IATSS Research xxx (2017) xxx–xxx

Contents lists available at ScienceDirect

IATSS Research

Research article

Association of reduced visibility with crash outcomes

Subasish Das [a,*], Bradford K. Brimley [b], Tomás E. Lindheimer [a], Michelle Zupancich [a]

[a] Texas A&M Transportation Institute, 3135 TAMU, College Station, TX 77843-3135, United States
[b] Binkley and Barfield, Inc., 3452 Rock Prairie Road West, College Station, TX 77845, United States

ARTICLE INFO

Article history:
Received 10 November 2016
Received in revised form 25 September 2017
Accepted 30 October 2017
Available online xxxx

Keywords:
Inclement weather
Variable importance
Random forest
Ordinal logistic

ABSTRACT

Most of the information necessary for driving a vehicle is regarded as visual information. In spite of its importance, visibility conditions at the time of a crash are often not documented at a high level of detail. Past studies have not examined the quantified values of visibility and its association with crashes. The current study merged data collected from the National Oceanic and Atmospheric Administration (NOAA) with 2010–2012 Florida crash data. From the thousands of logged weather events compiled by the NOAA, the researchers isolated periods of normal visibility and comparable periods of reduced visibility in a matched-pairs study. The NOAA data provided real visibility score based on the spatiotemporal data of the crashes. Additionally, the crash data, obtained from Roadway Information Database (RID), contains several geometric and traffic variables that allow for effects of factors and visibility. The study aims to associate crash occurrence under different levels of visibility with factors included in the crash database by developing ordinal logistic regression. The intent is to observe how different visibility conditions contribute to a crash occurrence. The findings indicate that the likelihood of a crash increase during periods of low visibility, despite the tendency for less traffic and for lower speeds to prevail during these times. The findings of this study will add valuable knowledge to the realm of the impact of visibility in the way of using and designing appropriate countermeasures.

1. Introduction

Human error is the most dominant factor in traffic crashes. This error ranges from complete negligence (e.g., distracted or impaired driving) to limitations of human abilities (e.g., slower reflexes with age, low visibility in inclement weather). One limitation that is often neglected is reduced visibility during inclement weather. To reduce the frequency of crashes that occur in inclement weather, it is necessary to investigate the key factors associated with these crashes thoroughly. This study seeks to identify the effects of reduced visibility on the likelihood of crashes and the factors that influence crashes during periods of reduced visibility.

Inclement or adverse weather presents a safety concern for vehicular traffic from multiple perspectives. One is the moisture on the road that reduces friction. Friction is reduced even more if the temperature is near or below freezing. Theofilatos and Yannis summarized studies presented at conferences or published in international journals that focus on the effects of traffic, weather, and the combined effect of traffic and weather on road safety. The research synthesis showed that precipitation has been widely investigated, and there is a constant effect of the increase in crashes. The risk posed by rainfall is due to a combination of poor road friction and low visibility. Other studies have found that injury severity in low visibility conditions was higher and head-on and rear-end collisions were the common types of crashes [1].

Visibility can be measured with specialized instrumentation. The National Oceanic and Atmospheric Administration (NOAA) regularly measures atmospheric and ground-level conditions, including visibility, at airports so pilots and air traffic control can make informed decisions for flying. These readings are stored in historical databases maintained by the NOAA. The key objective of this research is to identify the impact of visibility on crash outcomes so that appropriate countermeasures can be used to reduce visibility related traffic crashes. To accomplish the research goals, the study used historical weather data to determine the exact visibility score. The spatiotemporally determined visibility scores were merged with crash data to determine the association.

1.1. Driver adaptation in inclement weather

Drivers tend to change their behavior to adapt to new conditions presented by inclement weather. Studies evaluating driver behavior in inclement weather have been oriented toward large-scale observations rather than changes at the level of the individual driver. For example, researchers identified how inclement weather affects traffic speed, flow, and density on freeways in Seattle, Baltimore, and Minneapolis-Saint Paul [2]. Light rain was found to reduce free-flow speed 2 to 3.6%,

* Corresponding author.
E-mail addresses: s-das@tti.tamu.edu (S. Das), bbrimley@binkleybarfield.com (B.K. Brimley), t-lindheimer@tti.tamu.edu (T.E. Lindheimer), m-zupancich@tti.tamu.edu (M. Zupancich).
Peer review under responsibility of International Association of Traffic and Safety Sciences.

https://doi.org/10.1016/j.iatssr.2017.10.003
0386-1112/

capacity 10 to 11%, and the speed at capacity 8 to 10%. The reductions in free-flow speed and capacity increased with rain intensity, and snow had a larger impact on free-flow speed and capacity than rain. The researchers also found that visibility affects traffic conditions. Free-flow speeds on freeways in inclement weather were also studied in Spain [3], identifying a reduction in speed by 5.5 to 7 km/h for rain and 9 to 13.7 km/h for snow. The wind affects free-flow speed only when the wind speed is above 8 m/s. Visibility affects free-flow speed only when visibility is <2000 m. In a separate study [4], the wind speed was also shown to have a small effect on traffic compared to other weather factors. Analyzing traffic volumes, the researchers identified that volume is reduced 13 to 34% during inclement weather. Traffic volume and speed have also been observed to decrease during inclement weather in China [5].

In support of drivers responding to inclement weather by driving more conservatively, the research indicates that there is a change in traffic during inclement weather, not only in characteristics such as speed and capacity but also in the number of vehicles on roads. The reduction in traffic volumes during inclement weather is a reflection of driving as a derived demand. Depending on the driver and the severity of the weather, it is worth postponing the commercial and recreational activities that would have occurred to a later time. While less driving occurs during inclement weather, there are also reductions in free-flow and operating speeds. It suggests that, while some drivers forego driving altogether, those that still choose to drive do so with some level of caution.

While many drivers naturally drive slower during periods of inclement weather and low visibility, some agencies use variable speed limits on changeable message signs to encourage all drivers to respond similarly. Hassan et al. [6] surveyed drivers in Florida to investigate whether drivers respond to reduced speed limits in low visibility. Their models indicate that drivers 18 to 25 years and female drivers 51 years and older are more likely to reduce their speed in response to a variable speed limit when it is used during fog in low to medium-high traffic. Drivers are also more likely to reduce their speed if they are on a two-lane road.

1.2. Previous safety findings

Despite the behavioral changes triggered by inclement weather, there are safety concerns with driving in poor conditions. Several studies have focused on the effects of weather on crashes, including those related to visibility. One notable area for studying weather and visibility effects has been the state of Florida, which has several locations that frequently experience fog. Abdel-Aty et al. [7] compared crashes occurring in Florida during periods of fog and smoke with crashes during periods of clear visibility. They identified that there are a disproportionate number of crashes in fog and smoke when the speed limit is 55 mph or higher, light conditions are dark, and there is no street lighting. An odds ratio analysis showed that the probability of a crash in fog or smoke is 3.24 times more likely to result in a severe injury and is 1.53 times more likely to be a multiple vehicle crash. Head-on collisions are also more likely to occur. One interesting note is that the likelihood of young and middle-aged drivers to be in a crash increased during fog and smoke, but not for old drivers, suggesting that old drivers appropriately apply more caution in periods of low visibility. Also in Florida, Wang et al. [8] studied crashes on expressway ramps during periods of low visibility, finding an increase in the likelihood of a crash as visibility decreases. Though not focused on visibility, Sun et al. [9] calculated an increase in crash risk for rainy weather compared to dry weather. Das and Sun used crash data of Louisiana to investigate the pattern of crashes under the rainy weather [10].

Hassan and Abdel-Aty investigated whether real-time traffic flow data can be used to predict crashes under reduced visibility conditions. Researchers collected traffic flow data from loop detectors and crash records from two freeways in Florida during December 2007 and March 2009. There were 2984 total crashes. Of these crashes, 125 occurred under reduced visibility conditions, and only 67 reduced visibility crashes had traffic flow data. The study used a Random Forest analysis to identify significant variables, and a log odds ratio was used for predicting crashes. Three significant variables found were nearest upstream station speed, nearest downstream station speed (both 5 to 10 min before crash), and average occupancy at nearest downstream station 10 to 15 min before crash. The authors interpreted these results to mean that higher occupancy coupled with the increase of speed increases the likelihood of a crash between the upstream and downstream points [11]. In another study, Abdel-Aty and Hassan used loop detector data from interstate freeways and expressways in Florida to further predict reduced visibility related crashes using real-time data. Researchers adopted a Bayesian matched case-control logistic regression to analyze the crash data. When comparing the model that used loop detector data to automatic vehicle detection data, the study concludes that it is better to use loop detector data to predict crashes due to reduced visibility. The models developed in the study show that variation in speed can increase the risk of a crash when visibility is reduced [12].

The intent is to associate crash injury severity with visibility issues. Many studies claimed that visibility has somewhat association with injury severity. However, the quantification was not often conducted at the disaggregate level. This study aims to mitigate the current research gap.

2. Data collection and processing

The research team assembled a database collected in Florida from two different sources. These two sources are 1) the National Oceanic and Atmospheric Administration (NOAA) airport weather station data, and 2) Strategic Highway Research Program 2 (SHRP-2) Roadway Information Database (RID) data.

Fig. 1 illustrates a framework of the data compilation work to prepare the final dataset [13]. While low-visibility events tend to be associated with weather, there is one notable exception. Smoke is not a direct result of weather or precipitation, but usually a consequence of human dealings. The impact of smoke on crashes, in addition to the effects of fog, has been studied in previous work [7]. Smoke-related visibility events are infrequent and have a dramatically different quality than periods of reduced visibility caused by precipitation and moisture (such as fog or rain) because there is no water on the windshield nor need to use wipers [7].

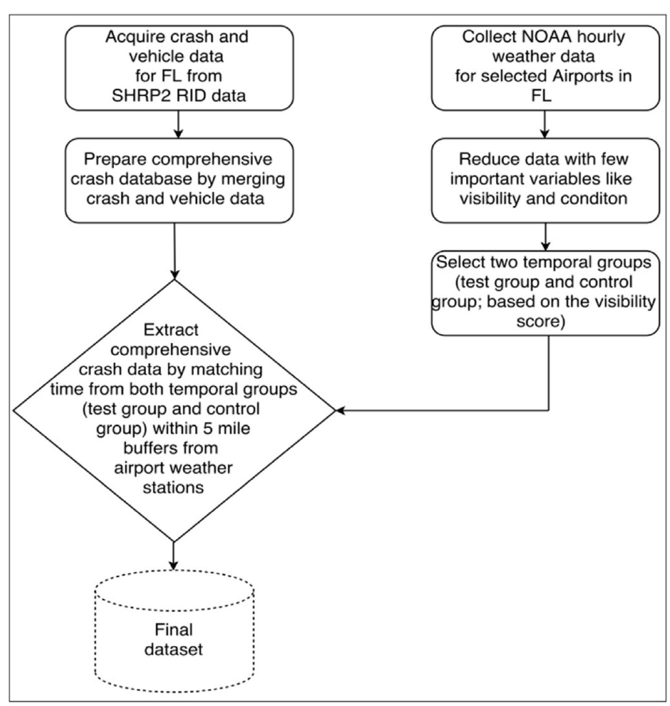

Fig. 1. Flowchart of data compilation (FL indicates Florida).

Please cite this article as: S. Das, et al., Association of reduced visibility with crash outcomes, IATSS Research (2017), https://doi.org/10.1016/j.iatssr.2017.10.003

2.1. NOAA weather data

The weather data used in the analysis was extracted from data collected by weather stations at airports in Florida and available in databases maintained by the NOAA. The equipment at each weather station reports hourly measurements but increases the frequency when there is a change in reported visibility. This results in accurate visibility information in real time to provide necessary alerts for periods of low visibility. Each entry in the original weather data included a visibility level (in miles), temperature, and a code for the type of weather. Researchers sifted through the weather data to identify periods when the visibility was within a defined range of poor visibility (0–0.5 miles) and medium visibility (0.5–4.0 miles).

The categorical options for the type of weather include events such as rain, thunderstorm, drizzle, mist, hail, snow, and fog. Modifiers "light" or "heavy" can be applied to indicate a heavy thunderstorm or light rain. Because different types of weather events cause reduced visibility, it was critical to have consistency in the types of weather observed within the same category of visibility. Fog is, therefore, the primary weather event associated with poor visibility; mist and rain are the primary weather events for medium visibility. Any events of snow, freezing temperatures, thunderstorms, and hail were removed in an effort to focus on reduced visibility. These other events are less frequent than fog, mist, and rain, and they are accompanied by characteristics such as reduced friction, wind speed, debris, or a need for rapid windshield wipers, thus adding variables that may hinder discoveries specific to a measure of atmospheric visibility. Windshield wipers, despite their utility in removing the precipitation, can be distracting and intermittently impede the driver's vision [14]. Smoke was also occasionally the cause of low visibility. Periods with smoke were removed from the analysis to focus exclusively on reduced visibility associated with moisture.

Each period of reduced visibility was matched with a period of normal visibility (9–10 mi) precisely one week earlier or one week later. This matching procedure, similar to the method used by Sun et al. [9], produced a control sample. This approach addresses spatial and temporal variations that are found in crash frequencies, where crashes can be more frequent during certain times of the year or in certain locations. Time periods with reduced visibility that could not be matched with a control period of reasonable visibility (either one week earlier or one week later) were removed. This reduction is around 7% of the total data. Both matches were removed if either the control or the test period included a state or national holiday or other day near the holiday when travel would be different from that of a typical day. The reduced or final data is around 90% of the total data. If visibility is assumed to have no effect on crash frequencies, the number of crashes observed during the tested periods of reduced visibility should be equivalent to the number of crashes observed during the matched control period.

2.2. SHRP2 RID crash data

Crash records from 2010 to 2012 were acquired from the SHRP-2 RID data and included both the crash-level (e.g. Annual Average Daily Traffic or AADT, the percentage of trucks, shoulder width, skid number, lighting condition, and facility type) and vehicle-level (e.g. number of vehicles, driver age, severity) fields. Crash data and vehicle data were merged to develop a general dataset involving required variables. Person level information was provided in vehicle level data. The final dataset contained unique information on person level. The recorded locations of the crashes were used to isolate the crashes to those occurring 5 miles within an airport, and the recorded times of the crashes were used to further reduce the crash dataset to those crashes occurring during a period of interest, whether during a test period of reduced visibility or a control period of normal visibility. A map graphically showing the airports, the 5-mile buffers, and observed crashes is shown in Fig. 2.

A preliminary data exploration was first conducted to examine the significant variables that may contribute to crash occurrence due to reduced visibility. The raw crash data, compiled from state maintained police records and supplemented with roadway information from SHRP-2, contain several fields that can be tested as variables. While the focus of this study is visibility, the influence of the road features and conditions at the time of the crash should be accounted for in studying how visibility impacts crash frequencies. There were several variables in the crash data that were often not reported. These variables (e.g. blood alcohol level, shoulder width, distance from the intersection, driveway density) were excluded from the analysis so efforts could be directed at the variables that were consistently reported. Descriptive statistics of the variables that were tested are shown in Table 1. The information in Table 1 shows how often a particular variable category is observed in the crash records during a control period (excellent visibility) or one of the test periods (medium visibility or poor visibility). From the descriptive statistics, for example, 16.6% of crashes in excellent visibility occurs when the maximum speed (speed limit) is 45 to 60 mph. Crashes increased to 19.7% when visibility is poor. For injury severity, there appears to be an increase in the percentage of crashes that were recorded as injury crashes, from nearly 48% of crashes in excellent visibility to 50% in poor visibility. The percentage of crashes reported as property damage only decreases from over 51% in excellent visibility to <49% in poor visibility. It suggests that crash severity is likely to increase as visibility decreases. The analyses that follow can account for other variables to test that observation.

2.3. Multicollinearity check

The variance inflation factor (VIF) is used to detect collinearity (strong correlation between two or more predictor variables) that causes instability in parameter estimation in regression models. VIF can be defined as:

$$VIF = 1/\left(1-R_i^2\right) \tag{1}$$

where:$R_i^2 =$ co-efficient of determination of ith variable on all other variables.

A general rule of thumb for multi-collinearity is to check whether VIF is >10. No variable has a VIF value >10 in the final dataset. The correlation values of Lighting, Weather, and Number of vehicles are closer to 1, suggesting that these variables have strong linear relationship with at least one other variables among these three variables.

2.4. Variable importance

The research team used a random forest algorithm assess variance importance using a package with the R statistical software [15]. With random forest algorithms, a randomly selected vector of input variables ($\boldsymbol{X}=X_1,\ldots,X_n$) to a random response variable $Y\in\mathcal{Y}$ is considered for analysis. The importance of a variable X_k while predicting or estimating Y is calculated by the Gini index, calculated from adding the impurity decreases in the following equation:

$$Importance(X_k) = \frac{1}{N_T}\sum_T \sum_{t\in T:v(s_t)=X_k} p(t).d \tag{2}$$

where:t = nodeT = all nodes$p(t)$ = proportion N_t/N of sample for node $t.s_t$= split for which all variables are sampled into two major nodes t_L and t_R to maximize the decrease, $dv(s_t)$ = variable used in split s_t

$$d = i(t) - p_L i(t_L) - p_R i(t_R) = \text{decrease}$$

N_T = all variables.

If the Gini index is considered an impurity function, the measurement is known as Mean Decrease Gini. Variables with higher values of Mean Decrease Gini are considered more important to the model. Fig. 3 shows the variable importance plot for the selected variables. Each

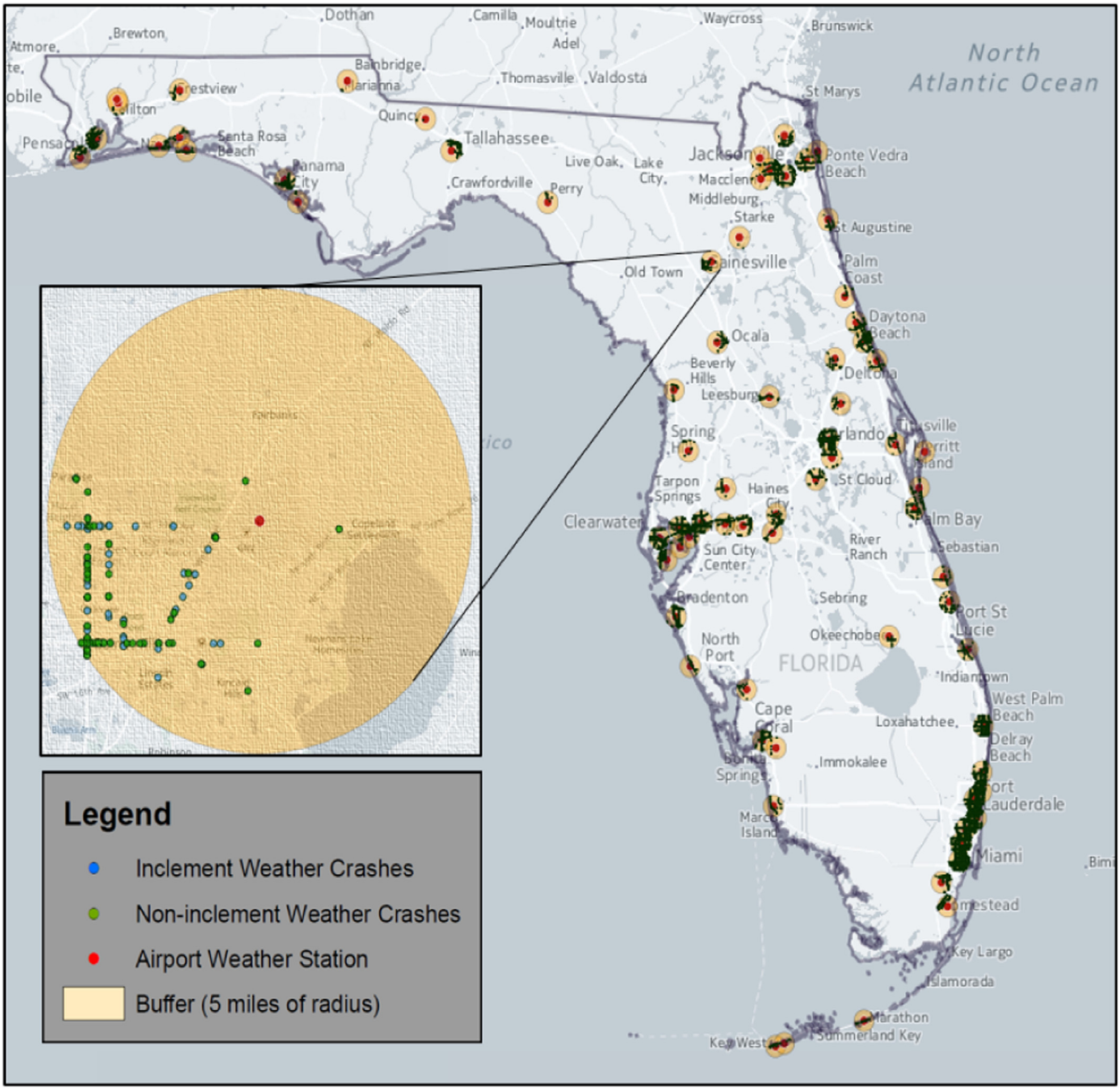

Fig. 2. Airports, buffer areas, and crashes in Florida.

variable is shown on the y-axis and variable importance on the x-axis. The variables on the y-axis are ordered from most to least important. Facility type and lighting condition are the two variables with lower Gini index. These two variables are correlated with average shoulder width. Due to lower significance and correlation with average shoulder width, the authors omitted these variables for model development.

2.5. Selected variables

Seven variables were selected for testing in models based on values of VIF, correlation, and Mean Decrease Gini. The selected variables are:

- Annual Average Daily Traffic (AADT)
- Driver Age
- Percentage of Trucks
- Skid Number
- Average Shoulder Width
- Maximum Speed
- Crash Severity

Histograms of the first six of these variables are shown in Fig. 4, divided by crashes occurring in the three visibility groups. The distributions appear to be quite similar for all variables. The percentages in each of the groups indicate relative percentages of the attributes in that particular group. The most notable differences are for skid number and driver age where there are peaks in the distributions for excellent visibility that are not seen in the other visibility levels. For driver age, higher numbers of crashes in excellent visibility condition are skewed toward the younger drivers. No significant differences are visible in

Please cite this article as: S. Das, et al., Association of reduced visibility with crash outcomes, IATSS Research (2017), https://doi.org/10.1016/j.iatssr.2017.10.003

Table 1
Descriptive statistics.

Category	Percentage of crashes		
	Excellent visibility	Medium visibility	Poor visibility
Functional class			
Rural major collector	0.02%	0.01%	0.00%
Rural minor arterial	0.17%	0.17%	0.14%
Rural principal arterial	0.82%	0.79%	0.75%
Urban collector	0.57%	0.60%	0.23%
Urban local	0.01%	0.01%	0.00%
Urban minor arterial	21.22%	18.71%	18.62%
Urban other principal	53.91%	53.48%	56.69%
Urban principal arterial	19.71%	22.09%	19.93%
Unknown	3.57%	4.14%	3.65%
Maximum speed (mph)			
0–30	7.89%	7.55%	7.76%
30–45	65.24%	61.73%	62.82%
45–60	16.62%	18.52%	19.74%
>60	10.25%	12.20%	9.68%
AADT (vehicle per day, vpd)			
0–9999	3.76%	3.66%	3.41%
10,000–34,999	35.08%	33.46%	35.08%
35,000–54,999	30.92%	30.14%	30.92%
55,000–124,999	16.47%	18.03%	16.79%
>124,999	12.85%	13.81%	12.96%
Unknown	0.92%	0.91%	0.84%
Percentage of trucks			
0–5	58.91%	56.90%	60.62%
5–10	35.19%	36.25%	32.32%
10–20	5.75%	6.55%	6.78%
>20	0.15%	0.30%	0.28%
Avg. shoulder width (ft.)			
0.00–1.00	39.41%	38.73%	40.74%
1.01–3.00	16.67%	15.88%	15.34%
3.01–5.00	12.44%	11.22%	11.46%
5.01–10.00	23.72%	26.00%	25.12%
>10.00	6.19%	7.35%	6.50%
Unknown	1.57%	0.83%	0.84%
Divided	67.21%	69.08%	69.69%
Undivided	31.42%	29.54%	29.33%
Unknown	1.36%	1.39%	0.98%
Skid number			
20–30	6.90%	7.80%	7.81%
30–40	67.73%	67.46%	65.39%
40–50	20.24%	19.27%	21.05%
>50	1.23%	0.83%	1.22%
Unknown	3.91%	4.65%	4.54%
Lighting condition			
Dark(no street light)	3.42%	3.02%	5.75%
Dark(street light)	28.29%	23.34%	38.63%
Dawn/dusk	5.01%	4.87%	6.22%
Daylight	63.29%	68.77%	49.39%
Number of vehicles			
Multi-vehicle	92.31%	91.66%	91.25%
Single vehicle	7.69%	8.34%	8.75%
Severity			
Fatal	0.90%	0.56%	0.98%
Injury	47.96%	47.14%	50.09%
No injury	51.14%	52.30%	48.92%
Driver age			
15–19	6.24%	6.58%	6.74%
20–29	26.14%	26.10%	27.08%
30–39	19.78%	18.70%	19.27%
40–49	18.63%	18.28%	18.29%
50–59	15.56%	13.89%	12.96%
60–69	8.37%	7.94%	7.39%
>70	5.19%	4.84%	4.07%
Unknown	0.09%	3.68%	4.21%

the histograms of maximum speed, AADT, and percentage of trucks. For shoulder width, the patterns are not similar for different levels of visibility.

3. Analysis

3.1. Model development with ordinal logistic regression

The research team used ordinal logistic regression to perform the analysis. The ordinal logistic regression models are also known as cumulative link models. In this analysis, the response variable is the crash occurrence due to visibility condition. Three ordinal response categories are considered: excellent visibility, medium visibility, and poor visibility. A cumulative link model will be developed based on the ordinal response variable, for example, Y_i with $k=1,\ldots,K$ categories [where $K \geq 2$]. Then Y_i follows a multinomial distribution with parameter π. The cumulative probabilities can be defined as:

$$Y_{ik} = P(\,Y_i \leq k) = \pi_{i1} + \ldots + \pi_{ik} \tag{3}$$

where:π_{ik} = probability of *ith* observation for response category k.

The cumulative logistic function is:

$$logit(y_{ik}) = logit(P(\,Y_i \leq k)) = \log\left[\frac{P(\,Y_i \leq k)}{1-P(\,Y_i \leq k)}\right] \quad where : k = 1,\ldots,k-1 \tag{4}$$

Note that the logit functions are defined as $\text{logit}(\pi) = \log[\frac{\pi}{1-\pi}]$. A cumulative link with a logistic link can be written as:

$$logit(y_{ik}) = \theta_k - X\mu \tag{5}$$

where:θ_k = parameters act as intercepts or horizontal displacements$X$ = transpose of a vector of predictor variables for the *ith* observationμ = matching set of regression parameters.

It is important to note that $X\mu$ is dependent on k categories. Therefore, it is considered that μ has the same effect for each of the $K-1$ cumulative logits. The odds ratio (OR) of the event $Y(\leq k)$ at x_1 relative to event $Y(\leq k)$ at x_2 is:

$$Odds\ Ratio\ (OR) = \frac{\frac{y_k(x_1)}{[1-y_k(x_1)]}}{\frac{y_k(x_2)}{[1-y_k(x_2)]}} = \frac{exp(\theta_k - X\mu)}{exp(\theta_k - X\mu)} = exp\left[(x_2^T - x_1^T)\mu\right] \tag{6}$$

Here, the odds ratio is independent of k. Thus, the cumulative odds ratio is proportional to the distance between x_1 and x_2. Therefore, the cumulative logit model is also known as proportional odds model. The analysis was performed with open source statistical software *R* [15,16,17]. Table 2 lists the values of the estimates of the first model. The first part of the outputs lists the regression coefficient values, standard errors, and p- values. By observing the *p*-values, it is found that maximum speed, skid number, and driver age have higher significance ($p < 0.05$). Severity is significant when the threshold of p is lower than 0.1. AADT is not significant by considering $p < 0.1$. One reason is that the impact of AADT is not significant in the model, as it is not considered in log scale. The 2.5% and 97.5% confidence interval values of these variables do not contain zero, which is also a good indicator of significance. The next part of Table 2 lists the estimates for the two intercepts (intercept between excellent visibility and medium visibility and intercept between medium visibility and poor visibility) to form three response categories. The intercepts indicate where the latent variable is cut to make the ordered subdivision in visibility score. The final part of the outputs provides -2*loglikelihood* of the model as well as the AIC value, used later for model selection.

The research team developed another model (model 2) by omitting percentage of trucks and average shoulder width because both of these

Please cite this article as: S. Das, et al., Association of reduced visibility with crash outcomes, IATSS Research (2017), https://doi.org/10.1016/j.iatssr.2017.10.003

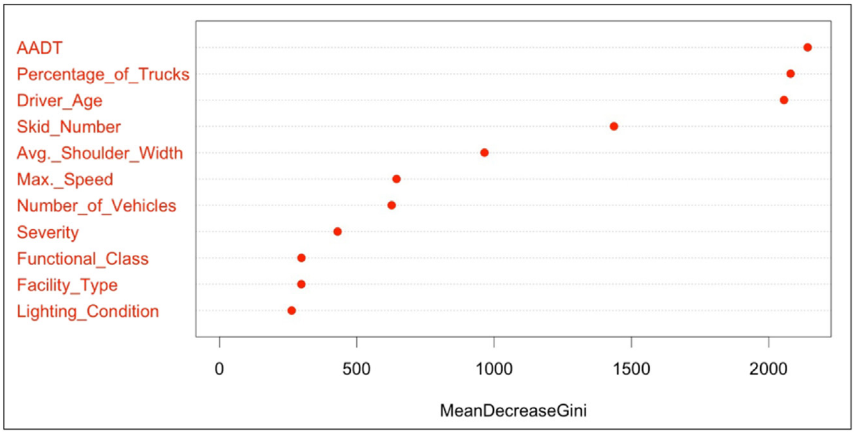

Fig. 3. Variable importance from random forest algorithm.

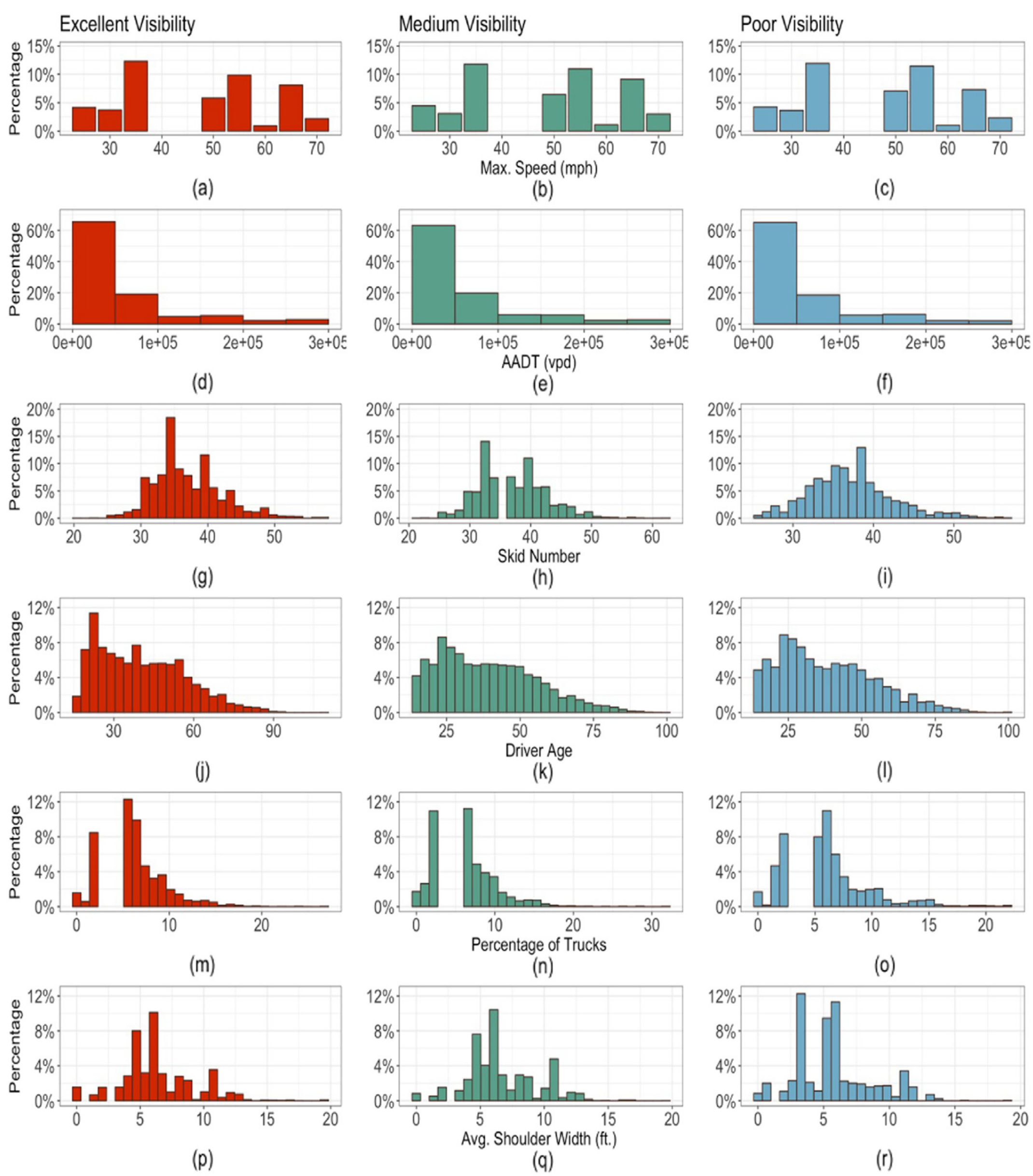

Fig. 4. Histogram of key variables.

Please cite this article as: S. Das, et al., Association of reduced visibility with crash outcomes, IATSS Research (2017), https://doi.org/10.1016/j.iatssr.2017.10.003

Table 2
Estimates from ordinal cumulative link model (model 1).

Variables and statistical measures	Combined						
	Estimate	2.50%	97.50%	St. Er.	z	Pr(>|z|)	Significance
Variables							
Severity: injury	0.265	−0.032	0.562	0.152	1.750	0.080	.
Severity: noinjury	0.263	−0.034	0.560	0.151	1.736	0.083	.
Maximum speed	0.008	0.004	0.012	0.002	4.048	0.000	***
AADT	0.000	0.000	0.000	0.000	−2.536	0.011	*
Skid number	−0.006	−0.010	−0.003	0.002	−4.209	0.000	***
Percentage of trucks	0.001	−0.008	0.010	0.005	0.197	0.844	
Avg. shoulder width	0.003	−0.007	0.013	0.005	0.536	0.592	
Driver age	−0.006	−0.007	−0.004	0.001	−7.273	0.000	***
Intercepts							
Visibility: excellent|medium	0.124			0.171	0.722		
Visibility: medium|poor	2.462			0.172	14.305		
Statistical measures							
AIC	43,888.17						
Log likelihood	−21,934.08						
Maximum gradient	6.38E−07						
Conditional H	4.60E+12						

Note: .: $p < 0.10$, *: $p < 0.05$, **: $p < 0.01$, ***: $p < 0.001$.

factors were insignificant in Model 1. The two models are compared in Table 3. The *p*-value of the test is 0.836, meaning that the Model 2 is not significantly different from Model 1. As Model 2 is not significantly different from Model 1, the research team used Model 1 as the final model.

3.2. Odds ratios

The interpretation of proportional odds ratios in cumulative link models is same as the odds ratios from a binary logistic regression model. Table 4 lists the odds ratio of the variables for the comparison of poor visibility to excellent and medium visibility. For injury crashes, the odds of injury severity in poor visibility versus medium visibility or excellent visibility combined is 1.304, indicating a 30% greater chance of a crash resulting in injury during poor visibility (assuming all other variables are held constant). For continuous variables, the odds ratio indicates the change in likelihood of a poor visibility crash for each unit change in the listed variable. For the maximum speed (odds ratio is 1.008 per 5 mph), the probability that the visibility at the time of the crash is poor (instead of medium or excellent visibility) increases 1.008 times higher for each 5-mph increase in speed. The odds ratio for skid number indicates there is a decrease in the likelihood of a poor visibility crash for increases in skid number. The odds ratio for AADT is 1, which indicates that the effect of AADT is negligible when considering the effect of visibility on crashes.

3.3. Crash probabilities by visibility level

Using the model results, the probabilities of a crash with a specific visibility level are plotted in Fig. 5, divided by the injury severity and independent variable. Plots for four independent variables are shown (the variables shoulder width and percentage of trucks are not shown as they were not significant in the original model). These plots can be used to examine how the probability that a particular visibility level is represented in the crash data changes as each independent variable changes.

Table 3
Model comparison.

Variables	AIC	Loglikehood	LR Stat	Pr (>Chisq)
Model 2	43,885	−21,934		
Model 1	43,888	−21,934	0.3582	0.836

In Fig. 5a, there were not enough crashes to identify a connection between maximum speed and visibility level for fatal crashes. However, as the maximum speed increases, there is a decrease in the likelihood that an observed crash occurred in excellent visibility compared to medium and poor visibility for both the injury and no injury crashes. There appear to be no common trends for AADT in Fig. 5b, but Fig. 5c suggests that increases in measured friction lead to crash reductions in medium and poor visibility. The probability that a crash from the dataset is coded as occurring in medium and poor visibility decreases while the probability for excellent visibility increases. This happens for all severity levels. Finally, driver age is shown as affecting the crash probabilities of all severity levels as well. The probability of a crash occurring in medium and poor visibility decreases compared to that of excellent visibility as age increases. This should not be surprising as drivers are less likely to drive in inclement weather as they age.

4. Conclusion

There is literature gap in determining the quantification of visibility and its association with crash or crash severity. Past studies show that there is a possible association between reduced visibility and crashes. The quantification part requires more exploration on real-time visibility score and its association with the severity of crashes. This study considered visibility status as response and quantified its association with visibility and other significant variables. The current study evaluated several of the factors that influence crashes during reduced periods of visibility during inclement weather. While previous research confirms that driver behavior changes during inclement weather, it is clear that there remain safety concerns due to reduced visibility. This study aims to investigate the impact of the level of visibility while driving with

Table 4
Odds ratios for model variables.

Variables	Odds ratio
Severity: injury	1.304
Severity: no injury	1.301
Maximum speed	1.008 (per 5 mph)
AADT	1.000 (per 10,000 vpd)
Skid number	0.994 (per 5 units of skid resistance)
Percentage of trucks	1.001
Average shoulder width	1.002
Driver age	0.994 (per 10 years of age)

Please cite this article as: S. Das, et al., Association of reduced visibility with crash outcomes, IATSS Research (2017), https://doi.org/10.1016/j.iatssr.2017.10.003

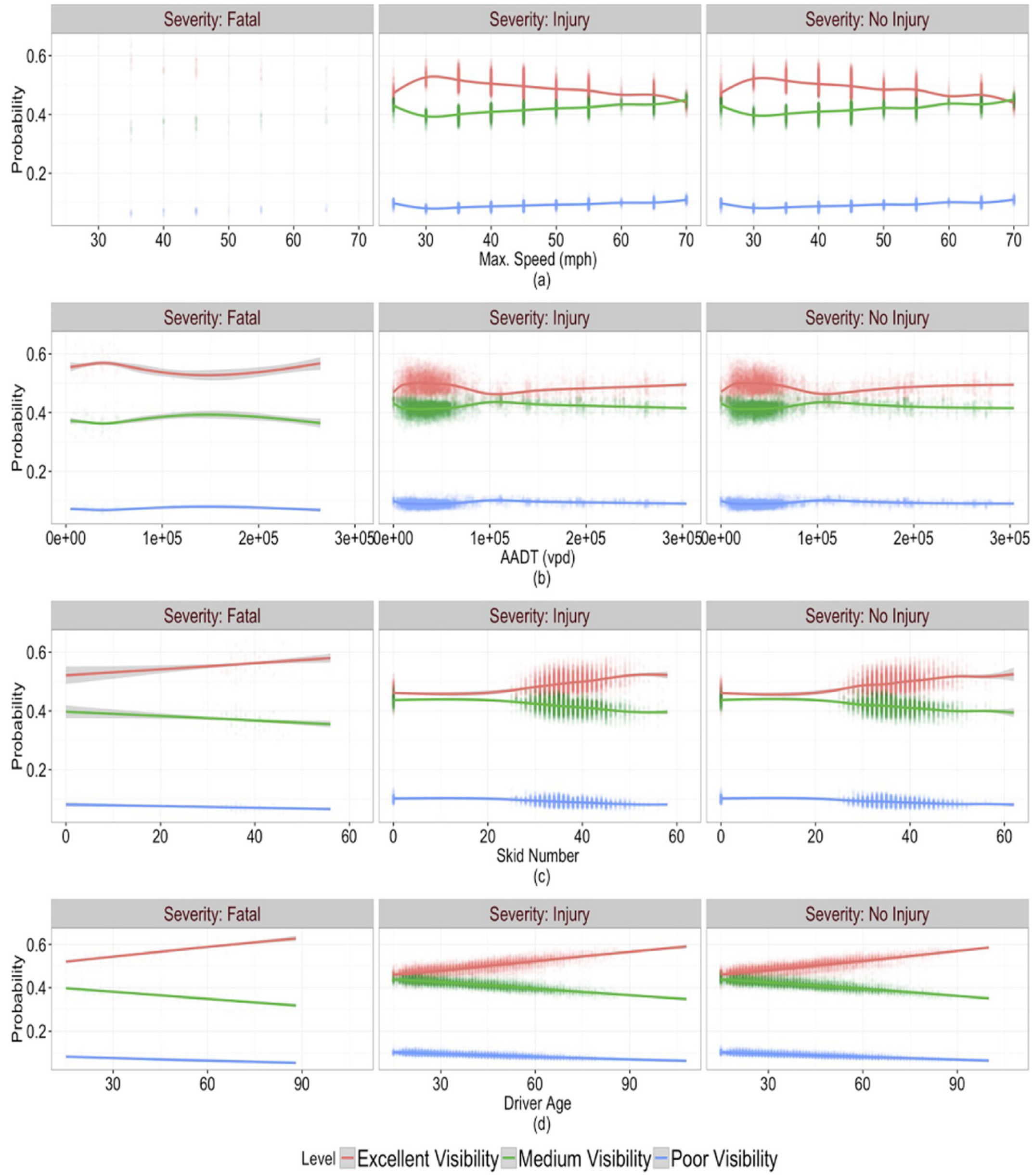

Fig. 5. Probability of occurrence for different variable levels.

other associated factors. Unless technology is able to adapt to the conditions drivers regularly drive in and overcome the limitations that human drivers have, the crash patterns are likely to continue with the same effects identified in this study.

Ahmed et al. [18] were unable to connect roadway geometry and traffic to models that identify effects of visibility conditions. The findings of this study concur that there is difficulty finding geometric or traffic effects when focusing on visibility. Shoulder width and truck percentage do not have significant effects, and AADT has small consequence. During low visibility, people tend to drive less. Thus, volumes are not positively associated with low visibility. The variables that appear most interesting are skid number and driver age, suggesting that higher friction reduces crashes in inclement weather and the old drivers are less likely to be in crashes in inclement weather.

The findings suggest that drivers are less likely to be involved in a crash during poor visibility conditions, as they get older. This is not surprising as older drivers are likely to apply caution and simply not drive in inclement weather [19]. A large part of the reduced traffic volumes in inclement weather as observed by Bartlett et al. [4] is likely to first come from the older drivers that tend to be more cautious, as suggested by Abdel-Aty et al. [7].

This study has several limitations. The current study does not consider the inclusion of naturalistic driving study (NDS) data to determine the association of human behavioral aspects.

Although this study developed statistically significant models, more advanced formulations may be applied with the inclusion of different conditions like roadway geometry, weather patterns, driver behavioral characteristics, and demographics in different regions. Another limitation of the study is the unavailability of the exposure matrices. Moreover, visibility scores were determined based on the proximity to the airport weather station data. This limitation can be reduced by using roadside weather station data [20], which was not available for Florida. Due to these limitations, the current findings would be limited to the regions with similar weather patterns. The authors recommend that future studies should focus on the current limitations. The authors also recommend that future studies may incorporate SHRP-2 naturalistic driving study (NDS) to determine the effect of driver behavior during adverse weather condition.

Please cite this article as: S. Das, et al., Association of reduced visibility with crash outcomes, IATSS Research (2017), https://doi.org/10.1016/j.iatssr.2017.10.003

The findings of this study would be useful for safety professionals to reemphasize the impact of visibility while driving. As low visibility can be considered as a difficult driving situation, the outcomes of this research can be used for the safety threshold determination of self-driving cars. If the performance matrices of the self-driving cars can overcome the performance matrices of human driving during difficult driving conditions like low visibility, the real-world integration of self-driving cars would be evident.

Acknowledgements

The research reported here was sponsored by the Atlas Center under Texas A&M (DTRT13-G-UTC54) Transportation Institute (TTI) Competitive Research Program 2016.

References

[1] A. Theofilatos, G. Yannis, A review of the effect of traffic and weather characteristics on road safety, Accid. Anal. Prev. 72 (2014) 244–256.

[2] H. Rakha, M. Farzaneh, M. Arafeh, E. Sterzin, Inclement weather impacts on freeway traffic stream behavior, Transportation Research Record: Journal of the Transportation Research Board, No. 2071, Transportation Research Board of the National Academies, Washington, D.C. 2008, pp. 8–18.

[3] F.J. Camacho, A. Garcia, E. Belda, Analysis of Impact of Adverse Weather on Freeway Free-Flow Speed in Spain, Transportation Research Record: Journal of the Transportation Research Board, No. 2169, Transportation Research Board of the National Academies, Washington, D.C., 2010 150–159.

[4] A. Bartlett, W. Lao, Y. Zhao, A.W. Sadek, Impact of Inclement Weather on Hourly Traffic Volumes in Buffalo, New York, Transportation Research Board 92nd Annual Meeting, No. 13-3240, 2013.

[5] P. Shang, R. Li, Z. Liu, X. Li, Inclement weather impacts on urban traffic conditions15th COTA International Conference of Transportation Professionals, 2015.

[6] H. Hassan, M. Abdel-Aty, A. Oloufa, The Effect of Warning Messages and Variable Speeds in Different Visibility Conditions, Transportation Research Board 90th Annual Meeting, No. 11-0354, 2011.

[7] M. Abdel-Aty, A. Ekram, H. Huang, K. Choi, A study on crashes related to visibility obstruction due to fog and smoke, Accid. Anal. Prev. 43 (2011) 1730–1737.

[8] L. Wang, Q. Shi, M. Abdel-Aty, P. Kuo, Predicting Crashes on Expressway Ramps With Real-Time Traffic and Weather Data, Transportation Research Record: Journal of the Transportation Research Board, No. 2514, Transportation Research Board of the National Academies, Washington, D.C., 2015 32–38.

[9] X. Sun, H. Hu, E. Habib, D. Magri, Quantifying crash risk under inclement weather with radar rainfall data and matched-pair method, Journal of Transportation Safety & Security 3 (2011) 1–14.

[10] S. Das, X. Sun, Investigating the Pattern of Traffic Crashes under Rainy Weather by Association Rules in Data Mining, The Proceedings of the 93rd TRB Annual Meetings, Washington D.C., 2014.

[11] H. Hassan, M. Abdel-Aty, Predicting reduced visibility related crashes on freeways using real-time traffic flow data, J. Saf. Res. 45 (2013) 29–36.

[12] M. Abdel-Aty, H. Hassan, M. Ahmed, A. Al-Ghamdi, Real-time prediction of visibility related crashes, Transp. Res. C 24 (2012) 288–298.

[13] S. Das, B. Brimley, T. Lindheimer, A. Pant, Safety Impacts of Reduced Visibility in Inclement Weather, Report No. ATLAS-2017-19, 2017.

[14] J. Sayer, J. Wetzel, D. Funkhouser, Examination of Naturalistic Windshield Wiper Usage, UMTRI, 2004.

[15] Liaw, A. and M. Wiener (2002). Classification and regression by random forest. R News 2(3), pp. 18–22, 2002.

[16] R Development Core Team, R: A Language and Environment for Statistical. Computing. Version 2.10.1, R Foundation for Statistical Computing, Vienna, Austria, 2013http://www.R-project.org.

[17] Christensen, R. Analysis of ordinal data with cumulative link models-estimation with the R-package ordinal. Vignette of 'ordinal' package. https://cran.r-project.org/web/packages/ordinal/vignettes/clm_intro.pdf Accessed: July, 2016.

[18] M.M. Ahmed, M. Abdel-Aty, J. Lee, R. Yu, Exploring the Feasibility of Using Airport Data in Real-Time Risk Assessment, Transportation Research Board 92nd Annual Meeting, No. 13-0834, 2013.

[19] New York State of Opportunity, Are You Concerned About an Older Driver? 2015.

[20] S. Chien, J. Meegoda, J. Luo, P. Corrigan, L. Zhao, Road Weather Information System Statewide Implementation Plan, New York State Department of Transportation, 2014.

Please cite this article as: S. Das, et al., Association of reduced visibility with crash outcomes, IATSS Research (2017), https://doi.org/10.1016/j.iatssr.2017.10.003