A COMPARITIVE STUDY ON CRASH CHARACTERISTICS BETWEEN URBAN AND INTERCITY EXPRESSWAY BASIC SEGMENTS

Yong WU. Doctor Candidate, Department of Civil Engineering, Nagoya University, Nagoya, 464-8603, Japan. E-mail: yong@genv.nagoya-u.ac.jp

Hideki NAKAMURA. Professor, Department of Civil Engineering, Nagoya University, Nagoya, 464-8603, Japan. E-mail: nakamura@genv.nagoya-u.ac.jp

Miho ASANO. Assistant Professor, Department of Civil Engineering, Nagoya University, Nagoya, 464-8603, Japan. E-mail: asano@genv.nagoya-u.ac.jp

ABSTRACT

This study aims at identifying the differences of crash characteristics at basic segments between urban and intercity expressways. Crash rate (CR) model is firstly developed as a function of traffic density. The causes leading to different CR characteristics by expressway type are then examined considering the interaction of geometry, traffic flow and ambient conditions. Affecting mechanisms of those factors are finally analyzed through using principle component analysis. The results reveal that, urban expressway has significantly higher CR in low-density uncongested flow due to its compact design relative to intercity expressway. As traffic flow increases, the significance of traffic-related variables is increasing as opposed to the decreasing significance of geometry affecting crashes. Meanwhile, intercity expressway then starts to have higher CR compared to urban expressway, which is possibly related to its vehicle composition characterized by more heavy vehicles (HV). Since HV can interrupt their surrounding traffic, more HV may result in higher frequency of “disruptive” traffic.

Keywords: crash characteristics, crash rate (CR), basic segments, principle component analysis (PCA), urban expressway, intercity expressway

INTRODUCTION

Understanding crash characteristics and their influencing factors is critical to safer geometric design and traffic control strategy. Several studies have asserted that crashes are associated with the interaction among geometry, traffic flow and ambient conditions (Rengarasu *et al.,* 2009; Bajwa *et al.,* 2010). However, previous analyses primarily examine those factors in separate models. Meanwhile, the affecting mechanisms of traffic flow to crashes may differ with the change of traffic conditions (Wu *et al.,* 2012), while existing studies have paid little attention to this regard. Besides, by expressway type, the characteristics of geometric and traffic flow are virtually different. Correspondingly, crash characteristics may be different as well. Nevertheless, most previous studies are focused on a single type of expressway. As a result, the nature of crash influenced by the variation in geometry and traffic characteristics cannot be comprehensively identified.

Comparing to intercity expressway, urban expressway is often constructed in tight geometric features due to urban limited space, such as higher access density, smaller curve radius and narrower lane width. Furthermore, traffic characteristics, *e.g.,* vehicle composition and driver population are actually different by expressway type. Necessarily, crash characteristics and their related influencing factors may be also different. It is necessary to make a distinction between urban and intercity expressways with the purpose to provide more appropriate measures for safer geometric design and traffic control strategy by expressway type.

Therefore, this paper aims at identifying the differences of crash characteristics between urban and intercity expressway basic segments through modelling crash rate (CR). Crash influencing factors are further analyzed considering the interaction among geometry, traffic flow and ambient conditions. Through adopting principle component analysis (PCA), the significances of those factors affecting crashes are investigated. Meanwhile, their influencing mechanisms are compared between urban and intercity expressways.

LITERATURE REVIEW

Traffic safety research includes an extensive array of research areas and the most prominent of them is crash data analysis (Abdel-Aty and Pande, 2007).The conventional approach has established statistical links between CR and individual explanatory factors (Kopelias *et al.,* 2007; Rengarasu *et al.,* 2009). In those studies, traffic conditions are generally represented by low-resolution traffic data, such as hourly and daily flows. Geometric features are primarily reflected in terms of the hierarchy of radius or slope (Fu *et al.,* 2010; Shively *et al.,* 2011). Besides, most existing models are separately developed based on a single factor. Even so, few studies have made a clear distinction between different types of expressways.

Given the insufficiency of aggregated statistics in reflecting the nature of individual crashes, some studies have tried to analyze crash characteristics at individual level, in an effort to predict crash likelihood on a real-time basis (Abdel-Aty and Pemmanabonia, 2006; Zheng *et al.,* 2010; Christoforou *et al.,* 2011). This new concept of real-time crash prediction exhibits huge promise for the application of dynamic traffic control strategy for safety.

To date, those existing crash prediction models are not perfect in view of their predictive performances (Hossain and Muromachi, 2012). Regardless of the limitation of data collection, inadequate modelling process is another potential issue to undermine the validity of models. Most existing models are developed without classifying traffic conditions, or based on the whole routes and placed little concern to facility type-specific crash characteristics. What is more important, for ensuring the reliability of statistics, the significances and independences of explanatory factors should be investigated in advance, while few studies have performed this kind of proactive analysis before crash modelling.

Given the problems of previous studies, this study designs a methodology to investigate the differences of crash characteristics at basic segments and their causes by expressway type. CR models as a function of traffic density in 5 minutes are developed at first. Due to the insufficient crash samples, CR analysis is incapable to examine the combined effects of geometry, traffic flow and ambient conditions on crashes by a single model. Thereby, PCA is employed instead to qualitatively identify the affecting mechanisms of those factors, which are further compared between urban and intercity expressways.

STUDY SITES AND DATASETS

As shown in Figure 1, Nagoya Urban Expressway network (NEX) and the section of Tomei-Meishin Expressway from Mikkabi Interchange (IC) to Yokaichi Interchange (IC) are involved in this study. Up to December 31, 2009, NEX was about 69.2km in total length with over 250 ultrasonic detectors installed in approximately 500m. Most roads are 4-lane roadway (2-lane/dir) except for Inner ring (route No.R) which is one-way and where the number of lane differs (2~5) with the change of ramp-junctions. Tomei-Meishin Expressway is one of the main arteries of communications in Japan. The selected section is 183.6km long including nearly 180 loop detectors (for 2-direction) in a spacing of 2km. Bedsides, 4-lane roadway is designed along the mainline excluding some areas where an auxiliary lane is located.

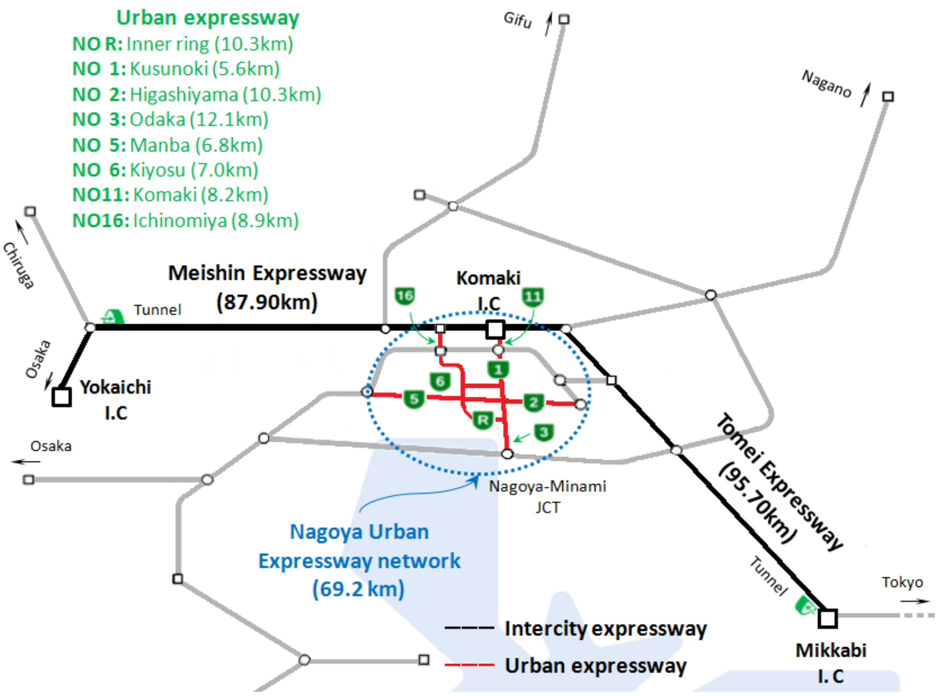

Figure 1 – Schematic map of study sites

Basic segment is selected for the following analysis. It is defined as the segments that are outside the influence areas of merging, diverging and weaving maneuvers. In this study, basic segment in NEX is extracted outside the 500m up- and downstream of ramp-junctions. Meanwhile, a special geometric design, tight curve with radius lower than 100m, is excluded in advance considering its higher CR relative to the other segments (Wu *et al.* 2012). For Tomei-Meishin Expressway, a midpoint between two neighbouring detectors, one of which is nearest to interchange/junction, is regarded as the boundary of basic segment separating from other facility types. Since the 3-lane/dir segments are limited, only 2-lane/dir segments are analyzed in this study. By this means, a total length of 56.63km basic segments in NEX and 154.90km on Tomei-Meishin Expressway can be successfully extracted.

Figure 2 – Standard cross section of NEX
Layout: 1.5m(shoulder)+3.25m(lane)×2+0.5m(curbside)

Figure 3 – Standard cross section of Tomei Expressway
Layout: 3.0m(shoulder)+3.6m(lane)×2+0.75m(curbside)

The standard cross-section layouts of both types of expressways are shown in Figure 2 and Figure 3, respectively. Urban expressway is often constructed as viaducts and generally designed with smaller radius and higher access density in contrast to intercity expressway. Referring to the above two figures, smaller lane width coupled with higher roadside barrier can be observed for NEX. Comparatively, the vertical alignment along NEX is designed more gently owing to the structure of viaducts: its maximal gradient of 2.8% relative to the value of 5.0% for Tomei-Meishin Expressway. With respect to traffic characteristics, Tomei-Meishin Expressway is a main transportation link between Tokyo and Osaka, two major metropolis of Japan. Hence, it is assumed that Tomei-Meishin Expressway generally carries large number of heavy vehicles (HV) and long-distance trips (LDT). In contrast, NEX bears more intra-city transportation that is composed of high population of commuters. In addition, the visibility on NEX is restricted more seriously in comparison with Tomei-Meishin Expressway due to its compact geometric design and higher roadside barrier. Those geometry- and traffic-related characteristics for two types of expressways are summarized in Table 1.

Table 1 – Summary statistics of geometry- and traffic-related characteristics

Index		Urban expressway		Intercity expressway
Traffic -related[a]	Vehicle composition	-		More heavy vehicles More long-distance trips
	Driver familiarity	More commuters		-
Geometry -related	Design speed[b]	60km/h	80km/h	80-120km/h
	Standard cross section	1.5m+3.25m×2	1.5m+3.5m×2	3.0m+3.5m×2
	Alignment feature[c]	R_{min}=200m i_{max}=±2.8%	R_{min}=1000m i_{max}=±2.0%	R_{min}=410m i_{max}=±5.0%
	Roadside barrier	Higher		-

[a] Assumed differences in this study and require related data to be proved in future work.
[b] 80km/h for Komaki (route No.11) and Ichinomiya (route No.16) routes only.
[c] R_{min} means the minimal value of radius; i_{max} corresponds to the maximal value of slope.

Five databases are prepared in this study; 1) crash records with the occurrence time in minute and the location in 0.1km, 2) detector data including flow rate, speed and occupancy per 5 minutes, 3) geometric design and detector allocation in a unit of 0.01km, 4) the locations and periods of temporal lane/cross-section closures, and 5) daily sunrise/sunset time records in Nagoya. The period of the data above is over three years (2007-2009) except for those on Kiyosu route (route No.6 in NEX) that opened from December 1, 2007.

METHODOLOGY

Data processing

Detector data

In principle, detectors can count the number of vehicles at their locations only. In such case, the “coverage area” of two neighbouring detectors should be separated for matching traffic conditions with individual crashes by detector data. The boundary of consecutive coverage areas is defined at the midpoint between two neighbouring detectors. Note that, the time of crash is not exact occurrence time, since it is recorded by road administrators after crash occurrence. For this reason, detector data within small time before crash should be rejected to avoid mixing up crash-influencing and crash-influenced data. Regarding this, the latest data at least 5 minutes before the recorder time are accepted in this study. The invalid data and the data within lane/cross-section closure intervals are excluded from the dataset.

The detector data on Tomei-Meishin Expressway is lane-based, while it is cross section-based in NEX. Considering the reliability of comparison by expressway type, this study converts the lane-based data into cross section-based data by the following equations.

$$q_S = \sum q_i \tag{1}$$

$$v_S = \frac{\sum q_i \times v_i}{\sum q_i} \tag{2}$$

$$k_S = \frac{12 \times q_S}{v_S} \tag{3}$$

Where, q_i and v_i are flow rate and speed on lane # i, respectively; q_s, v_s and k_s denote the converted flow rate, speed and traffic density for the whole cross-section, respectively.

Geometric features

Design consistency is the conformance of geometry of a highway with driver expectancy, and its importance and significant contribution to road safety is justified by understanding the driver-vehicle-roadway interaction (Ng and Sayed, 2004). Hikosaka and Nakamura (2001) proposed using geometric variation in the upstream distance from crash locations to reflect the safety benefit of design consistency measured by focusing on locations. In view of the length of coverage area, the variation in 500m long distance is calculated.

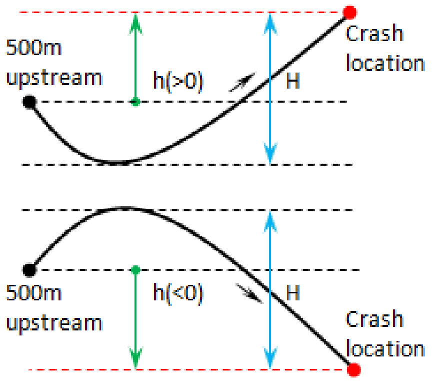

Figure 4 – Calculation of variation in road elevation

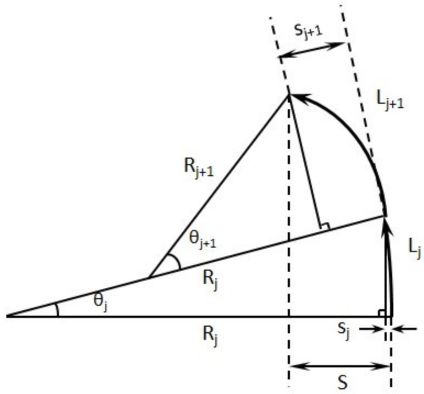

Figure 5 – Calculation of horizontal displacement

1. Variation in road elevation h between one crash location and its 500m upstream point, and the maximal elevation difference H during this distance (see Figure 4).
2. Horizontal displacement S. Radius is impossible to describe a section composed of various curves. In essence, centrifugal force is also associated with the horizontal displacement S in the direction of tangent to curve (see Figure 5). As an alternative index, S in the 500m section is adopted and calculated by the following equations.

$$\theta_j = \frac{L_j}{R_j} \qquad (0 < \theta_j < \frac{\pi}{2}) \tag{4}$$

$$s_j = R_j(1 - \cos\theta_j) \tag{5}$$

$$S = \sum s_j \tag{6}$$

Where, j is the ID of curve; R_j, θ_j, L_j and s_j correspond to the radius, centre angle, arc length and horizontal displacement of curve # j, respectively.

3. Index of centrifugal force I_{CF}. Speed v always has a square relation with centrifugal force. This study designs I_{CF} to reflect the combined effect of speed v along with horizontal displacement S, while it is not the correct value of centrifugal force.

$$I_{CF} = S \times v^2 \tag{7}$$

4. Index of space displacement I_{SD}. I_{SD} is employed to describe the comprehensive geometric features induced by the horizontal and vertical alignment variations.

$$I_{SD} = S \times H \tag{8}$$

Those data above are collected every 0.1km since crashes are recorded in a unit of 0.1km. In the meantime, they are also extracted at each detector location that is the common link between crash and detector data. Table 2 summarizes the process of those data collection.

Table 2 – Example of geometric variation collection

Route #	Direction	Location (km)	h (m)	H (m)	S (m)	I_{SD} (m^2)	Note
1	Southbound	0.0	-4.63	5.49	0.78	4.30	
1	…	0.1	-7.90	8.49	3.91	33.2	
1	…	0.2	-10.6	11.5	6.08	69.9	
1	…	0.21	-11.5	11.8	8.88	104.7	For detector #0101
1	…	0.3	-15.3	15.3	9.60	146.9	

Ambient conditions

Commonly prevailing and uncontrolled environment and weather for pre-crash conditions are defined as ambient conditions. They are 1) ambient light classified into daytime/nighttime that is the time period from sunrise to sunset and from sunset to sunrise, respectively, 2) sunny/cloudy/rainy weather conditions at the time of crash, 3) dry/wet pavement conditions at crash locations, and 4) day type on crash days that is composed of holiday/weekday. Here, holiday includes all weekends, all national and traditional holidays like Golden Week in May and the Obon Week in August in Japan. Correspondingly, other days belong to weekday.

Data matching

The related detector data, geometric variations and ambient conditions for individual crashes can be matched as demonstrated in Table 3. The crashes matched with invalid detector data and within lane/cross-section closure intervals are excluded as well. As a result, a total of 457 and 1496 crashes remain in NEX and on Tomei-Meishin Expressway, respectively.

Table 3 – Example of data matching for individual crashes

Crash ID	Detector data			Geometric features				Ambient conditions			
	q_s (veh/5min)	v_s (km/h)	k_s (veh/km)	h (m)	H (m)	I_{CF} (km^3/h^2)	I_{SD} (m^2)	Light	Weather	Pavement	Day type
1	58	86.4	8	4.50	4.85	57.4	37	Night	Sunny	Dry	Holiday
2	267	38.6	83	1.55	3.69	4.93	12	Day	Rainy	Dry	Weekday
3	60	77.0	9	4.44	4.44	61.9	46	Night	Sunny	Wet	Holiday
4	2	50.4	1	1.64	2.14	87.3	74	Night	Cloudy	Dry	Weekday

Classification of traffic conditions

Congested flow, as a typical oscillated traffic, has different characteristics from uncongested flow. It is necessary to make a distinction between the two traffic conditions. For this purpose, the diagrams of traffic flow-speed are analyzed at two bottlenecks: Figure 6 for Horita on-ramp junction of NEX and Figure 7 for Toyota junction (JCT) of Tomei-Meishin Expressway. Corresponding to the maximum flow rate, the boundaries of speed between the two traffic conditions can be found as 60km/h and 70km/h, defined as critical speed (Kobayashi *et al.,* 2011), for urban and intercity expressways, respectively. Those values can be generally accepted as the related index for NEX and Tomei-Meishin Expressway, since the critical speeds at other bottlenecks are observed to be around 60km/h (Wu *et al.,* 2012) and 70km/h (Kobayashi *et al.,* 2011) for the two types of expressways, respectively.

Towards reflecting the variation in traffic characteristics, traffic conditions are further sub-classified. In Figure 6, it is evident that speed has a high variance at low flow rates. Besides, occupancy isn't a commonly used variable. In this case, the estimated traffic density k_S calculated by Equation (3) is proposed to be the measure of effectiveness for further differentiating traffic conditions. Considering the number of crash samples available, the aggregation intervals of k_S are finally set to be 5veh/km and 30veh/km for un- and congested flows, respectively.

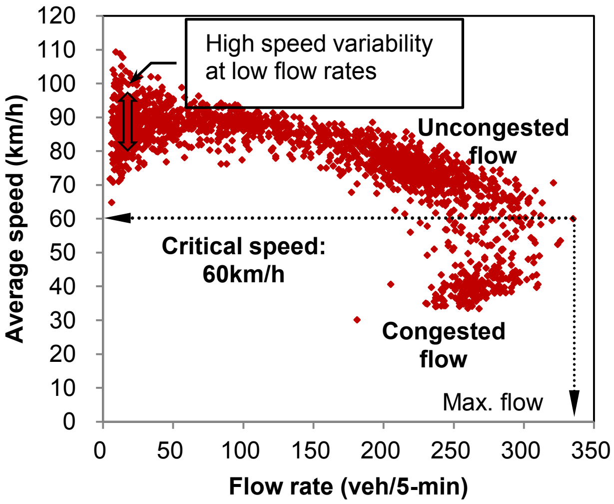

Figure 6 – Traffic flow-speed diagram at Horita on-ramp

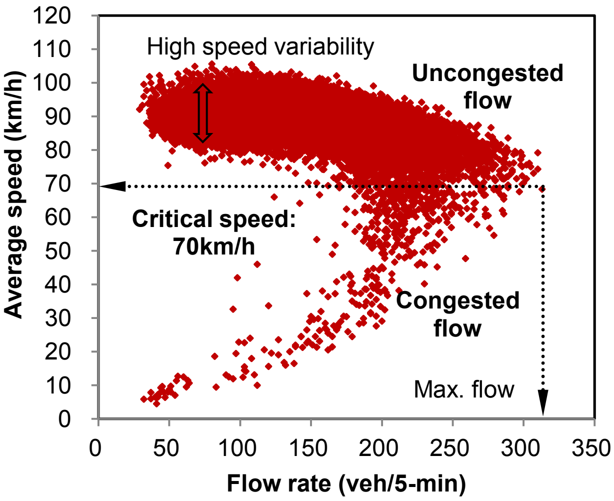

Figure 7 – Traffic flow-speed diagram at Toyota JCT

Calculation of crash rates

Crash rate (CR) for traffic condition # m can be calculated by the following equation.

$$CR_m = \frac{NOC_m \times 10^6}{\sum Q_{mn} L_n} \tag{9}$$

Where, m and n are the ID of traffic condition and detector coverage area, respectively; NOC_m is the number of crashes for traffic condition # m; $Q_{mn}L_n$ is the value of vehicle km of travelled (VKMT) in detector coverage area # n for traffic flow condition # m.

Principle component analysis

Principle component analysis (PCA) is a powerful tool for reducing a number of observed variables into a small number of artificial variables that account for most of the variance in the dataset (Sanguansat, 2012). Essentially, through orthogonal transformation, a set of observations of possibly correlated variables can be converted into a set of values of linearly uncorrelated variables. Those converted values are defined to be principle components. Technically, a principle component can be regarded as a linear combination of optimally-weighted observed variables (Sanguansat, 2012). This transformation is demonstrated in the following way: the first principal component has the largest possible variance, and accounts for as much of the variability in the data as possible; each succeeding component in turn has the highest variance possible and accounts for as much of the remaining variability as possible. Then, two criteria are generally available towards selecting how many components should be extracted; 1) 80% rule, the extracted components should be capable to explain at least 80% of the variance in the original dataset, 2) Eigen value rule, which means only components whose Eigen values are greater than 1.0 can be chosen.

DIFFERENCES OF CRASH RATE MODELS

Uncongested flow

Figure 8 gives CR tendencies following traffic density k_s by expressway type in uncongested flow. Generally, CR is convex downward to k_s, and the quadratic function seems to be sound for modelling these tendencies, as summarized in Table 4. By expressway type, it is obvious that, urban expressway has higher CR relative to intercity expressway in low-density conditions. As traffic flow increases, the CR on intercity expressway increases rapidly and gets much higher than that on urban expressway. For verifying these differences, a paired t-test of CR is exerted between two types of expressways (see Table 5). The results further demonstrate a significant difference of CR by expressway type separately for low- and high-density conditions even the CR is not significantly different for the whole uncongested flow.

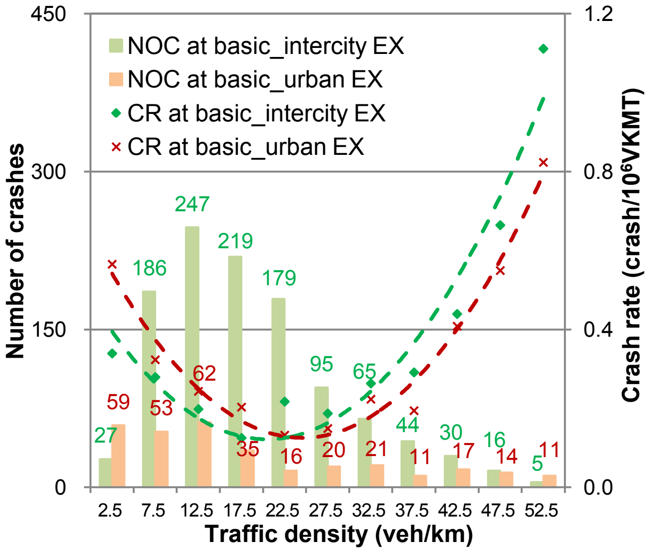

Figure 8 – CR regression models in uncongested flow

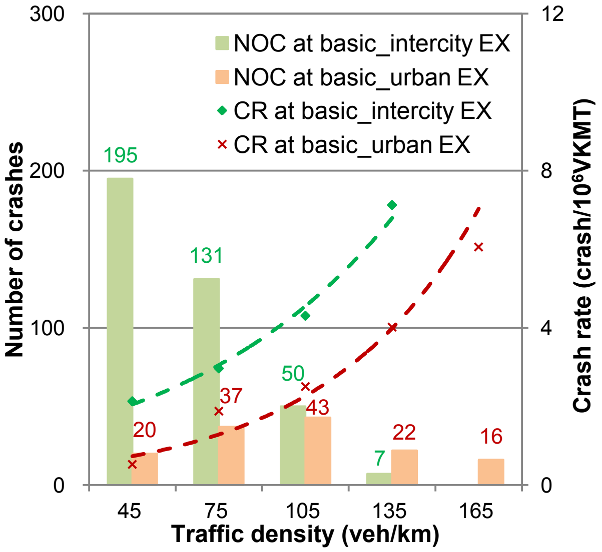

Figure 9 – CR regression models in congested flow

Table 4 – Summary of CR regression models

Traffic conditions	Segments [a]	Number of crashes	Models [b]
Uncongested flow	Basic_urban	319	$CR=8.55\times10^{-4}k_S^2-4.20\times10^{-2}k_S+6.41\times10^{-1}$ $R^2=0.971$ $k_S(CR_{min})=23$
	Basic_intercity	1113	$CR=8.43\times10^{-4}k_S^2-3.46\times10^{-2}k_S+4.76\times10^{-1}$ $R^2=0.941$ $k_S(CR_{min})=22$
Congested flow	Basic_urban	138	$CR=3.12\times10^{-1}e^{0.0189ks}$ $R^2=0.918$
	Basic_intercity	383	$CR=1.13e^{0.0133ks}$ $R^2=0.990$

[a] Basic_urban/Basic_intercity: urban/intercity expressway basic segments (same to following tables).
[b] k_S (CR_{min}): the value of traffic density corresponding to the estimated minimum CR.

Table 5 – t-test of CR between urban and intercity expressway basic segments

Traffic conditions [a]	t-value	df	Sig.	Note
Uncongested flow	-0.697	10	0.502	Low- and high-density congested flow are bounded at traffic density of 20veh/km in terms of $k_S(CR_{min})$
Low-density uncongested flow	2.841	4	0.047	
High-density uncongested flow	-2.713	5	0.035	
Congested flow	-4.426	4	0.021	

Congested flow

Figure 9 describes the differences of CR distribution to traffic density k_s by expressway type in congested flow. It appears that CR follows an increasing tendency to k_S on both types of expressways. Accordingly, this study applies the exponential function for modelling these CR tendencies (see Table 4 as well). Compared to urban expressway, it is clear that intercity expressway has much higher CR in congested flow. A paired t-test of CR between two types of expressways can also confirm this finding (see Table 5).

DIFFERENCES OF CRASH INFLUENCING FACTORS

The findings above reveal that CR characteristics are significantly different by expressway type, which may be related to the varied geometric and traffic characteristics between urban and intercity expressways. Next part thereby investigates the effects of individual factors on crashes in an effort to identify those different natures of crashes by expressway type. Given the insufficient crash samples, CR analysis is inappropriate to examine a variety of factors by a single model. Instead, principle component analysis (PCA) is employed and the affecting mechanisms of various factors are further analyzed.

Introduction of variables

Table 6 introduces individual variables combining with its type and some summary statistics. Theoretically, traffic flow diagram is two-dimensional, and thus, k_S and v_S are used together to describe traffic conditions. Towards reflecting geometric features, h, I_{CF} and I_{SD} are picked out to reflect the vertical, horizontal and the comprehensive alignment variations, respectively. Dummy variable is referred to incorporate ambient conditions into PCA. Since a dummy variable usually takes two values of 1 and 0, weather are replaced by pavement conditions.

Table 6 – Introduction of explanatory variables

Variables	Basic_urban[a]		Basic_intercity		Description
	Max	Min	Max	Min	
k_S	238	1	190	1	Traffic density (veh/km)
v_S	128.3	4.70	106.5	4.86	Average speed (km/h)
I_{CF}	1997.4	0	523.3	1	Index of centrifugal force (km^3/h^2)
h	12.1	0.04	24.2	0	Variation in road elevation (m)
I_{SD}	1113	0	1045	0	Index of horizontal displacement (m^2)
Pave	F(1)=24.5%		F(1)=19.6%		=1 if wet pavement, 0 otherwise
Light	F(1)=29.1%		F(1)=41.0%		=1 if nighttime, 0 otherwise
Day	F(1)=26.9%		F(1)=37.6%		=1 if holiday, 0 otherwise

[a] Max/min: the maximum/minimum values in statistics; F: frequency.

Analysis results of PCA

PCA essentially rotates data through using a linear transformation. Consequently, only the monotonic loadings of factors can be reflected reliably. In such case, uncongested flow is further classified into low- and high-density conditions at approximately 20veh/km in view of the value of $k_S(CR_{min})$ as shown in Figure 10, since different monotonicities of CR exist in the two conditions. As a result, the following three traffic conditions are analyzed, *i.e.,* low- and high-density uncongested flow as well as congested flow.

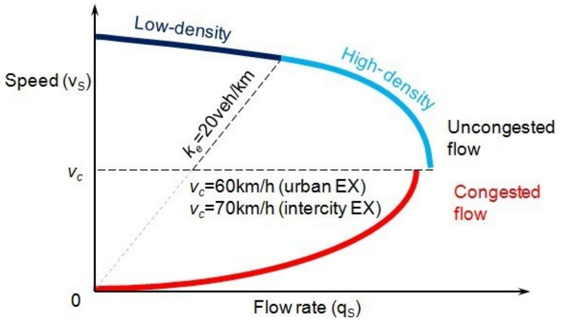

Figure 10 – Classification of traffic conditions

Low-density uncongested flow

Table 7 provides the PCA results in low-density uncongested flow for urban expressway basic segments. Based on the criteria aforementioned, four components are extracted in terms of their corresponding Eigen values. Furthermore, these selected components can explain at least 80% of the variance in the original datasets.

Crash occurrence is found to be significantly associated with geometric variation (I_{CF} and I_{SD}), traffic density k_S along with nighttime, speed v_S coupled with wet pavement and vertical variation h. Geometric variation is the 1st component, as greater variation may result in more frequent speed reduction. Accordingly, the difficulty for drivers to control vehicle behaviours increases. Low k_S can reduce drivers' attention, and tempt them take discretionary driving. Such condition combining with poor ambient light is possible to increase crash risk. Due to the reduced tire-pavement friction, wet pavement can negatively affect roadability, especially for high-speed running traffic. In addition, the vertical variation h has a positive loading, as a result of visibility restriction and the difficulty in safe driving with the increase of h. Note that, k_S and v_S are discovered to belong to different principal components, since both variables are not highly interrelated in this discretionary driving condition.

Table 7 – PCA results at urban expressway basic segments in low-density uncongested flow (225 crash samples)

Variables	Components			
	1st	2nd	3rd	4th
Traffic density (k_S)	-0.194	-0.852	-0.119	0.103
Average speed (v_S)	0.285	0.182	0.798	-0.086
Index of centrifugal force (I_{CF})	0.953	0.005	0.053	-0.122
Index of horizontal displacement (I_{SD})	0.959	0.011	-0.044	0.090
Variation in road elevation (h)	0.119	0.149	0.228	0.973
Pavement (Pave)	0.294	0.214	0.783	-0.095
Day type (Day)	0.139	0.093	0.190	-0.468
Ambient light (Light)	-0.164	0.838	-0.130	0.143
Initial Eigen values	2.12	1.54	1.37	1.12
% of Variance	30.3	20.2	18.2	15.0
Cumulative %	30.3	50.5	68.7	83.7

Table 8 – Principle components in low-density uncongested flow

Segments (Number of crashes)	Item	Principle component							
		1st		2nd		3rd		4th	
		Factor	Loading	Factor	Loading	Factor	Loading	Factor	Loading
Basic_urban (225)	Component	I_{CF} I_{SD}	0.953 0.959	k_S Light	-0.852 0.838	v_S Pave	0.798 0.783	h	0.973
	% of variance	30.3		20.2		18.2		15.0	
Basic_intercity (765)	Component	I_{CF} I_{SD} h	0.832 0.958 0.861	k_S Light	-0.881 0.730	v_S Pave	0.844 0.942	Day	0.979
	% of variance	26.3		20.6		18.6		14.6	

The principle components on intercity expressway are summarized in Table 8. The variables that are significantly related to each principle component are selected according to their loadings. For judging the relative significance of the same component by expressway type, the % of variance explained by individual components is provided as well.

Compared to urban expressway, h becomes one variable related to the 1st component on intercity expressway, while its geometric features are less significant in terms of the % of variance. Besides, Day is a significant influencing factor, which is not on urban expressway. As stated before, the vertical gradient on Tomei-Meishin Expressway is much higher relative to NEX. Hence, the comprehensive alignment variation is fairly related with h. However, the compact design of NEX can lead to greater loss of visibility and higher frequency of speed reduction. As a consequence, urban expressway has significant higher CR in low-density uncongested flow. In contrast to normal weekdays, holidays have more LDT, more travels in unfamiliar conditions and more drinking driving (Anowar *et al.,* 2013). These differences by day type on intercity expressway may be more significant in contrast with urban expressway.

High-density uncongested flow

As traffic flow increases, the inter-vehicle interaction gets more intensive. The corresponding PCA results in high-density uncongested flow are summarized in Table 9. Compared to low-density uncongested flow, it is clear that the traffic-related variables including k_S and v_S belong to the same component, as a reflection of increased inter-vehicle interaction. Crashes are found to be more probable to occur as k_S increases. Such findings can accord with the developed CR models before: CR is decreasing to k_S in low-density conditions as opposed to increasing in high-density conditions. In terms of the related % of variance, the significance of geometry is virtually decreasing with the increase of traffic density.

On intercity expressway, h is still a significant index of geometric features. Meanwhile, in terms of the % of variance, its significance of traffic flow affecting crashes gets up in contrast to urban expressway. Another different influencing factor by expressway type is pavement condition that is significant on intercity expressway. These above differences may account for higher CR on intercity expressway in high-density congested flow. Relative to urban expressway, more population of HV and LDT exist on intercity expressway. Since the inter-vehicle interaction gets more intensive, higher percentage of HV (HV%) may induce more frequent interruption to their surrounding traffic, and result in more serious fluctuation of traffic flow (*e.g.,* forced lane-changing behaviours). As suggested in Oh *et al.* (2005), crash occurrence is significantly affected by "disruptive" traffic flow as opposed to a "normal" traffic flow. More LDT can further aggravate such interruption since their higher expected speed would conflict with the slow speed of HV in downstream. With respect to wet pavement, it can negatively affect roadability, which is critical to safe lane-changing behaviours.

Table 9 – Principle components in high-density uncongested flow

Segments (Number of crashes)	Item	Principle component							
		1st		2nd		3rd		4th	
		Factor	Loading	Factor	Loading	Factor	Loading	Factor	Loading
Basic_urban (94)	Component	I_{CF} I_{SD}	0.983 0.974	k_S v_S	0.858 -0.885	Day Light	0.776 0.781	h	0.925
	% of variance	28.9		21.8		17.1		14.4	
Basic_intercity (445)	Component	I_{CF} I_{SD} h	0.797 0.957 0.649	k_S v_S	0.885 -0.891	Day Light	0.894 -0.973	Pave	0.956
	% of variance	25.2		24.0		18.4		14.2	

Congested flow

With the further increase of traffic density, congested flow appears. Table 10 summarizes its PCA results by expressway type. Crashes are still prone to higher k_S, and it confirms the increasing tendency of CR to k_S in this condition. Meanwhile, the traffic-related variables (k_S and v_S) get most important. By contrast, the significance of geometry is further decreased compared to uncongested flow. Besides, ambient light is no longer significant, which is possibly due to the insufficient crash samples collected during nighttime in congested flow.

Table 10 – Principle components in congested flow

Segments (Number of crashes)	Item	Principle component							
		1st		2nd		3rd		4th	
		Factor	Loading	Factor	Loading	Factor	Loading	Factor	Loading
Basic_urban (138)	Component	k_S v_S	0.950 -0.947	I_{CF} I_{SD}	0.842 0.743	h Pave	0.699 0.814	Day	0.942
	% of variance	26.0		20.3		18.9		15.4	
Basic_intercity (383)	Component	k_S v_S	0.923 -0.792	I_{CF} I_{SD} h	0.735 0.886 0.797	Day	0.809	Pave	0.883
	% of variance	26.3		22.2		17.0		16.5	

By expressway type, h is still a significant geometric index for intercity expressway. Besides, the geometric features on intercity expressway are found out to play a more important role to crash occurrence relative to urban expressway. As we know, higher HV% and steeper slope are the distinct traffic and geometric features of intercity expressway different from urban expressway. Due to poor dynamics, heavy vehicles may throw more serious interruption to other traffic while driving on steeper slope.

Given the above analyses, geometry is found to be the most significant influencing factor in uncongested flow, especially for low-density conditions. In this sense, geometric features are a major cause leading to the different crash characteristics by expressway type, *e.g.,* the compact design relating to higher CR on urban expressway. As traffic flow increases, the traffic-related variables get more important as opposed to the decreasing significance of geometry. In view of the interruption of HV to other traffic, the greater frequency of disruptive traffic may emerge on intercity expressway since its vehicle composition is characterized by higher HV%. As a result, the CR on intercity expressway increases rapidly and gets higher relative to urban expressway. In congested flow, the design of larger slope may intensify the above interruption. Hence, the tendency of higher CR on intercity expressway gets clearer.

CONCLUSION AND FUTURE WORK

This paper identified the differences of crash characteristics through CR model at basic segments by expressway type. In low-density uncongested flow, urban expressway has significantly higher CR compared to intercity expressway. As traffic flow increases, the CR on intercity expressway increases more rapidly and becomes much higher in contrast to urban expressway. The causes leading to those differences were further identified. The compact design is related to higher CR on urban expressway in low-density uncongested flow. When traffic density increases, HV can interrupt the surrounding traffic, and more HV may result in higher frequency of “disruptive” traffic flow. As a result, the CR on intercity expressway then gets much higher since its vehicle composition is characterized by higher HV%.

The potential benefits of integrating the above results in geometric design and traffic control are numerous. Based on those estimated CR models, road administrators can easily image safety performance with the variation in traffic conditions. Furthermore, PCA results may help prioritize countermeasures, and further estimate the safety performance of an adopted countermeasure. The quality of design consistency needs to be examined in terms of fitness between visibility restriction and speed reduction at high speed. Besides, the design of vertical slope should be verified considering the safety performance in congested flow.

For more accurate investigation of crash characteristics, further studies are required by using high-sample-size data. Since crash occurrence is significantly associated with the short-term turbulence of traffic, such kind of variables (*e.g.,* speed variance) is needed. As suggested by Christoforou *et al.* (2011), conditions preceding crashes are different by type of crash. Future analysis in this regard is necessary. Finally, an analysis to quantify the combined effects of various factors on crashes is highly expected considering the more effective application.

ACKNOWLEDGMENTS

Authors are very grateful to Nagoya Expressway Public Corporation and Central Nippon Expressway Company Limited for their data provision and other supports for this analysis.

REFERENCES

Abdel-Aty, M and Pemmanaboina, R (2006). Calibrating a real-time traffic crash prediction model using archived weather and ITS traffic data. IEEE Transactions on Intelligent Transportation System, 7(2), 167-174.

Abdel-Aty, M and Pande, A (2007). Crash data analysis: Collective vs. Individual crash level approach. Journal of Safety Research, 38(5), 581-587.

Anowar, S., Yasmin, S and Tay, R (2013). Comparison of crashes during public holidays and regular weekends. Accident Analysis & Prevention, 51, 93-97.

Bajwa, S., Warita, M and Kuwahara, M (2010). Effects of road geometry, weather and traffic flow on road safety. Proceedings of the 15th International Conference of Hong Kong Society for Transportation Study, 773-780.

Christoforou, Z., Cohen, S and Karlaftis, M. G (2011). Identifying crash type's propensity using real-time traffic data on freeways. Journal of Safety Research, 42, 43-50.

Fu, R., Guo, Y. S., Yuan, W., Feng, H. Y and Ma, Y (2011). The correlation between gradient of descending roads and accident rates. Safety Science, 49, 416-423.

Hikosaka, T and Nakamura, H (2001). Statistical analysis on the relationship between traffic accident rate and traffic flow condition in basic expressway sections. Proceedings of the 21st Japan Society of Traffic Engineering Meeting (JSTE), 173-176.

Hossain, M and Muromachi, Y (2012). A bayesian network based on framework for real-time crash prediction on basic freeway segments of urban expressways. Accident Analysis & Prevention, 45, 373-381.

Kopelias, P., Papadimitriou, F., Papandreou, K and Prevedouros, P (2007). Urban freeway crash analysis - geometric, operational and weather effects on crash number and severity. TRB Annual Meeting CD-ROM, Washington D.C.

Kobayashi, M., Nakamura, H., Asano, M and Komekawa, E (2011). Analysis on stochastic characteristics of capacities at intercity expressway bottlenecks. Proceedings of the 31st Japan Society of Traffic Engineering Meeting (JSTE), 133-138.

Ng, J. C. W and Sayed, T (2004). Effect of geometric design consistency on road safety. Canadian Journal of Civil Engineering, 31, 218-227.

Oh, C., Oh, J and Ritchie, S (2005). Real-time hazardous traffic condition warning system. IEEE Transactions on Intelligent Transportation System, 6(3), 8.

Rengarasu, T. M., Hagiwara, T and Hirasawa, M (2009). Effects of road geometry and cross section variables on traffic accidents using homogeneous road segments. Transportation Research Record: Journal of Transportation Research Board, 34-42.

Sanguansat, P (2012). Principle Component Analysis - Engineering Application. Intech.

Shively, T. S., Kockelman, K and Damien, P (2010). A bayesian semi-parametric model to estimate relationships between crash counts and roadway characteristics. Transportation Research Part B: Methodological, 44(5), 699-715.

Wu, Y., Nakamura, H and Asano, M (2012). A comparative study on crash rate characteristics at different intercity expressway facility types. Proceedings of the 46th IP meeting of Japan Society of Civil Engineering, CD-ROM.

Zheng, Z. D., Ahn, S and Monsere, C. M (2010). Impact of traffic oscillation on freeway crash occurrences. Accident Analysis and Prevention, 42, 626-636.