Physica A 509 (2018) 937–944

Contents lists available at ScienceDirect

Physica A

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

Application of a random effects negative binomial model to examine crash frequency for freeways in China

Qinzhong Hou [a], Xianghai Meng [a], Junqiang Leng [b,*], Lu Yu [b,*]

[a] School of Transportation Science and Engineering, Harbin Institute of Technology, Harbin, 150090, China
[b] School of Automobile Engineering, Harbin Institute of Technology at Weihai, Weihai, 264209, China

HIGHLIGHTS

- Factors affecting freeway safety were investigated.
- Advantages of RENB model were further confirmed.
- Climbing lane was demonstrated to be safety beneficial.
- Increasing median barrier offset improved safety.

ARTICLE INFO

Article history:
Received 27 January 2018
Received in revised form 8 May 2018
Available online xxxx

Keywords:
Traffic safety
Geometric design
Random effects model
Freeway

ABSTRACT

The freeway system of China is the longest in the world and planned to expand in the following years. However, the safety conditions of freeways are drawing increasing concerns both from the authorities and the public in the country. The safety design and management for freeways are urgently needed considering the current safety situation and great demands of new freeways. The study presented in this paper thoroughly investigated factors affecting safety using detailed data of crashes, traffic characteristics and freeway geometry. A random effects negative binomial (RENB) model was applied to account for spatial variations within groups together with a negative binomial (NB) model. The results indicated a better goodness of fit of RENB model than the NB model. In addition, a good number of factors significantly contributed to crash, such as truck proportion, presence of climbing lane, median barrier offset, curvature and longitudinal grade, were identified. This study was expected to provide a better understanding of how traffic condition and freeway design affect safety and should be useful to freeway engineers to design safe freeways, or develop effective safety countermeasures.

1. Introduction

With the fast economic growth in China, its freeway system became the longest in the word in 2011 after surpassing by length the U.S. interstate highway system. The total length of freeway network in China grew thirteen times from 11,600 km in 1999 to 131,000 km in 2016 [1], and it is planned to expand to 150,000 km by 2020 according to the 13th Five-year Plan of Modern Transportation System in China [2]. In addition, the freeway network aims to connect all of the cities with population of 200,000 and larger by the end of 2030 [3], implying considerable new construction of freeways in the next few years.

* Corresponding authors.
E-mail addresses: houqinzhong@126.com (Q. Hou), mengxianghai100@126.com (X. Meng), lengjunqiang0724@163.com (J. Leng), yuloo@yeah.net (L. Yu).

https://doi.org/10.1016/j.physa.2018.06.088
0378-4371/

Freeways played and continue to play a crucial role in transportation and economic development in China. Unfortunately, the fast development of freeway transportation in China has been accompanied with unacceptable safety issues and high crash rate (3.7 crashes/year in 2008–2012 period based on the studied freeways in this paper), which draw more and more attentions and concerns from both the road authorities and drivers. The effects of traffic and freeway design elements on crashes have been thoroughly investigated in other countries over the past few decades [4–8], such as the safety effects of truck proportion, horizontal curvature, and longitudinal grade. However, safety countermeasures developed in other countries may not be transformable to China due to differences in driving cultures and freeway characteristics. Additionally, studies in the topic of safety of geometric design are few in numbers currently for freeways in China. To the author's best knowledge, only three studies investigated the safety effects of freeway geometric design. Authors in [9] investigated the effects of gradients of descending segments on crash rate, and found that the crash rate was positively correlated with the average vertical gradient. A study by Wang et al. [10] explored the impacts of combined alignment on safety for mountainous freeways through a driving simulator. The minimum radii of curves combined with different vertical alignments were determined and recommended by the authors from the aspect of safety. Authors in [11] investigated the relationship between crash frequency and freeway design elements. They found that the longitudinal grade, the horizontal curvature, the road width, and the ratio of longitudinal grade to curve radius significantly affect safety. The small number of studies on safety design of freeways in China indicates a gap in the knowledge of this subject and more effort is urgently needed to better improve the design of freeways.

In the aspect of analytical methodology for crash frequency, negative binomial (NB) model was widely used as they can properly represent crash counts as non-negative, integer and often over-dispersed [12–14]. The NB model was proved to be an effective method to deal with over-dispersed crashes, but it also has limitations in addressing the temporal effects, i.e., time variations are not considered in NB model. Consequently, the standard error of regression coefficient can be underestimated and the model estimations may be biased. One way to overcome the shortcoming of NB model is treating the crash data in a time-series cross-sectional panel data structure and introducing individual effects to NB model, especially when crash data were collected from N (number of segments) freeway sections for a T (number of years) time period [15]. The individual effects can be defined as fixed or random resulting in fixed effects negative binomial (FENB) model and random effects negative binomial (RENB) model. However, the FENB model does not allow variations of location-specific effects, and the RENB model treats the location-specific effects as randomly distributed. RENB model was demonstrated to be more suitable for variables which have potential location-specific effects (for example, geometric design elements in this study) [11,16].

In summary, this research is motivated by the urgent need to better understand the safety effects of traffic condition and freeway design elements using a RENB model together with a traditional NB model. The potential contributions of this study can be summarized as: (1) the effects of a wide range of freeway design characteristics were investigated and quantified, including the rarely studied factors in the past such as freeway climbing lane; (2) the results from this study should provide better understanding of the safety effects of various factors that are believed to be crucial for freeway management and design, especially for freeways in China; (3) the advantages and superiorities of the RENB model were further confirmed. In light of that, this paper continuous in the following manner. First, the methodology used to analyze the safety factors is presented. Second, the data for modeling and analyzing is described, followed by modeling results and discussion of the results. Finally, conclusions from the presented study are summarized.

2. Methodology

2.1. Model structure

A wide range of crash frequency models have been developed over the past decades to gain the knowledge of factors contributing to crash. Due to the nonnegative integer nature of crash and the over-dispersion characteristics of crash data, negative binomial (NB) model was extensively used in the past studies. Over-dispersion indicates that the variance of crash exceeds the mean, and is usually caused by unobserved heterogeneity.

For a NB model, the probability of segment i having n_{it} crashes in year t is given by:

$$P(n_{it}) = \left(\frac{1/\alpha}{(1/\alpha)+\lambda_{it}}\right)^{1/\alpha} \frac{\Gamma[(1/\alpha)+n_{it}]}{\Gamma(1/\alpha)n_{it}!} \left(\frac{\lambda_{it}}{(1/\alpha)+\lambda_{it}}\right)^{n_{it}} \tag{1}$$

where $\Gamma(\cdot)$ is a Gamma function, and λ_{it} is expressed as:

$$E(n_{it}) = \lambda_{it} = \exp(\boldsymbol{\beta X}_{it} + \varepsilon_{it}) \tag{2}$$

where $\boldsymbol{X}_{it}$ is a vector of explanatory variables; β is a vector of estimable parameters; $\exp(\varepsilon_{it})$ is a Gamma distributed error term with mean 1 and variance α (also called over dispersion parameter). The NB model assumes that the mean value of crash frequency λ_{it} is independently distributed over time, thus the variance for a NB model is:

$$VAR(n_{it}) = E(n_{it})[1 + \alpha E(n_{it})] \tag{3}$$

Note that a NB model reduces to a Poisson model when α is not significantly different from 0.