A STUDY ON CRASH RISK AT EXPRESSWAY BASIC SEGMENTS AND ITS INFLUENCING FACTORS

WU, Yong

A STUDY ON CRASH RISK AT EXPRESSWAY BASIC SEGMENTS AND ITS INFLUENCING FACTORS

Doctoral Dissertation

Submitted in Partial Fulfillment of the
Requirements for the Degree of
Doctor of Engineering

by
WU, Yong

Academic Adviser:

Prof. Hideki Nakamura

Department of Civil Engineering
Nagoya University
October, 2013

Acknowledgements

Without the support of several individuals, this dissertation would not have been possible. It is a pleasure to express my gratitude to them all in my humble acknowledgment.

First of all, I am deeply indebted to my supervisor, Prof. Hideki Nakamura for providing me with an opportunity to study in Japan. From the very early stage of my study, his invaluable guidance has helped me to understand the truly scientific research, which makes my PhD experience memorable and productive. His working spirit and insightful guidance have been inspiring and enriching my growth. I am sure they would cause me to profit lifetime. Anyway, I am indebted to him very much.

My sincerest appreciation also goes to Prof. Takayuki Morikawa, Prof. Yukimasa Matsumoto and Prof. Toshiyuki Yamamoto for their crucial advice and comments as the members of doctoral examination committee. Their constructive comments helped me a lot to consolidate this dissertation.

Many thanks go in particular to Assistant Prof. Miho Asano, Dr. Danpeng Ma and other seniors. I am much indebted to them for their valuable advice in technical discussions, which helped me to finalize the research methodology and objectives.

It is a pleasure to pay tribute to my friends as well as the members of our laboratory who have shared their time with me and kindly supported me a lot. It is a nice experience for me to go through my study in Japan with them full of happiness and great memories.

I am very grateful to the financial support of China Scholarship Council (CSC). With the granted scholarship, I can concentrate better on my research. I am also indebted to Prof. Xinrong Liu of Chongqing University and Prof. Naixing Liang of Chongqing Jiaotong University, China for their encouragement for my study in Japan.

Finally, I would like to convey my utmost gratitude to my family especially to my wife for their understanding and patience during these years. It is their love and inspiration that support me to overcome the difficulty. The dissertation would not have been possible without them being with me in this journey. I dedicate this work to them all.

Abstract

In modern society, expressway serves as the main arteries for daily commuting/business trips and its performance is critical to quality of service (QOS) of the whole transportation network. Traffic safety, as one crucial aspect affecting QOS, its evaluation methodology remains absent at the stages of road planning and design. Without necessary performance check, improper road design and management may give rise to less safe operation.

For safer road design and traffic operation, it is required to identify crash characteristics and their influencing factors. With the realization of data collection preceding crashes, proactive traffic management strategies seem available through predicting the probability of crash occurrence in advance. However, most existing models are not perfect regarding their predictive powers. Despite the insufficiency of modeling methods, few studies have incorporated geometry, traffic flow and ambient conditions in a single model. Meanwhile, previous studies paid little attention to the different impact of traffic flow on crashes with the change of traffic conditions. Besides, crashes are characterized in facility/expressway type-specific, while existing models are focused on the whole network of expressway.

This study aims to develop a crash risk estimation model (CREM) to identify the affecting mechanisms of crash influencing factors at basic segments, considering the interaction of geometry, traffic flow and ambient conditions. CREM is developed separately for urban and intercity expressways and then is further compared between the two types of expressways, with the purpose to more comprehensively understand the cause of crash occurrence. The final objective of this study is to apply the measure of crash risk for traffic safety evaluation of geometric design and for proactive traffic management strategies.

In **Chapter 1,** background statements about the losses to society induced by expressway crashes, the importance of explanatory factor identification for geometric design/traffic management, and several significant characteristics of crashes on expressway are discussed. The related problems of existing studies are analyzed and the objectives of this study are hereby provided. Finally, the research outline is graphically illustrated.

Chapter 2 shows the state-of-the-art review on crash analysis. Fundamental characteristics of crash data and explanatory factors involved in existing studies are introduced. The strengths and weaknesses of popular crash analysis methods are discussed. Considering the nature of crash event as well as the objectives of this study, the matched case-control logistic regression is proposed as an appropriate method to measure the effects of various independent variables on crash occurrence, which is a binary outcome event in essence.

Chapter 3 describes the study sites including the section from Mikkabi interchange (I.C) to Yokaichi I.C of Tomei-Meishin Expressway referring to intercity expressway and Nagoya Urban Expressway corresponding to urban expressway. Five datasets over the year from 2007 to 2009 that consist of crash records, detector data, geometric design, traffic regulation records and daily sunrise/sunset time records, are utilized. Preliminary analyses on the differences of geometric design and traffic characteristics between urban and intercity expressways are conducted.

Following, **Chapter 4** explains the development of CREM for urban expressway. Crash rate (*CR*) statistics and Principal Component Analysis (PCA), as two proactive analyses, may identify the significantly independent variables through focusing on traffic conditions. Based on those variables, a matched case-control study is designed and then conditional logistic regression is applied for quantifying the effects of these variables on crash risk. The model demonstrates that 1) horizontal alignment is the most significant factor related to crashes, while its significance is on the decrease with the increase in traffic density; 2) in contrast, the effect of vertical alignment on crashes gets more important; and 3) owing to the more powerful inter-vehicle interaction, speed becomes more sensitive to crash risk as traffic density increases. Ambient conditions are not negligible exposures, since nighttime and holiday may increase crash risk compared to daytime and weekday, respectively.

By the similar process, CREM for intercity expressway is established in **Chapter 5**, and its crash characteristics including *CR* statistics and the sensitivities of variables to crash risk different from urban expressway are analyzed. By expressway type, geometric design is a major cause leading to higher *CR* and crash risk on urban expressway in low-density uncongested flow, such as poor geometric consistency induced by small curves and heavy driver workload caused by narrower cross section compared to intercity expressway. When traffic density increases, the inter-vehicle interaction gets more intensive, and then traffic conditions on intercity expressway get less safe for its vehicle composition characterized by higher percentage of heavy vehicle (HV). In congested flow, the variation in speed is quite sensitive to crash risk, and intercity expressway still has worse safety situation due to the interruption of HV to other traffic, especially driving on steep vertical slopes.

An extensive analysis on the applications of CREM is executed in **Chapter 6** regarding 1) the evolving process of crash risk with the change of traffic conditions and 2) the quality of geometric design, *e.g.,* assessing safety benefit and identifying crash-prone locations. It indicates that crash risk is convex downward to traffic density in uncongested flow, and following a decreasing trend in congested flow. Safety benefit of horizontal alignment can be more reliably measured through the prediction of crash risk in low-density uncongested flow. Comparatively, safety benefit of vertical alignment seems highly related to the prediction of crash risk in congested flow. Meanwhile, the potential crash-prone locations would be identified and they are further found out to be traffic condition dependent.

Finally, **Chapter 7** summarizes research conclusions and provides some recommendations for future works. For proactive traffic management strategies, the model in present study may provide leverage to predict hazardous conditions and avoid an impending crash. Crash risk estimation can be applied for assessing safety benefit of geometric design and the results of geometric variables related to crash risk may supply a benchmark for the improvement of performance-check expressway design. Besides, the concept of identifying crash-prone location based on crash risk estimation is considered more applicable for operational applications at the view point of traffic management. Concerning the better application for proactive traffic management and geometric design, directions for future works are addressed regarding the applicability to other facility types, the improvement of the model validity and the implementation of traffic control measures.

Table of Contents

List of Tables

List of Figures

Chapter 1

INTRODUCTION

1.1 Background

Expressway network plays an important role for daily commuting/business trips in modern society. National expressways, making up the majority of expressways in Japan, act as the arteries between metropolises. Inside the metropolises, urban expressways often serve as the channels for mass transit and high-speed transportation. In this sense, the performance of expressway is critical to quality of service (QOS) of the whole transportation network.

Traffic safety, as one crucial aspect affecting QOS, deserves more attention at the stages of expressway planning, design and operation, considering enormous losses to society caused by crashes in expressway operation. For safer geometric design and traffic operation, it is required to identify crash characteristics and their influencing factors. Meanwhile, with the realization of data collection immediately before crashes, growing concern over traffic safety has lead to research efforts directed to predicting crashes in advance, regarding the application of proactive strategies, *e.g.,* the Advanced Traffic Management Systems (hereinafter ATMS) and the Intelligent Transportation Systems (abbreviated into ITS).

1.1.1 Traffic crashes on expressways

In Japan, the most comprehensive set of annual statistics on traffic crashes is provided by National Police Agency (NPA). According to its annual report in 2012, a total of 665,138 crashes occurred, and induced 4,411 fatalities and 825,396 injuries in the whole country. The national annual social cost of traffic crashes is up to millions of dollars, corresponding to about 2% of the gross domestic product of Japan (Noked, 2010).

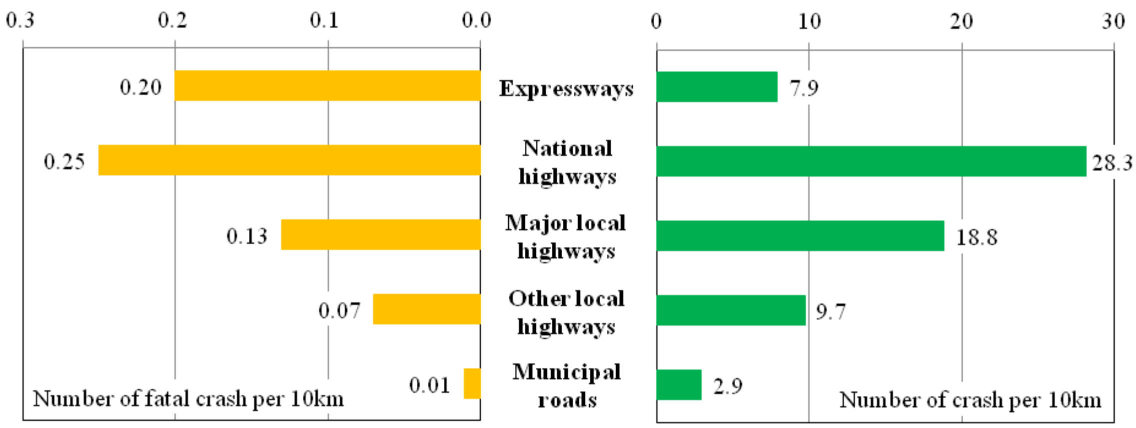


Figure 1.1 Number of crashes per 10km by road type in 2012
(Source: Annual Report of NPA, 2013)

Figure 1.1 demonstrates the number of total/fatal (at least one death involved) crashes per 10km by road type in 2012. It indicates that the value of total crashes on expressways is not as high as other road types. However, on expressways, the value of fatal crashes gets up to 0.2 per 10km, which is much high relative to other road types. It is worth noting that, the proposed threshold for classifying fatal crashes should be different to achieve desirable evaluation of crash severity by road type considering their different operating speeds.

Crashes on expressways not only impact traffic safety, but can also result in non-recurring congestion. In this regard, expressway crashes are also critical at the view point of traffic operation. Based on a report published by Japan Highway Public Corporation (JH) in 2003, the share of congestion induced by crashes is 19% on expressways. By comparison, more than 50% congestion is non-recurrent on expressways in US, and up to 55% of these congestions are caused by crashes (Chin *et al.,* 2004). In China, if crashes may be solved in time, more than 30% congestion can be potentially reduced (tranbbs.com, 2012).

1.1.2 Safety performance evaluation for geometric design

Operational performance, using to be an invaluable source in assessing the effectiveness of an investment plan, has served as an essential element in determining the most effective allocation of limited resources. Developing performance measure has gained momentum, and it further becomes a well-understood and widely applied procedure. Much research has been carried out for evaluating the quality of service from the view points of efficiency and smoothness of traffic flow. By contrast, less similar approaches exist for the evaluation of safety performance, which is another crucial aspect affecting QOS provided by expressway systems. Without necessary performance check, improper road design and management may give rise to less safe expressway operation.

In Japan, a series of Road Structure Ordinance have been published as a nation-wide road design standard, in which more concerns have been paid to transport efficiency and capacity provided by road structure. Even crash rate (hereinafter *CR*) statistics are provided for some combinations of vertical slopes with horizontal curves by road type, the reliability of these results is not ideal enough due to the limited samples of crash events. Furthermore, traffic demand has not been considered for those statistics, while the affecting mechanisms of geometric design to crashes may differ with the change of traffic conditions (Wu *et al.,* 2012a). Hence, safety performance has not been considered well in that standard, and the related evaluation measures remain absent at the stages of road planning and design.

In practice, owing to the technology and richer experience in geometric design, road design has become more of an engineering procedure with optimization. The new approach is called as an assessment of objectives, one of which is design consistency (Ng and Sayed, 2004). It has been suggested that identifying and treating any inconsistency of a highway can significantly improve its safety performance. Considerable research has been operated to explore that concept including measures and models to estimate design consistency, while little attention has been paid to quantify the safety benefit of consistency. In such case, the reliability to regard these existing design consistency measures as the way of safety performance evaluation deserves to be further verified.

1.1.3 Necessity of crash analysis for traffic control strategies

Traditional research on traffic management has focused on incident detection, whose idea involves analysis of patterns in traffic surveillance data observed just after the incident, named post-incident detection, as shown in Figure 1.2. However, for incidents which may induce serious harms or even result in life-threatening aftereffects, *e.g.,* crash and disease, post-incident detection has limited practical utility. With the revolution of information technology, the data immediately before incidents is available for traffic management. Spurred by these progresses, a new concept of incident detection is proposed, and expressway authorities are becoming more interested in proactive strategies, as shown in

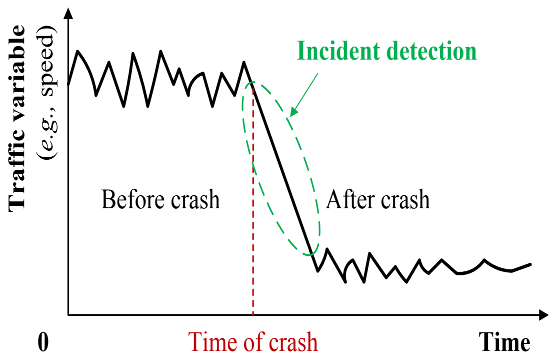


Figure 1.2 Idea of incident detection

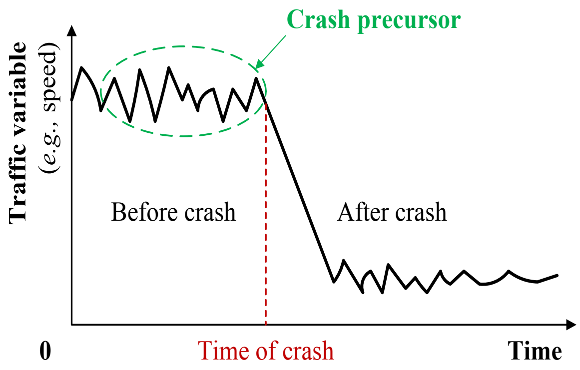


Figure 1.3 Idea of proactive traffic management

Figure 1.3 (Abdel-Aty and Pande, 2007). For traffic management, these strategies would involve anticipating incidents along with strategies to avoid them together.

In order to develop proactive traffic management strategies, traffic data prior to individual historical crashes should be extracted and analyzed, called as crash precursor. Meanwhile, the affecting mechanisms of explanatory variables on crash occurrence should be identified in order to quantify the influences of these variables on crash occurrence. In this way, the probability of crash occurrence under a given condition would be predicted for a very small time window using the short-term traffic data. These findings may provide leverage towards predicting and avoiding an impending crash, with the overall purpose to serve the proactive traffic management strategies (*e.g.,* ATMS and ITS) well.

1.1.4 Several significant characteristics of crash occurrence

A growing concept on the cause of crashes is that "disruptive" traffic conditions contribute much more to crash occurrence relative to "normal" traffic conditions (Oh *et al.,* 2001). Other than traffic demand, inadequate geometric design and the alteration in ambient conditions are the other reasons leading to the variation in traffic conditions in different ways. In this sense, crash events are essentially associated with the interaction of geometry, traffic flow and ambient conditions (Bajwa *et al.,* 2010). Therefore, any analysis towards identifying the cause of crash occurrence should incorporate those factors.

In Highway Capacity Manual (HCM, 2010), it suggests divide expressways into various uniform segments, defined as facility types, for identifying traffic characteristics such as capacity and level of service (LOS). Other than capacity and LOS, vehicle maneuvers may be facility type-specific as well. For example, different from basic segment, there is a ramp added at merge segment, and vehicle maneuvers are primarily characterized by frequent merging and lane-changing behaviors, which is much different from basic segment where car-following behavior is dominant. Given these situations, it is reasonable to assume that crash characteristics would be different during various facility types.

For different types of expressways, the standards of geometric design are different. Besides, various transportation functions are provided by individual expressways. Correspondingly, traffic characteristics such as speed, vehicle composition and driver population are actually different by expressway type. It is also necessary to make a distinction between different types of expressways for more comprehensive understanding the effects of variation in geometric design and traffic characteristics on crash occurrence.

By experience, with the change of traffic conditions, driving conditions are varied, and the mechanisms of traffic flow affecting crashes may be different (Wu *et al.,* 2012a). In the meantime, the way how geometric design and ambient conditions are related to crash characteristics seems different as well in various traffic conditions. Hence, crash modeling by traffic conditions may be more effective for identifying the mechanisms of explanatory factors on crash characteristics compared to these do not categorize traffic conditions.

1.2 Problem Statement

Even numerous studies have established statistical links between crashes and various explanatory factors, few studies have incorporated geometry, traffic flow and ambient conditions in a single model to investigate their combined effects on crashes. Meanwhile, one other limitation of existing crash models is data collection. In those studies, traffic conditions are generally represented by low-resolution data that is collected at a highly aggregated level (*e.g.,* hourly and daily flow). Besides, geometric features are primarily reflected in terms of the hierarchy of radius or gradient. However, these approaches have been criticized for "ecological fallacy" in data collection that highly-aggregated and static statistics are insufficient to investigate the natures of individuals (Golob *et al.,* 2004).

Another concern is the applicability of crash risk prediction models. To date, most of the existing models are not perfect regarding their predictive powers (Hossain *et al.,* 2012). Despite the limitation of data collection, another potential cause to undermine the validity of model is the inadequate modeling process. Although the mechanism of traffic flow affecting crashes may be different with the change of traffic conditions, most previous studies paid little attention to this regard. Furthermore, these models are focused on the whole road network, while crash characteristics are varied by facility/expressway types. Given these problems, the reliability of existing models is required to be improved. Therefore, a methodology which can capture the natural cause of crashes considering the interaction of geometry, traffic flow and ambient conditions through focusing on traffic conditions is highly expected for specific facility/expressway types.

1.3 Research Objectives

This study aims to develop a crash risk estimation model (hereinafter CREM) to identify the effects of explanatory factors on crash occurrence considering the interaction of geometry, traffic flow and ambient conditions. Regarding the available data, basic segments are focused on, where the percentage of crash events taking out of the total value in whole network is highest by facility type. For example, it can arrive at 31.9% and 38.3% for Nagoya Urban Expressway network and Tomei-Meishin expressways (intercity expressway), respectively (will state in Chapter 3). CREM is then further compared between urban and intercity expressway, with the purpose to more comprehensively understand the cause of crash occurrence. The overall objective of this study is to apply CREM for safer geometric design and proactive traffic control strategies. Such an objective is expected to be achieved through the following steps:

- Operating two proactive analyses: 1) *CR* statistics to identify the relationships between *CR* and traffic density, in order to categorize traffic flow in terms of traffic safety, and 2) Principal Component Analysis (PCA) for identifying the correlations of explanatory factors and their significances affecting crashes.
- Developing CREM according to the significant and independent variables through incorporating geometry, traffic flow and ambient conditions.
- Comparing crash characteristics and the sensitivities of explanatory factors to crash risk between urban and intercity expressways.

- Demonstrating the applicability of CREM for measuring the quality of geometric design and predicting the evolving process of crash risk with the variation in traffic conditions based on a subject basic expressway segment.

1.4 Research Outline

After clearly defining research objectives, a brief review of the major issues shared by previous studies is given in **Chapter 2**. Two types of expressways, Tomei-Meishin Expressway representing intercity expressway and Nagoya Urban Expressway referring to urban expressway are involved. Their geometric features are described in **Chapter 3**, and the different traffic characteristics between the two types of expressways are generally analyzed. Following, **Chapter 4** demonstrates the process of CREM at the basic segments of urban expressway. *CR* model and PCA, as two proactive analyses, may identify the significant and independent variables through focusing on traffic conditions. Based on those variables, a matched case-control study is designed and then conditional logistical regression is applied to quantify the impacts of those variables on crash risk. By the same process in Chapter 4, CREM for intercity expressway is established in **Chapter 5**. Meanwhile, crash characteristics and the sensitivities of explanatory variables to crash risk are comparatively analyzed between urban and intercity expressways. An analysis on the applicability of CREM is conducted in **Chapter 6** regarding the evolving process of crash risk along with traffic conditions and the quality of geometric design. Finally, conclusions and some perspectives for future research are provided in **Chapter 7**. The general research outline is presented in Figure 1.4.

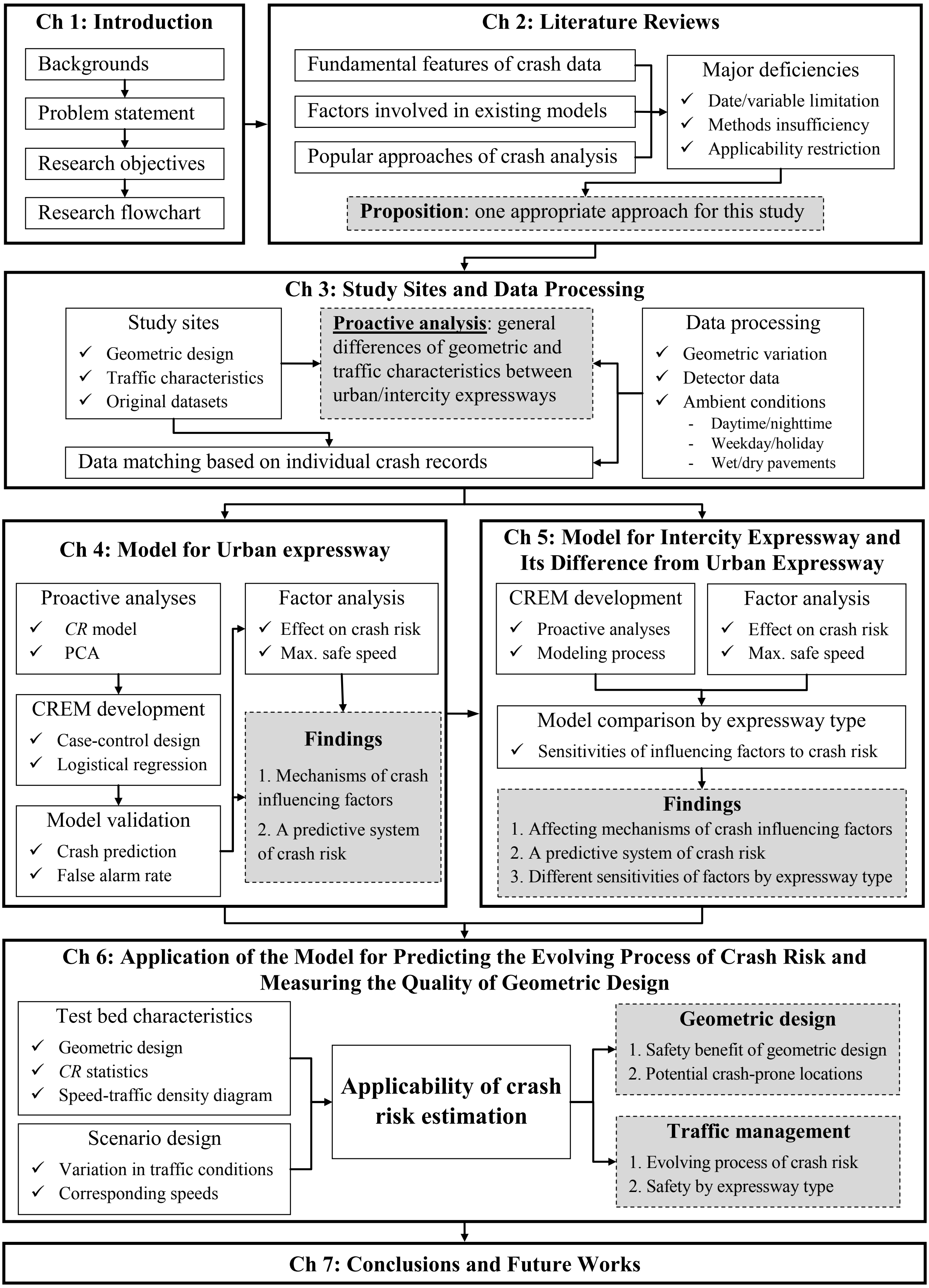


Figure 1.4 General research flow of the dissertation

Chapter 2

LITERATURE REVIEWS

2.1 Introduction

Improving traffic safety is a worldwide issue to be relieved urgently. Crash characteristics and their influencing factors, as the theoretical basis for safety improvement, may provide direction for policies and countermeasures aimed at smoothing hazardous conditions. For a better understanding of crash influencing factors, researchers have continually sought ways through an extensive array of approaches, such as traffic conflict technique and crash analysis. As a direct measure, crash analysis is still the most widely adopted approach to assess safety of a transportation facility (*e.g.,* expressways and intersections).

Since traffic crash itself is a kind of rare event, to assure the validity of crash analysis, it is necessary to select a reliable methodology to identify the affecting mechanisms of various factors. For this purpose, this chapter comprehensively reviews the state of the art of the related studies on the nature of crashes and the involved methods.

In the following, section 2.2 firstly discusses the fundamental characteristics of crash data. Then, explanatory factors and their associations with crashes are presented in section 2.3. Next, section 2.4 reviews the currently popular crash analysis methods. Meanwhile, their strengths and weaknesses to reveal the nature of crash occurrence are also discussed. Based on these reviews, an appropriate approach for this study is finally proposed in section 2.5 considering the crash data available and the objective of this study.

2.2 Fundamental Characteristics of Crash Data

After a long period of practice, the characteristics of crashes are gradually accepted to be discrete, rare and non-negative (Garber and Wu, 2001). The issue of crash data may be a potential source of error regarding the correctness of statistical models to reveal crash characteristics. Some characteristics of crash data have been summarized in Lord and Mannering (2010), while their related methodological problems are evaluated in terms of crash-frequency predictions. For this reason, the fundamental characteristics of crash data and their potential insufficiency to identify the above nature of crashes are discussed in this section, as summarized in Table 2.1 generally.

Table 2.1 Fundamental characteristics of crash data

Characteristics	Associated methodological problems
Limited crash samples	• The desirable large-sample crash properties cannot be realized, which would lead to biased estimation on crash influencing factors
Over- and under-dispersion	• Incorrect factors estimates for crash frequency analysis • Over- and under estimation of the effects of factors on crash risk
Different styles of crash records by road agency	• Reducing the availability of analysis through combining datasets • Variable space for crash models developed based on different records • Loss of important information related to difference by road type
Facility/road type-specific crash characteristics	• Under estimated crash influencing factors for a specific facility type as well as for a given road type
Crash characteristics by type of crash	• Type-specific crash characteristics and their related influencing factors cannot be revealed reliably through total crash models

2.2.1 Limited crash sample size available

Crash data collection, as a process highly related to collecting technique, its efficiency is determined by the improvement of collecting system. Meanwhile, that process may require the coordination of various departments and cause large costs. Thus, some existing models are developed based on limited crash samples. In such case, the desirable large-sample properties of some parameter-estimation techniques (*e.g.,* maximum likelihood estimation) would be not realized (Lord and Mannering, 2010).

On the other side, comparing to serious crashes, crashes with low severity are likely to be missing in crash datasets, which is known as incomplete reporting of crash data and has been a major problem in highway safety analysis for many years (Elvik and Mysen, 1999). Several researches have suggested that the crash count-related models are likely to produce biased estimations when incomplete reporting is not considered in the process of crash modeling (Kumara and Chin, 2005; Ma, 2009).

2.2.2 Over-dispersion and under-dispersion

As a rare event, crash can happen at any location in theory, while crash counts may allocate heterogeneously along roads due to the variations in geometric design and other features. Correspondingly, in a given period, the variance of crash frequency by roadway sections may exceed the mean value of crash frequency along the whole network. Such case is defined as over-dispersion. By contrast, crash data can sometimes be characterized by under-dispersion in which the mean of crash counts is greater than the variance, especially when the sample mean value is very low (Oh *et al.,* 2001).

As is known, some crash frequency modeling approaches, *e.g.,* Poisson regression model, restrict the mean and variance to be equal. If over-dispersed data are present, a common Poisson regression model can result in biased and inconsistent parameter estimates which in turn could lead to erroneous inferences regarding the factors that are used to determine crash frequencies (Park *et al,* 2007). Likewise, the presence of under-dispersed data can also induce many traditional deterministic models and cause incorrect parameter estimates. Even for crash models at individual level, over-dispersed data can result in overestimation on the influence of the dominant factors to crash occurrence. By contrast, the effects of other variables, primary exist in under-dispersion sections, would be underestimated.

2.2.3 Different styles of crash records by road agencies

Road administrators often vary by cities, regions and departments. Correspondingly, crash events will be reported by different agencies. One potential result from the non-unity crash records is the diversity of variable styles. Such characteristics reduce the availability of crash analysis through combining various datasets in the purpose to increase crash samples. Furthermore, it also induces a variable space for the existing crash risk prediction models (Hossain and Muromachi, 2012), since the related variables recorded in different ways. Even for the comparison on crash characteristics, such analysis should be based on the data recorded in the same style. In such case, some important information would be lost, which may be critical to find the differences by cities, regions and road types.

2.2.4 Crash characteristics by road facilities as well as by road types

As a general rule, Highway Capacity Manual (HCM, 2010) suggests divide expressways into various uniform segments based on their own structure for the analysis of some traffic characteristics such as capacity and level of service (LOS). Those segments are named as facility types, where vehicle maneuvers are practically different as well. Such as merge segment, there are merging behaviors and more frequent lane-changing behaviors, which is different from those at basic segment where car-following behaviors are dominant. In this regard, it is reasonably assumed that crash characteristics are facility type-specific. In view of this concern, Wu *et al.* (2012a, 2012b) identified the difference of *CR* models by facility type on urban and intercity expressways, respectively. As a conclusion, they finally suggested to analyze influencing factors separately based on individual facility types. Correspondingly, it may provide more effective direction for policies and countermeasures aimed at evolving hazardous conditions for a specific facility type.

By road type, geometric and traffic characteristics, such as speed, vehicle composition and driver population are actually different as well. The differences imply the necessity of crash analysis through differentiating road types for more comprehensive understanding the crash characteristics. For this reason, Wu *et al,* (2013a) compared *CR* characteristics and the related influencing factors between urban and intercity expressways. As an expectation, the differences enable road authorities to investigate safer geometric design and traffic control strategies more practicable for each type of expressway.

2.2.5 Crash characteristics by type of crash

The related analyses have demonstrated that different types of crashes may occur under substantially different conditions and associated with explanatory variables in different ways (Christoforou *et al.,* 2011). Kim *et al.* (2006) further argued crash type models are useful for at least three reasons: 1) the need to identify sites that are high risk with respect to a specific crash type but that may not be revealed through total crash modeling, 2) countermeasures are likely to affect only a subset of all crashes, and 3) different crash types are usually associated with geometry, traffic flow and ambient conditions in different ways. Pande and Abdel-Aty (2006) underlined the significance of by-crash-type analysis, particularly when it comes to real-time risk prediction. Based on their analysis, it is shown that the conditions preceding crashes differ by type of crash and. Therefore, any approach towards proactive traffic control strategies for safety should be type-specific in nature.

However, due to the difficulty in collecting the necessary data, most existing studies have not distinguished crash types well. In such context, several researches are performed after segregating crashes into two large categories: single- and multi-vehicle crashes (Qin, *et al.,* 2004; Christoforou *et al.,* 2011). In other studies, crash influencing factors are explored just considering the types of primary crashes, which are often distinguished between rear-ends and side-swipes (Golob *et al.,* 2004; Lee *et al.,* 2006).

2.3 Crash Influencing Factors

Due to the complexity of crash occurrence, many explanatory factors have been considered and applied in crash modeling. These factors generally include geometric design, traffic flow and ambient conditions like lighting and weather. This section discusses the factors and their association with crash characteristics through reviewing the past studies.

2.3.1 Road geometry

Studies focused on geometry and safety aim to improve geometric design and eliminate crash-prone locations. In the studies, the effects of elements such as horizontal curve and vertical grade on safety have been primarily analyzed. Since geometric design consistency is emerging as an important rule in highway design, some studies pay attention to this regard. They finally find that the measure of design consistency is more effective to reveal the driver-vehicle-roadway interaction than that just involves single design element.

1) Horizontal curve

Horizontal curve is one of the most important geometric factors since vehicle maneuvers would be strongly impacted by the variation in curvature (Gao *et al.,* 2004). Hauer (1999) investigated the correlation of safety with the characteristics of horizontal curves. One clear finding is that the change of crash rate is proportional to the change of radius length. Besides, larger radii correlate with fewer crashes. Garber and Kassebaum (2008) studied nearly 10, 000 crashes on urban and rural 2-lane highways in Virginia and found that the predominant type of crash is run-off-road crash in curve sections. Hummer *et al.* (2010) identified crash characteristics in curves through 51,000 crash samples on 2-lane road sections in North Carolina. Their results show that rural horizontal curves are particularly susceptible to crashes. Meanwhile, curve crashes occur more often on sections with a grade rather than on tangent sections on a grade. As a typical study on the effects of geometry, Pei and Ma (2003) audited 141,812 crash events on ShenDa Expressway in China (348km long in total) and traced back to investigate how horizontal curve relates to *CR*. They finally developed a power exponent model as a function of curve radius (*R*), which is demonstrated in Figure 2.2.

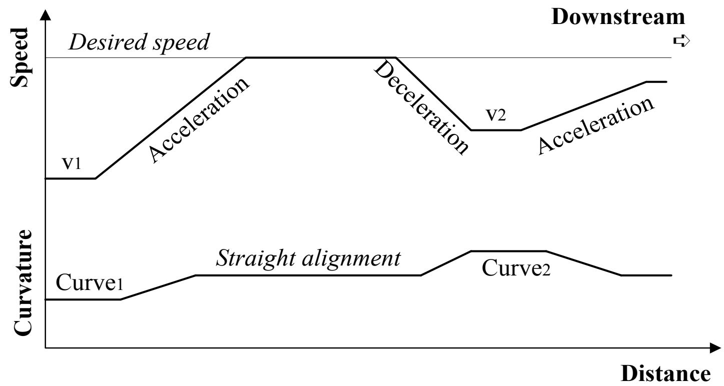


Figure 2.1 Operating speed along the variation in horizontal curvature
(Source: Gao *et al.*, 2004)

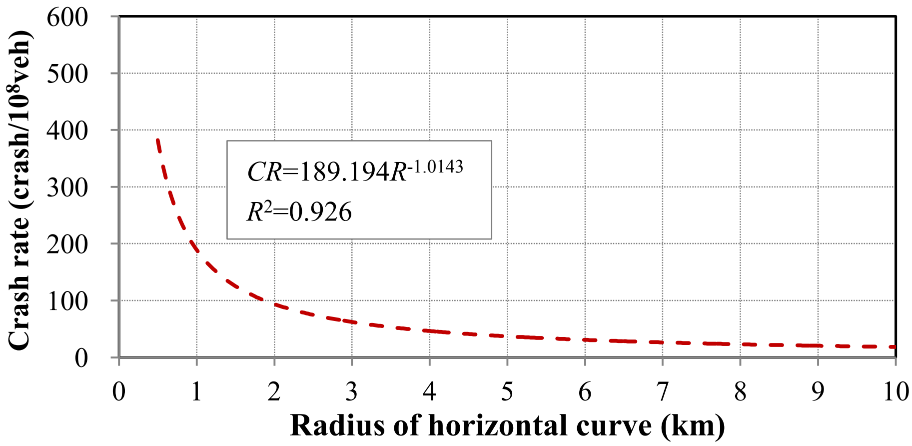


Figure 2.2 Relationship between *CR* and curve radius (*R*)
(Source: Pei and Ma, 2003)

2) Vertical grade

Numerous studies have investigated the influences of vertical slope on crashes, especially for heavy vehicle (hereinafter HV) due to its poor dynamic when driving on slopes. The general finding is that effect of HV related to *CR* would be higher with the increase in vertical gradient. Later, several studies (Pei and Ma, 2003; Hauer, 2006) concluded that between uphill and downhill grade impacts, not same correlation with crash occurrence can be concluded. With respect to this regard, Wang (2005) further found no relationship between vertical grades and crashes in an uphill direction, while in a downhill direction, such kind of relationship does exist though the correlation is no more than 0.4. To quantify the effects of vertical gradient on *CR* at descending sections, Fu *et al.,* (2011) extracted 1413 crash samples over an 85.43km section of expressway in western China. In that study, *CR* appears a good exponential relationship with vertical gradient, as illustrated in Figure 2.3. For crash risk estimation, Rengarasu *et al.,* (2009) and Bajwa *et al.,* (2010) represented vertical slope in terms of gradient, which is positive value for uphill and negative value for downhill slopes. Their findings imply that vertical slope is negative to crash occurrence, in other words, crash risk would increase on downhill sections.

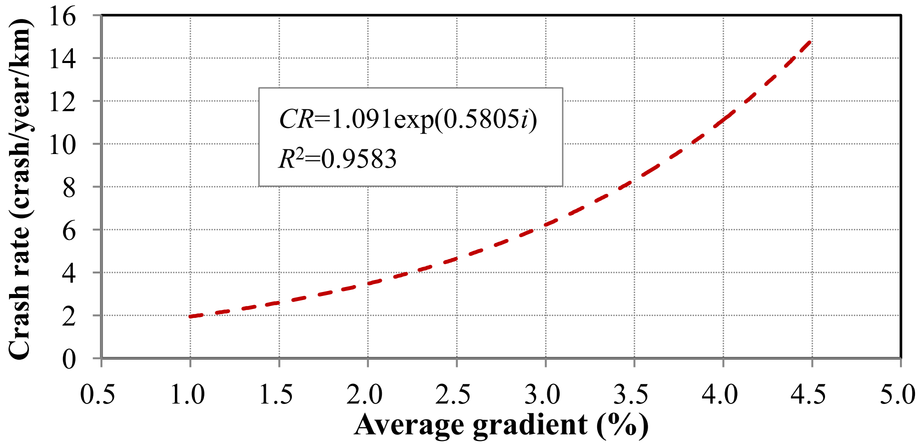


Figure 2.3 Relationship between *CR* and vertical gradient (*i%*)
(Source: Fu *et al,* 2011)

3) Geometric design consistency

As suggested in Ng and Sayed (2004), 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. The poor consistency can violate driver's expectation, and the driver may adopt an inappropriate maneuver adjustment (*e.g.,* speed reduction), potentially leading to crashes.

Several measures of design consistency have been available: operating speed, vehicle stability, driver workload and alignment indices. Amongst them, operating speed is widely considered to be the most notable and straightforward design consistency measure (Watters and O'Mahony, 2007). Operating speed is defined as the speed selected by highway users when not restricted by other users, and is normally represented by the 85% percentile operating speed (v_{85}). Based on the idea that the variation in speed is a visible indicator of design consistency, Lamm *et al.* (1999) developed design evaluation criteria to examine the consistency of neighboring road sections, as summarized in Table 2.2.

Table 2.2 Evaluation criteria of geometric design

(Source: Lamm *et al.,* 1999)

Evaluation	Criterion I* (km/h)	Criterion II** (km/h)	Criterion III***
Good	$\|v_{85}-v_d\|\leq 10$	$\Delta v_{85}\leq 10$	$\Delta f_R\geq 0.01$
Fair	$10<\|v_{85}-v_d\|\leq 20$	$10<\Delta v_{85}\leq 20$	$0.01>\Delta f_R\geq -0.04$
Poor	$\|v_{85}-v_d\|>20$	$\Delta v_{85}>20$	$\Delta f_R<-0.04$

* Operating speed at a single section; v_d stands for design speed value at one section.
** Operating speed at two successive sections; $\Delta v_{85} = |v_{85i}-v_{85i+1}|$, where v_{85i} and v_{85i+1} are operating speed on section i and $i+1$, respectively.
*** Vehicle stability; $\Delta f_R = f_R-f_{RD}$, where f_R and f_{RD} separately refer to side friction assumed and demand at section i.

A vehicle negotiating a horizontal curve can experience excessive centripetal force. Thus, Criterion III is suggested to evaluate design consistency through ensuring that enough side friction supply (f_R) is available to meet the side friction demand (f_{RD}). Both indices are calculated through the following formulas (Lamm et al., 1999).

$$f_R = 0.25 - (2.04\times 10^{-3} v_d) + (0.63\times 10^{-5} v_d^2) \tag{2.1}$$

$$f_{RD} = (v_{85}^2 / 127R) - e \tag{2.2}$$

Where, R is curve radius (m); e refers to superelevation rate.

Paul and O'Mahony (2007) qualitatively indentified the safety benefits of geometric design consistency on rural single carriageways in Ireland. They confirmed that design evaluation can be used to pin point locations on highways where crashes could conceivably be higher. Through using a comprehensive crash and geometric design database of two-lane rural highways in the province of British Columbia, Canada, Ng and Sayed (2004) incorporated seven design consistency measures to quantify the impacts on crash rate. It was concluded that models explicitly considering design consistency may identify the inconsistency more effectively and reflect the impacts on safety more accurately than those that do not.

However, these measures in nature are still focusing on the variation in curve sections, and are inadequate to reveal the true driver-vehicle-roadway interaction that is virtually varied by location. Furthermore, the above studies are primarily based on rural highways. Given these problems, Hikosaka and Nakamura (2001) employed the indices of geometric variation in road elevation and horizontal displacement in the direction of tangent to curve (will introduce in Chapter 3), along traffic direction for intercity expressways in Japan. Both indices are finally found out to be the most significant crash influencing factors.

2.3.2 Traffic characteristics

Researchers over the past several decades have conducted significant number of studies to identify the effects of traffic flow on road crashes. Existing studies in this area have tried to analyze the relationship between traffic conditions, as represented by long term traffic data, and crash rate, traditionally helping in identifying where "more crashes are likely to occur" (Abdel-Aty and Panda, 2007). In the recent years, surveillance apparatus that continuously records traffic data have become possible on instrumented roads. Availability of these data has inspired a new series of studies in traffic safety, which attempt to identify conditions where "a crash is more likely to occur" (Abdel-Aty and Pande, 2007). The distinction between two groups of studies is that the latter group of conditions would change based on the varying traffic patterns over the courses of day or even within hour. In view of their different modes of data processing, the two groups of studies are defined as aggregated and disaggregated analyses in the following reviews, respectively.

1) Traffic analysis at aggregated level

As a conventional approach, previous studies have established statistical links between crash frequency and traffic characteristics. Tracing back to 1964, Solomon found a U-shaped curve for the relationship between *CR* and speed difference (difference between speed and average speed). Several follow-up studies were completed by Lord *et al.,* (2004) and Hauer (2009). They revealed nonlinear relationships between *CR* and speed variables (*e.g.,* average speed and speed difference). Other than speed, some articles discussed flow rate-*CR* relationship (Zhou *et al.,* 1997; Zhang *et al.,* 2012). Most of those researches used the Average Annual Daily Traffic (AADT), and reported a downward convex relationship between AADT and *CR*. The usage of AADT, a highly aggregated measure of exposure, might conceal some traffic characteristics that may be critical to crash occurrence, and result in underdispersion on variables (Pasupathy *et al.,* 2000). Hence, other studies moved toward hourly crash analysis, and some hourly flow variables (v/c ratio in Chang *et al.,* 2000; level of service in Pasupathy *et al.,* 2000; hourly volume in Martin, 2002) replaced. However, hourly crash analysis is still an aggregated methodology. Furthermore, too many traffic variables do not help the clear indication of the effects of traffic flow on crashes.

In view of the problems above, some studies moved towards update *CR* models by using high-resolution data that is collected in a short time intervals like 5 minutes (Hikosaka and Nakamura, 2001; Zhong *et al.* 2011; Wu *et al.* 2011). In these studies, traffic conditions preceding crashes would be reflected much more reliable than previous ones. However, for *CR* statistics, traffic characteristics for individual crashes are further re-aggregated albeit the classification of traffic conditions getting smaller, inevitably susceptible to the problem of ecological fallacy (Golob *et al.,* 2004). The ecological fallacy is a widely recognized error in the interpretation of statistics, whereby inferences on the nature of individuals are based solely upon aggregated statistics for the group to which the individuals belong.

2) Traffic analysis on disaggregated level

A growing concept on the cause of crashes is that “disruptive” traffic conditions contribute much more to crash occurrence relative to “normal” traffic conditions (Oh *et al.,* 2000). This concept contrives the opportunity to improve the conventional aggregated approach to identify hazardous conditions with the advance of data collection, storage and analysis techniques. Furthermore, it also renders incident detection algorithm somewhat irrelevant and traffic control departments become more interested in proactive strategies. For this purpose, traffic data prior to individual historical crashes should be extracted and analyzed (see Figure 1.3). Correspondingly, several studies have proposed a serial of real-time crash prediction models based on the hypothesis that the probability of crash occurrence under a given condition can be predicted for a very short time window using the instantaneous traffic data (Lee *at al.,* 2002; Hossain and Muromachi, 2012).

Regarding the estimated variables, Oh *et al.* (2005) introduced a real-time hazardous traffic condition warning system developed by the standard deviation of speed, and suggested that variable is the best indicator of a "disruptive" traffic flow leading to crash. Abdel-Aty and Pemmanabonia (2006) found that the 5-min average occupancy, standard deviation of volume and the coefficient of variation in speed can affect crash most significantly. Dias *et al.* (2009) applied level of congestion rather than the speed as a predictor and affirmed a positive correlation between congestion and crash risk. Based on two routes (Shibuya 3 and Shin-juku 4) from Tokyo Metropolitan Expressway network, Hossain and Muromachi (2012) developed a crash prediction model considering congestion index and the difference in speed and occupancy between up- and downstream location of crash points.

However, in existing models, diverse variables are involved, which would undermine the applicability of these models. Such problem may be related to the nature of crash that is a complex event and accounts a wider range of variables. Another cause is the crash records different by road agency, a fundamental characteristic of crash data introduced in section 2.2. In this sense, such problem would be remained in the near future. On the other hand, the above studies are more concerned on identifying the factors and pay little attention to the affecting mechanisms of individual factors. Furthermore, the facility/expressway type-specific crash characteristics have not been concerned well, and thus the relevance and transferability of these findings from the above studies to a specific expressway facility type may be not be justified adequately.

2.3.3 Ambient conditions

Ambient conditions are another not neglectable explanatory factor, and sometimes, they would be critical to crash occurrence. Previous studies have tentatively investigated the influences of weather/pavement conditions, ambient light and day types (*e.g.* holiday and weekday), on crash characteristics.

1) Weather/pavement conditions

Inclement weather conditions (*e.g.* rainy and snowy) would constrain visibility and reduce the rolling tire-pavement friction (combining with wet pavement) that impairs roadability. Accordingly, traffic characteristics are affected, as demonstrated in Figure 2.4 (HCM, 2010), which would throw a negative influence on traffic safety (Chung *et al.,* 2005). Kopelias *et al.* (2007) discussed in detail the influence of weather on crash frequency and severity based on Attica Tollway in metropolitan Athens, Greece. In that study, a contribution about 5% to 10% taken by the presence of rain and wet pavements to crash counts and severity observed in 2004 and 2005 on that highway.

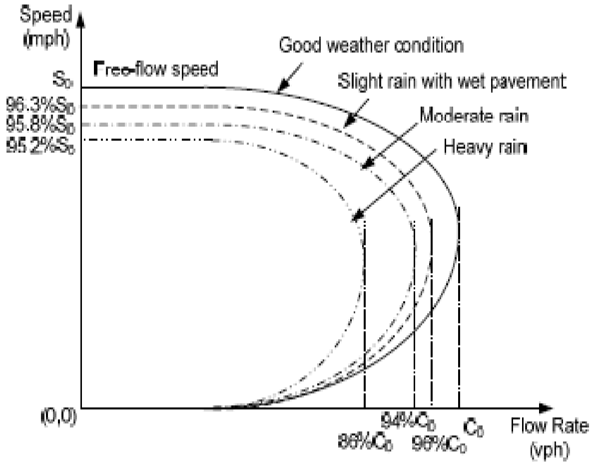


Figure 2.4 Traffic characteristics between different weather conditions
(Source: HCM, 2010)

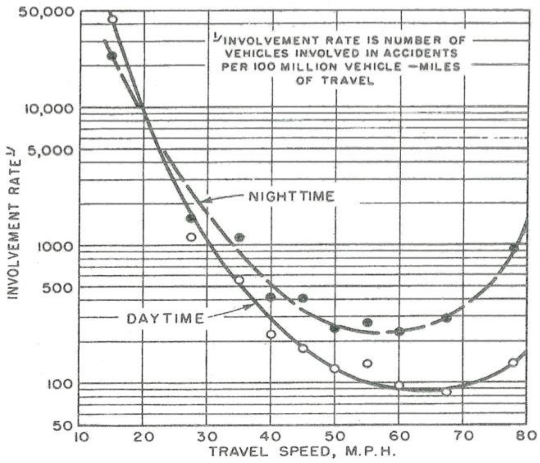


Figure 2.5 Relationships between *CR* and travel speed by ambient light condition
(Source: Solomon, 1964)

2) Ambient light

Comparing to daytime, the darkness in nighttime can result in great visibility loss, and correspondingly affects traffic characteristics such as speed/capacity drops and speed variation rise (Chung *et al.,* 2006; Shi *et al.,* 2011). Necessarily, crash characteristics differ with the variation in ambient light (Martin *et al.,* 2002; Sivak *et al.,* 2007). Generally, it is accepted that driving in nighttime is substantially riskier than driving during daytime. As shown in Figure 2.5, as a typical crash analysis by Solomon (1964) that can be referred from many literatures, crash frequency in daytime is clearly lower relative to nighttime, especially when driving at high speed. Sivak *et al.* (2007) also found out that the current fatality rate in U.S is 2.27 times higher at night than during the day.

3) Day type

In general, holidays refer to a large increase in recreational private travel, which may result in more long-distance trips, more travels in unfamiliar conditions and more drinking driving in contrast with normal weekdays (Anowar *et al.,* 2013). Based on a statistics of Alberta Transportation (2008), crashes during holidays represent only a small percentage (less than 10%) of total crashes, while the number of fatal crashes occurring on holidays is much higher than these during non-holidays (18% higher). Through analyzing fatal crash samples during major American holidays, Framer and Williams (2005) attributed those crashes to the probable combination of increased recreational travel, alcohol consumption and excessive speeding during holidays. Amongst other possible reasons, travel on rural unfamiliar roads, driver distractions and fatigue are suggested as causes resulted in the increased likelihood of driver committing errors.

2.4 Approaches of Crash Analysis

As stated before, in terms of the unit of analysis, the popular crash analysis approaches can be classified into aggregated and disaggregated levels. Generally, two approaches are represented by crash frequency modeling and crash risk prediction modeling, respectively. Focus of the models is generally two-fold on: 1) modeling methodology and 2) the parameters used as dependent and independent variables (Chang, 2005). Empirically, previous studies have used varied sets of variables depending upon the scope of research, which have been enumerated in section 2.3 in detail. Following, this section discusses the methodological advances of the existing crash models.

2.4.1 Crash frequency modeling

To deal with the data and methodological issues associated with crash frequency modeling, a wide variety of methods have been applied in the past decades. Initially, the linear and multiple linear regressions are used as modeling function. With the profundity of know in nature of crashes, it is found that this kind of regression cannot describe the characteristics of crashes adequately. Instead, Poisson or Negative Binomial (NB) regressions are considered to be better suited for defining the random, discrete and nonnegative nature of crash event (Milton and Mannering, 1998). For Poisson regression, it requires the mean and variance of the crash data to be equal, while crash data have the fundamental characteristics of over- and under-dispersion that impairs the hypothesis of Poisson model. In this case, some studies underline the application of NB model, and regard it as the most popular method for crash frequency modeling since NB model has all the desirable statistical properties and also relax the above hypothesis (Miaou, 1994). Those models mentioned above are essentially based on the ability of model function to capture the underlying distribution of crash frequency data (Abdel-Aty and Pande, 2007). To avoid any pre-defined assumption about the distribution of crash data, some researchers recently have proposed ‘distribution free’ methods for crash frequency modeling. Chang and Chen

(2005) adopted Classification and Regression Tree (CART), the most commonly used data mining technique, to analyze freeway crash frequency. Meanwhile, Neural and Bayesian Neural Network (BNN) models have been utilized in highway safety analysis mainly as an estimation tool for crash frequency (Chang, 2005; Xie *et al.,* 2007). Overall, these models exhibit better linear/non-linear approximation properties than traditional count-model approaches (Xie *et al.,* 2007). A comparison of the results from the CART, BNN and NB regression models demonstrated that both CART and BNN are good alternatives to NB regression for crash frequency modeling (Chang, 2005).

2.4.2 Crash risk prediction modeling

Through using individual crashes as analysis unit, in other words, disaggregating the data, a series of crash prediction models on real-time basis have been developed. The typical modeling methods can be generally classified into statistical methods and artificial intelligence or data mining based methods (Hossain and Muromachi, 2012). The former includes matched case-control logistic regression (Bajwa *et al.,* 2010; Zheng *et al.,* 2010; Wu *et al.,* 2013b), probit model (Christoforou et al., 2011), log linear model (Lee *et al.,* 2003) and Bayesian statistics (Oh, *et al.,* 2001). Amongst them, the matched case-control logistical model is a popular statistical approach. Because crash is a binary outcome event (occurrence *vs.* no occurrence), and the goal of logistical regression is to identify the best fitting model that describes the relationship between a binary dependent variable and a set of independent variables (Abdel-Aty and Pemmanaboina, 2006). The latter methods are composed of different kinds of neural network (Oh *et al.,* 2005; Abdel-Aty *et al.,* 2008) and classification trees (Pande and Abdel-Aty, 2006).

Traffic-related variables, *e.g.,* flow rate, speed and occupancy are highly correlated with each other. During statistical analysis, most of these variables get dropped as part of modeling process. Hence, it is important to employ suited methods that can accommodate correlated variables and make best use of every variable to improve the prediction success (Hossain and Muromachi, 2012). For this purpose, principal component analysis (PCA) is performed for selecting lesser indecent variables while accounting for most of the variance in the dataset (Golob and Recker, 2004; Bajwa *et al.,* 2010; Wu *et al.,* 2013a). Neural network based modeling methods can accommodate correlated variables. However, they expect sufficient prior knowledge regarding the problem domain exhibited through the interrelationship of the predictors, whereas, studies on that nature are highly resource demanding and often not available during modeling (Hossain and Muromachi, 2012).

2.5 Analysis Approach Proposed

Based on the above reviews, major deficiencies related to current studies are the following:

- Crash models developed by few explanatory variables (*e.g.,* using traffic flow as the only explanatory variable in a model) are not adequate to identify the combined effects of geometry, traffic flow and ambient conditions on crash occurrence.
- Modeling process without separating facility types may result in biased estimation on explanatory variables for a specific facility type.
- Crash analysis just focusing on a single type of expressway would reduce the availability for comprehensively understanding the cause of crashes affected by the variation in geometric design and traffic characteristics.
- Crash models based on low-resolution traffic data (*e.g.,* daily and hourly flows) are insufficient to reveal the nature of crashes that is highly associated with the short-term turbulence of traffic conditions.
- Models indentified the safety benefits of geometric design, without consideration of design consistency, are difficult to reflect the driver-vehicle-roadway interaction reliably during crash occurrence.
- Comparing to disaggregated analysis, aggregated analysis is performed through using static and aggregate measures, which cannot accord to the concept that disruptive traffic contributes more to traffic crashes as opposed to normal traffic. Hence, has limited applicability to predict crashes on a real-time basis.

With the advent of ATMIS, much attention has been paid to incident detection and traffic management, and thus the proactive strategies are proposed. As one of the prerequisite of the strategies for traffic safety management, crash risk prediction on real-time basis by using high-resolution data has exhibited huge promise in the future.

However, as discussed before, the methods of existing crash prediction models essentially have their corresponding assumptions that should be satisfied at first. For example, if more complicate correlations exist between explanatory variables and the dependent response, it is not reliable to apply the logistical regression. Because that method just has statistical significance for linear interrelationship between the independent and predictive variables or the relationship in a single monotonicity at least. Therefore, some prior knowledge regarding the above interrelationships is required. Furthermore, for ensuring the reliability of statistics, the relativity of explanatory variables should be investigated in advance, with the purpose to select significant and uncorrelated variables for crash modeling.

Given the above problems, this study designs a methodology to develop a crash risk estimation mode considering the interaction among geometry, traffic flow and ambient conditions. This model is separately developed for urban and intercity expressways, while only basic segments of the two types of expressways are focused on, in view of the availability of datasets (will introduced in Chapter 3). Based on the available data including archived detector data and the variables of design consistency, the matched case-control logistic regression is adopted for modeling. To intensify the validity of this method, two types of proactive analyses are required in advance: 1) *CR* statistics to investigate the safety performance of traffic conditions preliminarily and 2) PCA for identifying the correlations of explanatory variables and their significance affecting crashes. If *CR* tendency appears in different monotonicities, traffic conditions should be categorized, and the model would be developed in piecewise functions.

2.6 Summary

In this chapter, a general view on the history of crash data analysis including aggregated and disaggregated approaches was presented. For better understanding advantage and disadvantage of two studies, the fundamental characteristics of crash data were described, their potential methodological problems were briefly analyzed. Since researchers have conducted a lot of studies to identify crash influencing factors, the characteristics of those factors and their correlations with crashes were reviewed in details. Then, a discussion about the existing aggregated and disaggregated approaches including their merits and demerits was performed. Based on these reviews, a methodology for crash prediction that seems to be feasible was finally proposed for this study.

Chapter 3

STUDY SITES AND DATA PROCESSING

3.1 Study Sites

The findings expected in this study will be applied to all expressways in general. Thus, despite the availability of data, the selected study sites should approximately represent most of expressways in Japan in terms of traffic characteristics.

3.1.1 Nagoya Urban Expressway network

Urban expressways are intra-city expressways which currently exist in the metropolises of Japan. Due to the lack of urban space, many of those expressways are constructed as viaducts running above urban streets. Nagoya, as the third-largest city in Japan, its urban expressway network, namely Nagoya Urban Expressway (NEX), is one of the relatively complete networks in Japan currently. Except some unique characteristics inherently exist for other networks, NEX is assumed to be an acceptable study site for this research.

NEX, serving the great Nagoya area including Nagoya metropolitan district and its outskirt areas, is managed by Nagoya Expressway Public Corporation. Up to December 31, 2009, eight routes with a total length of 69.2km have been completed. Along the mainline of the routes, more than 250 ultrasonic detectors install in approximately 500m intervals to count the number of vehicles passing through. Other than these routes, another route named by Toukai (No.4) is still under construction. The schematic map of NEX is shown in Figure 3.1, including the name/ID of each route and its length in use, and a site plan in the lower left demonstrates the connections of other routes to Inner ring (No. R).

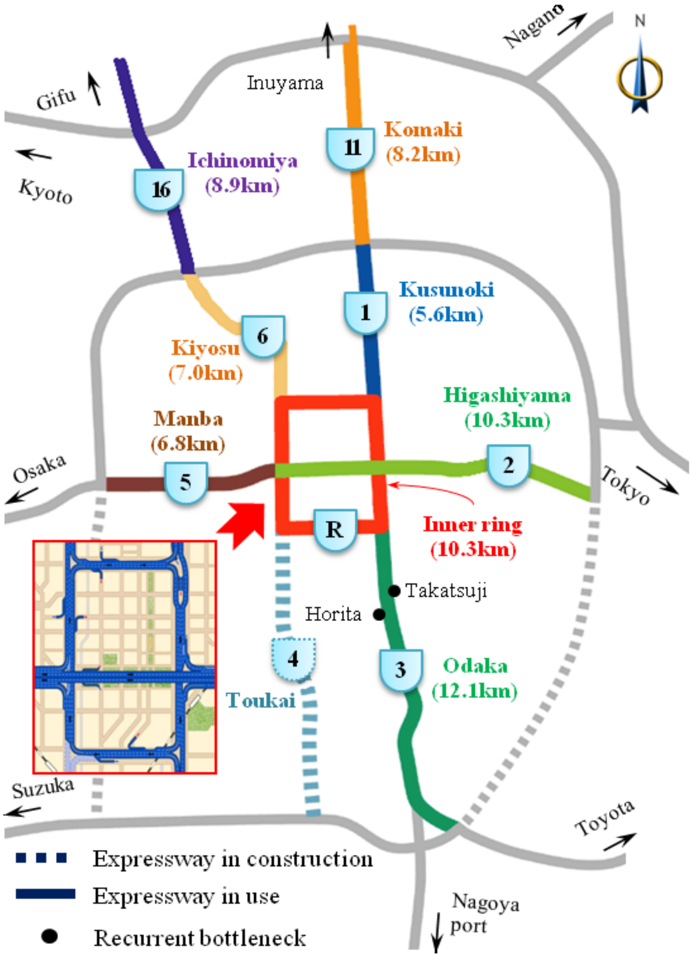


Figure 3.1 Schematic map of NEX (2009)
(Source: Nagoya Expressway Public Corporation, modified by author)

Most routes are 4-lane roadway (2-lane/direction) except for Inner Ring at the center of Nagoya, which is one-way, flowing clockwise and where the number of lane differs (2~5) with the change of ramp-junctions. Routes 1 through 6 with Route 4 absent extend radially from Inner Ring excluding Route 2 which bisects Inner Ring. In the sections, *e.g.,* the connections of other routes to Inner Ring and the sharp turns of Inner Ring, much small curves are adopted owing to the limited space. These routes above are designed at 60km/h and each lane on the mainline is wide in 3.25m. Comparatively, Route 11 and Route 16, which are in fact the extensions of Route 1 and Route 6, respectively, have a design speed of 80km/h and the corresponding lane-width is 3.5m. In this regard, there are three types of cross-section layouts in NEX, which are generally summarized in Table 3.1, and the standard cross-section layout of Route 1 through 6 is shown in Figure 3.2.

Figure 3.2 Standard cross-section layout of NEX (Route 1 through 6)

Figure 3.3 Standard cross-section layout on Tomei-Meishin Expressway

Source: http://fumi.ninja-x.jp/TOMEI6.html.

Table 3.1 Three types of cross-section layouts in NEX

ID of route	Cross-section layout*	Note
No.R	1.0m+3.25m×3+1.0m	Single-way roadway
No.11 and No.16	1.5m+3.5m×2+0.5m	Each direction for double-way roadway
Other routes	1.5m+3.25m×2+0.5m	Each direction for double-way roadway

* 1.0m(shoulder)+3.25m(lane)×3(number of lane)+1.0m(curbside).

3.1.2 Tomei and Meishin Expressways

Tomei Expressway is an important national expressway linking Tokyo and Nagoya, two major metropolises of Japan. Meishin Expressway actually extends Tomei Expressway to Osaka, another major metropolis, from Komaki junction near Nagoya. Tomei-Meishin Expressway is considered to typify intercity expressways. To better understand Japanese intercity expressway, the following vocabulary is provided (www.Japan-guide.com):

- Interchange (I.C): an interchange in expressway network refers to an expressway entrance and exit in the form of ramps.
- Junction (JCT): a junction in expressway network means the structure provided for the crossing among multiple routes each other.
- Parking Area (PA): a parking area comes with toilets and some vending machines, which may also feature a restaurant, sometimes.

- Service Area (SA): a service area is generally larger than parking area, which may provide toilet, shop, restaurant and a gasoline stand.

In view of the availability of data, the section of Tomei-Meishin Expressway from Mikkabi I.C to Yokaichi I.C is selected in this study, which is managed by Nagoya branch of Central Nippon Expressway Company Limited (NEXCO). The schematic map of the selected section is illustrated in Figure 3.4. The total length of this section is 183.6km including 95.7km for Tomei Expressway and 87.9km for Meishin Expressway, which are bounded at Komaki I.C. In this distance, more than 180 loop detectors are installed in approximately 2km intervals for both traffic directions.

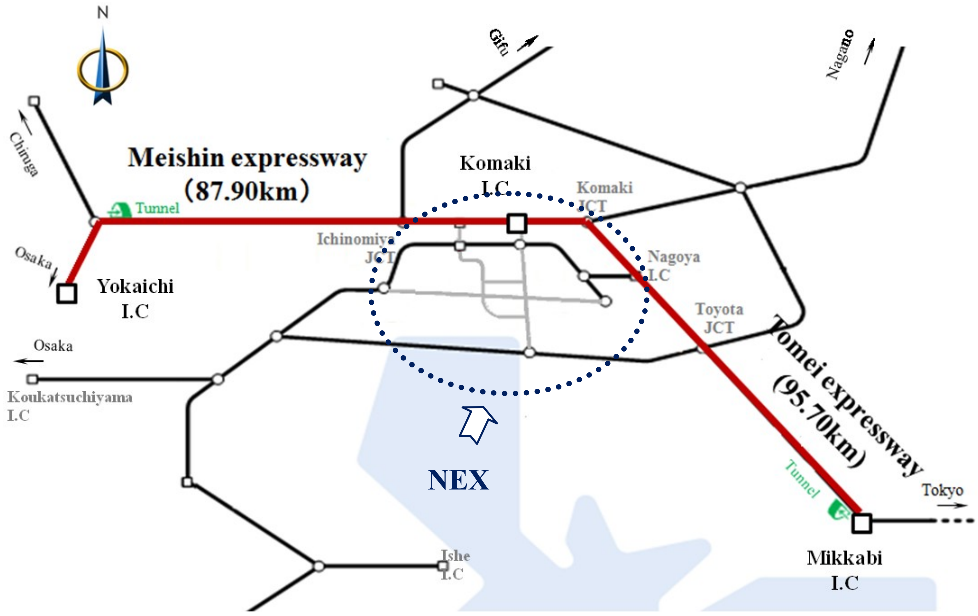


Figure 3.4 Schematic map of the study site on Tomei-Meishin Expressway (2009)

The alignment of this section is designed at varied speeds (80~120km/h) dependent on the terrain features, while the standard cross-section layout is unified: both directions originally carry 2 lanes separately with a width of 3.6m per lane. On the side of the outest lane, a hard shoulder in 3.0m wide is designed. In some areas, an additional lane with a width in 3.0m is constructed parallel to the existing lanes to serve the passing vehicles. The two types of cross-section layouts are summarized in Table 3.2, and the standard cross-section layout for one direction is shown in Figure 3.3.

Table 3.2 Two types of cross-section layouts on Tomei-Meishin Expressway

Types	Cross-section layout	Length (km)
2-lane roadway	3.0m+3.6m×2+0.75m*	179.2
3-lane roadway	1.5m+3.6m×2+3.0m+0.75m**	4.35

* 3.0m(shoulder)+3.6m(lane)×2(number of lane)+0.75m(curbside).
** 1.5m(shoulder)+3.6m(lane)×2(number of lane)+3.0m(additional lane)+0.75m(curbside).

3.1.3 Extraction of basic segments

Basic segments, as the focused facility type in this study, should be extracted from the above two networks separately. They are defined as the segments that are outside the influence of merging, diverging and weaving maneuvers (HCM, 2010).

In NEX, extraction of basic segments primarily considers the variation in traffic flow characteristics near to ramps. Referring to the related studies (Chen *et al.,* 2008; Zhou *et al.,* 2010; Wu *et al.,* 2013a, 2013b), basic segments are extracted outside 500m up- and downstream of ramp junctions, in consideration of the average space intervals between two neighboring detectors in 500m, as illustrated in Figure 3.5.

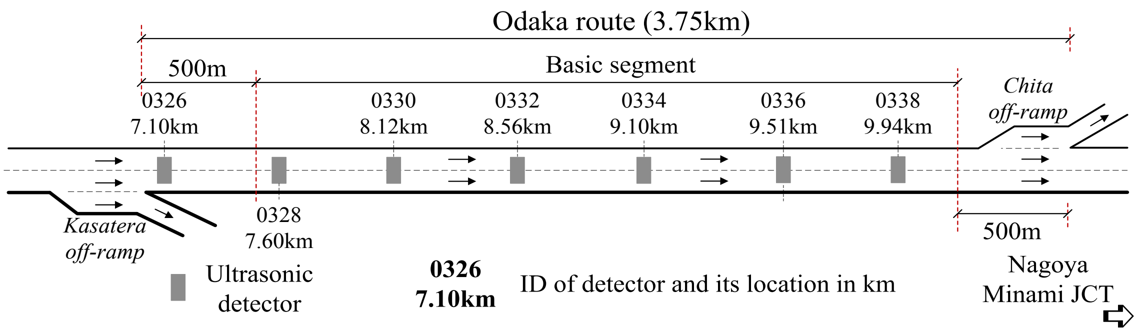


Figure 3.5 Example of basic segment extraction in NEX

As stated before, some much small curves are designed for the limited space of NEX. To ensure the reliability of extraction, this study also tentatively investigates *CR* distribution dependent on radius (the equation of *CR* will be introduced in Chapter 4), as represented in Figure 3.6. For the curve sections whose radii are larger than 100m, we cannot see a clear regular *CR* distribution to radius. By contrast, when the radii are below 100m, much higher *CR* can be observed, which may imply the distinct crash characteristics in those sections relative to other ones. Since the available small curve samples in NEX are limited, these sections are excluded in this study considering the validity of the expected crash model.

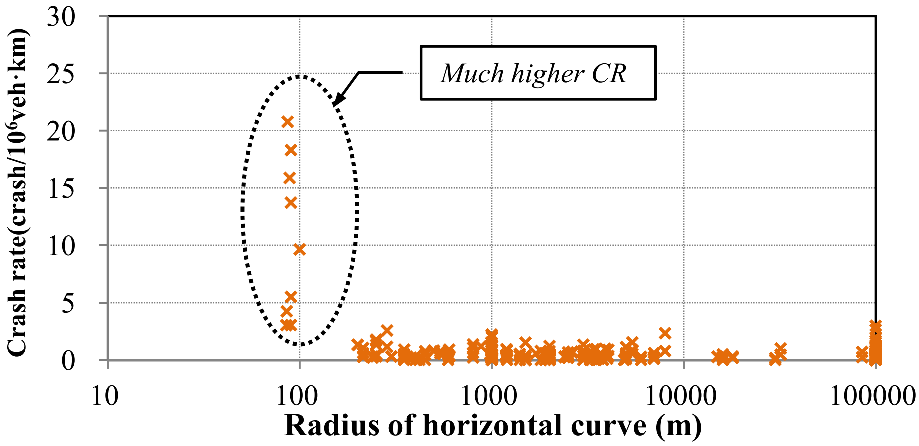


Figure 3.6 *CR* distributions dependent on radius of NEX

On Tomei-Meishin Expressway, the space interval of neighboring detectors is around 2km, which is 3 times longer than that in NEX. In this regard, basic segments are extracted in terms of the relative reliability of two neighboring detectors to represent the traffic characteristics near to the two detector locations (Wu *et al.,* 2003a). However, the exact boundary to differentiate that relativity is actually difficult based on the data in this study. In such case, the midpoints between two neighboring detectors are regarded as the boundary. Thus, the two detectors nearest to interchange/junction in both up- and downstream are focused on, and the midpoints between the two detectors are considered as the boundary of basic segments outside the influence areas of interchanges and junctions. One case about this segmentation method is shown in Figure 3.7.

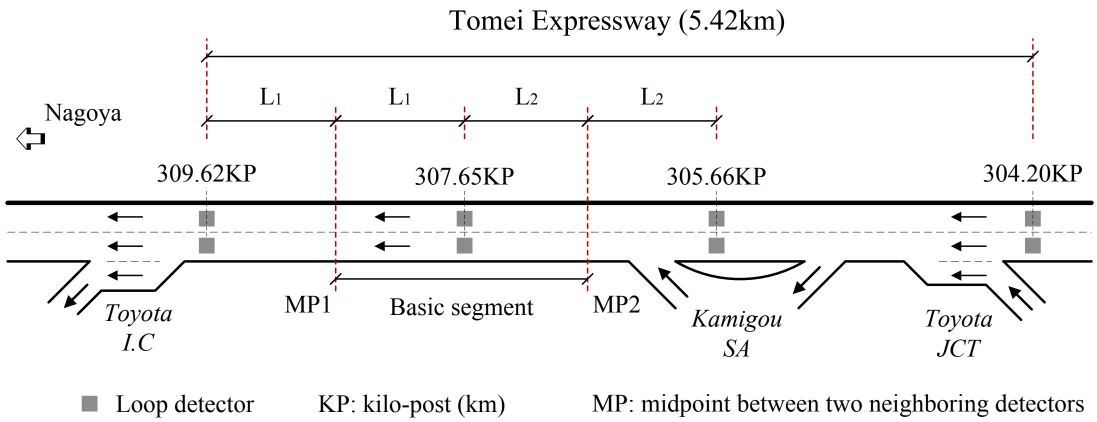


Figure 3.7 Example of basic segment extraction on Tomei-Meishin Expressway

On Tomei-Meishin Expressway, *CR* distribution dependent on radius is also identified as described in Figure 3.8. However, *CR* seems to be characterized by random distribution, and no specific characteristics of *CR* can be virtually observed from the study site.

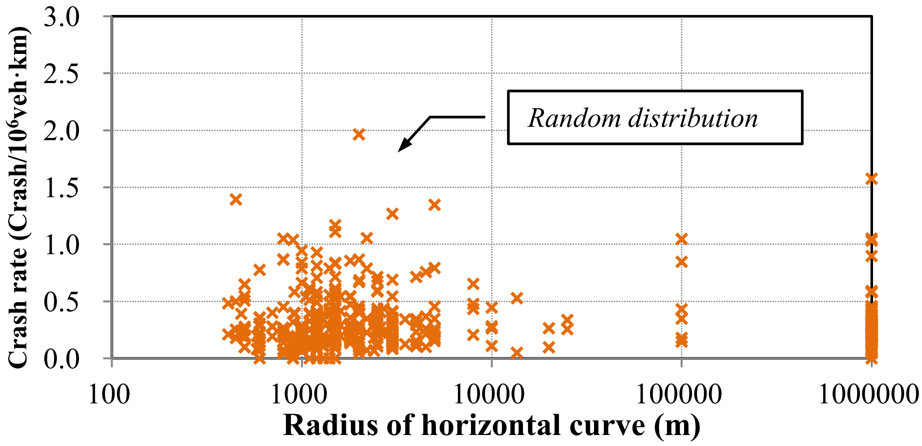


Figure 3.8 *CR* distributions dependent on radius of Tomei-Meishin Expressway

A total length of 57.02km and 164.60km basic segments can be successfully collected from NEX and the study sites of Tomei-Meishin Expressway, respectively. Table 3.3 summarizes the statistical results of those segments by cross-section layout.

Table 3.3 Length of basic segments by the type of cross-section layout

Expressway type	Basic segment by type of cross-section	Length (km)*
NEX	Inner Ring (No.R)	0.39
	Komaki (No.11) and Ichinomiya (No.16)	11.96
	Other routes	44.67
Tomei-Meishin	2-lane roadway	154.90
	3-lane roadway	9.70

* Length in two directions while it is in single direction for the segments of Inner Ring.

3.1.4 Statistics of crash events at basic segments

Other than basic segments, merge/diverge/weaving segments and others can be divided through referring to the methods introduced in Wu *et al.,* (2012a, 2012b). By facility type, the percentage of crash events taking out of the total value in whole network is highest at basic segments. As shown in Figure 3.9 and 3.10, for NEX and Tomei-Meishin Expressway, the values are 31.9% and 38.3%, respectively. Meanwhile, focusing on basic segments, it can be found that the dominant types of crashes for urban expressway (NEX) are vehicle-facility sideswipe and rear-end collision. Comparatively, for intercity expressway (Tomei-Meishin), the predominant one is rear-end. Such differences may be associated with the different geometric design and traffic characteristics by expressway type.

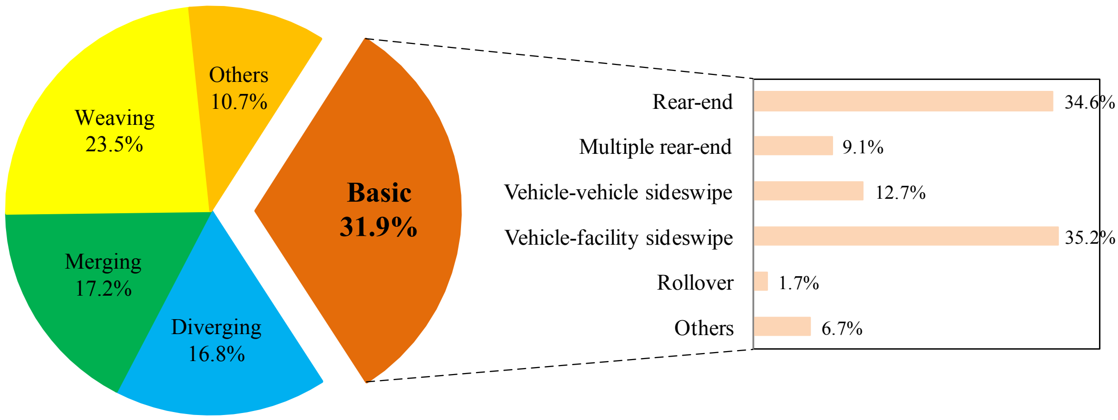


Figure 3.9 Statistics of crash events at basic segment for NEX

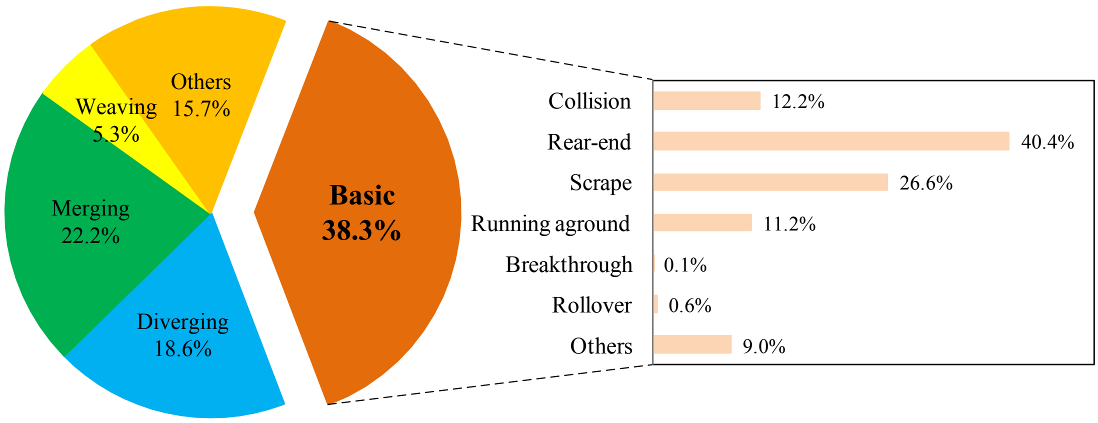


Figure 3.10 Statistics of crash events at basic segment for Tomei-Meishin Expressway

3.2 Original Databases

Four databases are utilized in this study, which are kindly provided by Nagoya Expressway Public Corporation and Central Nippon Expressway Company Limited for NEX and Tomei-Meishin Expressway, respectively. The period is over three years (2007-2009) except for those on Kiyosu (Route No.6) in NEX that opened from December 1st, 2007.

- Crash records. They are recorded by road administrators after crash occurrence with the time in minute and the location in the unit of 0.1km. Meanwhile, weather and pavement conditions corresponding to each crash record are available. The type of crash has been classified in the two crash records, while the standards of classification are different for NEX and Tomei-Meishin Expressway.

- Detector data. Traffic volume q, average speed v and occupancy occ can be automatically processed per 5 minutes from the detectors. Between the two types of expressways, the data are recorded in different ways due to the type of detector: in NEX, the data are cross-section based; by contrast, they are processed on each lane for Tomei-Meishin Expressway and include some other information about vehicle composition, *i.e.* heavy vehicle flow and its related percentage taking out of the whole traffic volume (abbreviated to HV %) every 5 minutes.

- Geometric design and detector locations. The databases mainly provide the design information of road alignment. Along the mainline, ID of each detector and its related location in the unit of 0.01km are provided.

- Traffic regulation records for incidents (*e.g.,* crash, working, inclement weather) including the locations and periods of temporal lane/cross-section closures can be found out from the datasets.

Other than these data above, the daily sunrise/sunset time in Nagoya can be referred from the website of GAISMA. The records of dawn, sunrise, sunset and dusk time are collected, towards reflecting the variation in ambient light conditions.

3.3 Data Collection and Processing

3.3.1 Detector data

In principle, detectors can count the number of vehicles passing through their locations only. For estimating traffic conditions at the locations by detector data, the "coverage area" of detector should be defined. The boundary of consecutive coverage area is defined at the midpoint between two neighboring detectors. Note that, the time of crash in the study dataset is not the exact occurrence time, since it was 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. In this case, the latest data at least 5 minutes before the recorded time are accepted in this study. In view of the reliability of detector data, the invalid data and the data within lane and cross-section closure intervals are excluded in advance from the datasets.

Detector data on Tomei-Meishin Expressway is lane-based, while they are cross section-based in NEX. Considering the comparison by expressway type, this study coverts the lane-based data into cross section-based through the following three equations.

$$q_{Si} = \sum q_{Li} \tag{3.1}$$

$$v_{Si} = \frac{\sum q_{Li} \times v_{Li}}{\sum q_{Li}} \tag{3.2}$$

$$k_{Si} = \frac{12 \times q_{Si}}{v_{Si}} \tag{3.3}$$

Where, q_{Li}, v_{Li} separately denote flow rate and speed on individual lanes in 5 minutes of ID i; q_{Si}, v_{Si} and k_{Si} are the converted flow rate, speed and traffic density for the whole cross section (one direction), respectively.

3.3.2 Geometric features

Crash models explicitly considering geometric design consistency may identify the driver-vehicle-roadway interaction more effectively and reflect the impacts of geometry on safety more accurately than those that did not. Since the driver-vehicle-roadway interaction may be varied by locations, this study employs the concept of geometric variation along traffic direction to represent design consistency (Hikosaka and Nakamura, 2001). In view of the length of detector coverage area, the following geometric variables in 500m upstream from crash locations are proposed in this study.

- Variation in road elevation h between crash locations and their 500m upstream locations, and the maximal elevation H during those 500m upstream distance. The detailed definitions are explained in Figure 3.11.

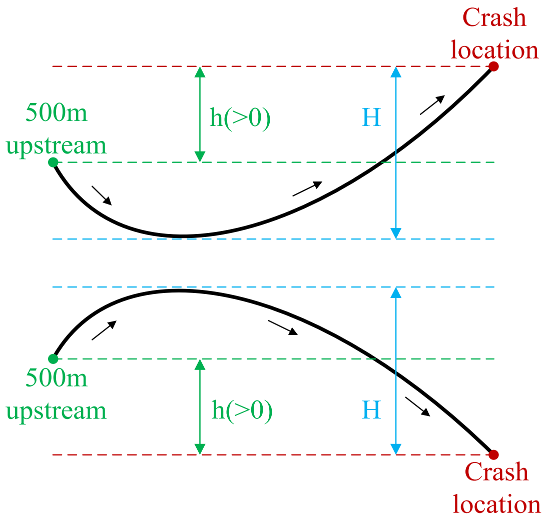


Figure 3.11 Variation in road elevation

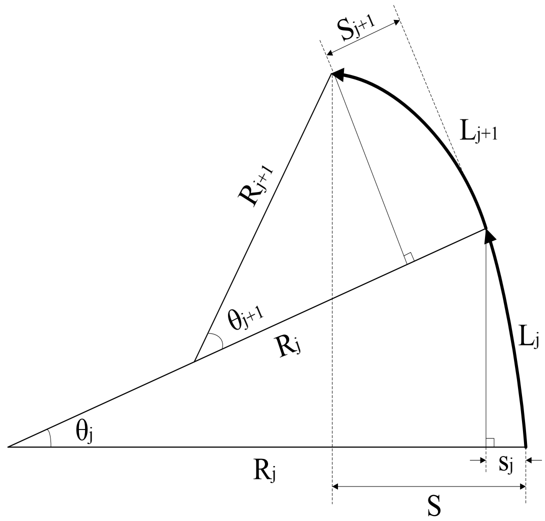


Figure 3.12 Horizontal displacement

- Horizontal displacement *S*. Radius is impossible to describe the variation of one section composed of different curves. Furthermore, the centrifugal force is essentially associated with the horizontal displacement *S* in the direction of tangent to curves, as indicated in Figure 3.12. Besides, the value of *S* can also reflect the combined effects of radius and length of individual curves. Instead, *S* in the 500m distance is adopted and calculated by the following equations.

$$\theta_j = \frac{L_j}{R_j} \qquad (0<\theta_j<\frac{\pi}{2}) \tag{3.4}$$

$$s_j = R_j(1-\cos\theta_j) \tag{3.5}$$

$$S = \sum s_j \tag{3.6}$$

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

- Index of centrifugal force I_{CF}. Speed v always has a square relation with centrifugal force. In support of this concept, I_{CF} calculated by formula (3.7) is designed for reflecting the combined effect of speed v along with horizontal displacement *S*, while it is not the correct value of centrifugal force.

$$I_{CF} = S \times v^2 \tag{3.7}$$

$$I_{SD} = S \times H \tag{3.8}$$

- Index of space displacement I_{SD}. I_{SD} is utilized to reveal the comprehensive geometric features induced by the horizontal and vertical variation in this study.

The geometric data above are collected every 0.1km as crash is recorded in a unit of 0.1km. Meanwhile, the data are also extracted at the location of detector, since it is the common link between crash and detector data. Table 3.4 exemplifies the process of geometric variations collection for individual crashes and detector locations.

Table 3.4 Examples of geometric variation collection

Route #	Direction*	KP**(km)	*h*(m)	*H*(m)	*S*(m)	I_{SD}(m^2)	Note
1	SB	0.0	-4.63	5.49	0.78	4.30	
1	SB	0.1	-7.90	8.49	3.91	33.2	
1	SB	0.2	-10.6	11.5	6.08	69.9	
1	SB	0.21	-11.5	11.8	8.88	104.7	Detector #0101
1	SB	0.3	-15.3	15.3	9.60	146.9	
1	SB	…	…	…	…	…	
1	SB	6.4	10.2	10.9	5.15	56.1	
1	NB	0.0	0.53	0.98	12.6	12.3	
1	NB	0.1	1.05	1.28	24.5	31.4	
1	NB	0.11	0.75	0.94	24.7	23.2	Detector #0100
1	NB	…	…	…	…	…	
1	NB	5.7	0.66	5.09	15.8	80.4	
1	NB	5.74	-0.87	5.21	15.8	82.3	Detector #0124
2	NB	0.0	1.25	1.63	4.55	7.42	
…	…	…	…	…	…	…	
16	NB	8.1	-1.31	1.40	17.0	23.8	

*SB-southbound, NB-northbound.
**KP-kilopost (the same in the following).

3.3.3 Ambient conditions

Commonly prevailing and uncontrolled environment/weather conditions are defined to be ambient conditions. In this study, the following conditions are available.

- Ambient light. With the alternation of sunrise and sunset, a day can be divided into daytime, twilight and nighttime in general. Daytime and nighttime correspond to the periods from sunrise to sunset and from dusk to dawn, respectively. Other periods, as the interim between daytime and nighttime, are defined as twilight. Actually, the period of twilight is very short and its brightness is losing while not zero. Thereby, twilight is combined into daytime, and thus ambient light can be classified into nighttime (N) and daytime (D), as demonstrated in Figure 3.13.

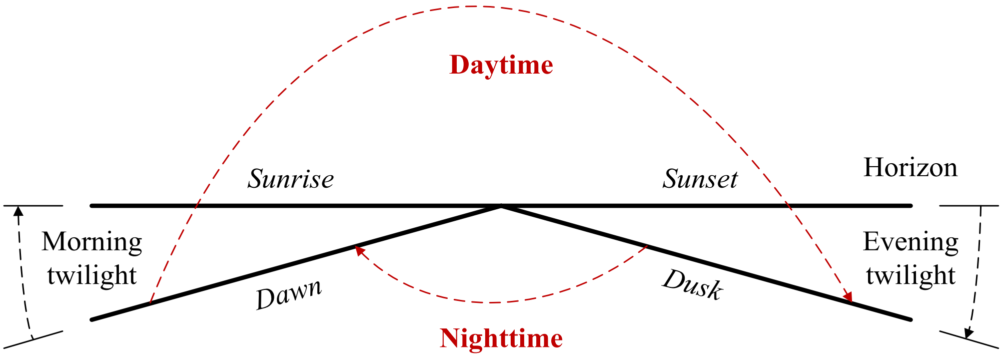


Figure 3.13 Classification of ambient light conditions
(Source from Wikipedia and modified by authors)

- Weather/pavement conditions. According to the original crash records, weather condition for individual crashes can be known as sunny (SU), cloudy (C), rainy (R) or snowy (SN). The dataset also identified the pavement conditions, *i.e.* dry (D) and wet (W) at crash locations, which are highly related to weather conditions. Such as on rainy days, not only the visibility is constrained, but wet pavement is induced, which reduces the rolling tire-pavement frictions.

- Day type. As stated in Chapter 2, holiday traffic often has different characteristics relative to normal weekdays. To allow for this situation, this study makes a distinction between holiday (H) and weekday (W). Here, holiday includes all weekends, national holidays and days during the Golden Week in May and the Obon Week in August. Correspondingly, the other days are defined as weekday.

3.4 Data Matching Based on Individual Crash Records

The related detector data, geometric variations and ambient conditions for individual crashes are matched, as exemplified in Table 3.5. The crashes matched with invalid detector data and within traffic regulation period such as lane and cross-section closure intervals are also excluded in advance. As a result, a total of 457 and 1496 crash records remain for NEX and the study sites of Tomei-Meishin Expressway, respectively.

Table 3.5 Examples of data matching for individual crash records

Crash ID	Detector data			Geometric design				Ambient condition*			
	q_{Si} (veh/5min)	v_{Si} (km/h)	k_{Si} (veh/km)	h (m)	H (m)	I_{CF} (km^3/h^2)	I_{SD} (m^2)	L	W	P	D
1	58	86.4	8	4.50	4.85	57.4	37.3	N	SU	D	H
2	267	38.6	83	-1.55	3.69	4.93	12.2	D	R	D	W
3	2	50.4	1	1.64	2.14	87.3	73.7	N	C	D	W
4	60	77.0	9	-4.44	4.44	61.9	46.4	N	SU	W	H
5	86	84.6	12	0.930	3.86	558.3	300.8	D	SU	D	W
6	6	85.7	1	0.786	1.14	48.8	7.59	N	R	W	H
7	266	60.6	53	-1.28	2.79	160.5	122.2	D	C	W	W
8	249	73.2	41	1.50	1.50	2.70	0.757	D	C	D	H
9	62	56.8	13	-1.28	2.79	140.9	122.2	D	C	D	H
10	213	36.2	71	0.819	2.23	3.89	6.62	D	SU	D	W

*L/W/P/D separately refers to light, weather and pavement conditions as well as day type.

3.5 Geometric Features and Traffic Characteristics by Expressway Type

One purpose of this study is to identify the different mechanisms of crash influencing factors between urban and intercity expressways, which may be caused by the variation in geometric design and traffic characteristics by expressway type. In view of this objective, this section tentatively identifies the differences of traffic characteristics and geometric design between NEX and Tomei-Meishin Expressway.

3.5.1 Geometric design

As demonstrated in section 3.1, NEX is generally designed with smaller curve radii, narrower cross section and higher access density compared to Tomei-Meishin Expressway.

Nevertheless, the vertical slope along the mainline of NEX is designed more gently owing to the structure of viaducts: its maximum gradient is 2.8% as opposed to the value of 5.0% on the study sites of Tomei-Meishin Expressway. Besides, by comparing Figure 3.2 and 3.4, it can be found that there is much high roadside barrier on urban expressway, which implies that the visibility may be restricted more seriously while driving in curve sections. The geometric features by expressway type are summarized in Table 3.6.

Table 3.6 Different features of geometric design by expressway type

Items	NEX		Tomei-Meishin
Design speed*	60km/h	80km/h	80~120km/h
Standard cross section**	1.5m+3.25m×2+0.5m	1.5m+3.5m×2+0.5m	3.0m+3.6m×2+0.75m
Alignment design***	R_{min}=200m i_{max}=±2.8%	R_{min}=1000m i_{max}=±2.0%	R_{min}=410m i_{max}=±5.0%
Roadside barrier	Higher		-

* 80km/h is for Komaki (Route No. 11) and Ichinomiya (Route No.16) only.
** 1.5m(shoulder)+3.25m(lane-width)×2(number of lane)+0.5m (curbside).
*** R_{min} means the minimum value of radius; i_{max} refers to the maximum value of slope.

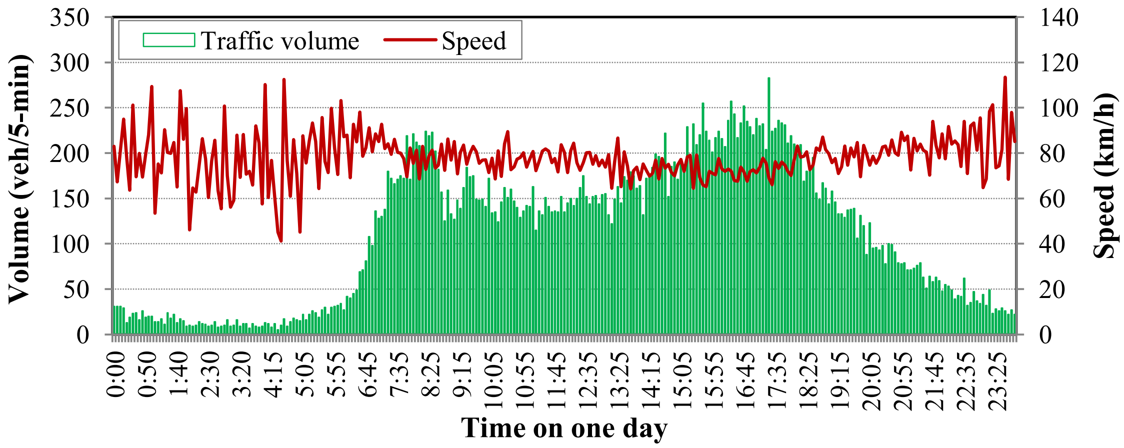
3.5.2 Traffic characteristics

1) Vehicle composition

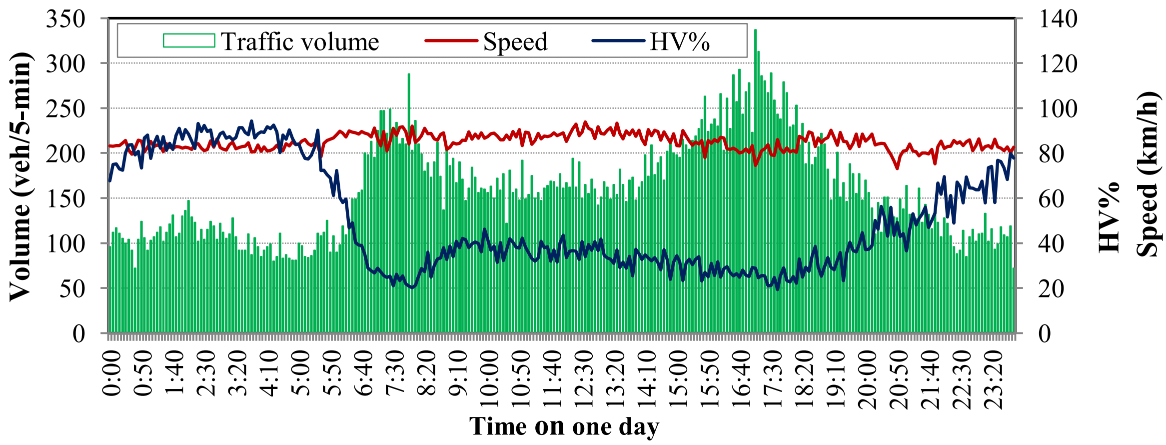
Tomei-Meishin Expressway is a main transportation link between Tokyo and Osaka, two major metropolis of Japan. In contrast, NEX bears much intra-city transportation that is composed of high population of commuters. Hence, it is considered reliable to assume that Tomei-Meishin Expressway carries much more long-distance trips (abbreviated to LDT). Meanwhile, although the data of vehicle composition is not available from detector data of NEX, this information based on monthly flow can be referred from the website of NEX. It reveals that HV% in NEX is generally around 3%~6%, much lower than the value on Tomei-Meishin Expressway, which is often over 20%.

2) Time serious variation in traffic flow

Based on detector data, the time-serious variation in speed and traffic volume are analyzed for the two types of expressways. The measured detectors are selected far from small curve sections, to avoid the influence of that geometric design. In this way, the detector #0339 and the detector at 296.44km are employed for NEX and Tomei-Meishin Expressway, respectively, since the daily traffic at the two detectors is large and frequent traffic crashes happen near these detectors. Besides, the distance of each detector from the nearest up- and downstream ramps is at least 1.4km. The variations in traffic flow on one weekday (July 1st, 2009, Wednesday) and one holiday (July 26th, 2009, Sunday) are shown in Figure 3.14 and 3.15, respectively. Weather/pavement conditions are just available for crash events, and thus the related analysis on those conditions is not performed in this study.



a) Odaka route (detector #0339)



b) Tomei Expressway (detector at 296.44KP)

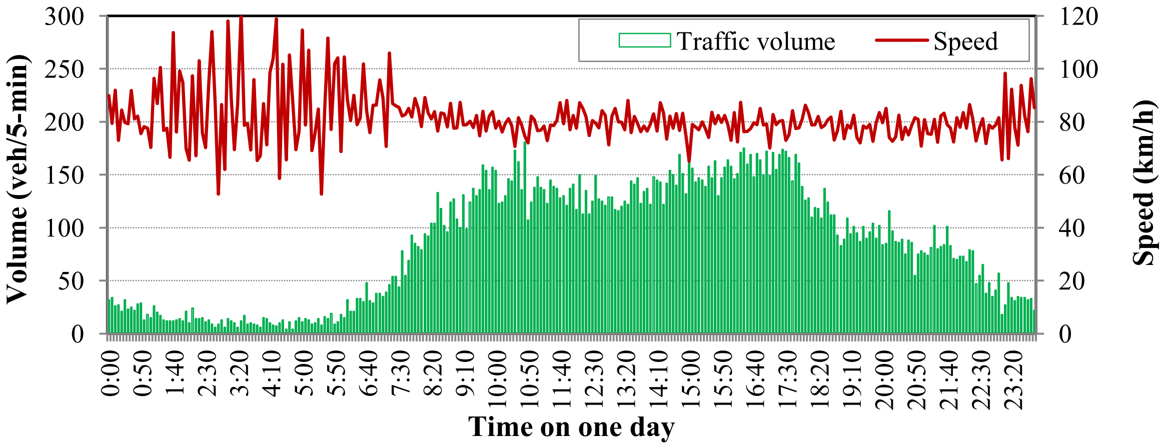
Figure 3.14 Time-serious variation in traffic flow on one weekday

The data of HV% is available on Tomei Expressway, which is one important influencing factor for traffic characteristics (Al-Kaisy and Jung, 2004). Its time-serious variation is thereby also analyzed for the following discussions. Through referring to Figure 3.12, the difference of weekday traffic by expressway type can be observed as follows:

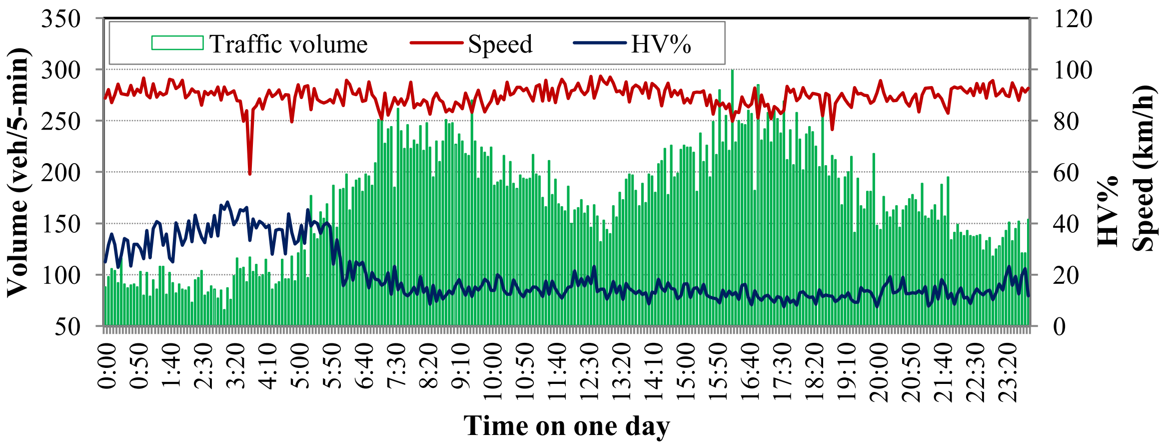
- Tomei Expressway carries more traffic volume compared to NEX.

- The speed on cross-section basis of Tomei Expressway mostly exceeds 80km/h. By contrast, the speed on Odaka route is primarily under 80km/h.
- Compared to Tomei Expressway, a higher speed variation exists on Odaka route during nighttime (20:00-24:00 and 0:00-6:00).

Speed variance of nighttime traffic on Odaka route is much higher in contrast to Tomei Expressway, which may be induced by the low traffic volume. Another potential cause is that the reduced visibility in nighttime would enforce some cautious drivers take slow speed while driving on narrower cross sections, which may aggravate the discretionary of driving conditions. If focus on Tomei Expressway, it is clear that much higher HV% exists in nighttime, often more than 60% as opposed to lower 40% in daytime.



a) Odaka route (detector #0339)



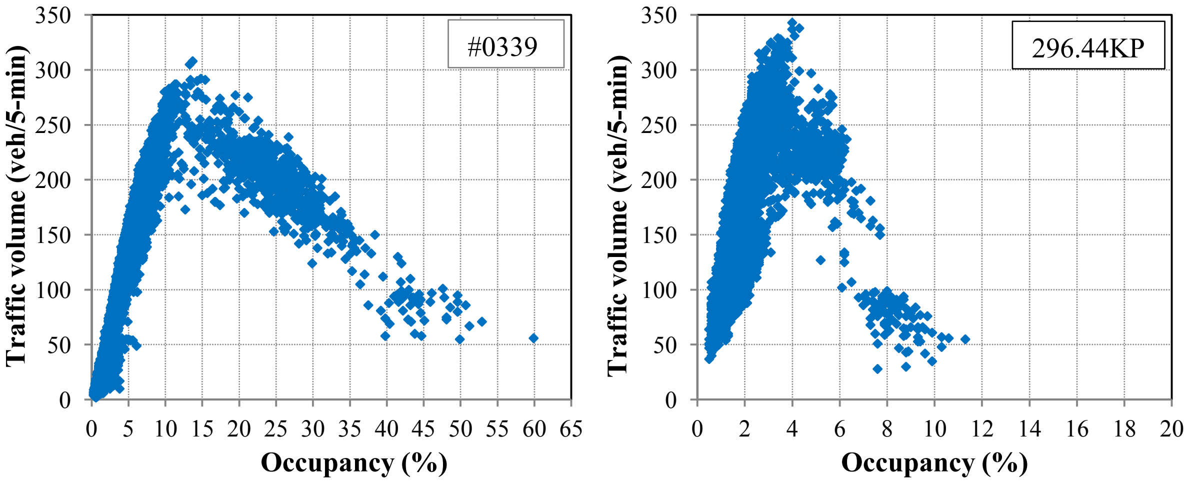
b) Tomei Expressway (detector at 296.44KP)

Figure 3.15 Time-serious variation in traffic flow on one holiday

With the alteration from weekday to holiday, traffic volume on Odaka route decreases significantly (see Figure 3.14(a) and 3.15(a)). Furthermore, the operating speed on holiday can be found out to be higher in contrast to weekday. Through comparing Figure 3.14(b) and 3.15(b), the maximum value of traffic volume on holiday is not distinctly reduced on Tomei Expressway. Instead, the period of peak hours last longer on holiday in contrast to weekday. Meanwhile, the HV% on holiday is much lower than the value on weekday. In view of this regard, the above phenomenon of traffic volume may reflect the increase of recreational private travel, which is mostly composed of cars or other small-sized vehicles. Furthermore, due to the reduced HV%, speed on holiday of Tomei Expressway is mostly around 90km/h, slightly higher than that on weekday. By expressway type, smaller traffic volume, lower speed and more serious speed variance of nighttime traffic still appear on Odaka route comparing to Tomei Expressway on holiday.

3) Interrelationships of traffic variables

Traffic density has been proposed as the service measure of traffic flow for basic segments in some literatures (HCM, 2010). Fortunately, occupancy as a dimensionless measure of density is available in the original datasets. Based on the data in July, 2009, the diagrams of traffic volume-occupancy are further analyzed at both detectors with the purpose to investigate the interrelationships of traffic variables, as indicated in Figure 3.16.



a) Odaka Route **b) Tomei Expressway**

Figure 3.16 Traffic volume-occupancy diagrams by expressway type

The diagram of intercity expressway seems more acute comparing to urban expressway: with the increase in occupancy, traffic volume of Tomei Expressway increases more rapidly to a higher value than that of Odaka Route. Then, as occupancy further increases, traffic volume of Tomei Expressway still reduces more rapidly. As a reflection, intercity expressway seems to be unable to contain congested traffic as heavy as urban expressway. Combined with the analyses on vehicle composition and time-serious variation in traffic flow, traffic characteristics by expressway type are summarized in Table 3.7.

Table 3.7 Different traffic characteristics by expressway type

Items	Odaka route	Tomei Expressway*
Vehicle composition	HV% is 3%~6%	HV% is over 20%
Driver familiarity	More commuters	More LDT
Speed limit	60 km/h	80-100km/h
q-occ diagram	Heavier congested flow	More rapid variation in traffic volume with the increase in occupancy
Traffic in nighttime in contrast to daytime	Higher speed variance; Lower traffic demand.	Lower traffic demand; Rise in HV%.
Traffic on holiday in contrast to weekday	Lower traffic demand; Rise in operating speed.	Longer duration of peak hours; Rise in operating speed; Obvious increase in recreational private travel; Decrease in HV%.

* HV%: the percentage of heavy vehicle, LDT: long-distance trip.

The vehicle composition on Tomei Expressway is characterized in higher HV% relative to NEX. On the increase of traffic density, the inter-vehicle interaction gets more intensive. For the sake of safety, the upstream vehicles would take some risk-avoiding behaviors, *e.g.,* slow-down and lane changing, in an effort to keep sufficient spacing from the downstream HV. These behaviors may affect the average speed of traffic streams. By contrast, the speed in the same traffic stream or during various lanes may be easier to get harmonization on urban expressway, and finally it could contain more heavy congested flow. Since the low-density traffic often exists during nighttime, as demonstrated before, the traffic on Odaka route is of higher speed variance. As Tomei Expressway is a main transportation link, even in nighttime, the minimum value of traffic density is not seriously low.

3.6 Summary

The suitable selection of study sites and the appropriate explanatory variables collected for crash analysis are the basic requirements for crash modeling. Hence, this chapter firstly introduced the prepared study sites and their original datasets available for this study. Based on these datasets, the measures to collect geometric features, traffic flow and ambient conditions for individual crashes are then discussed. One purpose of this study is to identify the affecting mechanisms of explanatory variables on crashes by expressway type considering the interaction of geometry, traffic flow and ambient conditions. Thus, as a preliminary analysis, the differences of geometric features and traffic characteristics between NEX and Tomei-Meishin Expressway are finally measured.

Chapter 4

CRASH RISK ESTIMATION MODEL FOR URBAN EXPRESSWAY

4.1 Modeling Flowchart

As discussed in Chapter 2, most existing statistical approaches can reasonably investigate the correlations between explanatory variables and responses on linear or monotonic basis. Meanwhile, for the reliability of statistics, it is necessary to examine the significances and interrelations of individual explanatory variables in advance. In essence, crash occurrence is a rare event, and has binary outcomes (crash *vs.* non-crash). Furthermore, a crash event may be associated with a variety of factors, such as geometry, traffic flow and ambient conditions. Given the above situations, the following methods are adopted for quantifying the effects of various influencing factors on crash occurrence. The way how these methods correlate with each other is demonstrated in Figure 4.1. The theories of these methods are introduced in detailed in their corresponding sections.

- *CR* tendency with the increase in traffic density, in the purpose to categorize traffic conditions in terms of safety performance.
- By focusing on the categories of traffic conditions, PCA is utilized for identifying the interrelation of explanatory variables and their significance affecting crashes.
- A case-control study is design based on the significant and independent variables in an effort to investigate the exposures which are related to crash occurrence.
- According to the case-control samples, conditional logistic regression is applied for quantifying the relationship between crash risk and the selected variables.

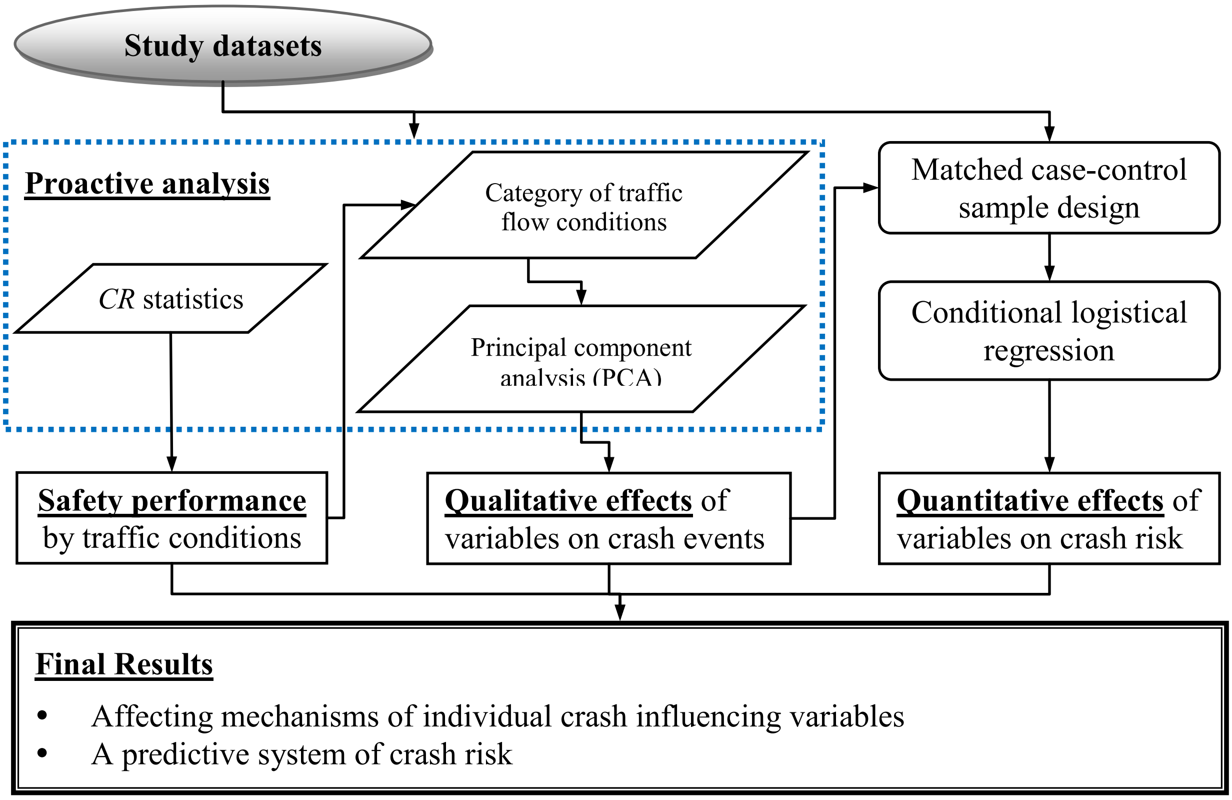


Figure 4.1 Flowchart of crash risk estimation model

In theory, safety performance of geometric features and ambient conditions should also be investigated in advance. However, due to the insufficiency of data collection, it is hard to detect a consecutive variation in *CR* tendency along with geometry and ambient conditions. Furthermore, as introduced in Chapter 2, *CR* tendency with the variation in the hierarchy of curve or slope has been identified in numerous studies. Generally, nighttime/wet pavement/holiday can aggravate crash risk compared to daytime/dry pavement/weekday, respectively. Hence, only *CR* tendency considering the variation in traffic conditions is investigated in Section 4.2. In terms of the monotonicity of *CR* tendency, traffic conditions are hereby categorized in Section 4.3, and then PCA is adopted through focusing on the categories of traffic conditions to examine the interrelation of explanatory factors and their related significances. Section 4.4 demonstrates the process of crash risk estimation model (CREM) based on the selected significant and independent variables. Next, Section 4.5 analyzes the contributions of independent variables to crash risk, and finally a brief summary of this chapter and motivation for this thesis is given in Section 4.6.

4.2 *CR* Statistics

In view of the purpose of *CR* statistics, traffic flow is firstly divided into several conditions to describe the variation of traffic characteristics.

4.2.1 Classification of traffic conditions

Congested flow, as a typical oscillated traffic, has distinct traffic characteristics from uncongested flow (Zheng, *et al.,* 2010). For the two traffic conditions, driving conditions and their impacts on crashes may be different, while crash characteristics between un- and congested flow are yet to be well identified through previous studies. Thus, it is necessary to make a distinction between the two traffic conditions.

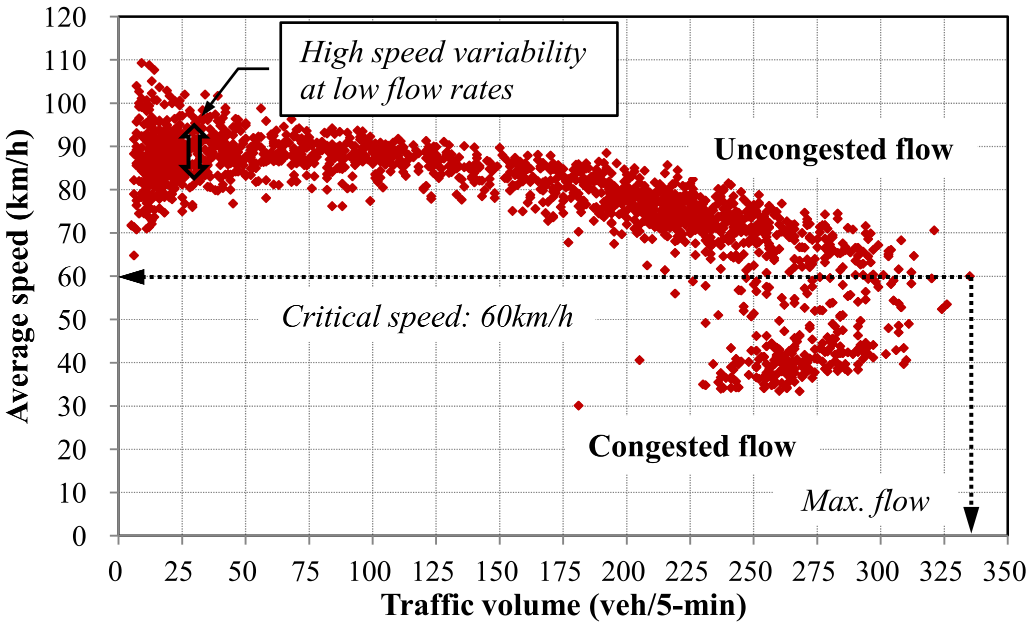


Figure 4.2 Traffic flow-speed diagram at Horita on-ramp junction

For classifying un- and congested flows, the diagram of traffic flow-speed is analyzed at a typical bottleneck, Horita on-ramp junction in Odaka route (No.3 in NEX), as shown in Figure 4.2. Corresponding to the maximum flow rate, the boundary of speed between the two traffic conditions can be observed as 60km/h, which is defined to be critical speed (Shawky and Nakamura, 2007). Since no bottleneck can be virtually observed at basic segments, the speed of 60km/h is generally regarded as the corresponding index at basic segments of urban expressway in this study.

Towards reflecting the variation in traffic characteristics, traffic conditions are further sub-classified into a serious of groups. In Figure 4.2, it is clear that speed has a high variability at low flow rates. Meanwhile, occupancy is not a commonly used traffic index, and *v/c* (traffic volume to capacity) ratio is generally not reliable for congested flow. In this case, traffic density k_S calculated by Formula (3.3) is proposed as the measure of effectiveness for sub-classifying 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, as summarized in Table 4.1.

Table 4.1 Classification of traffic conditions

ID of traffic conditions	**Uncongested flow**		**Congested flow**	
	Intervals of k_s (veh/km)	Median value (veh/km)	Intervals of k_s (veh/km)	Median value (veh/km)
1	0~5	2.5	<60	45
2	5~10	7.5	60~90	75
3	10~15	12.5	90~120	105
4	15~20	17.5	120~150	135
5	20~25	22.5	150~180	165
6	25~30	27.5	>180	195
7	30~35	32.5		
8	35~40	37.5		
9	40~45	42.5		
10	45~50	47.5		
11	>50	52.5		

4.2.2 Calculation of *CR*

CR as the result of crash frequency per million vehicle kilometer traveled (VKMT) for traffic condition #*i* is calculated by the following formula.

$$CR_i = \frac{NOC_i \times 10^6}{\sum Q_{il} L_l} \tag{4.1}$$

Where, *i* and *l* denote the ID of traffic conditions and detector coverage area, respectively. NOC_i refers to number of crashes for traffic condition #*i*. Q_{il} and L_l separately correspond to total traffic volume in detector coverage area #*l* for traffic condition #*i* and the length of detector coverage area #*l*.

4.2.3 *CR* tendency

1) The whole traffic conditions

For un- and congested flows, 319 crashes and 138 crashes are available for statistics, and Figure 4.3 gives the distributions of total crash rate *CR* following traffic density k_S. In general, *CR* is convex downward to k_S in uncongested flow and on the rise with the increase of k_S in congested flow. The quadratic function and exponential function can generally be fitted to the tendency of *CR* in the two traffic conditions, respectively. The figure also points out that congested flow has exceedingly higher *CR*, more than 60 times of the max *CR* in congested flow than the min *CR* in uncongested flow.

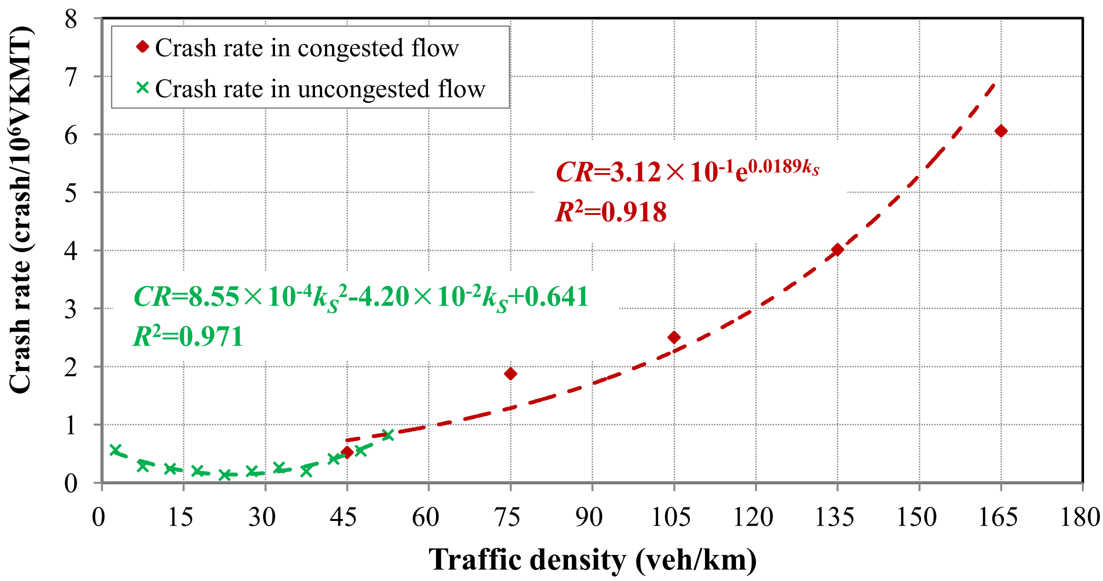


Figure 4.3 Total *CR* regression model (urban expressway)

2) Uncongested flow

Due to the small scale, the above figure is not clear enough to explain the characteristics of *CR*. In the following, uncongested flow is focused on and its *CR* tendency is further show in Figure 4.4. Meanwhile, single- and multiple-vehicle crashes are regarded to be two different crash modes, since conditions preceding crashes are identified to differ by type of crash (Christoforou *et al.,* 2011). This study also makes a distinction between two crash modes in that such classification can generally reveal the different effects of traffic flow on crashes along with the variation in traffic conditions (Toshio *et al.,* 2009). Figure 4.5 demonstrates the tendencies of *NOC* and *CR* by crash mode to k_S in uncongested flow.

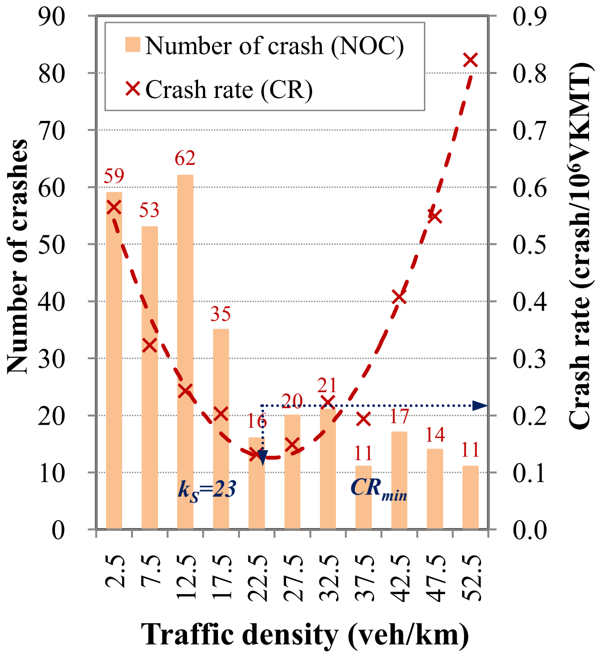


Figure 4.4 ***CR*** **model in uncongested flow (urban expressway)**

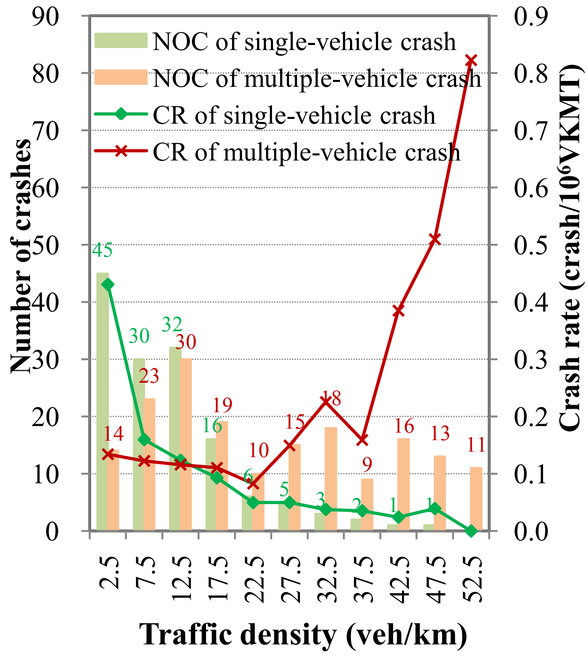


Figure 4.5 ***CR*** **by crash mode in uncongested flow (urban expressway)**

As k_S increase, total *CR* is decreasing in low-density area as opposed to increasing if k_S get over 23veh/km. In view of the different *CR* tendencies, traffic conditions are called as low-density and high-density uncongested flow, respectively. Regarding crash modes, as expected, single- and multiple-vehicle crashes indeed have distinct characteristics. The predominant crash mode near free-flow condition is single-vehicle crash, while its *CR* gets decreased with the increase of k_S and the decreasing tendency becomes milder at high-density stage. Compared to single-vehicle crashes, *CR* of multiple-vehicle crashes keeps stable at a lower value at low-density stage. At high-density stage, the value mushrooms with the increase of k_S, and finally it gets extremely higher in contrast to single-vehicle crashes near congested flow. A compared t-test of *CR* between single- and multiple-vehicle crashes is also performed as summarized in Table 4.2. The results also reveal a significant difference of *CR* by crash mode at high-density stage. Even if no significant difference can be examined at low-density stage, a distinct *CR* tendency can be observed for single- and multiple-vehicle crashes from Figure 4.5, respectively.

Table 4.2 *CR* comparison between single-/multiple crashes (urban expressway)

Traffic conditions	*t-value*	*df*	*Sig.*	Note*
Uncongested flow	-1.786	10	0.104	Low- and high-density stages are bounded at traffic density of 25veh/km in terms of $k_S(CR_{min})$
Low-density stage	0.964	4	0.390	
High-density stage	-3.062	5	0.028	
Congested flow	-2.921	4	0.043	

*The boundary value is rounded off based on $k_S(CR_{min})$.

Single-vehicle crashes appear to be largely geometry-dependent (Christoforou *et al.,* 2012), and significantly associated with speed relative to traffic density (Toshio *et al.,* 2009). On the contrary, empirical results indicate that multiple-vehicle crashes tend to occur under high traffic density (Christoforou *et al.,* 2011). Referring to the above analyses, it can be generally accepted that low-density and high-density uncongested flow stand for different conditions which affect crash characteristics in different ways.

3) Congested flow

In the same way, Figure 4.6 and Figure 4.7 show the tendencies of total *CR* and the value by crash mode following k_S in congested flow, respectively. A compared t-test of *CR* between single- and multiple-vehicle crashes is also carried out (Table 4.2).

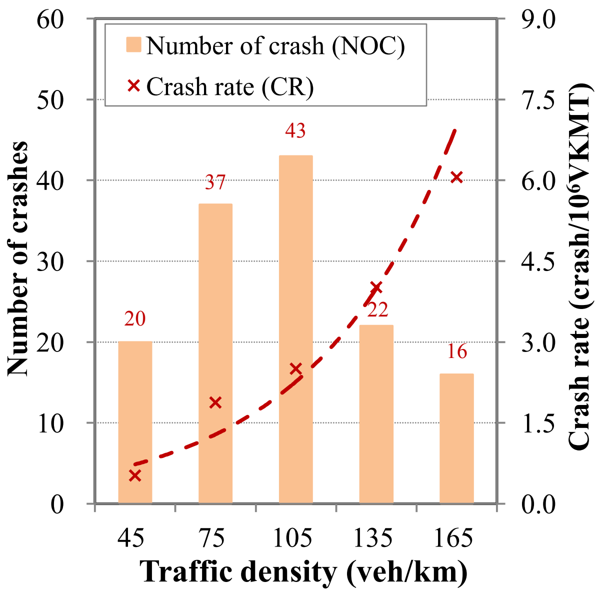


Figure 4.6 *CR* model in congested flow (urban expressway)

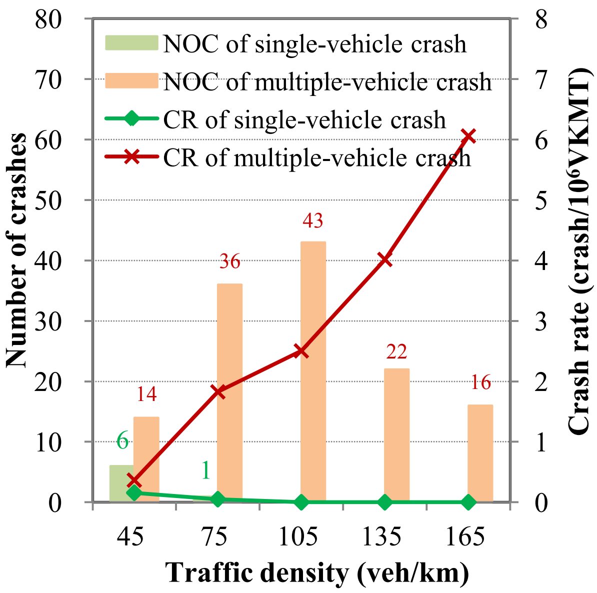


Figure 4.7 *CR* by crash mode in congested flow (urban expressway)

With the increase of k_S, a clearly increasing tendency is characterized for total *CR* model in congested flow. On the other side, are illustrated in Figure 4.7, in congested flow, the uppermost crash mode is multiple-vehicle crashes, and its *CR* can be strongly intensified by the increasing k_S. By comparison, the *CR* of single-vehicle crashes is as trifling as it is. Since the inter-vehicle interaction gets much intensive in congested flow, any change of vehicle behavior may throw an important interruption to the neighboring traffic. As a result, crashes would be mostly involved multiple vehicles. Considering the *CR* characteristics reflected in Figure 4.6 and 4.7, it may be reliable to regard congested flow as a distinct traffic condition that affect crash characteristics in different ways from uncongested flow.

4.3 Qualitative Analysis of Crash Influencing Factors

CR analysis, as an aggregated approach, is inappropriate to examine a variety of factors by a single model given the insufficient crash samples. However, crash occurrence is in truth associated with other explanatory factors, *e.g.,* geometry and ambient conditions, beside traffic flow. On the other hand, not all of the factors are significantly related to crashes and some of them may be conjugated with each other. If incorporate the conjugated factors into crash modeling, it can undermine the validity of crash models for identifying the fundamental mechanisms of these factors. In such case, principal component analysis (PCA) is employed instead and the relative significances of individual variables along with their interrelations with each other are identified.

4.3.1 Theory of PCA

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. Essentially, through orthogonal transformation, a set of observations of possibly correlated variables can be converted into a set of values of linearly uncorrelated variables, which is called principal components (Sanguansat, 2012).

Technically, a principal component can be regarded as a linear combination of optimally-weighted observed variables. 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.

In order to decide upon the number of principal components that may be used as input(s) for crash modeling, three rules are often available (Bajwa *et al.,* 2009). 1) 80% rule: the extracted components should be capable to explain at least 80% of the variance in original dataset; 2) Average eigenvalue rule: all those principal components whose eigenvalues are less than the average may be excluded; 3) Scree plot: the principal components that lie on the flat portion of the curve are ignored and those lie on the steeper slope are retained. Here, Scree plot is the plot of eigenvalues versus the number of eigenvalues. In this study, principal components are selected based on their eigenvalues, and the other two rules are used to check whether the selected ones satisfy or not to statistics significance.

4.3.2 Categorizing traffic conditions

Generally, PCA rotates data through using a linear transformation. Consequently, only the monotonic loading of factors can be reflected reliably. In support of this regard, low- and high-density uncongested flow is further categorized at approximately 25veh/km in terms of the value of $k_S(CR_{min})$, as shown in Figure 4.8, since different monotonicities of CR exist in the two conditions. As a result, the following three traffic conditions will be analyzed, *i.e.,* low-density and high-density uncongested flow as well as congested flow.

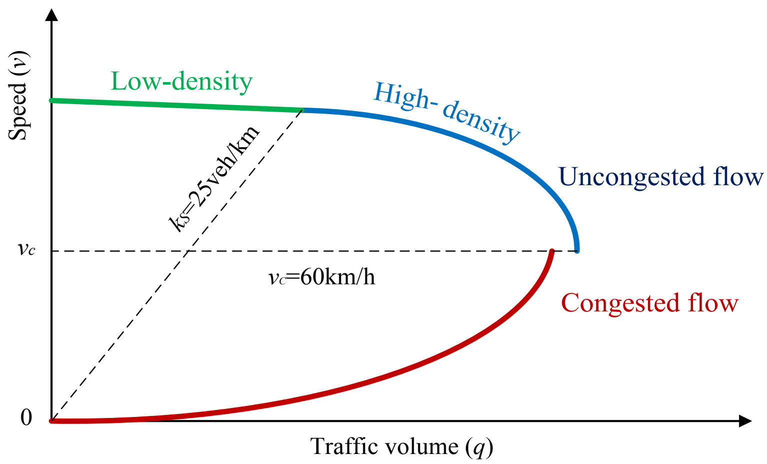


Figure 4.8 Categorizing traffic conditions on urban expressway

4.3.3 Introduction of variables

Table 4.3 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 characteristics. Towards reflection geometric features, h,

I_{CF} and I_{SD} are picked out to demonstrate the vertical, horizontal and the comprehensive alignment variations, respectively. Dummy variables are referred to incorporate ambient conditions into PCA. A dummy variable usually takes two values of 1 and 0, while weather conditions are recorded into more than 4 types. Instead, weather is replaced by pavement conditions since they are highly correlated each other.

Table 4.3 Statistics of explanatory variables (urban expressway)

Variables	Statistics*			Description
	Max.	Ave.	Min.	
k_S	238	34	1	Traffic density (veh/km)
v_S	128.3	69.6	4.70	Average speed (km/h)
I_{CF}	1997.4	106.2	0	Index of centrifugal force (km^3/h^2)
h	12.1	2.47	0.04	Variation in road elevation (m)
I_{SD}	1113	59.9	0	Index of horizontal displacement (m^2)
Pave	F(1)=24.5%			=1 if wet pavement, 0 otherwise
Light	F(1)=29.1%			=1 if nighttime, 0 otherwise
Day	F(1)=26.9%			=1 if holiday, 0 otherwise

* Max./Ave./Min.: the maximum/average/minimum values in statistics; F: frequency.

4.3.4 PCA results by traffic conditions

Through categorizing traffic conditions, 225/94/138 crash samples can be extracted for low-density, high-density uncongested flow and congested flow, respectively.

1) Low-density uncongested flow

Table 4.4 provides PCA results in low-density uncongested flow. Based on the criteria aforementioned, four principal components are remained in terms of their corresponding eigenvalues (>1.0). Furthermore, these principal components can explain at least 80% of the variance in the original datasets (83.7%), and lie on the steeper slope as shown in Figure 4.9. As a result, the selected components have statistical significance and can generally explain the variance in the original dataset.

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 behavior 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 in h. Here, it is worth noting that, k_S and v_S are discovered to belong to different principal components, since both variables are not highly interrelated with each other in such discretionary driving condition.

Table 4.4 PCA results in low-density uncongested flow (225 crash records)

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	-0.368
Ambient light (*Light*)	-0.164	**0.838**	-0.130	0.143
Initial Eigen values	2.31	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

*The variables highly related to each principal component are in bold.

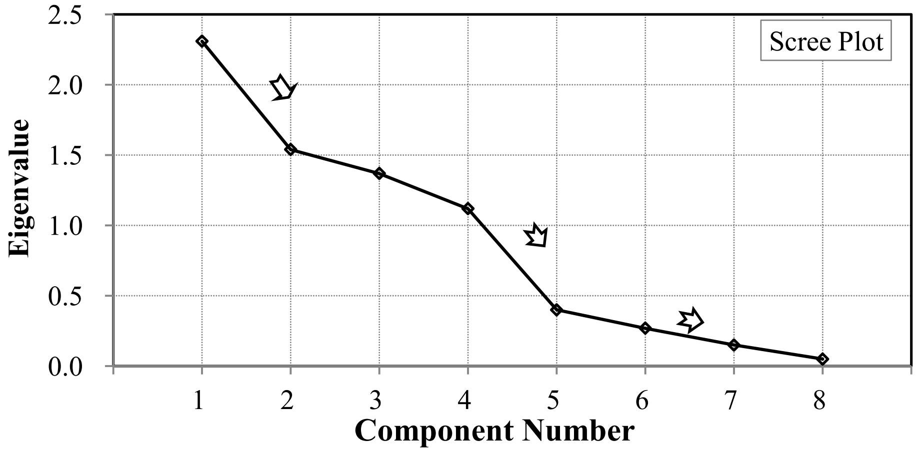


Figure 4.9 Scree Plot for PCA in low-density uncongested flow

2) High-density uncongested flow

As traffic flow increases, the inter-vehicle interaction gets intensive. The corresponding PCA results in high-density uncongested flow are summarized in Table 4.5. The variables that are significantly related to each principal component are selected in terms of their loadings. For judging the relative significance of the same component by traffic conditions, the % of variance explained by individual components is provided as well. Likewise, all of the selected components can be found out to be significant in statistics.

Table 4.5 Summary of principal components by traffic conditions

Traffic conditions	Items	Principal components*							
		1st		2nd		3rd		4th	
		V	L	V	L	V	L	V	L
Low-density uncongested	Value	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	
High-density uncongested	Value	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	
Congested	Value	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	

*V: variable; L: loading.

Compared to low-density uncongested flow, it is clear that the traffic-related variables including k_S and v_S belong to the same principal component, as a reflection of the increased inter-vehicle interaction. Meanwhile, crashes are found to be more probable to occur as k_S increases. Such finding 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. Furthermore, compared to low-density conditions, the significance of geometric features affecting crashes is found to be decreasing during the transformation of traffic conditions, in terms of the related % of variance. Besides, *Day* type gets to be a significant variable of ambient conditions. In contrast to normal weekdays, there are more inexperienced drivers and more travels in unfamiliar conditions on holidays (Anowar *et al.,* 2012). Since the enhanced inter-vehicle interaction may cause a serious sense to inexperienced drivers, if they feel being forced, they would take inappropriate maneuvers.

3) Congested flow

With the further increase of traffic density, congested flow appears. In the same way, Table 4.5 also demonstrates PCA results for the traffic condition. Obviously, crashes are still prone to higher k_S, which confirms the increasing tendency of *CR* to k_S in this condition. Meanwhile, the traffic-related variables (k_S and v_S) become most important. Due to the strongly intensive inter-vehicle interactions, any change in vehicle behavior may seriously affect the surrounding traffic. This phenomenon can account for why multiple-vehicle crashes are predominant in congested flow. Besides, compared to uncongested flow, the significance of road geometry is further reduced. In addition, ambient light is no longer found to be significant, which is likely induced by the insufficient crash samples collected during nighttime in congested flow. Under the force of the congested downstream traffic, the upstream traffic may take lane-changing behaviors frequently. As a result, even if speed slow down significantly from uncongested flow to congested flow, *Pave* is again to be significant since the roadability is critical to safe lane-changing behaviors.

4.4 Crash Risk Estimation Model (CREM)

4.4.1 Matched case-control study

The case-control design is an efficient method to study rare events that is particularly prevalent in epidemiology due to its simplicity, cost-effectiveness, and theoretical soundness (Zheng *et al.,* 2010). In theory, it is an observational-retrospective study: it identifies the cases (a group with outcome) and the controls (a group without outcome), and then traces back to investigate the exposures which are related to outcomes (Lewallen and Courtright, 1998). The theoretical details of this method are beyond the scope of this study, and the interested readers can refer to Armenian (2009) for more in-depth discussion.

For crash analysis, the case is a crash event which may be associated with several exposure factors. The matched controls correspond to the crash scenes or similar conditions but not involved in a crash. Once the factors interested for analysis are identified, they would be incorporated in analysis, and other factors which are not interested should be controlled in the following matched case-control design. Even if the controls should like the cases in many ways, it is possible to over-match, where the factor interested for analysis is also controlled. Over-matching can result in under- estimation on influences. Another important technique for adding power to this method is to enroll more than one control for each case (Lewallen and Courtright, 1998).

As a rule of thumb, a case-to-control ratio around 1:4 is recommended as the statistical power generally does not increase significantly under a 1:4 ratio (Abdel-Aty *et al.,* 2004 and Zheng *et al.,* 2010). Number of factors analyzed in those studies is smaller than 4 while more than 4 variables may be involved in this study. In such case, another way is proposed to decide the reliable case-to-control ratio through examining the contributions of individual factors to crash occurrence whether they significantly change or not as control samples increase (see Section 4.4.5).

4.4.2 Modeling methodology

Crash is a binary outcome event (occurrence *vs.* non-occurrence). If the outcome is binary, the prevalent method to measure the effects of several independent variables on it is logistic regression (Hosmer and Lemeshow, 2004). Besides, conditional logistic regression is a popular method to investigate the relationship between an outcome and a set of explanatory variables in matched case-control studies. Therefore, this study adopts logistic regression to predict the probability of crash occurrence: Y=1 for crash, and Y=0 for non-crash. The relevant theory of this method is briefly introduced below.

Generally, the probability of crash occurrence (P) considering various impact factors on crashes ($X=x_1, x_2, \ldots, x_n$) can be expressed as:

$$P = P(Y = 1|x_1, x_2, ..., x_n) \qquad (0 < P < 1) \tag{4.2}$$

$$P(Y = 1) \sim \beta_0 + \beta_1 x_1 + \beta_2 x_2 + ... + \beta_n x_n \tag{4.3}$$

$$P(Y = 1) \sim exp(\beta_0 + \beta_1 x_1 + \beta_2 x_2 + ... + \beta_n x_n) \tag{4.4}$$

Where, β_0 is the intercept value and β_1, β_2… β_n correspond to the coefficients estimated for individual variables.

In theory, the ranges for both sides of equation (4.3) are the intervals (0~1) and (-∞~+∞), respectively. Meanwhile, *P(Y=1)* is more popular a non-linear function through experience feedback. To solve these problems, the exponential function is adjusted as follows. The related theory for predicting crash probability is shown in Figure 4.10.

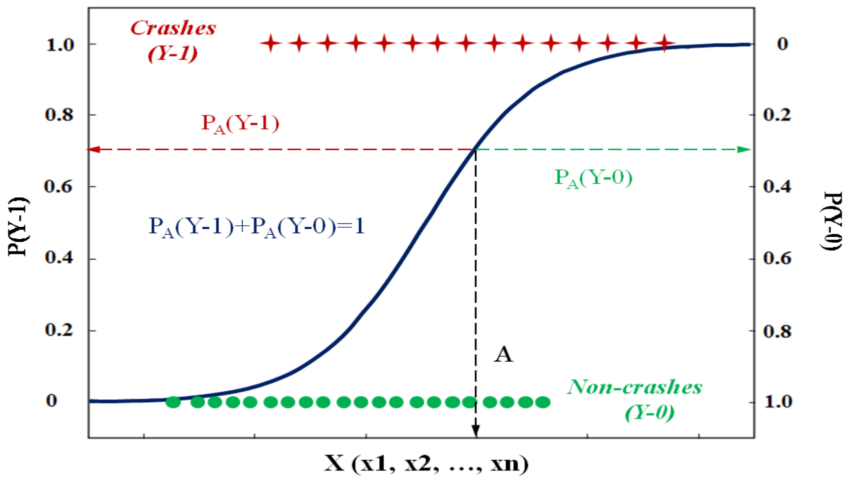


Figure 4.10 Theory of the probability of crash occurrence

4.4.3 Odds value

As shown in Figure 4.10, the probability value (*P*) for the given condition (denoted A in the figure) can be estimated by Formula (4.4), while the value is actually a pseudo-value in practice, since a crash either occur (*Y*=1) or does not (*Y*=0). However, theoretically, this value may be used to reveal the relative risk of crash occurrence for condition A relative to the conditions involved in crash occurrence. For this purpose, logit transformation is applied, and a new index, namely *Odds* of crash occurrence, is employed. As *P* is defined as the probability of crash occurrence, 1-*P* can be regarded as the probability of crash not occurring. Then, the *Odds* of a crash can be defined as *P*/(1-*P*). As stated before, several factors should be involved to crash occurrence. As a result, the joint effects of all of the factors on the *Odds* of crash occurrence put together can be expressed mathematically as:

$$Odds = \frac{P}{1-P} = \exp(\beta_0 + \beta_1 x_1 + \beta_2 x_2 + ... + \beta_n x_n) \tag{4.5}$$

In the purpose to reveal the contribution of each factor on the *Odds* of crash occurrence, *Odds ratio* (*OR*) is utilized. *OR* is defined as the ratio of the *Odds* in favor of a crash from one factor (x_k) to the *Odds* in favor of a crash from another factor (x_j).

$$Odds\ \ ratio = \frac{Odds\ (x_k)}{Odds\ (x_j)} \tag{4.6}$$

To reduce the number of comparison, *Odds* contributed by x_j is taken as 1.0, and thus x_j is actually an assumed factor that is designed to serve simplifying comparison. In this way, *OR* of x_k is equal to its corresponding value of *exp* (β_k). Therefore, this value can be regarded as the variation in *Odds* of crash induced by the increase in a unit of x_k. In essence, *OR* implies the relative contribution of individual factors to crash risk in a given condition.

4.4.4 Variable selection

The significances of individual variables and the interrelation within any two variables have been identified through PCA in section 4.3. Followed those findings, explanatory variables are filtrated for the following crash modeling. As for traffic-related variables, speed v_S and traffic density k_S are not highly related in low-density uncongested flow, and both variables are selected for this traffic. In the other conditions, a single index is applied since both variables belong to the same principal component. At the viewpoint of application in traffic control strategies (*e.g.,* ATMIS), speed v_S is selected. Concerning the independence of geometric factor, the comprehensive index I_{SD} is rejected since it is essentially related to the horizontal and vertical alignment variations. Because v_S has been accepted as an input, the index of centrifugal force I_{CF} ($I_{CF}=S\times v_S$) has to be replaced by horizontal displacement *S,* while the variation in road elevation *h* is still kept as a vertical index. Regarding ambient conditions, all of variables are independent in nature. Hence, those variables are chosen based on their related significances excluding pavement conditions that are not adopted in the following analysis, since those data are not virtually available for non-crash days. The selected variables as inputs for the following analyses are summarized in Table 4.6.

Table 4.6 Independent variables for crash modeling (urban expressway)

Traffic conditions		Traffic flow	Road geometry	Ambient conditions
Uncongested flow	Low-density	k_S and v_S	S and h	*Light*
	High-density	v_S	S and h	*Light* and *Day*
Congested flow		v_S	S and h	*Day*

4.4.5 Case-control samples design

The variables selected above should be matched for crash records and the related non-crash samples. For this purpose, detector data are extracted from the day of crash and from all corresponding non-crash days. The correspondence here means that non-crash days around the day of crash on the same day-of-week, in order to control the monthly variation in traffic characteristics. Meanwhile, in view of the daily variation in traffic demand, these data are collected around the time of crash, in an effort to design control samples within the same traffic conditions of crash events. In this study, half hour prior to the time of crash is accepted. Towards reflecting the variation in geometric features, detector data are also collected from other detector locations of basic segments. Besides, some data should be extracted in the other ambient conditions relative to the conditions at time of crash in the purpose to reveal the effects of variation in ambient conditions on crash risk.

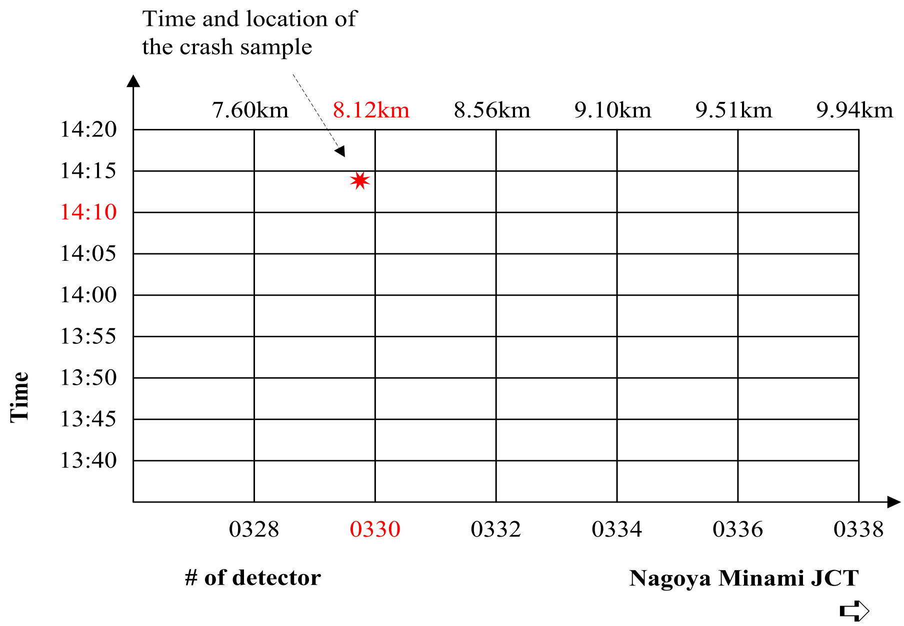


Figure 4.11 Time-space allocation of the crash example

Table 4.7 Example of case-control design
(Low-density uncongested flow)

Y^*	Day	Time	Detector #	v_S(km/h)	k_S(veh/km)	S(m)	h(m)	*Light*
1	04/04/2009	14:10	0330	73.4	19	22.2	1.00	0
0	03/28/2009	13:50	0312	89.4	22	1.78	4.60	0
0	03/28/2009	13:45	0314	90	20	0.11	0.00	0
…	…	…	…	…	…	…	…	…
0	04/18/2009	13:40	0316	89.3	20	3.31	1.55	0
0	04/18/2009	2:10	0318	111.8	3	0.00	-5.98	1
…	…	…	…	…	…	…	…	…
0	04/25/2009	2:05	0328	84.8	5	39.9	2.75	1
…	…	…	…	…	…	…	…	…
0	05/02/2009	14:10	0338	80.7	21	4.17	-0.55	0

* Y=1: crash event; Y=0: the matched non-crash samples.

As illustrated in Figure 4.11, a crash occurred at 14:15 on April 4th, 2009 (Wednesday) at 8.1km in Southbound of Odaka route. Referring to the method explained in Chapter 3, data from detector #0330 (nearest to 8.1km) at 14:10 (at least 5 minutes before 14:15) can be used to represent traffic condition prior to this crash. The geometric variation at the location of crash sample is extracted from geometric database, and the related ambient conditions preceding crash can be referred to crash records. Regarding its control samples, the corresponding days are regarded as other Wednesdays (and non-crash days) before/after the day of crash in one month. Then, traffic data are randomly collected from the detectors on Odaka route from 13:40 to 14:10, and some data are extracted in nighttime. Of course, those control samples are extracted at basic segment only and the data should also be examined whether they belong to low-density uncongested flow or not. The matched case-control design process above is exemplified in Table 4.7.

4.4.6 Development of CREM

1) Minimal required control sample size

This study designs a method to examine the minimal required control samples to simplify the above control sample design process. If the *OR* of individual variables (known as its contribution to crash risk) does not significantly change after inputting more than *m* control samples, m is regarded as the minimal required control sample size. Correspondingly, the maximal required case-to-control ratio can be taken as 1: m.

Table 4.8 *OR* of individual variables based on varied case-to-control ratios
(Low-density uncongested flow)

Case-to-control ratio (1:m)	v_S	k_S	S	h	*Light*
1:3	1.15	0.988	1.18	1.15	1.35
1:5	1.10	0.987	1.12	1.09	1.32
1:7	1.07	0.985	1.08	1.05	1.29
1:9	1.05	0.985	1.03	1.03	1.27
1:10	1.03	0.985	1.02	1.02	1.24
1:15	1.03	0.985	1.01	1.01	1.23
1:20	1.02	0.984	1.01	1.01	1.23
…	…	…	…	…	…
1:30	1.02	0.984	1.01	1.01	1.23
…	…	…	…	…	…
1:40	1.02	0.983	1.01	1.01	1.22
…	…	…	…	…	…
1:50	1.02	0.983	1.01	1.01	1.22

Table 4.8 provides the *OR* of various variables corresponding to individual statistic turns through increasing control samples for low-density uncongested flow. Even the *OR* values vary among different variables, they do not significantly change for the same variable after inputting more than 10 control samples. For the sake of security, the case-to-control ratio is finally defined as 1:20. Based on this ratio, a total of 193, 87 and 79 crashes along with their corresponding control samples are designed successfully in low-density, high-density uncongested and congested flow, respectively.

2) Modeling results

Table 4.9 demonstrates the results of CREM in low-density uncongested flow. All of the independent variables are found to be significant at 95% level. In terms of the results of model test and the value of R^2, which is summarized in Table 4.10, the developed model seems to be of statistical significance.

Table 4.9 Parameter estimation for CREM in low-density uncongested flow

Exposure	Variable	Parameter estimate	Standard error	Wals	*df*	*Sig.*	*OR**
Traffic flow	v_S	1.83×10^{-2}	7.56×10^{-3}	5.40	1	0.020	1.02
	k_S	-1.64×10^{-2}	1.69×10^{-2}	4.91	1	0.026	0.984
Geometric features	S	1.09×10^{-2}	2.24×10^{-3}	20.4	1	0.000	1.01
	h	$\pm7.48\times10^{-3}$	3.31×10^{-2}	4.99	1	0.025	1.01
Ambient conditions	*Light*	2.08×10^{-1}	1.97×10^{-1}	4.43	1	0.035	1.23
Intercept value		-1.5	6.97×10^{-1}	13.4	1	0.000	-

**OR*: Odds ratio.

For the other traffic conditions, Table 4.10 summarizes their corresponding CREM. The related results of model test can reveal the significances of the models. Besides, all of the independent variables are also significant at 95% level.

Table 4.10 Summary of CREM (urban expressway)

Exposures	Variables	Uncongested flow				Congested flow	
		Low-density		High-density			
		Coef. (β)*	*Sig.*	*Coef.* (β)	*Sig.*	*Coef.* (β)	*Sig.*
Traffic flow	v_S	1.83×10^{-2}	0.020	-2.82×10^{-2}	0.000	5.39×10^{-2}	0.001
	k_S	-1.64×10^{-2}	0.026				
Geometric features	S	1.09×10^{-2}	0.000	2.24×10^{-3}	0.036	1.25×10^{-3}	0.045
	h^{**}	$\pm7.48\times10^{-3}$	0.025	$\pm1.21\times10^{-2}$	0.024	$\pm2.18\times10^{-2}$	0.001
Ambient conditions	*Light*	2.08×10^{-1}	0.035	1.55×10^{-1}	0.000		
	Day			1.63×10^{-1}	0.007	1.85×10^{-1}	0.014
Intercept value		-1.5	0.000	1.65	0.000	-2.45	0.000
Model test	-2Log likelihood	1828.2		515.9		407.9	
	Chi-square	51.13		83.59		11.05	
	Significance	0.000		0.000		0.036	
Pseudo R^2	Cox and Snell	0.718		0.693		0.630	
	Nagelkerke	0.957		0.924		0.841	
	McFadden	0.913		0.851		0.718	

* *Coef.*: estimated coefficient.
** +/- correspond to upgrade and downgrade, respectively.

However, as shown in Table 4.10, the statistical significance of CREM in congested flow seems poor in contrast to uncongested flow. As it is known, traffic oscillation is a typical characteristic in congested flow, which is often characterized by recurring decelerations followed by accelerations. In this regard, average traffic variables may not be the optimal index to reveal the short-term turbulence of traffic conditions.

4.4.7 Model validation

As explained by Hosmer and Lemeshow (2004), R^2 of logistic regression is a pseudo value and not used to assess goodness-of-fit. For identifying the model fit, that book also proposes to compare the observed value with the predict value. This study thereby validate CREM through calculating the relative risk of observing a crash versus not. To minimize the misclassification rate, the *Odds* value (for Y=1) of 1.0 is regarded as a threshold value of hazardous conditions (hazardous if *Odds*≥1.0 and vice versa). In hazardous conditions, the probability of crash occurrence is larger than the probability of crash not occurring.

Based on the coefficients in Table 4.10 and the equation in Formula (4.5), the *Odds* of crash occurrence before the time of crash sample are estimated as shown in Figure 4.12. In the figure, geometric design of the related segment is roughly demonstrated, and the estimated *Odds* values are classified into several levels represented by various colors.

14:10	1.054	1.021	0.437	0.942	0.732	0.704
14:05	1.093	1.010	0.471	1.047	0.777	0.686
14:00	1.085	0.943	0.451	0.892	0.797	0.779
13:55	0.994	0.891	0.443	0.983	0.743	0.693
13:50	0.974	0.790	0.368	0.947	0.642	0.669
13:45	0.977	0.892	0.464	0.894	0.657	0.648
13:40	0.948	0.785	0.396	0.915	0.671	0.587
Odds	0328	0330	0332	0334	0336	0338

1.0-1.1
0.9-1.0
0.8-0.9
0.7-0.8
0.6-0.7
0.5-0.6
0.4-0.5
0.3-0.4

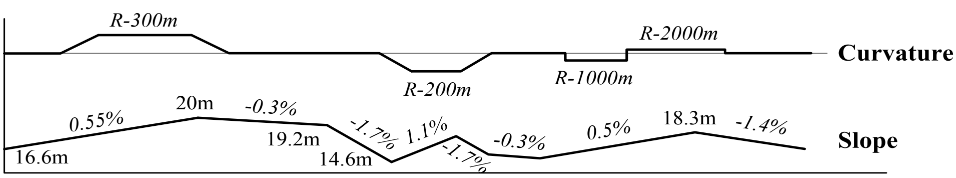


Figure 4.12 *Odds* of crash occurrence preceding the crash sample

It is obvious that the *Odds* value at detector #0330 gets higher than 1.0 from 14:05. If consider detector #0328 together since the traffic characteristics at the two detectors are influenced by the same curve, the *Odds* is observed over 1.0 from 14:00. Compared to the recorded time (14:15), traffic at most 15 minutes before crash occurrence can be found to be hazardous conditions. For traffic management, this finding is beneficial since it may provide leverage in terms of time to be able to predict and avoid an impending crash. Combined these estimated *Odds* values at individual detectors with their related geometric features, it is indicated that the detectors which locate in curves (*e.g.,* #0328) or in the nearly downstream of curves (*e.g.,* #0330 and #0334) have high crash risk. Furthermore, smaller radii of curves, and higher crash risk is caused (*e.g.,* #0328 *vs.* #0338). Hence, the potential crash-prone sections can be predicted through the *Odds* value estimations.

In order to reveal the cause of crash occurrence, the temporal-variations of traffic variables (*i.e.,* v_S and k_S) at detector #0330 which is closest to the location of crash are exhibited in Figure 4.13. As reflected in that figure, from 14:00 on, speed at detector #0330 actually exceeds 70km/h, and traffic density gets down correspondingly. The modeling results in Table 4.9 illustrates that higher speed combining with lower traffic density are positive to crash risk. Therefore, we can generally conclude that the crash occurrence is significantly associated with the increase in speed at such location of poor design consistency.

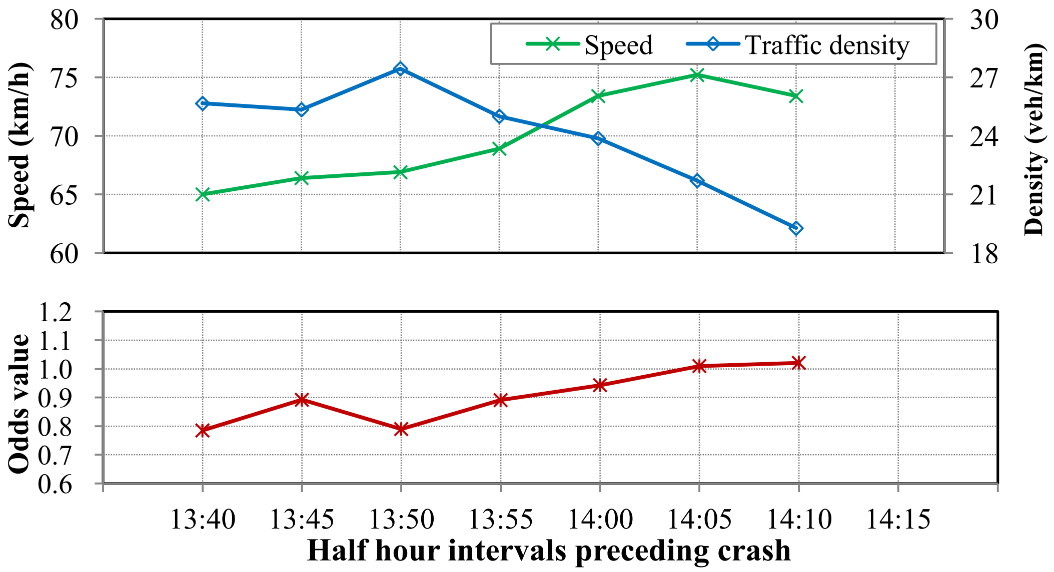


Figure 4.13 Time-serious variations in traffic flow and crash risk at detector 0330

Table 4.11 Crash risk estimation for crash events
(Low-density uncongested flow)

ID of crash	Observed values					Estimated *Odds*	Hazardous condition (*Odds*>1.0)
	v_S(km/h)	k_S(veh/km)	S(m)	h(m)	*Light*		
1	86.4	8	7.69	-4.50	1	1.23	Yes
2	97.5	10	12.1	2.50	1	1.61	Yes
3	92.0	13	11.1	0.28	0	1.10	Yes
4	62.7	20	0.00	-0.77	0	0.620	No
5	87.3	13	15.1	1.37	0	1.06	Yes
6	88.7	19	20.3	1.50	0	1.05	Yes
7	103.4	7	30.2	1.14	1	2.28	Yes
8	80.1	20	5.97	-0.40	0	0.741	No
…	…	…	…	…	…	…	…
Total number of crash			Estimated hazardous conditions				Percentage*
209			181				86.6%

* Regarded as the accuracy of the model for predicting crash events.

For other crash events, the *Odds* values are estimated in the same way. An example is illustrated in Table 4.11 for low-density uncongested flow. Through comparing the relative risk of a crash event to the boundary *Odds* value of 1.0, it is observed that 86.6% of total crash events can be correctly predicted in hazardous conditions.

Figure 4.14 illustrates the time-space variation of *Odds* values on July 1st, 2009 (non-crash day) at the same section during 6:00 to 6:30. As expected, most traffics are found to be not hazardous (*Odds*<1.0) except the one at 6:00 at detector #0328 that locates in a small curve. Such case is regarded as a false prediction. Meanwhile, the *Odds* values are also observed to be relatively lower at sections far from small curves. In this way, the potential hazardous locations can be identified on non-crash days, and they may be flagged with warnings by variable message signs (VMS) in order to remind drivers to pay attention. Likewise, the temporal variations in speed and traffic density at detector #0328, where some high *Odds* values are observed, are also analyzed, as shown in Figure 4.15.

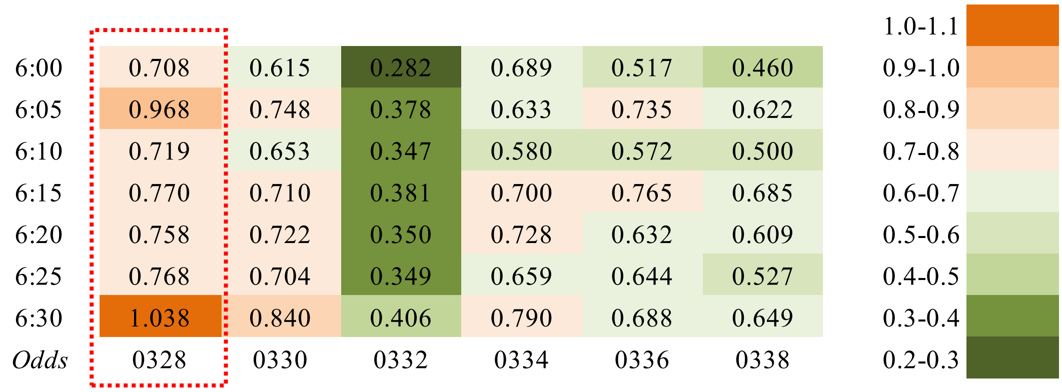


Figure 4.14 *Odds* of crash occurrence on a non-crash day

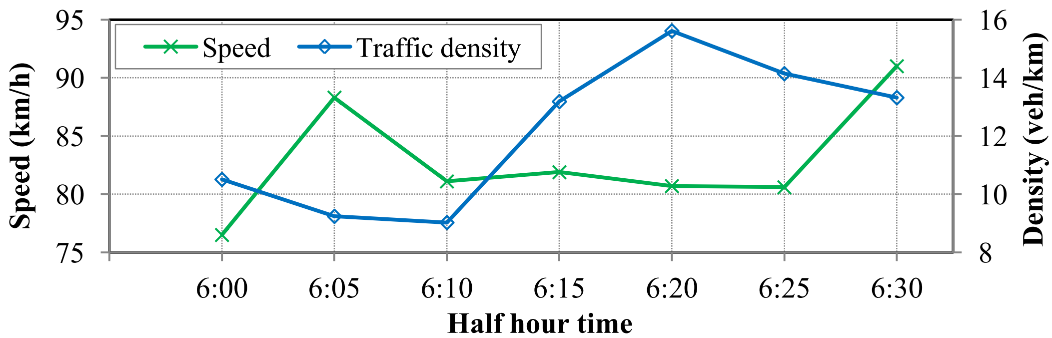


Figure 4.15 Time-serious variations in traffic flow at detector #0328

In the same way, *Odds* for other traffic conditions are estimated. The ratio of the number of crashes that can be predicted in hazardous traffic to the number of total crashes is defined as accuracy. Meanwhile, the ratio of number of non-crash samples being predicted in hazardous traffic to the number of total non-crash events is regarded as false alarm rate.

Table 4.12 Validation results of the developed CREM (urban expressway)

Samples		**Predictive performance**			
		(a) Odds≥1 (hazardous)	(b) Accuracy $\frac{(a)}{(a)+(c)} \times 100\%$	(c) Odds<1 (not hazardous)	(d) False alarm rate $\frac{(a)}{(a)+(c)} \times 100\%$
Low-density uncongested	0 (non-crash)	845		3335	20.2
	1 (crash)	181	86.6	28	
High-density uncongested	0 (non-crash)	351		1204	23.3
	1 (crash)	62	80.5	15	
Congested	0 (non-crash)	410		930	30.6
	1 (crash)	47	70.2	20	

The validation results are summarized in Table 4.12, and it points out that 80.5% and 70.2% of crash events can be identified as hazardous traffic for high-density uncongested and congested flow, respectively. However, in the three traffic conditions, there are still 20.2%, 23.3% and 30.6% of non-crash events that are misclassified as hazardous conditions, respectively. Note that, the proposed threshold value (*Odds*=1) in this study may be varied in an effort to achieve desirable validation in terms of the tradeoff between overall classification accuracy and crash identification.

Predictive performance of the above CREM may be not so perfect but it is worth mentioning that the models do not consider any variables related to driver factors and errors. In this study, it is difficult to obtain those variables based on the original data which are mostly collected at aggregated level. For the same reason, the short-term turbulence of traffic flow cannot be reflected appropriately by average variables in 5-min. Given these backgrounds, the predictive powers of the models seem to be acceptable. Furthermore, it is evident that the modeling strategy by focusing on traffic conditions is more reliable in comparison with the previous studies in terms of predictive performance (*e.g.,* the model of 59% predictive power in Abdel-Aty and Pemmanaboina, 2006), not to mention more influencing factors are involved in this study. Therefore, it is reasonable to believe that the analysis in this study including concept and technique is promising considering the application in proactive traffic management strategies, even if a substantial effort is further required to adapt such analysis into practice.

4.5 Contributions of Explanatory Factors on Crash Risk

In the three traffic conditions, the coefficients of individual variables involved in CREM are analyzed in Table 4.13. Based on that table, the contribution of one unit of each variable in favor of the *Odds* of crash occurrence can be observed directly. Here, it is worth noting that except *Light* and *Day*, which are two dummy variables, other variables are continuous ones. Thus, the relative contributions between variables should consider the domains of variability of individual variables.

Table 4.13 Model parameters for evaluating contributions (urban expressway)

Variable*	Uncongested flow		Congested flow
	Low-density	High-density	
v_S	0.0183	-0.0282	0.0539
k_S	-0.0164		
S	0.0109	0.00224	0.00125
h	±0.00748	±0.0121	±0.0218
Light	0.208	0.155	
Day		0.163	0.185

* *Light* and *Day* are dummy variables, while others are continuous ones.

4.5.1 Low-density uncongested flow

In low-density uncongested flow, CREM involves five variables: speed v_S, traffic density k_S, horizontal displacement S and the variation in road elevation h as well as ambient *Light*. In terms of coefficients, other variables except of k_S have positive contribution on the *Odds* of crash occurrence. Under the interaction of these variables, the variations in *Odds* value are demonstrated in Appendix A. Through referring to Table 4.13 and these figures, the contributions of individual variables on crash risk can be found out:

- Geometry: the rise in S and h indicates the increase in the difficulty for drivers to control vehicle behaviors. Meanwhile, the contribution of S is more important than that of h on crash risk in high-speed driving conditions.

- Traffic flow: due to the difficulty for drivers to control vehicle behaviors and the centrifugal force in curve sections at high speed, the larger v_S is and the higher crash risk is. In contrast, crash risk would be on the decrease with the increase in k_S owing to the rise of drivers' attention. Comparatively, the contribution of v_S is more notable relative to k_S in such discretionary driving conditions.

- Ambient conditions: only *Light* is involved and the traffic in nighttime is found out less safe than the traffic in daytime because of the reduced visibility.

These findings above indicate that high speed coupled with nighttime and large variation in geometric design can increase crash risk. At the view point of traffic management, it is important to control the operating speed to avoid the potential hazardous conditions. In this regard, by referring to the developed model, the maximum safe speeds corresponding to individual driving conditions are proposed in Table 4.14. In practice, these values also provide leverage for drivers to check driving maneuvers by themselves.

Table 4.14 Maximum safe speed proposed (km/h)
(Low-density uncongested flow on urban expressway)

Horizontal displacement (m)		0	10	20	30	40	50	60	70	80	90	100
Daytime	k_S=10veh/km, h=0m	118	112	106	100	94	88	82	76	70	64	58
	k_S=10veh/km, h=10m	114	108	102	96	90	84	78	72	66	60	54
	k_S=20veh/km, h=0m*					102	96	90	84	78	72	66
	k_S=20veh/km, h=10m				104	98	92	86	80	74	68	62
Nighttime**	k_S=10veh/km, h=0m	106	100	94	88	82	76	70	64	58	52	46
	k_S=10veh/km, h=10m	102	96	90	84	78	72	66	60	54	48	42

* The operating speed at the density of 20veh/km is mostly below 100km/h.
** Traffic density in nighttime is often smaller than 10veh/km.

4.5.2 High-density uncongested flow

In high-density uncongested flow, the inter-vehicle interaction is intensive. For this traffic condition, the involved variables for CREM include speed v_S, horizontal displacement S, the variation in road elevation h, ambient *Light* and *Day* type. Regarding the contributions of individual variables to crash risk, the following characteristics can be achieved based on the coefficient in Table 4.13 and the figures in Appendix A:

- Geometry: S and h still have positive contributions on the *Odds* value, while the significances of the two variables changes. The importance of S in favor of a crash is reduced as opposed to the increased importance of h.

- Traffic flow: in high-density uncongested flow*,* v_S has a negative contribution on crash risk. Meanwhile, in terms of the absolute value of coefficient, the effect of v_S on crash risk gets more important in contrast to low-density uncongested flow.

- Ambient conditions: nighttime still has positive contribution on crash risk, while its related significance is slightly reduced. *Day* is an added involved variable in high-density uncongested flow compared to low-density uncongested flow.

In high-density uncongested flow, the reduced contribution of *S* on crash risk may be associated with the reduced centurial force. The force is affected by horizontal geometry combining with speed, while speed is actually on the decrease as traffic density increases.

On the contrary, the contribution of *h* to crash risk is found out to be on the rise. Vertical alignment is highly related to the behaviors of HV. At downhill section, the increase in downgrade would aggravate the difficulty for drivers to control vehicles, which is possible to induce rear-end collision. Even at uphill section, the rise of upgrade may worsen the driving conditions for HV. The effect can further interrupt the upstream vehicles and such interruption may be aggravated when traffic density increases.

With the increase in traffic density, the change of speed becomes more sensitive to crash risk. Since the inter-vehicle interaction gets intensive, for downstream vehicles, a small change of speed would induce instability in traffic conditions in the upstream. Considering the contributions of traffic flow in low-density uncongested flow together, the findings can support the developed U-shaped *CR* model in uncongested flow.

On holiday, a large increase in recreational private travel exists. As a result, there are more travels in unfamiliar conditions and more inexperienced drivers. Given these situations, the inappropriate reactions of drivers to the unexpected variation in traffic conditions may be more frequent in contrast to normal weekdays.

4.5.3 Congested flow

The developed CREM in congested flow incorporates four variables; speed v_S, horizontal displacement *S*, the variation in road elevation *h* and *Day* type. Even if ambient *Light* is not involved, which is likely induced by the limited samples including congested flow and crash events in nighttime, it is not exact to conclude that the effect of ambient *Light* on crash risk is not significant. Because these variables are selected based on the results of principal component analysis that is just focusing on crash events. Through referring to the findings in Table 4.13 and the figures in Appendix A, the contributions of individual variables to crash risk are found out to be characterized:

- Geometry: compared to uncongested flow, the contribution of *S* on crash risk is continuously reduced, whereas, the related contribution of *h* is further on the rise.

- Traffic flow: v_S appears positive effect on crash risk, and its coefficient is observed to increase significantly in contrast to uncongested flow.

- Ambient conditions: the traffic on holidays is less safe than that on weekdays, and the rising trend of *Odds* value induced by the alteration of *Day* type is more distinct compared to uncongested flow.

Since operating speed is further reduced in congested flow, the contribution of S on crash risk is accordingly going on decreasing. On the other side, the inter-vehicle interaction gets more powerful in that traffic condition, and the interruption of HV to the surrounding traffic is stronger. As a result, the sensitivity of v_S to crash risk gets up and the contribution of h to crash risk becomes greater in comparison with uncongested flow.

Given the above analyses, the contributions of individual variables on crash risk with the variation in traffic conditions become explicit. The horizontal geometry, as the most significant explanatory variables in low-density uncongested flow as shown in Table 4.5, its effect on crash risk gets less significant with the increase in traffic density. On the contrary, the impact of vertical geometry on crash risk is actually on the increase. Meanwhile, the change of speed becomes more sensitive to crash risk while traffic density increasing. Ambient conditions are another non-negligible exposure. Generally, nighttime and holiday may aggregate crash risk relative to daytime and weekday, respectively.

4.6 Summary

This chapter presented a crash risk estimation model (CREM) at basic segments of NEX, considering the interaction of geometry, traffic flow and ambient conditions. *CR* model was firstly developed for understanding the general safety performance with the variation in traffic conditions. In terms of the safety performance, traffic conditions were categorized into low-density, high-density uncongested flow and congested flow. Then, by focusing on traffic conditions, PCA was employed, and the correlations of explanatory variables and their significances affecting crashes were identified, in order to select the appropriate independent variables for crash modeling. According to the findings of PCA, a matched case-control study was adopted for identifying the effects of independent variables on crashes. Based on these case-control samples, the CREM was finally developed through applying conditional logistic regression. The model was further found to be significant in statistics and of accepted goodness-of-fit, with 86.6%, 80.5% and 70.2% of predictive performance in low-density, high-density uncongested and congested flows, respectively.

Through referring to CREM, the quantitative effects of individual variables on crash risk can be identified. With the increase in traffic density, the significance of horizontal geometry affecting crash is on the decrease. In contrast, the effect of vertical geometry on crash risk becomes more significant. Due to the more powerful inter-vehicle interaction, operating speed gets more sensitive to crash risk when traffic density increases. Ambient conditions are another non-negligible exposure. Generally, nighttime and holiday may increase crash risk relative to daytime and weekday, respectively.

The potential benefits of integrating the model in proactive traffic management for safety are numerous. Based on the model by traffic conditions, a predictive system of crash risk can be developed on a real-time basis. Once a condition is identified to be hazardous, it may be flagged with warnings by variable message signs (VMS). Meanwhile, the findings of quantitative effects of various independent variables on crash risk may help prioritize countermeasures and recommend some specific measures for smoothing hazardous conditions. In addition, the safety performance of an adopted countermeasure may be estimated in advance by referring to this model.

Chapter 5

CRASH RISK ESTIMATION MODEL FOR INTERCITY EXPRESSWAY AND ITS DIFFERENCE FROM URBAN EXPRESSWAY

5.1 Introduction

As concluded in Chapter 3, geometric design and traffic characteristics between urban and intercity expressways are virtually different. Thus, crash characteristics may be varied by expressway type. Thereby, crash analysis for urban and intercity expressways is separately exerted, and this chapter will focus on intercity expressway to develop CREM at basic segments. Meanwhile, the contributions of explanatory variables on crash risk are comparatively analyzed between two types of expressways, in order to comprehensively identify the crash characteristics caused by the variation in geometry and traffic characteristics. For these purposes, this chapter is organized as follows. The following section analyzes *CR* tendency with the increase of traffic density. By focusing on traffic conditions, PCA is applied to qualitatively analyze explanatory factors regarding their related significances and interrelations with each other in section 5.3. Based on the significant and independent variables, CREM is developed in section 5.4. Next, section 5.5 analyzes the contribution of various variables on crash risk. Later, the different sensitivities of these variables to crash risk are identified between urban and intercity expressways in section 5.6. Finally, section 5.7 offers a brief summary and motivation for this thesis.

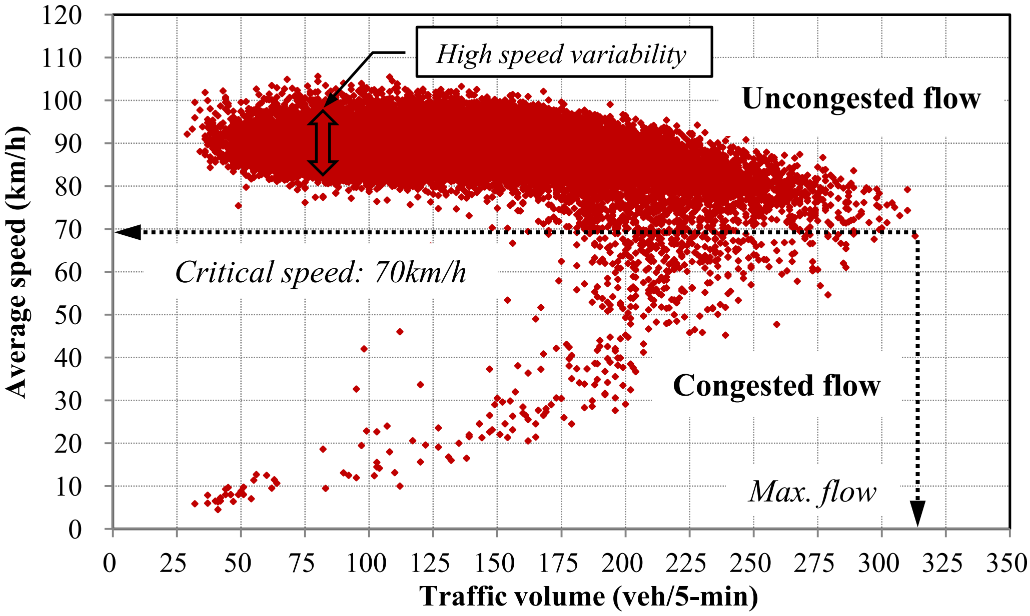


Figure 5.1 Traffic flow-speed diagram at Toyota JCT

5.2 Analysis on *CR* Tendency

5.2.1 Classification of traffic conditions

In the purpose to find out the critical speed for classifying un- and congested flows, the diagram of traffic flow-speed at Toyota junction (JCT) on Tomei Expressway (one typical bottleneck at basic segments) is analyzed as displayed in Figure 5.1. It illustrates that 70km/h can be regarded as the related value. The corresponding values at other bottlenecks are detected to be around 70km/h as well through referring to Kobayashi *et al.* (2011).

Based on traffic density k_S calculated by Formula (3.3), this chapter also classifies traffic conditions into several groups to reflect the variation in traffic characteristics. Considering the number of crash samples available as well as the comparative analysis with urban expressway, 5veh/km and 30veh/km are considered to be the aggregation intervals of k_S for un- and congested flows, respectively, as summarized in Table 4.1 as well.

5.2.2 *CR* tendency

For individual traffic conditions, *CR* is statistically calculated through Formula (4.1) as well. Then, the characteristics of *CR* following traffic density are separately analyzed in un- and congested flows. Furthermore, the *CR* characteristics on intercity expressway different from urban expressway are investigated.

1) The whole traffic conditions

For un- and congested flows, 1113 crashes and 383 crashes are available for statistics, and Figure 5.2 shows the distributions of total crash rate *CR* following traffic density k_S in the same way used in Chapter 4. Generally, *CR* is convex downward to k_S in uncongested flow and on the rise with the increase of k_S in congested flow. The quadratic function and exponential function can roughfully be fitted to the tendency of *CR* in the two traffic conditions, respectively. The figure further points out that congested flow has exceedingly higher *CR* in contrast to uncongested flow.

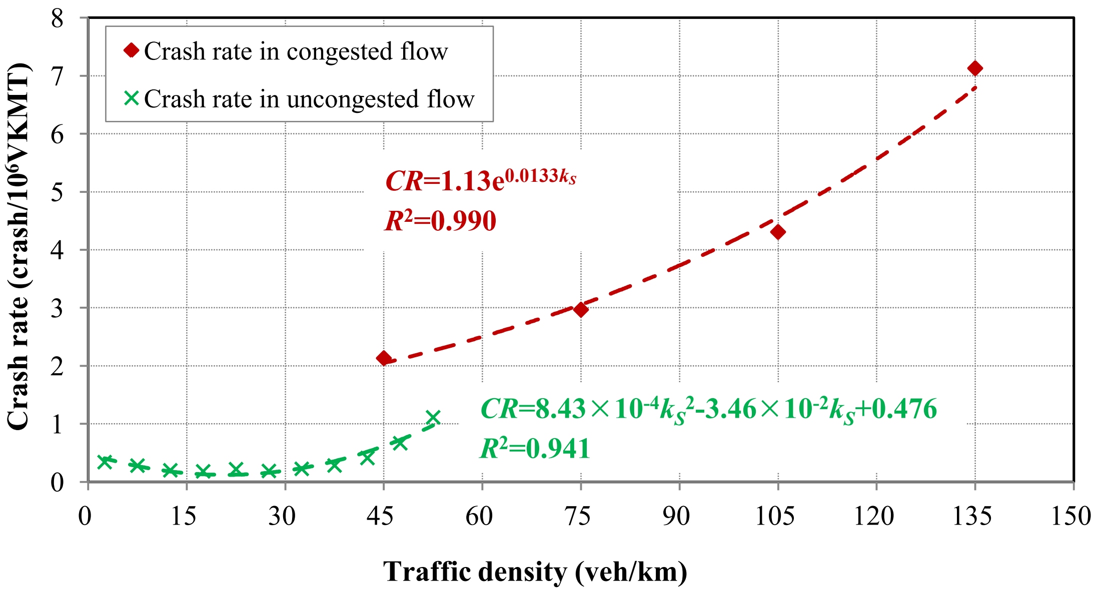


Figure 5.2 Total *CR* regression model (intercity expressway)

2) Uncongested flow

In the following, uncongested flow is also focused on and its *CR* model is further show in Figure 5.3. Meanwhile, the characteristics of *CR* model for intercity expressway different from urban expressway are also demonstrated in the figure. Besides, this study also makes a distinction between single- and multiple-vehicle crashes considering the variation in traffic conditions, as illustrated in Figure 5.4.

Regarding the difference by expressway type, Figure 5.3 implies that, at low-density stage, intercity expressway has lower *CR* compared to urban expressway. When traffic flow increases, *CR* on intercity expressway increases rapidly and becomes much higher than that on urban expressway. For verifying these differences, a paired t-test of *CR* between two types of expressways is exerted, as indicated in Table 5.1. The findings further demonstrate that *CR* is not significantly different by expressway type for the whole uncongested flow. However, if sub-classify uncongested flow into low- and high-density stages based on the monotonicity of *CR* tendency, a significant difference of *CR* between urban and intercity expressway can be found out separately.

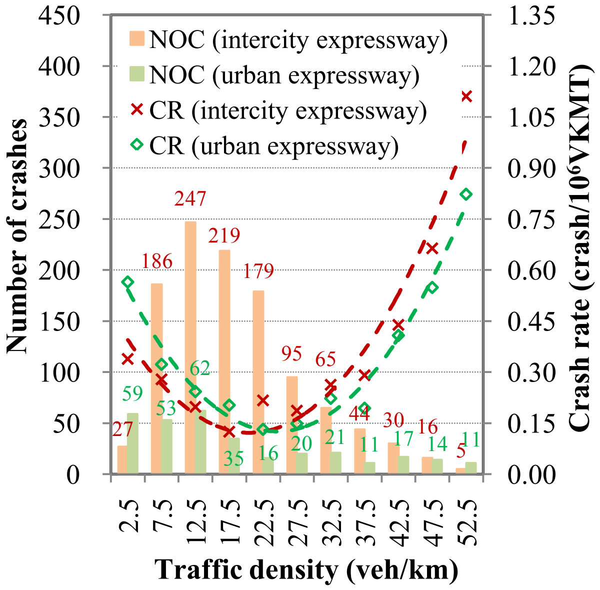


Figure 5.3 *CR* models by expressway type in uncongested flow

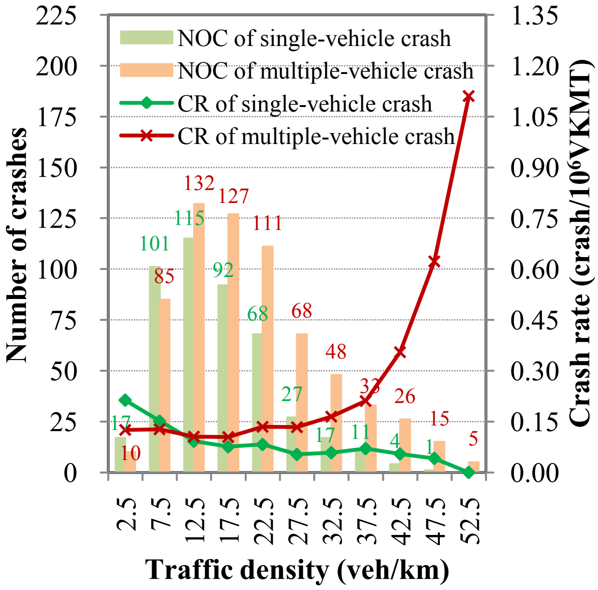


Figure 5.4 *CR* by crash mode in uncongested flow (intercity expressway)

Table 5.1 *CR* comparison between urban and intercity expressways

Traffic conditions	*t-value*	*df*	*Sig.*	Note*
Uncongested flow	-0.697	10	0.502	Low- and high-density stages are bounded at traffic density around 25veh/km
Low-density stage	2.841	4	0.047	
High-density stage	-2.713	5	0.035	
Congested flow	-4.426	4	0.021	

* The boundary value is rounded off based on $k_S(CR_{min})$.

As shown in Figure 5.4, *CR* tendencies are characterize in different features by crash mode as expected. Single-vehicle crash is the dominant crash mode near free-flow stage, and its *CR* descends with the increase of k_S, while the tendency gets to be milder at high-density stage of uncongested flow. By contrast, the *CR* of multiple-vehicle crash is in stability with a relatively small value at low-density stage. At high-density stage, with the increase of k_S, the value rises up and its rising trend gets to be more significant. Near congested flow, it becomes extremely higher compared to single-vehicle crash. To confirm the differences, Table 5.2 carries out a paired t-test and it implies that *CR* by crash mode is in truth significantly different at high-density stage. At low-density stage, although the difference is not significant in statistic, the different tendencies of *CR* can be observed between single- and multiple-crashes. Based on the above findings, it sounds that crash rate characteristics are different at low- and high-density stages. Thus, traffic conditions are further sub-classified into low- and high-density uncongested flow.

Table 5.2 *CR* comparison between single-/multiple-vehicle crashes
(Intercity expressway)

Traffic conditions	*t-value*	*df*	*Sig.*	Note*
Uncongested flow	-1.980	10	0.076	Low- and high-density stages are bounded at traffic density of 25veh/km
Low-density stage	0.139	4	0.897	
High-density stage	-3.364	5	0.021	
Congested flow	-3.251	3	0.047	

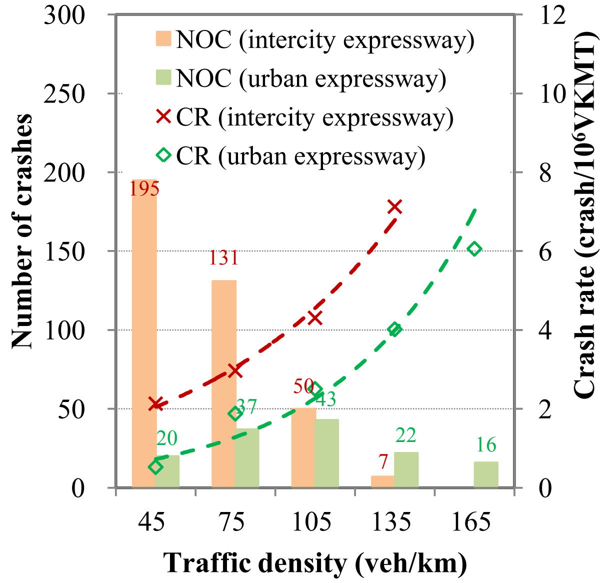
* The boundary value is rounded off based on $k_S(CR_{min})$.

3) Congested flow

In the same way, Figure 5.5 and Figure 5.6 give the tendencies of total *CR* and the value by crash mode following k_S in congested flow, respectively. A compared t-test of *CR* between single- and multiple-vehicle crashes is also carried out, as shown in Table 5.2.

In congested flow, *CR* on intercity expressway is found out to be higher in contrast with urban expressway. The results of a paired t-test of *CR* by expressway type can also confirm such relativity, as exhibited in Table 5.1 as well.

As shown in Figure 5.6, the tendencies of *NOC* and *CR* by crash mode, in congested flow, multiple-vehicle crash is uppermost as opposed to a trifling value of *CR* for single-vehicle crash. The risk of multiple-vehicle crash can rush up to an exceedingly high level if traffic congestion gets to be much heavy. In consideration of the *CR* characteristics reflected in Figure 5.5 and Figure 5.6 together, congested flow can be generally regarded as a distinct traffic condition affecting crash characteristics in different ways from uncongested flow.



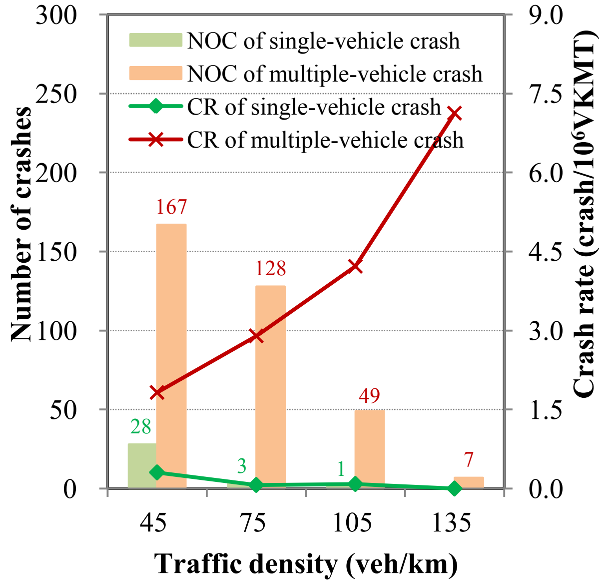


Figure 5.5 *CR* model by expressway type in congested flow

Figure 5.6 *CR* by crash mode in congested flow (intercity expressway)

5.3 Qualitative Analysis of Explanatory Variables

5.3.1 Categorizing traffic conditions

Figure 5.7 reveals the ways to categorize traffic conditions in terms of the monotonicity of *CR* tendency following k_S. According to the index of critical speed v_c, un- and congested flows are firstly differentiated. Then, based on the value of $k_S(CR_{min})$ in Table 5.1, k_S of 25veh/km is regarded as the boundary between low- and high-density uncongested flow.

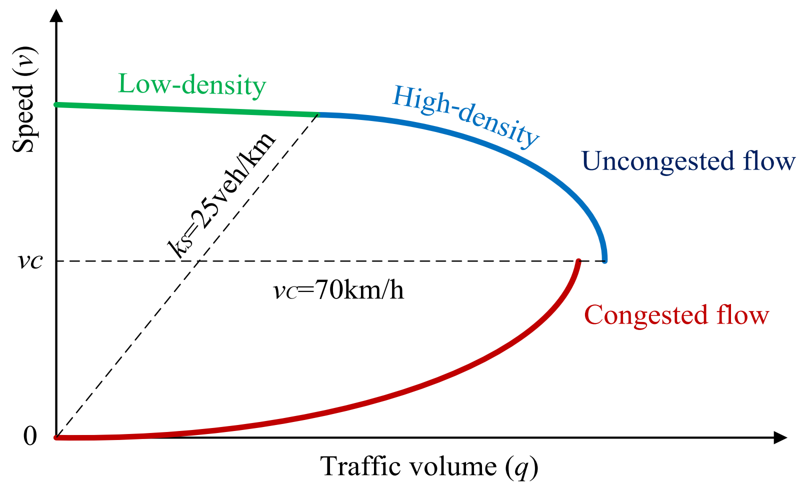


Figure 5.7 Categorizing traffic conditions on intercity expressway

5.3.2 Introduction of variables

In view of the comparability by expressway type, the similar types of explanatory variables for the analysis on urban expressway are selected for PCA on intercity expressway. Table 5.3 summarizes the statistical results of these variables.

Table 5.3 Statistics of explanatory variables (intercity expressway)

Variables	Statistics*			Description
	Max.	Ave.	Min.	
k_S	150	35	1	Traffic density (veh/km)
v_S	106.5	73.2	4.86	Average speed (km/h)
I_{CF}	523.3	71.3	0.59	Index of centrifugal force (km^3/h^2)
h	24.2	4.85	0.00	Variation in road elevation (m)
I_{SD}	1045.4	69.9	0.00	Index of horizontal displacement (m^2)
Pave	F(1)=19.9%			=1 if wet pavement, 0 otherwise
Light	F(1)=41.2%			=1 if nighttime, 0 otherwise
Day	F(1)=37.5%			=1 if holiday, 0 otherwise

* Max./Ave./Min.: the maximal/average/minimal values in statistics; F: frequency.

5.3.3 PCA results by traffic conditions

In low-density, high-density uncongested flow and congested flow, 765/445/383 crash samples can be extracted, respectively.

1) Low-density uncongested flow

PCA results in low-density uncongested flow are summarized in Table 5.4. Meanwhile, with the purpose to investigate the differences of influencing factors by expressway type, the related results for urban expressway are also provided.

Significant influencing factors in turn are found to be geometric variation (I_{CF}, I_{SD} and h), traffic density k_S along with nighttime, speed v_S coupled with wet pavement and holiday. Compared to urban expressway, h becomes one variable related to the 1st component on intercity expressway, while the geometric variations are less significant relative to urban expressway in terms of the % of variance. Besides, *Day* type gets to be a significant crash influencing factor while it is not on urban expressway.

Table 5.4 Principal components in low-density uncongested flow

Expressway types	Items	Principal components*							
		1st		2nd		3rd		4th	
		V	L	V	L	V	L	V	L
Intercity expressway	Value	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	
Urban expressway	Value	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	

*V: variable; L: loading.

2) High-density uncongested flow

Table 5.5 offers the results in high-density uncongested flow. Similar to urban expressway, the significance of geometry affecting crashes on intercity expressway is decreasing. Meanwhile, k_S and v_S also belong to the same principal component. Besides, crashes are prone to occur for the increase in k_S as opposed to the decrease in k_S in low-density uncongested flow, as a reflection of the *CR* model that is convex downward to k_S.

Table 5.5 Principal components in high-density uncongested flow

Expressway types	Items	Principal components *							
		1st		2nd		3rd		4th	
		V	L	V	L	V	L	V	L
Intercity expressway	Values	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	
Urban expressway	Values	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	

*V: variable; L: loading.

On intercity expressway, h is still a significant index of geometric features. Furthermore, the significance of traffic flow related to crashes gets up in terms of its % of variance in comparison with urban expressway. Another difference by expressway type is pavement condition that is significant on intercity expressway.

3) Congested flow

In the same way, the related PCA results on intercity expressway in congested flow are demonstrated in Table 5.6. Crashes are still found out to be associated with higher k_S. Besides, compared to uncongested flow, traffic-related variables, *e.g.,* k_S and v_S become most important. On the contrary, the significances of geometric variations are further decreasing in view of the % of variance.

Table 5.6 Principal components in congested flow

Expressway types	Items	Principal components *							
		1st		2nd		3rd		4th	
		V	L	V	L	V	L	V	L
Intercity expressway	Values	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	
Urban expressway	Values	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	

*V: variable; L: loading.

By expressway type, h is still a significant geometric index for intercity expressway. In addition, despite the reduced significance in congested flow, geometric features of intercity expressway play a more important role to crashes compared to urban expressway.

5.4 Crash Risk Estimation Model (CREM)

5.4.1 Variable selection

Table 5.7 summarizes the selected variables for crash modeling on intercity expressway. Speed v_S and traffic density k_S are remained in low-density uncongested flow. For other traffic conditions, just v_S is selected on behalf of the variation in traffic characteristics. Concerning geometry, horizontal displacement S and the variation in road elevation h are kept to serve as the alignment variation in horizontal and in vertical, respectively. With respect to ambient conditions, the variables are selected in terms of their significance. Since pavement conditions are still not available for non-crash days, the type of condition is not incorporated in the following crash modeling.

Table 5.7 Independent variables for crash modeling (intercity expressway)

Traffic conditions		Traffic flow	Road geometry	Ambient conditions
Uncongested flow	Low-density	k_S and v_S	S and h	*Light* and *Day*
	High-density	v_S	S and h	*Light* and *Day*
Congested flow		v_S	S and h	*Day*

5.4.2 Development of CREM

The matched case-control samples are designed in the similar way in chapter 4. Since the variables between urban and intercity expressways are not different so much, it is reliable to assume 1:20 as the maximum case-to-control ratio on intercity expressway. Then, a total of 521, 267 and 180 crashes along with their corresponding control samples are designed successfully in low-/high-density uncongested flow and congested flow, respectively.

Table 5.8 Summary of CREM (intercity expressway)

Exposures	Variables	Uncongested flow				Congested flow	
		Low-density		High-density			
		Coef. (β)*	*Sig.*	*Coef.* (β)	*Sig.*	*Coef.* (β)	*Sig.*
Traffic flow	v_S	2.48×10^{-2}	0.016	-3.06×10^{-2}	0.000	5.02×10^{-2}	0.000
	k_S	-1.13×10^{-2}	0.035				
Geometric features	S	1.09×10^{-2}	0.000	8.57×10^{-3}	0.000	2.62×10^{-3}	0.000
	h**	$\pm9.65\times10^{-3}$	0.018	$\pm1.92\times10^{-2}$	0.037	$\pm3.77\times10^{-2}$	0.028
Ambient	*Light*	1.12×10^{-1}	0.022	1.02×10^{-1}	0.000		
	Day	1.24×10^{-1}	0.007	1.84×10^{-1}	0.033	2.18×10^{-1}	0.028
Intercept value		-2.60	0.000	1.75	0.000	-2.25	0.033
Model test	-2Log likelihood	3174.6		875.9		320.7	
	Chi-square	166.61		160.16		142.86	
	Significance	0.000		0.000		0.005	
Pseudo R^2	Cox and Snell	0.700		0.683		0.663	
	Nagelkerke	0.933		0.910		0.884	
	McFadden	0.868		0.828		0.785	

* *Coef.*: estimated coefficient.
** +/- correspond to upgrade and downgrade, respectively.

Through focusing on the three traffic conditions, conditional logistic regression is employed for CREM, and the results for basic segments of intercity expressway is summarized in Table 5.8. Meanwhile, the test about coefficients and model are also provides, and it indicates that the CREM is of statistical significance, and all of the independent variables are significant at 95% level (*Sig.* <0.05). However, the statistical significance of CREM in congested flow still seems poor in contrast to uncongested flow, which may be also related to the insufficiency of variables.

5.4.3 Model validation

In the same way in Chapter 4, the predictive performance of CREM developed above is validated as well. The results are summarized in Table 5.9. The accuracy for predicting crash events is found out to be 82.5%, 78.1% and 71.4% for low-density, high-density uncongested flow and congested flow, respectively. Besides, there are still 25.2%, 27% and 31.1% false alarm rates for the three traffic conditions, respectively.

Table 5.9 Validation results of the developed CREM (intercity expressway)

Samples		Predictive performance			
		(a) Odds≥1 (hazardous)	(b) Accuracy $\frac{(a)}{(a)+(c)} \times 100\%$	(c) Odds<1 (not hazardous)	(d) False alarm rate $\frac{(a)}{(a)+(c)} \times 100\%$
Low- density uncongested	0 (non-crash)	2214		6586	25.2
	1 (crash)	363	82.5	77	
High- density uncongested	0 (non-crash)	1030		2790	27.0
	1 (crash)	150	78.1	42	
Congested	0 (non-crash)	565		1255	31.1
	1 (crash)	65	71.4	26	

The functions of CREM between urban and intercity expressways are actually similar each other, while their predictive performances are found out to be different by expressway type. Compared to urban expressway, the goodness-of-fit of CREM for intercity expressway is somehow poor. For the fundamental characteristics of original datasets, the relativity may be resulted by the longer space intervals of detector locations on intercity expressway. In one coverage area of detectors, traffic characteristics are virtually varied by locations while only one detector is available to collect traffic data. As a result, a biased estimation of traffic characteristics for pre-crash conditions would be caused. Since the average coverage area is about 2km intercity expressway, longer than that for urban expressway in 0.5km, and thus, such kind of bias may be more serious for intercity expressway.

5.5 Contributions of Explanatory Factors on Crash Risk

The coefficients of individual variables involved in CREM are demonstrated in Table 5.10 for the three traffic conditions. Likewise, the contribution of one unit of each variable in favor of the *Odds* of crash occurrence can be detected based on that table. The relative contributions between variables are also required to consider the domains of variability of individual variables, since these variables are represented in two types of values: *Light* and *Day* are two dummy variables, and others are described by continuous values.

Table 5.10 Model parameters for evaluating contributions (intercity expressway)

Variable*	Uncongested flow		Congested flow
	Low-density	High-density	
v_S	0.0248	-0.0306	0.0502
k_S	-0.0113		
S	0.0109	0.00857	0.00262
h	±0.00965	±0.0192	±0.0377
*Light**	0.112	0.102	
*Day**	0.124	0.184	0.218

* Light and Day are dummy variables, while others are continuous variables.

5.5.1 Low-density uncongested flow

As shown in Table 5.10, CREM in low-density uncongested flow involves six variables: speed v_S, traffic density k_S, horizontal displacement S, the variation in road elevation h and ambient *Light* as well as *Day* type. In terms of the coefficients, except of k_S, which is negatively affecting crash risk, other variables have positive contributions on the *Odds* of crash occurrence. Concerning the model parameters in Table 5.10 and the figures shown in Appendix A, the contributions of individual variables on crash risk can be found out to be characterized by the following features:

- Geometry: *Odds* value is on the rise with the increases in S and h, while the contributions of both variables are not as different as those for urban expressway.

- Traffic flow: higher v_S would give rise in crash risk as opposed to the decrease in *Odds* value with the increase in k_S. Furthermore, the contribution of v_S in favor of a crash is more significant compared to k_S.

- Ambient Conditions: both *Light* and *Day* are involved in low-density uncongested flow and the traffics in nighttime/on holidays are generally of higher crash risk than the traffics in daytime/on weekdays, respectively.

Table 5.11 Maximum safe speed proposed (km/h)
(Low-density uncongested flow on intercity expressway)

Horizontal displacement (m)			0	10	20	30	40	50
Week-day	**Daytime**	k_S=10veh/km, h=0m	108	104	100	96	92	88
		k_S=10veh/km, h=10m	104	100	96	92	88	84
		k_S=10veh/km, h=20m	100	96	92	88	84	80
		k_S=20veh/km, h=0m*				102	98	94
		k_S=20veh/km, h=10m			102	98	94	90
		k_S=20veh/km, h=20m		102	98	94	90	86
	Nighttime	k_S=10veh/km, h=0m	104	100	96	92	88	84
		k_S=10veh/km, h=10m	100	96	92	88	84	80
		k_S=10veh/km, h=20m	96	92	88	84	80	76
Holiday	**Daytime**	k_S=10veh/km, h=0m	104	100	96	92	88	84
		k_S=10veh/km, h=10m	100	96	92	88	84	80
		k_S=10veh/km, h=20m	96	92	88	84	80	76
		k_S=20veh/km, h=0m			100	96	92	88
		k_S=20veh/km, h=10m		100	96	92	88	84
		k_S=20veh/km, h=20m	100	96	92	88	84	80
	Nighttime**	k_S=10veh/km, h=0m	98	94	90	86	82	78
		k_S=10veh/km, h=10m	94	90	86	82	78	74
		k_S=10veh/km, h=20m	90	86	82	78	74	70

* The operating speed at the density of 20veh/km is mostly below 100km/h.
** Traffic density in nighttime is often smaller than 10veh/km.

On intercity expressway, the vehicle composition is characterized by higher HV% in contrast to urban expressway. As a result, the contribution of *h* on crash risk is increased on that type of expressway compared to urban expressway. By the same reason, the effect of v_S gets more important regarding the speed inharmonization of vehicle types, which would be more serious with the increase in average speed. As stated before, holiday traffic on intercity expressway is more distinct considering the vehicle composition and driver population. Compared to urban expressway, it is self-evident that *Day* type is an added involved variable on intercity expressway. As indicated by these findings above, the high speed coupled with frequent variation in geometric design and holiday/nighttime is an important crash influencing factor in low-density uncongested flow. In this sense, the way to remind driver attention (*e.g.,* driver warning system) and to control speed (*e.g.,* variable speed limit) can be regarded as two effective traffic management measures. In this purpose, the maximum safe speeds for individual driving conditions are proposed in Table 5.11.

5.5.2 High-density uncongested flow

In high-density uncongested flow, the involved variables for CREM include speed v_S, the horizontal displacement *S*, the variation in road elevation *h*, ambient *Light* and *Day* type. The following characteristics about the contributions of individual variables on crash risk can be observed through referring to Table 5.10 and the figures in Appendix A:

- Geometry: although *S* and *h* have positive contributions on the *Odds* value, the contribution of *S* in favor of a crash is reduced as opposed to the increased importance of *h* with the increase in traffic density.

- Traffic flow: with the transformation of traffic conditions, v_S starts to have a negative contribution on *Odds* of crash occurrence. Meanwhile, its effect on crash risk gets more important in terms of the absolute value of coefficient.

- Ambient conditions: both *Light* and *Day* still have positive contribution on crash risk. Compared to low-density uncongested flow, the effect of *Light* is slightly reduced by contrast to an increase in significance of *Day* type.

With the increase in traffic density, the contribution of *S* on crash risk is decreased due to the reduced speed as discussed in Chapter 4. By contrast, the contribution of *h* is actually increased, since the interruption of HV to the surrounding traffic would be stronger with the intensification of the inter-vehicle interaction. For the same reason, the increase in traffic density, in other word, the decrease in speed may result in the rise in crash risk. Combining the contribution of traffic conditions to crash risk in low-density uncongested flow, the above findings can also account for the cause of the U-shaped *CR* model in uncongested flow, as shown in Figure 5.2. As stated before, nighttime may aggravate crash risk of driving conditions due to the reduced visibility, and on holidays, the inappropriate reaction of drivers to the variation in traffic conditions may be more frequent, since there may be more drivers unfamiliar with traffic conditions on those days.

5.5.3 Congested flow

With the further increase in traffic density, congested flow is emergent and characterized by traffic oscillation. In such traffic condition, the developed CREM incorporates four variables; speed v_S, horizontal displacement *S*, the variation in road elevation *h* and *Day* type. Although ambient *Light* is not involved, it is still not exact to conclude that the effect of ambient *Light* on crashes is not significant, considering the limited samples of congested flow and crash events in nighttime. By referring to Table 5.10 and the related figures in Appendix A, the contributions of variables on crash risk can be achieved as follows:

- Geometry: compared to uncongested flow, the contribution of *S* on crash risk is continuously reduced and the related contribution of *h* is further on the rise.
- Traffic flow: v_S gets a positive contribution on crash risk again in congested flow and its coefficient is increasing significantly in contrast to uncongested flow.
- Ambient conditions: the traffic on holidays is less safe than that on weekdays, and the rising trend of *Odds* value induced by the alteration of *Day* type is more distinct relative to uncongested flow.

The contribution of *S* on crash risk is continuously decreasing due to the further reduced operating speed with the variation in traffic conditions. On the contrary, the sensitivity of v_S to crash risk is increased and the contribution of *h* on crash risk becomes greater in congested flow, since the inter-vehicle interaction gets more powerful and the interruption of HV to the surrounding traffic becomes stronger in congested flow. On holidays, owing to the serious fluctuation in traffic conditions, the inappropriate reaction of unfamiliar and inexperience drivers to the variation in traffic conditions may be much more frequent.

According to the above analyses, the contributions of individual variables with the variation in traffic conditions on intercity expressway can be achieved. With the increase in traffic density, the contribution of the horizontal geometry on crash risk would also be reduced. By contrast, as the intensification of the inter-vehicle interaction, the effect of vertical geometry on crash risk is on the increase. Besides, the sensitivity of operating speed to crash risk get higher while traffic density is increasing. In addition, *Day* type and ambient *Light* are not negligible influencing factors, since holiday and nighttime could increase crash risk relative to weekday and daytime, respectively.

5.6 Different Effects of Crash Influencing Factors by Expressway Type

Expressway type is also another aspect affecting crash characteristics while its impact on crash risk has not been considered in the above models. In order to more comprehensively understand the cause of crash occurrence, this part focuses on the comparison of crash risk and its influencing factors between the two types of expressways.

5.6.1 Low-density uncongested flow

Based on the speed-traffic density relationships (Figures 6.4 and 6.5), 10veh/km is regarded as an accepted traffic density for urban and intercity expressways in low-density uncongested flow. The overlapped scope of speed is correspondingly from 80 to 100km/h.

1) Analysis on geometry and traffic flow

Figure 5.8 appears the variation in *Odds* value with the increase in speed v_S, horizontal displacement S and the variation in road elevation h on urban and intercity expressways, leaving out ambient conditions. Generally, crash risk is on the rise when each of the indices increases. On urban expressway, the maximum value of S is nearly one time greater than that on intercity expressway. As a result, its maximum *Odds* value is approximately 1.8 times of that on intercity expressway. Within the same area of S, there are primarily three differences: 1) the rising trend of crash risk following S on urban expressway is slightly more distinct compared to intercity expressway at the speed of 80km/h; 2) with the increase in speed, a more notable increase in *Odds* value exists on intercity expressway, and finally, crash risk becomes to be 10%~15% higher than that on urban expressway at the speed of 100km/h; 3) at the same speed, the rise in *Odds* value induced by the increase in *h* is slightly more dramatic on intercity expressway: when *h* increases from 0 to 10m, the increased scope of Odds value is about 20% higher than that on urban expressway.

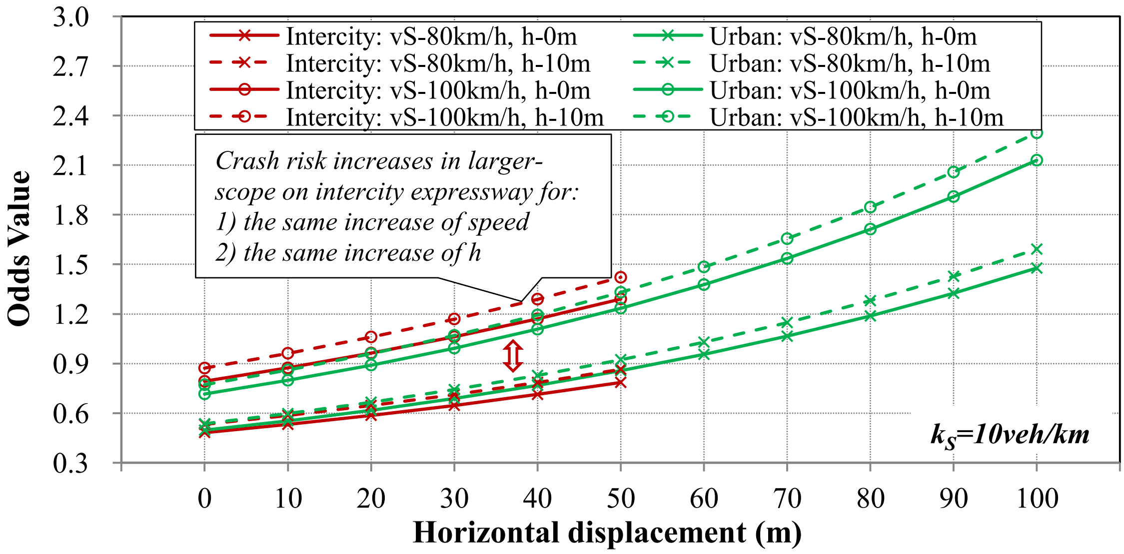


Figure 5.8 Analysis of geometry and traffic flow in low-density uncongested flow

The larger *S* is and the poorer design consistency is. As a result, more frequent speed reduction and driving direction adjustment would be induced. Even for the same scope of *S*, on urban expressway, the drivers' loading may be heavier due to the narrower cross section layout. Furthermore, in contrast to intercity expressway, its higher roadside barrier may aggravate the restriction of visibility in curve sections.

On the other side, when speed is high, the inter-vehicle interaction would be not negligible, since long inter-vehicle spacing is necessary for safety. Generally, the operating speed of HV is often below 100km/h. In such case, the vehicles driving at high speed have to take risk-avoiding behaviors if drivers feel the interruption of the downstream HV. As stated before, compared to urban expressway, the vehicle composition on intercity expressway is characterized by high HV%. Furthermore, low-density uncongested flow mostly exists in nighttime, when HV% is much higher than that on daytime. It may imply there are more frequent risk-avoiding behaviors on intercity expressway than that on urban expressway. In this case, it is better to control speed considering HV% on intercity expressway.

Meanwhile, since vertical alignment is highly related to the behaviors of HV, the increase in *h* would be more sensitive to the rise in crash risk on intercity expressway.

2) Analysis focusing on ambient conditions

In Figure 5.9, the variation in crash risk with the alteration of ambient *Light* is shown. Due to the reduced visibility, crash risk in nighttime is obviously higher than the value in daytime. Regarding expressway type, with the alteration of ambient *Light*, the rising scope of *Odds* value is much larger on urban expressway in contrast to intercity expressway. In brief, it seems that nighttime is more sensitive to crash risk on urban expressway.

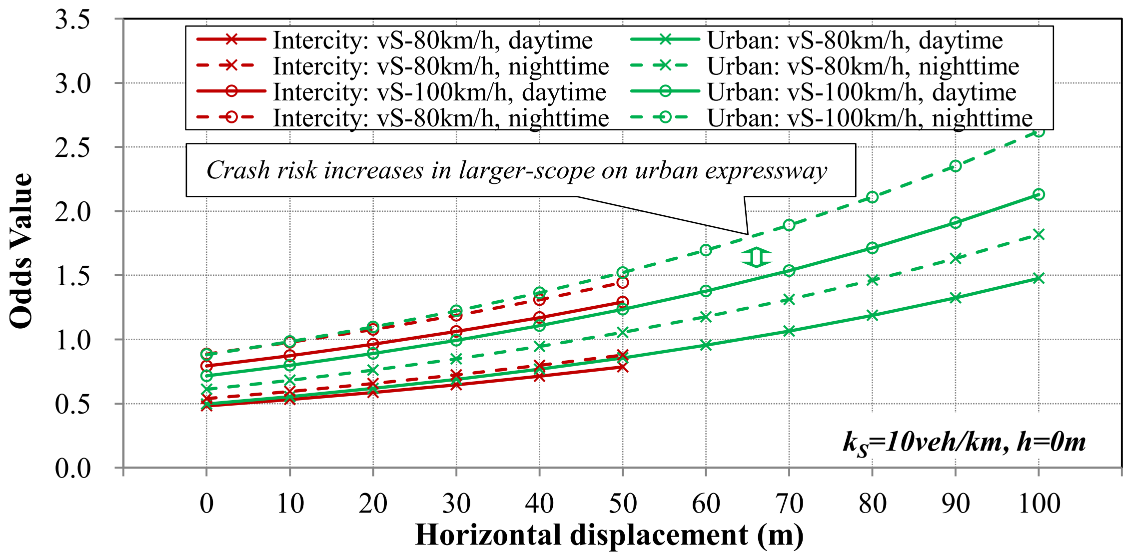


Figure 5.9 Analysis focusing on ambient light in low-density uncongested flow

Although the exact causes remain absent (small sample-size or the conflict between serious visibility reduction and narrow cross section), a higher speed variation indeed exists in nighttime on urban expressway with comparison to intercity expressway. Crash occurrence is significantly associated with the turbulence of traffic conditions. Thus, the distinct speed variation would be a cause of high crash risk on urban expressway in nighttime.

5.6.2 High-density uncongested flow

In high-density uncongested flow, the overlapping scope of speed on urban and intercity expressways is regarded from 70km/h to 90km/h (Figures 6.4 and 6.5).

1) Analysis on geometry and traffic flow

When traffic density increases, the inter-vehicle interaction gets intensive. The variation in *Odds* value with the combined effects of v_S, S and h is exhibited in Figure 5.10, and the following characteristics can be observed by comparing with low-density uncongested flow: 1) the rising trend of *Odds* value gets lower due to the reduced effect of horizontal geometry on crashes; and 2) owing to the negative effect of speed, the increase in speed can result in a drop in crash risk. By expressway type, the rising trend of *Odds* value on intercity expressway is generally steeper than that on urban expressway. Furthermore, when h increases from 0m to 10m, *Odds* value is on a higher rise on intercity expressway in contrast to urban expressway. In conclusion, traffic in high-density uncongested flow on intercity expressway is less safe relative to urban expressway, which can accounts for the feature that higher *CR* exists on intercity expressway as shown in Figure 5.2.

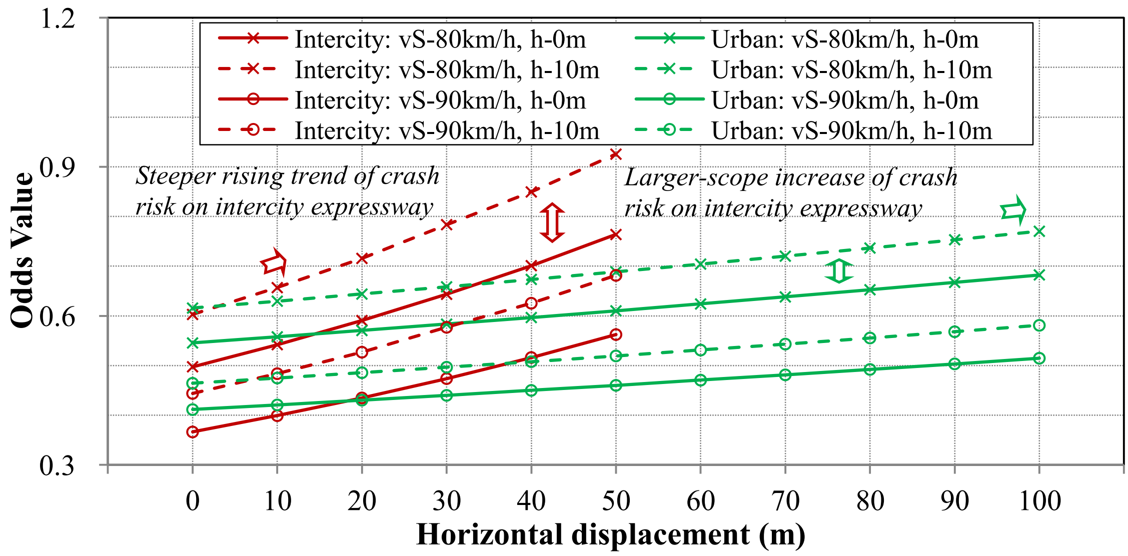


Figure 5.10 Analysis of geometry and traffic flow in high-density uncongested flow

Vehicle composition is a potential cause of the above crash risk characteristics. Because of the serious interruption of HV to its surrounding traffic, the upstream vehicles may take avoiding behaviors, *e.g.,* slow-down and lane changing, in an effort to keep sufficient spacing from HV for the sake of safety. Those behaviors often result in instability in traffic conditions. In comparison with urban expressway, there is higher HV% on intercity expressway, and thus, the instability in traffic conditions may be much more serious. In the meantime, vertical slope is highly related to the behaviors of HV, and the increase in h may worsen the driving conditions for HV. In this sense, the characteristic of higher HV% may play an important role to the higher crash risk on intercity expressway relative to urban expressway in high-density uncongested flow. In this regard, the measures to open the special lane for HV (HVL) or variable speed limits (VSLs) may be effective in achieving speed harmonization among various lanes (Duret *et al.,* 2012), in the purpose to reduce crash risk on intercity expressway.

2) Analysis focusing on ambient *Light*

Figure 5.11 reveals the variation in *Odds* value with the alteration from daytime to nighttime, and it still reveals that traffic flow in nighttime get less safe relative to daytime. By expressway type, the rising trend of crash risk is more notable on urban expressway.

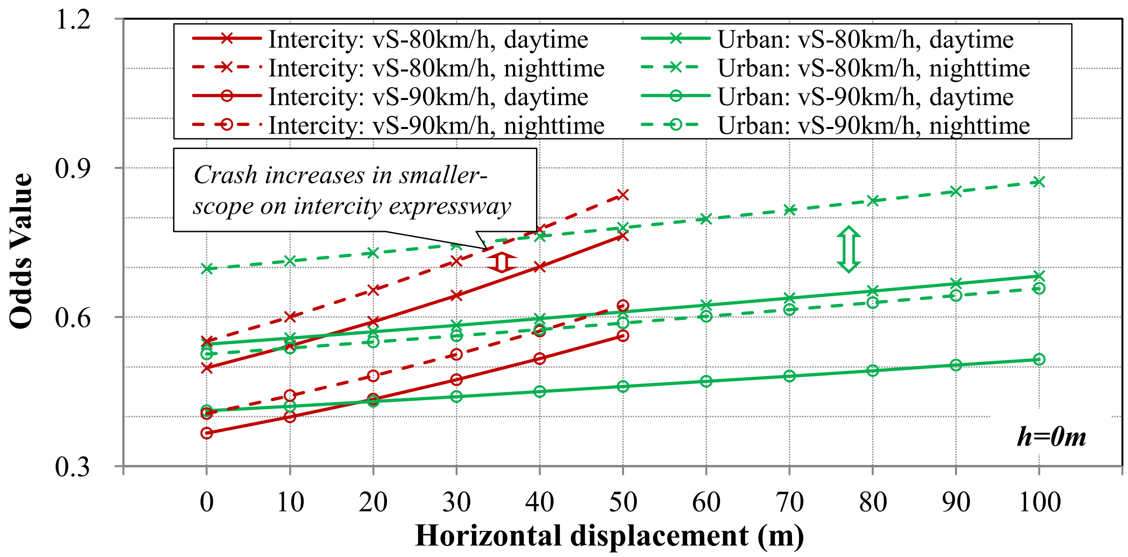


Figure 5.11 Analysis focusing on ambient *Light* in high-density uncongested flow

On urban expressway, high-density uncongested flow in nighttime is often during the evening, which is also characterized by high speed variance (Figure 3.14). In practice, the traffic seems to be the extension of evening peak traffic, and the share of private travel is high, such as homeward-bound and recreational travels. In this case, higher operating speed is expected, as reflected by the slight increase in speed. By contrast, the characteristics are not obvious on intercity expressway, as illustrated in Figure 3.14.

3) Analysis focusing on *Day* type

Day type is another influencing factor and Figure 5.12 illustrates its different effects on crash risk between urban and intercity expressways. During the alteration from weekday to holiday, the increasing tendency of crash risk is more remarkable on intercity expressway.

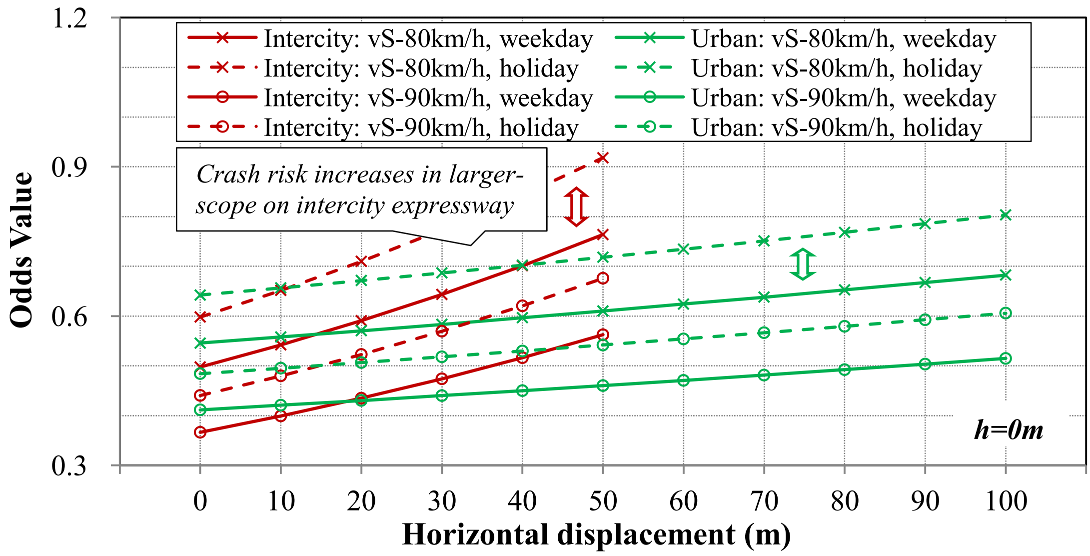


Figure 5.12 Analysis focusing on *Day* type in high-density uncongested flow

Generally, holidays refer to a large increase in recreational private travel, which may result in more LDT, more travels in unfamiliar conditions compared to normal weekdays (Anowar *et al.,* 2013). Necessarily, crash risk on holidays is higher than the risk on normal weekdays. Regarding expressway type, since intercity expressway is the main artery of LDT, the above characteristics of holiday traffic different from weekdays may be more distinguishing on that type of expressway in contrast to urban expressway.

5.6.3 Congested flow

With the further increase in traffic density, congested flow appears. As stated before, the average traffic variables are insufficient to reveal the natural traffic characteristics in congested flow where traffic oscillation is the typical state characterized by recurring decelerations followed by accelerations. In this regard, the reliability of CREM for congested flow deserves to be improved by using the data of short-term turbulence of traffic flow in the future. To serve as a modest spur for the future work, the characteristics of crash risk by expressway type are roughfully identified. On both types of expressways, the overlapping scope of speed is observed below 60km/h, while un- and congested flows are bounded at 70km/h on intercity expressway.

1) Analysis on geometry and traffic flow

The tendencies of *Odds* value following the interaction of v_S, S and h are described in Figure 5.13. One distinct difference from uncontested flow is the much milder rise in crash risk with the increase in S, as a result of the further reduced effect of geometry on crashes. As opposed to S, the change of v_S gets more sensitive to crash risk: when speed decreases from 60km/h to 50km/h, the reduction in *Odds* value can be up to one unit. Furthermore, the rising trend of *Odds* value induced by the increase in h gets more notable in congested flow relative to uncongested flow. Regarding expressway type, as stated before, congested flow on intercity expressway actually starts at the speed of 70km/h. According to the tendency shown in Figure 5.13, it is considered reliable that congested flow would be in higher-risk on intercity expressway. Caused by the same increase in h, the rising trend of *Odds* value is more prominent on intercity expressway in contrast to urban expressway.

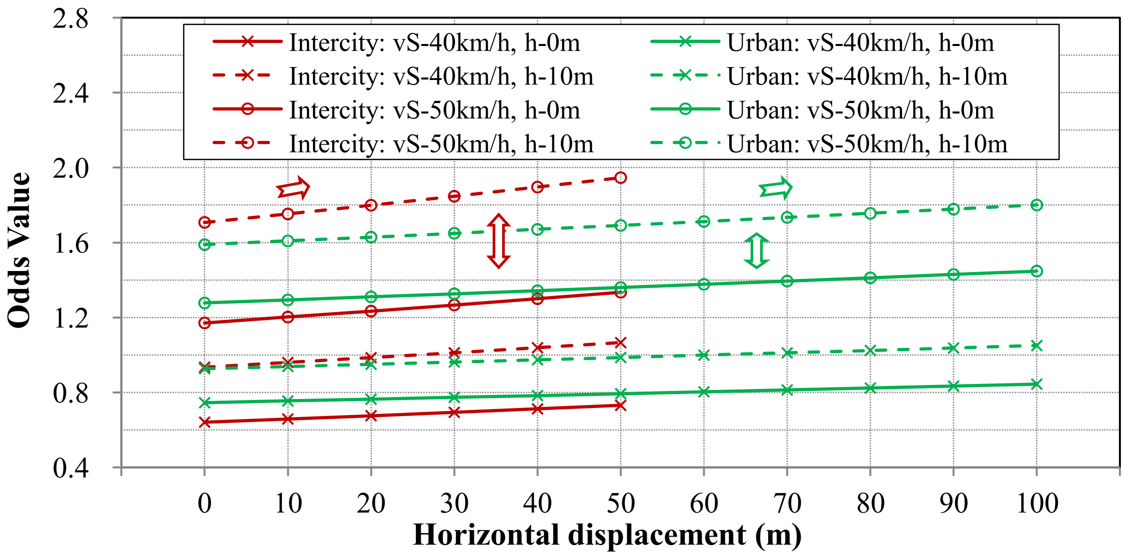


Figure 5.13 Analysis of geometry and traffic flow in congested flow

In congested flow, the inter-vehicle interaction gets much more intensive, and thus the instability in traffic flow induced by HV may be more serious on intercity expressway compared with urban expressway in view of their vehicle compositions. Meanwhile, as concluded in Chapter 3, high HV% and steeper slope design are two distinct features of intercity expressway different from urban expressway. In this regard, HV may throw much serious interruption to the surrounding traffic when that vehicle drives on a steep slope. In such case, the measures to control HV (*e.g.,* driving ban for trucks (DBTs)) or to open HVL are two reliable ways to improve traffic safety (Wu *et al.,* 2012b).

However, the above findings seem to be in conflict with the *CR* model in congested flow, which indicates an increasing tendency of *CR* following the decrease in speed. One possible reason is, as demonstrated in Figure 5.4, most crashes occur at the early stage of congested flow, where average speed is actually still high. As introduced in Chapter 2, unequal dispersion of crash events would result in biased estimation on explanatory factors. On the other side, by experience, traffic oscillation may be reduced with the growth of traffic congestion, which is combining with the decrease in speed and the increase in traffic density simultaneously. In this sense, it is considered reliable that crash risk would be reduced with the decrease in speed in congested flow.

2) Analysis focusing on ambient conditions

Figure 5.14 indicates the different effects of *Day* type on crash risk between urban and intercity expressways. On both types of expressways, holiday is found out to play a significant role to the rise in crash risk in congested flow. Compared to urban expressway, the increase in crash risk with the alteration from weekday to holiday can be observed more significant on intercity expressway.

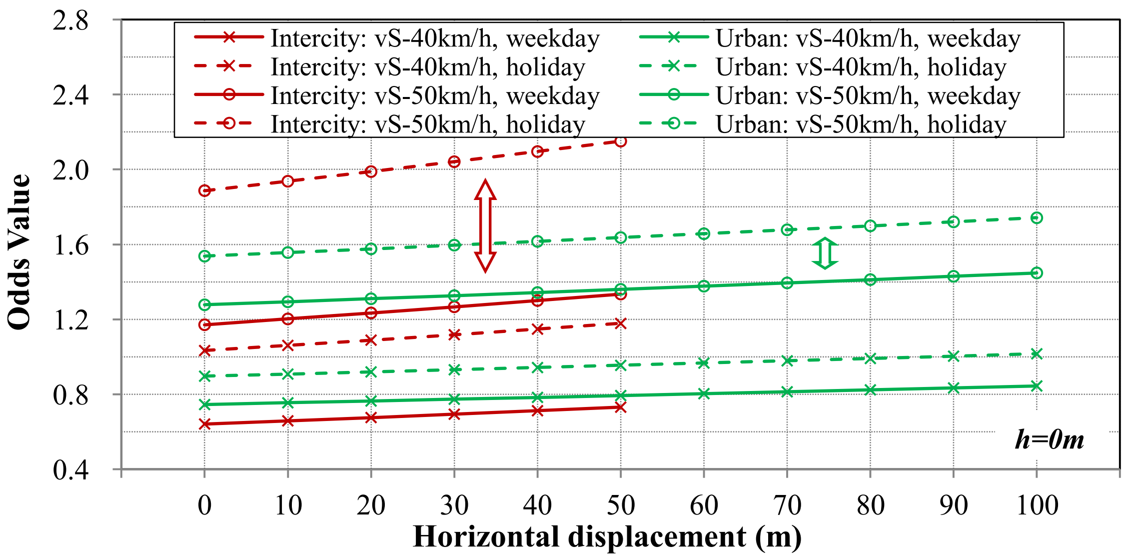


Figure 5.14 Analysis focusing on *Day* type in congested flow

5.7 Summary

This chapter developed a CREM at basic segments of intercity expressway, through following the similar process in chapter 4. The model was found to be significant in statistics and of accepted goodness-of-fit, with 82.5%, 78.1% and 71.4% of predictive performance in low-, high-density uncongested flow and congested flow, respectively. Meanwhile, crash characteristics between urban and intercity expressways were identified in terms of *CR* and principal components. In order to reveal the causes of the different crash characteristics, the impacts of explanatory variables on crash risk were comparatively analyzed between the two types of expressways. The results are summarized in Table 5.12.

Table 5.12 Relative significances of variables by expressway type

Exposures	**Variables**	**Traffic categories**					
		Low-density uncongested		High-density uncongested		Congested	
		Urban	Intercity	Urban	Intercity	Urban	Intercity
Traffic flow	v_S	++* (80km/h)	++ (100km/h)	+	++	+	++
	k_S	-- (80km/h)	-- (100km/h)				
Geometric variation	*S*	++	+	+	++	+	++
	h	+	++	+	++	+	++
Ambient conditions	*Light*	++	+	++	+	++	+
	Day	+	++	+	++	+	++

* +/- means the contribution of each variable to crash risk; ++/-- refers to the relatively significant effect of the same variable to crash risk by expressway type.

In low-density uncongested flow, geometric design is a major cause leading to different crash risk characteristics by expressway type. The compacted design, *e.g.,* small radii of curves and narrower cross section layout, is significantly related to higher crash risk on urban expressway. At much high speed, the inter-vehicle interaction cannot be negligible considering the speed inharmonicity between HV and other vehicles. As a result, much high speed-driving on intercity expressway gets less safe than that on urban expressway.

In high-density uncongested flow, the influence of traffic flow becomes more important as opposed to a decreasing significance of horizontal geometry affecting crashes. Since the inter-vehicle interaction gets intensive, the interruption of HV to its surrounding traffic plays an important role to crash risk. Regarding the vehicle composition *e.g.,* higher HV%, traffic condition on intercity expressway is less safe compared to urban expressway.

In congested flow, the significance of horizontal geometry is further reduced, while the variation in speed gets much more sensitive to crash risk, in contrast to uncongested flow. Intercity expressway still has worse traffic safety situation due to the interruption of HV.

As for other variables, the effect of variation in road elevation h on crash risk is more significant on intercity expressway than that on urban expressway, no matter how traffic conditions are, since h is highly related to the behaviors of HV. Nighttime traffic on urban expressway is of higher risk, due to the larger speed variance in nighttime on that type of expressway in contrast to intercity expressway. By contrast, the rise in crash risk with the variation from weekday to holiday is more remarkable on intercity expressway in view of the more distinct holiday traffic on that type of expressway.

Chapter 6

APPLICATIONS OF MODEL FOR PREDICTING THE EVOLVING PROCESS OF CRASH RISK AND MEASURING THE QUALITY OF GEOMETRIC DESIGN

6.1 Introduction

At the view point of traffic management, crash risk prediction is just one application of proactive strategy. For the better applicability of that strategy, understanding the evolving process of crash risk with the variation in traffic conditions is also important, with the purpose to advance the efficiency of traffic control measures. Meanwhile, if the safety benefit of geometric design can be identified through referring to CREM, it is reliable to consider crash risk as an applicable measure of safety evaluation. Besides, once the potential crash-prone sections are identified, drivers traveling in those sections would be warned for attention, and countermeasures including the improvement of geometric design would be adopted in time for safer expressway. Given these points, some applications of the developed model will be demonstrated in this chapter through focusing on the subject basic segment of urban and intercity expressways, respectively. The expected findings are further compared between the two types of expressways, in order to identify the influence of expressway type on crash characteristics.

- Predicting the evolving process of crash risk with the variation in traffic conditions.
- Measuring the safety performance of geometric design and detecting the potential crash-prone locations.

- Identifying the different sensitivities of traffic conditions to crash risk between urban and intercity expressways.

Odaka route and the section of Tomei Expressway from Okazaki I.C to Toyota JCT are involved for the following analyses, where a lot of daily traffic is carried and crashes are frequently occurred. The locations of the two segments in the expressway network around Nagoya are demonstrated in Figure 6.1. To achieve the objective of this chapter, traffic characteristics, geometric design and *CR* statistics of the two test beds are first analyzed. Next, various scenarios of driving conditions are designed through categorizing traffic flow, and crash risk is predicted based on the developed model. Further, the values of crash risk at individual locations are comparatively discussed considering their geometric design. Finally, this study makes a distinction between the two test beds in the purpose to identify the sensitivities of traffic conditions to crash risk by expressway type.

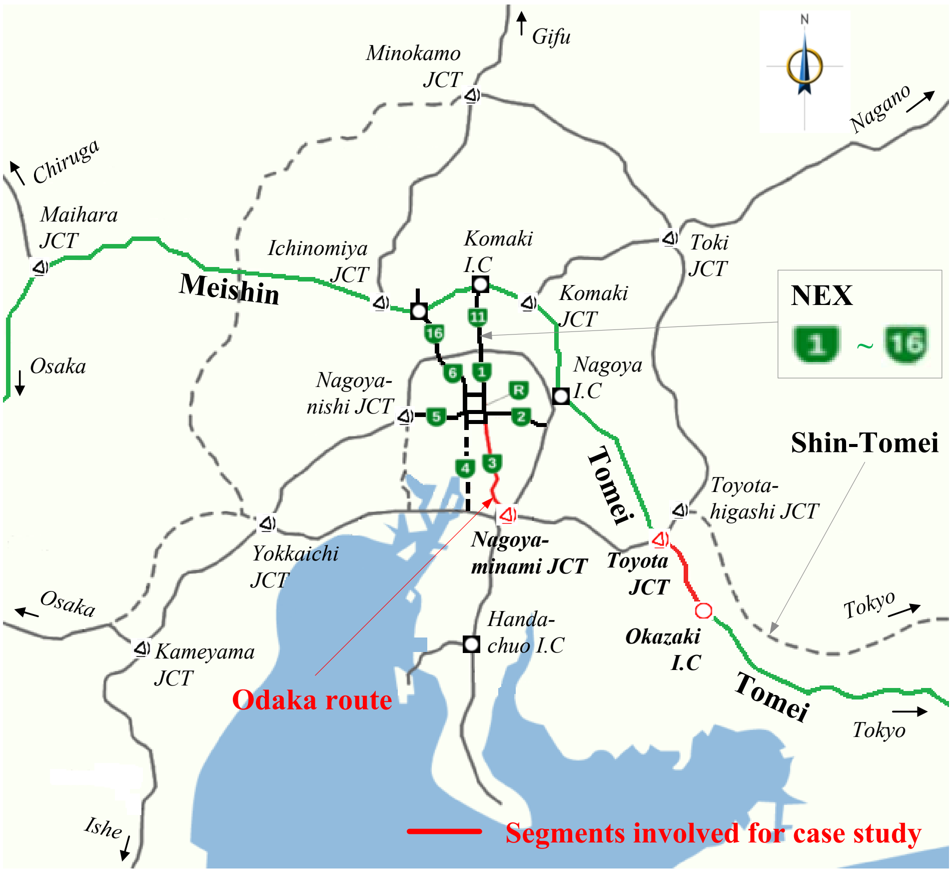


Figure 6.1 Locations of involved segments in expressway network around Nagoya
(Source: Central Nippon Expressway Company Limited and modified by author)

6.2 Test Bed Characteristics

6.2.1 Geometric design

1) Urban expressway

Compared to intercity expressway, access density along the mainline of Odaka route is much higher, as shown in Figure 6.2. In such case, the segment from Kasatera off-ramp to Chita off-ramp, southbound to Nagoya Minami JCT, is selected as the test bed for urban expressway, since the segment between the two ramps is relatively longer in contrast to other two neighboring ramps and the geometric features is more diverse. Figure 6.2 also illustrates the detector locations and the geometric design of the test bed, which is 2-lane roadway and 3.75km in total length. Excluding the influence areas of ramps, a length of 2.75km of basic segment can be extracted, and six ultrasonic detectors are installed in approximately 500m within this distance.

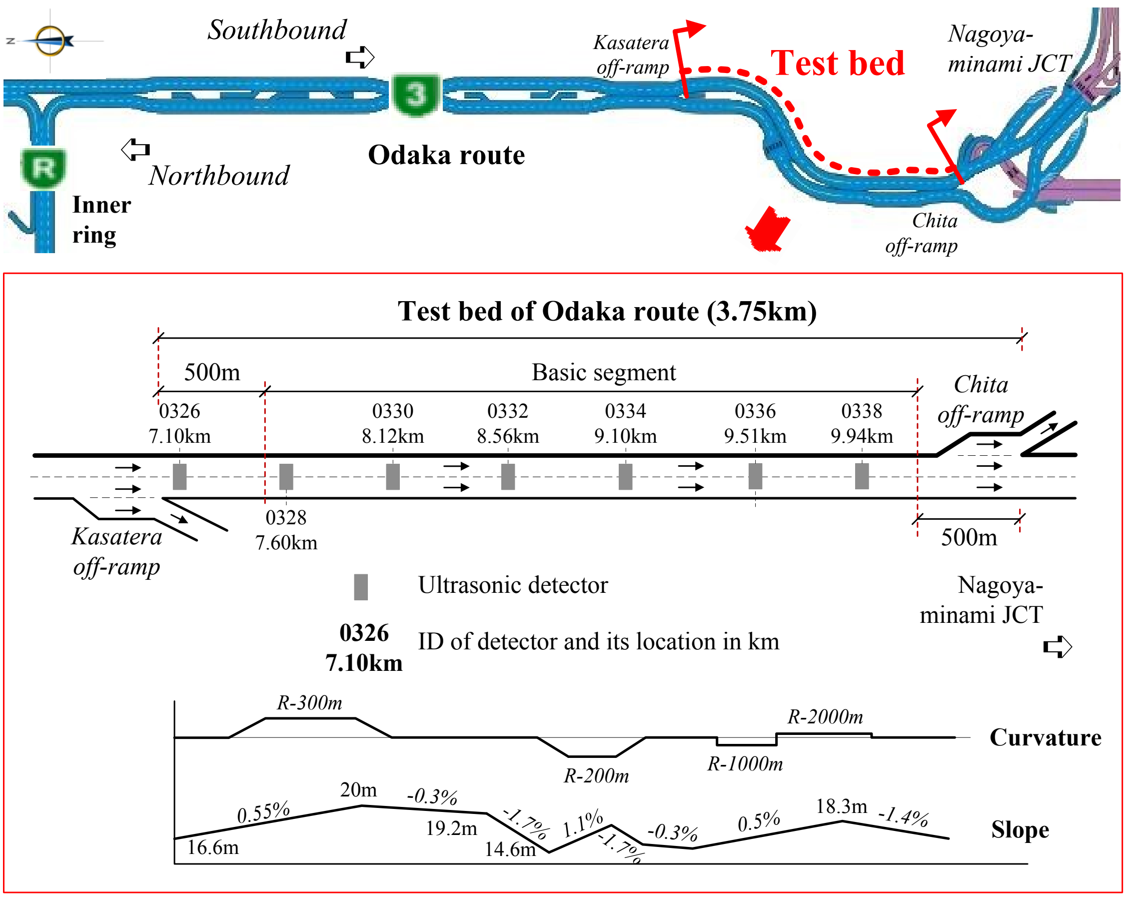


Figure 6.2 Geometric design of the test bed on Odaka route

2) Intercity expressway

As shown in Figure 6.3, the section between Okazaki I.C and Toyota JCT (Nagoya bound, 9.9 km) is selected as the test bed of Tomei Expressway. It is a two-lane roadway and five loop detectors locate in the test bed in spacing of 2km. The nearest detector locations to the two junctions can be regarded as outside the influence of merging/diverging, since each one virtually leaves at least 1km to the corresponding junction.

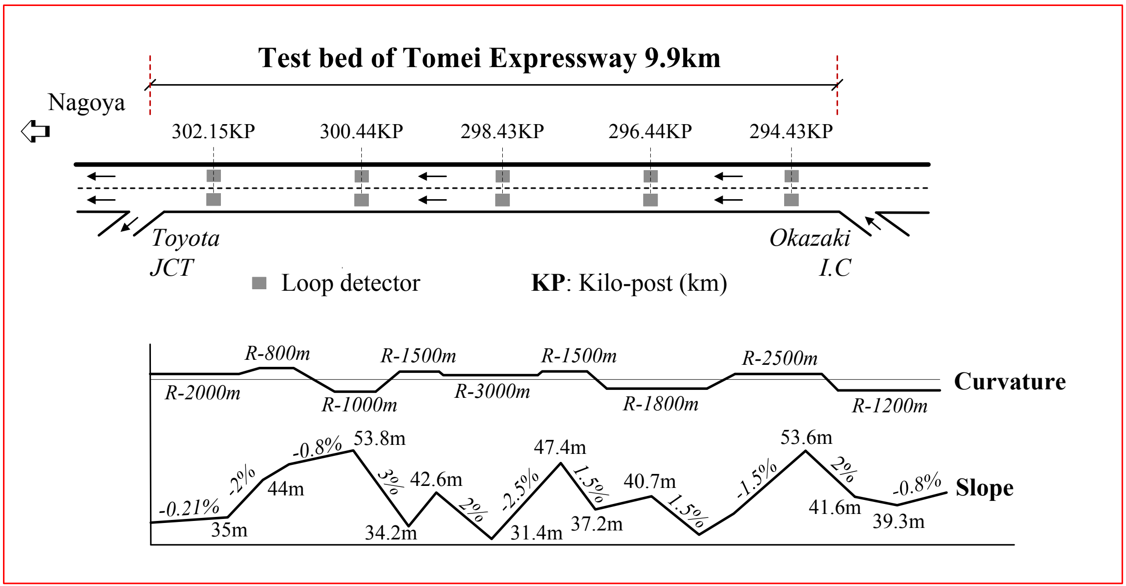


Figure 6.3 Geometric design of the test bed on Tomei Expressway

6.2.2 Speed-traffic density relationships

For the objective of this chapter, various scenarios of driving conditions will be designed. The above discussions indicate that the features of geometric design at individual detector locations are different, thus traffic characteristics may be varied at those locations. As a result, even for the same category of traffic condition, some traffic variables (*e.g.,* speed) would be different each other at individual locations. To support this assumption, this section aims to find the related variables at detector locations for different scenarios.

1) Scenarios of traffic conditions

As discussed in Chapter 3, several studies have proposed traffic density as the service measure of traffic flow for basic segments, considering the insufficiency of speed and the *v/c* ratio. Based on the original datasets, the minimum density is selected as 10vhe/km. Through referring to the previously studies on categorizing traffic flow, five scenarios of driving conditions are designed for uncongested flow as shown in Table 6.1 (HCM, 2010; Wu *et al.,* 2012a, 2012b). For congested flow, just three scenarios are generally designed to demonstrate the variation in traffic congestion, considering the data available for this study and the previous researches, *e.g.,* Brilon and Estel (2009).

Table 6.1 Scenarios of driving conditions

Scenario #	Traffic density (veh/km/2ln)*	Related traffic conditions**
1	10	Low-density uncongested flow
2	20	Low-density uncongested flow
3	30	High-density uncongested flow
4	40	High-density uncongested flow
5	50	High-density uncongested flow
6	70	Congested flow
7	100	Congested flow
8	130	Congested flow

* Min./max. values are decided through referring to the speed-traffic density diagrams in Figure 6.4.
** Association with the categories of traffic conditions (see Figure 4.8 and 5.7).

2) Urban expressway

Detector data in original datasets for crash modeling are also utilized for the following analyses. Here, it worth mentioning that, data based on cross-section is adopted for the two test beds considering the relativity by expressway type. In the following, the data in one month (July, 2009) excluding the invalid data and the data within lane/cross-section closure intervals is employed, and the diagrams of speed-traffic density for the test bed of Odaka route are demonstrated in Figure 6.4.

As expected, the diagrams indeed appear different characteristics at various detector locations. According to these diagrams, the mean value of speed at each location, which corresponds to individual scenarios of driving conditions, can be observed. The results are summarized in Table 6.2. Note that, in contrast to the traffic in uncongested flow, the distribution of traffic flow at the stage of density over 70veh/km is not distinct because of the limited samples. In such case, the related values of speeds for congested flow are estimated in light of the tendencies of these diagrams.

Table 6.2 Speed regarding driving conditions in the test bed on Odaka route

Traffic density (veh/km)	#0328	#0330	#0332	#0334	#0336	#0338
10	78*	75	68	78	70	78
20	74	72	66	76	68	75
30	68	67	65	72	65	71
40	65	63	62	68	61	65
50	62	60	58	64	57	60
70	46	45	44	46	45	45
100	35	35	35	35	35	35
130	25	25	25	25	25	25

* Unit of speed is km/h.

3) Intercity expressway

In the same way, the diagrams of speed-traffic density are investigated in the test bed of Tomei Expressway, as shown in Figure 6.5. Generally, these diagrams are characterized in different features at individual locations. Corresponding to individual scenarios of driving conditions illustrated in Table 6.1, the mean value of speed is also calculated and summarized in Table 6.3. Excluding 294.43KP and 296.44KP, the distributions of traffic flow at the stage of density larger than 70veh/km are also indistinct at other locations. Consequently, the speeds at these locations for congested flow are also estimated values.

Table 6.3 Speed regarding driving conditions in the test bed on Tomei Expressway

Traffic density (veh/km)	294.43KP	296.44KP	298.43KP	300.44KP	302.15KP
10	88*	90	89	88	89
20	86	88	85	87	86
30	80	83	81	82	83
40	72	74	73	74	75
50	62	63	63	69	64
70	44	45	45	45	46
100	35	35	35	35	35
130	25	25	25	25	25

* Unit of speed is km/h.

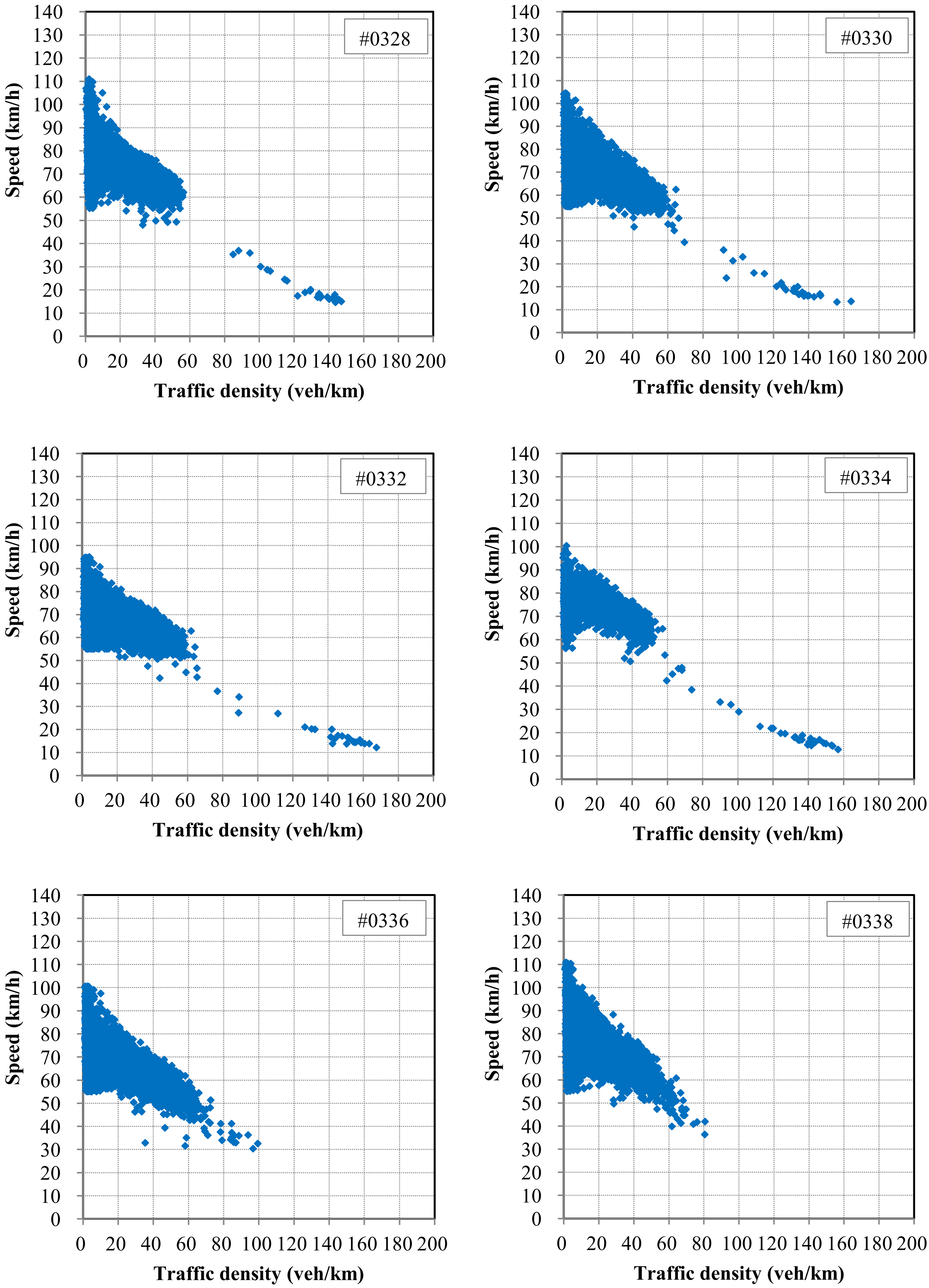


Figure 6.4 Speed-traffic density diagrams in the test bed on Odaka Route

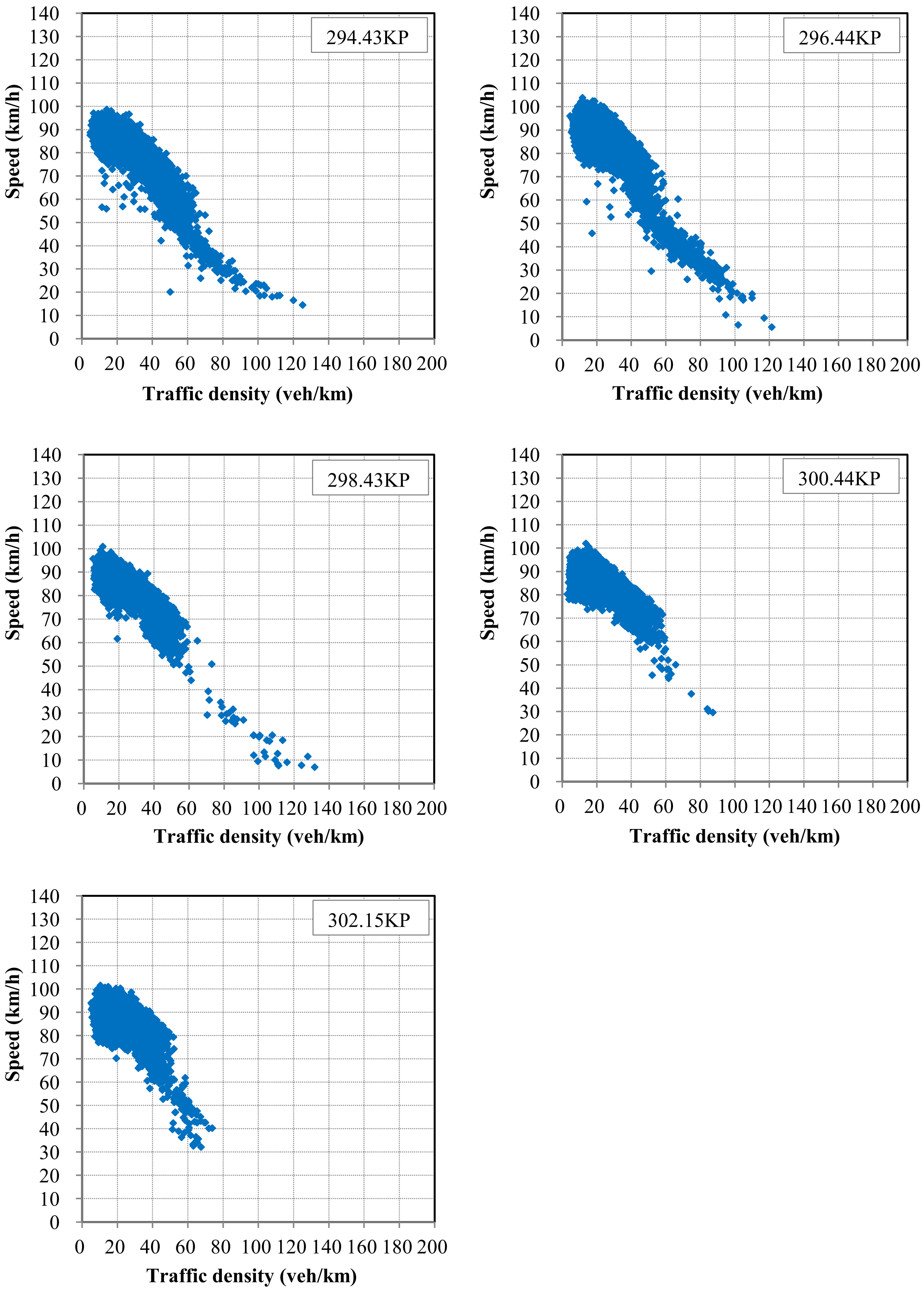


Figure 6.5 Speed-traffic density diagrams in the test bed on Tomei Expressway

6.2.3 *CR* statistics

Based on the crash records over the three years (2007-2009), *CR* in the two test beds are statistically calculated, as a reference object for the identification of potential crash-prone locations through using CREM. The statistical results are demonstrated in Figure 6.6 for Odaka route and Figure 6.7 for Tomei Expressway, respectively. The locations with high *CR* are observed at detector #0328 and #0334 in the test bed of Odaka route. For the test bed of Tomei Expressway, the corresponding location is 302.15KP. The three locations can be regarded as the so-called crash-prone locations by the statistics of *CR*.

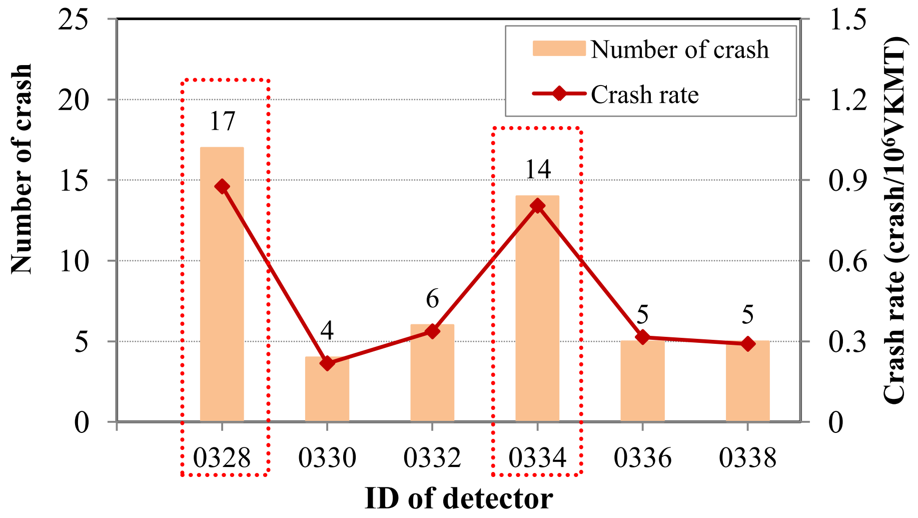


Figure 6.6 *CR* statistics in the test bed on Odaka route

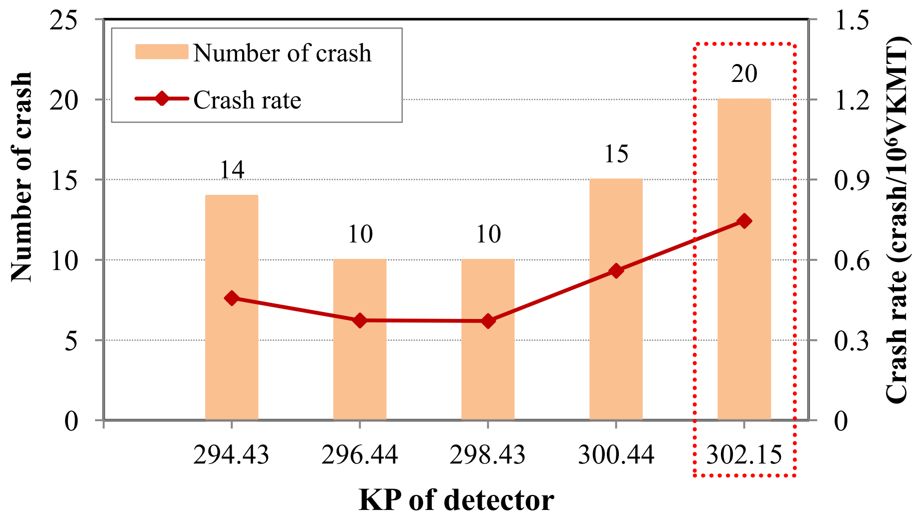


Figure 6.7 *CR* statistics in the test bed on Tomei Expressway

6.3 Crash Risk Prediction in Test Beds

6.3.1 Urban expressway

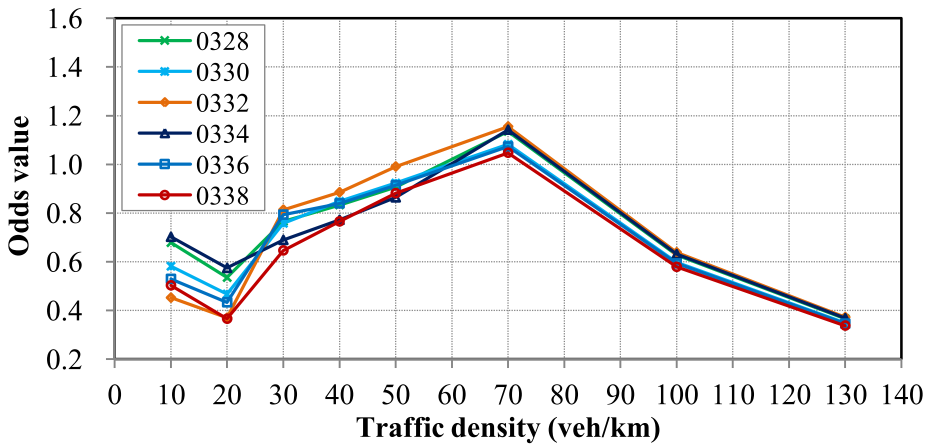
1) Crash risk tendency following traffic conditions

The tendencies of crash risk following traffic density are demonstrated in Figure 6.8. To discern the relativity of ambient conditions, crash risks on holiday and in nighttime are also illustrated. However, for some traffic categories, ambient *Light* and *Day* type are not both incorporated into CREM (Table 4.14). In these scenarios, crash risk in the normal ambient conditions is instead used to observe a complete process of the variation in crash risk.

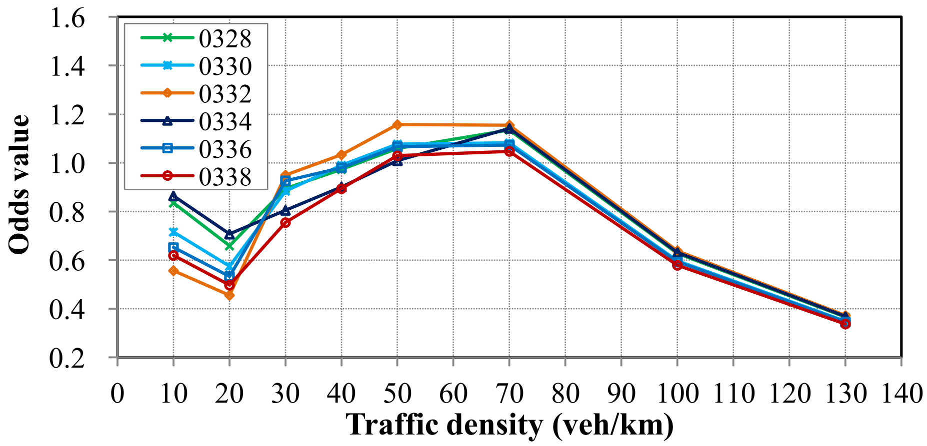
Crash risk generally appears decreasing tendencies at the early stages, then increases and finally decreases again in congested flow. Comparing to *CR* tendency shown in Figure 4.3, the tendency of crash risk may accord with that of *CR* in uncongested flow, while it is virtually in conflict with the *CR* tendency in congested flow.

When traffic density is much low, speed is high and the attention of drivers may be not sufficient in the discretionary driving conditions. Furthermore, such driving conditions often take place in nighttime, when visibility is reduced and it may conflict with the required long inter-vehicle spacing for safety, which is easy to induce crashes. As traffic density increases, speed will be reduced and drivers' caution would be enhanced. As a result, crash risk and the value of *CR* statistics may be on the decrease. With the further increase in traffic density, the impacts of inter-vehicles become severe and the demand of lane-changing behaviors for overtaking may increase due to the speed inharmonicity among vehicle types. Correspondingly, crash risk is on the up and the *CR* ascends.

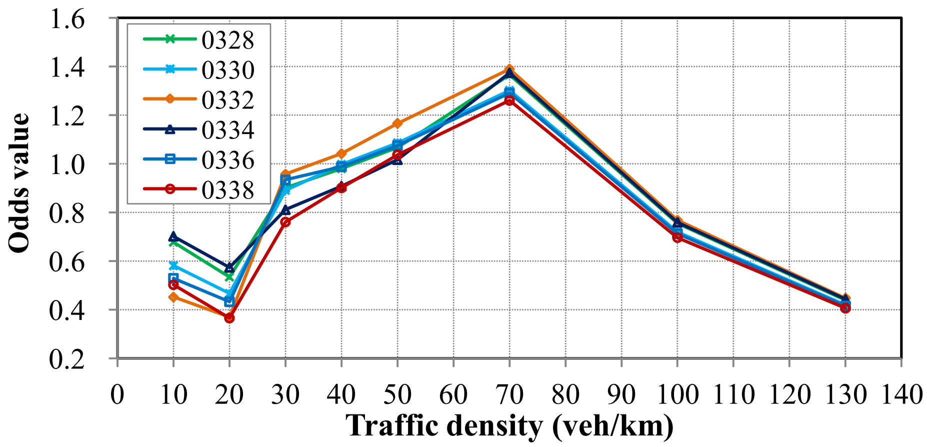
In congested flow, by experience, traffic oscillation is much more distinct immediately after the occurrence of breakdown, where operating speed is still high. With the growth in traffic congestion, traffic oscillation may be reduced combining with the decrease in speed. Since crash occurrence is significantly affected by short-term turbulence of traffic flow, it is considered reliable that crash risk could be reduced with the decrease in speed. In this sense, the confliction between crash risk and *CR* may be induced by the fundamental characteristics of crash records: insufficient crash samples and unequal dispersion (over- or under-dispersion) of crash events, as discussed in Chapter 2.



a) Daytime and weekday



b) Nighttime and weekday



c) Daytime and holiday

Figure 6.8 Crash risk following traffic conditions in the test bed on Odaka route

Between the detector locations, the differences of crash risk are observed getting smaller with the increase in traffic density, as the reflection of the reduced significance of geometric design affecting crashes. Through comparing the figures between daytime and nighttime, excluding congested flow where ambient *Light* is not incorporated, crash risks in uncongested flow are increasing in nighttime, and the differences of crash risk at these locations are slightly more distinct in comparison with daytime. The similar findings can be observed on holidays compared to weekdays, despite low-density uncongested flow where *Day* type is not considered for crash modeling. As a conclusion, nighttime/holiday may decrease the level of traffic safety and intensify the sensitivity of geometric design to crash risk in contrast to daytime/weekday, respectively.

2) Crash risk at individual locations

Through separating the categories of traffic conditions, this study predicts the values of crash risk per 0.1km for individual scenarios, as shown in Figure 6.9. Towards revealing the effect of geometric design on crash risk, the related geometric variations in the test bed are also demonstrated. For individual traffic conditions, the unit of vertical axis that is for crash risk is unified, in order to indicate the relativity between traffic conditions. This section focuses on the analysis in daytime on weekdays, and the related figures to illustrate the situations in nighttime and on holidays are described in the Appendix B. In the three figures, safety performance is investigated to be different by location, and the following characteristics of crash risk can be observed:

- The extent of fluctuation in *Odds* value along the test bed is highest in low-density uncongested flow. With the increase in traffic density, that extent is reduced.
- Compared with uncongested flow, the variation in crash risk with the change of scenario is more notable in congested flow.
- Generally, the high-crash risk locations can be observed to be varied with the transformation of traffic conditions.

The first finding reflects that the sensitivity of horizontal displacement to crash risk may be reduced with the increase in traffic density. The second one reveals the increased sensitivity of traffic conditions to crash risk in congested flow relative to uncongested flow. Both characteristics have been discussed in Figure 6.8. With respect to the third one, it may illustrate the phenomenon of crash-prone locations dependent on traffic conditions.

In low-density uncongested flow, the high-crash risk locations are observed around the locations at 8.0km and 8.9km, which are located at the end of a small curve (*R*-300m) and inside a small curve (*R*-200m) through referring to Figure 6.2, respectively. As it is known, low traffic density corresponds to high operating speed, which would induce large centrifugal force when drive inside a small curve. Thus, it is reliable to regard the location at 8.9km as one crash-prone location. At the end of a small curve, most drivers may take acceleration once they feel the decrease in curvature. In practice, such driving behavior is highly related to the experience of drivers. In this context, for the inexperienced drivers and the travelers unfamiliar with driving conditions, it is potential to operate an inadequate acceleration for the variation in alignment.

In high-density uncongested flow, the high-crash risk locations around the small curve with 200m of radius are found out to be moved to locations at 8.7km and 9.5km, which locate at the beginning of the small curve and inside a large curve (*R*-1000m) that follows the small curve. Once a vehicle is heading into a small curve, it has to slow down, and this behavior may throw an interruption to the upstream traffic due to the more intensive inter-vehicle interaction. As for the location around 9.5km, the speed of traffic has to be adjusted following the acceleration at the end of that small curve. In such case, if the drivers in the upstream cannot perceive the variation in vehicle maneuver of the downstream traffic in time, crashes such as vehicle-to-vehicle collision are prone to occur.

Through matching traffic conditions with crash records, up to 14 and 12 crash events occurred in low-density uncongested flow around the location at 8.0km and 8.9km. In this regard, the findings above can reliably account for the reason why so many crashes happened surrounding the two locations. Furthermore, the locations around 8.7km and 9.5km may be regarded as the potential crash-prone locations for high-density uncongested flows through crash risk prediction, which can't be discovered based on *CR* statistics. At these potential crash-prone locations, some countermeasures would be carried out in time, and thus the hazardous conditions may be smoothed appropriately.

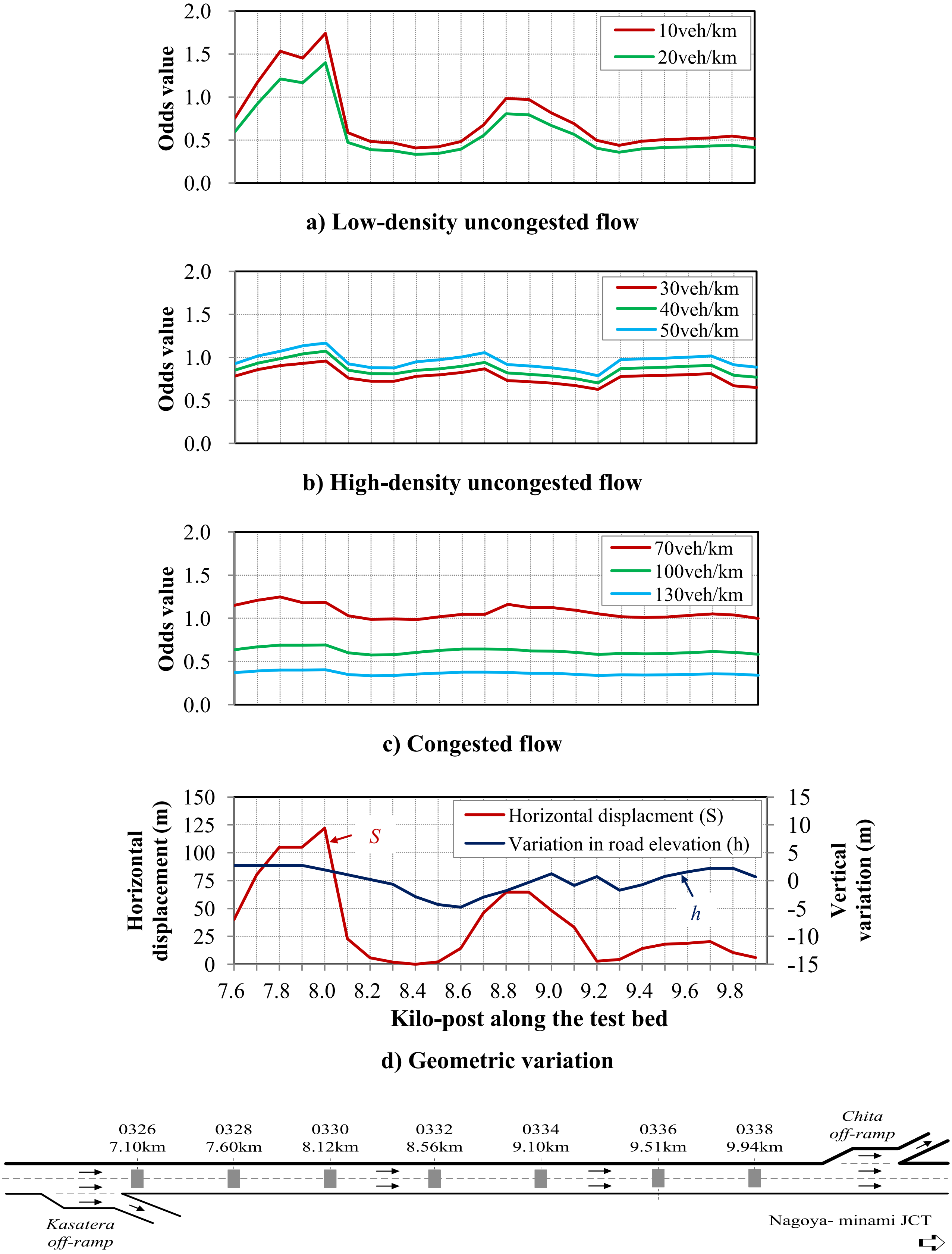


Figure 6.9 Crash risk prediction in the test bed on Odaka route

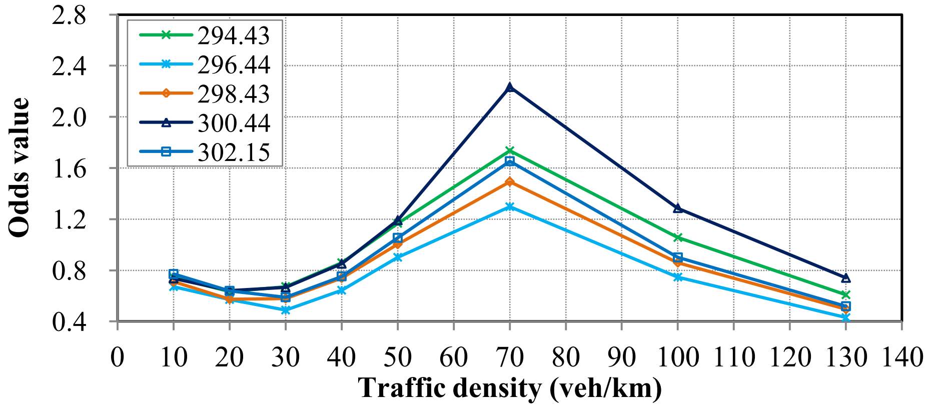
Other than the above findings, the safety benefit of geometric design in the test bed can also be evaluated. In low-density uncongested flow where the extent of fluctuation in *Odds* is distinct, the different crash risk near to the locations at 8.0km and 8.9km are up to over 1.0 and 0.8, respectively. It may be one potential cause that high *CR* occurs around these two locations. As shown in Figure 6.2, two small curves are virtually located around the two locations, and the geometric variations surrounding the two curves are much higher in contrast to other locations (see Figure 6.9d). To some extent, the predicted crash risk may imply the safety benefit of design consistency. Thus, it is reasonable to believe that crash risk estimation is an applicable measure to assess the safety performance of geometric design, even if a substantial effort is further required to adapt such analysis into practice, such as the development of assessing criterion based on crash risk.

6.3.2 Intercity expressway

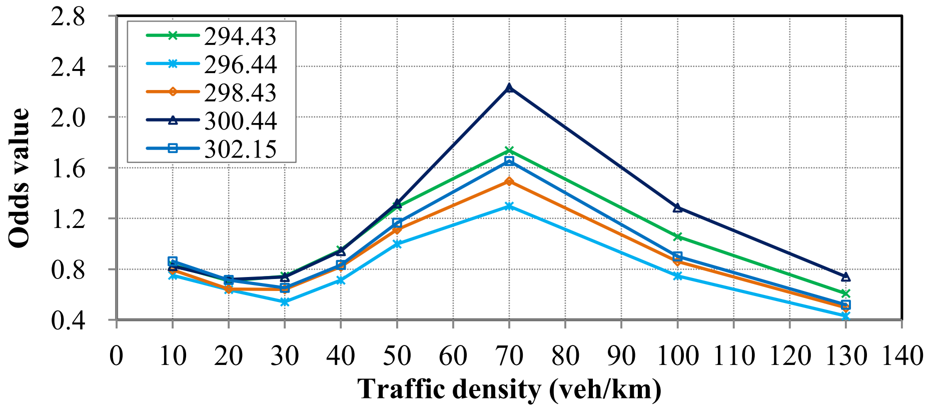
1) Crash risk tendency following traffic conditions

In the test bed on Tomei Expressway, Figure 6.10 illustrates the tendency of crash risk with the change of traffic conditions. For the same reason in section 6.3.1, the tendencies on holiday and in nighttime are also demonstrated, and the values of crash risk in normal ambient conditions are also utilized in some traffic conditions where ambient *Light* and *Day* type are not both incorporated into CREM, indicated in Table 5.11.

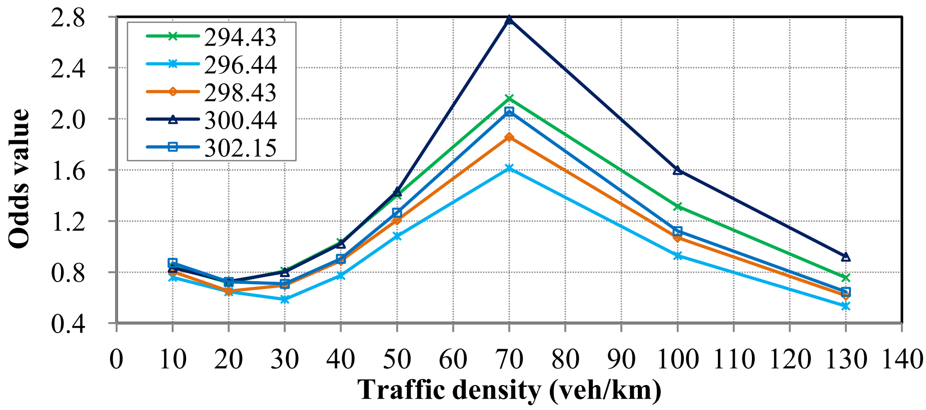
As similar to urban expressway, crash risk is observed to follow a decreasing tendency in low-density uncongested flow, then increases in high-density uncongested flow, and finally decreases with the increase in traffic density in congested flow. The features can also account for the U-shaped *CR* tendency in uncongested flow, as shown in Figure 5.2. However, in congested flow, the decreasing tendency of crash risk is also in conflict with the related *CR* tendency, due to the fundamental characteristics of crash records, *e.g.,* limited crash samples and over-/under-dispersions of crash events. Other than the findings, the differences of crash risk between individual locations are found out to be increased with the increase in traffic density in uncongested flow while decreased as traffic density increases in congested flow, different from urban expressway. Through referring to the three figures, it is found that the differences of crash risk between individual detector locations are slightly getting larger with the alterations of ambient *Light* and *Day* type, respectively. Besides, after the variation in ambient conditions, especially from weekday to holiday, the rise in crash risk gets more noticeable.



a) Daytime and weekday



b) Nighttime and weekday



c) Daytime and holiday

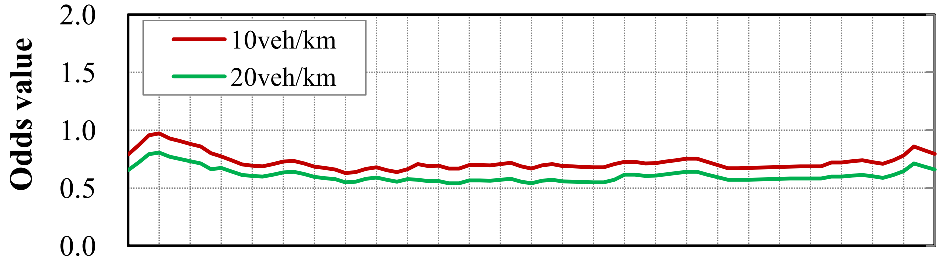
Figure 6.10 Crash risk following traffic conditions in the test bed on Tomei Expressway

As discussed in Chapter 3, comparing to urban expressway, the design of horizontal alignment on intercity expressway is better in view of the combination of curves. However, the vertical alignment on intercity expressway is poor regarding the value of gradient. As stated before, although the sensitivity of horizontal displacement S to crash risk is reduced with the increase in traffic density, the contribution of the variation in road elevation h would be on the rise. Especially for intercity expressway, where high HV% exists, the effect of h on vehicle maneuvers is critical to crash risk. Consequently, the differences of crash risk by location are increased as traffic density increases in uncongested flow. Once breakdown occurs, traffic oscillation gets to be maximal and later it would be reduced with the growth of traffic congestion. Therefore, even if the sensitivity of h on crash risk keeps higher relative to uncongested flow, the differences of crash risk by locations are on the decrease along with traffic density in congested flow.

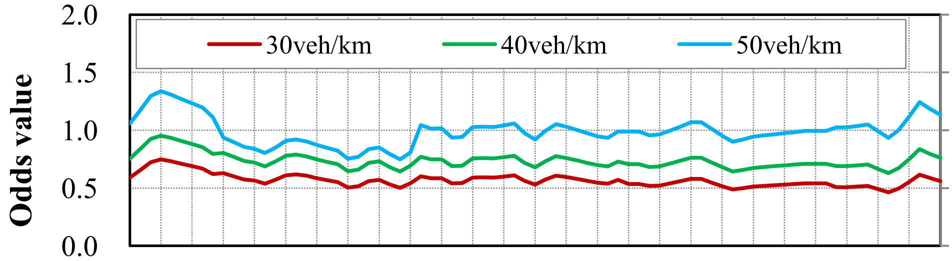
2) Crash risk at individual locations

Figure 6.11 shows the predicted values of crash risk per 0.1km for individual categories of traffic conditions. The related geometric variations in the test bed on Tomei Expressway are also demonstrated. The tendencies are also provided separately for each scenario. Likewise, the related figures in nighttime and on holiday are described in Appendix B of this dissertation. From these figures, the following characteristics can be observed:

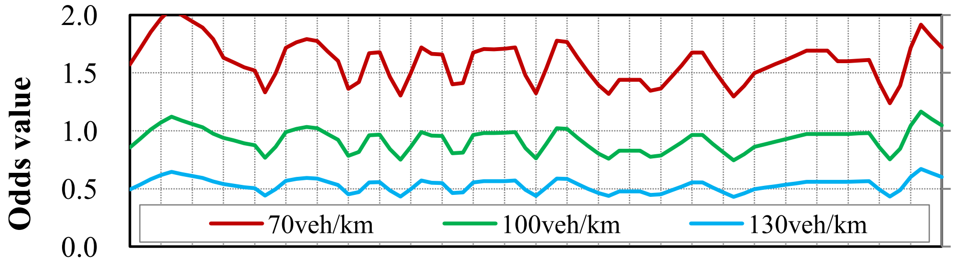
- As opposed to Odaka route, the extent of fluctuation in *Odds* value by location is on the rise with the variation in traffic conditions.
- As similar to Odaka route, the variation in *Odds* value caused by the change of scenario in congested flow is more distinct compared to uncongested flow.
- The relativity of crash risk by location also illustrates different crash characteristics for individual categories of traffic conditions.



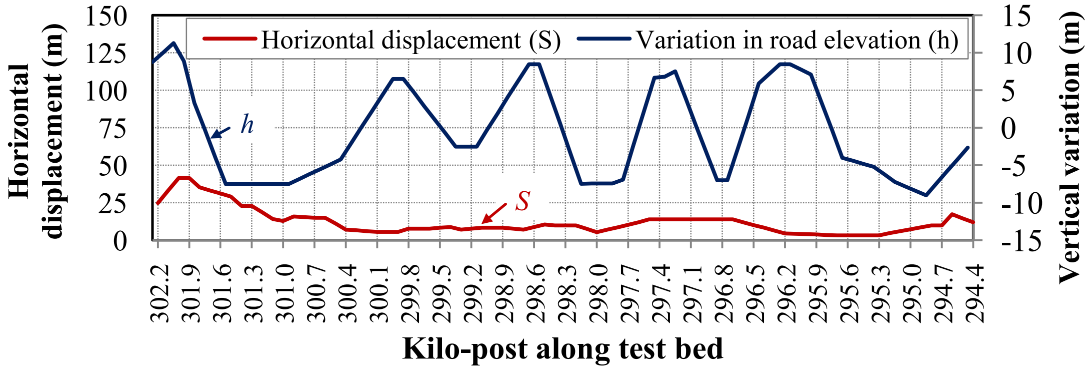
a) Low-density uncongested flow



b) High-density uncongested flow



c) Congested flow



d) Geometric variation

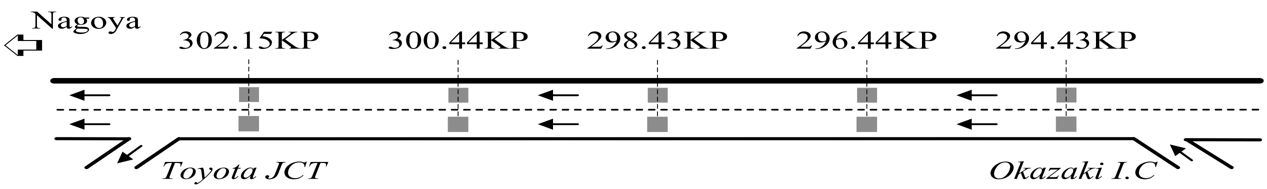


Figure 6.11 Crash risk prediction in the test bed on Tomei Expressway

As proved by the first finding, on Tomei Expressway, the sensitivity of geometric design to crash risk is increased with the increase in traffic density, which is contrary to the finding on Odaka route. As reflected by Figure 6.9d and 6.11d, compared to Odaka route, the horizontal variation *S* on Tomei Expressway is much smaller, while the variation of road elevation *h* is much higher. As discussed before, when traffic density increases, the effect of *h* on crash risk is virtually on the rise regarding the intensified inter-vehicle interaction. The second one indicates that the sensitivity of traffic conditions to crash risk gets more significant in congested flow relative to uncongested flow. On intercity expressway, more serious instability in traffic conditions may be causes with the increase in traffic density, due to its vehicle composition characterized in higher HV%. For Tomei Expressway, the third one also signifies the feature of crash-prone location dependent on traffic conditions.

For the whole traffic conditions, the location at 302km can be regarded as a crash-prone. As shown in Figure 6.11d, the horizontal variation near to that location is much larger in contrast to other locations, and the location is inside a downward section. As discussed in Chapter 2, relative to upward slope, downward slope may increase the difficulty for drivers to control vehicle maneuvers, which is more serious for HV. Such finding can further explain why there is highest *CR* surrounding 302.15km (Figure 6.7). In high-density uncongested flow, the other high-crash risk location is observed around the location at 294.6km. Corresponding to Figure 6.11d, *S* around 294.6km is the second largest value while the related *h* at the location is the largest one. Geometric variation especially *h* plays a significant role for crash risk if the inter-vehicle interaction arrive a relatively high value. The high crash risk at 294.6km can be confirmed by the result of *CR* statistics as shown in Figure 6.7. In congested flow, other than the high crash risk at locations at 294.6km and 302km, a more frequent fluctuation of *Odds* value can be observed in the section from 298km to 301km. In the distance, the variation in vertical slope is found out to be more frequent in contrast with other sections. In this regard, the variation in crash risk in congested flow is better to reveal the safety performance of vertical alignment.

Given the above discussions, since high speed-related crashes are significantly associated with the geometric design of horizontal alignment in view of the centrifugal force, the variation in crash risk by location in low-density uncongested flow is found out to be strongly correlated with the value of horizontal displacement *S*. Comparatively, the design of vertical alignment seems highly related to the variation in crash risk in congested flow, due to the increased impact of inter-vehicle interaction on crashes on vertical slopes. The findings may indicate the direction of safety performance evaluation for geometric design of horizontal/vertical alignment by crash risk prediction. In this way, due to the more distinct *S* on Odaka route, the safety performance of horizontal design can be more directly evaluated by crash risk prediction in the test bed on that route. In contrast, considering the more distinct *h* on Tomei Expressway, the safety performance of vertical design would be assessed more effectively in the test bed on that expressway.

6.4 Crash Risk Comparison between the Two Test Beds

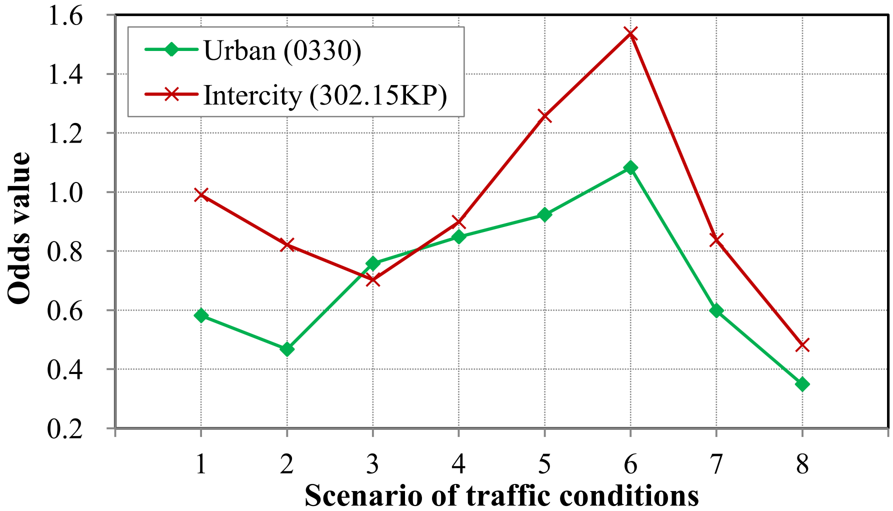
Through comparing the locations with the similar geometric variations between the two test beds, this section tries to identify the relative sensitivity of traffic conditions to crash risk by expressway type. Meanwhile, the effects of ambient *Light* and *Day* type on such kind of sensitivity by expressway type are also examined.

6.4.1 Analysis without ambient conditions consideration

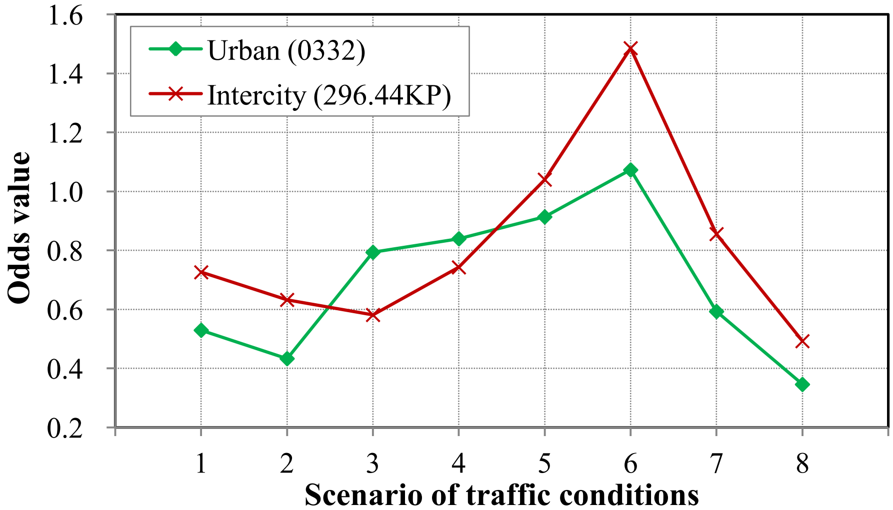
In view of the horizontal variation *S* referring to Figure 6.9d and 6.11d, two pairs of detectors are available to make a comparison by expressway type. Detectors #0330 and 302.15KP, whose horizontal displacements correspond to 22.2m and 20.9m, respectively, are regarded as the first pair between the two test beds. The other pair is composed of detectors #0332 and 296.44KP, and their related displacements are approximately 8.33m.

Figure 6.12 describes the tendencies of crash risk following traffic conditions for the two pairs of detectors (in daytime and on weekday). By comparison, there is higher crash risk on intercity expressway in scenario 1 and 2. With the increase in traffic density, a more notable rise in crash risk can be observed on urban expressway, and its value generally exceeds the risk on intercity expressway in scenario 3 (the first pair) or in both scenario 3 and 4 (the second pair). As traffic density further increases, the rising trend in crash risk on intercity expressway gets more significant, and its value turns to be much higher than that on urban expressway. If bring the two pairs into comparison, the difference of crash risk between the two test beds in scenario 3 can be found out in larger-scope for the second pair, whose horizontal displacement *S* is much smaller than that of the first one.

The differences above may be highly associated with traffic characteristics by expressway type. In low-density uncongested flow, the drivers on intercity expressway may operate in higher speed, due to the design of wider cross section. Meanwhile, the vehicle composition on that type of expressway is often characterized in high HV%, whose operating speed is often lower than other traffic. In such case, some crash risk-avoiding behaviors such as slow down and lane-changing would be caused to keep a safe inter-vehicle spacing. As a result, the traffic conditions on intercity expressway at speed over 100km/h is less safe than the conditions on urban expressway, as illustrated in Figure 5.8.



a) 20m<S<25m



b) 5m<S<10m

Figure 6.12 Crash risk comparison between the two test beds

When traffic density increases, the inter-vehicle interaction gets more intensive in scenario 3 and 4. The design of narrower cross section may cause a higher sensitivity of traffic density to the inter-vehicle interaction. In this regard, it is considered reliable that the rise in crash risk is more distinct on urban expressway from scenario 2 to 3. By contrast, owing to the wider cross section, intercity expressway may bear higher traffic density until the inter-vehicle interaction start to play a significant role to crash risk.

With the further increase in traffic flow, the inter-vehicle interaction is strengthened as well. Thereby, HV may give full interruption to the surrounding traffic, and result in frequent turbulence in traffic conditions. By reason of the high HV%, such instability in traffic conditions may be stronger on intercity expressway. Necessarily, much higher crash risk is performed on that type of expressway.

As regards the larger-scope difference of crash risk in scenario 3 for the second pair compared to the first pair, the primary cause is the much lower risk at 296.44KP, which belong to the second pair, in contrast to 302.15KP, whose horizontal displacement S is larger than the value at 296.44KP. When S is more significant, the increased traffic density would throw a serious impact on the inter-vehicle interaction, and enhance the role of that interaction for crash occurrence.

6.4.2 Analysis with ambient conditions consideration

Through focusing on the second pair of detector locations between urban and intercity expressways, the influences of ambient *Light* and *Day* type on the sensitivity of traffic conditions to crash risk are demonstrated in Figure 6.13 and 6.14, respectively. Given the variables incorporated into crash modeling for urban and intercity expressways, traffic conditions leaving out scenario 6 to 8 are analyzed for ambient *Light*. For *Day* type, the traffic conditions from scenario 3 to 8 are selected.

Regarding ambient *Light*, for the two types of expressways, crash risk of traffic flow in nighttime can be enhanced in contrast with daytime, as the result of the reduced visibility. By expressway type, the related increasing tendency is found out to be more substantial on urban expressway relative to intercity expressway. As analyzed in Chapter 3, on urban expressway, the speed variance in nighttime is more distinct. Furthermore, due to the design of narrower cross section, the reduced visibility would be more adverse to driving conditions for traffic safety. Correspondingly, the rising trend of crash risk in nighttime is more notable in the test bed of Odaka route.

With respect to *Day* type, the alternation from weekday to holiday may increase the crash risk for the reason of the increasing recreational travel on holiday. In the meantime, the related rise in crash risk seems more obvious in the test bed on Tomei Expressway, as the reflection of more distinct holiday traffic on that type of expressway, which is one distinct traffic characteristics of intercity expressway from urban expressway.

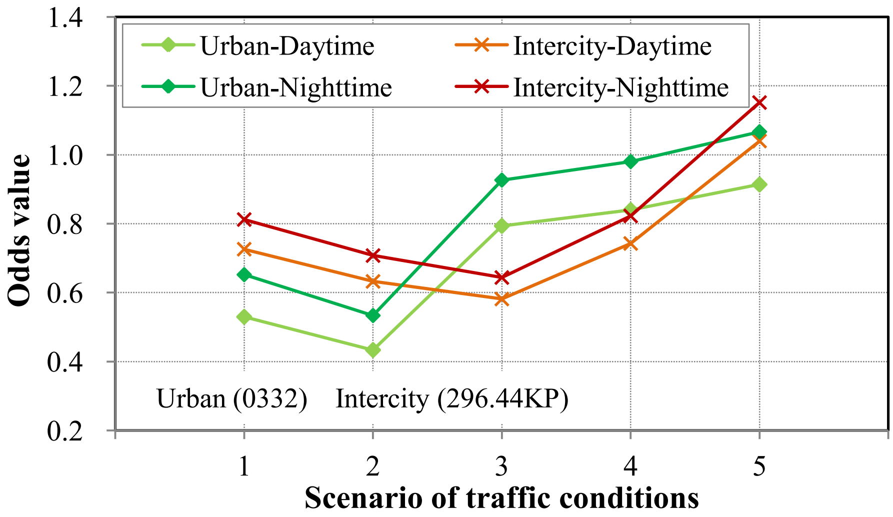


Figure 6.13 Crash risk comparison between daytime and nighttime

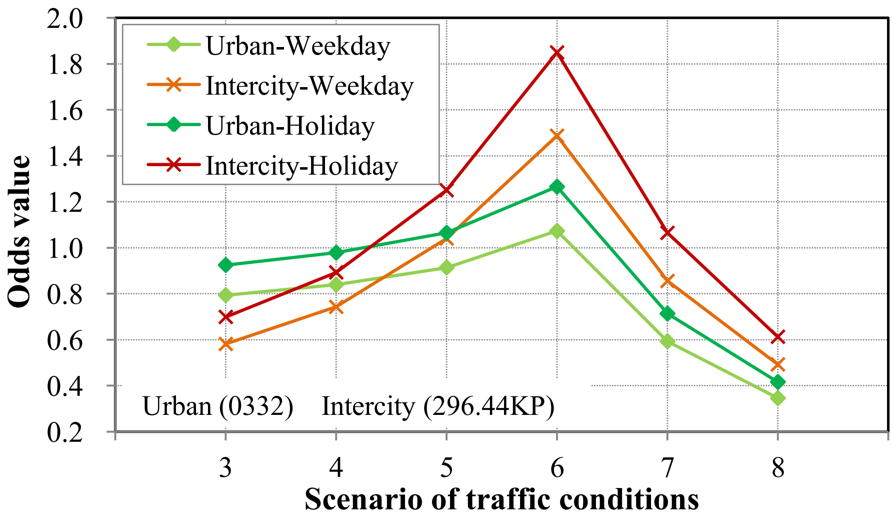


Figure 6.14 Crash risk comparison between weekday and holiday

6.5 Summary

The developed CREM was demonstrated for several applications in this chapter through focusing on the basic segments on Odaka route and Tomei Expressway, respectively. Through these applications, the evolving process of crash risk following traffic conditions was investigated. Meanwhile, crash risk was comparatively analyzed at individual locations in order to validate the safety benefit of design consistency. Besides, a new concept, traffic condition-dependent crash-prone location, was proposed. These findings were finally distinguished between urban and intercity expressways, and the sensitivity of traffic flow to crash risk was revealed to be different on the two types of expressways.

The first application showed that crash risk is convex downward to traffic density in uncongested flow, and follows a decreasing tendency in congested flow. With the increase in traffic density, the impact of geometry on crash risk is gradually reduced as opposed to a rise in the sensitivity of traffic conditions to crash risk.

Through the next application, the safety benefit of geometric design between neighboring locations can be directly indicated by the difference of crash risk at the two locations. Furthermore, crash-prone locations were found out to be traffic condition dependent. In this way, not only the locations where high *CR* is virtually observed, but also the potential crash-prone locations by traffic conditions can be measured. In this sense, this method may renew the traditional concept of crash-prone sections.

The different sensitivities of traffic flow to crash risk by expressway type are the major aim of the third application. With the change of traffic conditions, higher sensitivity was found on intercity expressway in low-density uncongested flow. Then, traffic becomes more sensitive to crash risk on urban expressway, before inter-vehicle interaction fully playing impacts on crash risk. Later, traffic flow on intercity expressway gets more sensitive again. Besides, in contrast to weekday and daytime, holiday and nighttime may increase crash risk, respectively. The rising trend caused by the alteration of *Day* type was more tremendous on intercity expressway. In contrast, the rising trend was observed more distinct on urban expressway with the alteration from daytime to nighttime.

Chapter 7

CONCLUSIONS AND FUTURE WORKS

7.1 Conclusions

Crash characteristics and the related influencing factors, as the theoretical basis for safety improvement, are critical to safer geometric design and traffic control strategy. Up to now, the conventional studies primarily identify the relationships between crashes and traffic variables or geometric elements in separate models. Meanwhile, crash characteristics on traffic condition-dependent and facility/expressway type-specific bases have not been paid sufficient attention. Given these problems, a crash risk estimation model through focusing on traffic conditions was developed for basic segments of urban/intercity expressways, respectively, considering the interaction of geometry, traffic flow and ambient conditions. Conclusions and results of the study are briefly described in the following sections.

7.1.1 Effects of explanatory variables on crash occurrence

Crash occurrence is a complex phenomenon and it is associated with the interaction of geometry, traffic flow and ambient conditions in nature. This study identified the effects of these explanatory variables on crashes dependent on traffic conditions.

Horizontal displacement S is the most significant influencing factor in low-density uncongested flow. As traffic density increases, the contribution of S to crashes is gradually reduced. On the contrary, the variation in h becomes more sensitive to crash risk. In the whole traffic conditions, the two variables have positive contribution to crash risk. With the increase in S and h, the operating speed has to change more frequently. As a result, much more serious instability in traffic conditions would be induced.

As opposed to geometric variations, the sensitivity of traffic variables to crash risk is on the rise when traffic density increases. In low-density uncongested flow, the driving condition is discretionary, and traffic density k_S and speed v_S are incorporated into crash modeling in an effort to more reliably represent traffic characteristics. It was found that crash risk is rising as v_S increases. By contrast, k_S has a negative effect on crash risk. In other traffic conditions, only v_S that is more applicable for traffic control is involved in view of the high correlation of the two traffic variables. In high-density uncongested flow, v_S has a negative contribution to crash risk, since the inter-vehicle interaction becomes more intensive with the decrease in speed. In congested flow, the coefficient of v_S related to crash risk gets positive, because traffic oscillation is more serious at the early stage of traffic congestion and major crash events are virtually observed at that stage.

Ambient conditions are another non-negligible explanatory factor. Nighttime and holiday may increase crash risk compared with daytime and weekday, respectively. Besides, the two ambient conditions would raise the sensitivities of geometric variations to crash risk in contrast to their corresponding opposing conditions.

7.1.2 Different affecting mechanisms of explanatory variables by expressway type

With the purpose to comprehensively identified crash influencing factors, a comparative analysis on these variables was operated between urban and intercity expressways.

In low-density uncongested flow, geometric design is a major cause leading to different crash characteristics by expressway type. The compacted designs, *e.g.,* small radius of curve and narrower cross section layout, are significantly related to higher crash risk on urban expressway. Within the same scope of geometric variation, high speed is more sensitive to crash risk on intercity expressway in contrast to urban expressway. Most low-density traffic is in nighttime, where much high HV% is characterized on intercity expressway. The high speed of other traffic would be in conflict with the slower operating speed of HV. As a result, the turbulence of traffic flow is potentially induced.

In high-density uncongested flow, since the inter-vehicle interaction gets more intensive, the interruption of neighboring traffic plays an important role in crash characteristics. In view of the higher HV%, the interruption of HV to its surrounding traffic would be stronger on intercity expressway. Besides, the more LDT may aggravate such interruption due to its high expected speed conflicting with slower speed of HV. Thus, the driving conditions on intercity expressway are then less safe than those on urban expressway.

Compared to uncongested flow, the variation in speed is much more sensitive to crash risk in congested flow. In virtue of the high HV%, intercity expressway still has worse traffic safety situation in contrast to urban expressway, since the variation in speed surrounding HV would be stronger and more frequent.

For both types of expressways, higher crash risk would exist in nighttime and on holidays compared to to daytime and weekdays, respectively, while the growth rate is different by expressway type. On urban expressway, traffic flow in nighttime is of much notable speed variance. Meanwhile, the reduced visibility would be more adverse to driving conditions on narrower cross section. Correspondingly, the rise in crash risk during nighttime is more significant on urban expressway. Comparatively, the characteristics of holiday traffic are considered to be more distinct on intercity expressway. As a result, the rise in crash risk caused by the alternation of *Day* type is in a larger scope on that type of expressway.

7.1.3 Evolving process of crash risk with the variation in traffic conditions

To improve the efficiency of proactive traffic management, the evolving process of crash risk with the variation in traffic conditions was investigated by applying the developed model on a subject basic segment of urban and intercity expressway, respectively. It is found out 1) crash risk is convex downward to traffic density in uncongested flow and follows a decreasing tendency in congested flow; 2) the sensitivity of traffic conditions to crash risk is different for urban and intercity expressways.

When traffic density is much low, discretionary driving conditions (*e.g.,* high speed and lower attention of drivers) along with the reduced visibility in nighttime where low-density uncontested flow often exists are the significant cause of high crash risk. With the increase in traffic density, the inter-vehicle interaction is increasing, and crash risk may be on the decrease due to the reduced speed and the enhanced drivers’ caution. As traffic density further increases, the impacts of inter-vehicles get severe and the demand of lane-changing behaviors may increase because of the speed inharmonicity of various vehicle types. As a result, crash risk is on the rise. In congested flow, with the increase in traffic congestion, traffic oscillation may be decreased, and thus crash risk would be reduced.

By expressway type, due to the wider cross section, driving conditions may be more discretionary on intercity expressway in low-density uncongested flow. Thus, traffic flow is more sensitive to crash risk on intercity expressway relative to urban expressway, disregarding the difference of geometric variation. When traffic density increases, the narrower cross section layout is, the smaller traffic volume can be contained until the inter-vehicle interaction starts an important role in crash risk. As a result, the sensitivity of traffic conditions to crash risk gets higher on urban expressway with the increase in traffic density. Once the inter-vehicle interaction plays a fully important role, the sensitivity of traffic conditions to crash risk would be more notable on intercity expressway in virtue of its vehicle composition characterized by higher HV% and more LDT.

7.1.4 Measuring the quality of geometric design

The safety evaluation of geometric design is critical for the improvement of expressway planning and design for safer expressway. Since high speed-related crashes are associated with the design of horizontal alignment in view of the centrifugal force, the variation in crash risk by location in low-density uncongested flow may be strongly correlated with the value of horizontal displacement *S*. Comparatively, the design of vertical alignment seems highly related to the variation in crash risk in congested flow, due to the increased impact of inter-vehicle interaction on crashes on vertical slopes. In this regard, the measure of crash risk is an applicable way to assess safety performance of geometric design.

Identifying crash-prone location and its cause is very important for road administration to enable effective countermeasures. However, the traditional identification is based on *CR* statistics without traffic conditions consideration. In this study, a new concept is proposed regarding the interaction of geometry, traffic flow and ambient conditions on crash risk. It points out that crash-prone locations are traffic condition dependent. In this way, not only the locations where high *CR* exists, but also the potential crash-prone locations by traffic conditions can be identified. At the viewpoint of proactive traffic management, the concept is considered more applicable for operational applications.

7.1.5 Applicability of proactive strategies

One of the overall purposes of this study is to serve more applicable application for safer geometric design. The major applicability is summarized in the following.

- For driving conditions in low-density uncongested flow, the maximum safe speed was proposed corresponding to individual exposures, which could be regarded as a benchmark for geometric design.
- The measure of crash risk estimation can be applied for the evaluation of safety performance of geometric design.
- The concept of traffic condition-dependent crash-prone locations is more effective comparing to the traditional methods in view of the applicability for adopting countermeasures of geometric design for traffic safety.

The other application is for proactive traffic management strategies in the purpose to find a more applicable way for controlling traffic flow.

- Based on the model by traffic conditions, a real-time predictive system for crash risk was developed. For proactive traffic management strategies, it may provide leverage to predict hazardous conditions and avoid an impending crash.
- The quantitative effects of explanatory variables on crash risk may help prioritize countermeasures. Meanwhile, the efficiency of an adopted countermeasure may be predicted in advance through referring to the quantitative relationships.
- The evolving process of crash risk with the variation in traffic conditions was predicted, and the mechanism of hazardous conditions can be understood more sufficiently, supporting for more efficient traffic control measures.
- Different effects of explanatory variables on crash risk by expressway type may help comprehensively understanding the causes of crash occurrence. Therefore, for each type of expressway, more effective countermeasures may be applied.

7.2 Future Works

The current research was achieved based on several assumptions which need to be improved. Furthermore, the scope of this study could be extended to a more universal methodology. Based on the primary conclusions summarized in the previous section, some directions for future research are addressed in the following sections.

7.2.1 Applicability to other facility types

The developed model is just focused on basic segment, and it is difficult to image to what extent the findings can be transferable to other facility types, since crash characteristics are facility type-specific in nature. The vehicle behaviors are generally complex on other facility types, which is often impacted by mainline and ramp traffic streams. If the related data are available, it would be interesting to quantify the effects of geometry, traffic flow and ambient conditions on crash risk for other facility types, through using the same methodology of this study. In this way, a real-time traffic flow monitoring system in terms of safety can be completely developed for the whole network of expressway.

7.2.2 Improvement of data collection

Since crash occurrence is significantly associated with the short-term turbulence in traffic conditions, the variables representing average conditions even in 5 minutes are inadequate to reflect the momentarily disruptive traffic flow. In this sense, the related variables are quite required to improve the performance of the current CREM. Meanwhile, this study did not distinguish individual lanes, while the variations in speed and density separately across lanes are significantly related to crashes. Besides, due to the speed inharmonization, HV plays an important role in crash risk. However, considering the relativity by expressway type, the two kinds of data above have not been incorporated into crash modeling.

Ambient conditions are another not negligible influencing factor. For the current model, just two types of ambient conditions were incorporated in the form of dummy variables. The method may imply the effect of ambient conditions on crash characteristics to some extent but not perfectly. The detailed ambient data, such as precipitation and visibility, may improve the accuracy for assessing the contribution of ambient conditions to crash risk.

7.2.3 Crash type-specific analysis

Individual crash types may occur under substantially various circumstances and may be associated with explanatory variables in different ways. In this regard, crash modeling by focusing on total crash events may be not exactly to reveal the affecting mechanisms of explanatory variables to different types of crashes. In other words, the countermeasures to reduce crash risk obtained from CREM could be more reliable for crashes of a specific type than those without differentiating crash type.

7.2.4 Implementation of traffic control measures

Once a given condition is identified to be hazardous, the next step is to adopt various traffic smoothing measures to countervail crash risk. Based on the current model, the direction of traffic control may be explicit, while the way how to apply those measures in efficiency is not entirely clear. Furthermore, after the measures at that given location, how the measures affect the driving conditions at the upstream locations remains unclear. In future, this research will be extended to develop traffic control strategies.

References

Abdel-Aty, M., Uddin, N., Abdalla, F., Pande, A. and **Hsia, L.** (2004). Predicting Freeway Crashes Based on Loop Detector Data Using Matched Case-control Logistic Regression. *Transportation Research Record 1897:* 88-95.

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.

Abdel-Aty, M., Pande, A., Das, A. and **Knibbe, W. J.** (2008). Assessing Safety on Dutch Freeways with Data from Infrastructural-based Intelligent Transportation Systems. *Transportation Research Record 2083,* 153-161.

Alberta Transportation. (2008). Alberta Traffic Collision Statistics. Office of Traffic Safety, Alberta Transportation, Edmonton, Canada.

Al-Kaisy, A. and **Jung, Y.** (2004). Examining the Effect of Heavy Vehicles during Congestion Using Passenger Car Equivalents. *Proceedings of the 8th International Symposium on Heavy Vehicle Weights and Dimensions 'Loads, Roads and the Information Highway',* Johannesburg, South Africa.

Anowar, S., Yasmin, S. and **Tay, R.** (2012). Severity of single vehicle crashes during holidays. *TRB Annual Meeting CD-ROM,* Washington, D.C.

Anowar, S., Yasmin, S. and **Tay, R.** (2013). Comparison of Crashes during Public Holidays and Regular Weekends. *Accident Analysis & Prevention,* 51, 93-97.

Armenian, H. K. (2009). The Case-control Method: Design and Applications. Oxford University Press, New York.

Bajwa, S., Warita, H. 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.

Brilon, W. and **Estel, A.** (2009). A Different Analysis of Level of Service F within the German Highway Capacity Manual. *TRB Annual meeting CD-ROM,* Washington D.C.

Chang, J., Oh, C. and **Chang, M.** (2000). Effects of Traffic Condition (v/c) on Safety at Freeway Facility Sections. *Transportation Research Circular E-C018: 4th International Symposium on Highway Capacity,* 200-208.

Chang, L. Y. (2005). Analysis of Freeway Accident Frequencies: Negative Binomial Regression versus Artificial Neural Network. *Safety Science,* 43 (8), 541-557.

Chang, L. Y. and **Chen, W. C.** (2005). Data Mining of Tree-based Models to Analyze Freeway Accident Frequency. *Journal of Safety Research,* 36, 365-375.

Chen, H. Y., Liu, P., Behzadi, B. and **Lu, J. J.** (2008). Impacts of Exit Ramp Type on the Safety Performance of Freeway Diverge Areas, *TRB Annual meeting CD-ROM,* Washington D.C.

Chin, S. M., Franzese, O., Greene, D. L., Hwang, H .L. and Gibson, R. C. (2004). Temporary Loss of Highway Capacity and Impacts on Performance: Phase 2. Oak Ridge, Tennessee: Oak Ridge National Laboratory.

Chung, E., Ohtani, O., Warita, H., Kuwahara, M. and **Morita, H.** (2005). Effect of Rain on Travel Demand and Traffic Accidents. *Proceedings of the 8th International IEEE Conference on Intelligent Transpiration System,* 13-16, Vienna, Austria.

Chung, E., Ohtani, O., Warita, H., Kuwahara, M. and **Morita, H.** (2006). Does Weather Affects Highway Capacity. *5th International Symposium on Highway Capacity and Quality of Service,* Vol.1, 139-146.

Christoforou, Z., Cohen, S. and **Karlaftis, M. G.** (2011). Identifying Crash Types Propensity Using Real-time Traffic Data on Freeways, *Journal of Safety Research,* 42, 43-50.

Christoforou, Z., Cohen, S. and **Karlaftis, M. G.** (2012). Integrating Real-time Traffic Data in Road Safety Analysis, *Procedia-Social and Behavior Sciences,* 40, 2454-2463.

Dias, C., Miska, M., Kuwahara, M. and Warita, H. (2009). Relationship between Congestion and Traffic Accidents on Expressways: an Investigation with bayesian Belief Networks. Proceedings of the 40th *IP meeting of Japan Society of Civil Engineering CD-ROM.*

Duret, A., Ahn, S. and **Buisson, C.** (2012). Lane Flow Distribution on a Three-lane Freeway: General Features and the Effects of Traffic Controls, *Transportation Research Part C,* 24, 157-167.

Elvik, R. and Mysen, A. B. (1999). Incomplete Accident Reporting: Meta-analysis of Studies Made in 13 Countries. *Transportation Research Record 1665,* 133-140.

Farmer, C. and Williams, A. (2005). Temporal Factors in Motor vehicle Crash Deaths. Injury Prevention, 11(1), 18-23.

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.

GAISMA. (2013). Nagoya, Japan - Sunrise, sunset, dawn and dusk times, table. http://www.gaisma.com/en/location/nagoya.html.

Gao, J. P., Kong, L. Q., Guo, Z. Y. and **Wu, Z. H.** (2004). Study on Operating Speed of Vehicles on Expressway. *Journal of Chongqing Jiaotong University,* 04, 80-83. (In Chinese)

Garber, N. J. and Wu, L. (2001). Stochastic Models Relating Crash Probabilities with Geometric and Corresponding Traffic Characteristics Data. Research Report No. UVACTS-5-15-74.

Garber, N. J. and Kassebaum, E. A. (2008). Evaluation of Crash Rates and Causal Factors for High-Risk Locations on Rural and Urban Two-lane Highways in Virginia. Publication VTRC 09-R1. Virginia Transportation Research Council.

Golob, T. F. and **Recker, W. W.** (2003). Relationships among Urban Freeway Accidents, Traffic Flow, Weather and Lighting Conditions. *Journal of Transportation Engineering*, 129(4), 342-353.

Golob, T. F. and **Recker, W. W.** (2004). A Method for Relating Type of Crash to Traffic Flow Characteristics on Urban Freeways. *Transportation Research Part A,* 38(1), 53-80.

Golob, T., Recker, W. and **Alvarez, V.** (2004). Freeway Safety as a Function of Traffic Flow. *Accident Analysis & Prevention,* 36 (6), 933-946.

Hall, R. W. (1993). Non-recurrent Congestion: How big is the Problem? Are Traveller Information Systems the Solution? *Transportation Research Part C: Emerging Technologies,* 1(1), 89-103.

Hauer, E. (1999). Safety and the Choice of Degrees of Curve. *Transportation Research Record: Journal of the Transportation Research Board,* 1665.

Hauer, E. (2006). The Frequency-Severity Indeterminacy. *Accident Analysis & Prevention,* 38, 78-83.

Hauer, E. (2009). Speed and Safety. *TRB Annual meeting CD-ROM,* Washington D.C.

Hikosaka, T. and Nakamura, H. (2001). Statistical Analysis on the Relationship between Traffic Accident Rate and Traffic Flow Condition in Basic Expressway Sections, *21st Japan Society of traffic Engineers (JSTE) Meeting*. (In Japanese)

Hosmer, D. W. and **Lemeshow, S.** (2004). Applied Logistic Regression. *Wiley-Interscience*.

Hossain, M. and **Muromachi, Y.** (2012). A Bayesian Network based Framework for Real-time Crash Prediction on the Basic Freeway Segments of Urban Expressways. *Accident Analysis & Prevention,* 45 (2), 373-381.

Hummer, J. E., Rasdorf, W., Findley, D. J., Zegeer, C. V. and **Sundstrom, C. A.** (2010). An Examination of Horizontal Curve Collisions Characteristics and Corresponding Countermeasures. *TRB Annual meeting CD-ROM,* Washington D.C.

Japan-guide. (2013). Travel-Transportation-Expressway. http://www.japan-guide.com/e/e2354.html.

Japan National Police Agency (NPA). (2013). Annual Report of Traffic Crash Statistics for 2012. http://www.e-stat.go.jp/SG1/estat/List.do?lid=000001108012.

JH. (2004). Traffic Statistics of Japan Highway Public Corporation 2004. Japan Highway Public Corporation. (in Japanese)

Kim, D. G., Washington, S. and **Oh, J.** (2006). Modelling Crash Types: New Insight into the Effects of Covariates on Crashes at Rural Intersections. *Journal of Transportation Engineering,* 132, 282-292.

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.

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.

Kumara, S. S. P. and **Chin, H. C.** (2005). Application of Poisson Underreporting Model to Examine Crash Frequencies at Signalized Three-legged Intersections. *Transportation Research Record 1908,* 46-50.

Lamm, R., Psarianos, B. and **Mailaender, T.** (1999). Highway Design and Traffic Safety Engineering Handbook. McGraw-Hill Companies, Inc., New York.

Lee, C., Saccomanno, F. and **Hellinga, B.** (2002). Analysis of Crash Precursors on Instrumented Freeways. *Transportation Research Record 1784,* 1-8.

Lee, C., Abdel-Aty, M. and **Hsia, L.** (2006). Potential Real-time Indicators of Sideswipe Crashes on Freeway. *TRB Annual meeting CD-ROM,* Washington D.C.

Lewallen, S. and **Courtright, P.** (1998). Epidemiology in Practice: Case-control Studies. *Community Eye Health,* 11(28), 57-58.

Lord, D., Manar, A. and **Vizioli, A.** (2005). Modeling Crash-flow-density and Crash-flow-v/c ratio Relationships for Rural and Urban Freeway Segments. *Accident Analysis & Prevention*, Vol. 37, 185-199.

Lord, D., Guikema, S. and **Geedipally, S. R.** (2008). Application of the Conway-Maxwell-Poisson Generalized Linear Model for Analyzing Motor Vehicle Crashes. *Accident Analysis & Prevention,* 40 (3), 1123–1134.

Lord, D. and **Mannering, F.** (2010). The Statistical Analysis of Crash-Frequency Data: A Review and Assessment of Methodological Alternatives. *Transportation Research Part A,* 44, 291-305.

Ma, J. (2009). Bayesian Analysis of Underreporting Poisson Regression Model with an Application to Traffic Crashes on Two-lane Highway. *TRB Annual meeting CD-ROM,* Washington D.C.

Martin, J, L. (2002). Relationship between Crash Rate and Hourly Traffic Flow on Interurban Motorways. *Accident Analysis & Prevention,* 34, 619-629.

Mayora, J. P. and Pina, R. J. (2013). An Assessment of the Skid Resistance Effects on Traffic Safety under Wet Pavement Conditions. *Accident Analysis & Prevention,* 41 (4), 881-886.

McGuigan, D. R. D. (1981). The Use of Relationships between Road Accidents and Traffic Flow in "Black-Spot" Identification. *Traffic Engineering and Control,* Aug.–Sept, 448-453.

Miaou, S. P. (1994). The Relationship between Truck Accidents and Geometric Design of Road Sections: Poisson versus Negative Binomial Regression. *Accident Analysis & Prevention,* 26(4), 471-482.

Milton, J. and **Mannering, F.** (1998). The Relationship among Highway Geometrics, Traffic-related Elements and Motor Vehicle Accident Frequencies. *Transportation,* 25(4), 395-413.

Nagoya Expressway Public Corporation. (2012). Map of Nagoya Urban Expressway network. http://www.nagoya-expressway.or.jp/guide/rosenzu/index.html.

Nagoya Expressway Public Corporation. (2013). Data of AADT, Change tendency of AADT . http://www.nagoya-expressway.or.jp/kosya/jigyo/tuukoudaisuu.html.

Ng, J. C. W. and **Sayed, T.** (2004). Effects of Geometric Design Consistency on Road Safety. *Canadian Journal of Civil Engineering,* 31(2), 218-227.

Noked, N. (2010). Providing a Corrective Subsidy to Insurers for Success in Reducing Traffic Accidents. http://nrs.harvard.edu/urn-3:HUL.InstRepos:4889453.

Oh, C., Oh, J. S. and Ritchie, S. G. (2001). Real-time Estimation of Freeway Accident Likelihood. *TRB Annual Meeting,* Washington, D.C.

Oh, C., Oh, J.S. and **Ritchie, S. G.** (2005). Real-time Hazardous Traffic Condition Warning System: Framework and Evaluation. *IEEE Transactions on Intelligent Transportation System*, 6(3), 8.

Orme, J. G. and **Combs-Orme, T.** (2009). Multiple Regressions with Discrete Dependent Variables. *Oxford University Press,* USA.

Panda, A. and **Abdel-Aty, M.** (2006). Comprehensive Analysis of the Relationship between Real-time Traffic Surveillance Data and Rear-end Crashes on Freeways. *Transportation Research Record 1953,* 31-40.

Park, E. S. and **Lord, D.** (2007). Multivariate Poisson-lognormal Models for Jointly Modeling Crash Frequency by Severity. *Transportation Research Record 2019,* 1-6.

Pasupathy, R., Ivan, J. and **Ossenmbruggen, P.** (2000). Single and Multi-vehicle Crash Prediction Models for Two-lane Roadway. *Final Report of Project UCNR9-8,* USDoT, Region I University Transportation Center.

Pei, Y, L. and **Ma, J.** (2003). Research on Countermeasures for Road Condition Causes of Traffic Accidents. *Journal of China Highway and Transport,* 16(4), 77-81. (In Chinese)

Qin, X., Ivan, J. and **Ravishanker, N.** (2004). Selecting Exposure Measures in Crash Rates Prediction for Two-lane Highway Segments. *Accident Analysis & Prevention,* 36, 183-191.

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*, 2102, 34-42.

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

Sarvi, M. and **Kuwahara, M.** (2007). Microsimulation of Freeway Ramp Merging Processes under Congested Traffic Conditions. *IEEE Transaction on Intelligent Transportation Systems,* 8(3), 470-479.

Shawky, M. and **Nakamura, H.** (2007). Characteristics of Breakdown Phenomena in Japan Urban Expressway Merging Sections, *TRB Annual meeting CD-ROM,* Washington D.C.

Shi, L. J., Cheng, Y., Jin, J., Ran, B. and **Chen, X. H.** (2011). Effects of Rainfall and Environmental Factors on Traffic Flow Characteristics on Urban Freeway. *TRB Annual meeting CD-ROM,* Washington D.C.

Shively, T. S., Kockelman, K. and **Damien, K.** (2010). A Bayesian Semi-parametric Model to Estimate Relationships between Crash Counts and Roadway Characteristics. *Transportation Research Part B,* 44(5), 699-715.

Solomon, D. (1964). Accidents on Main Rural Highways Related to Speed, Driver and Vehicle. U.S. DOT/FHWA.

Tranbbs. (2012). Please Analyze the Causes of Congestion on Expressway in China. http://bbs.tranbbs.com/dispbbs.asp?boardid=28&Id=54294.

Transport Bureau of Japanese National Police Agency. (2013). Traffic Accidents Statistics in 2012. http://www.e-stat.go.jp/SG1/estat/eStatTopPortal.do.

Transportation Research Board. (2010). Highway Capacity Manual. Washington D.C.

Watters, P. and **O'Mahony, M.** (2007). The Relationship between Geometric Design Consistency and Safety on Rural Signal Carriageways in Ireland. *European Transport Conference,* Producing Safer Transport Infrastructure.

WIKIPEDIA. (2011). Types of twilight, http://en.wikipedia.org/wiki/File: Types-of-twilight-en2.svg.

Wang, A. (2005). Bachelor Thesis: Using GIS to Explore the Relationship between Road Grade and Accidents. University of Southern Queensland, Queensland, Australia.

Wu, Y., Nakamura, H. and **Asano, M.** (2011). Relationship between Traffic Conditions and Crash Rates in Urban Expressway Basic Segments. *Proceedings of the 13th International Summer Symposium of JSCE,* 271-274.

Wu, Y., Nakamura, H. and **Asano, M.** (2012a). 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.*

Wu, Y., Nakamura, H. and **Asano, M.** (2012b). Effects of Traffic Flow Conditions on Crash Rates at Different Urban Expressway Facility Types, *Proceedings of the 14th International Summer Symposium of JSCE,* 111-112.

Wu, Y., Nakamura, H. and **Asano, M.** (2013a). A Comparative Study on Crash Characteristics between Urban and Intercity Expressway Basic Segments. *Selected Proceedings of the 13th Word Conference on Transport Research.*(Web-Journal)

Wu, Y., Nakamura, H. and **Asano, M.** (2013b). A Crash Risk Estimation Model for Urban Expressway Basic Segments Considering Geometry, Traffic Flow and Ambient Conditions. *Asian Transport Studies* (Accepted).

Xie, Y., Lord, D. and **Zhang, Y.** (2007). Predicting Motor Vehicle Collisions Using Bayesian Neural Network: an Empirical Analysis. *Accident Analysis & Prevention,* 39(5), 922-933.

Yeo, H., Jang, K. and **Skabaronis, A.** (2010). Impact of Traffic States on Freeway Collision Frequency, *TRB Annual meeting CD-ROM,* Washington D.C.

Yoshii, T., Hyodo, T. and **Endo, K.** (2010). Analysis on the Risk of Traffic Crashes Occurrence, *2nd seminar about utilization of traffic flow data on expressways*. (In Japanese)

Zheng, Z., Ahn, S. and **Monsere, C. M.** (2010). Impacts of Traffic Oscillations on Freeway Crash Occurrences. *Accident Analysis & Prevention,* 42 (2), 626-636.

Zhong, L. D. and **Sun, X. D.** (2011). Relationship between Traffic Flow and Safety of Freeways in China: A Case Study of Jingjintang Freeway. *The 3rd International Conference on Road Safety and Simulation,* Indianapolis, USA.

Zhou, M. and **Sisiopiku, V.** (1997). Relationship between Volume-to-capacity Ratios and Accident Rates. *Transportation Research Record 1581*: 47-52.

Zhou, H. G., Chen, H. Y., Zhao, J. G. and **Hsu, P.** (2010). Operational and Safety Performance of Left-side Off-ramps at Freeway Diverge Areas, *TRB 2010 Annual meeting CD-ROM,* Washington D.C.

Appendix A

VARIATION IN ODDS VALUE

A.1 Urban Expressway

The following figures demonstrate the variation in *Odds* value caused by the interaction of geometry, traffic flow and ambient conditions, for different categories of traffic conditions. Figure A.1 describes the variation in *Odds* value caused by the interaction of v_S and S for daytime and nighttime considering the rise in k_S and h in low-density uncongested flow, in the purpose to identify the contributions of those factors to crash risk. In view of traffic characteristics shown in Figure 3.12, only low-density traffic (k_S=10veh/km) is analyzed for nighttime. By analogy, Figure A.2 and A.3 illustrate the related variation in *Odds* value in high-density uncongested flow and congested flow, respectively.

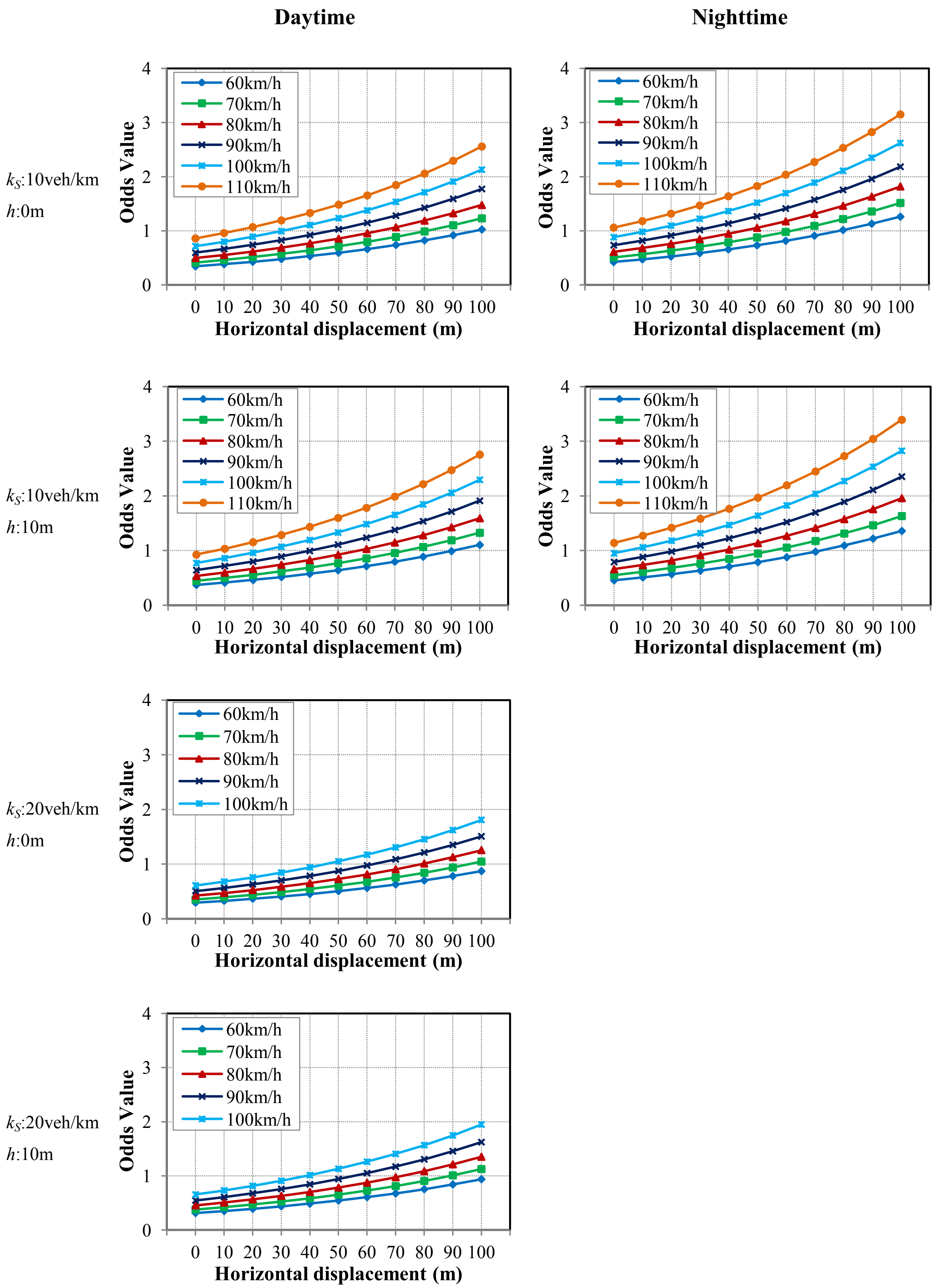


Figure A.1 Variation in *Odds* value in low-density uncongested flow

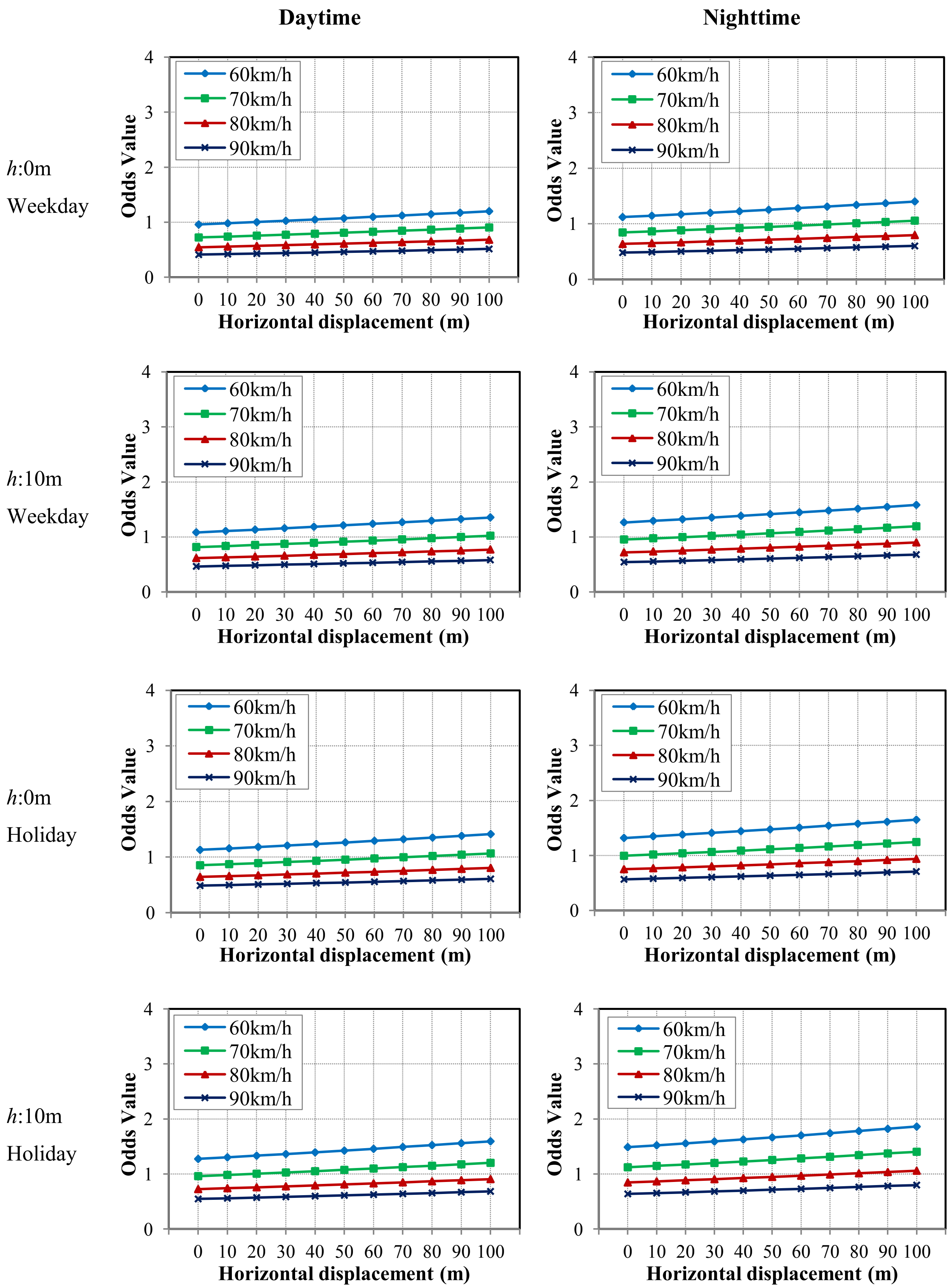


Figure A.2 Variation in *Odds* value in high-density uncongested flow

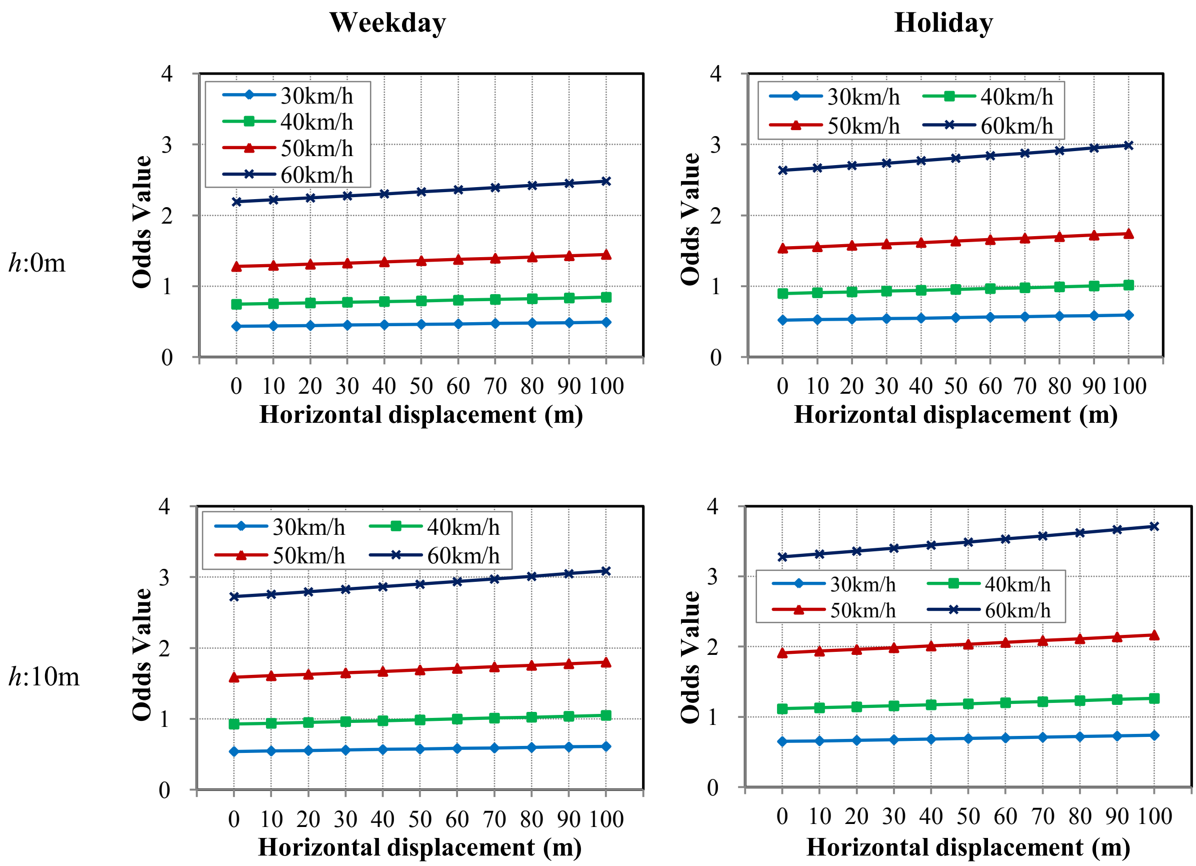


Figure A.3 Variation in *Odds* value in congested flow

A.2 Intercity Expressway

In the same way, in order to identify the contributions of individual variables to crash risk, Figure A.4 to A.6 demonstrate the variation in *Odds* value on intercity expressway caused by the interaction of geometry, traffic flow and ambient conditions, for low-density, high-density uncongested flow and congested flow, respectively. As analyzed in Chapter 3, by expressway type, holiday traffic is more notable on intercity expressway, and nighttime traffic is more distinct on urban expressway. Since the effect of nighttime on crashes has been analyzed on urban expressway above, this section will focus on *Day* type.

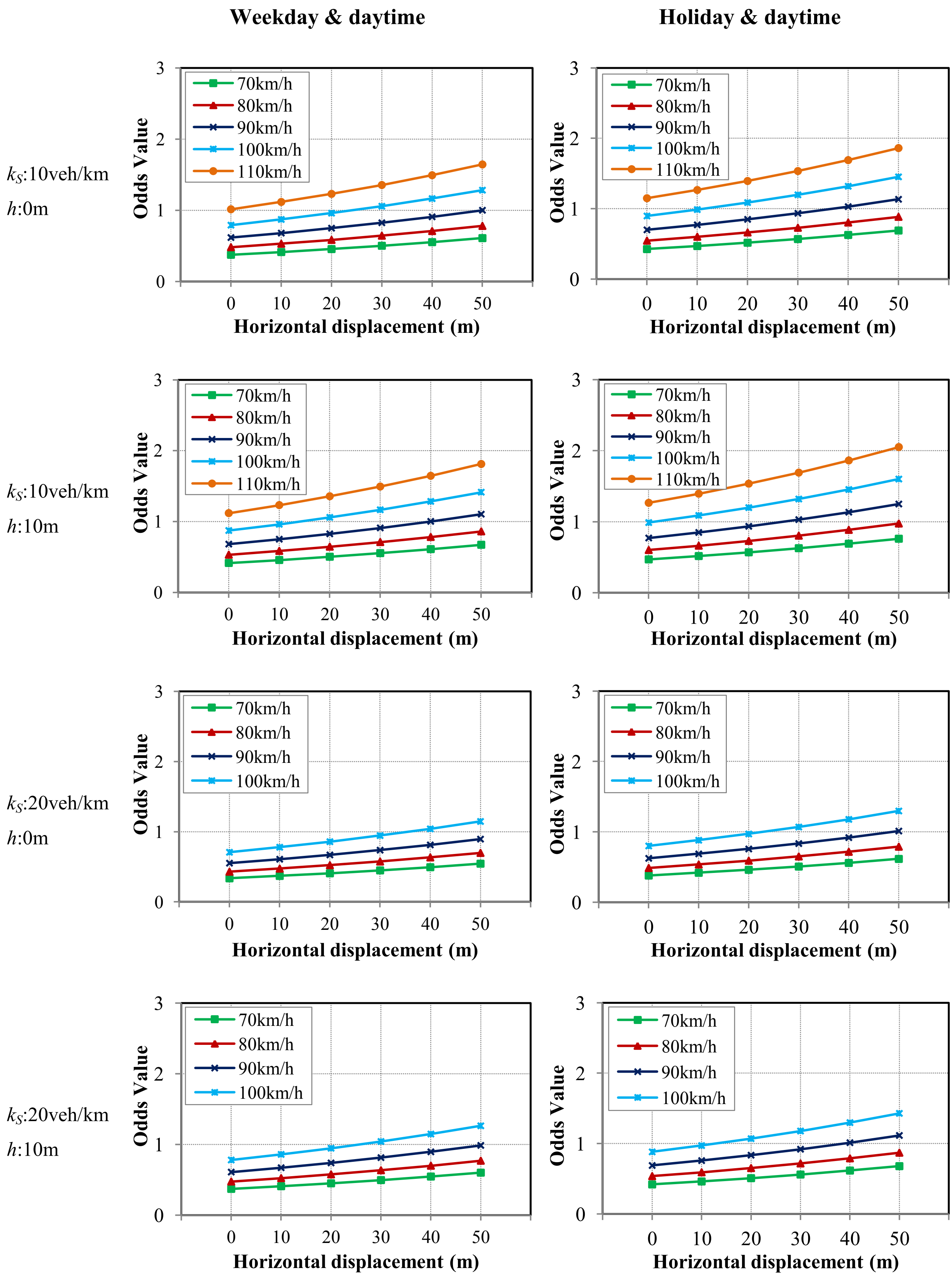


Figure A.4 Variation in *Odds* value in low-density uncongested flow

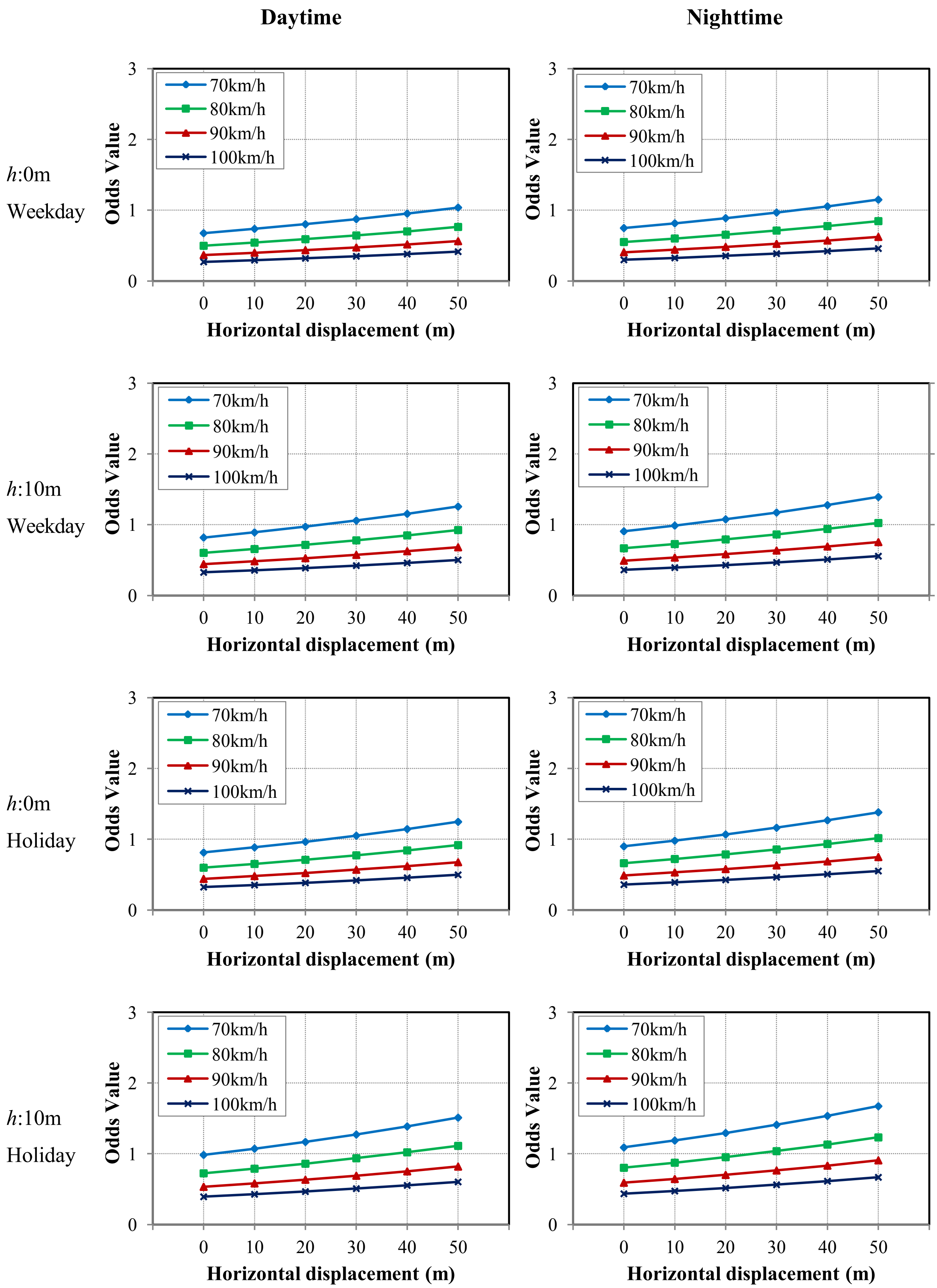


Figure A.5 Variation in *Odds* value in high-density uncongested flow

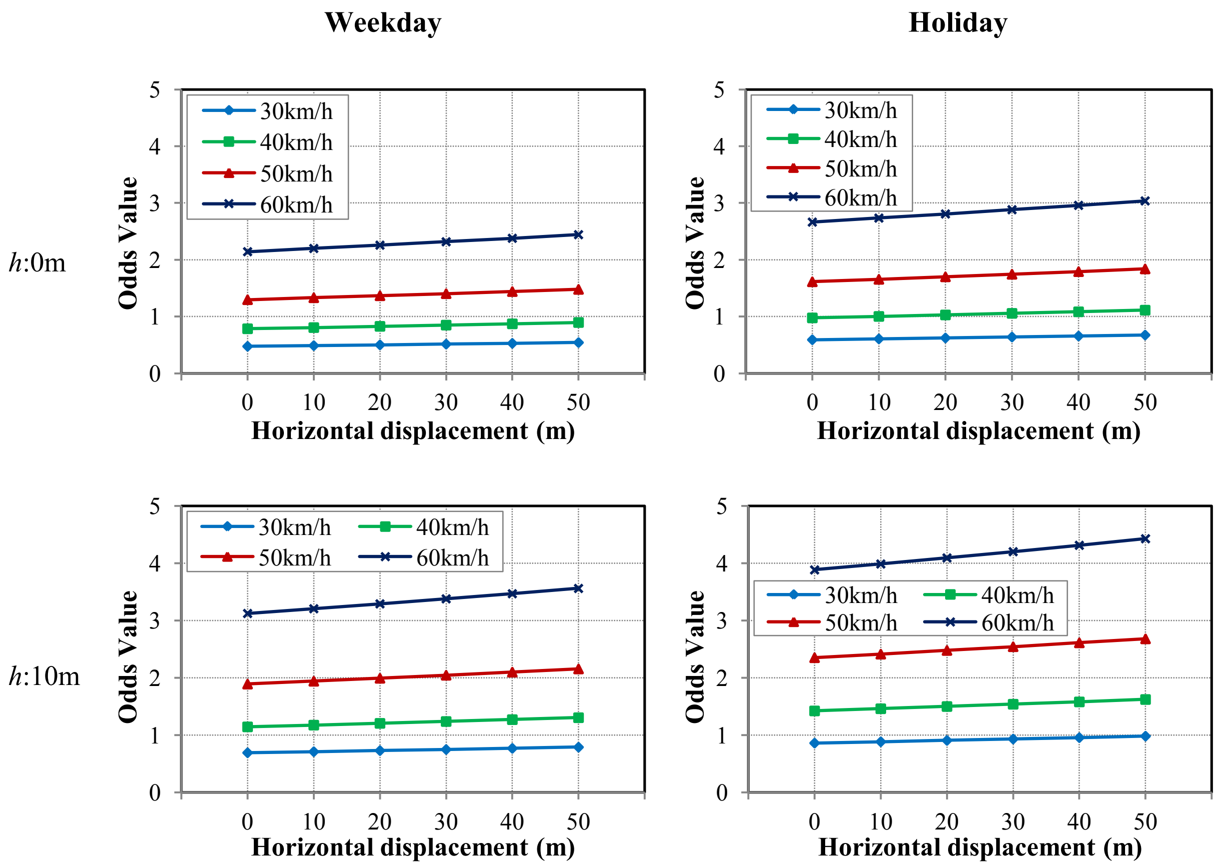


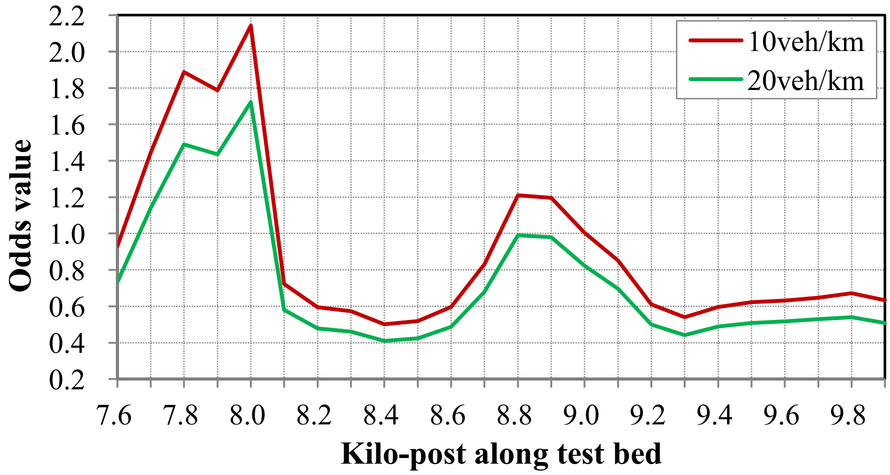
Figure A.6 Variation in *Odds* value in congested flow

Appendix B

CRASH RISK ESTIMATION AT THE TEST BEDS CONCERNING AMBIENT CONDITIONS

B.1 Crash Risk on Odaka Route

The values of crash risk per 0.1km along the test bed of Odaka route are predicted through separating traffic conditions considering ambient *Light* as shown in Figure B.1 and *Day* type as shown in Figure B.2. Considering the variables incorporated into crash modeling, traffic conditions leaving out scenario 6 to 8 are analyzed for ambient *Light*. For *Day* type, the traffic conditions from scenario 3 to 8 are selected.



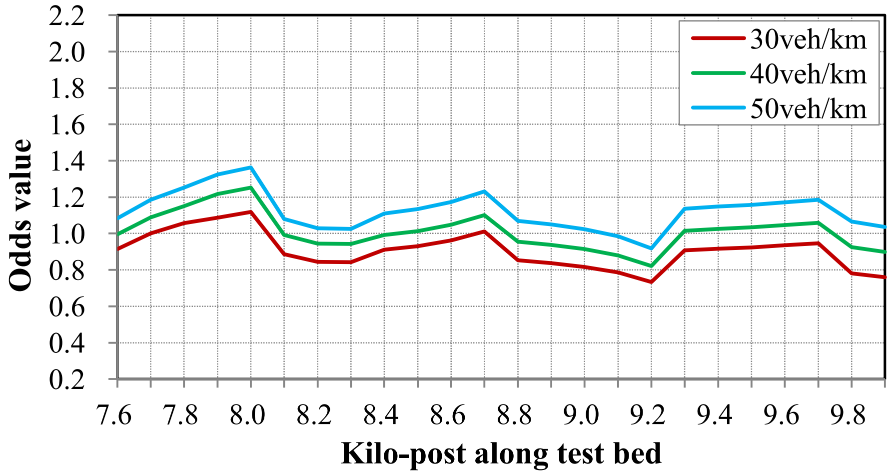


Figure B.1 Crash risk in the test bed of Odaka route (Nighttime)

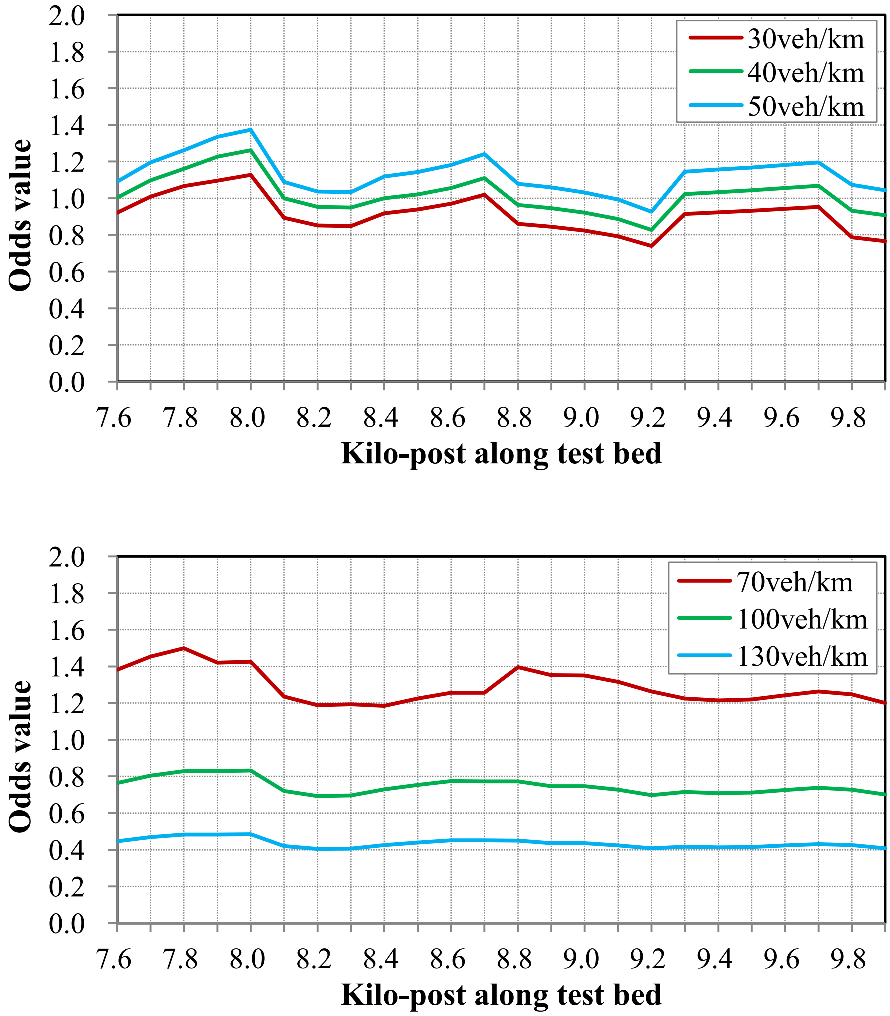


Figure B.2 Crash risk in the test bed of Odaka route (Holiday)

B.2 Crash Risk on Tomei Expressway

In the following figures, the values of crash risk per 0.1km along the test bed of Tomei Expressway are predicted through separating traffic conditions considering ambient *Light* as shown in Figure B.3 and *Day* type as shown in Figure B.4. Since ambient *Light* is not involved in CREM of congested flow, the following figures about the traffic in nighttime are also not considered the situations of scenario 6 to 8. On intercity expressway, traffic on holidays is quite distinguishing from that on weekdays, and *Day* type is involved in CREM for the whole traffic conditions.

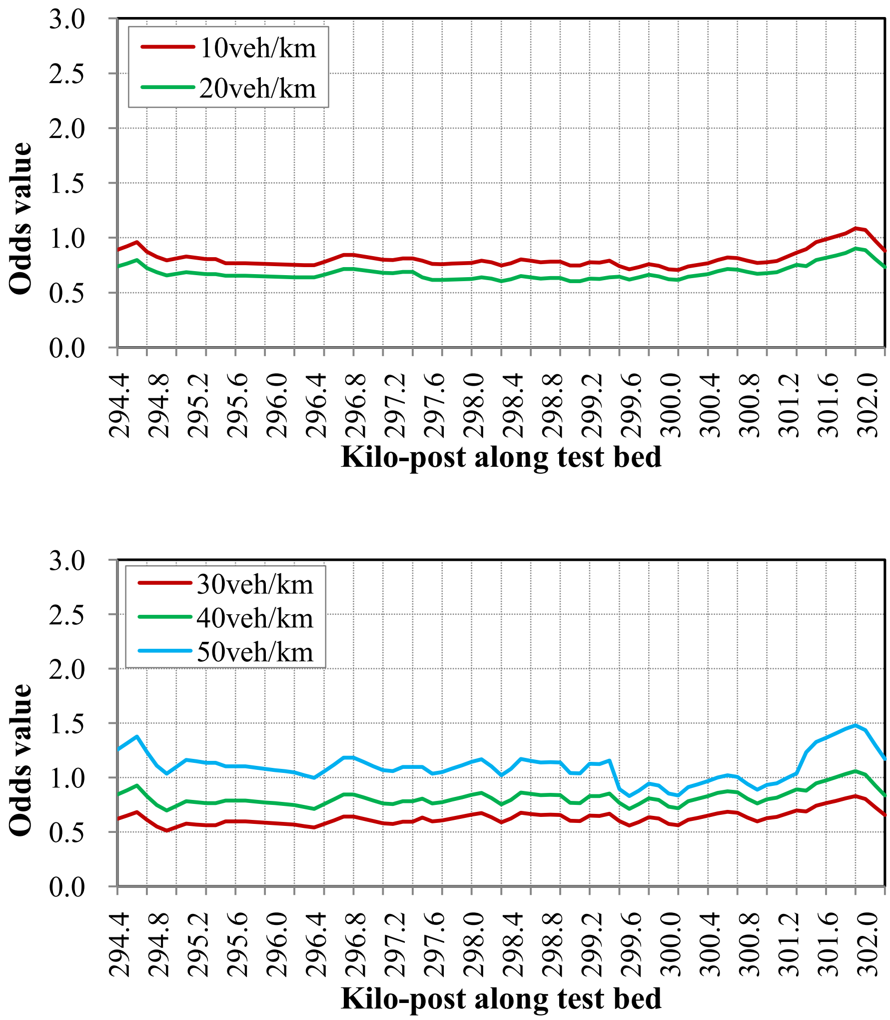
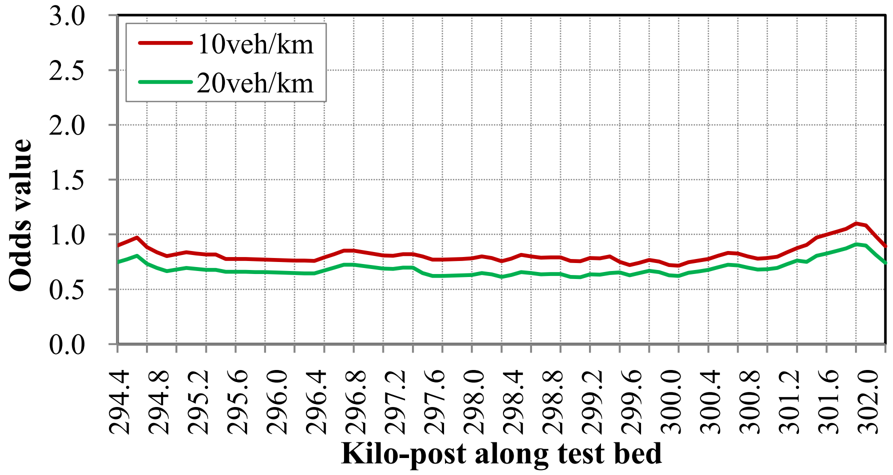
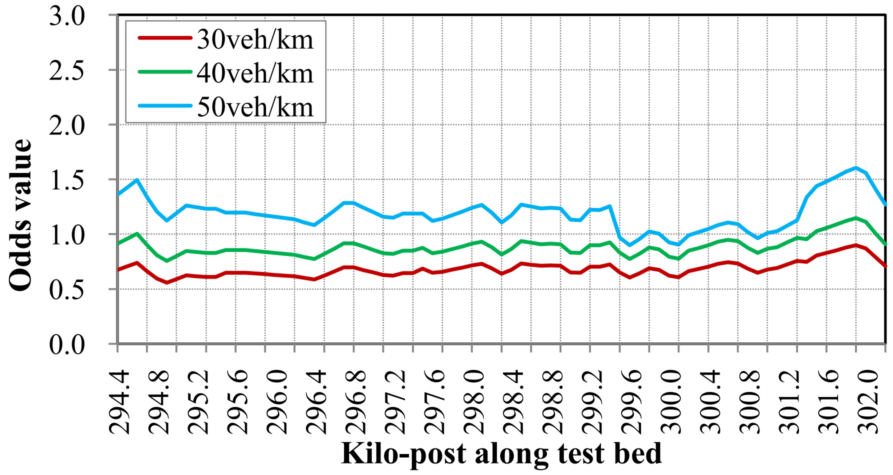


Figure B.3 Crash risk in the test bed of Tomei Expressway (Nighttime)





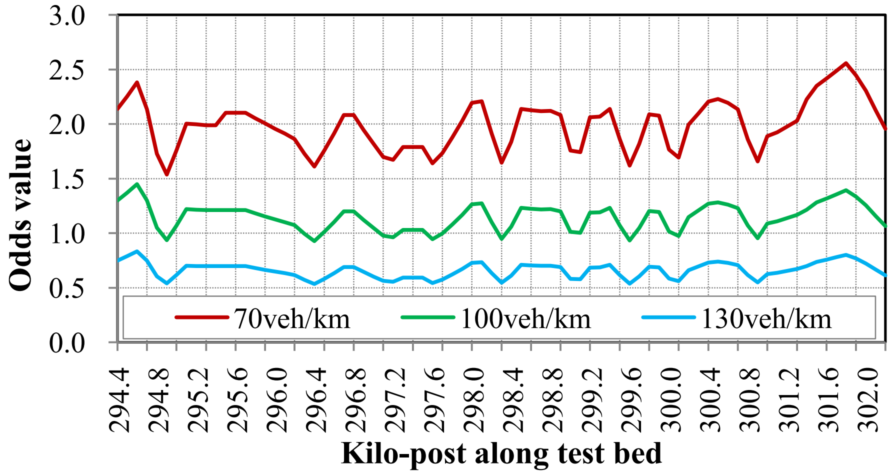


Figure B.4 Crash risk in the test bed of Tomei Expressway (Holiday)