Accident prediction modelling for expressways: a review

P Kumar¹, J K Jain¹ and G Singh²

¹Civil Engineering Department, Malaviya National Institute of Technology, Jaipur, India
²Civil Engineering Department, Deenbandhu Chhotu Ram University of Science and Technology, Murthal, Sonipat, Haryana, India
{parveen.rh.kumar, jkjain05, singhgyan27}@gmail.com

Abstract. Expressways are the need for fast movement of goods and human being for long distances in current scenario for every country, but at the same time these facilities have the highest severity rate among all the road categories. Various researchers around the world have tried to find out the effect of traffic volume, road geometry and environmental factors on frequency of accident on expressway by using different accident prediction modeling methods during last three decades. The purpose of this review paper is to find out the appropriate modeling method which can be used to predict accident frequency on the expressways in the developing country like India where traffic conditions, vehicular characteristics, and the driver behaviour are very different from the developing countries and to find out the research areas where the findings are inconclusive. Literature review suggests that among various models used so far, correlated random parameter model found to be the most advanced model to simultaneously account for both the heterogeneous effects of explanatory factors across the road segments and the cross correlations among the random parameter estimates. Findings related to variables like percentage of heavy vehicles, vertical gradient and lane width are not conclusive. However, safety effect of variables like speed limit, roadway lighting, pavement type and fog are less studied on expressway. The effect of allowing two wheelers on the expressways safety has not been studied yet, and the accidents related to fatigue or drowsiness also need examination.

Keywords: Accident Frequency Models, Explanatory Variables, Review, Expressway.

1. Introduction
Road traffic accidents are the important cause of death and injury in all countries, causing more than 1.35 million deaths each year and about 50 million injuries globally. About ninety percent of these deaths and injuries are being happened in developing countries [1]. Among the developing countries, India have the highest numbers of traffic deaths in the world with 1,51,113 persons killed in 2019 [2]. Road traffic deaths and injuries become unnecessary burden on the health system and as well as on the economy of the developing countries.
Road safety experts, researchers and various agencies related to road transportation all over the world are continuously trying to understand various factors affecting accident frequency and their severity outcomes so that proper countermeasures and policies can be adopted to reduce these unwanted incidences. In line with these efforts, accident prediction models (APMs) have been proved an important tool to develop statistical relationship or algorithm between accident frequency or severity and various independent variables like traffic volume, road geometry and road environmental factors etc. Literature suggests that selection of a particular type of model is based upon the requirement of the planners or engineers and also effected by various data and methodological issues like “over-dispersion, under-dispersion, time varying explanatory variables, temporal and spatial correlation, low sample mean and small sample size, injury-severity and accident-type correlation, under-reporting, omitted variables, endogenous variables, functional form and fixed parameters” [3]. Other than these, unobserved heterogeneity is also an important issue that affects the choice of APM [4]. These APMs can be divided broadly into parametric models and non-parametric models [5]. Parametric models have major advantage of quantifying the effects of different causal factors on road traffic accidents. However, at the same time these models require assumption of predefined relationship between accidents and their causal factors. On the other hand, new generation of non-parametric methods requires no such assumptions and their predictive power is much better than statistical models. A major limitation of non-parametric models is that these are not capable to quantifying the effects of contributory factors related to road traffic accidents [6].

Expressways are the controlled access highest class of roads being constructed for fast movement of people and goods from one place to another. About 1642 km length of expressway is in operation at present in India. However, Government of India planned to construct 23 more expressways of about 7800 km length in next five years. Expressways have highest accident severity rate due to their high speed limit and diversity in traffic composition as compared to other types of highway facility [7–12]. Similarly, in India also expressways have high accident severity rate as compared to national highways and state highways [8]. Some of the probable causes of this high accident severity rate on expressways in India have been reported as driver’s fatigue and over-speeding [13], lack of foot over bridges for pedestrian crossing, non-availability of vehicle lay-by and after crash care facility [14], practice of raised median instead of flushed or depressed median, non-standard guard rail installation, unsafe entry and exit ramps, non-standard crash cushions and chevron markings [15]. But none of the research tries to quantify effect of these factors on accidents.

Keeping the above in mind, this paper reviews more than 60 papers comprising different accident prediction modeling methods, and various explanatory variables used for the estimation of frequency of accidents on expressways all around the world to understand advantages and shortcomings of the adopted methods and to find out gaps in research work.

2. Literature Review

2.1. Parametric Accident Prediction Models
Accident is a discreet variable and a rare event. Hence, Poisson regression found as an appropriate method to model the accident frequency data on expressways as compared to linear models [16–20]. However, data related to accident frequency generally found to be over-dispersed, thus violating basic assumption of Poisson model (mean and variance are equal). So, Poisson model is not capable of
handling over-dispersion and under-dispersion. Its results are seriously affected when size of the sample as well as the sample mean are small [3]. To overcome these limitations and enhance the predictive power of the Poisson model many variants of Poisson model are in use.

The zero-truncated Poisson (ZTP) model [21], also known as conditional Poisson distribution or the positive Poisson distribution, is used when due to limitations of observability and omission of certain accident data, the sample has to be truncated. Hosseinlou et al.[19] developed a zero-truncated Poisson (ZTP) model for the estimation of accidents and potential violations on Iran’s expressway and compared the results with Poisson model. The results show that ZTP could fit data better than Poisson model for truncated sample.

The negative binomial/Poisson gamma (NB) model has been widely used during last three decades for expressways by various researchers. It is an extension of Poisson model [3], effectively dealing with over-dispersion issue of accident data. However, it has limitations in addressing the correlation in data due to temporal effects and location specific effects as it assumes accident frequency on a road segment as independent [7, 22–24].

Poisson lognormal model can account over-dispersion issue [25] and found more flexible than NB model [3]. However, it has limitation of complex model estimation. Caliendo et al. [17] developed negative multinomial model for expressway to address the problem of correlation among observations. Similar to Poisson and NB model, negative multinomial model also do not have the capacity of handling under-dispersion [3].

A univariate negative binomial (UVNB) model was developed by Li et al.[26] for the estimation of accident likelihood by collision type on the expressway in the state of Florida. Univariate or independent count NB models do not address correlation problem between different severity levels [26]. To address correlation problem, bivariate or multivariate approaches can be applied. However, these models are complex to estimate due to formulation of correlation matrix [3]. A bivariate negative binomial (BVNB) model was applied by Chen et al. [27] to analyze the effect of geometric elements on the safety of expressway. A multivariate negative binomial (MVNB) model was applied by Anastasopoulos [28] to estimate injury-severity frequency for expressway in Indiana.

A multivariate Poisson lognormal (MVPLN) model was estimated by Li et al. [26] to estimate the impact of independent variables simultaneously on various types of accidents, and the results showed that prediction accuracy of MVPLN model was found high as compared to the UVNB model. The MVPLN model can handle negative correlations, whereas MVNB model is not capable of handling this issue [29]. Wang et al. [30] was also recommended the use of the MVPLN in estimating weather related accidents by accident type and severity, and results showed more unbiased parameter estimates as compared to the UVPLN model.

Haq et al. [31] applied a zero inflated negative binomial (ZINB) model for the expressway in the state of Wyoming to develop APMs for truck accidents. The ZINB model is good for over-dispersed accident frequency data that have excess zeros and offers enhanced goodness-of-fit than the NB model [32]. This model assumes that there are two different state of accident generating process: an accident free state, and a normal accident prone state. However, the ZINB model does not correctly reproduce the accident-data generating process [32]. Multivariate zero inflated negative binomial (MVZINB) model can handle correlation issue among different injury-severity levels and types of accident [28, 33].

The unobserved heterogeneity arising from temporal and spatial effects, results biases in parameter
estimation in Poisson and NB model, as the accident distributions for similar road sections are considered the same in these models [34]. Many researchers [7, 35–38] have used random effect negative binomial (RENB) model to address the problem of unobserved heterogeneity while estimating accident frequency on expressways. The results showed that RENB model performed better as compared to NB model in handling unobserved heterogeneity issue. However, if the assumptions regarding variability of spatial and temporal effects are not correct, RENB model may be insufficient [37]. Some researchers [11, 18, 39] developed random effect model for expressways to address the unobserved variations at different sites. Zeng et al. [18] developed Random effect model with autoregression-1 (AR-1) correlation structure (REAR-1) to address chronological association among accidents occurred in the consecutive months and also unseen variations and temporal correlation present in the data.

Many researchers have used random parameter negative binomial (RPNB) model to address the issue of unobserved heterogeneity for the estimation of accidents on expressways [35, 37, 38, 40–43]. RPNB model is an extension of the RENB model [37], and have more flexibility in handling unobserved heterogeneity as compared to the later one. However, parameter estimation process for RPNB model is more complicated. The RPNB model cannot address the problem of probable correlation between the varied effects of explanatory factors [38]. Hou et al.[38] developed a correlated random parameter negative binomial (CRPNB) model for expressway tunnel safety in China and the results showed better goodness of fit as compared to RPNB model. CRPNB model estimates were found better accounting for cross correlations among the random-parameter estimates as well as heterogeneous effects of explanatory factors [43].

Due to the increasing availability of geo-referenced accident data and increasing computational power, Liu et al. [44] have estimated a geographically weighted negative binomial (GWNB) model to predict accident frequencies on expressway. GWNB model is capable of handling spatial heterogeneity and outperforms the traditional NB model. However, it does not address temporal heterogeneity. GWNB models are highly localized and they cannot be generalized to other locations [44].

Random parameters negative binomial-lindley (RPNB-L) model has been used by Rahman Shaon et al. [42] for handling expressway accident data with unobserved heterogeneity and high over-dispersion, and the results showed that this model outperformed RPNB model.

APMs for expressways have also been described in Highway Safety Manual (HSM) supplement [45]. These prediction models consist of a base safety performance function (SPF) with accident modification factors (AMFs) and a calibration factor to estimate the predicted average accident frequency. Using HSM approach La Torre et al. [9] analyzed single and multiple vehicle casualty and severe accidents for Italian expressways and provided a trustworthy tool for expressway agencies to deal with potential safety issues.

The generalized estimating equation (GEE) is an extension of generalized linear models to accommodate correlated dependent variable [46]. There are different approaches in this method to handle serial correlation including independent, dependent, autoregressive 1, exchangeable and unstructured correlation structure [3]. The GEE method was used by Mohammadi et al. [47] and Zheng et al. [46] to model the yearly accident counts on expressways to accommodate the temporal correlation problem.

To account for the effect of unobserved heterogeneity and spatio-temporal correlation in the data, Bayesian spatial and temporal approaches have also been proposed for the accident prediction [11].
Conditionally auto regressive (CAR) model is most common Bayesian spatial model, capable of handling correlation issue among the consecutive segment on the expressway [11]. Spatial Poisson-lognormal model is useful for addressing unobserved heterogeneity and spatial correlation among neighboring road segments of expressway [35].

2.2. Non-Parametric Accident Prediction Models
Non-parametric models do not require any prior assumption of the relationship between dependent variable and explanatory variables [20, 48]. Artificial Neural Network (ANN) model is one of the commonly used models suitable for nonlinear interactions among variables and it allows the inclusion of a large number of variables. However, ANN does not explain parameter estimates associated to the explanatory factors [20] and the results cannot be applied to the data sets with different specifications [49].

Classification and Regression Tree (CART) is a good tool for prediction and classification problems. This model can effectively handle collinearity and outliers problem which may be an important issue with regression models [50]. It was found a better substitute for NB regression models to analyze expressway accidents. However, CART model has its own limitations of not make use of continuous, ordinal variables, have problems in handling the interactions between independent variables and in doing elasticity analysis or sensitivity analysis [50].

2.3. Significant Explanatory Variables
The literature suggests that the choice of explanatory variables into the model has been based upon availability of data, significant variables established in earlier researches and the new variables which can be accurately calculated [51]. Significant explanatory variables found in previous studies and their impacts on accident frequency are discussed in the following paragraphs.

2.3.1. Exposure Variables. The main exposure variables selected are Traffic volume and segment length [27]. Almost all researchers found traffic volume as a significant independent variable affecting accident frequency. Most of the researchers [7, 16, 17, 20, 28, 30, 40–43, 46, 47, 52–55] found that annual average daily traffic (AADT) or ADT was positively correlated with accident frequency. Zeng et al. [18] used monthly total traffic (MTT) instead of AADT to avoid information loss in time-varying variable and found positive association with accident frequency. Persaud and Dzbik [53] used hourly traffic volume for microscopic model to capture daily variation in traffic volume. Many researchers [17, 18, 40, 41, 43, 46, 55] found the expressway segment length to be positively correlated with number of accidents. Wen et al. [56] used variable daily vehicle kilometer travel (DVKT) and found it positively associated with accident frequencies. Similarly, Wen et al. [11] used variable monthly vehicle kilometer travel (MVKT) and found positively correlated with the accident frequency.

2.3.2. Traffic Conditions. Traffic composition was used as an important independent variable by various researchers to identify the safety effects on expressway. Some researchers [20, 24, 31] found that higher percentage of trucks in traffic volume on expressway causes more number of accidents, and this could be attributed to the difference of speed between light vehicles and heavy vehicles. However, other researchers [11, 47, 57] found that high fraction of trucks and trailers in traffic volume result in decreasing accident risk. This contradictory finding could be associated with cautious driving of light vehicle drivers close to heavy vehicles, better driving capabilities of heavy vehicle drivers as compared
to light vehicle drivers, and increased homogeneity in traffic due to higher percentage of heavy
due to escalating lane change opportunities and consequently rising conflict between the traffic.

Saeed et al. [37] found that lane width have negative correlation with the injury accidents. Ture
Kibar et al. [14] concluded that truck involvement in accidents reduces with increased lane width.

However, Mohammadi et al. [47] found positive correlation among the lane width and accident
occurrence. Thus, this contradictory finding requires further investigation. Undersized and oversized
lane factors were found statistically significant and positively related to accident frequency for urban
expressway[55]. Lane width less than 3.25 m and more than 3.75 m were considered as undersized and
oversized segments respectively [55].

Some researchers found a decreasing impact of adding the climbing lane on the accident frequency [24, 31, 37] along section with steep grades. This could be attributed to the reduction in interaction among light vehicles and heavy vehicles after adding climbing lane.

Hou et al. [37] found that increase in the median barrier offset can significantly lower the accidents as the larger median barrier offset reduces driver’s tension and anxiety of hitting the barrier. It was found that the segments with median barriers witness decreased number of accidents [40, 42].

Median width was found significant in affecting safety for expressways [16]. A negative correlation was found between median width and number of accidents on expressway [27, 42, 54].

Both inside and outside shoulder widening were found associated with decreased accident count [30, 44, 46, 51, 52].

2.3.5. **Category of Segment.** Higher accident frequency was observed on the segments in service areas, tunnel and within interchanges as compared to the basic segments under similar conditions, probably caused by frequent braking, accelerating, and lane changing maneuvers on these segments [43].

2.3.6. **Pavement Related Variables.** Pavement condition factors significantly associated with accident frequency on expressway include distress ratio, rutting depth, and International Roughness Index (IRI) [37]. Friction and pavement type were also significantly affected accident frequency. Saeed et al. [43] found that the segments with excellent pavement condition (IRI ≤ 100) had a higher number of injury accidents due to associate higher speeds. Chen et al. [27] found that 1 percent increase in pavement serviceability index would decrease no-casualty accidents on expressway by 0.26 percent. Increasing IRI was found associated positively with the accident frequency [40]. Increased friction had a positive impact on the accident frequency [42]. Change in pavement type (from concrete to asphalt) almost always increases the probability of accidents [42]. No explanation of this phenomenon is found in the literature and hence need to be investigated further.

2.3.7. **Some Other Variables.** Traffic Sign: Installation of more traffic signs up to a specific number would have positive impact on accident frequency [46].

Ramp Density: Ramp density and ramp combination are statistically significant and positively related to accident frequency for urban expressway [55].

Lighting: In comparison to both side lighting, point lighting, no lighting, and continuous lighting were found to have a negative effect on accidents [41]. However, the effect of this variable was less studied.

Weather Related Variables: Haq et al. [31] found a increasing impact of rainy day per year on the truck related accidents on expressway. Strong winds increased the problem of controlling heavy vehicles on curves and consequently the frequency of accidents [37].

Pavement Marking: Zhao et al. [59] found raised pavement marking as an effective measure for reducing accidents in expressway tunnels. Lee et al. [60] observed a significant reduction in the average driving speed of the vehicles at the ramps of expressways by using transverse pavement marking and thus consequently found a reduction in accident frequency. Yang et al. [61] compared the effect of traditional marking, fishbone-shaped marking and edge-rate marking on the speed reduction of vehicles at the expressway’s deceleration lane and found that fishbone-shaped marking was effective in reducing vehicles speed from 12.3 km/h to 15.2 km/h.
2.3.8. Interaction Effect Related Variables. It is a well known fact that the road accident occurs due to interaction of different risk factors. Some researchers [11, 30, 47] have tried to understand the interaction effect of risk factors on accident frequency. Mohammadi et al. [47] observed that accident frequency on an intersection in urban area decreases with increase in speed limit and at a speed limit of 55 mph, by increasing the number of lanes. Wang et al. [30] found that total number of rain related accidents reduce with increase in median width, but, single-vehicle rain related accidents increase. This might be due to reduced driving visibility and the driver’s lane control ability, which results in median-related run-off-road accidents in presence of a wider median [30]. Interaction between curve and rain increased accident frequency on expressway and it might be attributed due to decline in skid resistance in rainy weather [11, 30]. Interaction between rain and vertical gradient had reduced the number of accidents and it might be due to alerted driving behavior under difficult climatic situations [11]. Wen et al. [11] also found that interaction of wind speed and vertical gradient could increase accident likelihood.

3. Conclusions

This paper attempts to review findings of more than 60 research articles related to accident prediction modeling on expressways to find out suitable methods of modeling, specific data requirements and issues related to data quality and quantity, and criteria for selection of a variable into the model. One of the major objectives of this work was to identify the areas where the results are still inconclusive. According to the literature review, it has been found that the most of the APMs have been developed for the expressways in the developed countries and only a few researches available on APMs for the expressways in the developing countries, which may be attributed to the late development of the expressways type facilities in these countries as compared to the developed one. However, the traffic composition, vehicle characteristics, driver behavior, and environmental characteristics in developing countries like India are very much different from those in developed countries. So, the APMs developed for expressways in developed countries could not be directly applied in developing countries.

The correlated random-parameter model is the most advanced approach in all the parametric modeling methods to date, as it can account for both the heterogeneous effects of explanatory factors across the road segments and the cross correlations among the random-parameter estimates, but at the same time this approach is computationally expensive.

Non parametric methods of accident prediction modeling for expressways have better predicting capabilities as compared to the parametric one, however, at the same time these models do not provide the quantification of the effect of various explanatory variables on the accident frequency.

Various independent variables (exposure, traffic condition, expressway design, cross-section, segment category, pavement condition, weather condition, and interaction effect related) have been used to develop the APMs for expressway during last three decades. Findings related to variables like percentage of heavy vehicles, vertical gradient, lane width are not conclusive requiring further investigation. However, safety effect of variables like speed limit, roadway lighting, pavement type, and fog are less studied in the literature, and should be investigated further. Literature review suggests that no study has been conducted so far on the accident frequency due to the effect of allowing two wheelers on the expressways. Also, the accidents related to fatigue or drowsiness also needs examination which is one of the leading causes of accidents on expressways in India.
References

1. WHO, “GLOBAL STATUS REPORT ON ROAD SAFETY 2018,” 2018.
2. MoRTH, “Road Accidents in India [2019],” Transp. Res. Wing, Minist. Road Transp. Highw. Gov. India, New Delhi, India., 2020, [Online]. Available: https://morth.nic.in/road-accident-in-india.
3. D. Lord and F. Mannering, “The statistical analysis of crash-frequency data: A review and assessment of methodological alternatives,” Transp. Res. Part A Policy Pract., vol. 44, no. 5, pp. 291–305, 2010, doi: 10.1016/j.tra.2010.02.001.
4. F. L. Mannering, V. Shankar, and C. R. Bhat, “Unobserved heterogeneity and the statistical analysis of highway accident data,” Anal. Methods Accid. Res., vol. 11, pp. 1–16, 2016, doi: 10.1016/j.amar.2016.04.001.
5. G. Singh, M. Pal, Y. Yadav, and T. Singla, “Deep neural network-based predictive modeling of road accidents,” Neural Comput. Appl., vol. 8, 2020, doi: 10.1007/s00521-019-04695-8.
6. G. Singh, S. N. Sachdeva, and M. Pal, “Comparison of three parametric and machine learning approaches for modeling accident severity on non-urban sections of Indian highways,” Adv. Transp. Stud., vol. 45, pp. 123–140, 2018.
7. Z. Ma, H. Zhang, S. I. J. Chien, J. Wang, and C. Dong, “Predicting expressway crash frequency using a random effect negative binomial model: A case study in China,” Accid. Anal. Prev., vol. 98, no. January, pp. 214–222, 2017, doi: 10.1016/j.aap.2016.10.012.
8. D. Mohan, G. Tiwari, and K. Bhatta, “Road safety in India: Status report 2017,” Transp. Res. Inj. Prev. Program. Indian Inst. Technol. Delhi., 2017, [Online]. Available: www.iitd.ac.in/-tripp.
9. F. La Torre, M. Meocci, L. Domenichini, V. Branzi, N. Tanzi, and A. Paliotto, “Development of an accident prediction model for Italian freeways,” Accid. Anal. Prev., vol. 124, no. March, pp. 1–11, 2019, doi: 10.1016/j.aap.2018.12.023.
10. J. Wang, Y. Kong, and T. Fu, “Expressway crash risk prediction using back propagation neural network: A brief investigation on safety resilience,” Accid. Anal. Prev., vol. 124, no. January, pp. 180–192, 2019, doi: 10.1016/j.aap.2019.01.007.
11. H. Wen, X. Zhang, Q. Zeng, and N. N. Sze, “Bayesian spatial-temporal model for the main and interaction effects of roadway and weather characteristics on freeway crash incidence,” Accid. Anal. Prev., vol. 132, no. November, p. 105249, 2019, doi: 10.1016/j.aap.2019.07.025.
12. Q. Zeng, W. Gu, X. Zhang, H. Wen, J. Lee, and W. Hao, “Analyzing freeway crash severity using a Bayesian spatial generalized ordered logit model with conditional autoregressive priors,” Accid. Anal. Prev., vol. 127, no. June, pp. 87–95, 2019, doi: 10.1016/j.aap.2019.02.029.
13. JP Research India Pvt Ltd, “Mumbai – Pune Expressway Road Accident Study,” 2019.[Online]. Available: http://www.jpri.in/publications/.
14. P. Parida, H. S. Nagar, and N. Chakravarthy, “Causes of accidents on Yamuna Expressway (Noida to Greater Noida) and suggested remedial measures,” BVAAP, vol. 20, no. 2, pp. 212–220, 2012, [Online]. Available: http://hdl.handle.net/123456789/18663.
15. L. S. Bisht and G. Tiwari, “Outcomes of Road Safety Audit of Yamuna Expressway,” IIT Delhi Industry Day, 2019.
16. M. A. Hadi, J. Aruldhas, L. F. Chow, and J. A. Wattleworth, “Estimating safety effects of cross-section design for various highway types using negative binomial regression,” Transp.
17. C. Caliendo, M. Guida, and A. Parisi, “A crash-prediction model for multilane roads,” *Accid. Anal. Prev.*, vol. 39, pp. 657–670, 2007, doi: 10.1016/j.aap.2006.10.012.

18. Q. Zeng, J. Sun, and H. Wen, “Bayesian Hierarchical Modeling Monthly Crash Counts on Freeway Segments with Temporal Correlation,” *J. Adv. Transp.*, vol. 2017, no. 24 October, 2017, doi: https://doi.org/10.1155/2017/5391054.

19. M. H. Hosseinlou, A. Mahdavi, and M. J. Nooghabi, “Validation of the influencing factors associated with traffic violations and crashes on freeways of developing countries : A case study of Iran,” *Accid. Anal. Prev.*, vol. 121, no. December, pp. 358–366, 2018, doi: 10.1016/j.aap.2018.06.009.

20. F. Ture Kibar, F. Celik, and F. Wegman, “Analyzing truck accident data on the interurban road Ankara–Aksaray–Eregli in Turkey: Comparing the performances of negative binomial regression and the artificial neural networks models,” *J. Transp. Saf. Secur.*, vol. 11, no. 2, pp. 129–149, 2019, doi: 10.1080/19439962.2017.1363841.

21. D. Kim and Y. Lee, “Modelling crash frequencies at signalized intersections with a truncated count data model,” *Int. J. Urban Sci.*, vol. 17, no. 1, pp. 85–94, 2013, doi: 10.1080/12265934.2013.774702.

22. Y. Zou, L. Wu, and D. Lord, “Modeling over-dispersed crash data with a long tail : Examining the accuracy of the dispersion parameter in Negative Binomial models,” *Anal. Methods Accid. Res.*, vol. 5–6, pp. 1–16, 2015, doi: 10.1016/j.amar.2014.12.002.

23. M. Shirazi, D. Lord, S. S. Dhavala, and S. R. Geedipally, “A semiparametric negative binomial generalized linear model for modeling over-dispersed count data with a heavy tail : Characteristics and applications to crash data,” *Accid. Anal. Prev.*, vol. 91, pp. 10–18, 2016, doi: 10.1016/j.aap.2016.02.020.

24. Q. Hou, X. Meng, J. Leng, and L. Yu, “Application of a random effects negative binomial model to examine crash frequency for freeways in China,” *Physica A*, vol. 509, pp. 937–944, 2018, doi: 10.1016/j.physa.2018.06.088.

25. K. Wang, J. N. Ivan, N. Ravishanker, and E. Jackson, “Multivariate poisson lognormal modeling of crashes by type and severity on rural two lane highways,” *Accid. Anal. Prev.*, vol. 99, pp. 6–19, 2017, doi: 10.1016/j.aap.2016.11.006.

26. Z. Li, W. Wang, P. Liu, L. Bai, and M. Du, “Analysis of Crash Risks by Collision Type at Freeway Diverge Area Using Multivariate Modeling Technique,” *J. Transp. Eng.*, vol. 141, no. 6, 2015, doi: 10.1061/(ASCE)TE.1943-5436.0000757.

27. S. Chen, T. Usman, M. Alinizzi, S. Lavrenz, and S. Labi, “Safety sensitivity to roadway characteristics : A comparison across highway classes,” *Accid. Anal. Prev.*, vol. 123, pp. 39–50, 2019, doi: 10.1016/j.aap.2018.10.020.

28. P. C. Anastasopoulos, “Random parameters multivariate tobit and zero-inflated count data models : Addressing unobserved and zero-state heterogeneity in accident injury-severity rate and frequency analysis,” *Anal. Methods Accid. Res.*, vol. 11, pp. 17–32, 2016, doi: 10.1016/j.amar.2016.06.001.

29. E. S. Park and D. Lord, “Multivariate Poisson – Lognormal Models for Jointly Modeling Crash Frequency by Severity,” *Transp. Res. Rec. J. Transp. Res. Board, No. 2019, Transp. Res. Board Natl. Acad. Washington, D.C.*, pp. 1–6, 2007, doi: 10.3141/2019-01.

30. K. Wang, S. Zhao, and E. Jackson, “Multivariate Poisson Lognormal Modeling of Weather-
Related Crashes on Freeways,” Transp. Res. Rec. J. Transp. Res. Board, No. 2672, Transp. Res. Board Natl. Acad. Washington, D.C., no. 38, pp. 184–198, 2018, doi: 10.1177/0361198118776523.

31. M. T. Haq, M. Zlatkovic, and K. Ksaibati, “Freeway Truck Traffic Safety in Wyoming: Crash Characteristics and Prediction Models,” Transp. Res. Rec. J. Transp. Res. Board, No. 2673, Transp. Res. Board Natl. Acad. Washington, D.C., no. 10, pp. 333–342, 2019, doi: 10.1177/0361198119847980.

32. D. Lord, S. P. Washington, and J. N. Ivan, “Poisson, Poisson-gamma and zero-inflated regression models of motor vehicle crashes: balancing statistical fit and theory,” Accid. Anal. Prev., vol. 37, pp. 35–46, 2005, doi: 10.1016/j.aap.2004.02.004.

33. C. Liu, M. Zhao, W. Li, and A. Sharma, “Multivariate random parameters zero-inflated negative binomial regression for analyzing urban midblock crashes,” Anal. Methods Accid. Res., vol. 17, pp. 32–46, 2018, doi: 10.1016/j.amar.2016.08.004.

34. G. Singh, S. N. Sachdeva, and M. Pal, “M5 model tree based predictive modeling of road accidents on non-urban sections of highways in India,” Accid. Anal. Prev., vol. 96, pp. 108–117, 2016, doi: 10.1016/j.aap.2016.07.010.

35. M. Quddus, “Exploring the Relationship Between Average Speed, Speed Variation, and Accident Rates Using Spatial Statistical Models and GIS Exploring the Relationship Between Average Speed, Speed Variation, and Accident Rates Using Spatial Statistical Models and G,” J. Transp. Saf. Secur., vol. 5, no. 1, pp. 27–45, 2013, doi: 10.1080/19439962.2012.705232.

36. G. I. M. A. Mothafer, T. Yamamoto, and V. N. Shankar, “A negative binomial crash sum model for time invariant heterogeneity in panel crash data: Some insights,” Anal. Methods Accid. Res., vol. 14, no. June, pp. 1–9, 2017, doi: 10.1016/j.amar.2016.12.003.

37. Q. Hou, A. P. Tarko, and X. Meng, “Investigating factors of crash frequency with random effects and random parameters models: New insights from Chinese freeway study,” Accid. Anal. Prev., vol. 120, no. November, pp. 1–12, 2018, doi: 10.1016/j.aap.2018.07.010.

38. Q. Hou, A. P. Tarko, and X. Meng, “Analyzing crash frequency in freeway tunnels: A correlated random parameters approach,” Accid. Anal. Prev., vol. 111, pp. 94–100, 2018, doi: 10.1016/j.aap.2017.11.018.

39. M. Ahmed, H. Huang, M. Abdel-aty, B. Guevara, and M. Marker, “Exploring a Bayesian hierarchical approach for developing safety performance functions for a mountainous freeway,” Accid. Anal. Prev., vol. 43, pp. 1581–1589, 2011, doi: 10.1016/j.aap.2011.03.021.

40. P. C. Anastasopoulos and F. L. Mannering, “A note on modeling vehicle accident frequencies with random-parameters count models,” Accid. Anal. Prev., vol. 41, pp. 153–159, 2009, doi: 10.1016/j.aap.2008.10.005.

41. N. S. Venkataraman, G. F. Ulfarsson, V. Shankar, J. Oh, and M. Park, “Model of Relationship Between Interstate Crash Occurrence and Geometrics Negative Binomial Approach,” Transp. Res. Rec. J. Transp. Res. Board, No. 2236, Transp. Res. Board Natl. Acad. Washington, D.C., pp. 41–48, 2011, doi: 10.3141/2236-05.

42. M. R. Rahman Shao, X. Qin, M. Shirazi, D. Lord, and S. R. Geedipally, “Developing a Random Parameters Negative Binomial-Lindley Model to analyze highly over-dispersed crash count data,” Anal. Methods Accid. Res., vol. 18, no. June, pp. 33–44, 2018, doi: 10.1016/j.amar.2018.04.002.
43. T. U. Saeed, T. Hall, H. Baroud, and M. J. Volovski, “Analyzing road crash frequencies with uncorrelated and correlated random-parameters count models: An empirical assessment of multilane highways,” Anal. Methods Accid. Res., vol. 23, no. September, 2019, doi: 10.1016/j.amar.2019.100101.

44. J. Liu, A. J. Khattak, and B. Wali, “Do safety performance functions used for predicting crash frequency vary across space? Applying geographically weighted regressions to account for spatial heterogeneity,” Accid. Anal. Prev., vol. 109, no. December, pp. 132–142, 2017, doi: 10.1016/j.aap.2017.10.012.

45. AASHTO, “Highway Safety Manual, first edition. 2014 Supplement,” Am. Assoc. State Highw. Transp. Off., 2014.

46. L. Zheng, J. Sun, and X. Meng, “Crash prediction model for basic freeway segments incorporating influence of road geometrics and traffic signs,” J. Transp. Eng. Part A Syst., vol. 144, no. 7, 2018, doi: 10.1061/JTEPBS.000155.

47. M. A. Mohammadi, V. A. Samaranayake, and G. H. Bham, “Crash frequency modeling using negative binomial models: An application of generalized estimating equation to longitudinal data,” Anal. Methods Accid. Res., vol. 2, pp. 52–69, 2014, doi: 10.1016/j.amar.2014.07.001.

48. L. Chang, “Analysis of freeway accident frequencies: Negative binomial regression versus artificial neural network,” Saf. Sci., vol. 43, pp. 541–557, 2005, doi: 10.1016/j.ssci.2005.04.004.

49. Y. Xie, D. Lord, and Y. Zhang, “Predicting motor vehicle collisions using Bayesian neural network models: An empirical analysis,” Accid. Anal. Prev., vol. 39, pp. 922–933, 2007, doi: 10.1016/j.aap.2006.12.014.

50. L. Chang and W. Chen, “Data mining of tree-based models to analyze freeway accident frequency,” J. Safety Res., vol. 36, pp. 365–375, 2005, doi: 10.1016/j.jsr.2005.06.013.

51. R. Eenink, M. Reurings, R. Elvik, J. Cardoso, W. Wichert, and C. Stefan, “Accident Prediction Models and Road Safety Impact Assessment: recommendations for using these tools. Report RI-SWOV-WP2-D2-F. SWOV Institute for Road Safety Research, Leidschendam,” 2007.

52. Q. Hou, X. Huo, and J. Leng, “A correlated random parameters tobit model to analyze the safety effects and temporal instability of factors affecting crash rates,” Accid. Anal. Prev., vol. 134, no. 105326, 2020, doi: 10.1016/j.aap.2019.105326.

53. B. Persaud and L. Dzbik, “Accident prediction models for freeways,” Transp. Res. Rec. 1401, TRB, Natl. Res. Counc. Washington, D.C., vol. 1401, pp. 55–60, 1993.

54. B.-J. Park, K. Fitzpatrick, and D. Lord, “Evaluating the effects of freeway design elements on safety,” Transp. Res. Rec. 2195, TRB, Natl. Res. Counc. Washington, D.C., no. 1, pp. 58–69, 2010, doi: 10.3141/2195-07.

55. L. Wu, J. Sun, and T. Li, “Relationship between Lane Width and Safety along Urban Expressways in Shanghai,” J. Transp. Eng. Part A Syst., vol. 145, no. 3, 2019, doi: 10.1061/JTEPBS.0000216.

56. H. Wen, J. Sun, Q. Zeng, X. Zhang, and Q. Yuan, “The effects of traffic composition on freeway crash frequency by injury severity: A Bayesian multivariate spatial modeling approach,” J. Adv. Transp., 2018, doi: https://doi.org/10.1155/2018/6964828.

57. Q. Zeng and H. Huang, “Bayesian spatial joint modeling of traffic crashes on an urban road network,” Accid. Anal. Prev., vol. 67, pp. 105–112, 2014, doi: 10.1016/j.aap.2014.02.018.

58. A. Montella, L. Colantuoni, and R. Lamberti, “Crash prediction models for rural motorways,”
59. X. Zhao, Y. Ju, H. Li, C. Zhang, and J. Ma, “Safety of Raised Pavement Markers in Freeway Tunnels Based on Driving Behavior,” Accid. Anal. Prev., vol. 145, no. July, 2020, doi: 10.1016/j.aap.2020.105708.

60. S. M. Lee, K. Lee, and S. Kim, “Effectiveness of the Peripheral Transverse Line as Speed-Reduction Treatment on Korean Expressway Ramps,” Int. J. Highw. Eng., vol. 14, no. 6, pp. 85–92, 2012, [Online]. Available: http://www.ksre.or.kr/.

61. Y. Yang, S. M. Easa, X. Zheng, A. Hu, F. Liu, and M. Chen, “Evaluation effects of two types of freeway deceleration markings in China,” PLoS One, vol. 14, no. 8, pp. 1–18, 2019, doi: 10.1371/journal.pone.0220811.