Assessing the Crash Risks of Evacuation: A Matched Case-Control Approach Applied over Data Collected during Hurricane Irma

Rezaur Rahman, Tanmoy Bhowmik, Naveen Eluru, and Samiul Hasan

ABSTRACT

Recent hurricane experiences have created concerns for transportation agencies and policymakers to find better evacuation strategies, especially after Hurricane Irma—which forced about 6.5 million Floridians to evacuate and caused a significant amount of delay due to heavy congestion. A major concern for issuing an evacuation order is that it may involve a high number of crashes in highways. In this study, we present a matched case-control based approach to understand the factors contributing to the increase in the number of crashes during evacuation. We use traffic data for a period of 5 to 10 min just before the crash occurred. For each crash observation, traffic data are collected from two upstream and two downstream detectors of the crash location. We estimate models for three different conditions: regular period, evacuation period, and combining both evacuation and regular period data. Model results show that, if there exist a high volume of traffic at an upstream station and a high variation of speed at a downstream station, the likelihood of crash occurrence increases. Using a panel mixed binary logit model, we also estimate the effect of evacuation itself on crash risk and find that, after controlling for traffic characteristics, during evacuation the chance of a crash is higher than in a regular period. Our findings have implications for evacuation declarations and highlight the need for better traffic management strategies during evacuation. Future studies may develop advanced real-time crash prediction models which would allow us to deploy proactive countermeasures to reduce crash occurrences during evacuation.

Keywords: Evacuation, crash, case-control, variation of speed, crash prediction.

INTRODUCTION

Evacuation has become a major issue for coastal residents, particularly after recent hurricanes in the United States. During Hurricane Irma, about 6.5 million residents of Florida evacuated from major cities including Key West, Miami, and Tampa. With only two major interstate highways (I-75 and I-95) available for leaving Florida, evacuation caused a significant amount of traffic congestion and crashes affecting the physical and mental health of evacuees. To deal with this condition, advanced traffic management strategies are needed to ensure safety and better mobility for the evacuees. Thus, evacuation traffic management has been a major concern for transportation agencies and policymakers (House of Representatives Florida, 2018) and several strategies have been deployed (Murray-Tuite et al., 2017; Murray-Tuite and Wolshon, 2013). However, these strategies seem to be less effective to reduce the number of crashes during an evacuation. For instance, during Hurricane Irma's evacuation, about 221 crashes occurred on I-75 between September 7 and September 9, 2017 (close to Irma's landfall day)—which also caused significant delay for the evacuees to reach a safe destination. As such, to ensure a safe and efficient evacuation of a large number of people, we need to assess the contributing factors causing an increase in the number of crashes during evacuation.

[This is the pre-print version of the following article: Rahman, Bhowmik, Eluru, and Hasan (2021). Assessing the Crash Risks of Evacuation: A Matched Case-Control Approach Applied over Data Collected during Hurricane Irma. Accident Analysis and Prevention, 159, 106260, which has been published in final form at https://doi.org/10.1016/j.aap.2021.106260*].*

Researchers have proposed several state-of-the-art methods to assess crash risks with applications of proactive traffic management strategies (e.g., variable speed limit, ramp metering, etc.)—to reduce the number of crashes (Hossain et al., 2019b). However, these studies investigated only regular traffic conditions where traffic flows show predictable patterns. Although a significant number of studies has been made for crash risk analysis and prevention, studies recognizing and investigating safety issues associated with evacuations hardly exist. No study has investigated the factors which cause an increase in the number of crashes during a hurricane evacuation. In this study, we seek to fill this gap by assessing evacuation crash risks based on real-world hurricane evacuation data.

During an evacuation period, traffic flow does not show any regular time-dependent variations such as heavy traffic during morning and evening peak hours and weekdays. Instead, it shows unpredictable patterns, and most of the time major evacuation routes sustain a heavy demand all day, causing drastic change in speed and unstable traffic stream. Previous studies have found that unstable traffic flow and large variation of traffic speed are two key factors contributing to highway crashes (Golob et al., 2004; Tanishita and van Wee, 2017). However, these studies are limited to the peak hour traffic analysis where traffic congestion is predictable and sustains over a short period of time. Moreover, during hurricane evacuation, a large number of people are forced to reach a safe destination as early as possible causing long-term congestion, in such conditions drivers are more likely to make perception-related errors. In this study, we are conducting a macro-level analysis to understand the influence of different factors on crash risk rather than individual driver-level behavior.

This study seeks to understand the impact of hurricane evacuation on crash risks. We analyze the relationship between traffic state (speed, occupancy, and volume) variations with the likelihood of crash occurrence during hurricane evacuation. We also estimate the impact of these variables on crash risk in a non-evacuation period and compare the result with an evacuation period. To conduct the study, we have collected traffic and crash data for Interstate 75 (I-75) between September 3 and September 16, 2017, including the evacuation period of Hurricane Irma, from the Regional Integrated Transportation Information System (RITIS) database. We have also collected the data for the non-evacuation period from August 1 to August 31, 2017.

First, we analyze evacuation traffic data, to understand the relationship between traffic state variables and crash occurrence at a macroscopic level. Next, we design two case-control studies for non-evacuation period and evacuation period data, and for each case, we estimate the impact of different contributing factors on crash risk. Case-control analysis using conditional logistic regression accounts for the confounding variables. However, if we control all the confounding variables, we are unable to estimate the impact of evacuation on crash risks. Therefore, we also apply a panel mixed binary logistic regression model on the combined dataset (includes both evacuation and non-evacuation period data) to estimate the impact of evacuation on crash risk.

This study helps us understand the factors contributing to highway crashes during evacuation. This study makes two major contributions to the literature. First, it combines data from multiple sources to create a database that helps us gain insights on crash risk of evacuation, using real-world hurricane evacuation data. Second, it reports the influence of evacuation on crash and

finds the relationship between traffic state variables and crash risks. As such, we expect that this study will significantly contribute to the literature and practice by guiding us towards a proactive evacuation management system that will reduce traffic-related incidents during evacuation.

LITERATURE REVIEW

Ensuring safer mobility during hurricane evacuation has become a major concern for emergency managers, especially due to high number of crashes on evacuation routes. Proactive evacuation traffic management can overcome this challenge using real-time traffic monitoring and crash risk assessment tools. Proactive traffic management largely depends on prior detection of the crash risk from real-time data and application of smart strategies (such as variable speed limit and ramp metering) to reduce cash occurrences (Abdel-Aty et al., 2010; Hossain et al., 2019a). However, recent practices in evacuation traffic management mostly focus on evacuation behavior analysis to understand different dimensions of evacuation decisions (Fu and Wilmot, 2004; Moynihan and Fonseca, 2016; Murray-Tuite and Wolshon, 2013; Pel et al., 2012; Sadri et al., 2013; Wong et al., 2018). Although a few studies have explored evacuation traffic behavior, these are limited to understanding operational capacity loss of highways during a hurricane evacuation (Dixit and Wolshon, 2014; Litman, 2006). Understanding the contributing factors of crashes during evacuation is the key to develop crash prediction model for real-time applications. However, there is hardly any study that explores such factors for evacuation period; hence we investigate some of the existing literature to highlight the concept and challenges in crash risk assessment during non-evacuation period.

In traffic safety research, a vast amount of studies have been conducted to understand the key factors contributing towards highway crashes such as driver errors, ambient traffic and environmental condition, and geometric characteristics of the highway segment (Abdel-aty et al., 2004). Among them, most of the early studies focus on post-mortem analysis based on historical accident data and combine them with driver behavior, traffic characteristics, vehicle, and environmental conditions (Chu and Zhang, 2018). These studies provide valuable insights regarding the contributing factors for real-time crash risk prediction. Earlier studies did not consider high-resolution traffic data for real-time crash risk modeling due to lack of such data.

In another direction, researchers have been developing real-time crash prediction models that estimate crash probability on specific road segments using real-time traffic data (Abdel-aty et al., 2004; Abdel-aty and Pande, 2005; Lee et al., 2003; Yu et al., 2014). In recent years, real-time crash prediction methods (*19*, *23–29*) have gained attention due to the widespread deployment of intelligent transportation systems (ITS) technologies and traffic sensors that allow us to obtain large-scale multi-resolution traffic data in real-time (Shi and Abdel-Aty, 2015). Availability of these data sources has encouraged researchers to develop real-time crash prediction models for proactive traffic management applications. One of the challenges for developing such models involves reducing the biases in model results due to the lower number of crash samples (imbalanced sample) compared to the total number of available data samples (Basso et al., 2018; You et al., 2017). To overcome the issue of imbalanced sampling, researchers have adopted a matched case-control logistic regression approach (Abdel-aty et al., 2004).

For instance, using data from multiple loop detectors and a case-control approach, Abdel-Aty et al. (Abdel-aty et al., 2004) found that traffic crashes at a particular location can be predicted based on variations in occupancy and speed at the upstream and downstream of that location. Zheng et al (Zheng et al., 2010) adopted a similar approach to evaluate the impact of traffic oscillations on crash risk. Considering only peak hour traffic data to represent the oscillatory traffic condition (higher traffic demand), they found that in stop-and-go traffic conditions, traffic speed variations significantly increase crash risks. Recent studies have applied synthetic data generation techniques such as Adaptive Boosting (AdaBoost)(Ariannezhad et al., 2021), adaptive synthetic sampling (ADASYN) (You et al., 2017), synthetic minority over-sampling (SMOTE) (Elamrani Abou Elassad et al., 2020) to balance data samples. However, these approaches may not work well when modeling for a special condition such as hurricane evacuation where the number of crashes are significantly lower (<1%) compared to the available data and we need to control the samples' distribution to create a representative dataset. In such condition, matched case-control sampling approach can overcome sampling bias by ensuring specific sampling ratio (crash and non-crash samples).

Although previous studies investigated the relationships between crash occurrence probability and real-time traffic conditions immediately preceding a crash, weather condition, and roadway geometry, these studies considered only regular traffic conditions with predictable traffic patterns (e.g., heavy flow during morning and evening peak hours). On the contrary, during evacuation, traffic demand is significantly higher than regular peak hour traffic and the demand remains higher for all the days starting from the declaration of evacuation orders until the landfall day. Thus, stop-and-go traffic conditions and speed variations are some of the common features of evacuation traffic and may significantly impact crash risks during evacuation. Moreover, evacuation itself could induce some impact on crash occurrences, since during evacuation people are forced to travel to a safe destination as early as possible and are frustrated to drive long hours through heavily congested highways. In such a condition, drivers are more likely to make perception related errors. However, to the best of our knowledge, no studies have investigated crash risk during evacuation period. Consequently, previous studies did not address the challenges of modeling crash likelihood in a rare event such as hurricane evacuation utilizing real-world data. Moreover, existing crash prediction models do not test the impact of any change in a variable on crash risk across two different traffic conditions such as regular condition and evacuation condition.

This study seeks to fill this significant research gap by assessing the crash risk of evacuation traffic based on real-world hurricane evacuation data. We aim to run a macro-level analysis to understand the aggregated influence of all the latent factors related to evacuation on crash risk. Here, we are dealing with a special condition where the number of crash cases are very low (< 1%) compared to total data sample over the study period. Hence, we adopt a similar approach as (Abdel-aty et al., 2004; Zheng et al., 2010) and developed a matched case-control study to assess the safety impact of different traffic state variables during the evacuation period of Hurricane Irma. We run a panel mixed binary logit model to estimate the impact of different variables on crash risk across evacuation and non-evacuation conditions thereby allowing us to

estimate the impact of evacuation on cash risk. We also address the impact of common unobserved factors that affect crash occurrence.

DATA COLLECTION AND ANALYSIS

Data Description and Preprocessing

We have collected traffic data from RITIS, for I-75 northbound direction from September 3, 2017 to September 17, 2017, which includes the evacuation period of Hurricane Irma. To select the study location, we have identified major evacuation routes in Florida and observed that a large portion of residents living in Florida evacuates to Georgia or adjacent states (Roy and Hasan, 2021). Hence, we have chosen the segment between Wildwood and Gainesville (about 50 miles long), which served a major portion of the evacuation traffic during Hurricane Irma. In addition, this segment was highly equipped with MVDS detectors, spaced approximately every 0.5-mile interval. Each detector provides speed, volume, and occupancy data at a very high resolution (every 20 to 30-sec).

We have also collected incident data for the study area from the RITIS incident database. The incident data cover four types of incidents: crash, weather-related incident, congestion, and other regular events (disabled vehicle, road construction related delay, etc.). We map each crash event into its exact location and identify two nearest upstream and downstream MVDS detectors (Figure 1 and Figure 2). From these detectors, we extract traffic speed, volume, and occupancy data for a period of 30 min just before the crash occurred. For example, if a crash occurred on 2 pm, we extract the data from 1:30 to 2:00 pm. Since a few detectors were not functioning during our study period, we could not obtain traffic data from those detectors. Therefore, we discard the crashes corresponding to these detectors. Finally, we create a dataset of 63 crashes during evacuation and extract the traffic data from their corresponding upstream and downstream detectors from September 4 to September 9, 2017. In figure 3, we demonstrate the workflow diagram for data preparation. To compare the traffic characteristics leading to a crash with non-crash traffic characteristics during the evacuation period, we also extract the traffic data which corresponds to a non-crash condition for the same location on the same day. Here, since we are interested to understand the influence of evacuation traffic on crashes, we do not have much flexibility to collect the data for a non-crash condition on different dates/times during the evacuation period. However, during an evacuation period, there is no peaking pattern in traffic flow, that means the time of the day or day of the week will not have any significant impact on traffic flow characteristics.

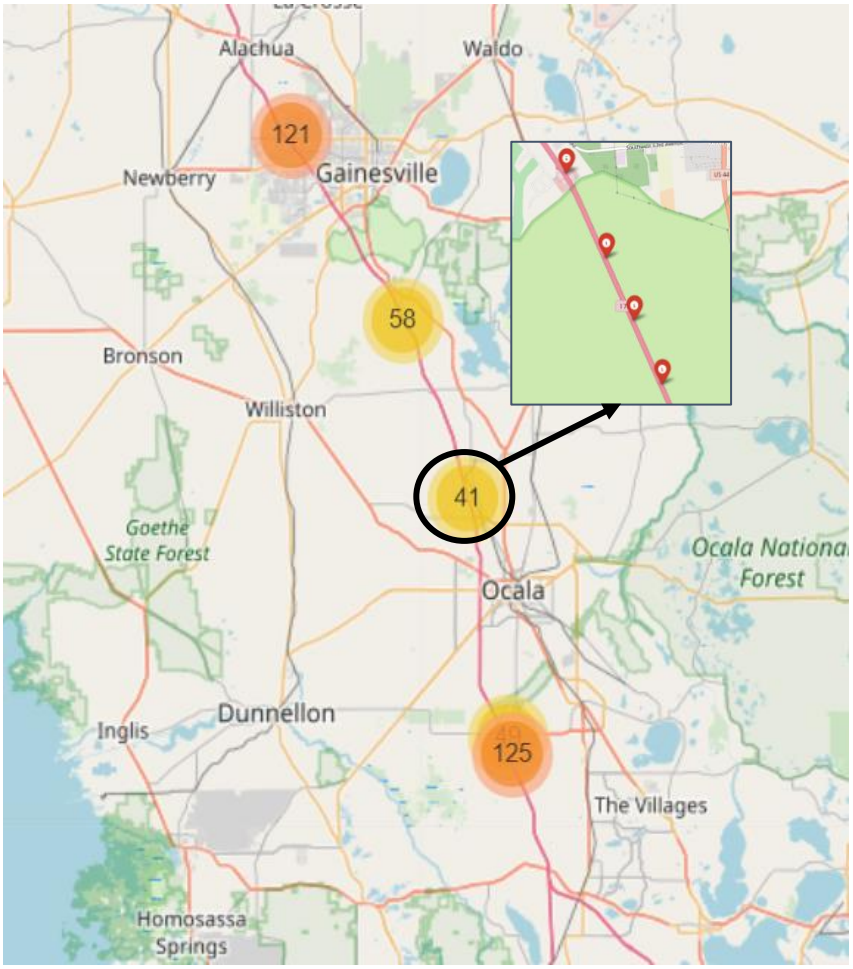


Figure 1: Crash locations aggregated in an Open Street Map based on the actual coordinates of the crashes

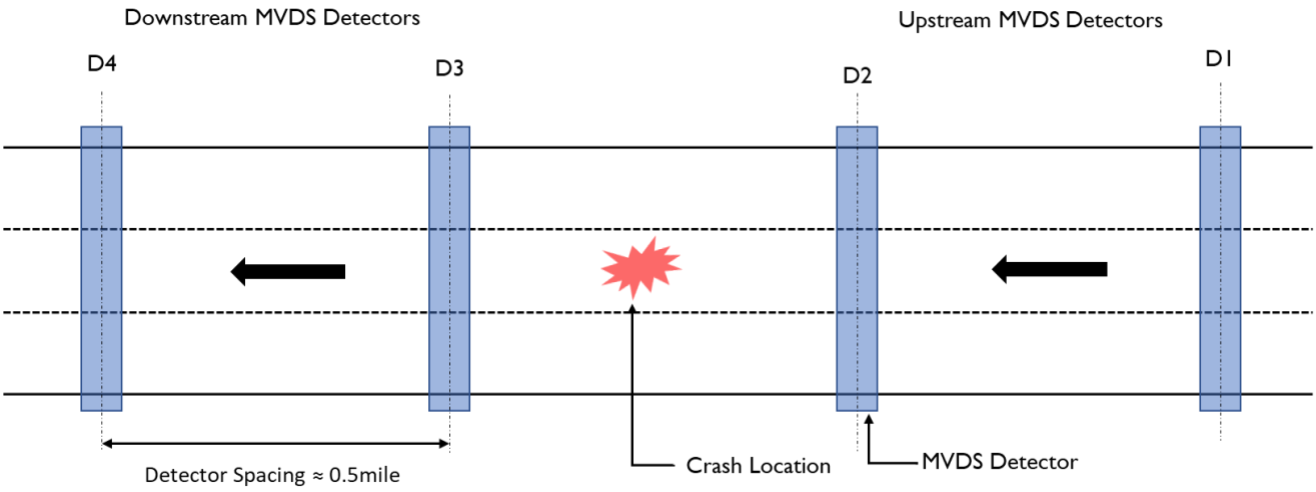


Figure 2: Layout of the segments and MVDS detectors

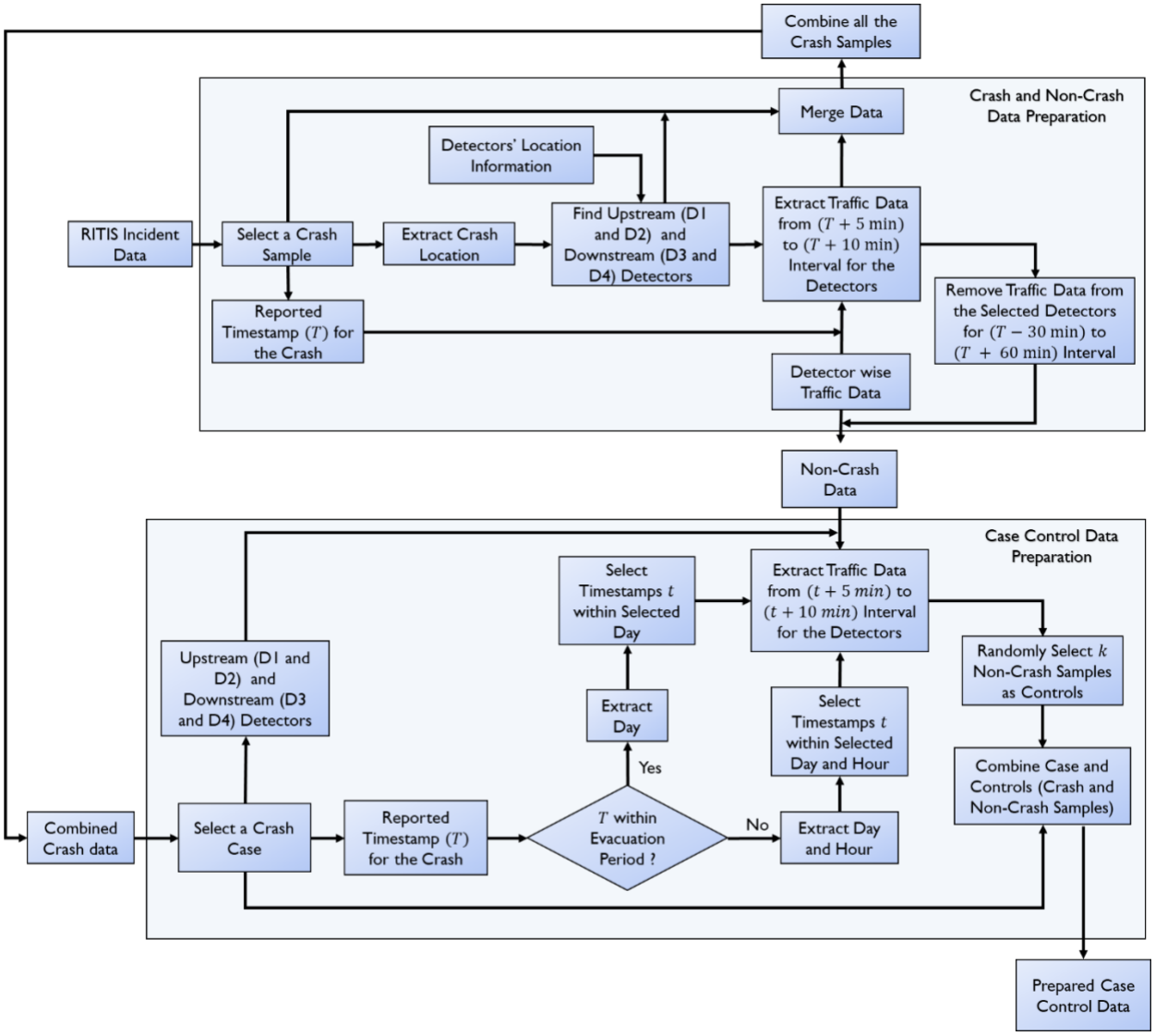


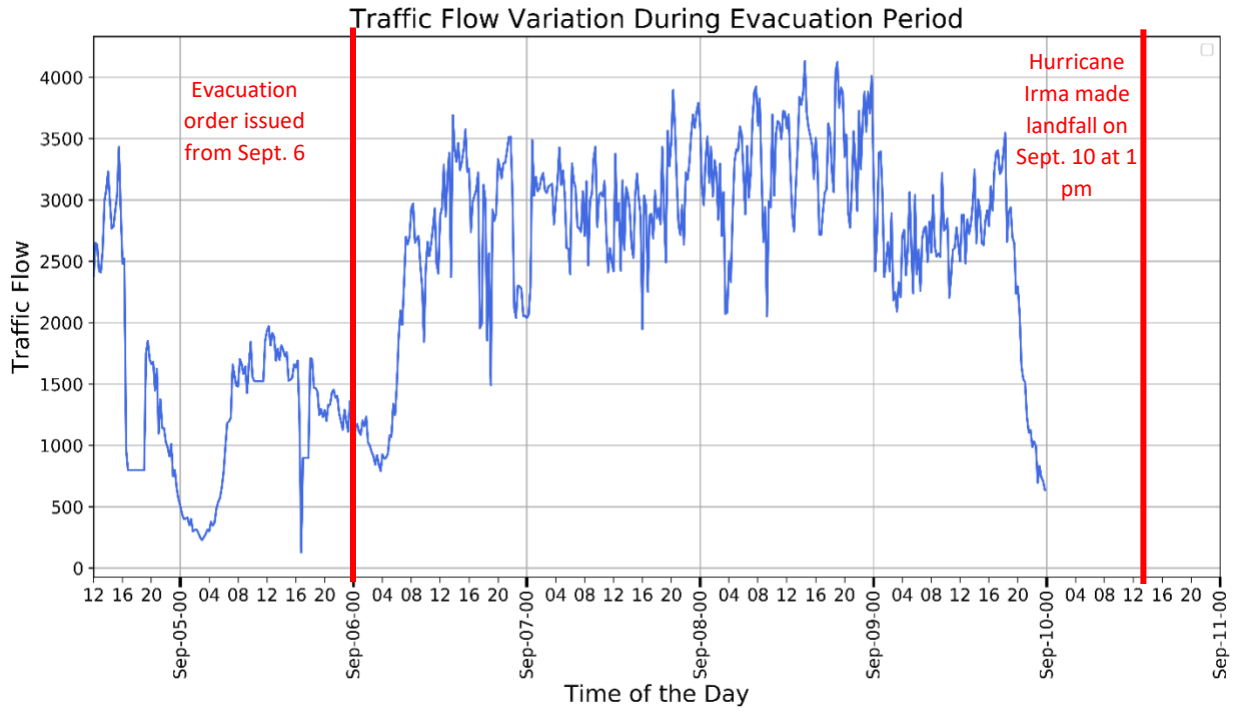
Figure 3: Workflow diagram for data processing

When matching non-crash data sample corresponding to each crash, we discard the traffic data which belongs to the 30 min period just before the crash occurrence and extract the data before that period (see Figure 3). For example, if a crash occurred on 2:00 pm on August 5, 2017, we discard the data from 1:30 to 2:00 pm, and collect the data before 1:30 pm. We have also ensured that there is no overlapping between two consecutive crash conditions in case of multiple crashes. We assume that when a crash occurs it takes at least one hour for the traffic to reach its normal operating condition (Ji et al., 2014; Tirtha et al., 2020; Xie et al., 2015), based on this assumption, we discard the data points which fell within this time period. For example, if two crashes occurred at the same location on 2:00 pm and 5:00 pm on the same day, we just extract the data from 3:00 pm to 4:30 pm as a non-crash condition. The 1hr time period between 2:00 to 3:00 pm is considered as the time required for the traffic to reach normal operating condition; we do not collect any data from this time period. We prepare the dataset in such a way that each matched set (i.e., crash (*1*): non-crash (*k*)) belongs to the evacuation period within the same day. For each matched set, we are controlling for the day and location when matching non-crash traffic observations with a crash

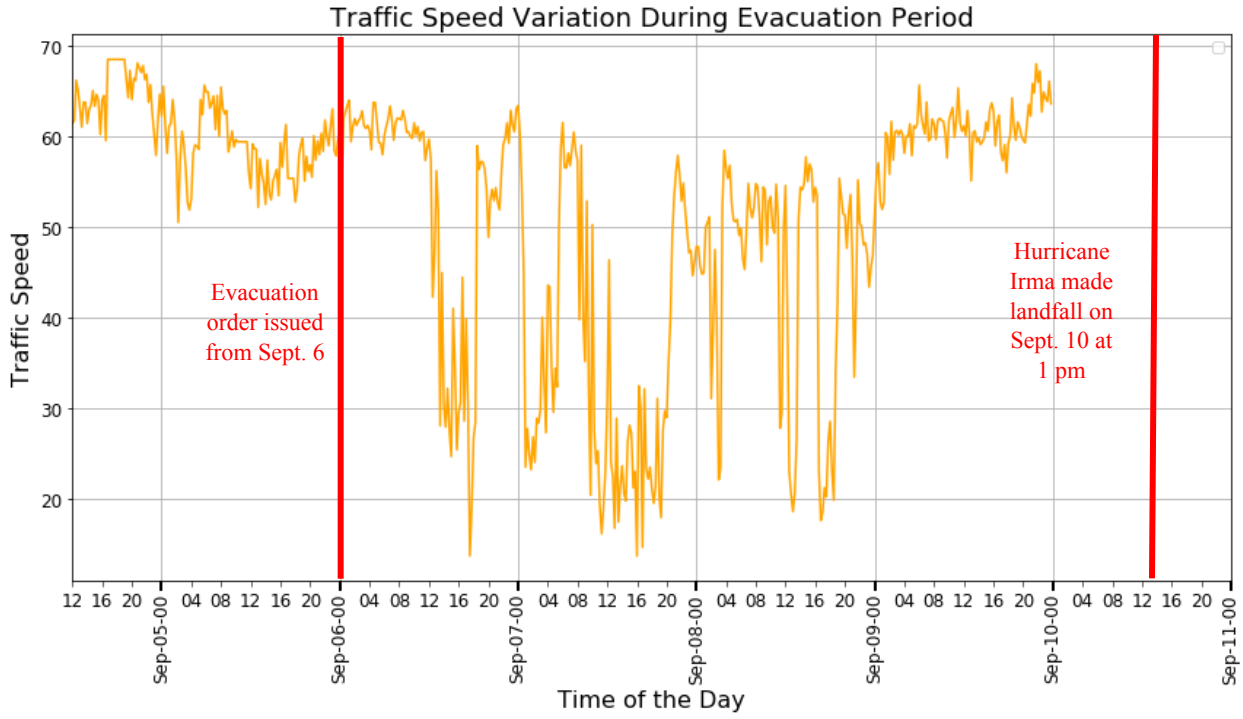
observation. In our final dataset, we have in total 63 crashes, as well as 63 matched non-crash data corresponding to each crash for the evacuation period.

We have also collected crash data for non-evacuation periods at the same locations. In total, we obtain 74 crashes from August 1 to August 31, 2017, and for each of these crashes, we collect traffic data for 30 min periods just before the crash occurrence. In this case, it is likely to have regular patterns in traffic flow variations (morning and evening peak, weekday, weekend, etc.). For matching the traffic data related to a non-crash period with a crash data sample, we have accounted for the time-dependent variations of traffic characteristics. While preparing the data sample for non-crash case corresponding to each crash, we control the location, time of the day (e.g., 3:00 pm to 4:00 pm, 4:00 pm to 5:00 pm, etc.) and day of the week (e.g., Sunday, Monday, etc.) (Figure 3). For example, if a crash occurred on 3:50 pm on August 5, 2017 (Monday) at a particular location then we search all the non-crash data for that location and select the data which correspond to any Monday of that month and within 3:00 to 4:00 pm. Similar to the previous case, we consider pre-crash condition as the 30-min period from the crash occurrence as well as control overlapping of multiple pre-crash conditions (i.e., 1 hr. time period after the crash).

To extract the variables, we divide the sampled traffic data (20 to 30-sec resolution) into 5 min time interval and aggregated them to estimate average speed ($\overline{s}$), standard deviation of speed (ss), coefficient of variation of speed ($cvs = (\overline{s}/ss)$), average volume ($\overline{v}$) and average occupancy ($\overline{oc}$) within that 5-min period. For each of the crashes, we have six-time slices $(1, 2, 3, 4, 5, 6)$, and each having 5 variables defining traffic states from four detectors (two upstream detectors and two downstream detectors). Though we have stated 30-min period before the crash occurrence as the pre-crash condition, previous studies have shown that the traffic characteristics just 5 min before the crash occurred and extending up to 10 min is the most significant period for predicting real-time crash risk (Abdel-Aty et al., 2009; Hossain and Muromachi, 2012). Based on these findings from previous studies, we use pre-crash condition as a 5 min time period ending at least 4 min before the recorded time of the crash. The time of the crash has been reported in the nearest 1 min and the detector data have been aggregated for 5 min. Thus, if a crash had occurred on September 6, 2017, at 4:39 pm then the corresponding pre-crash condition will be traffic data from 4:30 pm to 4:35 pm on that day.



(a)



(b)

Figure 4: Traffic state variation during evacuation (a) traffic flow (b) speed

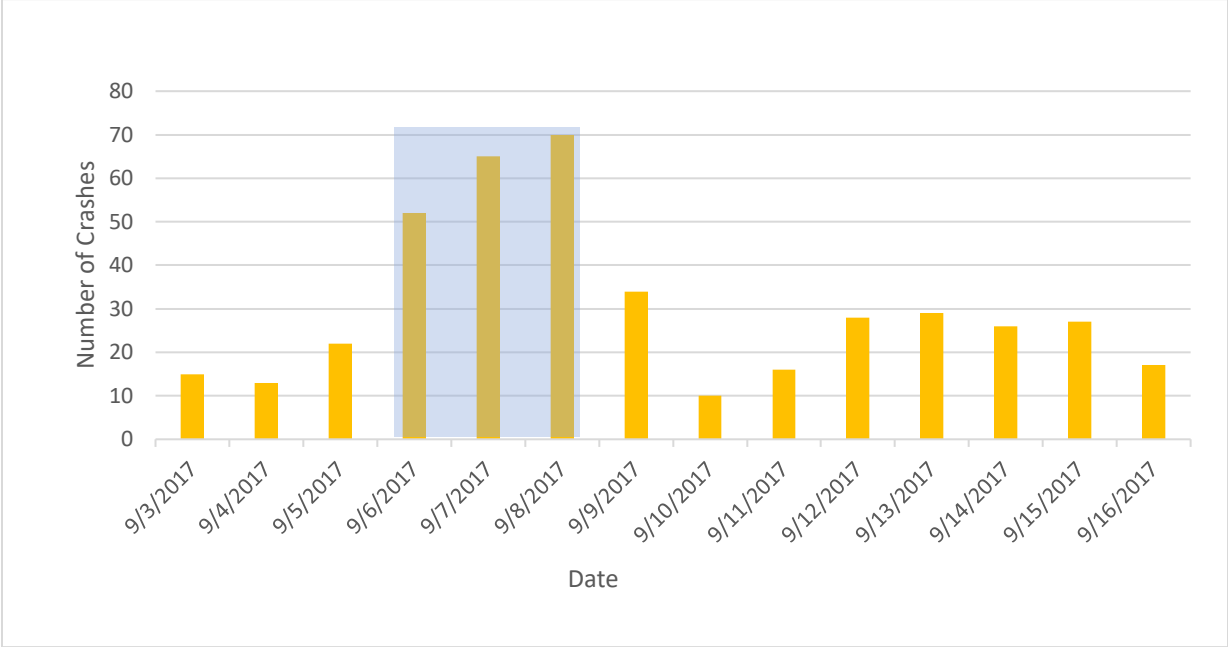


Figure 5: Number of Crashes on Interstate-75 during Hurricane Irma Evacuation (Data Source: RITIS)

Finally, to prepare the dataset for evacuation and non-evacuation period, we form strata of N crashes (i.e., N stratum), where each stratum has one crash as a case and corresponding k (i.e., k=1, 2, 3, 4, 5, etc.) non-crash samples as control. For each crash sample, we randomly choose k non-crash samples (including the traffic variables) corresponding to that crash sample. We control different confounding factors such as location, day of week, hour while matching the crash and non-crash samples.

Descriptive Analysis

In a normal operating condition, traffic flow shows predictable patterns such as heavy demand during peak hours resulting into high traffic flows. However, during an emergency event such as a hurricane evacuation, overall traffic condition has to bear severe disruption due to a drastic increase in traffic demand. Drastic oscillation and sudden flow breakdown are the common characteristics of evacuation traffic. Figure 4(a) shows the distribution of evacuation traffic from September 5, 2017 to September 9, 2017, which demonstrates traffic flow variation during the evacuation period of Hurricane Irma. We observe that during Hurricane Irma's evacuation, overall traffic flow is higher than a regular period and it shows irregular variations which means no distinctive morning or evening peak (see appendix A Figure A.1). Consequently, traffic speed variation is also irregular and significantly lower than normal operating speed (Figure 4(b)). To provide statistical evidence on the difference between evacuation and non- evacuation traffic flow pattern, we perform a two tail T test. From the estimation result, we find the T-Stat value as 22.393 (P-value <0.001) which means that during an evacuation period, traffic flow pattern is significantly different compared to a non-evacuation period.

Moreover, we find that traffic condition starts to deteriorate just after the declaration of evacuation order on September 6, 2017 and remain same till September 9, 2017. Since, we could

not extract any traffic data after September 9, 2017, we are unable to show the traffic flow variation after that time period. Hurricane Irma made its landfall at Florida Keys on September 10, 2017, as a category 4 storm. Then it passed over several regions of Florida from September 10, 2017 to September 12, 2017. It caused significant power outages, in its path, at several regions in Florida. It took about a week to restore the overall system. It is likely that the detectors were malfunctioning, or the data collection server could not retrieve any information during that period.

Figure 5 shows the distribution of crashes for Interstate-75 on different dates during evacuation period. There is a significant increase in the number of crashes on September 6, 7 and 8 which include the evacuation period after the declaration of the state of emergency. From the figures, we can intuitively state that crash occurrence may be associated with traffic flow and speed variation, and the chances of crashes increase if there is a large variation in traffic flow and speed.

METHODOLOGY

Conditional Logistic Regression

We design a matched control study for two separate conditions: evacuation period and regular period to explore the effects of traffic flow variables on crash risk while controlling for the effects of other confounding variables. The differences between crash and non-crash traffic flow characteristics within a stratum are used in the model. This is done under the conditional likelihood principle of statistical theory. For this part, we adopt the methodological approach followed in Aty et al. (Abdel-aty et al., 2004).

In this study, we have *N* strata where each stratum $(i = 1,2, \dots \dots \dots, N)$ has one crash and k non-crash samples$(j = 1,2, \dots \dots \dots, k)$. The conditional likelihood for the ith stratum is the probability of the observed data given the total number of observations and the number of crashes observed in the stratum. Let, $Pr_i(x_{ji})$ be the probability that the jth observation in the ith stratum is a crash where $x_{ji} = (x_{1ji}, x_{2ji}, \dots, \dots x_{pji})$ is the vector of p traffic flow variables $(x_1, x_2, \dots, \dots x_p)$. The probability of a crash can be estimated using a linear in parameter logistic regression model, as follows (Abdel-aty et al., 2004):

$$logit[Pr_i(X_{ji})] = \alpha_i + \beta_1 x_{1ji} + \beta_1 x_{1ji} + \dots\dots + \beta_k x_{pji} \quad (1)$$

The intercept term α captures the effect of controlled variables, which are used to form the strata. To consider the stratification impact into the analysis, conditional log-likelihood can be constructed. The conditional log-likelihood function is the product of N terms, where each term defines the conditional probability of a crash occurrence for a given stratum. The conditional likelihood function can be expressed as follows (Abdel-aty et al., 2004):

$$\mathcal{L}(\beta) = \prod_{i=1}^{N} \left[1 + \sum_{j=1}^{k} exp\left\{ \sum_{u=1}^{p} \beta_u \left(x_{uji} - x_{u0i} \right) \right\} \right]^{-1} \quad (2)$$

The likelihood function $L(\beta)$ in Equation (2) does not contain the intercept terms $\alpha_1, \alpha_2, \dots \dots \alpha_N$. Thus, the effects of matching variables cannot be estimated, and hence Equation 1 cannot be used to estimate crash probabilities. However, the values of the β parameters that maximize the likelihood function (given by Equation 2) are the same as the estimates of β coefficients in Equation 1. These estimates are log–odds ratios and can be used to approximate the

relative risk of a crash. They are also known as hazard ratio (ratio of odds for crash occurrence versus not, i.e., odds ratio). The hazard ratio is defined as e raised to the power of the value of a coefficient (β_u). In this study, we use the survival package in R programming (Therneau, 2020) to estimate the model parameters.

Panel Mixed Binary Logit Model

In this section, we provide the econometric formulation of the proposed panel mixed binary logit model. Let $q(q = 1,2,3, \dots \dots \dots \dots m; m = 6)$ represents the index for different samples for each stratum. With this notation, the formulation takes the following familiar form:

$$v_{iq}^{*} = \{(\alpha + \gamma_{iq})z_{iq} + \varepsilon_{iq} + \varrho_i\}, v_{iq} = 1, if\ v_{iq}^{*} > 0;\ v_{iq} = 0, otherwise \quad (3)$$

where, v_{iq}^{*} represents the propensity for crash occurrence for sample q in stratum i; $v_{iq}^{*}\ is$ 1 if sample specific to a given stratum indicate crash and 0 other wise. z_{iq} is a vector attributes associated with sample q in stratum i and α is the vector of corresponding mean effects. γ_{iq} is a vector of unobserved factors affecting probability of crash occurrence. ε_{iq} is an idiosyncratic error term assumed to be identically and independently standard logistic distributed. ϱ_i is a vector of unobserved effects specific to stratum i. In estimating the model, it is necessary to specify the structure for the unobserved vectors $\gamma\ and\ \varrho$ represented by Ω. In this paper, it is assumed that these elements are drawn from independent normal distribution: $\Omega \sim N(0, (\pi'^{2}, \Phi^{2}))$. Thus, the equation system for modeling the probability of crash takes the following form (conditional on Ω):

$$P_{iq} = p((v_{iq}^{*})|(\Omega)) = \frac{exp\{(\alpha + \gamma_{iq})z_{iq} + \varepsilon_{iq} + \varrho_i\}}{1 + exp\{(\alpha + \gamma_{iq})z_{iq} + \varepsilon_{iq} + \varrho_i\}} \quad (4)$$

The corresponding probability for non-crash is computed as

$$Q_{iq} = 1 - P_{iq} \quad (5)$$

Further, conditional on Ω, the joint probability (for each stratum i) can be expressed as:

$$JP_i = \left[\prod_{i=1}^{N}\{(P_{iq})^{v_{iq}} * (Q_{iq})^{(1-v_{iq})}\}\right]$$

The unconditional probability will be generated by assuming that the parameters in Ω follow a normal distribution (*f(.)*) and is provided as follows:

$$JP_i = \int \left[\prod_{i=1}^{N}\{(P_{iq})^{v_{iq}} * (Q_{iq})^{(1-v_{iq})}\}\right] f(\Omega) d\Omega \quad (7)$$

As the integral defined in Equation (7) cannot be analytically estimated, we employ the maximum simulated estimation approach. The simulated log-likelihood function is evaluated by replacing the integral in Equation (7) with a summation of the function for each realization *r* (*r* = 1, 2 … *R*) as defined below:

$$LL = \sum_i Ln(\frac{\sum_{r=1}^{R} JP_i}{R}) \quad (8)$$

The parameters to be estimated in the model are: $\alpha, \gamma, \varrho, \pi \; and \; \Phi$. To estimate the proposed model, we apply Quasi-Monte Carlo simulation techniques based on the scrambled Halton sequence with R set to 150 (see (Bhat, 2001; Eluru et al., 2008) for examples of Quasi-Monte Carlo approaches in literature). We tested the model with higher R values and found that the model estimation was stable.

We estimate this model using GAUSS matrix programming language. We code and optimize the log-likelihood function using the non-linear optimization routes within GAUSS. The code we employ in this paper has been tested over multiple other research contexts.

MODEL RESULTS

In our dataset, for each of the two upstream and two downstream detectors, we have 4 explanatory variables: 5-min aggregated mean values of occupancy ($\overline{oc}$), volume ($\overline{v}$), speed ($\overline{s}$), and standard deviation of speed (ss). So, in total, each dataset contains 16 variables. Previous studies (Abdel-aty et al., 2004; Lee et al., 2003) found that the coefficient of variation of speed (cvs) better captures the effects of speed and speed variation on crash risk. We combine the standard deviation of speed and mean speed to obtain cvs, which reduces the number of explanatory variables into 3. Now, we have 12 variables associated with four detectors for each data set. In table 1, we include descriptive statistics for all the variables.

We estimate the Pearson correlation coefficients between different pairs of variables which show that volume is highly correlated with occupancy. Therefore, in our final model, we use either occupancy or volume. Moreover, it appears that in some cases, the same variable (e.g., occupancy) over different detectors are also correlated with each other (Figure 6). To avoid using highly correlated exogenous variable (i.e., speed, volume, etc.) from multiple detectors, we have decided to use its value observed in one detector (instead of multiple detectors). We select these variables based on the hazard ratio and corresponding t statistics (T-Stat).

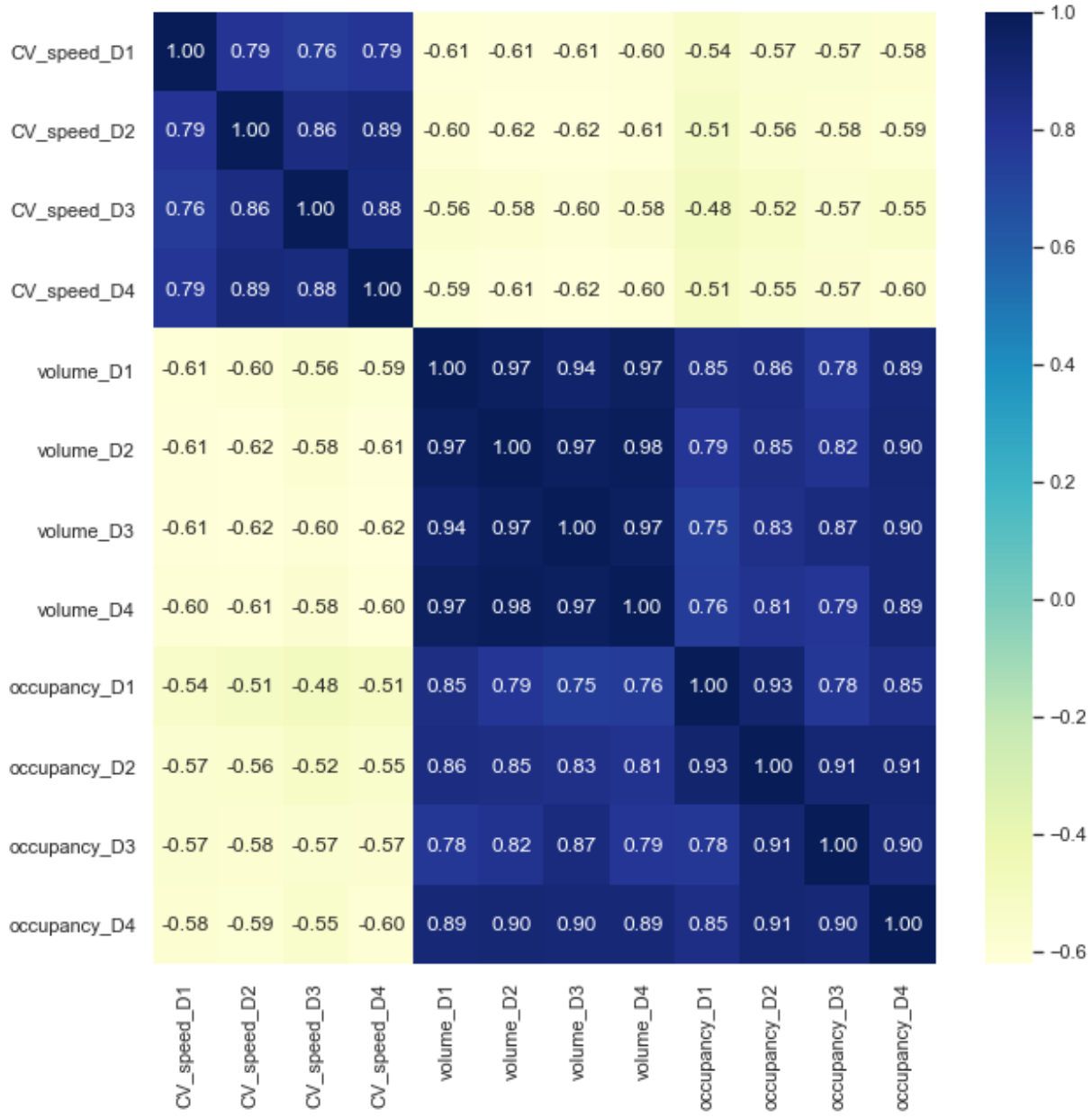


Figure 6: Pearson correlation values for different pairs of variables

TABLE 1: Descriptive statistics of all variables used in the models.

Description	Min	Max	Mean	Std. Deviation
Dependent variable				
Crash (1 if crash occurred, 0 otherwise)	0	1	0.167	0.373
Independent variables				
Upstream detector D1				
Coefficient of variation of speed (mph/mph)	0.012	1.409	0.113	0.134
Occupancy (sec/sec)	0.222	44.2	8.657	8.664
Volume (veh/5 min)	11	412	181.183	92.823
Upstream detector D2				
Coefficient of variation of speed (mph/mph)	0.005	0.971	0.108	0.139
Occupancy (sec/sec)	0.333	45.967	8.948	8.311
Volume (veh/5 min)	15	409	189.158	92.912
Downstream detector D3				
Coefficient of variation of speed (mph/mph)	0.011	0.973	0.107	.119
Occupancy (sec/sec)	0.178	41.756	9.055	8.139
Volume (veh/5 min)	15	388	187.854	93.583
Downstream detector D4				
Coefficient of variation of speed (mph/mph)	0.015	2.071	0.134	0.218
Occupancy (sec/sec)	0.133	47.222	8.324	7.943
Volume (veh/5 min)	3	423	184.645	95.268
Evacuation (1 for evacuation period data sample, 0 otherwise)	0	1	0.459	0.499

In our matched case-control sample, we have N strata (depends on the number of crashes. e.g., 74 for non-evacuation period) and each stratum has one crash and corresponding k non-crashes. The number of non-crash samples varies from 1 to k. To fix the number of controls (k) in a stratum we run the conditional logistic regression model with different number of control samples and check the estimates from each model. We find no significant changes in model estimates after k=5, hence, we use 5 control samples in our final matched case-control sample (see Appendix B for details). We apply this approach for both evacuation and non-evacuation periods and obtain similar results.

First, we run the conditional logistic regression model for evacuation and non-evacuation condition to understand the influence of the exogenous variables on crash risk for different traffic demand condition. Table 2 presents the final results for each model. Under regular condition, the final models include two variables: mean occupancy and coefficient of variation of speed at D1. Both variables have positive coefficient with hazard ratio greater than 1 (positive coefficient), indicating that the odds of a crash increase with the increase of these variables. Moreover, the mean occupancy variable associated with the upstream detector D1 is significant at 95% confidence interval while the other variable is significant at a 90% confidence interval. These estimates indicate if there is high occupancy of vehicles and large variation of speed at the upstream 5 to 10 min before the crash, the chance of a crash occurrence increases. Since the coefficient of variation of speed includes the average speed as the denominator, this also indicates that the average speed is lower in crash cases.

Similarly, under an evacuation condition, the final models have two significant variables coefficient of variation of speed for the downstream detector D3 and mean volume for the upstream detector D1. The value of hazard ratio for both of these variables is greater than 1 (positive coefficient), which means if there is a high volume of traffic at upstream and high variation of speed at downstream then the chances of crash occurrence is higher. Moreover, we can interpret the combined effects of these variables that higher volume of traffic at the upstream location coupled with high variation in the speed at the downstream location, increase the likelihood of crash occurrence, at a location in between these two zones. The detectors D1 at the upstream zone and D3 at the downstream zone are spaced 1 mile (approximately), that means during evacuation period this 1-mile segment experience high-speed variation, the high volume of traffic, and lower average speed, which indicates potential queue formation under oscillatory speed conditions. Consequently, this would have caused a significant increase in the number of crashes within this segment.

TABLE 2: Estimates for the final models for evacuation and regular period (sampling ratio for crash and non-crash is 1:5)

Traffic Condition	Variables	Conditional Logistic Regression		
		Coeff. **(Hazard Ratio)**	T-Stat	P value
Non-Evacuation Period	cvs_D1	2.69 (**14.74)**	2.76	0.0059
	Occupancy_D1	0.0628 (**1.0648)**	1.72	0.0858
Evacuation Period	cvs_D3	1.251 **(3.494)**	1.96	0.0503
	Volume_D1	0.0052 **(1.004)**	2.06	0.0397

In our analysis, we combine data from both evacuation and non-evacuation periods. We add a dummy indicator variable "Evc" (0 or 1) to separate the evacuation period data from the non-evacuation period data. This variable would indicate the impact of several latent factors associated with evacuation on crash risk. We have collected the non-evacuation period crash data from a 30-day data sample while for evacuation period we only have 6 days data sample. Therefore, before combining the data from two different groups, we need to address this imbalanced sampling. To account for the imbalance in the sampling process, we assign frequency basis weight (Johnson, 2008; Wicklin, 2017) on each crash case. We put a weight of $(74/30)$ or 2.5 with non-evacuation related crash and a weight of $(63/6)$ or 10.5 with evacuation related crash. Which means each evacuation related crash has a weight of $10.5/2.5 \cong 4$ while each of the non-evacuation related crash has a weight of 1. Since we are controlling all the confounding variables, we are unable to estimate Evc as well as other unobserved factors that affect crash occurrence during evacuation.

Hence, we apply a Panel Mixed Binary Logit model mainly for two reasons. First, the model allows us to pool the evacuation and non-evacuation cases in a single model allowing us to test if the impact of any variables changes across the two cases. Second, we estimate a panel mixed model to recognize the repeated observations of records at the same location. These repeated observations are likely to have common unobserved factors that affect crash occurrence. Ignoring the presence of such factors when present can result in incorrect or biased estimates. However, we implement the model just to understand the impact of evacuation on crash risk. Since matched case control sample does not provide true estimate of the constant term, we cannot use such a model for crash prediction. We apply similar approach as the conditional logistic regression model to fix the number of control samples (k) within each stratum, and employ the same value for k=5, there is no significant variations in the model estimates increasing the k value over 5 (see Appendix B).

Table 3 presents the estimates for Panel Mixed Binary Logit model. From the model estimate, we find that the coefficient for the variable coefficient of variation of speed for detector D3 (cvs_D3) is positive and significant at 95% confidence interval, that means the chances of crash occurrence increases with the increase in speed variations at the downstream location of a roadway segment. The coefficient for mean volume (volume_D1) variable is also positive which means with increase in traffic volume at the upstream location D1 increases the chances of crash occurrence. We also find that the variable Evc is highly significant and the coefficient associated with this variable is positive indicating that during evacuation the chances of crash occurrence is higher than in the non-evacuation period.

TABLE 3: Model estimates for the combined datasets including both evacuation and non-evacuation periods (sampling ratio for crash to non-crash events is 1:5)

Variables	Independent LR			Mixed LR		
	Coeff.	T-Stat	P value	Coeff.	T-Stat	P value
Constant	-2.019	-9.500	< 0.001	-2.078	-9.500	< 0.001
Volume_D1	0.002	2.281	0.0228	0.003	2.281	0.0228

cvs_D3	1.132	2.233	0.0258	1.271	2.233	0.0258
Evacuation	1.029	3.395	0.0007	0.909	3.395	0.0007
Standard deviation	--	--	--	0.412	1.760	0.078
Log-likelihood	-585.141			-581.971		

Several unobserved factors can potentially influence crash occurrence process. We tested for the influence of roadway specific unobserved factors, temporal unobserved factors, and presence of variation among estimated parameters (i.e., random parameters). In our estimation, only one random parameter offered statistically significant effect for the Evacuation variable.

The results highlight how crash propensity varies under evacuation conditions. The mean parameter is positive highlighting how the crash risk is higher during evacuation relative to non-evacuation time periods. However, it is also important to recognize that the increase in crash risk is not fixed and exhibits significant heterogeneity. The crash risk based on unobserved factors can actually be much higher or lower than the mean effect estimated. The actual probability of the crash risk will depend on the specific factors affecting the process. The result indicates that the probability of crash risk in an evacuation period is substantially higher than that of in non-evacuation periods (0.9863). Furthermore, based on the normal distribution, the 90th percentile, 75th percentile, 25th percentile, and 10th percentile values for the parameter are: 1.437, 1.187, 0.631, and 0.381. The distributional values further reinforce how the crash risk during evacuation is higher than the corresponding crash risk for non-evacuation periods. However, it is important to note that the results presented here are arrived at using one dataset only. Further investigation of data from multiple evacuation scenarios is necessary to generalize this finding and to determine how much higher of a crash risk an evacuation period is relative to a non-evacuation period.

DISCUSSION AND CONCLUSION

This study reveals traffic flow characteristics during hurricane evacuation using real-world data from Hurricane Irma's evacuation period. As expected, it shows that during evacuation overall traffic demand is higher than the regular traffic condition which causes irregular variation of traffic flow. Consequently, it leads to significant variations in traffic speed, resulting into a stop-and-go traffic situation.

Adopting a case-control analysis, we find that during evacuation the coefficient of variation of speed at the downstream station and average occupancy at the upstream station of a crash location significantly affect crash likelihood. This implies that higher occupancy rates at upstream coupled with high variation in speed at downstream locations, increase the likelihood of crash occurrence. Moreover, a panel mixed binary logit model applied over combined (including both evacuation and non-evacuation period) data showed that evacuation itself increases the chance of a crash occurrence, even after we account for traffic characteristics. This indicates that during evacuation the likelihood of crash occurrence increases compared to the regular period.

Although the implemented model cannot be directly used for crash prediction, the insights from this model will help us develop a crash prediction model which will work for evacuation traffic, and consequently, proactive measures can be developed to reduce crash risks during an emergency situation. Particularly, this method will help identify potential crash locations created by prevailing traffic conditions during an evacuation. This can be used to warn evacuee drivers about the impending crash risk and enforce them to reduce travel speed to a certain limit.

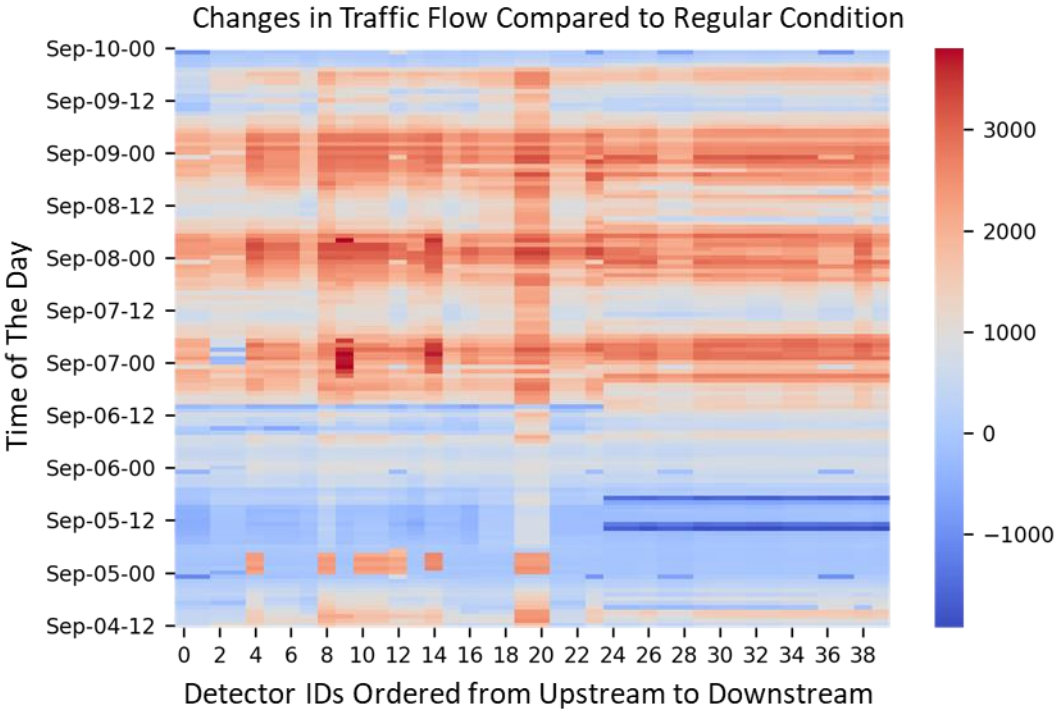
The study has further implications for evacuation declarations. Our result shows that high volume and occupancy of traffic during evacuation are key contributing factors for crashes. If the volume of traffic on the evacuation routes can be reduced, the chances of crash occurrence will significantly decrease. However, during evacuation, the traffic demand surge occurs just after the declaration of evacuation order due to evacuation of a large number of people at the same time from different zones. Therefore, one potential strategy should be to adopt a phased declaration of evacuation orders, which require identification of primary risk zones based on spatial and temporal information on hurricane landfall. Evacuation orders should be declared in a phased manner starting with the primary risk zone and then other zones based on potential hurricane threat.

The study also show that traffic speed variation causes a significant increase in the number of crashes during evacuation, which means we need adequate strategies to reduce the abrupt speed variation. Apart from infrastructure-based strategies such as variable speed limit and emergency shoulder use, in-vehicle controls using adaptive cruise control and cooperative cruise control, connected vehicles and vehicle platooning can play a vital role as well (Rahman et al., 2021). These are more proactive technologies which can assist drivers to maintain a constant cruising speed and gap, reducing overall speed variation. These technologies need to be field-tested for evacuation traffic to understand their impact on crash risk. In near future, however, microscopic simulation experiments can be conducted to understand the impact of different in-vehicle control systems on overall crash reduction during evacuation.

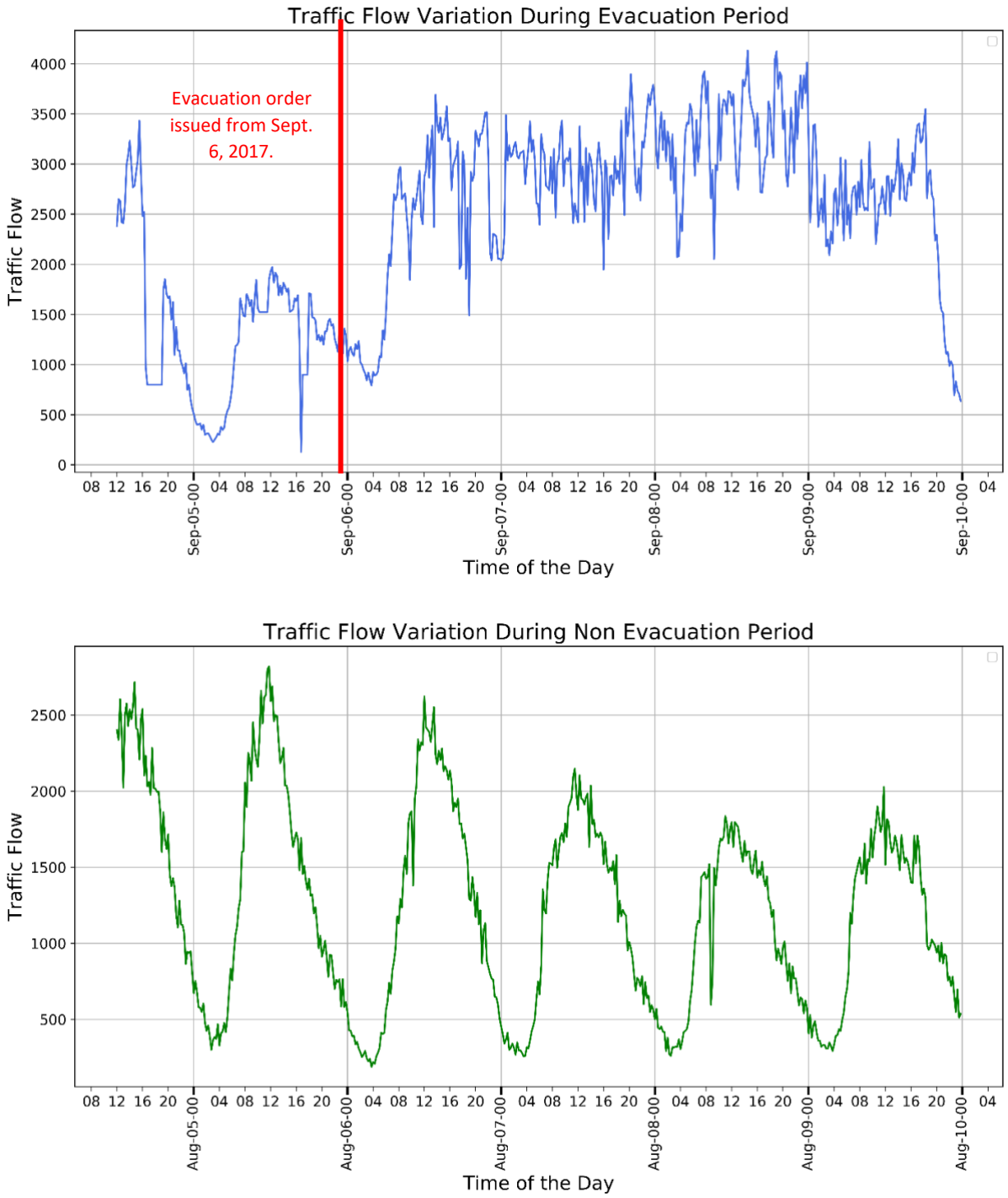
Although this study provides some key insights on crash risk during Hurricane Irma evacuation, it has several limitations. First, we have analyzed the data from one evacuation period only; thus, the conclusions drawn from this study cannot be generalized. We need crash data from multiple hurricanes to generalize these findings. Also, the transferability of this model in future hurricane evacuation scenarios needs to be tested. Another limitation is that we do not have enough data sample from off-peak hours of non-evacuation period (See Appendix A Figure A.2), thereby we are unable to compare crash risks in evacuation period with that of peak and off-peak hours during non-evacuation period. Future research may address these limitations.

APPENDIX A. Comparison between evacuation and non-evacuation period traffic flow.

To compare the traffic flow variation in a spatiotemporal resolution, we calculate the average hourly traffic volume of all the detector for regular period over different hours (1 to 24 hr.) and days (weekdays and weekends). Finally, we calculate the difference between hourly traffic volume during evacuation period and non-evacuation period. Figure A.1 shows the difference between evacuation traffic and regular traffic for the study location. From the figure, we find that during Hurricane Irma's evacuation from September 6 to September 8, 2017, overall traffic flow is higher all the time regardless of whether it is a peak hour or not; there is a significant amount of traffic congestion on the major interstates even after peak hours.



(a) Spatiotemporal evacuation traffic flow variation compared to average non-evacuation traffic flow



(b) Traffic flow variation evacuation and non-evacuation period for individual detector

Figure A.1: Comparison between evacuation and non-evacuation period traffic flow variations

Figure A.2 demonstrates the hourly variation of traffic crashes for evacuation period and non-evacuation period. In the figure, we also include the hourly variation of average traffic combining all the detectors, which guides us to differentiate between peak hour and off-peak hour traffic.

From the figure, we find that in regular period, the majority (60%) of the crashes occur during peak hours (i.e., 12 pm to 7 pm).

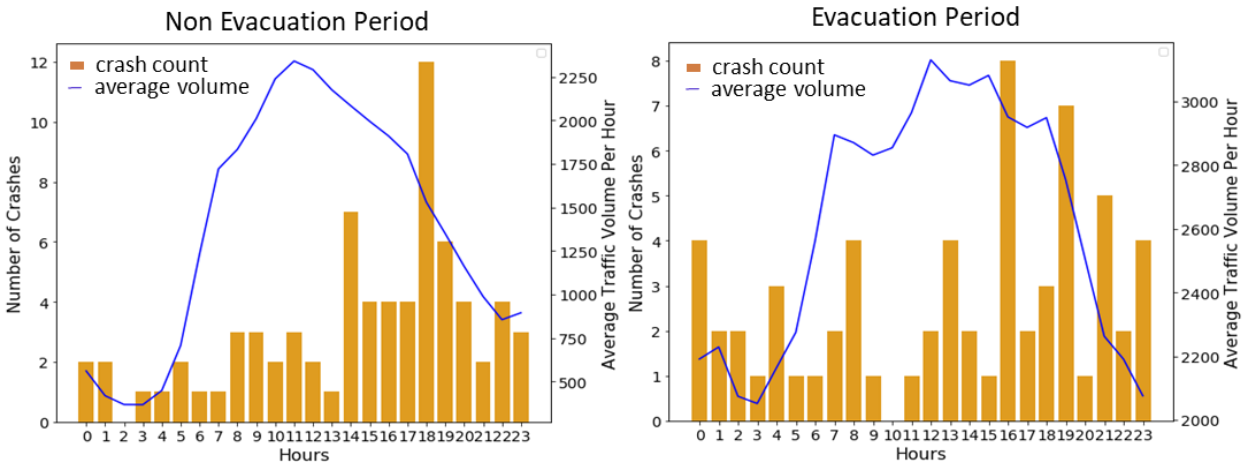


Figure A.2: Comparison between evacuation and non-evacuation period traffic volume and crashes at different time (hours of the day)

Appendix B. Selection procedure for optimal number of controls (k) for MBL

The idea is to start with crash sample and corresponding non-crash samples ratio as 1:1 and increase the ratio gradually (1:3, 1:5, 1:7, 1:9….) until the difference between the coefficients across the two successive models does not change significantly. As indicated by the figure, we can clearly see the model with samples ratio 1:5 and 1:6 do not depict any significant differences and hence, we select the 1:5 ratio for our analysis. We also look at the log-likelihood improvement across the models and the results further reinforces our hypothesis of selecting crash sample and non-crash samples ratio as the 1:5. The differences across the model coefficients for each variable is calculated following the formula stated in (Clogg, 1995).

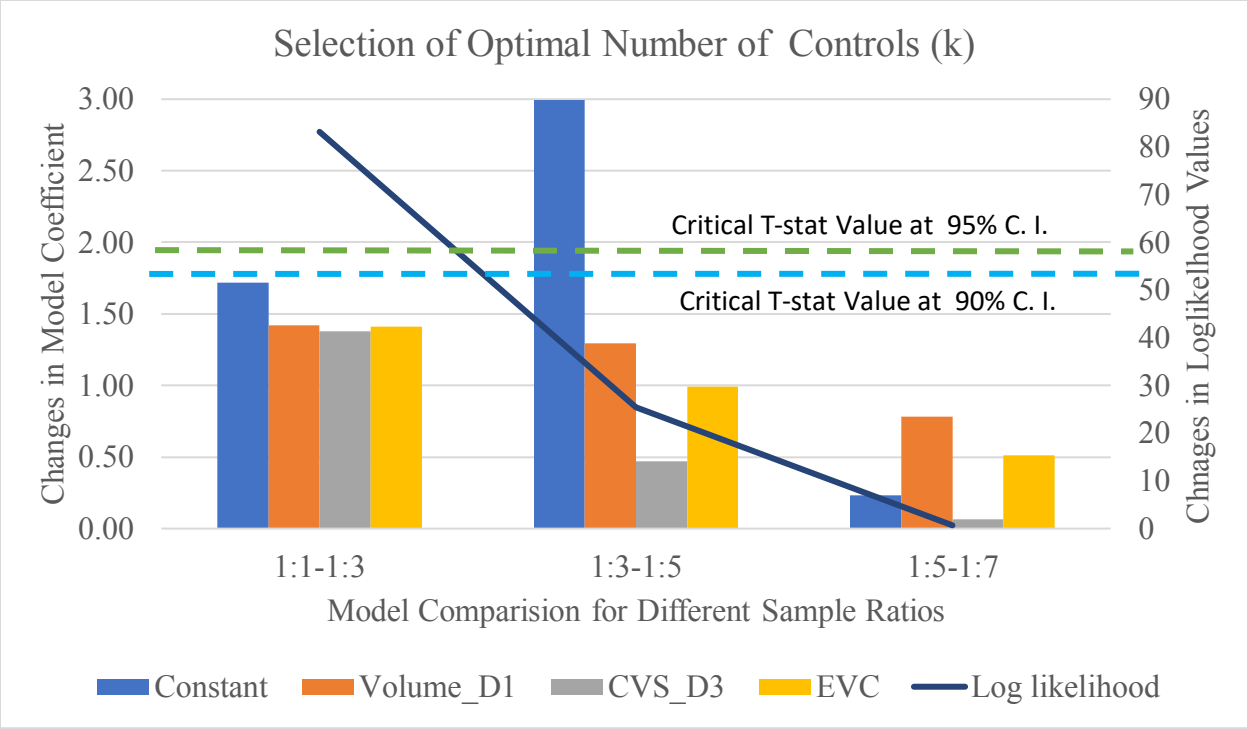


Figure B.1: Selection of optimal number of controls (k)

AUTHOR CONTRIBUTIONS

The authors confirm contribution to the paper as follows: study conception and design: R. Rahman, T. Bhowmik, S. Hasan, N. Eluru; data analysis and interpretation of results: R. Rahman, T. Bhowmik; draft manuscript preparation: R. Rahman, T. Bhowmik, S. Hasan, N. Eluru; All authors reviewed the results and approved the final version of the manuscript.

ACKNOWLEDGMENT

The research presented in this paper is supported partially by Safety Research using Simulation University Transportation Center (SAFER-SIM) and partially by the U.S. National Science Foundation grant CMMI-1917019. SAFER-SIM is funded by a grant from the U.S. Department of Transportation's University Transportation Centers Program (69A3551747131). However, the U.S. Government assumes no liability for the contents or use thereof.

REFERENCES

Abdel-aty, M., Pande, A., 2005. Identifying crash propensity using specific traffic speed conditions 36, 97–108. https://doi.org/10.1016/j.jsr.2004.11.002

Abdel-Aty, M., Pande, A., Das, A., Knibbe, W.J., 2009. Assessing Safety on Dutch Freeways

with Data from Infrastructure-Based Intelligent Transportation Systems. Transp. Res. Rec. J. Transp. Res. Board 2083, 153–161. https://doi.org/10.3141/2083-18

Abdel-Aty, M., Pande, A., Hsia, L., 2010. The concept of proactive traffic management for enhancing freeway safety and operation. ITE J. (Institute Transp. Eng. 80, 34–41.

Abdel-aty, M., Uddin, N., Pande, A., Abdalla, M.F., Hsia, L., 2004. Predicting Freeway Crashes from Loop Detector Data by Matched Case-Control Logistic Regression 88–95.

Ariannezhad, A., Karimpour, A., Qin, X., Wu, Y.-J., Salmani, Y., 2021. Handling Imbalanced Data for Real-Time Crash Prediction: Application of Boosting and Sampling Techniques. J. Transp. Eng. Part A Syst. 147, 04020165. https://doi.org/10.1061/jtepbs.0000499

Basso, F., Basso, L.J., Bravo, F., Pezoa, R., 2018. Real-time crash prediction in an urban expressway using disaggregated data. Transp. Res. Part C Emerg. Technol. 86, 202–219. https://doi.org/10.1016/j.trc.2017.11.014

Bhat, C.R., 2001. Quasi-random maximum simulated likelihood estimation of the mixed multinomial logit model. Transp. Res. Part B Methodol. 35, 677–693. https://doi.org/10.1016/S0191-2615(00)00014-X

Chu, W., Zhang, H., 2018. Real-Time Crash Prediction Estimation of Freeway Safety: A Review, in: CICTP 2017: Transportation Reform and Change—Equity, Inclusiveness, Sharing, and Innovation. American Society of Civil Engineers, Reston, VA, pp. 4475–4482.

Clogg, C.C., 1995. Statistical Methods for Comparing Regression Coefficients Between Models Author (s): Clifford C . Clogg , Eva Petkova and Adamantios Haritou Published by : The University of Chicago Press Stable URL : http://www.jstor.com/stable/2782277 Statistical Meth 100, 1261–1293.

Dixit, V., Wolshon, B., 2014. Evacuation traffic dynamics. Transp. Res. Part C Emerg. Technol. 49, 114–125. https://doi.org/10.1016/j.trc.2014.10.014

Elamrani Abou Elassad, Z., Mousannif, H., Al Moatassime, H., 2020. A proactive decision support system for predicting traffic crash events: A critical analysis of imbalanced class distribution. Knowledge-Based Syst. 205, 106314. https://doi.org/10.1016/j.knosys.2020.106314

Eluru, N., Bhat, C.R., Hensher, D.A., 2008. A mixed generalized ordered response model for examining pedestrian and bicyclist injury severity level in traffic crashes. Accid. Anal. Prev. 40, 1033–1054. https://doi.org/10.1016/j.aap.2007.11.010

Fu, H., Wilmot, C.G., 2004. Sequential logit dynamic travel demand model for hurricane evacuation. Transp. Res. Rec. 19–26. https://doi.org/10.3141/1882-03

Golob, T.F., Recker, W.W., Alvarez, V.M., 2004. Freeway safety as a function of traffic flow. Accid. Anal. Prev. 36, 933–946. https://doi.org/10.1016/j.aap.2003.09.006

Hossain, M., Abdel-aty, M., Quddus, M.A., Muromachi, Y., 2019a. Real-time crash prediction models : State-of-the-art , design pathways and ubiquitous requirements. Accid. Anal. Prev. 124, 66–84. https://doi.org/10.1016/j.aap.2018.12.022

Hossain, M., Abdel-Aty, M., Quddus, M.A., Muromachi, Y., Sadeek, S.N., 2019b. Real-time crash prediction models: State-of-the-art, design pathways and ubiquitous requirements. Accid. Anal. Prev. 124, 66–84. https://doi.org/10.1016/j.aap.2018.12.022

Hossain, M., Muromachi, Y., 2012. A Bayesian network based framework for real-time crash prediction on the basic freeway segments of urban expressways. Accid. Anal. Prev. 45, 373–381. https://doi.org/10.1016/j.aap.2011.08.004

House of Representatives Florida, 2018. Select Comittee on Hurricane Response & Preparedness Final Report [WWW Document]. URL https://www.myfloridahouse.gov/Sections/Documents/loaddoc.aspx?PublicationType=Committees&CommitteeId=2978&Session=2018&DocumentType=General Publications&FileName=SCHRP - Final Report online.pdf

Ji, Y.B., Jiang, R., Qu, M., Chung, E., 2014. Traffic incident clearance time and arrival time prediction based on hazard models. Math. Probl. Eng. 2014. https://doi.org/10.1155/2014/508039

Johnson, D.R., 2008. Using weights in the analysis of survey data. Present. Prep. Popul. Res. Institute, Pennsylvania State Univ. Novemb.

Lee, C., Hellinga, B., Saccomanno, F., 2003. Real-Time Crash Prediction Model for Application to Crash Prevention in Freeway Traffic. Transp. Res. Rec. J. Transp. Res. Board 1840, 67–77. https://doi.org/10.3141/1840-08

Litman, T., 2006. Lessons from Katrina and Rita: What major disasters can teach transportation planners. J. Transp. Eng. 132, 11–18. https://doi.org/doi.org/10.1061/(ASCE)0733-947X(2006)132:1(11)

Moynihan, G.P., Fonseca, D.J., 2016. Design of an Evacuation Demand Forecasting Module for Hurricane Planning Applications. J. Transp. Technol. 06, 257–276. https://doi.org/10.4236/jtts.2016.65024

Murray-Tuite, P., Wolshon, B., 2013. Evacuation transportation modeling: An overview of research, development, and practice. Transp. Res. Part C Emerg. Technol. 27, 25–45. https://doi.org/10.1016/j.trc.2012.11.005

Murray-Tuite, P., Wolshon, B., Matherly, D., 2017. Evacuation and Emergency Transportation: Techniques and Strategis for Systems Resilience. TR NEWS 311.

Pel, A.J., Bliemer, M.C.J., Hoogendoorn, S.P., 2012. A review on travel behaviour modelling in dynamic traffic simulation models for evacuations. Transportation (Amst). 39, 97–123. https://doi.org/10.1007/s11116-011-9320-6

Rahman, R., Hasan, S., Zaki, M.H., 2021. Towards reducing the number of crashes during hurricane evacuation: Assessing the potential safety impact of adaptive cruise control systems. Transp. Res. Part C Emerg. Technol. 128, 103188. https://doi.org/10.1016/j.trc.2021.103188

Roy, K.C., Hasan, S., 2021. Modeling the dynamics of hurricane evacuation decisions from twitter data: An input output hidden markov modeling approach. Transp. Res. Part C Emerg. Technol. 123, 102976. https://doi.org/10.1016/j.trc.2021.102976

Sadri, A.M., Ukkusuri, S. V., Murray-Tuite, P., 2013. A random parameter ordered probit model to understand the mobilization time during hurricane evacuation. Transp. Res. Part C Emerg. Technol. 32, 21–30.

Shi, Q., Abdel-Aty, M., 2015. Big Data applications in real-time traffic operation and safety monitoring and improvement on urban expressways. Transp. Res. Part C Emerg. Technol. 58, 380–394. https://doi.org/10.1016/j.trc.2015.02.022

Tanishita, M., van Wee, B., 2017. Impact of vehicle speeds and changes in mean speeds on per vehicle-kilometer traffic accident rates in Japan. IATSS Res. 41, 107–112. https://doi.org/10.1016/j.iatssr.2016.09.003

Therneau, T.M., 2020. A package for survival analyis in R [WWW Document]. URL https://cran.r-project.org/package=survival

Tirtha, S.D., Yasmin, S., Eluru, N., 2020. Modeling of incident type and incident duration using data from multiple years. Anal. Methods Accid. Res. 28, 100132. https://doi.org/10.1016/j.amar.2020.100132

Wicklin, R., 2017. How to understand weight variables in statistical analyses [WWW Document].

Wong, S., Shaheen, S., Walker, J., 2018. Understanding evacuee behavior: A case study of hurricane Irma. https://doi.org/10.7922/G2FJ2F00

Xie, K., Ozbay, K., Yang, H., 2015. Spatial analysis of highway incident durations in the context of Hurricane Sandy. Accid. Anal. Prev. 74, 77–86. https://doi.org/10.1016/j.aap.2014.10.015

You, J., Wang, J., Guo, J., 2017. Real-time crash prediction on freeways using data mining and emerging techniques. J. Mod. Transp. 25, 116–123. https://doi.org/10.1007/s40534-017-0129-7

Yu, R., Abdel-aty, M.A., Ahmed, M.M., Wang, X., 2014. Utilizing Microscopic Traffic and Weather Data to Analyze Real-Time Crash Patterns in the Context of Active Traffic Management. IEEE Trans. Intell. Transp. Syst. 15, 205–213. https://doi.org/10.1109/TITS.2013.2276089

Zheng, Z., Ahn, S., Monsere, C.M., 2010. Impact of traffic oscillations on freeway crash

occurrences. Accid. Anal. Prev. 42, 626–636. https://doi.org/10.1016/j.aap.2009.10.009