Real-time assessment of fog-related crashes using airport weather data: A feasibility analysis

Mohamed M. Ahmed [a],*, Mohamed Abdel-Aty [b], Jaeyoung Lee [b], Rongjie Yu [b]

[a] Department of Civil and Architectural Engineering, University of Wyoming, Laramie, WY 82071, United States
[b] Department of Civil, Environmental and Construction Engineering, University of Central Florida, Orlando, FL 32816, United States

ARTICLE INFO

ABSTRACT

The effect of reduction of visibility on crash occurrence has recently been a major concern. Although visibility detection systems can help to mitigate the increased hazard of limited-visibility, such systems are not widely implemented and many locations with no systems are experiencing considerable number of fatal crashes due to reduction in visibility caused by fog and inclement weather. On the other hand, airports' weather stations continuously monitor all climate parameters in real-time, and the gathered data may be utilized to mitigate the increased risk for the adjacent roadways. This study aims to examine the viability of using airport weather information in real-time road crash risk assessment in locations with recurrent fog problems. Bayesian logistic regression was utilized to link six years (2005–2010) of historical crash data to real-time weather information collected from eight airports in the State of Florida, roadway characteristics and aggregate traffic parameters. The results from this research indicate that real-time weather data collected from adjacent airports are good predictors to assess increased risk on highways.

© 2014 Elsevier Ltd. All rights reserved.

1. Introduction

Current statistics of crashes due to reduction in visibility and major inclement weather, particularly fog show that the fatal crashes in these weather conditions are certainly a major problem that needs to be addressed. According to the Federal Highway Administration (Goodwin, 2003), weather contributed to over 22% of the total crashes in 2001. This means that adverse weather likely increase the likelihood of crash occurrence. Several studies, in fact, concluded that crashes increase due to vision obstruction during rainfall by 100% or more (Brodsky and Hakkert, 1988; NTSB, 1980), while others found more moderate (but still statistically significant) increases (Andreescu and Frost, 1998; Andrey and Olley, 1990). Sudden reduction in visibility due to fog was found to increase severity level of crashes and tend to involve more vehicles. Statistics from the Fatality Analysis Reporting System (FARS) showed that fatal crashes during inclement weather events, i.e. rain, snow, fog and smoke, is certainly a major problem that needs to be mitigated. Inclement weather of rain, snow and fog/smoke resulted in a total of 31,514 fatal crashes between 2000 and 2007 in all 50 states, the District of Columbia, and Puerto Rico (FARS).

Koetse and Rietveld (2009) addressed the effect of the climate change on transportation in general, while others discussed the effect of the particular weather condition on traffic operations, safety, traffic demand, flow and traffic intensity, and operating speeds (Maze et al., 2006; Cools et al., 2008; Edwards, 1999). Visibility detection systems have been a successful element of existing variable speed limit (VSL) systems in some European countries; weather-controlled VSL system was found to reduce injury crashes on a German motorway by 20% (Baltz and Zhu, 1994), and an integrated slippery road warning system and a VSL system was found to be effective in crash reduction by 10% on Finish motorway (Rämä, 1997). Cooper and Sawyer (1993) examined an automatic fog-warning system in England, the results indicated a reduction in the net mean speed of approximately 3 km/h when "Fog" warning was displayed. Hogema et al. (1996) found higher speed reduction of 8–10 km/h on a Dutch fog warning system. Many studies indicated a reduction in the mean speed and the standard deviation of speed during adverse weather utilizing weather-controlled VSL systems. (Rämä et al., 2001; Rämä and Schirokoff, 2004). Várhelyi (2002) introduced an Intelligent Speed Adaptation (ISA) system based on road condition and visibility level, Várhelyi recommended using a vehicle-based visibility range meters instead of the expensive aerodromes visibility sensors used in fog warning systems on motorway

in Germany. Abdel-Aty and Pemmanabonia (2006) utilized rainfall data provided by Florida Automated Weather Network (FAWN) and the National Oceanic and Atmospheric Administration (NOAA) for 5 airports in vicinity of the 36-mile freeway section on I-4 in the Central Florida region in real-time crash prediction; they concluded that airport weather stations may provide suitable and relevant measure of rainfall indices to the nearby study freeway section. Ahmed et al. (2012) utilized real-time weather information in risk assessment on freeways; the results showed that the inclusion of weather information is essential in proactive traffic management systems.

Florida is considered among the top states in the United States regarding traffic safety problems resulting from adverse visibility conditions due to fog. Florida was the third after California and Texas with 299 fatal crashes occurring due to fog between 2002 and 2007. The most recent example for visibility related (VR) crashes in Florida is the 70 vehicle pileup on I-4 in Polk County, Florida in January 2008. This multi vehicle crash caused 5 fatalities, many injuries, and shutting down I-4 for extended time. Equally devastating was the major multi vehicle fatal crashes on I-75 in January 2012. Fig. 1 illustrates the distribution of injury severity for crashes related to vision obstruction in Florida. It depicts the increased severity levels of vision obstruction related crashes compared to crashes that are not related to vision obstruction.

The problem descends from the inadequacy of traffic control techniques to provide guidance for drivers and the unpredictability of locations and times of reduced visibility on highways. Although some transportation agencies developed and/or implemented some countermeasures e.g. visibility detection and warning systems to mitigate the limited-visibility problem, these countermeasures are limited at specific locations and expensive for installation. Ahmed et al. (2014) provided a recent synthesis about visibility detection systems in the United States and around the world, their study showed that visibility detection systems are rare and expensive to cover all roadways in fog-prone areas. Only 18 states with visibility detection systems were identified in the United States; Alabama, Arizona, California, Florida, Georgia, Idaho, Indiana, Louisiana, Nevada, New Jersey, North Carolina, Pennsylvania, South Carolina, Tennessee, Utah, Virginia, Washington, and Wisconsin. It is worth mentioning that each state has only one system which covered only a short length of one roadway. The cost of these systems ranged from $18,000 for a very simple and less accurate system installed at a spot location to $5,000,000 system covering only few miles of one roadway. On the other hand, there are existing weather stations at mostly all airports that can provide real-time weather data. The study areas considered in this
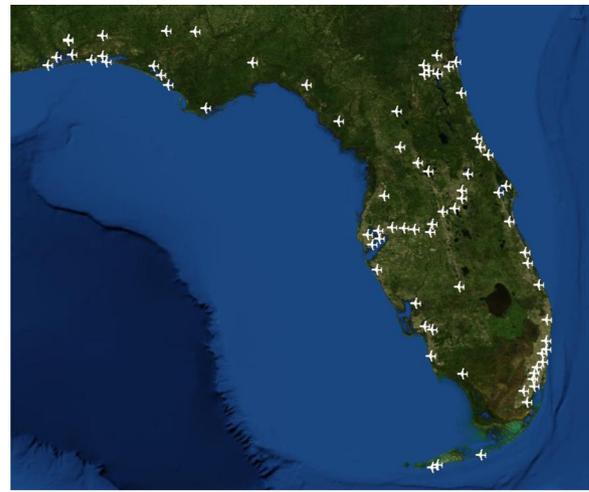

Fig. 2. Locations of airports and military airbases in Florida.

study are six counties in Florida that were identified as hotspots of fog related crashes. Weather data were collected from weather stations located at eight airports within the study areas. Considering the abovementioned facts and recognizing that fog occurrence can have significant effect on crash occurrence, this research aims at exploring the feasibility of using weather data collected from airports in real-time road crash risk assessment.

The following sections illustrate the procedures of preparing the data, modeling technique, interpretation and evaluation, and the conclusions.

2. Data description and preparation

There are a variety of methods that have been used for collecting weather data on highways. Crash and weather data can be collected from the long form crash reports; however, recorded weather conditions may be mistakenly reported (Shinar et al., 1983). Other studies collected data from weather stations installed within Advanced Traffic and Information System (ATIS) (Ahmed et al., 2012). Andrey et al. (2003) gathered weather data from the Meteorological Service of Canada (MSC), but found it difficult to compare MSC data with the weather data from crash reports and one study collected weather data from airports in the vicinity of the locations of the crashes (Abdel-Aty and Pemmanabonia, 2006).
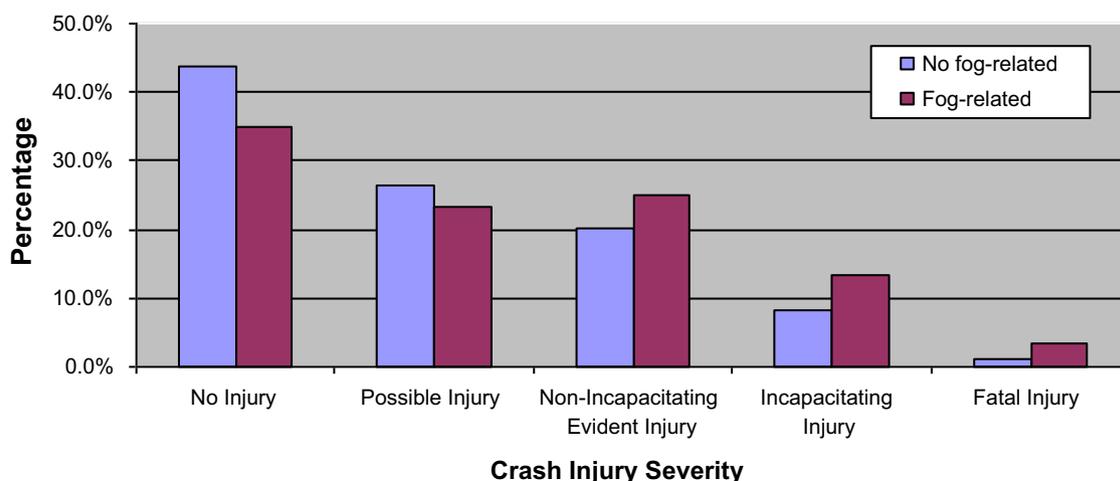

Fig. 1. Distribution of crash injury for vision and non vision obstruction.

Crash data were collected from Florida Department of Transportation's (FDOT) Crash Analysis Reporting (CAR) system. Only crashes on the state highway system such as Interstate, US Route and State Route highways were included since roadway characteristic data are only available for such road types. Airport weather data were collected from the National Climate Data Center (NCDC) under the umbrella of the National Oceanic and Atmospheric Administration (NOAA). NCDC archives weather data from various weather stations nationwide including radar, satellites, airport weather stations, US Navy, US Air force, etc.

As shown in Fig. 2, there are 76 airports, US Air Force/Navy bases and one space center in Florida. Airports' automated weather stations monitor the weather conditions continuously and the weather parameters are recorded according to a specific change in the reading threshold and hence they do not follow a specific time pattern. The stations report frequent readings as the weather conditions change within short time; if the weather conditions remain the same the station would not update the readings. These weather data include visibility, temperature, humidity, wind speed and direction, precipitation, etc. Among all these parameters, visibility is considered one of the most critical factors affecting crash occurrence. Visibility in general can be described as the maximum distance (in mile) that an object can be clearly perceived against the background sky, visibility impairment can be a result of both natural (e.g., fog, mist, haze, snow, rain, windblown dust, etc.) and human induced activities (transportation, agricultural activities, and fuel combustion). The automated weather stations do not directly measure the visibility but rather calculate it from a
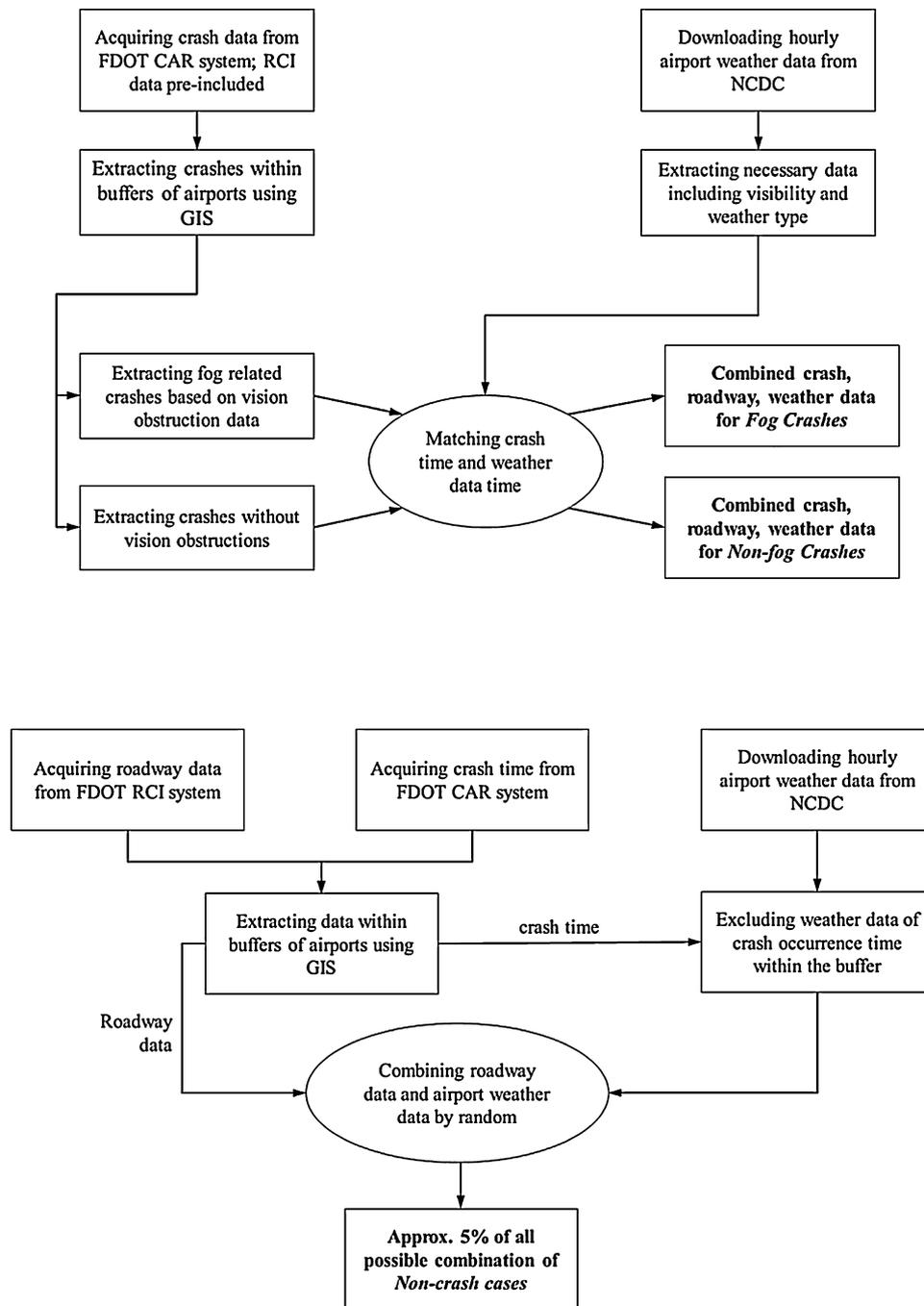

Fig. 3. Data preparation for fog/non-fog crashes, and non-crashes.

measurement of light extinction which includes the scattering and absorption of light by particles and gases. All airports' weather data are collected and maintained by NCDC.

Some airports have complete data since several decades ago. Unfortunately, some other airport weather data are not complete and there are several missing values. We selected airports with more complete weather data for the years 2005–2010. Finally eight airports were chosen in fog-prone areas identified from the GIS analysis.

As mentioned earlier that crash and the corresponding roadway data are collected from FDOT's CAR system, only crashes within the vicinity of airports were extracted using Geographic Information System (GIS). Fog crashes were extracted from the crash report and were matched with weather data closest time just before the crash. Meanwhile, crashes without any vision obstructions were also extracted and they were matched with the weather data from the closest airport. The modeling procedure required also non-crash data, a random selection from the whole remaining weather, traffic and roadway geometry datasets where there was no crash within 2-hour before the extraction time was utilized in the study to represent the whole population of different traffic, weather conditions and roadway characteristics. Fig. 3 illustrates the process of the data extraction. The first flow chart in Fig. 3 displays the data preparation for fog crash and non-fog crash cases. Non-fog crashes are all crashes not related to fog event, it can include other vision obstructed crashes due to smoke/smog, haze, etc. Crashes occurring within airport buffers were extracted and separated by the existence of vision obstruction. Then fog related crashes and non-fog related crashes were combined with airport weather data. As a result, 90 fog related crashes and 66,230 non fog-related crash data were prepared. These data were used to develop Fog vs. Non-Vision Obstructed Crashes Model. Non-obstructed vision crashes are all crashes occurred in clear weather. The second flowchart in Fig. 3 shows the data preparation process for extracting non-crash cases. A sample of non-crash cases was generated by randomly selecting roadway location (with traffic and roadway characteristics) and weather conditions (including visibility distance) from nearby airports. These randomly selected locations and time on the roadways are used to model "crash vs non-crash with different thresholds of visibility condition" (see Table 5). Among the all possible combination (13,287,320 cases), 5% (664,366 cases) were drawn randomly from the available options for analyzing effects of visibility distance on crash occurrence.

3. Preliminary analysis of NOAA airport data

The first step of investigating the fog related crashes was to examine the spatial distribution and as such the crash hotspots could be identified and focused on for further safety evaluation. The statewide map with frequent fog crash clusters was also presented for better visualization and understanding of the spatial distribution of fog crashes.

In order to determine the study region, priority areas with frequent fog crashes were identified in macroscopic analysis. Kernel density estimation (KDE) (see Chainey and Ratcliffe, 2005) was used to serve the purpose of clustering the crashes and identifying the hotspots. Schneider et al. (2004) adopted KDE and investigated pedestrian related crash hotspots on the university campus. Anderson (2009) explored the spatial distribution of injury crashes in London. The author adopted KDE technique to identify crash hotspots. Abdel-Aty et al. (2010) and Ekram (2009) examined fog related crashes in Florida. The author identified 10 hotspots for fog crashes in Florida using KDE. Recently, Kuo et al. (2013) also applied KDE for hotspots for traffic crashes and crimes. The authors provided a strategy to organize police patrol routes efficiently. Previous studies that adopted KDE agreed that KDE is a

useful for highlighting concentrations of traffic crashes, however, several researchers commonly pointed that KDE has drawbacks. First, KDE does not provide predicted probability of crashes in the future. Second, KDE does not account for exposure. In this study, KDE technique was applied to roughly identify hotspots for fog related crashes and select study areas.

The KDE defines the spread of risk as an area around a defined cluster in which there is an increased likelihood of a crash to occur based on spatial dependency. It places a symmetrical surface over each point and then evaluating the distance from the point to a reference location based on a mathematical function and then summing the value for all the surfaces for that reference location. This procedure is repeated for successive points, which allows us to place a kernel over each observation, and summing these individual kernels gives us the density estimate for the distribution of crash points (Fotheringham et al., 2000).

$$f(x, y) = \frac{1}{nh^2} \sum_{i=1}^{v} K \left(\frac{d_i}{h} \right)$$

where $f(x, y)$ is the density estimate at the location (x, y); n is the number of observations, h is the bandwidth or kernel size, K is the kernel function, and d_i is the distance between the location (x, y) and the location of the ith observation. The main objective of placing these kernels over the crash points is to create a smooth, continuous surface. Around each point at which the indicator is observed a circular area (the kernel) of defined bandwidth is created. This takes the value of the particular indicator at that particular point spread into it according to some appropriate function. Then it sums up all of these values at all places, including those at which no incidences of the indicator variable were recorded, gives a surface of density estimates.

The ArcGIS spatial analyst tool provides the features needed to do the cluster analysis by density estimation methods. The KDE process needs that the data-points be spatially jointed. For the points to be joined spatially, fishnet of square size cells was created using the "create fishnet" tool. The cell size (cell width and height) was selected in such a way that the area under
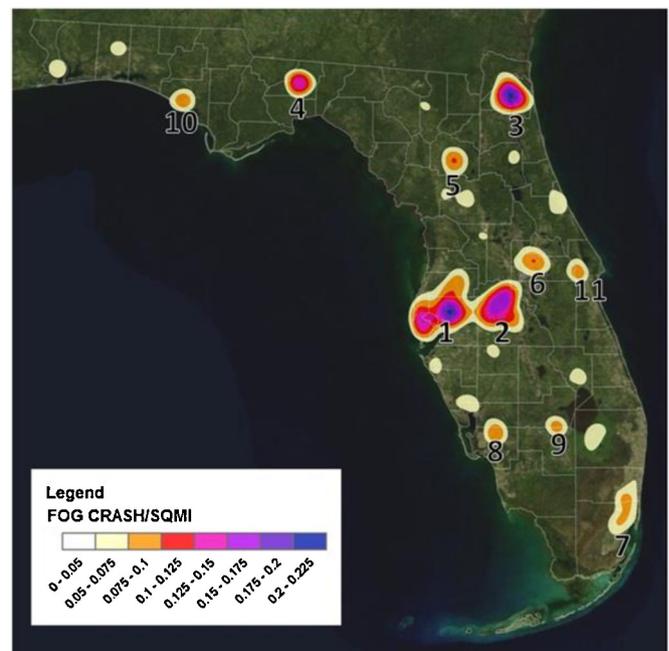

Fig. 4. Cluster analysis of fog related crashes on Florida state highway system (2003–2010).

Table 1
Matching result of fog crashes by buffer sizes.

Buffer size	Fog crashes within buffers	Matched	Not matched	Matching rate
2SM	11	10	1	90.9%
3SM	28	24	4	85.7%
4SM	51	42	9	82.4%
5SM	78	63	15	80.8%
5.8SM (5 NM)	90	74	16	82.2%
6SM	97	78	19	80.4%
7SM	113	92	21	81.4%
8SM	128	100	28	78.1%
9SM	152	118	34	77.6%
10SM	169	131	38	77.5%

Table 2
Classification table of the data.

	Fog crash	Non-fog crash	Sum
Airport weather: fog	74 82.22% sensitivity	3737 6.01%	3811 (no. of fog weather condition)
Airport weather: non-fog	16 17.78%	58,493 93.99%	58,509 (no. of non-fog weather condition)
Sum	90 (no. of fog crashes within the buffer)	62,230 (no. of non-fog crashes within the buffer)	62,320 (total no. of crashes within the buffer)

consideration is divided into a finite number of cells that can be calculated. Since the fog-related crashes are sparsely occurring, the fishnet cells were created such that the number of cells on each side does not exceed 100. The kernel density function was applied to calculate the boundaries of each cluster, with more number of points (crashes) within the center of each cluster.

Fig. 4 shows the statewide map with clustering output from the GIS analysis. The KDE technique presents 11 distinct fog crash hotspot areas on the Florida road network. The colors represent the density of crashes per square mile area. From the figure, the 11 clusters identified are associated with crash densities above 0.075 crashes per square mile. It is notable that several most hazardous areas have crash densities higher than 0.20 crashes per square mile, e.g. Pinellas, Hillsborough and Pasco.

3.1. Determining size of the airport buffer

Determining buffer size is important because the buffer size can be interpreted as the effective range of airport weather data for the adjacent highways (Fig. 5).

Fog related crashes were collected within buffers and were tested if they have occurred when adjacent airport reported a foggy weather condition. Matching rates by buffers were also calculated to determine the best buffer size. The matching rate refers to the ratio of number of fog crashes within the buffer that are matched with airport foggy weather condition to total fog crashes within the buffer. Matching process is illustrated in Fig. 6.

The number of fog related crashes within buffers by buffer sizes from 2 to 10 statute miles (SM) and matching results are presented in Table 1. 5 nautical mile (NM) buffer is also included since most airports have jurisdiction over surface area within 5NM radius surrounding the airport.

As a result, a trend was found that matching rate decreases as the buffer size increases; it seems reasonable because the airport weather information is more reliable for the nearer location whereas it is less reliable for the farther location. 2SM has the highest matching rate; however, it has only 11 fog crashes which are too small for the analysis. Finally 5NM was selected as the best buffer size because it has enough fog crashes and relatively higher matching rate.

In order to validate weather conditions before fog/non-fog crashes with weather reported by the adjacent airports, sensitivity analysis is conducted as shown in Table 2. Sensitivity is the proportion of fog crashes that are correctly matched in foggy conditions by the airport weather data while specificity is the proportion of non-fog crashes that are correctly matched clear weather conditions by the airport weather data. The sensitivity was found to be 82.22% while the specificity was found to be about 94%.

Therefore, it may be concluded that the fog crashes within 5NM buffers can be well explained by adjacent airport weather data. The

accuracy of the data was 0.9398, which means 94% of weather conditions from crash report within the buffers are concordant with the weather data from the airport. The false alarm rate which shows the rate of number of non-fog-related crashes not matched with the weather data to total number of non-fog crashes was about 6%. However, it is required to be cautious to determine airport buffer size. In this study, crash and weather data of eight airports in Florida were collected. However, Florida has unique geographical characteristics. For example, the mean elevation of Florida is very low (100 ft) compared to other states and the highest point is only 345 ft. Also, Florida does not have mountainous terrains. It is possible that 5NM buffer size was solely valid in Florida due to the unique features. Thus, if data are collected from other regions, the optimal buffer size should be re-calculated.

Moreover, it also needs to be careful for false alarms. In the false alarm situation, the weather system warns it is foggy condition but the actual weather is not foggy. The false alarm may be caused because sometimes fog is only concentrated very near the airport. Other possible reason is that the recorded weather conditions may be mistakenly reported by police officers (Shinar et al., 1983). Some may think it would be better to convey fog warning regardless of accuracy considering for the worst case. Nevertheless, frequent false alarms may make people to be insensible to the warnings. Therefore, it is desirable to consider both accuracy and false alarm rates to select the optimal buffer size.

4. Methodology

4.1. Bayesian logistic regression

The study utilized a Bayesian logistic regression approach to estimate the probability of crash occurrence. Bayesian logistic regression has the formulation of a logistic equation and can handle both continuous and categorical explanatory variables. The classical logistic regression treats the parameters of the models as fixed, unknown constants and the data is used solely to best estimate the unknown values of the parameters. In the Bayesian approach, the parameters are treated as random variables, and the data is used to update beliefs about the behavior of the parameters to assess their distributional properties. The interpretation of Bayesian inference is slightly different than the classical statistics; the Bayesian derives updated posterior probability of the parameters and construct credibility intervals that have a natural interpretation in terms of probabilities. Moreover, Bayesian inference can effectively avoid the problem of over fitting that occurs when the number of observations is limited and the number of variables is large.

The Bayesian logistic regression models the relationship between the dichotomy response variable (crash/no-crash) and the explanatory variables of roadway geometry, real-time weather and traffic. Suppose that the response variable y has the outcomes $y = 1$
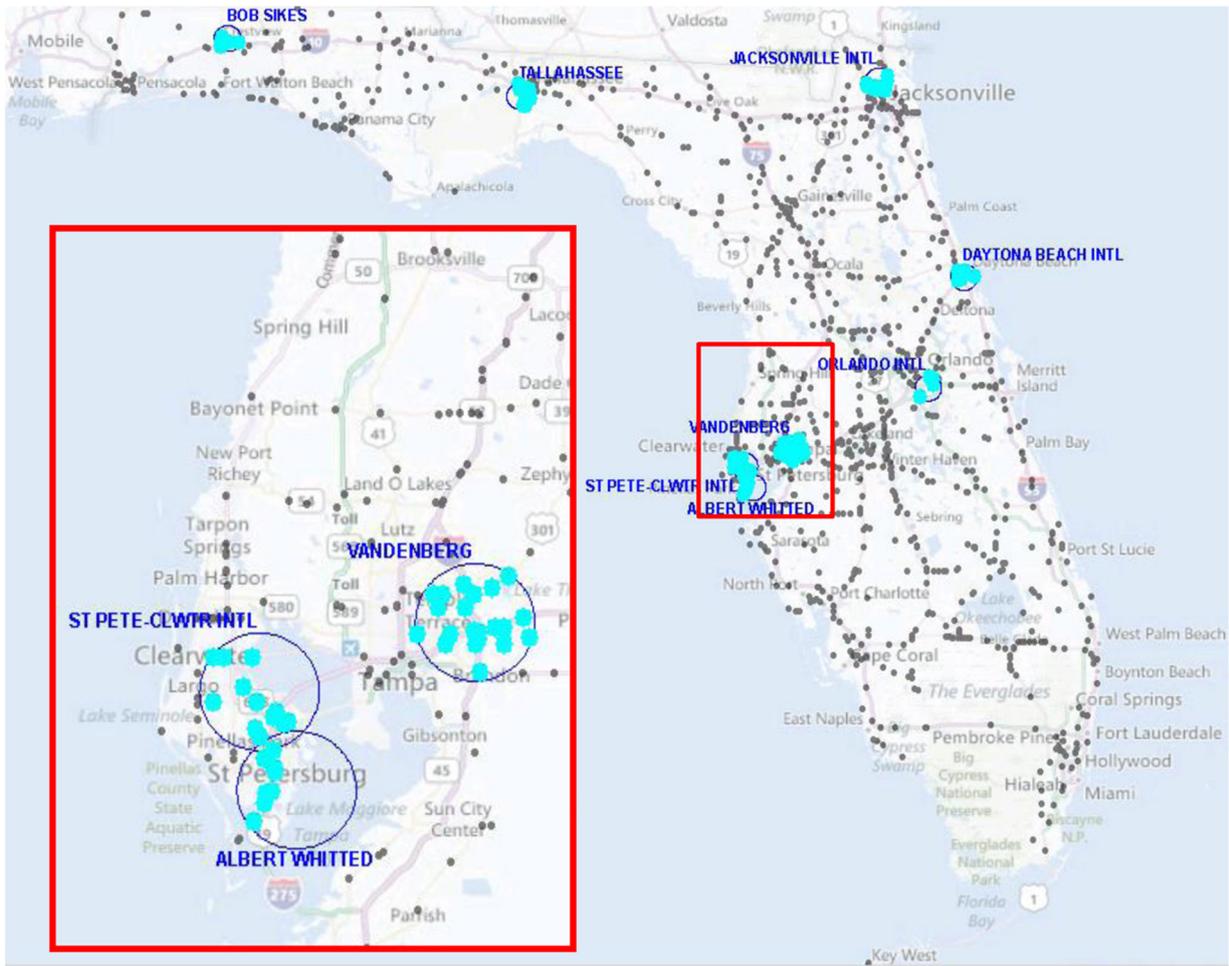
Fig. 5. Locations of eight chosen airports and fog crashes within the buffer (tourmaline blue dots) and fog crashes outside the buffer (black dots).

or $y = 0$ with respective probability p and $1 - p$. The logistic regression equation can be expressed as:

$$\log\left(\frac{p}{1-p}\right) = \beta_0 + \beta X \tag{1}$$

where β_0 is the intercept, β is the vector of coefficients for the explanatory variables, and X is the vector of

the explanatory variables. The logit function relates the explanatory variables to the probability of an outcome $y = 1$. The expected probability that $y = 1$ for a given value of the vector of explanatory variables X can be theoretically calculated as:

$$p(y = 1) = \frac{\exp(\beta_0 + \beta X)}{1 + \exp(\beta_0 + \beta X)} = \frac{e^{\beta_0 + \beta X}}{1 + e^{\beta_0 + \beta X}} \tag{2}$$
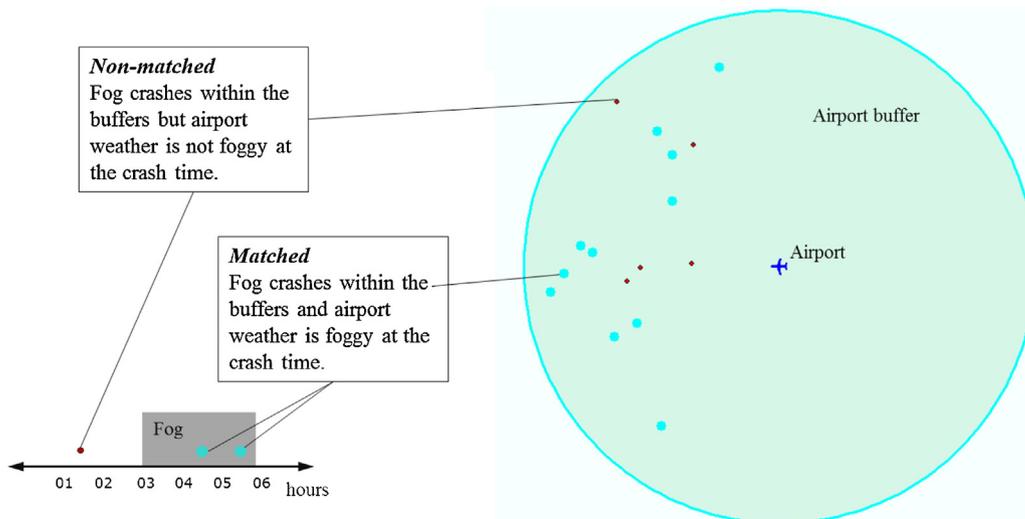

Fig. 6. Matched and non-matched fog crashes.

One advantage of the Bayesian approach over the classical model is the applicability of choosing the parametric family for prior probability distributions. There are three different priors that can be used (1) informative prior distributions based on the literature, experts' knowledge or explicitly from an earlier data analysis, (2) weak informative priors that do not supply any controversial information but are strong enough to pull the data away from inappropriate inferences, or (3) uniform priors or non-informative priors that basically allow the information from the likelihood to be interpreted probabilistically. In this study, uniform priors following normal distribution with initial values for the estimation of each parameter from the maximum likelihood method was used. Different types of prior distributions using the results from this study as prior could be considered for further research once more data become available to update the estimated models.

All models were estimated by Bayesian inference using the freeware Winbugs (Lunn et al., 2000). For each model, three chains of 10,000 iterations were set up in Winbugs based on the convergence speed and the magnitude of the dataset. The Deviance Information Criterion DIC, a Bayesian generalization of Akaike Information Criterion AIC, is used to measure the model complexity and fit. DIC is a combination of the deviance for the model and a penalty for the complexity of the model. The deviance is defined as $-2\log(\text{likelihood})$. The effective number of parameters, pD, is used as a measure of the complexity of the model, $pD = Dbar - Dhat$, where $Dbar$ is the posterior mean of the deviance, and $Dhat$ is a point estimate of the deviance for the posterior mean of the parameters. DIC is given by $\text{DIC} = Dhat + 2\ pD$ (Spiegelhalter et al., 2003). Moreover, receiver operating characteristic (ROC) curve analysis was used to compare across models.

5. Results and discussion

5.1. Model estimation, interpretation and diagnostics

As mentioned earlier that the main objective of this paper is to investigate whether the weather information from airport weather stations can be utilized as input for the nearby roadways' management systems. Two logistic regression models with Bayesian inference technique were calibrated. (i) The first model was meant to see if the visibility data from the airport weather stations can be used to detect the visibility conditions of the nearby roadways; the fog related crashes were compared to the non-obstructed vision (NOV) crashes. (ii) The second series of models were calibrated to check the influence of visibility conditions on crash occurrence; crashes with different thresholds for poor visibility were compared with randomly selected non-crash cases. Poor visibility crashes are defined as crashes occurred when the visibility was below 1-mile.

5.2. Fog vs. non-obstructed vision crashes

A total of 90 fog related crashes were documented within the study area and 360 non-obstructed vision crashes were randomly selected and matched to the fog related crashes. Bayesian logistic regression model was employed here along with weather information from the airport detectors and the geometry characteristics and traffic data e.g. AADT, roadway type, number of lanes, etc. Table 3 shows the model results and fit statistics. Only visibility came out to be significant with a negative sign, which demonstrated that poor visibility conditions may increase the likelihood of crash occurrence. None of the roadway geometry or traffic parameters came out to be significant in the model, however, the model has a large ROC area of 0.937 which indicates that the hazardous visibility condition can be easily classified using data collected from airport

Table 3
Fog vs. non-obstructed vision crashes model.

Variable	Estimate	Std. error	2.5%	97.5%
Intercept	2.639	0.377	1.983	3.471
Visibility	−0.644	0.056	−0.758	−0.539
ROC area	0.937			
DIC	179.2			

Table 4
Sample sizes of different thresholds for poor visibility condition.

Visibility threshold (mile)	Crash sample	Non-crash sample	Total
0.1	51	204	255
0.2	72	288	360
0.3	242	968	1210
0.4	242	968	1210
0.5	390	1560	1950
0.6	390	1560	1950
0.7	390	1560	1950
0.8	504	2016	2520
0.9	504	2016	2520
1	643	2572	3215

weather stations. It also indicates that in case of poor visibility, this factor dominates all other geometric and traffic factors.

5.3. Poor visibility crashes vs. non-crash cases

The preliminary analysis as well as the abovementioned model indicated that weather information and particularly visibility collected from the airport weather stations can be used to represent the nearby roadways visibility conditions with high confidence level. Here we make a step forward to investigate the influence of visibility conditions on crash occurrence. Different thresholds of poor visibility conditions have been considered by comparing poor visibility crashes and randomly select non-crash cases. Thresholds for poor visibility conditions were attempted from a range of 0.1 mile to 1 mile with 0.1 mile increment. Detailed sample sizes for each threshold were shown in Table 4. It can be seen that there are mainly six distinct datasets with different crash sample sizes. Logistic regression models with the Bayesian inference approach have been used for each specific data and the estimate coefficients, credible intervals and model fits are shown in Table 5. Visibility

Table 5
Crash vs. non-crash with different thresholds of visibility condition.

| Thresholds	Variable	Estimate	Std. error	2.5%	97.5%
0.3 Mile	Intercept	−11.96	0.942	−13.94	−10.04
	Visibility	−2.448	0.764	−3.899	−0.948
	Log (AADT)	1.089	0.091	0.906	1.283
	ROC	0.723			
	DIC	1088.01			
0.5 Mile	Intercept	−11.8	0.684	−13.15	−10.41
	Visibility	−1.36	0.358	−2.048	−0.670
	Log (AADT)	1.065	0.066	0.928	1.196
	ROC	0.720			
	DIC	1752.64			
0.8 Mile	Intercept	−12.04	0.522	−13.29	−10.86
	Visibility	−0.658	0.229	−1.102	−0.217
	Log (AADT)	1.071	0.059	0.957	1.19
	ROC	0.715			
	DIC	2269.96			
1 Mile	Intercept	−11.91	0.528	−13.03	−10.88
	Visibility	−0.492	0.149	−0.786	−0.202
	Log (AADT)	1.057	0.051	0.959	1.164
	ROC	0.715			
	DIC	2893.41			
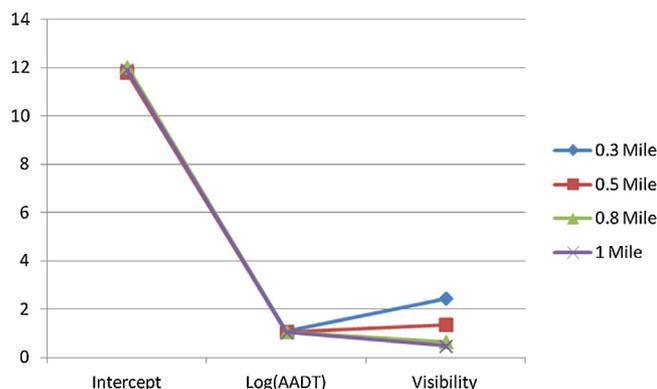
Fig. 7. Absolute values of estimated coefficients for different thresholds models.

thresholds of 0.1 mile and 0.2 mile were not considered because of the small sample size of crash cases. It is worth mentioning that there were no difference in the number of crashes at 0.4, 0.6, 0.7, and 0.9-mile than in previous levels and these cutoff values were excluded.

Results of the Bayesian logistic regression models showed that visibility conditions and Log (AADT) are significant related to the crash occurrence under the poor visibility conditions. Visibility conditions are negatively correlated with the crash occurrence which indicates that the increase of visibility would bring down the crash occurrence probability. Moreover, the Log (AADT) is significant with a positive sign which can be interpreted as higher exposure would result in a high risk of crash occurrence. The four models show similar classification ability since the ROC areas are almost the same. Fig. 7 shows the absolute values of the coefficient estimates for the four models which were presented as dots and linked with lines for each model. From the figure it can be illustrated that the estimated values of Intercept and coefficients of Log (AADT) barely change while the coefficients of the visibility parameter decrease significantly as the visibility threshold increases. This phenomenon can be understood as that with the poor visibility thresholds increasing, the effects of visibility condition's influence on crash occurrence decreases.

6. Conclusion

The main objective of this research was to explore the feasibility of utilizing the weather information collected from airports in crash risk assessment in real-time on highways. The study analyzed 6 years (2005–2010) of crash data in Florida, hazardous locations of fog-prone crashes were identified and ranked using GIS cluster analysis KDE. In the most hazardous counties in Florida, it was found that airports can provide more than 60% of spatial-temporal coverage of weather conditions in real-time for all crashes and more than 82% of the determined 5 nautical miles buffer zone.

Weather data collected from 8 airports in Florida were combined with traffic crashes within 5 nautical miles (NM) buffers of airports. Logistic regression models with Bayesian inference were calibrated to examine the effect of airport weather data, traffic, and roadway geometry data on crash occurrence. It was proven statistically that the reduction in visibility reported by airports' weather stations is associated with crash occurrence. This means that we can identify the crash-prone conditions if this reduction in visibility is observed from nearby airports. Other issues such as the comparison between data reported at different airports and the overlap between airports coverage may need further research.

The results from this study indicate that the available real-time airport weather data can be utilized by traffic management centers (TMC) to mitigate the increased risk of limited-visibility. Airport weather stations can be a reasonably reliable source to determine visibility conditions of the roadways within 5 nautical miles radius around airports, and may be more depending on the location, type and overlap of buffers around airports. TMCs can benefit from the availability of real-time weather data from airports and utilize it with relatively negligible cost.

References

Abdel-Aty, M., Pemmanabonia, R., 2006. Calibrating a real-time traffic crash-prediction model using archived weather and ITS traffic data. IEEE Trans. Intell. Transport. Syst. 7 (2), 167–174.

Abdel-Aty, M.A., Oloufa, H., Hassan, M., Ahmed, C., Siddiqui, A.E., Huang, H., 2010. Developing an Early Detection System for Reduced Visibility. (FDOT Report BDK78 562-1). University of Central Florida, Orlando, FL.

Ahmed, M.M., Abdel-Aty, M., Qi, S., Abuzwidah, M., 2014. Synthesis of state-of-the-art in visibility detection systems' applications and research. J. Transport. Saf. Secur. 6 (3), 2014.

Ahmed, M., Abdel-Aty, M., Yu, R., 2012. Assessment of the interaction between crash occurence, mountainous freeway geometry, real-time weather and avi traffic data. J. Transport. Res. Board 2280, 51–59.

Anderson, T.K., 2009. Kernel density estimation and K-means clustering to profile road accident hotspots. Accid. Anal. Prevent. 41 (3), 359–364.

Andreescu, P., Frost, D., 1998. Weather and traffic accidents in Montreal, Canada. Clim. Res. 9 (3), 225–230.

Andrey, J., Olley, R., 1990. Relationships between weather and road safety, past and future directions. Climatol. Bull. 24 (3), 123–137.

Andrey, J., Mills, B., Leahy, M., Suggett, J., 2003. Weather as a chronic hazard for road transportation in Canadian cities. Nat. Hazards 28, 319–343.

Baltz, W., Zhu, J., 1994. Nebelwarnsystem A8 Hohenstadt-Riedheim. Wirkungsanalyse. Landesamt für Strassenwesen. Baden-Württenberg & PTV Consult GmbH.

Brodsky, H., Hakkert, S., 1988. Risk of a road accident in rainy weather. Accid. Anal. Prevent. 20 (2), 161–176.

Chainey, S., Ratcliffe, J.H., 2005. GIS and Crime Mapping. John Wiley and Sons, London.

Cools, M., Moons, E., Wets, G.,2008. Assessing the impact of weather on traffic intensity. In: Proceedings of the 87th annual meeting of the Transportation Research Board. CD-ROM. Transportation Research Board of the National Academies, Washington, DC.

Cooper, B.R., Sawyer, H.E., 1993. Assessment of M25 automatic fog-warning system: final report. TRL Project report 16, Crowthorne, 11pp.

Edwards, J.B., 1999. Speed adjustment of motorway commuter traffic to inclement weather. Transport. Res. F 2, 1–14.

Ekram, A.A.M.Y., (Doctoral dissertation) 2009. Reduced visibility related crashes in Florida: crash characteristics, spatial analysis and injury severity. University of Central Florida, Orlando, FL.

Fotheringham, A.S., Brunsdon, C., Charlton, M., 2000. Quantitative Geography: Perspectives on Spatial Data Analysis. Sage, London.

Goodwin, L., 2003. Best Practices for Road Weather Management, Version 2.0, Report No. FHWA-OP-03-081.

Hogema, J.H., van der Horst, R., van Nifterick, W., 1996. Evaluation of an automatic fog-warning system. Traffic Eng. Control, 629–632.

Koetse, M.J., Rietveld, P., 2009. The impact of climate change and weather on transport: an overview of empirical findings. Transport. Res. D 14, 205–221.

Kuo, P.F., Lord, D., Walden, T.D., 2013. Using geographical information systems to organize police patrol routes effectively by grouping hotspots of crash and crime data. J. Transp. Geogr. 30, 138–148.

Lunn, D.J., Thomas, A., Best, N., Spiegelhalter, D., 2000. WinBUGS-a Bayesian modeling framework: concepts, structure, and extensibility. Statist. Comput. 10, 325–337.

Maze, H.T., Agarwal, M., Burchett, G., 2006. Whether weather matters to traffic demand, traffic safety, and traffic operations and flow. Transport. Res. Rec. 1948, 170–176.

National Traffic Safety Board (NTSB), 1980. Fatal Highway Accidents on Wet Pavement – The Magnitude Location and Characteristics. NTIS. HTSB-HSS-80-1, Springfield, VA.

Rämä, P., Schirokoff, A., 2004. Effects of weather-controlled variable speed limits on injury accidents. In: 11th World Congress and Exhibition on Intelligent Transport Systems and Services, Proceedings, CD-ROM. ERTICO, ITS Europe, Budapest 24–26 May.

Rämä, P., 1997. Sää-ja kelitietoon perustuvan liikenteen ohjausjärjestelmän vaikutukset Kotka-Hamina-moottoritiellä (effects of the weather-controlled traffic management system in the motorway section between Kotka and Hamina), Finnra reports 1/1997. Finnish National Road Administration, Helsinki.

Rämä, P., Raitio, J., Anttila, V., Schirokoff, A., 2001. Effects of weather controlled speed limits on driver behaviour on a two-lane road. In: Proceedings of Traffic Safety on Three Continents, International Conference, Moscow 19–21 September 2001, TRB, CSIR, VTI.

Schneider, R.J., Ryznar, R.M., Khattak, A.J., 2004. An accident waiting to happen: a spatial approach to proactive pedestrian planning. Accid. Anal. Prevent. 36 (2), 193–211.

Shinar, D., Treat, J.R., McDonald, S.T., 1983. The validity of police reported accident data. Accid. Anal. Prevent. 15 (3), 175–191.

Spiegelhalter, D., Best, N., Carlin, B., Linde, V., 2003. Bayesian measures of model complexity and fit (with discussion). J. R. Stat. Soc. B 64 (4), 583–616.

Várhelyi, A., 2002. Dynamic speed adaptation in adverse conditions. IATSS Res. 26 (2), 2002.