed values are considered better predicted than others. The observations that occur in the extreme ends of the graph are the ones experiencing significant impacts due to flooding. These flood events would need to be predicted with greater accuracy and would need special attention by city managers. To increase the prediction accuracy of these events, the ground truth would need to be strengthened with data that are highly impacted and have greater variability in their characteristic inputs into the model. Since the sample set for the test is so small, there also aren't any notable characteristic emerging from the test set yet. Since the model is

based on sequential learning, lack of a big dataset also hinders the prediction accuracy. With collection of more data, it will be possible to make the predictions better these characteristics become more prominent, and enable the classification of locations that are estimated to experience a significant change in volume or those locations that do not.

Once these predictions are made better, there can be a threshold selected around 0 which can separate events of low and high impact. Low impact events would be the events that have a Traffic Impact Index close to 0, implying not having a significant change in traffic from non-flood periods. High impact events would constitute of flood events that cause a high increase or decrease of traffic. Based on the input roadway features, topographic features and predicted water depth, this model can be helpful in predicting the traffic impact index of different locations in the future, which do not have observed traffic counts from crowdsourced data or estimated traffic volumes from the estimation model (*Praharaj et al., 2021a*). Based on these estimated traffic impact index, potential flood prone locations can be separated into categories in which they would fall into for emergency management services in the city. In accordance with that, decisions on traffic rerouting can also be made to reduce severe traffic impacts due to flooding.

CONCLUSION & FUTURE WORK

Nuisance flooding is becoming an increasingly common phenomena in coastal cities due to sea level rise and climate change. Knowledge of location of incidents and water levels on roads can serve as a very useful tool in reduction of travel delays caused to road users. Early assessment of water levels and potential flooding hotspots can be useful in early detection of flooding as well as rerouting to reduce traveler delays. In this study, we demonstrate the potential of several machine learning models which can relate locational characteristics to traffic impacts of flood incidents. Among the different kinds of regression-based machine learning models used, random forest (RF) and extreme gradient boosting (XGBoost) were extensively tested and compared. It was found that XGBoost model outperforms classification and regression tree model and random forest models, with model error (normalized RMSE) at 14%. It was found that AADT, tide level, rainfall and estimated water depth are few of the important features in predicting traffic impact index. One of the results observed when plotting estimated and model predicted traffic impact index is that the model predicts about 67% of the events in the one-standard deviation buffer. Adding more data to the dataset will help improve the model predictions, as well as enable setting a threshold for division of high and low impact events. Thus, these kinds of machine learning models can prove to be pivotal in estimating the traffic impacts on the flood incidents that can be predicted ahead of time.

The current study serves as a demonstrative model, which with a larger dataset and better predictive power, can serve as an important tool for the decision makers to identify flood prone hotspot locations that deeply impact traffic operations. The knowledge of these can be crucial in directing emergency services to impacted locations. This could also separate highly impacted locations, which can then be prioritized. In addition to this, preemptive knowledge of historically impacted locations can help in proactively assigning services to anticipated high impact locations.

The current study has its own set of limitations. Although this study can prove powerful in early assessment of flooding and forewarnings about estimated traffic impacts, the current dataset is relatively small. The current dataset has less than 200 observations with complete data from all the data sources used. Observations are unbalanced for a few classes in categorical variables, which can heavily impact the learning ability of the model. There are also different kinds of continuous variables in the model. More data can provide a better range of these values as well, which could greatly improve the predicting power of the model.

Nevertheless, this paper provides important guidance in how to utilize crowdsourced data to predict traffic impacts due to flooding on the roadway network. This methodology merges different hydrological, transportation and trustworthiness modeling together in an effort to reduce travel impacts due to nuisance flooding. With newer technology for draining out water from the roadway network, this technique can still be used for assessing hotspot locations for flooding. This can help the city officials to work out diversion routes in advance from the areas which would get highly impacted, and also direct temporary drainage measures to the highly impacted areas.

ACKNOWLEDGEMENTS

The authors would also like to thank City of Norfolk, Streetlight Data and Waze for facilitating data acquisition. This work is supported by the National Science Foundation's Critical Resilient Interdependent Infrastructure Systems and Processes program (Award 1735587).

REFERENCES

1. Appert, M., and Laurent, C. (2007). "Measuring urban road network vulnerability using graph theory: the case of Montpellier's road network." 23.
2. Chen, X.-Z., Lu, Q.-C., Peng, Z.-R., and Ash, J. E. (2015). "Analysis of Transportation Network Vulnerability under Flooding Disasters." Transportation Research Record: Journal of the Transportation Research Board, 2532(1), 37–44.
3. Jun-qiang, L., Long-hai, Y., Liu, W., and Zhao, L. (2017). "Measuring Road Network Vulnerability with Sensitivity Analysis." PLOS ONE, (T. Tang, ed.), 12(1), e0170292.
4. Kermanshah, A., Derrible, S., and Berkelhammer, M. (2017). "Using Climate Models to Estimate Urban Vulnerability to Flash Floods." Journal of Applied Meteorology and Climatology, 56(9), 2637–2650.
5. Kim, S., and Yeo, H. (2016). "A Flow-based Vulnerability Measure for the Resilience of Urban Road Network." Procedia - Social and Behavioral Sciences, 218, 13–23.
6. Pyatkova, K., Chen, A. S., Djordjević, S., Butler, D., Vojinović, Z., Hammond, M., Pyatkova, K., Chen, A. S., Djordjevic, S., Butler, D., Vojinovic, Z., Abebe, Y., and Hammond, M. (2015). "Flood Impacts on Road Transportation Using Microscopic Traffic Modelling Technique." 9.
7. Shekar, V., Fiondella, L., Chatterjee, S., and Halappanavar, M. (2017). "Quantitative assessment of transportation network vulnerability with dynamic traffic simulation methods." 2017 IEEE International Symposium on Technologies for Homeland Security (HST), IEEE, Waltham, MA, USA, 1–7.
8. Singh, P., Sinha, V. S. P., Vijhani, A., and Pahuja, N. (2018). "Vulnerability assessment of urban road network from urban flood." International Journal of Disaster Risk Reduction, 28, 237–250.
9. Yang, J., Sun, H., Wang, L., Li, L., and Wu, B. (2013). "Vulnerability Evaluation of the Highway Transportation System against Meteorological Disasters." Procedia - Social and Behavioral Sciences, 96, 280–293.
10. Yang, C., Tu, Y., and Chen, X. (2009). "Analysis Method for Topology Vulnerability of Transportation Network." International Conference on Transportation Engineering 2009, American Society of Civil Engineers, Southwest Jiaotong University, Chengdu, China, 3639–3644.
11. Iragavarapu, V., D. Lord, and K. Fitzpatrick. Analysis of Injury Severity in Pedestrian Crashes Using Classification Regression Trees. Presented at 94th Transportation Research Board Annual Meeting, Washington D.C., 2015.
12. Harb, R., X. Yan, E. Radwan, and X. Su. Exploring precrash maneuvers using classification trees and random forests. *Accident Analysis and Prevention*, Vol. 41(1), 2009, pp. 98-107.
13. Pande, A., M. Abdel-Aty, and A. Das. A classification tree based modeling approach for segment related crashes on multilane highways. *Journal of Safety Research*. Vol. 41(5), 2009, pp. 391-397.

14. de Ona, J., R. de Ona, and F. Calvo. A classification tree approach to identify key factors of transit service quality. *Expert Systems with Applications*. Vol. 39 (12), 2012, pp. 11164-11171.
15. Meng, Q., and J. Weng. Classification and Regression Tree Approach for Predicting Drivers: Merging Behavior in Short-Term Work Zone Merging Areas. *Journal of Transportation Engineering*. Vol. 138(8), 2012. http://dx.doi.org/10.1061/(ASCE)TE.1943-5436.0000412.
16. NOAA, 2018b. Sewells Point - Station Home Page - NOAA Tides & Currents [WWW Document]. URL https://tidesandcurrents.noaa.gov/waterlevels.html?id=8638610 (accessed 10.1.19).
17. Syme, W. J. TUFLOW - Two one-dimensional Unsteady FLOW Software for Rivers, Estuaries and Coastal Waters, 2001.
18. Zahura, F. T., Goodall, J. L., Sadler, J. M., Shen, Y., Morsy, M. M., & Behl, M. (2020). Training Machine Learning Surrogate Models from a High-Fidelity Physics-Based Model: Application for Real-Time Street-Scale Flood Prediction in an Urban Coastal Community. *Water Resources Research*, 2020. Volume 56(10).
19. Berdica, K. (2002). An introduction to road vulnerability: what has been done, is done and should be done. *Transport Policy*, 9, 117–127.
20. Haghighi N., Fayyaz S. K., Liu X. C., Grubesic T. H., & Wei R. (2018). A multi-scenario probabilistic approach for critical transportation network risk assessment. *Network and Spatial Economics*, 18(1), 181-203.
21. Jenelius E., & Mattsson L. G. (2012). Road network vulnerability analysis of area-covering disruptions: a grid-based approach with case study. *Transportation Research: Part A*, 46, 746–760.
22. Lu Q. C., & Peng Z. R. (2011). Vulnerability Analysis of Transportation Network under Scenarios of Sea Level Rise. *Transportation Research Record: Journal of the Transportation Research Board*, 2263, 174-181.
23. Rodriguez- Nunez E., & Palomares J. C. G. (2014). Measuring the vulnerability of public transport networks. *Journal of Transport Geography*, 35, 50-63.
24. Sohn J. (2006) Evaluating the significance of highway network links under the flood damage: an accessibility approach. *Transportation Research Part A: Policy*, 40, 491-506.
25. Lopez F. A., Paez A., Carrasco J. A., & Ruminot N. A. (2017). Vulnerability of nodes under controlled network topology and flow autocorrelation conditions. *Journal of Transport Geography*, 59, 77-87.
26. Praharaj, S., Chen, T. D., Zahura, F. T., Behl, M., & Goodall, J. L. (2021). Estimating impacts of recurring flooding on roadway networks: a Norfolk, Virginia case study. *Natural Hazards*, 1-25.
27. Praharaj, S., Zahura, F. T., Chen, T. D., Shen, Y., Zeng, L., and Goodall, J. L. (2021). Assessing Trustworthiness of Crowdsourced Flood Incident Reports using Waze Data: a Norfolk, Virginia case study. *Transportation Research Record: Journal of the Transportation Research Board* (accepted with minor revisions).
28. Chen, B. Y., Lam, W. H., Sumalee, A., Li, Q., & Li, Z. C. (2012). Vulnerability analysis for large-scale and congested road networks with demand uncertainty. Transportation Research Part A: *Policy and Practice*, 46(3), 501-516.
29. Taylor, M. (2017). *Vulnerability analysis for transportation networks*. Elsevier.
30. Koetse, M. J., and P. Rietveld. 2009. The impact of climate change and weather on transportation: An overview of empirical findings. *Transportation Research Part D*.
31. Breiman, L. (2001). Random forests. *Machine learning*, *45*(1), 5-32.
32. Herbert Pang, H., Lin, A., Holford, M., Enerson, B. E., Lu, B., Lawton, M. P., Floyd, E., Zhao, H. Pathway analysis using random forests classification and regression, *Bioinformatics*, Volume 22, Issue 16, 15 August 2006, Pages 2028–2036.

33. Feng, Y., & Wang, S. (2017, May). A forecast for bicycle rental demand based on random forests and multiple linear regression. In *2017 IEEE/ACIS 16th International Conference on Computer and Information Science (ICIS)* (pp. 101-105). IEEE.

34. Trigila, A., Iadanza, C., Esposito, C., & Scarascia-Mugnozza, G. (2015). Comparison of Logistic Regression and Random Forests techniques for shallow landslide susceptibility assessment in Giampilieri (NE Sicily, Italy). *Geomorphology*, *249*, 119-136.

35. Guidon, S., Reck, D. J., & Axhausen, K. (2020). Expanding a (n)(electric) bicycle-sharing system to a new city: Prediction of demand with spatial regression and random forests. *Journal of Transport Geography*, *84*, 102692.

36. Gong, H., Sun, Y., Shu, X., & Huang, B. (2018). Use of random forests regression for predicting IRI of asphalt pavements. *Construction and Building Materials*, *189*, 890-897.

37. Friedman, J. H. Greedy function approximation: A Gradient Boosting Machine. *Annals of Statistics.* Vol 29, No. 5, 2001, pp 1189-1232.

38. XGBoost 0.6 Documentation. Understand Your Dataset with XGBoost. http://xgboost.readthedocs.io/en/latest/R-package/ discoverYourData.html#feature-importance. Accessed March 20, 2021.

39. Wang F, Ross CL. Machine Learning Travel Mode Choices: Comparing the Performance of an Extreme Gradient Boosting Model with a Multinomial Logit Model. *Transportation Research Record*. 2018;2672(47):35-45.

40. Anwar, M. T., Winarno, E., Hadikurniawati, W., & Novita, M. (2021, April). Rainfall prediction using Extreme Gradient Boosting. In *Journal of Physics: Conference Series* (Vol. 1869, No. 1, p. 012078). IOP Publishing.

41. Menze, B.H., Kelm, B.M., Masuch, R. *et al.* A comparison of random forest and its Gini importance with standard chemometric methods for the feature selection and classification of spectral data. *BMC Bioinformatics* **10,** 213 (2009).

42. Rogers, J., & Gunn, S. (2005, February). Identifying feature relevance using a random forest. In *International Statistical and Optimization Perspectives Workshop" Subspace, Latent Structure and Feature Selection"* (pp. 173-184). Springer, Berlin, Heidelberg.

CHAPTER 5
OVERALL CONCLUSION AND FUTURE WORK

Sea level rise and climate change have caused urban coastal flooding to be a more frequentt and more intense disruption, which impacts the stormwater system and roadway networks. There has been significant progress made in the research area in terms of estimating how the system reacts to recurrent flooding and the system recovery time and effort. However, one of the major limitations is the lack of empirical flood incident and traffic data in these studies, which stems from not having adequate reliable data sources that can tell the complete story of when and where the flood happens, how much flooding occurs, how the traffic behaves, and what measures would reduce the stress on the transportation and stormwater systems. This can be a critical limitation due to the nature of urban coastal flooding, which distresses the system for several hours and then the system in due course returns to normal operations. Yet the increasing nature of these disruptions are causing significant impacts on the transportation system users. This dissertation explored how crowdsourced data can be used to overcome the limitation posed by limitations in agency data sources. Major contributions of this research, potential applications of the developed research, and future research are summarized in this chapter.

5.1 SUMMARY OF MAJOR FINDINGS AND CONCLUSIONS

This research was conducted with the data obtained from various city and state agencies for the City of Norfolk, Virginia. The models developed in this research primarily focus on freeways and major arterials. Summary of the major findings in this study are listed as follows:

- Flooding impact estimation:
    - Estimation of citywide traffic flow impacts in terms of vehicle-hours of travel (VHT) for recurrent flooding was found to be 3%, relatively not significant at a citywide level. It was found that the agency-based flood incident datasets were not comprehensive enough to capture the temporal and spatial disaggregation levels of nuisance flooding. In addition, VHT is a cumulative measure for speeds and volumes, and a simultaneous decrease in volumes and increase in travel times aggregated over the entire city is not a good measure for impact assessment.
    - Estimation of localized traffic flow impacts is conducted, and the results show that the majority of the links within the subnetwork of the impacted area show a reduction in speed and volumes, with an aggregate VMT reductions of 12% across all affected links. There

are significant VHT reductions also observed during specific periods in the day and in specific roadway classes, which is indicative of varying temporal impacts of flooding on the traffic flow.

- It was found from the first paper that recurrent flooding is a localized event, and does not impact the entire city network, and is thus recommended to assess only on subnetwork level, to find clearer impacts on the roadway network.

- Waze trustworthiness:
  - This study presented a framework to assert the trustworthiness of crowdsourced flood incident data. It was found that contextual factors such as topographic, environmental, and infrastructure variables can play an important role in determining the trustworthiness of crowdsourced flood incident reports.
  - This study uses a binary logistic regression model to separate trustworthy reports from untrustworthy ones. The preferred model developed in this study shows an accuracy of about 90%. When utilized on a larger Waze dataset, about 72% of the reports were classified as trustworthy. The most notable characteristics of the untrustworthy reports were low occurrence of peer reporting, reporting at higher elevation locations, or during low rainfall intensities.
- Traffic Impact Index assessment:
  - This study presents a framework to estimate the traffic impacts of locations in the roadway network due to recurrent flooding where verified crowdsourced traffic data might not be available. The model relies heavily on its dataset, and thus can currently only predict locations that have similar characteristics as the modeling dataset.
  - The study compares a few different machine learning techniques to find the model with the least errors for predicting traffic impact index. It was found that the extreme gradient boosting model showed the least errors, amounting to 14% normalized errors. About 67% of the dataset was predicted within one standard deviation of the dataset. The dataset used in this study was small, hence not many results were discussed in this section.

5.2 RESEARCH CONTRIBUTIONS

Major contributions of this study are listed as follows:

- The first study addresses the lack of near-real-time datasets which are temporally and spatially disaggregated, to help understand the impacts of recurrent flooding on the roadway network. Most of the studies in this domain use simulated data for analysis, which are not useful in providing a decision framework for reduction of disruption delays in real-time. Chapter 2 presents a data-predictive framework for using spatially and temporally disaggregate crowdsourced traffic volume-related datasets. With the increased availability of traffic flow and hydrological variables, this framework can be crucial for real-time assessment of recurrent flooding, which is increasingly being a concern in low-elevation coastal cities, especially on the east coast of the US. This framework can also be generalized to be used in areas with limited agency datasets. A similar framework can be used for accident or post-disruption analysis as well.
- The second study addresses the hesitancy of decision-makers on the usage of crowdsourced incident data due to its lack of reliability, which stems from unregulated data collection. Chapter 3 presents a methodology to assert trust in crowdsourced incident detection datasets by utilizing contextual variables. Waze is becoming widely popular all over the world, harnessing location-based services and user-generated incident detection data for finding incidents, traffic jams, irregularities, etc. The methodology presented in this study can be replicated for identifying trustworthy flood detection incidents, or a similar framework to identify trustworthy disruption detection with relevant geographical context, such as landslides, wildfires, etc.
- The third portion of this study builds on the previous models and defines a near-real-time traffic impact index. This data is then used to build a model that can predict traffic impact index. With this model, traffic impact index can be predicted of areas which do not have sufficient traffic information. The modeling framework used in this study can be used as preliminary work to estimating the degree of impact of various flooded locations. In addition, forecasting different variables such as water depth, tide level, and rainfall can have a potential in estimating the directionality of impact ahead of time, which can be useful in preemptively redirecting traffic away from the problematic locations.

5.3 POTENTIAL APPLICATIONS

A major gap in the existing literature (as illustrated throughout the dissertation) is the dearth of research utilizing empirical data sources to characterize flooding impacts, especially for recurrent flooding. The various models developed in this study will enable more familiarity with crowdsourced data and promote

its usage in various transportation applications. The models developed and findings from this research can be utilized in different applications such as:

- Traffic volume estimation: City and state agencies have a relatively small amount of continuous traffic volume collection stations available. These are usually expensive to maintain and agencies strategically place them at certain locations only. The data-predictive model built for traffic volume estimation can be used in estimating traffic volumes in any region with limited agency data and could use more comprehensive crowdsourced data for various traffic analyses. These estimated traffic volumes can be used in traffic impact studies, traffic delay assessment, traffic signal studies, safety studies, short-term forecasting, etc.
- Data predictive framework: The framework proposed in this dissertation for the assessment of flooding impacts is not limited to flooding only. This framework can be used for other disruption analyses as well, such as congestion or accident analyses, post-disruption analyses, etc. Assessment of these disruptions can be useful in understanding the intensity and frequency of disruptions, as well as allow decision-making measures to reduce the impact on the roadway network. With more disaggregated data, this framework can also be used to identify the time it takes the system to return to its normal functioning, such that steps can be taken to reduce this time.
- Incident data reliability: There have been various papers in the domain of crowdsourced incident data that showcase the amount of overlap with agency-reported data and the potential of obtaining a higher volume of incident data from locations where there is not enough agency presence for incident reporting. However, there is still hesitancy in the adoption of crowdsourced datasets for incident detection due to its innate nature of unregulated reporting. The framework described in this dissertation can prove useful to address this concern by tying the incidents with potential incident-causing contextual factors to assign a value of reliability to the incident reports. The current study deals with recurrent flooding; however, the framework can be extended to other location-dependent disruption events such as landslides, forest fires, fog, etc.
- Dynamic vulnerability assessment framework: There have been various studies in the transportation resilience domain that have assessed the impacts of minor disruptions using simulated data, and provided suggestions for reduction of traffic impacts. However, simulated scenarios are not the same as observed disruptions, and can make real-time decision-making difficult. Therefore, using a data predictive model which can learn from similar disruptions to

make near-real-time predictions on traffic impacts would be useful in preemptive decision-making for various types of disruptions. This preemptive decision-making can be used to select times and locations of planned disruptions like work zones, or other minor predicted natural disruptions, for which the decision making can be used as a forewarning.

5.4 FUTURE RESEARCH

While this research has made significant progress in enabling the use of crowdsourced data into flood assessment and mitigation research, several possible avenues of research have been identified, and are listed as follows:

1. **Expanding traffic volume estimation methodology to other facility types**: The current traffic volume estimation model in our study is restricted to freeways and major arterials only. This is due to the ground truth used to build the model only consisting of VDOT's continuous count station data. While the selected model has a very low error rate, this model can be expanded more. The inclusion of traffic volume data from temporary count stations across the study can increase the spatial footprint of the model to more local roads as well. In addition, this model can also be tested for different regions outside of Norfolk, VA to assess if the same model can be used for traffic volume predictions in other similar areas.
2. **Adjusting for other vehicle types**: The current traffic volume estimation model is restricted to passenger cars only due to the type of information from crowdsourced traffic counts data. With the use of different crowdsourced data sources, this framework can be used for heavy-duty vehicles as well. Heavy-duty vehicles usually follow a stricter schedule and delays in their schedule incur significant economic costs, and quantification of such costs would be essential in impact assessment.
3. **Inclusion of smart cities' data**: As the cities include smart sensor technology, there would be more disaggregate data available for rainfall, tide level, flood locations, traffic volumes, etc. The models used in this research could be updated to include more spatially and temporally disaggregate data and include potentially new variables. Smart cities can also enable obtaining more data, which can strengthen the ground truth datasets in all the models used, and thereby provide more accurate predictions.
4. **Improving the trustworthiness model**: Ground truth data used in the Waze trustworthiness model relied on the City of Norfolk flood incident data, which was biased more towards more intense flood events and rainfall-induced flood events. Upon installing smart sensors in select

locations of the city, ground truth data can be significantly improved with lesser bias. This will enable the Waze model significantly better in asserting trustworthiness to unregulated Waze flood incident data.

5. **Decision-making system framework**: This dissertation uses a modeling framework to identify incident locations that experience different levels of traffic impacts due to recurrent flooding. However, the dataset used in this study is very limited. Obtaining more data and building a dataset with variables having more diverse data, can help with the prediction of traffic impacts. This can improve the spatial footprint of the dataset, and enable in finding hotspot locations that have more deteriorating impacts, which can be used for rerouting to adjacent roads in the network that are not as impacted.
6. **Traffic rerouting for reduced vulnerability**: As the more vulnerable locations are identified in the network, this work can be coupled with other hydrological models like water depth prediction models, automated real-time stormwater, and real-time traffic flow. Combination of all these models can establish the resilience of the network to recurrent flooding, as well as enable the possibility of traffic rerouting to bring the system is back to normalcy at a faster rate.

PRESENTATIONS AND PUBLICATIONS

PUBLICATIONS

- Praharaj, S., Chen, T.D., Zahura, F.T., Behl, M., and Goodall, J. L. Estimating impacts of recurring flooding on roadway networks: a Norfolk, Virginia case study. *Nat Hazards* (2021). https://doi.org/10.1007/s11069-020-04427-5.
- Praharaj, S., Zahura, F. T., Chen, T. D., Shen, Y., Zeng, L., and Goodall, J. L. Assessing Trustworthiness of Crowdsourced Flood Incident Reports using Waze Data: a Norfolk, Virginia case study. *Journal of Transportation Research* Record (accepted with minor revisions).
- Praharaj, S., Chen, T. D., Zeng, L., Zahura, F. T., and Goodall, J. L. Vulnerability Index for coastal flooding: A Norfolk, Virginia case study. (to be submitted).

NATIONAL & INTERNATIONAL PRESENTATIONS

Estimating Impacts of Recurring Flooding on Roadway Networks: A Norfolk, Virginia Case Study

- Presented *at Bridging Transportation Researchers*, August 2020
- Presented *at Transportation Research Board 99th Annual Meeting*, January 2020
- Presented at 2nd *International Conference on Transportation System Resilience to Natural Hazards and Extreme Weather*, November 2019
- Presented at 3rd *World Conference for Transport Research*, Mumbai, India, May 2019
- Presented at *Engineering Sustainability*, April 2019

Assessing Trustworthiness of Crowdsourced Flood Incident Reports using Waze Data: A Norfolk, Virginia Case Study

- Presented at *ASCE International Conference for Transportation & Development*, June 2021
- To be presented *at Transportation Research Board 100th Annual Meeting*, January 2021