Please cite this paper as follows: Roshandel, Saman, Zheng, Zuduo, Washington, Simon, 2015. Impact of real-time traffic characteristics on freeway crash occurrence: Systematic review and meta-analysis. Accident Analysis and Prevention. Accepted.

Impact of Real-time Traffic Characteristics on Freeway Crash Occurrence: Systematic Review and Meta-analysis

Saman Roshandel
Research student
Civil Engineering & Built Environment School,
Queensland University of Technology
2 George St GPO Box 2434, Brisbane QLD 4001, Australia
Email: saman.roshandel@student.qut.edu.au

Dr Zuduo Zheng (Corresponding author)
Civil Engineering & Built Environment School,
Queensland University of Technology
2 George St GPO Box 2434, Brisbane QLD 4001, Australia
Tel.: +61 07 3138 9989
Fax: +61 07 3138 1170
Email: zuduo.zheng@qut.edu.au

Professor Simon Washington
Civil Engineering & Built Environment School,
Queensland University of Technology
2 George St GPO Box 2434, Brisbane QLD 4001, Australia
Tel.: +61 07 3138 9990
Fax: +61 073 138 1170
Email: simon.washington@qut.edu.au

Abstract: The development of methods for real-time crash prediction as a function of current or recent traffic and roadway conditions is gaining increasing attention in the literature. Numerous studies have modelled the relationships between traffic characteristics and crash occurrence, and significant progress has been made. Given the accumulated evidence on this topic and the lack of an articulate summary of research status, challenges, and opportunities, there is an urgent need to scientifically review these studies and to synthesize the existing state-of-the-art knowledge.

This paper addresses this need by undertaking a systematic literature review to identify current knowledge, challenges, and opportunities, and then conducts a meta-analysis of existing studies to provide a summary impact of traffic characteristics on crash occurrence. Sensitivity analyses were conducted to assess quality, publication bias, and outlier bias of the various studies; and the time intervals used to measure traffic characteristics were also considered. As a result of this comprehensive and systematic review, issues in study designs, traffic and crash data, and model development and validation are discussed. Outcomes of this study are intended to provide researchers focused on real-time crash prediction with greater insight into the modelling of this important but extremely challenging safety issue.

Keywords: crash prediction; traffic characteristics; road safety; systematic review; meta-analysis

1. Introduction

Because of its extreme importance, roadway safety is one of the most heavily studied topics in transport engineering, with the ultimate motivation of a majority of studies to reduce fatalities and injuries. There are a variety of research directions that may help to achieve this goal, including both reactive and proactive approaches, behavioral and engineering improvements, and vehicle design changes. A relatively recent pursuit has focused on potential relationships between roadway operational characteristics and temporally and spatially proximal crash risk. It is generally accepted that crash causes are complex and often the result of a confluence of numerous factors, including behavioral factors (e.g., a driver's mental state, fatigue, distraction, impairment, etc.), vehicle state of repair, traffic conditions (e.g., level of congestion, prevailing speeds), geometry (e.g., horizontal and vertical curves, sight distances, channelisation, etc.) and environmental factors (e.g., ice, snow, rain, etc.). Due to the relative ease of gaining information about real time roadway and operational factors relative to behavioural and vehicle factors — courtesy of electronic detection and control systems — there is interest in exploring whether relationships exist, and if so, how reliable and useful they might be for predicting crash risk.

Given the desire to develop crash prediction models that are responsive to real time (or nearly so) traffic conditions on freeways, it is worth acknowledging the challenges and opportunities that confront this research:

i) Given the absence of behavioural influences on crash risk known to contribute to upwards of 80% of all crashes, false negative and positive rates are *priori* threats to

such models. In other words, traffic conditions alone may be found to constitute an elevated crash risk, but without an additional behavioural factor to help differentiate the relative risk, the predicted crash risk shall remain low, giving rise to a high proportion of false positive predictions.

ii) A theoretical relationship between microscopic traffic characteristics and crashes is lacking; and thus it is not clear what traffic 'signature' should be associated with elevated crash risk. It is possible that data mining approaches will reveal 'spurious' relationships.

iii) Establishing crash risk relationships on a short time scale has great intuitive appeal. Traffic conditions immediately upstream and preceding a crash should have bearing on crash risk, whereas exposure over the last year (for example) has a less obvious direct linkage to crashes.

iv) Traffic measurements are determined by device locations and capabilities, and may not be consistently placed and therefore be subject to statistical noise.

Despite the relatively recent interest in real time crash risk modelling on freeways, most freeway crash models have aimed to predict crash frequency for a particular road segment and/or to identify crash black spots (Moore et al., 1995; Davis, 2002; Cheng and Washington, 2005; Cheng and Washington, 2008; Washington et al., 2010; Meng and Qu, 2012). These models can be used to identify black spots where crashes have frequently occurred, and to evaluate the impact of regulations and/or interventions on a freeway's safety performance (e.g., a new speed limit's impact on the annual crash rate). However, this type of model is reactive, focusing on its historic safety performance to determine if remedial actions are warranted. These models typically rely on aggregate data, whereby safety performance is characterized over the most recent one or two years. Thus, the models are insensitive to real-time operational features of freeways. A crash prediction model with the ability to predict the probability of crash occurrence based on temporally and spatially proximal measurements (e.g. 50 meters upstream within the most recent minute) could substantially complement existing aggregate level models, and potential serve real-time safety management objectives. With such temporally and spatially proximal models, crash avoidance systems could be developed and implemented based on these models (Hourdos *et al*., 2008).

Safety modelling in the current literature is predominantly focused on aggregate level crash forecasting, with one to three year accident histories. In contrast, proactive, real time crash prediction models began appearing in the 1990s (Preston, 1996). Given the appeal to predict crash risk in real time with the aim to more proactively manage safety, the latter models have received rapidly increased attention recently, and notable progress has been made in identifying significant factors contributing to crash occurrence. In this literature, researchers have developed relationships between real-time traffic conditions (e.g., speed, density, volume, and their combinations) immediately preceding a crash, weather (e.g., rain, snow), and geometric features (e.g., curves, on-/off-ramps) and probability of crash occurrence. An assumption underlying these studies is that certain combinations of traffic conditions are relatively more 'crash prone' than others. Thus the research has focused on detecting and

quantifying crash-prone traffic conditions, and establishing their association with crash occurrence.

Numerous studies have investigated the connection between crash occurrence and traffic characteristics, and much has been learned from these investigations. For example, using loop detector data and crash reports, Lee *et al.* (2002) developed a real-time crash prediction model for vehicles travelling on freeways. A logistic linear approach rather than a binary logistic regression model was employed to address the over-preponderance of observations without crashes. They (Lee *et al*., 2002) report that measures of CVS (i.e., standard deviation of speed divided by average speed) and average density were related to crash occurrence. Based on a matched case-control design, Abdel-Aty *et al.* (2004) developed logistic regression models to measure the relationship between traffic flow variables and crash occurrence in real-time. After controlling external causes such as roadway geometry and time of day, speed variation and occupancy at the site of crash were found to be significant.

Other studies have examined the connection between traffic conditions and crash occurrence and revealed that volume, median speed, and temporal variations in speed and volume impact the likelihood of a crash (Garber and Wu, 2001; Golob and Alvarez, 2004; Abdel-aty *et al*., 2005; Hourdos *et al.,* 2006; Christoforu *et al*., 2011; Kuang et al., 2014). While investigating stop-and-go traffic oscillation on freeways, Zheng *et al*. (2010) report that speed variation (i.e., the standard deviation of speed during a specific period of time) is related to crash occurrence with an average odds ratio of around 1.08.

Although the use of real-time traffic data to identify crash-prone conditions* and to predict crash occurrence is promising, inconsistent performance and high prediction errors (e.g., false positive rates of 38.8% and 15% were reported in Abdel-Aty et al., (2004) and Hourdos et al., (2006), respectively) mean that this method is currently unsuitable for implementation at the real world operational level.

Many previous studies did not rigorously assess and/or report their models' predictive performance, such as false positive and negative prediction rates. Moreover, inconsistent and sometimes contradictory conclusions have been reported. For instance, results in Lee et al. (2003) suggest that increasing the value of coefficient of variation of speed (CVS) will reduce crash risk, while Abdel-Aty et al. (2006) reports the opposite.. To shed light on the preponderance of evidence in the research area, there is a need to comprehensively and systematically review previous studies, summarize their common findings, highlight their differences, identify the issues raised, and determine where future research is needed. To addresses this need, this paper provides a systematic literature review and meta-analysis of the current literature on this topic.

The remainder of this paper is organized as follows. Section 2 provides details of the systematic literature review, which provides the basis for the meta-analysis that follows in Section 3; Section 4 discusses issues arising at different stages of modelling the association between traffic characteristics and crash occurrence, and describes where future research should be directed; and, finally, Section 5 concludes the paper by summarizing its main findings.

---

* Traffic conditions with a higher likelihood of leading to a crash (Abdel-Aty and Pande, 2006; Hourdos et al., 2008).

2. The systematic literature review

In this section a review of relevant papers is provided to catalogue the research progress made in the prediction of crashes on freeways. To ensure an exhaustive search, internet searches, backtracking references, and contacting authors were undertaken. A systematic literature search of five databases included *ScienceDirect, Scopus, MetaPress*, *ProQuest*, and *Google Scholar.* The keywords used in the study were "crash prediction", "crash precursor", "traffic flow", "traffic condition", and "real-time". Several studies did not report the statistical features of their results in sufficient detail. To address this issue, authors were contacted via email to obtain additional information. Papers for which sufficient details could not be obtained were excluded from the analysis.

2.1. Coding for the systematic review

To facilitate the systematic review, a coding system was developed to extract information from the relevant studies. This system is summarized in Table 1. More specifically, the following coding variables were selected:

i) *Publication year* Studies published between 1997 and 2012 are targeted in this study (No notable journal papers on predicting crash occurrence using real-time traffic characteristics were discovered prior to 1997).

ii) *Study design* This variable is used to capture information on the study design methods employed by the previous studies.

iii) *Data type* This variable is used to indicate whether loop detector data or vehicular trajectories were used in a study.

iv) *Traffic characteristics* These variables that used to measure traffic characteristics; e.g., speed (average, difference, variation, coefficient of variation), density (average and difference), and volume.

v) *Estimator of variable* This variable is used to indicate how a study reports its modelling results (Most studies report their results as logarithms of odds-ratios).

vi) *Controlled confounders* The main categories of confounding variables are geometry, traffic states, weather, and environmental conditions.

vii) *Model validation (*or *application)* indicates whether the model was validated.

2.2. Study selection for meta-analysis

In the initial literature review, 99 studies were identified as potentially relevant. After further screening, this number was reduced to 25. Among these 25 studies, 13 were selected for meta-analysis because they focused on real-time crash occurrence, and also provided sufficient statistical information about their models (as shown in Table 2). In total, these studies contain 46 estimated effects.

Over 60 per cent of the 25 studies were published between 2001 and 2006. Most studies used matched case-control design and logistic regression to assess possible relationships. Crash data were extracted from local police reports or databases. Loop detector data were used in most studies to measure traffic conditions for a specific freeway segment, while trajectories extracted from video surveillance facilities were used in two studies (see Table 2).

Table 1 *Coding of the studies*

Variable	Codes applied
Study identification	Studies were numbered from earliest to most recent
Publication year	1997 to 2012
Study design	Case-control (CC); Before-after (BA); Random sampling (RS)
Data type	Loop detector (LD); trajectory data (TD)
Traffic characteristics	Average speed (S); speed variation (SV); coefficient of variation of speed (CVS); speed difference (SD); average density (D); density difference (DD); average volume (V); volume difference (VD)
Estimator of variable	Regression coefficient (RC); odds-ratio (OR)
Confounders controlled	Time of day (T); location (L); geometry (G); weather conditions (W)
Model validation (application)	Performed (P); not performed (NP)

With regard to the association between crash occurrence and traffic characteristics, 12 studies reported more than two traffic variables influenced crash occurrence, while others reported more than one estimate for a specific traffic variable depending on traffic and environmental conditions (e.g., weather condition, time of the day, visibility and pavement condition) (Garber and Wu, 2001; Lee *et al*., 2002; Lee *et al*., 2003; Abdel-Aty *et al*., 2005; Hourdos *et al.*, 2006; Park and Oh, 2009; Christoforou *et al.,* 2011). One of the 13 studies reported effects using odds-ratios, while others reported the log of the odds-ratio. Potential confounding factors such as time of day, location, roadway geometric characteristics (e.g., curvature), traffic congestion and weather conditions (e.g., wet or dry pavement) were controlled in most studies. Eight (61%) of 13 studies validated the performance of their models (Garber and Wu, 2001; Lee *et al.*, 2002; Lee *et al.*, 2003; Abdel-Aty and Pande, 2006; Hourdos *et al.*, 2006; Hourdos *et al.*, 2008; Zheng *et al.,* 2010; Hossain and Muromachi, 2012) . (Note that 'model application' and 'model validation' are considered as the same process in this study.)

Table 2 *Studies included in the meta-analysis*

Study ID	Study title	Authors	Publication year	Study design	Traffic flow variables	Type of data used
1	Stochastic models relating crash probabilities with geometric and corresponding traffic characteristics data	Garber and Wu	2001	unclear	Speed, speed variation, volume	Loop detector data
2	Analysis of crash precursors on instrumented freeways	Lee, Saccomanno and Hellinga	2002	Case-control	Density, CVS	Loop detector data
3	Real-time crash prediction model for application to crash prevention in freeway traffic	Lee, Hellinga, and Saccomanno	2003	Case-control	Density, CVS	Loop detector data
4	Predicting freeway crashes from loop detector data by matched case-control logistic regression	Abdel-Aty, Uddin, Pande, Abdalla, and Hsia	2004	Case-control	Density, standard deviation of volume, CVS	Loop detector data
5	Split Models for predicting multi-vehicle crashes during high-speed and low-speed operating conditions on freeways	Abdel-Aty, Uddin and Pande	2005	Case-control	Density, volume, standard deviation of volume, CVS	Loop detector data
6	Calibrating a real-time traffic crash-prediction model using archived weather and ITS traffic data	Abdel-Aty and Pemmanaboina	2004	Case-control	Density, standard deviation of volume, CVS	Loop detector data
7	ATMS implementation system for identifying traffic conditions leading to potential crashes	Abdel-Aty and Pande	2006	Case-control	Density, speed variation, CVS	Loop detector data
8	Real-time detection of crash-prone conditions at high-crash freeway locations	Hourdos, Garg, Michalopoulos and Davis	2006	Case-control	Speed, speed difference, volume, density, density difference	Trajectory data
9	Accident prevention based on automatic detection of accident prone traffic conditions: Phase I	Hourdos, Garg, Michalopoulos and Davis	2008	Case-control	Speed, speed difference, volume, density,	Trajectory data

Study ID	Study title	Authors	Publication year	Study design	Traffic flow variables	Type of data used
					density difference	
10	Relating freeway traffic accidents to inductive loop detector data using logistic regression	Park and Oh	2009	Case-control	Speed, speed variation, volume, density	Loop detector data
11	Impact of traffic oscillations on freeway crash occurrences	Zheng, Ahn, and Monsere	2010	Case-control	Speed variation	Loop detector data
12	Identifying crash type propensity using real-time traffic data on freeways	Christoforou, Cohen and Karlaftis	2011	Case-control	Speed, density, volume	Loop detector data
13	A Bayesian network-based framework for real-time crash prediction on the basic freeway segments of urban expressways	Hossain and Muromachi	2012	Case-control	Speed difference and density difference	Loop detector data

Table 3 lists those studies excluded from our analysis. The main reasons for their exclusion are as follows:

i) Some studies (Lee *et al.,* 2006; Pande, 2005) estimated the impact of traffic flow conditions on a specific crash types (e.g., sideswipe, multi-vehicle, visibility-related). As this study focuses on all crash type rather than a specific type of crash, it is inappropriate to combine results from general crashes with those from a specific crash type as the effect sizes are incomparable.

ii) Some studies are based on aggregated loop detector data (e.g., AADT), which are inherently not suitable for testing the association between real-time traffic conditions and crash occurrence. Thus, these studies were excluded from the meta-analysis (Liu, 1997; Pei *et al.*, 2012).

Table 3 *Reason for excluding studies from the meta-analysis*

Study ID	Authors	Year	Reason for exclusion
1	Lee and Abdel-Aty	2008	Study area of this study was freeway ramps
3	Lee, Abdel-Aty and Hsia	2006	This study examines the relationship between traffic condition and a specific type of crash; i.e., sideswipe crashes
4	Abdel-Aty, Pemmanaboina, and Hsia	2006	The main objective of this study was to examine the relationship between crash frequency and geometric design
5	Kockelman and Ma	2007	Due to the methodology used in this study, estimates reported in this study cannot be combined with estimates from other studies
6	Moore, Dolinis and Woodward	1995	This study focused on the relationship between crash severity and traffic conditions
7	Pande, Abdel-Aty, and Hsia	2005	The main aim of this study was to investigate spatiotemporal variation of crash risk on freeways
8	Pande and Abdel-Aty	2006	Contains no specific model outputs
9	Oh, Oh, Ritchie and Chang	2000	The objective of this study was to demonstrate the feasibility of identifying crash-prone traffic conditions prior to a crash
10	Golob and Recker	2004	This study aimed to identify traffic regimes where crashes are likely to occur
11	Pande	2005	This study focused on rear-end and lane-changing-related crashes
12	Liu	1997	Highly aggregated data was used in this study

2.3. Evaluating quality of the selected studies

In a research synthesis, assessing the quality of studies prior to conducting a meta-analysis is a critical step. Although Elvik (2008) questions whether numerical scales are capable of reflecting the quality of road safety evaluation studies, Cooper *et al.* (2009) regard them as important for assisting researchers in deciding whether or not to include a study in a research synthesis. Furthermore, they suggest that examining the quality of studies prior to performing a systematic review increases the reliability of the review. Therefore, consistent with Cooper, a scoring system was developed (see Table 4) incorporating the following criteria:

i) Did the study control for potential confounders (e.g., geometric characteristics, weather and environmental conditions)?

ii) What types of data were used in the study?

iii) Was the model validated against other sites or another time period?

Information on these three criteria is reported in the selected studies; thus, there is no risk of wrongly scoring studies due to missing information (Elvik, 2013). Noteworthy is that control for external confounders is the most important criterion, and represents 42.8% of the maximum score

(see Table 4). To minimise the risk of being dominated by any single criterion, the minimum possible score from each criterion is the same (i.e., 1 unit), and the maximum possible score from each criterion is similar (as indicted in Table 4). To further ensure the objectivity of how a quality score is assigned to each study, each study has been assessed against these criteria by the authors independently, and the score assigned to each study has been cross checked.

The quality scores for the selected studies are summarized in Table 5. The relationship between the final quality score and the publication year was examined to detect how the quality of studies has changed over time. Figure 1 reveals this relationship, where the publication year is located on the horizontal axis and relative quality score is on the vertical axis. For each study, the relative score was computed as the quality score of a study divided by the maximum possible score (i.e., 7). The correlation was statistically tested, and no significant relationship was detected (p-value=0.604).

Table 4 *Quality assessment criteria*

ID	Criterion	Scores assigned	Maximum score and percentage
C1	Control for external confounders (e.g., geometric characteristics; weather and traffic regimes)	No confounder (1); one to three confounders (2); More than three confounders (3)	3 (42.8%)
C2	Data type used	Loop detector data(1); Trajectory data (2)	2 (28.5%)
C3	Was the model validated against other sites or another time period	Not performed (1); Performed (2)	2 (28.5%)

Table 5 *Final quality scores*

Study title	C1	C2	C3	Total Score
Stochastic models relating crash probabilities to geometric and corresponding traffic characteristics data	2	1	2	5
Analysis of crash precursors on instrumented freeways	2	1	2	5
Real-time crash prediction model for application to crash prevention in freeway traffic	2	1	1	4
Predicting freeway crashes from loop detector data by matched case-control logistic regression	3	1	2	6
Split models for predicting multivehicle crashes during high-speed and low-speed operating conditions on freeways	3	1	1	5
ATMS implementation system for identifying traffic conditions leading to potential crashes	3	1	2	6
Calibrating a real-time traffic crash-prediction model using archived weather and ITS traffic				

Study title	C1	C2	C3	Total Score
data	3	1	1	5
Real-time detection of crash-prone conditions at freeway high-crash locations	2	2	2	6
Accident prevention based on automatic detection of accident-prone traffic conditions: Phase I	3	1	1	5
Relating freeway traffic accidents to inductive loop detector data using logistic regression	1	1	1	3
Impact of traffic oscillations on freeway crash occurrence	3	1	2	6
Identifying crash-type propensity using real-time traffic data on freeways	3	1	1	5
A Bayesian network based framework for real-time crash prediction on the basic freeway segments of urban expressways	1	1	2	4

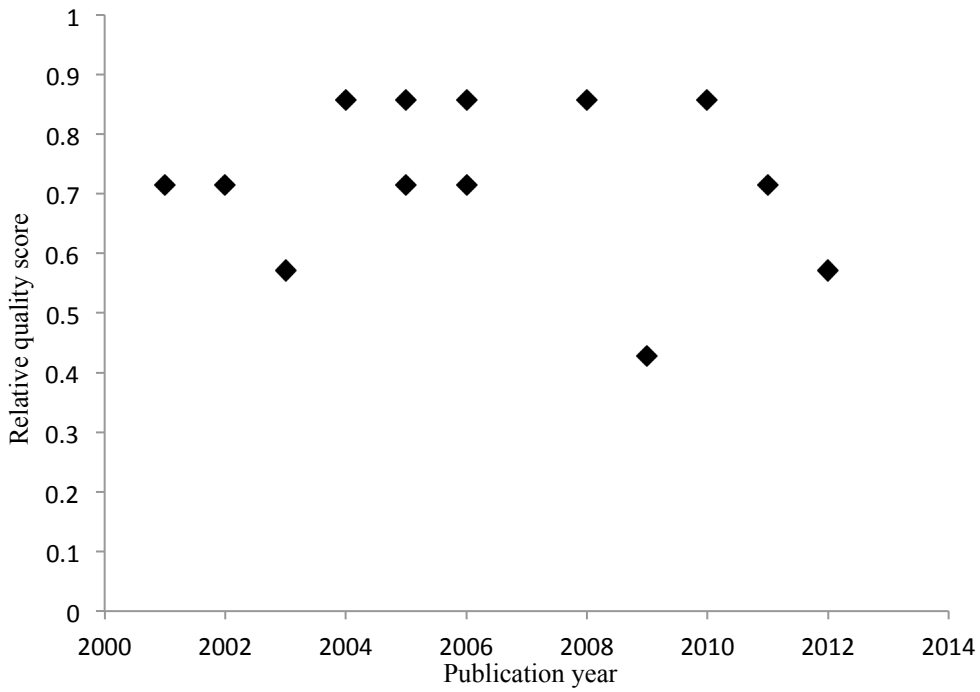


Figure 1. Relationship between the final quality scores and the publication year

3. Meta-analysis

Although a qualitative literature synthesis can enable researchers to better understand the results of one study in the context of other relevant studies, there is often a desire to quantitatively assess the magnitude and variance of effect size (i.e., the estimate of a factor's impact on crash occurrence) across studies. If consistent, a more accurate and precise effect size can be estimated; otherwise,

inconsistencies in effect are revealed[†]. Meta-analysis, applied here, is a well-established method for achieving these goals, whose two main objectives are to

i) obtain the summary effect size of the potential traffic flow variables estimated in prior studies, and
ii) assess the consistency of the variable estimates reported in the literature.

Table 6 summaries the design features of the meta-analysis conducted in this study. It shows three groups of traffic characteristics (i.e., speed, density, and volume) and various combinations. In total, 44 estimates were extracted.

3.1. Effect size and statistical weight

In order to conduct a meta-analysis, effect sizes and statistical weights are computed for each variable estimate.

Effect size indicates the strength of a relationship between two variables, and the summary effect size is obtained by combining different effect sizes for the same variable reported in different studies. Most of the studies listed in Table 2 applied matched case-control (binary data), and all of these studies reported either odds-ratios or log-odds ratios. An odds ratio (OR) is known as the ratio of the odds of an event (e.g., a crash) occurring in one group (e.g., the case group) to the odds of it occurring in another group (i.e., the control group).

Statistical weights were assigned to each variable as inversely proportional to the squared standard error of the estimate such that more precise estimates receive larger weights, as shown in Equation (1) (Borenstein *et al.,* 2011):

$$w = \frac{1}{SE^2} \tag{1}$$

where *w* stands for statistical weight and SE for standard error.

For studies that reported confidence intervals rather than standard errors, statistical weights are instead computed using Equation (2) (Borenstein *et al.*, 2011):

$$w = \frac{1}{(\ln\,(upper95\%) - \ln\,(lower95\%))/3.92)^2} \tag{2}$$

where upper (lower) 95% stands for the upper (lower) 95% confidence level.

3.2. Fixed-effect or random-effect model?

The next step in conducting a meta-analysis is to decide whether to develop a fixed-effect (FE) or a random-effect (RE) model. The FE model assumes that estimates (effect sizes) across studies

[†] As one referee pointed out, while inconsistencies might exist, they could be caused by a variety of reasons. And for each individual study, the conclusion about specific effect of variable could be convincing in context of its study object. Thus there is perhaps no need to have consistent effects in all cases. The inconsistency also does not necessarily impair the strength or quality of these studies.

share the same unobserved true value, and that all observed differences among effect sizes arise from sampling error. Alternatively, the RE model assumes that there are multiple unobserved true values that reflect unobserved differences across sites.

As shown in Equation (3), a statistical test (Q-test) is used to test heterogeneity, and helps to identify which model is more appropriate for each variable tested (Borenstein *et al.,* 2011).

$$Q = \sum_{i=1}^{g} w_i y_i^2 - \frac{\left(\sum_{i=1}^{g} w_i y_i\right)^2}{\sum_{i=1}^{g} w_i y_i} \tag{3}$$

where y_i is the estimate reported by study *i* (e.g., a log odds ratio from study *i*), w_i is the statistical weight assigned to study *i*, and *g* is the number of studies that are combined to compute the summary effect size. *Q* is approximately chi-square distributed with *(g-1)* degrees of freedom under the null hypothesis that the FE model is appropriate. If the *Q* does not arise by chance, the null hypothesis is rejected in favour of the RE model. The Q-test results for each variable in the meta-analysis are summarized in Table 6. The analysis suggests that the FE model is most appropriate for speed difference and density variation as Q-test conducted for estimates of these two variables was not statistically significant. However this test showed that there is considerable heterogeneity between estimates of other variables. Thus the RE model was selected for them.

For variables where an RE model is appropriate (see Table 6), variances (i.e., random effect variances) were calculated using Equation (4):

$$v_i^* = v_i + \tau^2 \tag{4}$$

where v_i is the within-study variance and τ^2 represents the between-studies variance, which is computed using Equations (5) and (6).

$$\tau^2 = \frac{Q-(g-1)}{C} \tag{5}$$

$$\mathrm{c} = \sum_{\mathrm{i}=1}^{\mathrm{g}} \mathrm{w_i} - \left[\frac{\sum_{\mathrm{i}=1}^{\mathrm{g}} \mathrm{w_i^2}}{\sum_{\mathrm{i}=1}^{\mathrm{g}} \mathrm{w_i}}\right] \tag{6}$$

Statistical weights were then updated accordingly, and the summary effect size of each variable was computed as a weighted average. This is shown in Equation (7)

$$\bar{y} = \exp\left(\frac{\sum_{i=1}^{g} w_i y_i}{\sum_{i=1}^{g} w_i}\right) \tag{7}$$

where $\bar{y}$ is the summary effect size, y_i and w_i are the estimates and statistical weights of each estimate respectively in Study *i*. The lower and upper limits that define 95% confidence interval boundaries of the summary effects were computed using Equations (8) and (9):

$$LL = \exp\left[\left(\frac{\sum_{i=1}^{g} w_i y_i}{\sum_{i=1}^{g} w_i}\right) - \left(\frac{1.96}{\sqrt{\sum_{i=1}^{g} w_i}}\right)\right] \tag{8}$$

$$UL = \exp\left[\left(\frac{\sum_{i=1}^{g} w_i y_i}{\sum_{i=1}^{g} w_i}\right) + \left(\frac{1.96}{\sqrt{\sum_{i=1}^{g} w_i}}\right)\right] \tag{9}$$

3.3. Publication bias and trim-and-fill approach

Publication bias occurs when a meta-analysis fails to include all relevant studies, and thus reduces the trustworthiness of the meta-analysis results. These undesired studies are either unpublished studies or studies with findings that are difficult to interpret (Hoye and Elvik, 2010). Thus, unpublished studies were sought to minimise publication bias. However, we have not obtained any unpublished studies from researchers who may have unpublished materials relevant to this topic, based on the information we gathered.

A funnel plot (Light and Pillemer, 1984) can be used to assess the existence of publication bias in the meta-analysis. In order to generate a funnel plot, the values of estimates and their related standard errors are placed on the *X* and *Y* axes respectively (Light and Pillemer 1984). In the absence of publication bias, the plot should by approximately shaped like an upside-down funnel, and reveal symmetry of data points with respect to the vertical axis.

Furthermore, to assess the magnitude of publication bias for each variable, trim-and-fill was implemented in the meta-analysis. As a non-parametric technique, trim-and-fill detects the possible existence of publication bias by examining the asymmetry in the funnel plot through three ranking-based estimators (Duval and Tweedie, 2000). A step-by-step description of this approach is provided in Hoye and Elvik (2010). This technique was performed for all variables included in this meta-analysis, the results of which are summarized in Table 6. As shown in this table, two cases of publication bias were detected. These were subsequently adjusted by adding extra data points, as recommended in Hoye and Elvik (2010). Additional details of this adjustment are discussed later in this paper.

Table 6 *Meta-analysis design of study*

Variable		Number of estimates	Test for heterogeneity	Model type	Application of trim-and-fill	Data points added
Speed	Average speed	10	Positive	RE	Yes	0
	Speed Variation[1]	5	Positive	RE	Yes	0
	CVS[2]	7	Positive	RE	Yes	1
	Speed difference[3]	2	Negative	FE	Yes	0
Density	Average density	11	Positive	RE	Yes	2
	Density variation[4]	3	Negative	FE	Yes	0
Volume	Average volume	6	Positive	RE	Yes	0

[1] The standard deviation of speed within a certain time interval right before the crash
[2] Standard deviation of speed divided by average speed ($CVS = \frac{SS}{AS}$)
[3] The difference in speed between one specific downstream and upstream loop detector position
[4] The absolute difference in density between the data for each time and the daily average

3.4. Main analysis and results

Table 7 presents the main results of the meta-analysis. It shows that three summary estimates (speed variation, speed difference and average volume) have statistically significant negative impacts on crash occurrence. This indicates that increasing values of these variables is associated with an elevated risk of crash. However, the summary estimate of speed difference should be interpreted with caution because of its small number of estimates (i.e., 2). More specifically, the tables shows that: i) the summary effect size (i.e., summary odds ratio) of speed variation is 1.225, which indicates that the odds ratio of a crash occurrence increases by 22.5% when speed variation increases by one additional unit; ii) the summary odds ratio of speed difference is 1.032, which indicates that the odds ratio of a crash occurrence increases by 3.2% when speed difference increases by one additional unit; and iii) the summary odds ratio of average volume is 1.001, which indicates that the odds ratio of a crash increases by 0.1% when average volume increases by one additional unit.

In contrast, average speed has a summary odds ratio of 0.952, which implies that increasing values of speed are associated with a reduced risk of a crash. More specifically, if average speed increases by one additional unit, the odds ratio of a crash decreases by 4.8%. This result seems counterintuitive. However, as explained in Zheng (2012), this result is not surprising because a stop and go driving conditions are associated with lower average speeds, and the outcome being assessed is crash occurrence and not severity. The summary odds ratio of density variation is 0.876, which implies that increasing density variation is associated with a reduced risk of a crash. This phenomenon is caused by the fact that density variation is measured as the absolute difference in density between the data for each time and the daily average. Thus, large density variations likely correspond to off peak hours.

As previously discussed, funnel plots and trim-and-fill techniques were used for each variable to detect and correct for potential publication bias. As a result, publication bias was detected in the estimates for CVS and average density. As shown in Figure 2, the estimate of CVS retrieved from Hourdos *et al.*, (2006) has a very large absolute value (-49.1) compared with other estimates; this causes asymmetry in the funnel plot. After applying the trim-and-fill technique, one data point (which was a mirror of the aforementioned estimate with the same standard error) – the triangular point in Figure 2 – was added to achieve symmetry. However, as shown in Table 7, even after correcting for publication bias, the summary estimate of this variable remained insignificant at a 5% level.

Similarly, publication bias was detected in average density, and two extra data points were added (the two triangular points in Figure 3). After correcting for publication bias, the summary odds ratio of average density was found to be significant. More specifically, if average density increases by one additional unit, the odds ratio of a crash occurrence decreases by 9.1%. This result is surprising because several studies (Abdel-Aty *et al.,* 2007; Golob and Recker, 2004; Golob *et al.,* 2008) report that congestion contributes to the occurrence of (rear-end) crashes. However, as discussed in the next section, this surprising result is due to location's confounding impact.

Noteworthy is that Table 7 shows that after accounting for publication bias, CVS is the only variable that is not statistically significant (at a 5% significance level). In addition, several studied have differentiated variables according to locations. Meta-analysis results for variables differentiated by locations are presented and discussed later.

Table 7 *Meta-analysis results*

Variable		Number of estimates	Summary odds ratio*	95% confidence interval	Summary odds ratio adjusted for publication bias*	95% confidence interval
Speed	Average speed	10	**0.952**	(0.909,0.996)		
	Speed variation	5	**1.225**	(1.132,1.326)		
	CVS	7	2.76	(0.969,7.863)	2.842	(0.985, 8.195)
	Speed difference	2	**1.032**	(1.026,1.038)		
Density	Average density	11	0.968	(0.902,1.039)	**0.909**	(0.831, 0.993)
	Density variation	3	**0.876**	(0.866,0.886)		
Volume	Average volume	6	**1.001**	(1.000,1.002)		

* The bolded numbers are statistically significant.

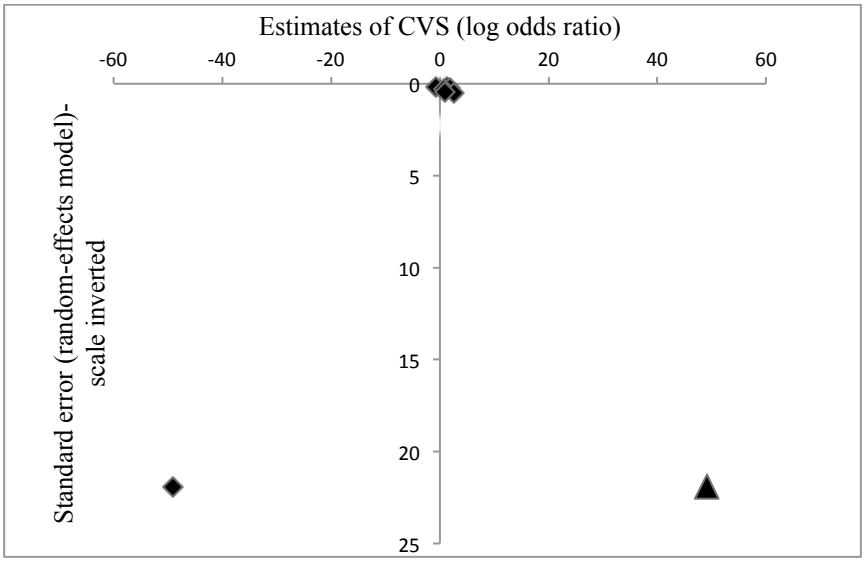


Figure 2. Funnel plot of estimates of CVS adjusted for publication bias (random effects)

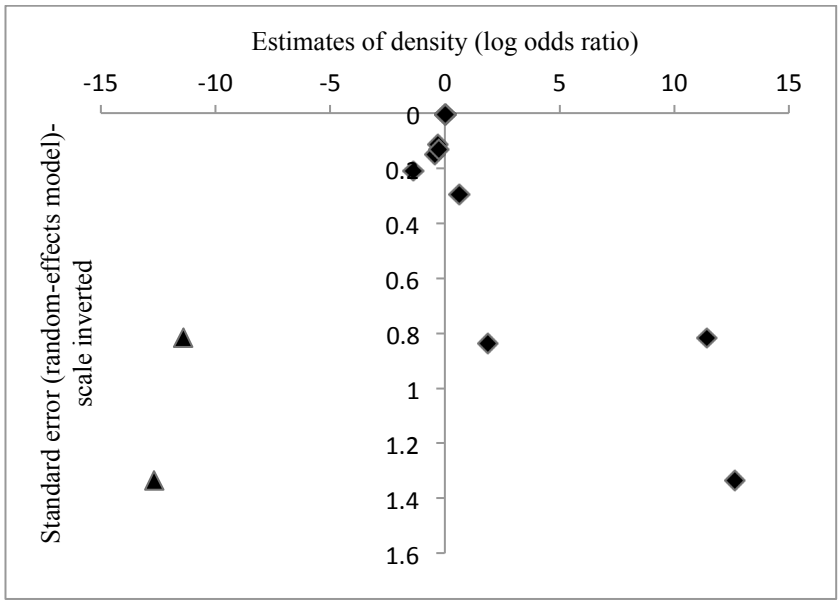
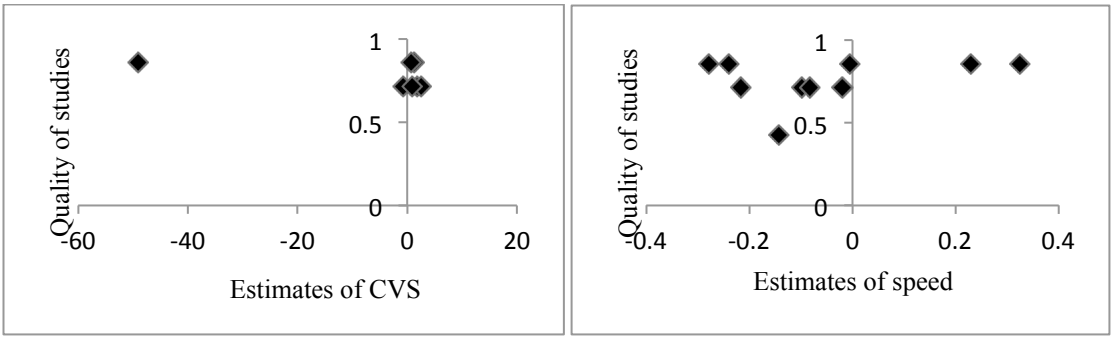


Figure 3. Funnel plot of estimates of density adjusted for publication bias (random effects)

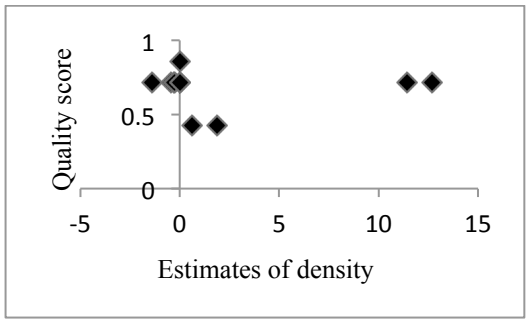
3.5. Sensitivity analysis

Moderators may affect the results of a meta-analysis; therefore, it is essential to assess the sensitivity of the results (Borenstein *et al.*, 2011). Accordingly, the sensitivities of the summary estimates reported in the previous section were tested with respect to quality, outlier bias, the time interval chosen for measuring traffic flow characteristics, and locations where variables were measured.

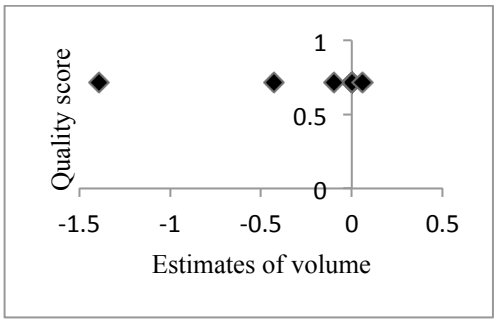
Low-quality studies can distort meta-analysis results. The relationship between study quality and the magnitude of estimates was separately examined for each traffic flow variable, as shown in Figure 4. This figure revealed no notable relationship between an estimate of a traffic characteristic variable and a study's quality. To confirm this result, the correlation matrix was produced and summarized. As seen in Table 8, the same conclusion was reached. Of note, this test was not applied to speed difference and density variation due to the small numbers of estimates (fewer than 5).



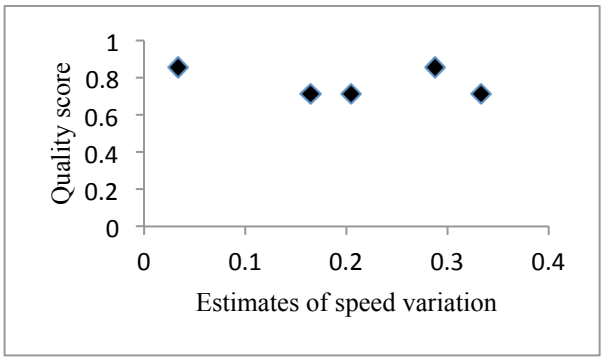
(a) (b)



(c)



(d)



(e)

Figure 4. Relationship between a study's relative quality score and magnitude of estimates (in log odds ratios)

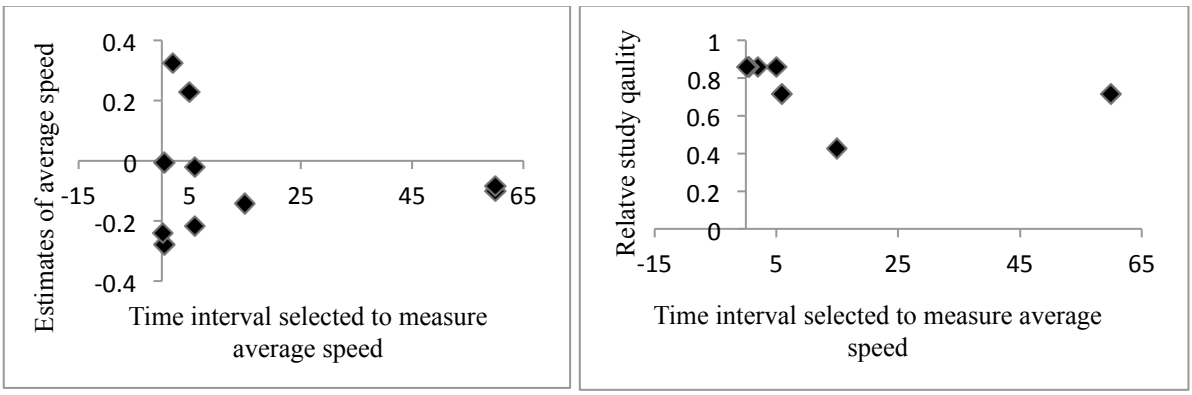
Table 8 *Coefficient of correlation between traffic flow variables and quality score*

	Speed	Speed variation	CVS	Density	Volume
Correlation between estimates and quality score	0.282	-0.347	-0.474	0.035	0

Outlier bias occurs when the summary estimate is greatly influenced by specific data points (Elvik, 2005). In order to determine the presence of outliers, each effect estimate was omitted from computation of the summary effect and a new summary effect (from *N-1* estimates) was calculated. If the new summary effect did not remain within the 95% confidence interval of the main summary (from *N* estimates), the omitted effect estimate was deemed an outlier. This test was applied to all variable estimates. Although some data points with large values were suspected of being outliers (e.g., the estimate of CVS reported in Hourdos *et al.*, 2006), they passed this test due to the large standard errors reported; therefore, no outliers were identified.

Studies included in the meta-analysis had different strategies for selecting the time interval for measuring traffic flow variables; however, most of these were arbitrarily selected (this issue is discussed further in the following section). The impact of the time interval on the

magnitude of estimates and a study's quality score were individually examined. For instance, both the quality score and estimated effect decreased as larger time intervals were selected for measuring average speed (see Figure 5). Relationships for other variables are summarized in Table 9, which shows that the time interval has a significant impact on the quality of a study where speed, speed variation, or CVS was used, and on estimates of speed variation, CVS, density, and volume (note that this test was not performed for speed difference and density variation due to the limited sample sizes).



(a) (b)

Figure 5. (a) Relationship between average speed estimates and time intervals; (b) Relationship between relative study qualities and time intervals

Table 9 *Relationship between time interval, quality score, and estimate value*

	Speed	Speed variation	CVS	Density	Volume
Relationship between time interval and quality score	-0.335	-0.612	-0.471	-0.02	0
Relationship between time interval and estimate value	-0.109	0.822	0.998	0.992	0.505

Finally, several studied have differentiated variables according to locations. To shed light on the potential confounding effect of locations on relationship between traffic characteristics and freeway crash occurrences, three additional meta analyses have been conducted: one for variables measured at upstream of crash locations, one for variables measured at downstream of crash locations, and one for variables in studies that did not distinguish locations. Results are summarized in Table 10. This table clearly shows location's confounding effect on relationship between traffic characteristics and freeway crash occurrences:

- Average speed: its impact on crash occurrences at upstream or in studies where locations are not distinguished is consistent with what is reported in the main meta analysis, while its impact at downstream becomes insignificant;

- CVS: its impact on crash occurrences is totally different depending on where it is measured, which causes the insignificance of this variable in the main meta analysis. More specifically, for CVS measured at downstream of crash locations, the odds ratio of a crash occurrence increases by about 191% when CVS increases by one additional unit (the large summary odds ratio is caused by the fact that one study reported a large value for CVS as discussed previously, i.e., 49.1 in Hourdos et al. (2006)); for CVS measured at upstream of crash locations, the odds ratio of a crash occurrence decreases by 18.3% when CVS increases by one additional unit.
- Volume: for volume in studies where locations are not distinguished, its impact on crash occurrences is consistent with what is reported in the main meta analysis, while for volume at upstream, opposite effect is detected, that is, the odds ratio of a crash occurrence decreases by 53% when volume at upstream increases by one additional unit.
- Density: for density that is measured at upstream or in studies where locations are not distinguished, its impact on crash occurrences is not significant. However, for density that is measured at downstream, its impact on crash occurrences is significant. More specifically, the odds ratio of a crash occurrence increases by 2.1% when density at downstream increases by one additional unit.

Note that speed variation, speed difference, and density variation are not listed in this table because of no location variation for these variables across the cited studies.

Table 10 Location's confounding effect on relationship between traffic characteristics and freeway crash occurrences

Average speed			
	Number of estimates	Summary odds ratio	Confidence interval
Downstream	5	0.933	[0.818,1.064]
Upstream	3	0.993	[0.988,0.998]
Not-distinguished	2	0.915	[0.907,0.923]
CVS			
	Number of estimates	Summary odds ratio	Confidence interval
Downstream	4	2.914	[1.937,4.385]
Upstream	3	0.817	[0.674,0.990]
Not-distinguished	0	-	-
Volume			
	Number of estimates	Summary odds ratio	Confidence interval
Downstream	0	-	-
Upstream	2	0.47	[0.370,0.597]
Not-distinguished	4	1.001	[1.0005,1.001]
Density			
	Number of estimates	Summary odds ratio	Confidence interval
Downstream	5	1.021	[1.014,1.027]
Upstream	4	0.956	[0.365,2.501]
Not-distinguished	2	1.0005	[0.380,2.628]

4. Discussion

Recently, many researchers have made significant attempts to investigate the connection between real-time traffic characteristics and freeway crashes. Owing to notable advances in data collection technologies, high-resolution traffic operations data are now widely accessible, and this availability has motivated many researchers to advance the common understanding in this area. Despite remarkable progress in the field, however, challenging issues remain that render models from being widely applicable to real-world scenarios. Therefore, there is value in reviewing the current state of knowledge and to identify where fruitful future research might be directed.

For the convenience of the following discussion, the studies included in the systematic review are referred to as 'selected studies'. The section outlines the shortcomings of and the common issues shared among the selected studies including general issues, study design, data constraints, model development concerns, and model validation.

4.1. General discussion

Identifying the causes of motor vehicle crashes is a complex phenomenon that involves the known conceptual interactions of behavioural, vehicle, and roadway factors (see Figure 6). A fundamental issue relating to the selected studies is that they have been aimed at measuring the relationships between traffic conditions and crash occurrence, with vehicle and most importantly behavioural factors omitted. This omission makes the critical assumption that crashes are related to the spatially and temporally proximal traffic data both in average and dynamic traffic conditions. In all likelihood, however, a high proportion of crashes are caused primarily by other factors, the most glaring of which is human error, which in the US in 2013 was estimated to account to around 93% of all crashes (NHTSA, 2013). Essentially, this is a model specification issue that gives rise to some undesirable model qualities. First, when important factors are ignored, their contribution to crash occurrence is incorrectly attributed to traffic flow characteristics that are correlated with the omitted factors. When omitted factors are not correlated with included variables, their effects will be revealed through increased model uncertainty. In the first case the estimates of effects attributed to traffic flow variables are biased, in the latter case the precision of effects are reduced. Second, as shown in Figure 6, interactions between variables are quite important. In particular, about 30% of crashes are caused by interactions between roadway and human factors, with 3% of these involving vehicle factors as well. When these interactions are omitted, the ability to identify critical pre-cursors and to discriminate between 'critical' and 'non-critical' pre-cursors are nearly impossible—leading to extremely large type I error rates or false positives. According to Lum and Reagan (1995), only 3% of all crashes are the result of roadway factors alone. This finding suggests that models intended to predict all crashes using only traffic factors will not have sufficient information to discriminate the pre-cursors for approximately 97% of cases.

Partly for these reasons, the performance of the existing crash prediction models are inaccurate, imprecise, and reveal inconsistent results as reported in the literature.

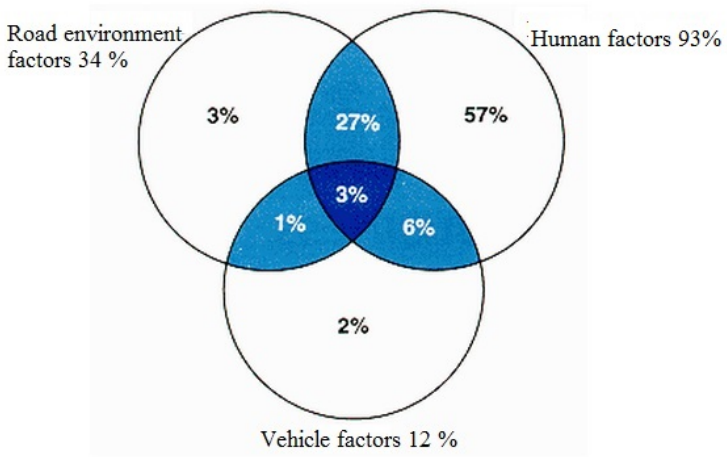


Figure 6. Factors contributing to crash occurrence in the United States (Lum and Reagan, 1995)

Perhaps more importantly, little has been offered to articulate a fundamental theory relating crashes to traffic operations, and results have been largely empirically driven. While the empirical research conducted to date is vitally important to both substantiate and refine a collective theory, a fundamental theory is lacking. Any theory that should evolve moving forward might be expected to capture the non-linearities known to exist between speed, flow, and density, as refined over the past three decades of transport research.

A fundamental articulated theory should also include explicit recognition of relationships between crashes and freeway segment types known to perform different functions, as the majority of selected studies to date have ignored the impact of different types of freeway segments and their associated traffic operations. As an example, because traffic characteristics on weaving segments are relatively more dynamic compared to other segments (e.g., in basic freeway, merge, and diverge segments), it is essential to consider how different free segments are expected to perform based on well understood operational theory (HCM, 2010). Although the resulting impact on crash occurrence of these segments has been postulated, no serious efforts to date have been made to articulate and test a theory.

A third issue arising from this systematic review is that although publications on this topic are rapidly increasing, a number of these papers are based on the same dataset. Studies by researchers using more diverse datasets could lead to a more robust collective inquiry.

4.2. Study design

Most of the selected studies applied the case-control design to investigate the significance of potential variables, while controlling external confounders such as weather conditions and road geometry (Abdel-Aty et al., 2004, 2005, 2005a, 2007; Pande *et al.*, 2005; Zheng *et al.,* 2010). Because of its simplicity, cost effectiveness and theoretical soundness, case-control design is an efficient method of studying the relative risks of rare events, and is widely used in epidemiology (Manski, 1995; Schlesselman and Stolley, 1982). It is natural to use the

case-control design at the individual vehicle level by treating a vehicle involved in a crash as 'a case', and other vehicles at the crash scene or in similar situations (but not involved in a crash) as 'controls'. Kloeden *et al*'s (1997) study is the first notable case-control study in the traffic literature. Zheng *et al*. (2010) later reviewed traffic safety studies using the case-control design.

While case-control design was predominant in the study of the link between traffic characteristics and crash occurrences, defining 'cases' and 'controls' is not straightforward: a 'case' should represent traffic conditions prior to a crash, and a 'control' should represent non-crash traffic conditions. Some researchers define the controls as the equivalent location, time, and weekday of other weeks of a crash throughout a dataset (Abdel-Aty and Pande, 2005; Abdel-Aty *et al*., 2008; Abdel-Aty *et al*., 2007; Hossain, 2011; Hossain and Muromachi, 2009; Lee *et al*., 2003; Pande and Abdel-Aty, 2005a). Pham *et al.* (2011) chose controls of a crash from the corresponding traffic regime, regardless of time and location of the traffic situations. As described previously, as only about 3% of all crashes are a function only of traffic conditions, choosing a 'control' is likely to produce a set of conditions that is very similar to crash-prone traffic conditions, since both would lack behavioural and driver factors.

While it is evident that the way in which controls are selected has a significant impact on modelling results, no study has comprehensively investigated this important issue. An appropriate methodology for selecting non-crash situations can lead to a better understanding of crash mechanisms, and further improve the predictive performance of models. Therefore, more research is needed to comprehensively investigate the effects of different approaches to the selection of non-crash situations on model performance.

More importantly, although case-control design is predominant in the literature, the validity of its use in the study of this topic needs to be scrutinized. In traditional case-control studies, the control sample is often unknown, or it is too expensive to recruit all legitimate controls. Thus, this type of study often uses a fixed case-to-control ratio such as 1:5, while a control-to-case ratio of around 4:1 is recommended since the statistical power generally does not increase significantly beyond that (Ahrens and Pigeot, 2005; Hennekens and Buring, 1987; Rothman and Greenland, 1998; Schlesselman and Stolley, 1982). However, in investigating the impact of real-time traffic characteristics on crash occurrences, once the criteria for selecting controls are determined, the total number of legitimate controls is known to researchers. Of course, this number can be large; for example, around 1000 candidate controls were available for each case in Zheng *et al.* (2010). However, the bottom-line question is, Why not use all the controls? Are there any undesirable consequences of doing so? Zheng *et al.* (2010) indirectly investigated this issue by experimenting with different ratios, and by re-sampling controls 20 times for each case to check consistency. In our view, there is clearly a need to rigorously investigate this important issue because of the predominance of the case-control design in the literature, despite the existence of ample data to use all possible "controls".

4.3. Data preparation

4.3.1. Traffic data

In recent decades, the availability of high-resolution vehicular data collected by loop detectors and video surveillance facilities has motivated researchers to specifically examine the connection between pre-crash traffic characteristics and crash occurrence. The primary features of these two data types are summarized below.

Loop detector data were predominantly used in the selected studies (85%) as they were widely available and accessible. However, there are several issues related to the general processing of loop detector data in the selected studies. The first is that the raw data (which is usually collected every 20 or 30 seconds) were often aggregated to longer periods (such as 5 or 10 minutes) to suppress noise. As pointed out by Davis (2002), this aggregation can lead to ecological fallacy because such data cannot reflect the trajectory of an individual vehicle (Zheng, 2012). It also emphasises the lack of a theory to guide the selection of time scale appropriate for capturing temporally and spatially appropriate levels of data aggregation. Different time intervals can have a significant impact on study quality and modelling results, as discussed previously in the sensitivity analysis. Based on the assumption that the traffic conditions prior to a crash are a direct contributor to that crash, traffic characteristics during a certain time interval immediately before the crash are often measured and linked to crash likelihood. Most of the selected studies split a certain period prior to a crash into equal time segments, and examine the linkage between the crash and traffic flow characteristics within these time slices, and traffic flow variables are measured for each time slice. The length of each time slice (e.g., 5, 6, 10 or 15 minutes) in most of these studies was, by and large, arbitrarily selected, without the guidance of a well-articulated hypothesis or theoretical justification.

There are, however, several notable exceptions where the selections were at least empirically motivated. Pande *et al*. (2005) and Abdel-Aty *et al*. (2005), for example, considered two different time intervals (3 and 5 minutes), and found that the 5 minute interval was an empirically superior choice. Lee *et al*. (2003) proposed an objective method to compute a proper time interval, assuming that the value chosen for the interval maximizes the difference between two estimates of variables in crash and non-crash cases. Eventually, three different time intervals for measurement (2, 3 and 8 minutes) were selected for each crash precursor. Zheng *et al*. (2010) selected 10 minutes – a typical period of traffic oscillation, and the focus of their study. As Zheng (2012) pointed out: if, indeed, there is a precursor traffic condition prior to a crash, the time period corresponding to that precursor traffic condition may be different for each individual crash. Data mining/pattern recognition techniques can be utilised to detect a unique precursor period for each crash.

Another issue in using loop detector data is limited and discontinuous data coverage, both spatially and temporally. Generally, an individual vehicle's trajectory cannot be reconstructed

from such data. Traffic characteristics derived from such data are less informative than those from the trajectories of an individual vehicle. An even more limiting factor is that researchers are often forced to use data collected from loop detectors far from crash locations and must estimate traffic characteristics prior to a crash.

A seemingly sensible approach for overcoming these shortcomings in loop detector data is to extract individual vehicle trajectories from video cameras. Two of the selected studies applied this approach (Hourdos *et al.,* 2006; Hourdos *et al.,* 2008). Video cameras are running continuously at black spots over a long time period, e.g., about one year in Hourdos *et al.* (2008). Any crashes occurred at these locations are captured in the video, e.g., 110 crashes were captured in Hourdos *et al.* (2008).Traffic characteristics (e.g., vehicle speed, and headway) and environmental conditions (e.g., weather) can also be obtained through the video footages. This enabled the research teams to gain a clearer understanding of vehicular interactions prior to a crash, and to obtain a more accurate and more reliable representation of traffic dynamics. Meanwhile, crash information in the police report, such as location and time, can be crosschecked by watching the video. More importantly, unreported crashes can be captured and retrieved (Hourdos *et al*., 2006; Hourdos *et al.,* 2008). However, significant disadvantages of extracting vehicular trajectories from video cameras include costly and intensive labour involved with data collection, difficulty in capturing a sufficient crash sample, and data noise that can affect a model's reliability. For example, the estimate of CVS in Hourdos *et al*. (2006) is 49.1, an absolute value more than 15 times larger than what are reported in other studies. This difference does highlight that different ways of measuring traffic can yield vastly different predictors.

4.3.2. Temporal Precision Issues

In order to extract traffic characteristics operating immediately before crash occurrence from the available traffic data, it is critical to obtain the exact time and location of the crash. Most of the selected studies relied on police reports to extract such information. However, for various reasons, many crashes are not recorded in police reports, and the inaccuracy of these reports is widely acknowledged (Hu et al., 1994; Oh *et al*., 2001, Abdel-Aty *et al*., 2004). For instance, the time of a crash is often reported within an hourly window, and this period is too imprecise to link with traffic characteristics. Alternatively, crash occurrence times are sometimes rounded to the nearest 5 minute time period (Golob and Recker, 2004; Kockelman and Ma, 2007). Obviously, the use of such information in police reports will be difficult to temporally link with traffic characteristics that 'caused' a crash versus crash characteristics caused by the crash, leading to the potential for "cause and effect" ambiguity.

Some researchers attempted to correct the time information in police reports prior to model development. For example, Abdel-Aty *et al.* (2005) and Zheng *et al.* (2010) used traffic flow data as a complementary source to check the accuracy of police reported crash occurrence times by detecting abrupt and dramatic changes in traffic conditions at the upstream and downstream detectors. Of course, using vehicle trajectory data (if available) is another way of addressing this issue, and one can identify the exact time and location of each crash by

reviewing video footage (Hourdos *et al.,* 2006).

4.4. Model development

The selected studies identified a diverse range of potential variables (predictors) to capture traffic dynamics prior to crash occurrences; for example, averages of speed, density, volume, speed variance, and CVS at different loop detector stations within the study area. Therefore, the methods used for predictor selection played an important role in the model development of the selected studies. Roughly, two approaches were used in these studies: statistical models and data mining techniques.

While most studies used statistical approaches, some applied data mining techniques, such as classification trees; Kohonen clustering algorithm; multi-layer perceptron (MLP); normalized radial basis function (NBFF); and Bayesian belief net (BBN) (Gholob and Recker, 2003, Pande and Abdel-Aty, 2006a, Hossain and Muromachi, 2011). Compared to traditional statistical methods, data mining techniques can generally easily handle correlated explanatory variables and high order interactions (this is the case for most of the potential variables relevant to traffic characteristics; they are often correlated) (Pande and Abdel-Aty, 2006; Christoforu *et al*., 2011; Hossain and Muromachi, 2012). However, using data mining techniques usually requires large amounts of information as input, and operates in a non-inferential way, making their results extremely difficult to interpret or to assist in refining an underlying theory of crashes caused by traffic—a deficiency described previously.

4.5. Model validation

Model validation is an important step in the development of all models. Unfortunately, it is frequently ignored or only partially discussed in the literature related to crash prediction modelling. Measures of model validation typically include prediction accuracy (i.e., the percentage of correct predictions of crashes), false positive rates (i.e., a non-crash traffic condition is identified incorrectly as a pre-crash situation), false negative rates (i.e., while a crash has actually occurred, the model has estimated it as a non-crash), and overall percent correctly predicted.

Only a few studies have provided model performance metrics, to their credit. Abdel-Aty et al. (2004) reported that their model predicted 69% of crashes correctly, with a false negative rate of 38.8% and a false positive rate of 5.39%. Hourdos et al. (2006) tested their model and indicated a prediction accuracy of 80% and false positive rate of 15%. In another study, the prediction and false positive rates from a Bayesian belief model were 66% and 20%, respectively (Hossain and Muromachi, 2012).

Overall, the prediction accuracy and the false positive rate are frequently used, while few studies have used all three measures to comprehensively validate their model's performance. However, such a comprehensive validation is critical before any prediction model is implemented. Meanwhile, although the ideal rates for prediction accuracy, false positive and false negative are 100%, 0%, and 0%, in practice improving one measure may compromise another. Thus, balancing prediction accuracy and false positive/negative is an important issue

that is not yet addressed in the literature. Moreover, many studies did not report any measure of their model's performance, thus making it difficult to assess how 'practice-ready' these models are or to compare with existing models.

In addition, numerous selected studies were based on a case-control design, and the case control ratio varied from 1:1 to 1:5. The use of these different case-control ratios can further complicate model comparison because it can affect false positive/negative errors, even when other conditions are the same. In other words, using a 1:1 case control design will bias false positive rates to be really low, because in practice there are far more 'real' controls then the one used in the study. The extent of this bias has not been reported in the literature, and is an important omission.

Finally, to facilitate an objective comparison of different models, a publicly accessible and well-structured dataset for benchmarking crash prediction models is highly desired. This benchmarking is a common practice in algorithm design for image processing, signal processing, and so on (Wang *et al.,* 2004; Hawang and Arakawa, 1996). This type of dataset would allow modellers to compare different model specifications objectively.

5. Conclusions

This paper presents a systematic review of the existing literature on the relationship between real-time traffic characteristics and crash occurrence on freeways. It then describes a meta-analysis undertaken to collate the findings from selected studies, and reports the summary effects of traffic characteristic impacts on crash occurrence determined by this analysis. Specifically, the paper reports that:

i) The summary effect size of speed variation is 1.226, which indicates that if speed variation increases by one additional unit, the odds ratio of a crash occurrence increases by 22.6%.
ii) The summary odds ratio of speed difference is 1.032, which indicates that if speed difference increases by one additional unit, the odds ratio of a crash occurrence increases by 3.2%.
iii) The summary odds ratio of average volume (when locations are not distinguished) is 1.001, which indicates that if average volume increases by one additional unit, the odds ratio of a crash occurrence increases by 0.1%.
iv) A larger value of average speed is associated with a lower risk of crash. More specifically, if average speed increases by one additional unit, the odds ratio of a crash occurrence decreases by about 4.8%.

Although existence of crash-prone conditions is generally acknowledged (e.g., Oh *et al.*, 2001; Abdel-Aty and Pande, 2006; Hourdos *et al.*, 2008), it is still unclear how 'risky' these pre-cursor conditions actually are. Our meta-analysis results indicate that speed variation and CVS can highly affect the likelihood of crash while average speed, average density, and speed difference have moderate impact. In contrast, traffic volume's effect on crash occurrence is considerably small (when locations are not distinguished).

Sensitivity analyses were conducted in our meta-analysis to investigate the effect of study quality, publication bias, outlier bias, and the time interval used to measure traffic characteristics. Overall, no notable trend between the variable estimate of a traffic characteristics and the quality of a study was observed. Publication bias was detected in CVS and average density, and no outliers were identified. However, the time interval selected for a study has a significant impact on the quality of a study where speed, speed variation, or CVS was used, and on estimates of speed variation, CVS, density, and volume.

Furthermore, the sensitivity analyses clearly show location's confounding effect on relationship between traffic characteristics and freeway crash occurrences. Notably, CVS' impact on crash occurrences is totally different depending on where it is measured. Similar conclusion is obtained for density. For density that is measured at upstream or in studies where locations are not distinguished, its impact on crash occurrences is not significant. However, for density that is measured at downstream, the odds ratio of a crash occurrence increases by 2.1% when density at downstream increases by one additional unit. In addition, for volume in studies where locations are not distinguished, its impact on crash occurrences is consistent with what is reported in the main meta-analysis (i.e., the odds ratio of a crash occurrence increases as average volume increases), while for volume at upstream opposite effect is detected.

Based on the comprehensive and systematic literature review and results from the meta-analysis, substantive issues in study design, traffic and crash data, model development, and model validation are discussed. Outcomes of this study are intended to both summarise the existing state of the knowledge in this fast growing research field, and to guide future research in the area of real-time crash prediction.

Acknowledgements: The authors were grateful for two anonymous reviewers' insightful and constructive comments. This research was partially funded by the Early Career Academic Development program at Queensland University of Technology.

References

Abdel-Aty, M., Uddin, N., Pande, A., Abdalla, F. M., Hsia, L., 2004. Predicting freeway crashes from loop detector data by matched case-control logistic regression. Transportation Research Record: Journal of the Transportation Research Board, 1897(1), 88-95.

Abdel-Aty, M., Pande, A., 2005. Identifying crash propensity using specific traffic speed conditions. Journal of safety Research, 36(1), 97-108.

Abdel-Aty, M., Uddin, N., Pande, A., 2005a. Split models for predicting multivehicle crashes during high-speed and low-speed operating conditions on freeways. Transportation Research Record: Journal of the Transportation Research Board, 1908(1), 51-58.

Abdel-Aty, M., Pande, A., 2006. ATMS implementation system for identifying traffic conditions leading to potential crashes. Intelligent Transportation Systems, IEEE Transactions on, 7(1), 78-91.

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

Abdel-Aty, M., Pande, A., Lee, C., Gayah, V., Santos, C. D., 2007. Crash risk assessment using intelligent transportation systems data and real-time intervention strategies to improve safety on freeways. Journal of Intelligent Transportation Systems, 11(3), 107-120.

Ahrens, W., Pigeot, I., 2005. Handbook of epidemiology. Springer.

Borenstein, M., Hedges, L. V., Higgins, J.P., Rothstein, H. R., 2011. Introduction to meta-analysis. John Wiley and Sons.

Cheng, W., Washington, S., 2005, Experimental evaluation of hotspot identification methods, Accident Analysis and Prevention 37 (5), 870 – 881.

Cheng, W., Washington, S., 2008, New criteria for evaluating methods of identifying hot spots, Transportation Research Record, 2083: 76-85.

Christoforou, Z., Cohen, S., Karlaftis, M. G., 2011. Identifying crash type propensity using Real-time traffic data on freeways. Journal of safety research,42(1), 43-50.

Cooper, H., Hedges, L. V., Valentine, J. C., 2009. The handbook of research synthesis and meta-analysis. Russell Sage Foundation.

Davis, G.A., 2002. Is the claim that 'Variance Kills' an ecological fallacy? Accident Analysis and Prevention 34 (3), 343–346.

Duval, S., Tweedie, R., 2000. A non-parametric trim and fill method of assessing publication bias in meta-analysis. Biometrics 56, 455–463.

Elvik, R. 2005. Can we trust the results of meta-analyses? A systematic approach to sensitivity analysis in meta-analyses. Transportation Research Record: Journal of the Transportation Research Board, 1908(1), 221-229.

Elvik, R., 2008. Making sense of road safety evaluation studies. Developing a quality scoring system. Report, 984.

Elvik, R., 2013. Risk of road accident associated with the use of drugs: A systematic review and meta-analysis of evidence from epidemiological studies. Accident Analysis and Prevention, 60, 254-267.

Garber, N. J., Wu, L., 2001. Stochastic models relating crash probabilities with geometric and corresponding traffic characteristics data. Publication UVACTS-5-15-74. Charlottesville, VA: Center for Transportation Studies, University of Virginia.

Golob, T. F., 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., Recker, W. W., Alvarez, V. M., 2004. Safety aspects of freeway weaving sections. Transportation Research Part A: Policy and Practice, 38(1), 35-51.

Hwang, K., Xu, Z., Arakawa, M., 1996. Benchmark evaluation of the IBM SP2 for parallel signal processing. Parallel and Distributed Systems, IEEE Transactions on, 7(5), 522-536.

Hennekens, C.H., Buring, J.E., 1987. Epidemiology in medicine. Lippincott Williams and Wilkins.

Hossain, M., Muromachi, Y., 2009. A framework for real-time crash prediction: Statistical

approach versus artificial intelligence. Journal of Infrastructure Planning Review (JSCE), 26 (5), 979-988.

Hossain, M., Muromachi, Y., 2011. Understanding crash mechanisms and selecting interventions to mitigate Real-time hazards on urban expressways. Transportation Research Record: Journal of the Transportation Research Board, 2213(1), 53-62.

Hossain, M., Muromachi, Y., 2012. A Bayesian network based framework for Real-time crash prediction on the basic freeway segments of urban expressways. Accident Analysis and Prevention, 45(0), 373-381.

Hourdos, J. N., Garg, V., Michalopoulos, P. G., Davis, G. A., 2006. Real-time detection of crash-prone conditions at freeway high-crash locations. Transportation Research Record: Journal of the Transportation Research Board, 1968(1), 83-91.

Hourdos, J. N., Garg, V., Michalopoulos, 2008. Accident Prevention Based on Automatic Detection of Accident Prone Traffic Conditions: Phase I. Intelligent Transportation Systems Institute Centre for Transportation Studies University of Minnesota.

Høye, A., Elvik, R., 2010. Publication bias in road safety evaluation. Transportation Research Record: Journal of the Transportation Research Board, 2147(1), 1-8

Hu, P., Young, J. 1994 . 1990 NPTS Databook: Volume II, Chapter 10". NPTS Databook, U.S. DOT, Office of highway Information Management.

Kloeden, C., McLean, A., Moore, V., Ponte, G., 1997. Travelling speed and the risk of crash involvement. In: NHMRC Road Accident Research Unit. University of Adelaide, Adelaide, Australia.

Kockelman, K. K., Ma, J., 2007. Freeway speeds and speed variations preceding crashes, within and across lanes. In Journal of the Transportation Research Forum (Vol. 46).

Kuang, Y., Qu, X., Wang, S., 2014. Propagation and dissipation of crash risk on saturated freeways. Transportmetrica B 2(3), 203-214.

Lee, C., Saccomanno, F., Hellinga, B., 2002. Analysis of crash precursors on instrumented on freeways. Transportation Research Record: Journal of the Transportation Research Board, 1784(1), 1-8.

Lee, C., Hellinga, B., Saccomanno, F., 2003. Proactive freeway crash prevention using Real-time traffic control. Canadian Journal of Civil Engineering, 30(6), 1034-1041.

Lee, C., Abdel-Aty, M., Hsia, L., 2006. Potential real-time indicators of sideswipe crashes on freeways. Transportation Research Record: Journal of the Transportation Research Board, 1953(1), 41-49.

Lee, C., Abdel-Aty, M., 2008. Two-level nested logit model to identify traffic flow parameters affecting crash occurrence on freeway ramps. Transportation Research Record: Journal of the Transportation Research Board, 2083(1), 145-152.

Light, R., Pillemer, D., 1984. Summing up: The science of reviewing research. Harvard University Press. Cambridge, MA.

Liu, P.C., 1997. Freeway incident likelihood prediction models: Development and application to traffic management. Doctoral dissertation, Purdue University.

Lum, H., Reagan, J.A., 1995. Interactive highway safety design model: Accident predictive module. Public roads, 58(3), 14.

Manski, C. F., 1995. Identification problems in the social sciences. Harvard University Press.

Meng, Q., Qu, X., 2012. Estimation of rear-end vehicle crash frequencies in urban road tunnels. Accident Analysis and Prevention 48, 254-263.

Moore, V. M., Dolinis, J., Woodward, A. J., 1995. Vehicle speed and risk of a severe crash. Epidemiology, 258-262.

NHTSA, 2013. Strategy for vehicle safety strategic planning for domestic and global integration of vehicle safety national highway traffic safety administration

Oh, C., Oh, J. S., Ritchie, S. G., Chang, M., 2001. Real-time estimation of freeway accident likelihood. In 80th Annual Meeting of the Transportation Research Board, Washington, DC.

Pande, A., 2005. Estimation of hybrid models for Real-time crash risk assessment on freeways. Doctoral dissertation, University of Central Florida Orlando, Florida.

Pande, A., Abdel-Aty, M., Hsia, L., 2005. Spatiotemporal variation of risk preceding crashes on freeways. Transportation Research Record: Journal of the Transportation Research Board, 1908(1), 26-36.

Pande, A., Abdel-Aty, M. 2006. Comprehensive analysis of the relationship between real-time traffic surveillance data and rear-end crashes on freeways. Transportation Research Record: Journal of the Transportation Research Board, 1953(1), 31-40.

Park, J.,Oh, C., 2009. Relating freeway traffic accidents to inductive loop detector data using logistic regression. 4th IRTAD conference 2009.

Pei, X., Wong, S., Sze, N., 2012. The roles of exposure and speed in road safety analysis. Accident Analysis & Prevention, 48, 464-471.

Pham, M. H., Dumont, A. G., Chung, E., 2011. Motorway traffic risks identification model - myTRIM methodology and application. Doctorate, École polytechnique fédérale de Lausanne, Lausanne (4998).

Preston, H. 1996. Potential safety benefits of intelligent transportation system (ITS) technologies. Semisequicentennial Transportation Conference Proceedings, Center for Transportation Research and Education, Iowa State University.

Rothman, K.J., Greenland, S., 1998. Modern epidemiology. Lippincott- Raven, Philadelphia.

Schlesselman, J.J., Stolley, P.D., 1982. Case-control studies: design, conduct, analysis. Oxford University Press.

Wang, Z., Bovik, A. C., Sheikh, H. R., Simoncelli, E. P., 2004. Image quality assessment: from error visibility to structural similarity. Image Processing, IEEE Transactions on, 13(4), 600-612.

Washington, S., Karlaftis M., Mannering, F., 2010, Statistical and econometric methods for transportation data analysis (2nd Edition), CRC Press.

Zheng, Z., Ahn, S., Monsere, C. M., 2010. Impact of traffic oscillations on freeway crash occurrences. Accident Analysis and Prevention, 42(2), 626-636.

Zheng, Z., 2012. Empirical analysis on relationship between traffic conditions and crash occurrences. Procedia-Social and Behavioural Sciences, 43, 302-312.