https://doi.org/10.3311/PPtr.16295 | 267
Creative Commons Attribution

Periodica Polytechnica Transportation Engineering, 50(3), pp. 267–272, 2022

Crash Prediction Models and Methodological Issues

Anteneh Afework Mekonnen[1*], Tibor Sipos[1]

[1] Department of Transport Technology and Economics, Faculty of Transportation and Vehicle Engineering, Budapest University of Technology and Economics, H-1521 Budapest, P.O.B. 91, Hungary
[*] Corresponding author, e-mail: anteneh.mekonnen@kjk.bme.hu

Received: 21 April 2020, Accepted: 24 May 2020, Published online: 09 February 2022

Abstract
The conducted literature review aimed to provide an overall perspective on the significant findings of past research works related to vehicle crashes and prediction models. The literature review also provided information concerning past road safety research methodology and viable statistical analysis and computing tools. Though the selection of a specific model hinges on the objective of the research and nature of the response, when compared to statistical modeling techniques, Artificial Neural Networks (ANNs), which can model complex nonlinear relationships among dependent and independent parameters, have been witnessed to be very powerful.

Keywords
crash prediction models, Artificial Neural Network, statistical methods, road safety, soft computing tools, methodological issues

1 Introduction

Mobility is essential for welfare, and passengers demand safe, efficient, and reliable transport systems (Kiss et al., 2013). Crash prediction models significantly contribute to road safety, from establishing relationships among crashes and different covariates to predicting values (Geedipally et al., 2012).

The article's main aim is to compile methodological components that can be used in further research and provide an overall perspective on the significant findings of past research works related to vehicle crashes and prediction models (Szabó and Török, 2020). The conducted literature review also aimed to provide an overall perspective on the significant findings of past research works related to vehicle crashes and the effect of geometric road features on the severity and the number of crashes. The literature review also provided information concerning past road safety research methodology viable statistical analysis techniques. It also aims to describe the Artificial Neural Network technique as a Cesrobust modeling tool.

In recent years, considerable research works have been conducted on crash prediction models. With the increase in computers' computing capability, many researchers attempted to mimic the human brain's functioning using Artificial Neural Network tools. Few experts believe that this approach is at the nascent stage but hold promise for the analysis of the future problems.

2 Crash data and methodological issues

In the past years, crash data and methodological issues were highly discussed. They were proven to be sources of errors in crash models, leading to erroneous forecasting results and wrong inferences related to factors contributing to a vehicle crash.

2.1 Over-dispersion and under-dispersion

Overdispersion means that variance is higher than the mean of the crash data. Over-dispersion is mainly caused by the heterogeneity of the subjects (Agresti, 2002). The number of crash frequencies is highly skewed. An estimated overdispersion parameter would determine a suitable modeling technique, and such type of data is usually modeled with a negative binomial method.

Crash data is sometimes characterized by under-dispersion. Under-dispersion is when a variance is lower than the mean of the crash dataset. It is not as common as overdispersion.

2.2 Temporal and spatial correlation

Crash counts are influenced by temporal factors such as economy, weather, traffic flow, and crash reporting practices (Lord and Persaud, 2000). Ignoring temporal correlation could lead to loss of information and models of distorted risk-factors (Usman et al., 2011).

Cite this article as: Mekonnen, A. A., Sipos, T. (2022) "Crash Prediction Models and Methodological Issues", Periodica Polytechnica Transportation Engineering, 50(3), pp. 267–272. https://doi.org/10.3311/PPtr.16295

Road vehicle crashes can occur at a specific time and location but are affected by comprehensive interaction among spatial factors like human factors, spatial road environments, and other factors (Hong et al., 2013).

2.3 Low sample mean and small sample size

Crash datasets are mostly characterized by small sample size and low sample mean values. (Lord and Miranda-Moreno, 2008) Crashes are considered rare events, and data collection is expensive because of these two main reasons; only a few crashes can be observed for any study. The low sample mean is caused by too many zero-crashes registered and high skewness.

3 Literature review

3.1 Works done based on statistical methods

Mountain et al. (1996) considered G.L.M. modeling to estimate common crashes, and an empirical Bayes procedure was applied to improve the estimates.

Turner and Nicholson (1998) described the implementation of G.L.M. models to forecast specific crash types at junctions. The result showed that generalized linear models of separate crashes are better than models of the total crashes. Jones and Jørgensen (2003) implemented a regression model developed recently, which measures multiple influences on casualty outcomes, to introduce the potential of multilevel models. Chin and Quddus (2003) described using the Random-Effect Negative-Binomial (RENB) model to identify elements that affect intersection safety.

Caliendo et al. (2007) developed crash-prediction models namely Negative-Binomial, Poisson, and Negative-Multinomial regression models to the model frequency of crash occurrence. The models appeared to be useful for many applications.

Milton et al. (2008) applied the mixed logit model and demonstrated a substantial promise as a modeling tool in road safety studies.

Berhanu (2004) carried out an extensive crash study on the Addis Ababa-Nazareth road. He tried to fit both quasi-Poisson and Negative-Binomial models, using only one-year crash data and concluded that the Negative-Binomial model was generally preferable.

Tuladhar and Justo (1981), Kadyali et al. (1984), and Chandra et al. (2004) modeled crashes using regression techniques.

3.2 Works done based on Artificial Neural Network

Hashemi et al. (1995) formulated a logistic regression model, multiple discriminant analysis, and a neural network to predict container-ship crashes. The neural network has shown a better performance than the others. Delen et al. (2006) employed Artificial Neural Networks and all models, compared to a model with a five-category outcome variable, were witnessed to have a better predictive capability.

Cansiz and Easa (2011) targeted their study mainly to forecast collision-frequency on the vertical curves combined with horizontal tangents using ANNs. It is witnessed that Artificial Neural Network models have the lowest mean-square-error value compared to the statistical models (Moghaddam et al., 2011).

Bayata et al. (2011), Akin et al. (2017), Moghaddam et al. (2011), and Chiou (2006) used Artificial Neural Networks to model crashes and identified crash-related factors.

Chang (2005) employed an Artificial Neural Network model and a negative-binomial model to analyze crash data. The Artificial Neural Network model was better for locations with one or more crashes when the negative-binomial model was slightly better for zero crash locations with no crashes.

4 Modeling techniques

The most used crash prediction models are explained in the following subtopics, from basic ones to mostly practiced and state-of-the-art ones.

4.1 Multiple linear regression model

Linear regression is the fundamental modeling technique in crash data modeling. It is not commonly utilized because crash data usually violate the linear regression analysis assumptions, for instance, normal error structure and constant error variance (Al-Qadi et al., 2008). Its general form is given by Eq. (1):

$$y_{it} = \beta_\circ + \sum_k \beta_k x_{itk} + \varepsilon_{it}, \tag{1}$$

where for $i = n$ observations:

- y_{it}: number of observed crashes in the segment i at time t;
- x_{it}: explanatory variables or factors contributing to crashes;
- $\beta_\circ$: intercept (constant term);
- β_k: regression coefficients;
- ε_{it}: the model's error term (also known as the residuals).

4.2 Poisson regression model

Crash data are count data, and the traditional regression models are not suitable because they do not account for heterogeneity among observations (Elvik, 2011). Poisson regression assumes that the observations' mean and variance are equal (Moksony and Hegedűs, 2014). The general form of Poisson regression is explained by Eq. (2):

$$P(n_i) = \lambda i e^{(-\lambda i)} / n!, \quad (2)$$

where:

- $P(n_i)$: probability of n number of crashes occurring at segment i;
- n_i: number of observations per time period;
- λi: expected crash-frequency on the segment i.

λi is expressed with Eq. (3):

$$\lambda i = EXP(\beta \boldsymbol{x}_i), \quad (3)$$

where:

- $\boldsymbol{x}_i$: a vector of independent variables;
- β: a vector of coefficients of the independent variables $\boldsymbol{x}_i$.

4.3 Negative binomial (Poisson-gamma) model

Crash data have unique behavior like over-dispersion (Shirazi et al., 2016). Overdispersion is when a variance is greater than the mean of the observations. Underestimation or deflation of standard errors of estimates is caused by overdispersion (Hilbe, 2011).

Because of overdispersion associated with Poisson model usage, researchers usually consider the negative binomial model a potential alternative (Agresti, 2007).

The negative binomial (Poisson-gamma) distribution has the same 'sample space' as the Poisson distribution and has an additional parameter used to model the variance. This parameter is referred to, unsurprisingly, as the dispersion parameter (Gardner et al., 1995). The model is specialized by Eq. (4):

$$y_{it} \sim \text{Poisson}(\theta_{it}), \quad (4)$$

where:

- y_{it}: the number of crashes observed in the segment i at time t.
- θ_{it}: expected Poisson-rate, which is the expected crash-frequency for the segment i at time t.

The Poisson-rate is modeled as a function of the covariates following the log link shown in Eq. (5):

$$\log(\theta_{it}) = \beta_\circ + \sum_k \beta_k x_{itk} + \varepsilon_{it}, \quad (5)$$

where $\beta_\circ$ is called the *intercept*. The regression coefficients β_k are unknown parameters that are estimated from a set of data. ε_{it} is the error-term for the segment i at time t.

4.4 Poisson log-normal model

Poisson log-normal regression models have been proposed recently in highway traffic safety analysis as a better means to handle low sample mean, compared to negative-binomial (Poisson–gamma) models (Aguero-Valverde and Jovanis, 2008; Lord and Miranda-Moreno, 2008; Ma et al., 2008; Miaou and Lord, 2003). Even though it gives greater flexibility, the marginal distribution Poisson-lognormal models don't have closed-form like Poisson-gamma models, making the model estimation more complex (Aguero-Valverde, 2013). The Poisson-log-normal model is like the Poisson-gamma model, except that the Poisson-rate is modeled by log-normal distribution given in Eq. (6) (Aguero-Valverde, 2013):

$$\log(\theta_{it}) = \beta_\circ + \sum_k \beta_k x_{itk} + v_{it}, \quad (6)$$

where v_{it} is the random effect.

4.5 Zero-inflated Poisson and negative-binomial model

It is ubiquitous to encounter crash data containing excessive zero observations. Since the number of zeros is greater than expected in using the negative binomial regression or Poisson regression models, they are called zero-inflated data (Jang et al., 2010). This method has been applied by many researchers recently, and the following are among them, Kumara and Chin (2003), Lord et al. (2007), Huang and Chin (2010), Jiang et al. (2013), and Xu et al. (2017).

4.6 Conway-Maxwell Poisson regression model

This method is an extension of the Poisson regression method and was introduced by researchers to analyze under-dispersed or over-dispersed count data (Lord et al., 2008). The method was applied by Boatwright et al. (2006) and Lord and Guikema (2012).

4.7 Generalized linear model

Generalized Linear Models with negative-binomial error distribution have been used for road safety analysis (Gardner et al., 1995; Geedipally et al., 2012; Hilbe, 2011; Shirazi et al., 2016; Wood, 2002). G.L.M.s with log link function and negative binomial distributions are widely used to relate crashes with explanatory variables (Wood, 2005; Maher and Summersgill, 1996).

The Generalized Linear Model consists of three elements. A linear predictor, an exponential family of probability distributions, and a link function (Aliakbar Golkar and Valizadeh Haghi, 2011).

4.8 Random parameter model

The random-parameters negative-binomial regression model allows to account and correct for heterogeneity that can occur from different factors (Anastasopoulos and Mannering, 2009).

Greene and Hensher (2007) have developed estimation procedures to incorporate random parameters in Poisson and negative binomial count-data models and to account for heterogeneity with random parameters.

4.9 Artificial Neural Networks (ANNs)

They are made of parallel distributed information processing systems, mimicking human brains' nature, which can model complex nonlinear relationships among dependent and independent parameters. Errors are very low as the function works based on error backpropagation (Moghaddam et al., 2011).

Numerous researchers applied the method in safety analysis and crash modeling. A few of them are Hashemi et al. (1995), Delen et al. (2006), Moghaddam et al. (2011), Bayata et al. (2011), Cansiz and Easa (2011), Zeng et al. (2016), and Chakraborty et al. (2019).

4.10 Hybrid intelligent genetic algorithms (G.A.s)

Liyan and Chunfu (2009) came to the conclusion in their study that the genetic algorithms' accuracy and generalization abilities are better than a neural network. Amiri et al. (2020) concluded that even though ANNs outsmart G.A.s, they model high severity crashes better than Artificial Neural Networks.

Many recent studies tried to combine the best out of these two methods in improving the learning of ANNs by G.A.s.

5 Conclusions

This paper has presented a review of crash prediction models. In the past few years, many statistical methodological innovations have been established, holding a great promise to enhance our understanding of the factors that affect crash frequencies.

In statistical modeling techniques, it is assumed that there is a relationship between the dependent and independent variables, which may sometimes be challenging to get in case of complex crashes. If this assumption is violated, it will lead to erroneous results.

The selection of a specific model hinges on the research's objective and the nature of the response compared to statistical modeling techniques. Artificial Neural Networks (ANNs), which can model complex nonlinear relationships among dependent and independent parameters, have been witnessed to be very powerful.

Few researchers believe that this approach is at a nascent stage but hold promise for the analysis of the future problems.

References

Agresti, A. (2002) "Categorical Data Analysis", Wiley-Interscience, John Wiley & Sons, INC., Hoboken, NJ, USA. https://doi.org/10.1002/0471249688

Agresti, A. (2007) "An Introduction to Categorical Data Analysis", Wiley-Interscience, John Wiley & Sons, INC., Hoboken, NJ, USA .

Aguero-Valverde, J. (2013) "Full Bayes Poisson gamma, Poisson log-normal, and zero inflated random effects models: Comparing the precision of crash frequency estimates", Accident Analysis & Prevention, 50, pp. 289–297. https://doi.org/10.1016/j.aap.2012.04.019

Aguero-Valverde, J., Jovanis, P. P. (2008) "Analysis of Road Crash Frequency with Spatial Models", Transportation Research Record: Journal of the Transportation Research Board, 2061(1), pp. 55–63. https://doi.org/10.3141/2061-07

Akin, M., Onderdonk, A., Dzakpasu, R., Guo, Y. (2017) "Functional Reconstruction of Dyadic and Triadic Subgraphs in Spiking Neural Networks", In: Cherifi, H., Gaito, S., Quattrociocchi, W., Sala, A. (eds.) Complex Networks & Their Applications V. COMPLEX NETWORKS 2016, Studies in Computational Intelligence, Springer, Cham, Switzerland, pp. 697–708. https://doi.org/10.1007/978-3-319-50901-3_55

Aliakbar Golkar, M., Valizadeh Haghi, H. (2011) "Using a Multivariate Doe Method for Congestion Study in Distribution Systems Under Impacts of Plug-in Electric Vehicles, In: 21st International Conference on Electricity Distribution, Frankfurt, Germany, Article number: 0718.

Al-Qadi, I., Sayed, T., Alnuaimi, N., Masad, E. (eds.) (2008) "Efficient Transportation and Pavement Systems: Characterization, Mechanisms, Simulation, and Modeling", CRC Press, London, UK. https://doi.org/10.1201/9780203881200

Amiri, A. M., Sadri, A., Nadimi, N., Shams, M. (2020) "A comparison between Artificial Neural Network and Hybrid Intelligent Genetic Algorithm in predicting the severity of fixed object crashes among elderly drivers", Accident Analysis & Prevention, 138, Article number: 105468.
https://doi.org/10.1016/j.aap.2020.105468

Anastasopoulos, P. C., Mannering, F. L. (2009) "A note on modeling vehicle accident frequencies with random-parameters count models", Accident Analysis & Prevention, 41(1), pp. 153–159.
https://doi.org/10.1016/j.aap.2008.10.005

Bayata, H. F., Hattatoglu, F., Karsli, N. (2011) "Modeling of monthly traffic accidents with the artificial neural network method", International Journal of the Physical Sciences, 6(2), pp. 244–254.
https://doi.org/10.5897/IJPS10.606

Berhanu, G. (2004) "Models relating traffic safety with road environment and traffic flows on arterial roads in Addis Ababa", Accident Analysis & Prevention, 36(5), pp. 697–704.
https://doi.org/10.1016/j.aap.2003.05.002

Boatwright, P., Borle, S., Kadane, J. B., Minka, T. P., Shmueli, G. (2006) "Conjugate analysis of the Conway-Maxwell-Poisson distribution", Bayesian Analysis, 1(2), pp. 363–374.
https://doi.org/10.1214/06-BA113

Caliendo, C., Guida, M., Parisi, A. (2007) "A crash-prediction model for multilane roads", Accident Analysis & Prevention, 39(4), pp. 657–670.
https://doi.org/10.1016/j.aap.2006.10.012

Cansiz, O. F., Easa, S. M. (2011) "Using Artificial Neural Network to Predict Collisions on Horizontal Tangents of 3D Two-Lane Highways", World Academy of Science, Engineering and Technology, International Journal of Civil and Environmental Engineering, 5(10), pp. 491–500.

Chakraborty, A., Mukherjee, D., Mitra, S. (2019) "Development of pedestrian crash prediction model for a developing country using artificial neural network", International Journal of Injury Control and Safety Promotion, 26(3), pp. 283–293.
https://doi.org/10.1080/17457300.2019.1627463

Chandra, S., Prashanth, K. (2004) "Accident Analysis on two lane roads", Highway Research Bulletin, Indian Road Congress, 70, pp. 77-92.

Chang, L. Y. (2005) "Analysis of freeway crash frequencies: Negative binomial regression versus artificial neural network", Safety Science, 43(8), pp. 541–557.
https://doi.org/10.1016/j.ssci.2005.04.004

Chin, H. C., Quddus, M. A. (2003) "Applying the random effect negative binomial model to examine traffic accident occurrence at signalized intersections", Accident Analysis & Prevention, 35(2), pp. 253–259.
https://doi.org/10.1016/S0001-4575(02)00003-9

Chiou, Y. C. (2006) "An artificial neural network-based expert system for the appraisal of two-car crash accidents", Accident Analysis & Prevention, 38(4), pp. 777–785.
https://doi.org/10.1016/j.aap.2006.02.006

Delen, D., Sharda, R., Bessonov, M. (2006) "Identifying significant predictors of injury severity in traffic accidents using a series of artificial neural networks", Accident Analysis & Prevention, 38(3), pp. 434–444.
https://doi.org/10.1016/j.aap.2005.06.024

Elvik, R. (2011) "Assessing causality in multivariate crash models", Accident Analysis & Prevention, 43(1), pp. 253–264.
https://doi.org/10.1016/j.aap.2010.08.018

Gardner, W., Mulvey, E. P., Shaw, E. C. (1995) "Regression analyses of counts and rates: Poisson, overdispersed Poisson, and negative binomial models", Psychological Bulletin, 118(3), pp. 392–404.
https://doi.org/10.1037/0033-2909.118.3.392

Geedipally, S. R., Lord, D., Dhavala, S. S. (2012) "The negative binomial-Lindley generalized linear model: Characteristics and application using crash data", Crash Analysis & Prevention, 45, pp. 258–265.
https://doi.org/10.1016/j.aap.2011.07.012

Greene, W. H., Hensher, D. A. (2007) "Heteroscedastic control for random coefficients and error components in mixed logit", Transportation Research Part E: Logistics and Transportation Review, 43(5), pp. 610–623.
https://doi.org/10.1016/j.tre.2006.02.001

Hashemi, R., Le Blanc, L. A., Rucks, C. T., Shearry, A. (1995) "A Neural Network for Transportation Safety Modeling", Expert Systems with Applications, 9(3), pp. 247–256.
https://doi.org/10.1016/0957-4174(95)00002-Q

Hilbe, J. M. (2011) "Negative Binomial Regression", Cambridge University Press, Cambridge, UK.

Hong, J. Y., Lee, S. B., Lim, J., Kim, J. (2013) "Application of Spatial Econometrics Analysis for Traffic Accident Prediction Models in Urban Areas", Proceedings of the Eastern Asia Society for Transportation Studies, 9, pp. 390–397.

Huang, H., Chin, H. C. (2010) "Modeling road traffic crashes with zero-inflation and site-specific random effects", Statistical Methods and Applications, 19(3), pp. 445–462.
https://doi.org/10.1007/s10260-010-0136-x

Jang, H., Lee, S., Kim, S. W. (2010) "Bayesian analysis for zero-inflated regression models with the power prior: Applications to road safety countermeasures", Accident Analysis & Prevention, 42(2), pp. 540–547.
https://doi.org/10.1016/j.aap.2009.08.022

Jiang, X., Huang, B., Zaretzki, R. L., Richards, S., Yan, X., Zhang, H. (2013) "Investigating the influence of curbs on single-vehicle crash injury severity utilizing zero-inflated ordered probit models", Accident Analysis & Prevention, 57, pp. 55–66.
https://doi.org/10.1016/j.aap.2013.03.018

Jones, A. P., Jørgensen, S. H. (2003) "The use of multilevel models for the prediction of road crash outcomes", Accident Analysis & Prevention, 35(1), pp. 59–69.
https://doi.org/10.1016/S0001-4575(01)00086-0

Kadiyali, L., Venketesan, S. (1984) "Traffic Accidents, forecast and remedies", Indian Highway, 12, 7–12.

Kiss, Á. O., Sesztakov, V., Török, Á. (2013) "Road Safety Analysis in Győr", Periodica Polytechnica Transportation Engineering, 41(1), pp. 51–56.
https://doi.org/10.3311/PPtr.7098

Kumara, S. S. P., Chin, H. C. (2003) "Modeling Accident Occurrence at Signalized Tee Intersections with Special Emphasis on Excess Zeros", Traffic Injury Prevention, 4(1), pp. 53–57.
https://doi.org/10.1080/15389580309852

Liyan, Q., Chunfu, S. (2009) "Macro Prediction Model of Road Traffic Accident Based on Neural Network and Genetic Algorithm", In: 2009 Second International Conference on Intelligent Computation Technology and Automation", Changsha, China, pp. 354–357. https://doi.org/10.1109/ICICTA.2009.93

Lord, D., Persaud, B. N. (2000) "Accident Prediction Models With and Without Trend: Application of the Generalized Estimating Equations Procedure", Transportation Research Record: Journal of the Transportation Research Board, 1717(1), pp. 102–108. https://doi.org/10.3141/1717-13

Lord, D., Washington, S., Ivan, J. N. (2007) "Further notes on the application of zero-inflated models in highway safety", Accident Analysis & Prevention, 39(1), pp. 53–57. https://doi.org/10.1016/j.aap.2006.06.004

Lord, D., Guikema, S. D., Geedipally, S. R. (2008) "Application of the Conway–Maxwell–Poisson generalized linear model for analyzing motor vehicle crashes", Accident Analysis & Prevention, 40(3), pp. 1123–1134. https://doi.org/10.1016/j.aap.2007.12.003

Lord, D., Miranda-Moreno, L. F. (2008) "Effects of the low sample mean values and small sample size on the estimation of the fixed dispersion parameter of Poisson-gamma models for modeling motor vehicle crashes: A Bayesian perspective", Safety Science, 46(5), pp. 751–770. https://doi.org/10.1016/j.ssci.2007.03.005

Lord, D., Guikema, S.D. (2012) "The Conway-Maxwell-Poisson model for analyzing crash data", Applied Stochastic Models in Business and Industry, 28(2), pp. 122–127. https://doi.org/10.1002/asmb.937

Ma, J., Kockelman, K. M., Damien, P. (2008) "A multivariate Poisson-lognormal regression model for prediction of crash counts by severity, using Bayesian methods", Accident Analysis & Prevention, 40(3), pp. 964–975. https://doi.org/10.1016/j.aap.2007.11.002

Maher, M. J., Summersgill, I. (1996) "A comprehensive methodology for the fitting of predictive accident models", Accident Analysis & Prevention 28(3), pp. 281–296. https://doi.org/10.1016/0001-4575(95)00059-3

Miaou, S. P., Lord, D. (2003) "Modeling Traffic Crash-Flow Relationships for Intersections: Dispersion Parameter, Functional Form, and Bayes Versus Empirical Bayes Methods", Transportation Research Record: Journal of the Transportation Research Board, 1840(1), pp. 31–40. https://doi.org/10.3141/1840-04

Milton, J. C., Shankar, V. N., Mannering, F. L. (2008) "Highway accident severities and the mixed logit model: An exploratory empirical analysis", Accident Analysis & Prevention, 40(1), pp. 260–266. https://doi.org/10.1016/j.aap.2007.06.006

Moghaddam, F. R., Afandizadeh, S., Ziyadi, M. (2011) "Prediction of accident severity using artificial neural networks", International Journal of Civil Engineering, 9(1), pp. 41–48.

Moksony, F., Hegedűs, R. (2014) "The use of Poisson regression in the sociological study of suicide", Corvinus Journal of Sociology and Social Policy, 5(2), pp. 97–114. https://doi.org/10.14267/cjssp.2014.02.04

Mountain, L., Fawaz, B., Jarrett, D. (1996) "Accident prediction models for roads with minor junctions", Accident Analysis & Prevention, 28(6), pp. 695–707. https://doi.org/10.1016/S0001-4575(96)00042-5

Shirazi, M., Lord, D., Dhavala, S. S., Geedipally, S. R. (2016) "A semi-parametric negative binomial generalized linear model for modeling over-dispersed count data with a heavy tail: Characteristics and applications to crash data", Accident Analysis & Prevention, 91, pp. 10–18. https://doi.org/10.1016/j.aap.2016.02.020

Szabó, Z., Török, Á. (2020) "Spatial Econometrics – Usage in Transportation Sciences: A Review Article", Periodica Polytechnica Transportation Engineering, 48(2), pp. 143–149. https://doi.org/10.3311/PPtr.12047

Tuladhar, S. B. S., Justo, C. E. G. (1981) "Analysis of Accident Rates: A Case Study", Highway Research Bulletin (New Delhi), 16, pp. 32–59.

Turner, S., Nicholson, A. (1998) "Intersection accident estimation: the role of intersection location and non-collision flows", Accident Analysis & Prevention, 30(4), pp. 505–517. https://doi.org/10.1016/S0001-4575(97)00095-X

Usman, T., Fu, L., Miranda-Moreno, L. F. (2011) "Accident Prediction Models for Winter Road Safety: Does Temporal Aggregation of Data Matter?", Transportation Research Record: Journal of the Transportation Research Board, 2237(1), pp. 144–151. https://doi.org/10.3141/2237-16

Wood, G. R. (2002) "Generalised linear accident models and goodness of fit testing", Accident Analysis & Prevention, 34(4), pp. 417–427. https://doi.org/10.1016/S0001-4575(01)00037-9

Wood, G. R. (2005) "Confidence and prediction intervals for generalised linear accident models", Accident Analysis & Prevention, 37(2), pp. 267–273. https://doi.org/10.1016/j.aap.2004.10.005

Xu, P., Huang, H., Dong, N., Wong, S. C. (2017) "Revisiting crash spatial heterogeneity: A Bayesian spatially varying coefficients approach", Accident Analysis & Prevention, 98, pp. 330–337. https://doi.org/10.1016/j.aap.2016.10.015

Zeng, Q., Huang, H., Pei, X., Wong, S. C. (2016) "Modeling nonlinear relationship between crash frequency by severity and contributing factors by neural networks", Analytic Methods in Accident Research, 10, pp. 12–25. https://doi.org/10.1016/j.amar.2016.03.002