Accident Analysis and Prevention 79 (2015) 198–211

Contents lists available at ScienceDirect

Accident Analysis and Prevention

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

Impact of real-time traffic characteristics on freeway crash occurrence: Systematic review and meta-analysis

Saman Roshandel, Zuduo Zheng *, Simon Washington [1]

Civil Engineering & Built Environment School, Queensland University of Technology, 2 George St., GPO Box 2434, Brisbane, QLD 4001, Australia

ARTICLE INFO

Article history:
Received 8 August 2014
Received in revised form 30 January 2015
Accepted 10 March 2015
Available online 2 April 2015

Keywords:
Crash prediction
Traffic characteristics
Road safety
Systematic review
Meta-analysis

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 modeled 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 modeling of this important but extremely challenging safety issue.

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 behavioral 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 behavioral 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 behavioral 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.

* Corresponding author. Tel.: +61 07 3138 9989; fax: +61 07 3138 1170.
E-mail addresses: saman.roshandel@student.qut.edu.au (S. Roshandel), zuduo.zheng@qut.edu.au (Z. Zheng), simon.washington@qut.edu.au (S. Washington).
[1] Tel.: +61 07 3138 9990; fax: +61 073 138 1170.

http://dx.doi.org/10.1016/j.aap.2015.03.013
0001-4575/

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 modeling 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, 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 modeling 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 traveling 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[1] 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 address 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 modeling 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.

2. The systematic literature review

In this section a review of relevant papers is provided to catalog 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

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