Accident Analysis and Prevention 128 (2019) 103–113

Contents lists available at ScienceDirect

Accident Analysis and Prevention

journal homepage: www.elsevier.com/locate/aap

Estimation of heavy vehicle-involved rear-end crash potential using WIM data

Young Jo[a], Cheol Oh[a], Seoungbum Kim[b,*]

[a] Department of Transportation and Logistics Engineering, Hanyang University Erica Campus, 55 Hanyangdaehak-ro, Sangnok-gu, Ansan, 15588, Republic of Korea
[b] Division of Architectural, Urban, and Civil Engineering, Engineering Research Institute, Gyeongsang National University, 501, Jinju-daero, Jinju-si, Gyeongnam, 52828, Republic of Korea

ARTICLE INFO

Keywords:
CRP
Rear-end crash
WIM
Heavy vehicle

ABSTRACT

Rear-end crash risk has been modeled and calibrated from various sources of data. The results can be used to develop traffic control strategies for avoiding collision. As the demand for freight transportation has increased in recent years, traffic agencies are more interested in monitoring heavy vehicle-involved crash risk because of the severity of accidents and high secondary crash risks. This study processes per-vehicle data from a WIM (weight-in-motion) sensor to investigate the heavy vehicle-involved crash risk potential (CRP) on freeways. To estimate rear-end CRP, one of the previously developed models is modified to use WIM data. In addition, this study proposes a technique for estimating the crash risk level, which is determined by comparing the estimated crash risk probability with the maximum expected risk probability at a given level of service (LOS). A time-series investigation reveals that there is clear difference in the CRP and crash risk level between lanes and there are two major time windows with high heavy vehicle-involved crash risk levels. After repeating the exercise for all available data, this study confirms that these observations are systematic and reproducible. Not only can the results in this study provide a local guidance for traffic agencies to prepare for heavy vehicle crash prevention strategies, but reiterating the investigation over multiple locations also allows them to monitor heavy vehicle crash risk levels at the network level.

1. Introduction

On freeways, drivers interact with adjacent drivers continuously by following other vehicles and conducting lane-changing maneuvers. A traffic accident occurs when the interaction becomes unbalanced and unstable. Since the demand for freight transportation has increased in recent years, the chances of conflict between a passenger vehicle and a heavy vehicle or heavy vehicles have increased. These increased conflicts involving heavy vehicles have resulted in an increased number of accidents, which are linked with fatal accidents or accidents with severe injuries. The relative risk of a crash involving a large truck is 1.2 to 1.55 times higher than that of a passenger car (Moonesinghe et al., 2003). Thus, the identification of microscopic driving patterns of heavy vehicles and the prediction of heavy vehicle involved accidents would be crucial resources for transportation agencies to develop proactive accident risk management strategies.

Crash risk has been investigated using various kinds of data. Kim et al. (2007) estimated rear-end crash risks using historical accident data. Wang and Abdel-Aty (2006) also employed data associated with intersection geometric features, control and operational features, traffic data, and crash data to estimate rear-end crashes at signalized intersections. Another branch of rear-end crash related studies employed a concurrent loop detector and historical crash records together to develop real time crash risk estimation or prediction models (Abdel-Aty et al., 2004; Pande and Abdel-Aty, 2006a; Pande and Abdel-Aty, 2006b; Hossain and Muromachi, 2011 and 2013, Xu et al., 2013a, b). However, measurements from conventional loop detector data, such as speed, occupancy, and volume, are aggregated over 1 min or 5 min. Thus, it may not be appropriate to analyze and capture crash events that occur instantaneously from the instability of driving behaviors among individual vehicles (Zhao, P., and Lee, C., 2018). To this end, many researchers prefer to develop various surrogate safety measures after extracting individual vehicle trajectories with the aid of traffic simulation or NGSIM (Next Generation Simulation). Few studies have tried to extract hazardous traffic events directly from surveillance systems (Oh et al., 2006; Ozbay et al., 2008; Saunier and Sayed, 2008; Saccomanno et al., 2008; Oh et al., 2009; Oh and Kim, 2010, Dimitriou et al., 2018, Park et al., 2018). There were other trials that investigated

* Corresponding author. Tel: +82 55 772 1778; fax: +82 55 772 1779
E-mail addresses: young2@hanyang.ac.kr (Y. Jo), cheolo@hanyang.ac.kr (C. Oh), kimsb@gnu.ac.kr (S. Kim).

https://doi.org/10.1016/j.aap.2019.04.005
Received 27 September 2018; Received in revised form 9 March 2019; Accepted 4 April 2019
Available online 13 April 2019
0001-4575/

the risk of rear-end collisions using an interactive driving simulator (Alicandri, 1994; Jenkins and Rilett, 2004; Broughton et al., 2007; Bella, 2009; Farah et al.,2009; Bella and Russo, 2011).

There are few efforts on estimating or investigating the risk of heavy vehicle-involved crashes. Abdel-Aty and Abdelwahab (2003) developed mathematical models to investigate how the limited visibility of a passenger driver obscured by a light truck vehicle is associated with a rear-end crash potential. Bachmann et al. (2012) investigated truck-related conflicts using a revised definition of TTC (time-to-collision) on a truck-only highway. It focused on the improvement of a revised TTC against a simplified TTC definition rather than truck related crash risk potential. Hjelkrem and Ryeng (2016) investigated drivers' risk perception under various adverse weather conditions by vehicle type. Rather than using TTC, they employed chosen risk index that is a function of speed, weight, and time gap between vehicles. Wu et al. (2018a,Wu et al., 2018b developed a mathematical model, which improves TTC, to reveal the relationship between the crash risk and the reduced visibility under fog conditions. Malin et al. (2019) also investigated accident risk under different weather conditions using the empirical relative accident risk index that is a function of crash records and road geometry information rather than TTC but this study did not consider vehicle type in the analysis. Gudes et al. (2017) investigated various historical crash data using a GIS approach to visualize the spatial-temporal patterns of articulated heavy vehicle crashes. Zhao and Lee (2018) revised the rear-end collision risk model of Cunto and Saccomanno (2008), such that it accounts for differences in driver behavior between passenger cars and heavy vehicles. One of difficulties in analyzing heavy vehicle crash risk is to obtain the required per-vehicle data, such speed and spacing, for different types of vehicles. To overcome this limitation, previous studies often used a microscopic traffic simulation such as VISSIM (e.g., Zhao and Lee, 2018) or made an extra effort to process video data to extract those measurements (e.g., Oh and Kim, 2010).

Unlike other studies, we examine WIM data to estimate the heavy vehicle-involved crash potential on freeways because it is capable of providing all the necessary inputs for crash risk potential estimation as well as vehicle class. The WIM (Weigh-In-Motion) system is designed to record vehicles' gross weight information based on the axle weight as vehicles pass over the system. There are two types of WIM systems: low-speed WIM (LS-WIM), which is capable of weighing a vehicle at speeds up to 35 kph (20 mph), and high-speed WIM (HS-WIM), which can weigh the vehicle from 35 kph to 120 kph. Transportation agencies install WIM systems and process the weight data for various reasons: to monitor volumes of vehicles per vehicle class, to evaluate transportation impacts on infrastructure and air quality, and to enforce overloaded vehicles. Recently, HS-WIM has become widely accepted for enforcement purposes on freeways. Even though the data collectable by HS-WIM is available for other purposes than overload vehicle enforcement, there is a lack of effort to enlarge its applicability.

This study seeks to estimate the heavy vehicle-involved rear-end crash potential leveraging a concurrent HS-WIM sensor. First, this study modified the model to estimate the rear-end crash potential, which was originally proposed by Oh and Kim (2010), so that it can export the heavy vehicle related risk estimation. The crash risk was often represented by a probability or a nonintuitive index. The output from the rear-end crash potential model in this study is also a probability ranging between 0 and 1, and the marginal probability value for two successive time frames may not be informative for transportation agencies. Therefore, we propose a technique to estimate the crash risk level, which is determined by comparing the estimated crash risk probability with the maximum expected risk probability at a given level of service (LOS). By converting the crash risk probability to the crash risk level, one can intuitively judge the status of a crash risk under the current traffic stream condition, which can help traffic agencies develop various proactive strategies and plan their time schedules to avoid or mitigate the potential of a heavy vehicle-involved crash.

The remainder of this paper is organized as follows. The next section briefly reviews the rear-crash risk potential estimation algorithm and introduces the technique to evaluate the risk potential. The following section introduces the analysis area and WIM data sets used in the study. The paper then presents the results after examining the data set using the proposed methodology and highlighting several findings that can be derived by the comparison between the estimated crash potential and the maximum expected probability for a given LOS. The last section presents the conclusions, summarizes the results of the study, and discusses future research.

2. Methodology

As discussed, the rear-end crash risk has been modeled in various mathematical ways. This study further extended the model proposed in 2010 by Oh and Kim (2010) to estimate the probability of heavy vehicle-involved rear-end crashes using WIM data.

More details of the model are elaborated with a hypothetical example of a two-lane freeway section as shown in Fig. 1. Suppose that a vehicle 'n-1' is followed by the subject vehicle 'n' in the starting lane and another group of vehicles in the target lane are placed as depicted in Fig. 1. In addition, assume that the speed of vehicle 'n' is faster than that of vehicle 'n-1' in Fig. 1 (as of this point, it is called condition 'X'). Under condition 'X' the spacing between vehicle 'n' and 'n-1' decreases and vehicle 'n' should either change lanes or reduce its speed to avoid a collision with vehicle 'n-1'. Similarly, vehicle 'n-1' should either change lanes or increase its speed to avoid collision with vehicle 'n'. If they endure the condition 'X' where the spacing is kept short, then a rear-end crash can occur after some duration Δ t, which is often called TTC (time-to-collision). If the vehicle trajectories in Fig. 1 are available, TTC

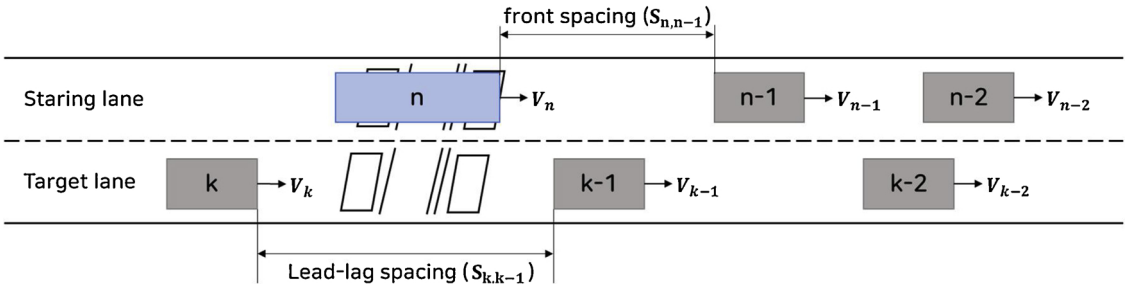

Fig. 1. Hypothetical example of two lane freeway section.

is calculated by Eq. (1).

$$\mathrm{TTC}_{n,n-1}(t) = \frac{S_{n,n-1}(t)}{V_n(t) - V_{n-1}(t)} \quad (1)$$

Where

$S_{n,n-1}(t)$: spacing between vehicle n and n-1 at time t

$V_n(t)$: speed of vehicle n at time t

Then, the probability that a rear-end crash would occur at time t can be estimated by Eq. (2) when the condition 'X' is maintained. Note that this study assumes that "lane change maneuvers for subject and front vehicles and collision are independent" at time t as described in Oh and Kim (2010).

$$P^{(t)}(Crash_{n,n-1}) = P^{(t)}(NLC_nX_n) \times P^{(t)}(NLC_{n-1}X_{n-1}) \times P^{(t)}(Crash_{n,n-1}TTC_{n,n-1}) \quad (2)$$

Where

$P^{(t)}(NLC_nX)$: Probability that the subject vehicle 'n' does not change the lane under a given condition X at time t

$P^{(t)}(Crash_{n,n-1}TTC_{n,n-1})$: Probability that vehicle 'n' collides with vehicle n-1 under a given TTC at time t

To estimate the probability of a crash on vehicle n in Eq. (2), calculations of $P^{(t)}(NLC_n|X)$ and $P^{(t)}(Crash_{n,n-1}|TTC_{n,n-1})$ are required. According to Oh and Kim (2010), the binary logistic regression model in Eq. (3a) has been adapted to calculate the first probability component, where the dependent variable should be '1' if the subject vehicle 'n' does not change lanes and '0' otherwise. The speed of vehicle 'n' and 'n-1,' the spacing between 'n' and 'n-1,' and the spacing between vehicle 'k' and 'k-1' are considered independent variables in this study.

$$P^{(t)}(NLC_nX_n) = \frac{exp[f(X_n, \beta)]}{1 + exp[f(X_n, \beta)]} \quad (3a)$$

Where

X_n: a vector of independent variables associated with the lane change

$f(X_n, \beta)$: a function of X_n and a parameter vector β to be estimated

$$\frac{\exp[f(X_n, \beta)]}{1 + \exp[f(X_n, \beta)]} = \frac{\exp(-0.045\mathrm{V_n} - 0.083\mathrm{V_k} - 0.046S_{n,n-1} + 0.023S_{k,k-1} + 11.476)}{1 + \exp(-0.045\mathrm{V_n} - 0.083\mathrm{V_k} - 0.046S_{n,n-1} + 0.023S_{k,k-1} + 11.476)} \quad (3b)$$

Where

$\mathrm{V_n}$: the speed of the subject vehicle

$\mathrm{V_k}$: the speed of the lag vehicle

$S_{n,n-1}$: the spacing between the subject and the front vehicle in the current lane

$S_{k,k-1}$: the spacing between the lag and the lead vehicle in the target lane

With regard to the second probability component in Eq. (2), a generalized exponential decay function (EDF) in Eq. 4 is calibrated to estimate the crash probability for a given TTC.

$$P^{(t)}(Crash_{n,n-1}TTC_{n,n-1}) = \mathrm{a} + \mathrm{b} \times \exp\left(-\frac{1}{c}\left(\frac{S_{n,n-1}(t)}{V_n(t) - V_{n-1}(t)}\right)\right) \quad (4a)$$

Where

$S_{n,n-1}(t)$: the front spacing that the subject vehicle follows the front vehicle at time step t

$V_n(t)$: the speed of the subject vehicle at time step t

$V_{n-1}(t)$: the speed of the front vehicle at time step t

a, b: initial values

c: coefficient of curvature determination

$$P^{(t)}(Crash_{n,n-1}TTC_{n,n-1}) = \exp\left(-\frac{1}{1.87}\left(\frac{S_{n,n-1}(t)}{V_n(t) - V_{n-1}(t)}\right)\right) \quad (4b)$$

In Oh and Kim, the models for $P^{(t)}(NLC_n|X)$ and $P^{(t)}(Crash_{n,n-1}|TTC_{n,n-1})$ were calibrated via vehicle trajectories obtained from a video data reduction process. They used 1 h of video footage recorded on a basic segment of a two lane freeway. Each vehicle trajectory is extracted every 1/10 s using the Premier software. Other variables, such as speed and spacing of vehicles in starting and target lanes, are also calculated from the vehicle trajectories. After taking a set of dependent and independent variables, the BLR and EDF models are calibrated as expressed in Eq. (3b) and Eq. (4b). After putting all components together, Eq. (2) is revised as follows.

$$P^{(t)}(Crash_{n,n-1}) = \left[\frac{\exp[f(X_n, \beta)]}{1 + \exp[f(X_n, \beta)]}\right] \times \left[\frac{\exp[f(X_{n-1}, \beta)]}{1 + \exp[f(X_{n-1}, \beta)]}\right] \times \left[\exp(-\frac{1}{c}\left(\frac{S_{n,n-1}(t)}{V_n(t) - V_{n-1}(t)}\right)\right] \quad (5a)$$

$$P^{(t)}(Crash_{n,n-1}) = 4\left[\frac{\exp[-0.045V_n - 0.083V_k - 0.046S_{n,n-1} + 0.023S_{k,k-1} + 11.476]}{1 + \exp[-0.045V_n - 0.083V_k - 0.046S_{n,n-1} + 0.023S_{k,k-1} + 11.476]}\right] \times \left[\frac{\exp[-0.045V_{n-1} - 0.083V_{k-1} - 0.046S_{n-1,n-2} + 0.023S_{k-1,k-2} + 11.476]}{1 + \exp[-0.045V_{n-1} - 0.083V_{k-1} - 0.046S_{n-1,n-2} + 0.023S_{k-1,k-2} + 11.476]}\right] \times \left[\exp(-\frac{1}{1.87}\left(\frac{S_{n,n-1}(t)}{V_n(t) - V_{n-1}(t)}\right)\right] \quad (5b)$$

For the goodness of fits for the BLR and other details associated with the model development and calibration, please refer Oh and Kim (2010).

The probability estimation itself may not be useful from a practical point of view and its marginal value between two successive time steps is not informative. Thus, previous studies have developed surrogate safety measurements, such as TTC, PET, or CRI (Crash Risk Index) to quantitatively measure the overall level of the crash risk potential. Those measurements are often aggregated over a certain time period (e.g., 1 min or 5 min) at a given location, and their trends are monitored as a function of time. However, it is not a convenient measure to investigate heavy vehicle-involved crash events because of their aggregation feature. There are other researches that employ vehicle trajectory data to derive surrogate safety measures without aggregating over time period (Oh and Kim, 2010; Dimitriou et al., 2018; Park et al., 2018) but the measure itself is still less informative and not intuitive. To this end, this study put extra effort towards developing another surrogate measurement, which is more intuitive and conveniently applicable for the investigation of heavy vehicle crash risk potential.

Revisiting Eq. (4), one can calculate the maximum expected probability of a crash for a given TTC and LOS. The numerator of the multiplier in Eq. (4) is the spacing between two consecutive vehicles at time t. To calculate the minimum and maximum expected spacing, one calculates the minimum and maximum expected headway from the maximum and minimum expected volume from Table 1 (tabulated in

Table 1
LOS Criteria for basic freeway segments (quoted from HCM2000).

Criteria (when speed limit is 100 kph)	LOS				
	A	B	C	D	E
Maximum Density (pc/km/ln)	7	11	16	22	28
Minimum Speed (km/h)	100	100	100	93.8	82.1
Maximum v/c	0.3	0.48	0.7	0.9	1.0
Maximum Service flow rate (pc/h/ln)	700	1,100	1,600	2,065	2,300

Table 2
Maximum expected probability per LOS.

LOS	Min Vol.	Max Vol.	Min H	Max H	Min S	Max S	Min $S_{n,n-1}/v_n$-v_{n-1}	Max $S_{n,n-1}/v_n$-v_{n-1}	Min P(Crash|TTC)	Max P(Crash|TTC)
A	0	700	5.14	Inf	0.14	Inf	12.86	Inf	0	0.001
B	700	1,100	3.27	5.14	0.09	Inf	8.18	Inf	0	0.01
C	1,100	1,600	2.25	3.27	0.06	Inf	5.63	Inf	0	0.05
D	1,600	2,065	1.74	2.25	0.05	Inf	4.36	Inf	0	0.10
E	2,065	2,300	1.57	1.74	0.04	Inf	3.91	Inf	0	0.12

the 4th and 5th column in Table 2). Then the minimum and maximum spacing can be calculated based on the speed limit (e.g., 100 km/h) at the study site (shown in the 6th and 7th column in Table 2). The denominator of the multiplier in Eq. (4) is the difference in speed between the two vehicles. While the denominator is an unknown variable, one can use a time-series speed from WIM data to determine the value (i.e., it turns out that the value is 40 kph in this study). After calculating the minimum and maximum multiplier (shown in 8th and 9th column), then the minimum and maximum expected probability of crash for a given LOS and TTC is calculated and tabulated in 10th and 11th column of the Table 2. In fact, the maximum expected probability of a crash for a given LOS and TTC in Table 2 is equal to the maximum expected probability of crash for a given LOS because the maximum value of the $P^{(t)}(NLC_nX)$ is equal to 1 in Eq. (2).

3. Analysis area and data collection

This study employed WIM data to estimate the probability of heavy vehicle-involved crash risk potential. The WIM data has been extracted from a WIM station located within a basic freeway section in the Jungbunaeryuk expressway in Korea. The expressway is one of the major freeways connecting the capital with a mid-size city, Changwon, and is 302.5 km long. Since Changwon is located near two metropolitan cities, Busan and Ulsan, which are the top two port cities in the country in terms of the volume of freights, the expressway is one of the major routes for trucks and truck-trailer units coming from or going into the port cities. Geographical information of the expressway and details of the WIM station is depicted in Fig. 2.

The WIM station consists of sensors in the pavement and cameras on the top of the roadway to identify vehicle license plates. The sensors installed in the pavement are dual loop detectors that measure the vehicle speed, a piezoelectric sensor to detect the vehicle axle, and an axle load sensor to weigh the vehicle. Data from the WIM system includes time, speed, vehicle ID, vehicle length, class, acceleration rate, weight on each axle, and gross weight of each vehicle. Each vehicle is sorted into one of 12 vehicle classes based on the axle weight and the number of axles. Unlike the FHWA's 13 vehicle classes, classes 2, 4, 5, 6, and 7 are either buses or single unit trucks and classes 8 through 12 denote semitrailer trucks. In this study, 'heavy vehicle' includes single unit trucks and semitrailer trucks.

This study employed 18 days of WIM data on March in 2015 for southbound traffic only. After eliminating some missing data, a total of 779,845 per-vehicle records (477,475 in lane 1 and 302,370 records in lane 2) are analyzed to estimate the heavy vehicle-involved crash potential.

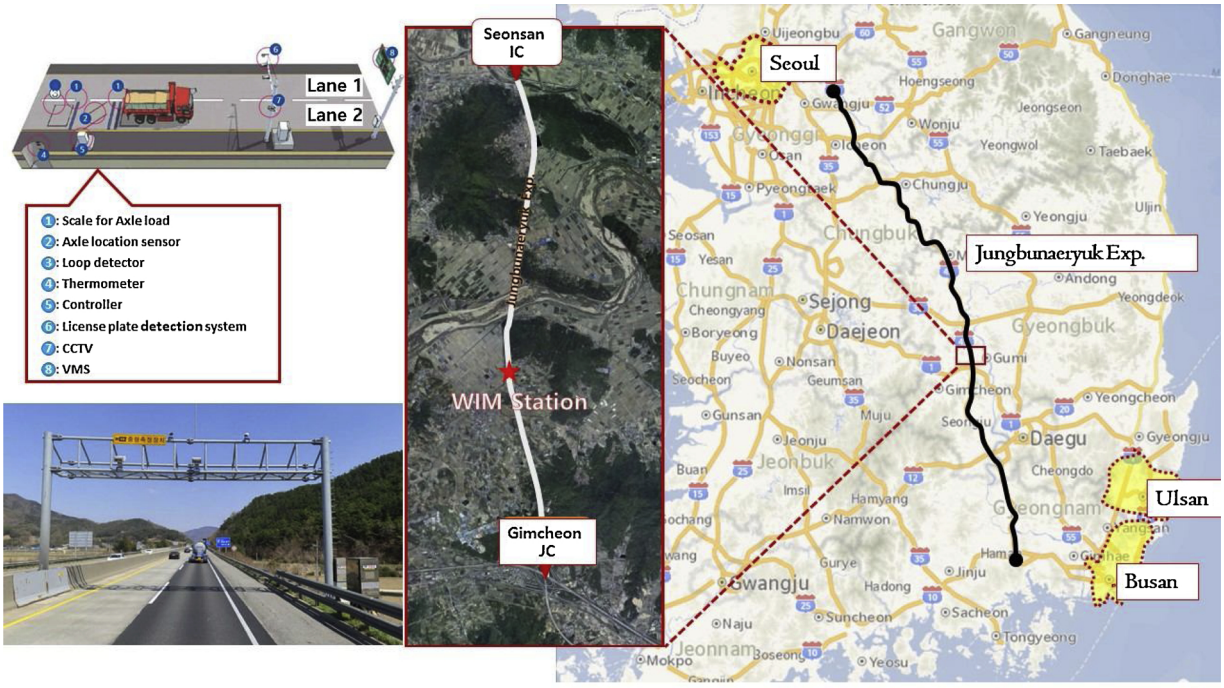

Fig. 2. Specifications and location of the WIM station on Jungbunaeryuk expressway in Korea.

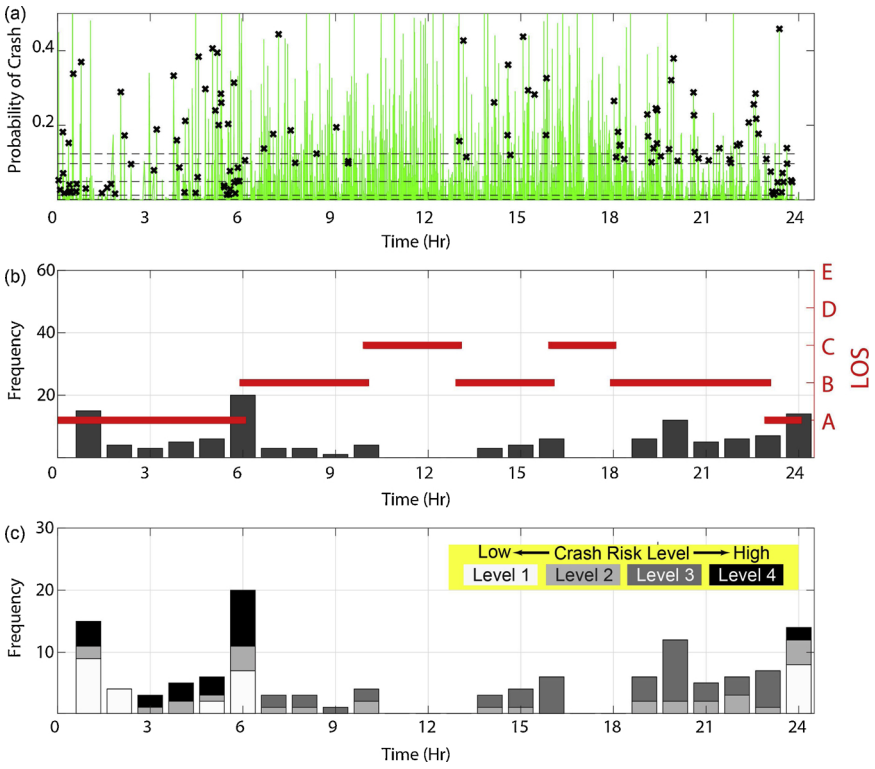

Fig. 3. (a) Probability of a crash higher than MECRP for a given LOS in lane 1, (b) Hourly frequency of the probability of a crash higher than MECRP for a given LOS and flow in lane 1, (c) Crash risk level in lane 1.

4. Results

First, the relative speed and spacing between the subject and the front vehicles in starting lane and the lag and lead vehicles in target lane are calculated from the WIM data set. They are then plugged into Eq. 5b to calculate the rear-end crash risk probability (CRP). The calculation is repeated for every pair of vehicles in lanes 1 and 2. Meanwhile, the vehicle class information in each pair is indexed with the CRP estimation to identify whether it has heavy vehicle-involved crash risk potential. To evaluate the heavy vehicle related crash risk status using the CRP in practice, this study proposed comparing the CRP with the maximum expected crash risk probability (MECRP) for a given LOS. If there is a case where the CRP is frequently higher than the MECRP within a short period of time, traffic agencies need to take special attention. However, the MECRP changes over time at a given location since the LOS changes over time. Thus, this study takes advantage of WIM data to determine LOS based on measured flow and calculates the MECRP.

Figs. 3a and 4 a show two time-series of the rear-end CRP estimated from Eq. 5b using one day of WIM data (March 2, 2015) in lanes 1 and 2, respectively. Five dashed horizontal lines indicate the MECRP in Table 2. We highlight the CRP with an 'x' shape if either one of vehicles in the pair is classified as a heavy vehicle by the WIM system and the estimated CRP is higher than MECRP for a given LOS.

Upon comparing lane 1 to lane 2, it is clear that the 'x' events occur more frequently in lane 2 because heavy vehicles are recommended to travel in outside lane. Regardless of lanes, the majority of the events of 'x' are distributed over nighttime. One can see these findings more easily in Fig. 3b, where the total number of 'x' events is tallied every hour. Each red horizontal line corresponds to the LOS derived from flow. In each 'x' event, the level of CRP varies. For example, there are some events where the CRP is barely above the MECRP but in other events the CRP is much higher than the MECRP. These two cases should be handled differently by traffic agencies. Figs. 3c and 4 c are the same as Figs. 3b and 4 b, respectively; they are just bar graphs with multiple stacks grouped together. The stacks are defined based on the vertical difference between the CRP and MECRP, which ranges from 1 to 4. (i.e., higher number indicates a higher crash risk). Each level in the stack is color coded differently (the darker color represents a higher crash risk). In Figs. 3c and 4 c, 'x' events with a high CRP level (e.g., levels 3–4) are focused during the nighttime, while they are observed more frequently in lane 2 than lane 1. Note that we don't barely see level 4 since the LOS before midnight is above B, but we do after midnight. The number of high CRP levels tends to be more stable during nighttime than daytime. This feature is more evident in lane 2. The following is a summarization of the results from the investigation with a time-series of CRP using one day of data.

- Heavy vehicle-involved CRP is higher in lane 2 than lane 1 due to lane utilization for different vehicle types.
- Heavy vehicle-involved CRP is higher during nighttime than daytime due to relatively less passenger vehicles but more heavy vehicles during nighttime.
- According to the crash risk level based upon the comparison between the CRP and MECRP, periods between 9 pm and 1 pm and between 6 a.m. and 9 a.m. are the time windows with the highest

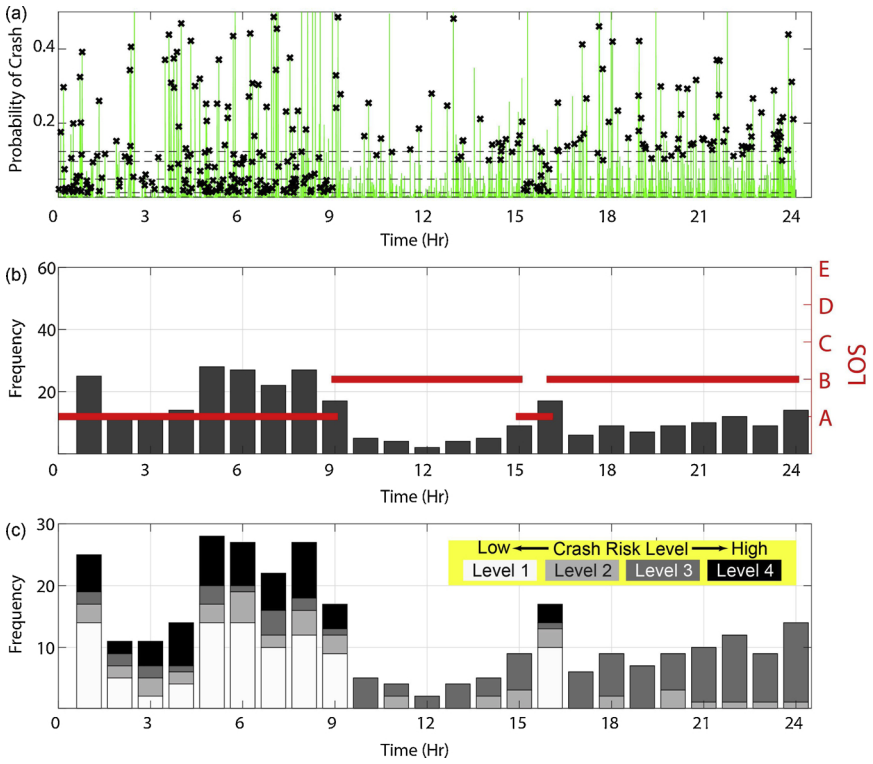

Fig. 4. (a) Probability of a crash higher than MECRP for a given LOS in lane 2, (b) Hourly frequency of the probability of a crash higher than MECRP for a given LOS and flow in lane 2, (c) Crash risk level in lane 2.

heavy vehicle-involved rear-end crash risk probability.

- The change in the total number of observations with high CRP levels is more stable during nighttime than daytime.

The high level of rear-end crash risk is a result of the large difference in relative speed while maintaining a short spacing between vehicles. To prevent such collisions or alleviate such crash risks, traffic agencies monitor the CRP over extended periods of time and determine hot spots in time to allocate their limited resources effectively for traffic safety surveillance. Thus, this study examines the remainder of the WIM data (17 days of data) by repeating the exercise for Figs. 3 and 4 to determine if the above results are reproducible over time. Fig. 5(a) and (b) shows contour plots of the hourly frequency for 'x' events after repeating the procedure to generate Figs. 3(b) and 4 (b) over 18 days of WIM data.

As shown in the color bar, the darker the color the higher the frequency of an 'x' event. Of those 18 days, 12 weekdays are included. Fig. 5(a) and (b) shows similar features that we derived from Figs. 3 and 4. The frequency of an 'x' event is much higher in lane 2 and the events are more concentrated during the nighttime regardless of the lanes. More specifically, there are two time windows when the events with a high CRP are evident: one between 9 pm and 1 a.m. and the second between 6 a.m. and 9 a.m. during the weekdays. The second period is not only the morning peak period but the time period when the proportion of heavy vehicles is still high in the freeway. In terms of the impact of a crash on the adjacent traffic stream, traffic agencies pay more attention to the later time window and potentially consider active strategies to enhance the road safety. In addition, this study regenerates Fig. 5a and b for a high CRP level (level 3–4) in Fig. 5c and d. The distribution of the high CRP in Fig. 5c and d are generally comparable with Fig. 5a and b. What we concluded in Fig. 5a and b does not change if we are more interested in a high CRP level.

In some rural areas, one can travel in a freeway section with more than two lanes. We suspect that crash risk characteristics in two lane section is different from three or four lane section. Thus, this study repeats the analysis in the two lane section in a three lane section as shown in Appendix A.

5. Conclusions

The WIM system is designed to classify a vehicle, evaluate the impact analysis on infrastructures or environment (e.g., air quality), or enforce overloaded vehicles on a freeway. We attempted to add one more dimension in terms of applicability of the WIM system. This study applied a mathematical model to estimate the heavy vehicle-involved rear-end crash potential by leveraging a concurrent HS-WIM sensor on a freeway. The probability was estimated from the model for every pair of vehicles passing the WIM sensor in a given lane. Using the per-vehicle record from the WIM, each pair is categorized according to whether it includes a heavy vehicle or more. Since the output of the model is a form of probability, this study developed a technique to convert the probability-based crash risk estimation to one of five crash risk levels. The current crash risk level is determined after comparing the current CRP against the MECRP for a given LOS. After investigating the time-series CRP estimated in a two lane freeway section, the heavy vehicle-involved CRP (called event 'x') in lane 2 was higher than lane 1, since

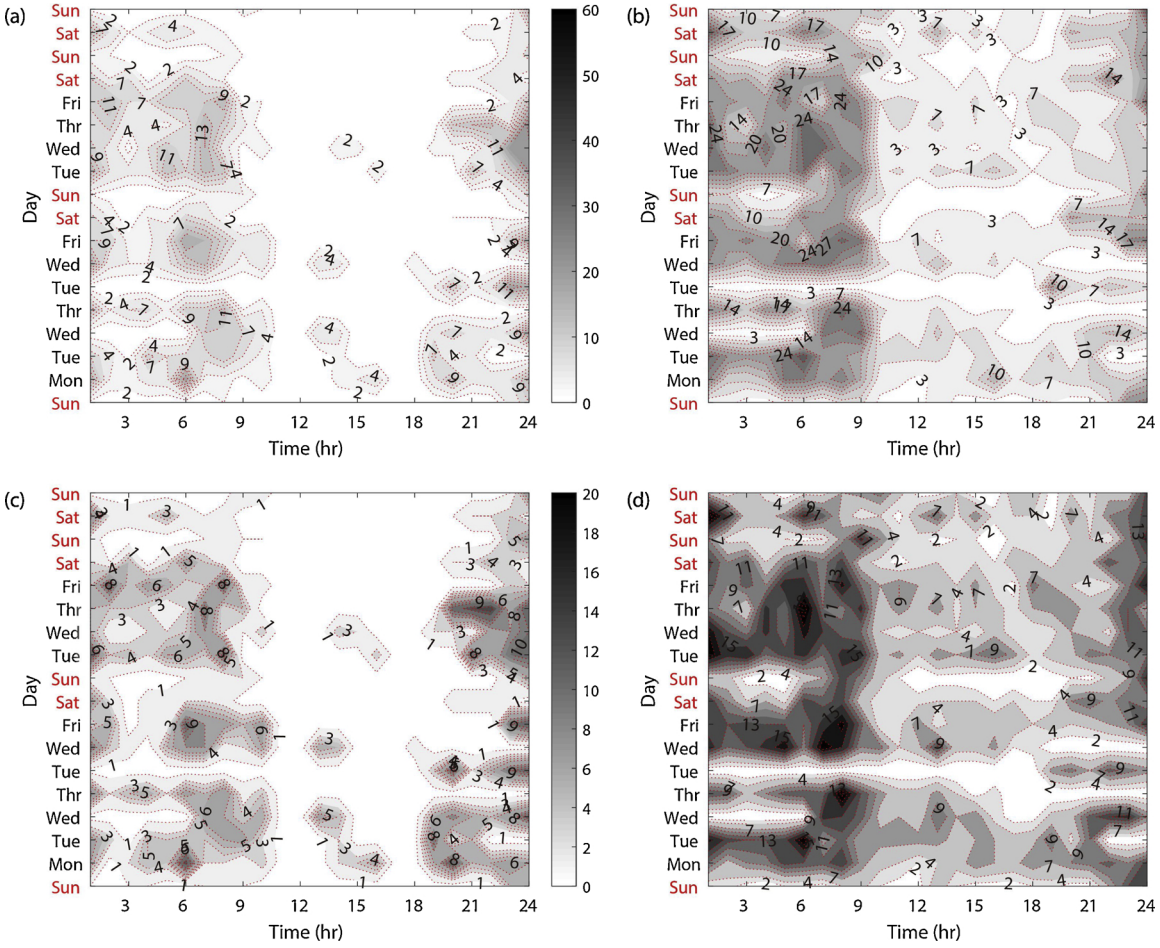

Fig. 5. (a) Contour plot of hourly frequency for all levels over 18 days of WIM data in lane 1, (b) Contour plot of hourly frequency for all levels over 18 days of WIM data in lane 2, (c) Contour plot of hourly frequency for high levels (level 3-4) over 18 days of WIM data in lane 1, (d) Contour plot of hourly frequency for high levels (level 3-4) over 18 days of WIM data in lane 2.

trucks tend to use the outside lane. A time-series analysis revealed the number of 'x' events increases gradually from 9 pm to 1 a.m. and then decreases for the next 3–4 h. It then increases again from 6 a.m. to 9 a.m. We suspect that the fluctuations over the two time periods are associated with lane utilization from different vehicle types and changes in vehicle compositions. After breaking the CRP into 4 crash risk levels, this study noticed that the proportion of the high risk levels (level 3–4) changes over time but it is more stabilized during the two time windows. All of these findings were reproducible for other weekdays. In addition, this study repeated the exercise conducted in two lane freeway section in a three lane freeway section. While we observed slightly different crash risk characteristics due to different traffic conditions between sites, they were reasonable explainable.

In summary, the time-series patterns in terms of heavy vehicle-involved CRP between lanes and the level of the crash risk are systematic and reproducible on each of those study sites. Therefore, we expect that this study can provide a local guidance to reduce heavy vehicle-involved crash risk when traffic agencies apply limited financial and labor resources. In methodological point of view, the technique proposed in this study has a few advantages when compared with other surrogate safety measures. While there are various surrogate safety measures conventionally used, the technique to calculate MECRP is differentiated from others in terms of defining crash risk potential levels. In the process of applying conventional surrogate safety measures, it is preceded that users classify the entire range of a measurement into several levels associated with different crash risks. In the SSAM (Surrogate Safety Assessment Model) developed by FHWA, for example, TTC is classified by 3 ranges whose boundaries are 0.5 s, 1.0 s, and 1.5 s. In general, enormous amount of actual collision data is necessary to validate those crash risk levels on a conventional surrogate safety measures but the validation task would be difficult "given the infrequent and random nature of crashes (Gettman et al., 2008)". Unlike other surrogate safety measures, MECRP is associated with a given LOS, which is one of widely accepted measures to evaluate freeway traffic streams. Thus, this technique does not need extra research to define a certain levels associated with crash risk potentials in application. More importantly it is not necessary to calibrate the level from one site to other sites. Secondly, suppose that the same crash risk potentials exist under different traffic stream conditions or different LOSs. Traffic operator would treat these cases differently because it more likely that a primary accident under worse traffic conditions leads to secondary accidents because of higher density. That is, the technique discussed here is more appropriate than conventional ones in terms of assessing the crash risk potential while considering underlying traffic conditions.

However, there are several tasks to be fulfilled in the future. First, we need to further evaluate the results here by comparing them against the historical crash records. Second, if more WIM data are available, this study can be extended to analyze monthly or seasonal crash risk patterns. Third, collision events and lane changing maneuvers could be correlated over time while we assumed they are independent in the calculation of crash risk potential. Thus, more in-depth studies applying time-series approach are required to make up for the limitation in future research. Finally, the WIM system is not available over a long freeway section; it is open operated on the periphery of freeway JC or IC. One alternative would be to employ other vehicle detection systems (e.g., inductive loop detectors) near a WIM station, which is generally installed densely in a freeway (e.g., every one-third of a mile within the urban area in Columbus, OH). Since one can classify vehicles based on magnetic profiles or vehicle length from the loop detector system, the methodology proposed in this study can be implemented.

Acknowledgement

This work was supported by the National Research Foundation of Korea grant funded by the KoreaGovernment(MSIP) (NRF-2017R1A2B4005835).

Appendix A

Repeating the crash risk analysis in a three lane freeway section

The three section freeway chosen for additional investigation is located in the Gyungbu expressway in Korea that connects between Seoul and Busan. Compared with the freeway section in Fig. 2, the volume of freights is lower so that less heavy vehicles are expected. Fig. A1 shows details of the three lane section. This study employed 27 days of WIM data on March in 2015 for northbound traffic only.

Figs. A2–A4 show a typical example in three lane freeway section corresponding to Figs. 3 and 4. Compared against two lane sections, it appears that events with high CRP (level 3 and 4) are distributed in lane 2 and lane 3 in three lane section. Unlike the two lane section, this location shows less truck volumes during nighttime so that we saw less heavy vehicle involved CRP samples. Investigating the results from the three lane section, Fig. A2 exhibits only few heavy vehicle involved observations where CRP is higher than MECRP in lane 1 because heavy vehicles are recommended to travel at outside lane (i.e., there are only 3% of heavy vehicles in lane 1). However, there are more such cases in lane 2 and lane 3 as shown in Figs. A3 and A4. Most of heavy vehicles traveling in lane 2 are heavy vehicles overtaking another heavy vehicles in lane 3. At the same time, majority of passenger vehicles are still traveling in lane 2 (52% of passenger vehicles in lane 2). Since heavy vehicles' speed is still slower than passenger vehicles' in lane 2, we suspect that the observations associated with higher CRP than MECRP in Fig. A3 result from the tailgating passenger vehicles behind heavy vehicles in lane 2. We do not see such cases in lane 1 due to few heavy vehicles. In fact, there are more heavy vehicles traveling in lane 3 (i.e., 72% in lane 3 and 25% in lane 2) and perhaps one expects to see higher risk of heavy vehicle involved crash in lane 3 than in lane 2. However, there are fewer observations with higher CRP in lane 3 because of lower percentage of passenger vehicles and more spacing between heavy vehicles due to overtaking heavy vehicles in lane 2. That is, heavy vehicle crash potential is not always proportional to heavy vehicle volume between lanes.

Fig. A5 illustrates contour plots of the hourly frequency for 'x' events in Figs. A2 through A4 over 27 days of WIM data. Fig. A5 generally supports the observations mentioned in Figs. A2 through A4 .

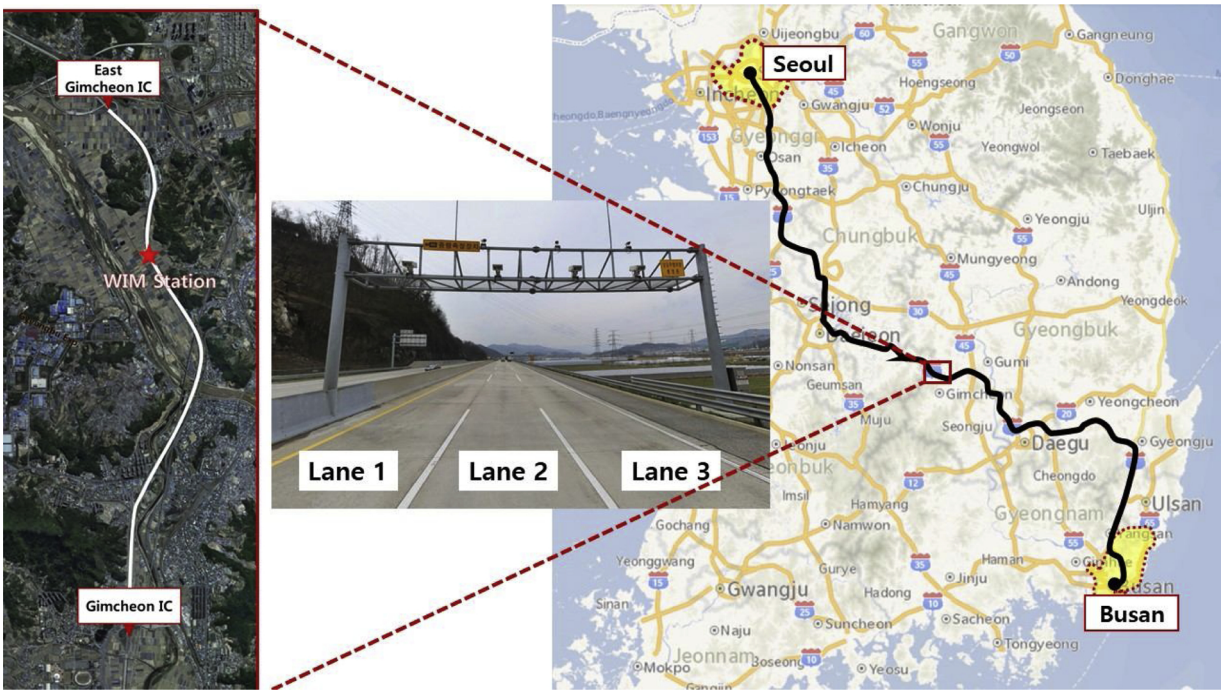

Fig. A1. Specifications and location of the WIM station on Jungbunaeryuk expressway in Korea.

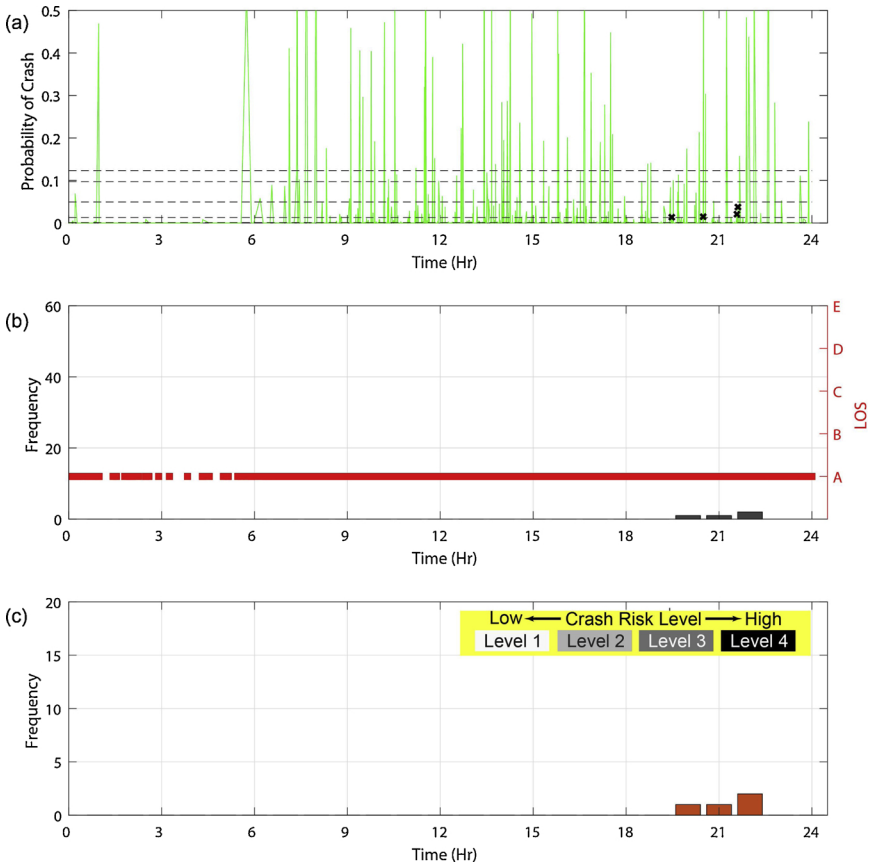

Fig. A2. (a) Probability of a crash higher than MECRP for a given LOS in lane 1, (b) Hourly frequency of the probability of a crash higher than MECRP for a given LOS and flow in lane 1, (c) Crash risk level in lane 1.

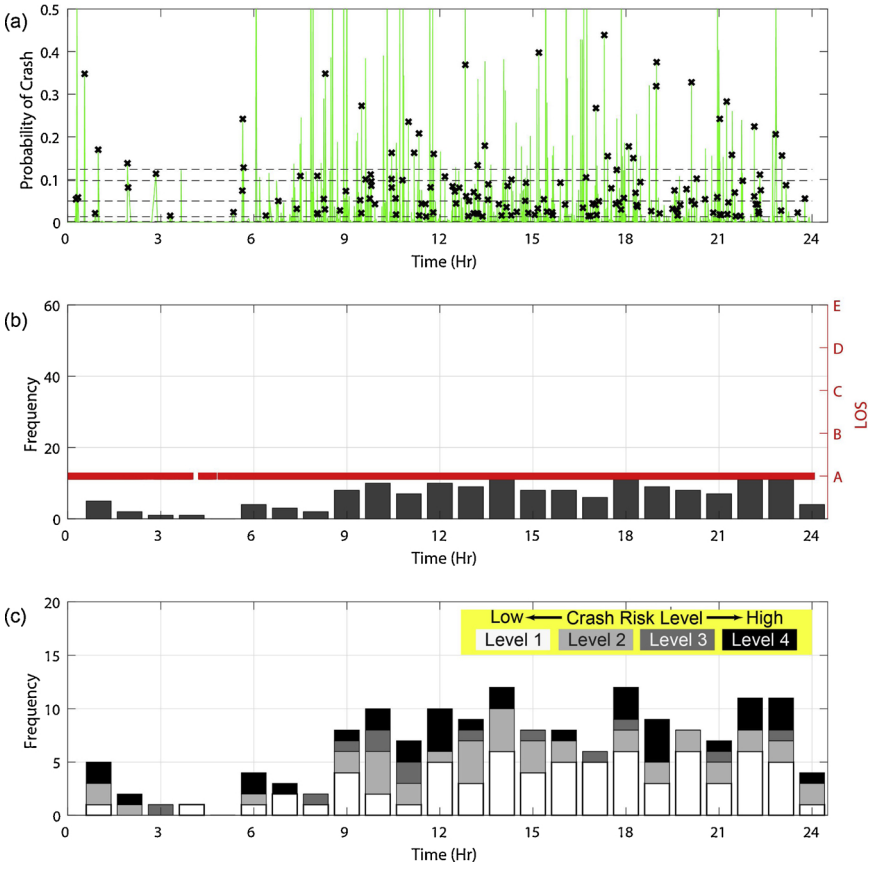

Fig. A3. (a) Probability of a crash higher than MECRP for a given LOS in lane 2, (b) Hourly frequency of the probability of a crash higher than MECRP for a given LOS and flow in lane 2, (c) Crash risk level in lane 2.

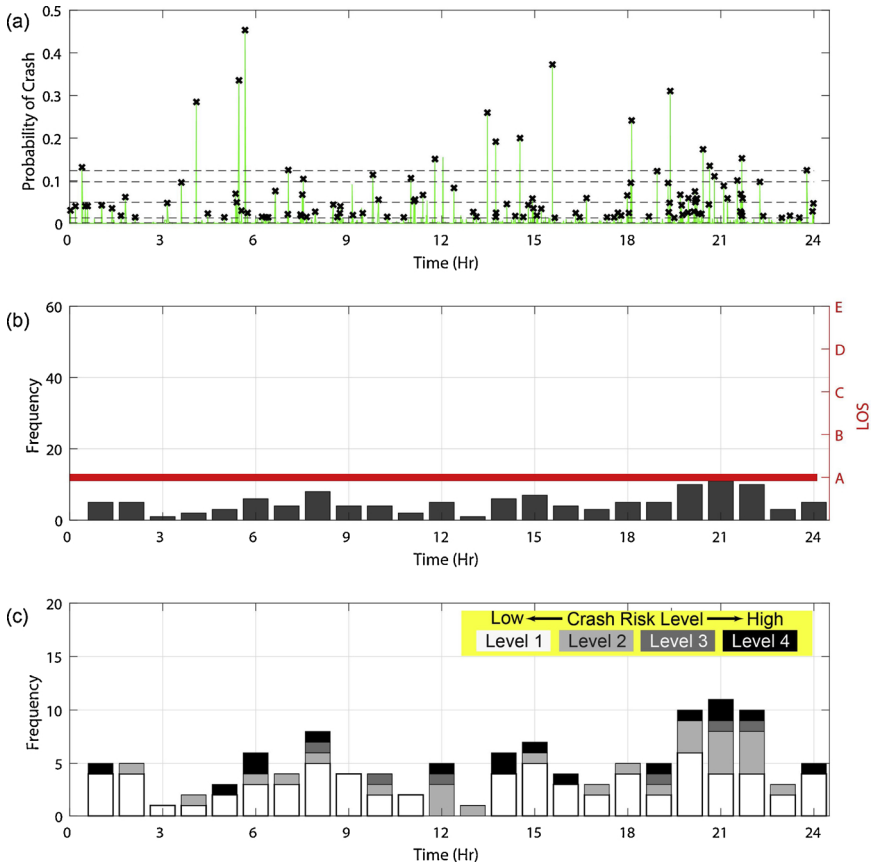

Fig. A4. (a) Probability of a crash higher than MECRP for a given LOS in lane 3, (b) Hourly frequency of the probability of a crash higher than MECRP for a given LOS and flow in lane 3, (c) Crash risk level in lane 3.

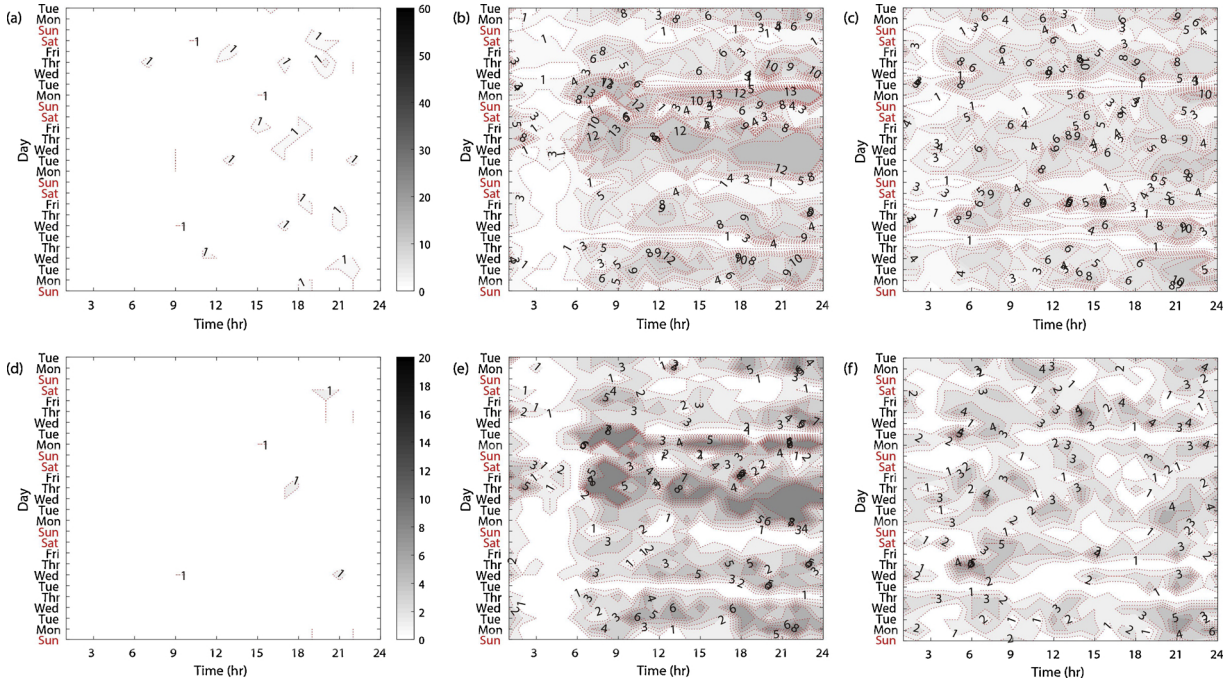

Fig. A5. (a) Contour plot of hourly frequency for all levels over 27 days of WIM data in lane 1, (b) Contour plot of hourly frequency for all levels over 27 days of WIM data in lane 2, (c) Contour plot of hourly frequency for all levels over 27 days of WIM data in lane 3, (d) Contour plot of hourly frequency for high levels (level 3-4) over 27 days of WIM data in lane 1, (e) Contour plot of hourly frequency for high levels (level 3-4) over 27 days of WIM data in lane 2, (f) Contour plot of hourly frequency for high levels (level 3-4) over 27 days of WIM data in lane 3.

References

Abdel-Aty, M., Abdelwahab, H.T., 2003. Configuration analysis of two vehicle rearend crashes. Transp. Res. Rec. 1840, 140–147.

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. Transp. Res. Rec. 1897, 88–95.

Alicandri, E., 1994. The highway driving simulator (HYSIM): the next best thing to being on the road. Public Roads 57, 19–23.

Bachmann, C., Roorda, M.J., Abdulhai, B., 2012. Improved time-to-collision definition for simulating traffic conflicts on truck-only infrastructure. Transp. Res. Rec. 2237, 31–40.

Bella, F., 2009. Can the driving simulators contribute to solving the critical issues in geometric design? Transp. Res. Rec. 2138, 120–126.

Bella, F., Russo, R., 2011. A collision warning system for rear-end collision: a driving simulator study. Procedia-social behav. Sci. 20, 676–686.

Broughton, K.L.M., Switzer, F., Scott, D., 2007. Car following decisions under three visibility conditions and two speeds tested with a driving simulator. Accid. Anal. Prev. 39, 106–116.

Cunto, F., Saccomanno, F.F., 2008. Calibration and validation of simulated vehicle safety performance at signalized intersections. Accid. Anal. Prev. 40, 1171–1179.

Dimitriou, L., Stylianou, K., Abdel-Aty, M.A., 2018. Assessing rear-end crash potential in urban locations based on vehicle-by-vehicle interactions, geometric characteristics and operational conditions. Accid. Anal. Prev. 118, 221–235.

Farah, H., Bekhor, S., Polus, A., 2009. Risk evaluation by modeling of passing behaviour on two-lane rural highways. Proceedings 88th Annual Meeting Transportation Research Board.

Gettman, D., Shelby, S., Siemens, I.T.S., 2008. Surrogate Safety Assessment Model (SSAM), TechBrief (No. FHWA-HRT-08-049). Federal Highway Administration. Office of Research, Development, and Technology, United States.

Gudes, O., Varhol, R., Sun, Q.C., Meuleners, L., 2017. Investigating articulated heavy-vehicle crashes in western Australia using a spatial approach. Accid. Anal. Prev. 106, 243–253.

Hjelkrem, O.A., Ryeng, E.O., 2016. Chosen risk level during car-following in adverse weather conditions. Accid. Anal. Prev. 95, 227–235.

Hossain, M., Muromachi, Y., 2011. Understanding crash mechanisms and selecting interven- tions to mitigate real-time hazards on urban expressways. Transp. Res. Rec. 2213, 53–62.

Hossain, M., Muromachi, Y., 2013. Understanding crash mechanism on urban expressways using high-resolution traffic data. Accid. Anal. Prev. 38, 936–948.

Jenkins, J.M., Rilett, L.R., 2004. Application of distributed traffic simulation for passing behavior study. Transp. Res. Rec. 1899, 11–18.

Kim, J., Wang, Y., Ulfarsson, G.F., 2007. Modeling the probability of freeway rearend crash occurrence. J. Transp. Eng. 133, 11–19.

Malin, F., Norros, I., Innamaa, S., 2019. Accident risk of road and weather conditions on different road types. Accid. Anal. Prev. 122, 181–188.

Moonesinghe, R., Longthorne, A., Shankar, U., Singh, S., Subramanian, R., Tessmer, J., 2003. An Analysis of Fatal Large Truck Crashes. No. HS-809 569. .

Oh, C., Kim, T., 2010. Estimation of rear-end crash potential using vehicle trajectory data. Accid. Anal. Prev. 42, 1888–1893.

Oh, C., Park, S., Ritchie, S.G., 2006. A method for identifying rear-end collision risks using inductive loop detectors. Accid. Anal. Prev. 38, 295–301.

Oh, C., Oh, J., Min, J., 2009. Real-time detection of hazardous traffic events on freeways. Transp. Res. Rec. 2129, 35–44.

Ozbay, K., Yang, H., Bartin, B., Mudigonda, S., 2008. Derivation and validation of new simulation-based surrogate safety measure. Transp. Res. Rec. 2083, 105–113.

Pande, A., Abdel-Aty, M., 2006a. A comprehensive analysis of the relationship between real- time traffic surveillance data and rear-end crashes on freeways. Transp. Res. Rec. 1953, 31–40.

Pande, A., Abdel-Aty, M., 2006b. Assessment of freeway traffic parameters leading to lane- change related collisions. Accid. Anal. Prev. 38, 936–948.

Park, H., Oh, C., Moon, J., Kim, S., 2018. Development of a lane change risk index using vehicle trajectory data. Accid. Anal. Prev. 110, 1–8.

Saccomanno, F.F., Cunto, F., Guido, G., Vitale, A., 2008. Comparing safety at signalized intersections and roundabouts using simulated rear-end conflicts. Transp. Res. Rec. 2078, 90–95.

Saunier, N., Sayed, T., 2008. Probabilistic framework for automated analysis of exposure to road collisions. Transp. Res. Rec. 2083, 96–104.

Wang, X., Abdel-Aty, M., 2006. Temporal and spatial analysis of rear-end crashes at signalized intersections. Accid. Anal. Prev. 38, 1137–1150.

Wu, Y., Abdel-Aty, M., Lee, J., 2018a. Crash risk analysis during fog conditions using real-time traffic data. Accid. Anal. Prev. 114, 4–11.

Wu, Y., Abdel-Aty, M., Cai, Q., Lee, J., Park, J., 2018b. Developing an algorithm to assess the rear-end collision risk under fog conditions using real-time data. Transp. Res. Part C Emerg. Technol. 87, 11–25.

Xu, C., Tarko, A.P., Wang, W., Liu, P., 2013a. Predicting crash likelihood and severity on freeways with real-time loop detector data. Accid. Anal. Prev. 67, 30–39.

Xu, C., Wang, W., Liu, P., 2013b. Identifying crash-prone traffic conditions under different weather on freeways. J. Safety Res. 46, 135–144.

Zhao, P., Lee, C., 2018. Assessing rear-end collision risk of cars and heavy vehicles on freeways using a surrogate safety measure. Accid. Anal. Prev. 113, 149–158.