Accident Analysis and Prevention 156 (2021) 106103

Contents lists available at ScienceDirect

Accident Analysis and Prevention

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

Application of different negative binomial parameterizations to develop safety performance functions for non-federal aid system roads

Ali Khodadadi[a,*], Ioannis Tsapakis[b], Subasish Das[b], Dominique Lord[a], Yingfeng Li[c]

[a] Texas A&M University, 3136 TAMU, College Station, TX 77843-3136, United States
[b] Texas A&M Transportation Institute, 3500 NW Loop 410, San Antonio, TX 78229, United States
[c] Virginia Tech Transportation Institute, 3500 Transportation Research Plaza, Building 1 R207, Blacksburg, VA 24061, United States

ARTICLE INFO

Keywords:
Non-federal aid system roads
Safety performance function
Negative binomial
Negative binomial-Lindley
Variance structure
Dispersion structure

ABSTRACT

Safety performance functions (SPFs) are the main building blocks in understanding the relationships between crash risk factors and crash frequencies. Many research efforts have focused on high-volume roadways that typically experience more crashes. A few studies have documented SPFs for non-federal aid system (NFAS) roads including rural minor collectors, rural local roads, and urban local roads. NFAS roads are characterized by unique features such as lower speeds, and shorter segment lengths, and they usually experience fewer crashes given the low exposure of these roads. As a result, there is a clear need to investigate the associated safety issues of NFAS roadways and generate distinct SPFs for them. The main objective of this study is to bridge the gap in the literature and develop SPFs for NFAS roads. This study examined the application of traditional negative binomial and zero-favored negative binomial models (i.e., negative binomial-Lindley). Both groups of models were formulated by different variance and dispersion structures. Using crash, roadway inventory, and traffic volume data from 2014 to 2018 in Virginia, the results showed that the NB-L models perform better than the traditional NB models. Furthermore, an appropriate variance structure along with a reasonably chosen dispersion function can further improve the model performance.

1. Introduction

Non-federal aid system (NFAS) roads are categorized into three functional classes: rural minor collectors (6R), rural local (7R), and urban local roads (7U). These roadways are not considered high-volume roads, but they account for more than 75% of the total roadway mileage in the country. As a result, evaluating the safety performance of these roadways is of high importance.

A safety performance function (SPF) is a statistical model (more specifically a crash-frequency model) that estimates the average crash frequency for a specific facility type under certain base conditions (HSM, 2010). In general, SPFs are developed using roadway characteristics and observed crash data at facilities of the same type with similar geographical and geometrical characteristics over a certain period of time. Materials included in part C of the Highway Safety Manual (HSM) intend to provide a basic understanding of predictive methods to estimate the expected average crash frequency of a facility (segment or intersection) using roadway characteristics, such as annual average daily traffic (AADT), segment length, number of lanes, etc. However, the HSM only provides SPFs for three facility types, (1) rural two-lane two-way roads, (2) rural multi-lane highways, and (3) urban and suburban arterials. These roadways are categorized as high-volume roads that are more likely to pose safety challenges. Compared to high-speed and high-volume roadways, fewer research studies have been done to develop SPFs for lower functional classes. This is primarily attributed to inadequate or unavailable traffic information about these roads. The high cost and time-consuming task of data collection can limit local agencies ability to conduct safety improvement programs for lower functional classes. However, as of 2016, the U.S. Department of Transportation requires states to collect traffic volume information for all public paved roads, including both federal aid system (FAS) and NFAS roads. Having AADT data collected or estimated from the short term count or permanent sites, regardless of the sampling techniques or estimation method, can significantly clear the way to conduct data-driven safety analysis and introduce advanced measures to evaluate the safety performance of NFAS roads.

* Corresponding author.
E-mail addresses: a.khodadadi1994@tamu.edu (A. Khodadadi), i-tsapakis@tti.tamu.edu (I. Tsapakis), s-das@tti.tamu.edu (S. Das), dlord@civil.tamu.edu (D. Lord), eli@vtti.vt.edu (Y. Li).

https://doi.org/10.1016/j.aap.2021.106103
Received 9 December 2020; Received in revised form 12 March 2021; Accepted 22 March 2021

Although not much research has focused on the development of the SPFs for NFAS roads, some studies have attempted to analyze the safety issues related to low-volume roads, which are a large part of NFAS roads (Das et al., 2019). Zegeer et al. (1994) attempted to quantify the effect of roadway width on low-volume (AADT < 2000 vehicle per day or vpd) rural roadway crashes. They found that wider roadways, the presence of a shoulder, and paving roadways with volume higher than 250 vpd can significantly improve safety on rural low-volume roads. Stamatiadis et al. (1999) tried to determine the influential factors of crashes on low-volume roadways. They observed that crash frequency on low-volume roadways is a function of the same parameters as found effective in other roadway functional classes. Cook (2010) employed three different segmentation methods, and for each method established four different SPFs for very low-volume roadways, split into four classes (paved 1–99 vpd, paved 100–400 vpd, unpaved 1–99 vpd, and unpaved 100–400 vpd). Roadway length, AADT, lane width, shoulder width, shoulder type, and terrain were the variables that the authors initially considered to develop SPFs. Dell'Acqua and Russo (2011) subcategorized low-volume (AADT < 1000 vpd) rural roads into rolling/flat and mountainous terrain and developed two separate models to predict the number of crashes as a function of environmental, geometrical, and roadway characteristics. The authors found that lowering speed limit and changing curvature rate and roadway width might lower crash density on low-volume roads. Stapleton et al. (2019) utilized the mixed effect negative binomial regression model to develop SPFs for rural low-volume intersections (major AADT < 2000 vpd). They compared the developed SPFs with the base models reported in the HSM and observed that the HSM models overpredict crash frequency at both four-leg and three-leg stop-controlled intersections. Using crash data and estimated traffic volumes from the North Carolina local network, Das et al. (2019) aimed to predict crash frequency for each category of NFAS roads, separately. This study developed SPFs for 6R, 7R, and 7U road types, as a function of roadway length and traffic volume data. They also found that AADT estimation error can affect the predicted number of crashes occurring on these roadways.

As mentioned above, the HSM's SPFs are limited to specific high-volume and high-speed roadways and may not be transferable to lower road classes. On the other hand, previous studies did not perform an in-depth investigation to quantify the safety performance of the NFAS roads. To the best of our knowledge, except for the study conducted by Das et al. (2019), no other studies have attempted to develop distinct SPFs for NFAS roads, but rather for low-volume roadways that may or may not be classified as the same road type as NFAS roadways. Furthermore, previous studies utilized traditional count models to develop SPFs for low-volume roadways and did not examine more advanced and innovative models that are capable of accounting for unique characteristics of crash data, such as having a large proportion of zeros in the dataset. Also, NFAS roads are characterized by unique features such as lower speeds (35–55 and 20–45 mph for collector and local roads, respectively), shorter segment lengths, and fewer crashes given the exposure (i.e., data characterized by low sample mean values). Therefore, there is a clear need to further investigate the safety performance of the NFAS roads and generate distinct SPFs for these roadways using more advanced models for comparison with traditional model formulations.

Generalized linear models, and more specifically, the negative binomial (NB) regression models have been extensively used in SPF development. The negative binomial or Poisson-Gamma mixture distribution is the generalization of the Poisson distribution and has been considered the most popular model in highway safety (Lord and Mannering, 2010; Lord et al., 2021). The NB distribution eases the assumption of equality of the mean and variance held in the Poisson regression model. In general, the negative binomial model allows for the variance to be higher than the mean in order to capture the variation in the dataset. This ability can be further improved by changing the variance and dispersion structure, or mixing the NB distribution with other distributions. In the following, each of these improvements is discussed in detail.

- Dispersion structure: the dispersion parameter in the NB model allows for the variance to be greater than the mean by forming a quadratic relationship between the mean and the variance. Although past studies assumed that the dispersion parameter is invariant of roadway features, recent findings confirm that the dispersion parameter varies from site to site and is dependent upon site characteristics such as segment length and AADT (Geedipally et al., 2009; Lord and Park, 2008; Cafiso et al., 2010). In another study, Lord and Park (2008) conducted an Empirical Bayes (EB) method using both fixed and varying dispersion parameters to rank hazardous sites. They employed different functional forms of the dispersion parameters in the NB models and concluded that the varying dispersion parameter provides a better statistical fit. Cafiso et al. (2010) investigated the association between rural roadway length and the dispersion parameter and found that the dispersion parameter variation is more significant for shorter segments. Also, Geedipally et al. (2009) used three different datasets to empirically examine the effect of the varying dispersion structures. They evaluated ten different structures as functions of roadway length and AADT. The authors concluded that selecting a suitable functional form and an appropriate combination of covariate sets for the dispersion structure greatly depends on the dataset being used. In another study, Meng et al. (2020) attempted to develop SPFs with both fixed and varying dispersion parameter for unsignalized intersections in Texas. The authors found that similar to the segment SPFs, intersection SPFs are also improved when using the varying dispersion parameters.
- Variance structure: besides the relationship between roadway characteristics and the dispersion parameter, the type of mean-variance relationship that the dispersion parameter makes is of importance as well. The dispersion parameter in the NB model allows for the variance to be higher than the mean by forming a quadratic relationship between the mean and variance. Different variance structures lead to different parameterizations of the NB distribution. Cameron and Trivedi (2013) proposed two popular forms of the NB distribution, called NB-1 and NB-2 (the latter usually referred to as NB), in which the digit refers to the exponent of the mean value multiplied by the dispersion parameter in the mean-variance equation. Pei et al. (2011), Mehta and Lou (2013) and Wang et al. (2019) employed both the NB-1 and NB-2 to model crash frequency for different severity levels and to develop SPFs, respectively. They all concluded that the NB-1 does not perform better than the NB-2, which supported the findings of Lord et al. (2012), indicating that the NB-1 is less flexible to capture the large variations existing in the crash data. Both NB-1 and NB-2 parameterizations are nested in an unrestricted general model, entitled as NB-P, which does not restrict the variance structure (Greene, 2008). Ismail and Zamani (2013) assessed the application of different variance structure of different count models, such as the Poisson and NB model, in the over and under dispersed data condition. They concluded that the NB-P model outperforms other NB parameterizations. Also, Wang et al. (2019) found that the NB-P provides more flexibility to the model, and hence is preferable over the NB-1 and NB-2 when developing SPFs for rural intersections. They concluded that the variance structure of the NB-P model could even capture some of the variations in the dispersion parameter.

- Mixture distribution: although the NB regression model accounts for the built-in dispersion in crash data, crash datasets are naturally characterized by unique features such as having a large number of zeros or a heavy tail, which the traditional NB models cannot efficiently deal with. To overcome this problem, Lord and Geedipally (2011) and Geedipally et al. (2012) examined the application of the NB and Lindley mixture distribution (NB-L) introduced by Zamani and Ismail (2010), in crash data analysis.[1] They found that the NB-L model, while preserving the NB characteristics, provides a better fit compared to the traditional NB models for the datasets suffering from a large proportion of zeros or high dispersion problems.

All the above-mentioned regression models could be estimated in either frequentist or Bayesian framework. The superiority of the Bayesian paradigm has been documented in the literature from different points of view. First of all, when limited data is available, the full Bayesian (FB) method can yield unbiased estimates by incorporating common Beliefs (prior distribution) about the variable of interest into the analysis (Heydari et al., 2014; Lord and Miranda-Moreno, 2008). Also, despite frequentist analysis, which requires a considerable number of repeated random trials to build the confidence intervals, Bayesian methods represent the hypothesis uncertainty in a natural probabilistic way and attach it to the modeling procedure. Moreover, as the hierarchy level grows and the data structure gets more complex, the frequentist method needs more computational effort to find a closed-form of the distribution or employ a simulation-based solution; however, Bayesian methods can take advantage of both, Bayes theorem and its hierarchical nature to easily draw samples from the posterior distribution of the parameter of interest using the Markov Chain Monte Carlo (MCMC) simulation (Heydari et al., 2014; Lord et al., 2021). Full Bayesian paradigm has been extensively used in various settings including the hierarchical Poisson model (Pawlovich et al., 2006), NB model (Heydari et al., 2014; Farid et al., 2017), Poisson log-normal model (Aguero-Valverde and Jovanis, 2009), and NB-L model (Lord and Geedipally, 2011; Geedipally et al., 2012) for various crash analysis such as crash frequency prediction, site ranking, and SPF development.

1.1. Study objective

The main objective of this study is to apply different forms of count models to develop SPFs for NFAS roads. Given the share of NFAS roads from the total roadway mileage, accurately quantifying the safety issues of these roadways can considerably contribute to more robust safety analysis and effective decision making. The next section describes the formulation and hierarchical representation of the NB models that were examined in this study.

2. Methodology

The most common methods that are used by researchers to develop SPFs are the Poisson and Poisson-gamma regression models (Lord et al., 2005). The Poisson-gamma mixture or NB distribution is the generalization of the Poisson distribution, which eases the assumption that the mean and variance are equal by introducing the dispersion parameter to the model. As mentioned in the previous section, both the variance structure and the dispersion parameter can be formulated in different ways leading to the various NB parameterizations. The following subsection discusses the various NB formulations derived from different variance structures. Then, the NB-L model is presented, which introduces more flexibility to the traditional NB model. The last subsection presents different functional forms of the dispersion structure to better capture the variation in dispersion parameter across the road segments.

2.1. NB-2

The NB-2 is the most common form of the NB models. The hierarchical representation of the NB-2 is described as follows (Heydari et al., 2014):

$$\begin{aligned} &Y_i \sim \text{Poisson}(\lambda_i) \\ &\lambda_i = \mu_i r_i \\ &r_i \sim \text{Gamma}(\phi, \phi) \end{aligned} \tag{1}$$

where ϕ is the inverse dispersion parameter (i.e., $\alpha = \frac{1}{\phi}$ is the dispersion parameter), and $\mu_i = exp(\beta_i X_i)$ is the mean response crash frequency which is an exponential function of roadway characteristics. As seen, the NB model allows for inter-observation heterogeneity by multiplying a gamma distributed error term, r_i, to the mean function. After integrating the prior out of the Poisson-gamma joint distribution, we obtain the following probability density function with the mean, and variance functions as follows:

$$P(Y|\phi,\mu) = \frac{\Gamma(\phi + y_i)}{\Gamma(\phi)y_i}\left(\frac{\phi}{\phi+\mu_i}\right)^{\phi}\left(\frac{\mu_i}{\mu_i+\phi}\right)^{y_i} \tag{2}$$

$$E(y_i) = \mu_i \tag{3}$$

$$\text{Var}(y_i) = \mu_i + \frac{\mu_i^2}{\phi} \tag{4}$$

As described in Eq. (4), the NB-2 model assumes that there is a quadratic association between the mean and the variance through the inverse dispersion parameter.

2.2. NB-1

The other commonly used formulation of the NB model is shown in Eq. (5). The NB-1 model assumes that there is a constant ratio linking the mean and the variance of the crash frequencies. This could be achieved by replacing the inverse dispersion parameter, ϕ, with $\phi\mu_i$ in Eq. (4). The NB-2 model assumes that there is only one fixed dispersion parameter in the entire dataset, while the NB-1 adjusts the dispersion parameter for each site, individually. This adjustment leads to a different parameterization of the NB model which also preserves the conditional mean. The probability density function, mean, and variance of the NB-1 model can be written as follows (Greene, 2008):

$$P(Y|\phi,\mu) = \frac{\Gamma(\phi\mu_i + y_i)}{\Gamma(\phi\mu_i)y_i}\left(\frac{\phi\mu_i}{\phi\mu_i+\mu_i}\right)^{\phi\mu_i}\left(\frac{\mu_i}{\mu_i+\phi\mu_i}\right)^{y_i} \tag{5}$$

$$E(y_i) = \mu_i \tag{6}$$

$$\text{Var}(y_i) = \mu_i + \frac{\mu_i^2}{\phi\mu_i} = \mu_i + \frac{\mu_i}{\phi} \tag{7}$$

2.3. NB-P

A more general type of the NB model is the NB-P model, which does not constrain the mean-variance relationship. As seen in Eq. (10), the exponent of the mean, μ_i, in the mean-variance relationship can take any value. Similar to the NB-1 model, this formulation also makes an adjustment to the dispersion parameter of each site while maintaining the conditional mean. The probability density function, mean and

[1] The NB-L has been proposed as an alternative to the application of zero-inflated (ZI) models for handling datasets with a large percentage of zero responses. The NB-L model offers a single mean function that is never equal zero, which is not the case for the ZI model. This and other limitations have not only been documented in highway safety (Lord et al., 2005, 2007), but in various other fields, such as environmental science, substance abuse, criminology and social sciences. Additional discussion can be found in Chapter 3 of Lord et al. (2021).

variance of the NB-P distribution can be derived as follows (Greene, 2008):

$$P(Y|\phi,\mu)=\frac{\Gamma(\phi\mu_i^{2-p}+y_i)}{\Gamma(\phi\mu_i^{2-p})y_i}\left(\frac{\phi\mu_i^{2-p}}{\phi\mu_i^{2-p}+\mu_i}\right)^{\phi\mu_i^{2-p}}\left(\frac{\mu_i}{\mu_i+\phi\mu_i^{2-p}}\right)^{y_i} \tag{8}$$

$$E(y_i)=\mu_i \tag{9}$$

$$\mathrm{Var}(y_i)=\mu_i+\frac{\mu_i^2}{\phi\mu_i^{2-p}}=\mu_i+\frac{\mu_i^p}{\phi} \tag{10}$$

2.4. NB-L

To deal with the unique and problematic characteristics of crash data such as excess zeros and having a long heavy tail, extensions of the NB model have been proposed, which offer a more flexible structure to the original model in order to deal with problematic datasets. In this regard, Zamani and Ismail (2010) proposed the use of the mixture of NB and Lindley distribution to analyze a highly dispersed dataset characterized by a large number of zeros and a heavy tail. This model, also referred to as a multi-parameter model (Lord and Geedipally, 2018), under a hierarchical Bayesian framework can be described as follows (Geedipally et al., 2012):

$$\begin{aligned}&P(Y=y,\mu_i,\phi|\epsilon)=\mathrm{NB}(y;\phi,\epsilon\mu_i)\\&\epsilon\sim\mathrm{Lindley}(\theta)\end{aligned} \tag{11}$$

where θ is the Lindley distribution parameter. The Lindley distribution is a mixture of the exponential and gamma distribution (Zamani and Ismail, 2010). The probability density function and the mean structure of the Lindley distribution can be written as follows:

$$P(E=\epsilon|\theta)\sim\frac{\theta^2}{\theta+1}(1+\epsilon)e^{-\theta x};\quad \theta>0 \tag{12}$$

$$E(\epsilon)=\frac{\theta+2}{\theta(\theta+1)} \tag{13}$$

The NB-L formulation then could be derived by integrating the Lindley prior out of the NB and Lindley joint distribution:

$$P(Y=y,\mu,\phi,\theta)=\int\mathrm{NB}(y;\phi,\epsilon\mu)\mathrm{Lindley}(\epsilon,\theta)d\epsilon \tag{14}$$

The conditional mean and variance of the NB-L distribution then can be given as:

$$E(y_i)=\mu_iE(\epsilon)=\mu_i\frac{\theta+2}{\theta(\theta+1)} \tag{15}$$

$$\mathrm{Var}(y_i)=\mu_i\frac{\theta+2}{\theta(\theta+1)}+\mu_i^2\left(\frac{2(\theta+3)}{\theta^2(\theta+1)}\right)\left(\frac{1+\phi}{\phi}\right)-\left(\mu_i\frac{\theta+2}{\theta(\theta+1)}\right)^2 \tag{16}$$

As seen in Eq. (16), despite the NB-1, NB-2, and NB-P models (from now on referred to as traditional NB models), in which the variation is only explained by the dispersion parameter, part of the variability in the NB-L model is captured by the mixed Lindley distribution. This could offer even more flexibility to the model to capture further variations in the dataset. Similar to the traditional NB models, the NB-L can also be formulated with different variance structures. The NB1-L, NB2-L, and NBP-L are the counterparts of the traditional NB models, which take advantage of two sources of variation, the dispersion parameter, and the mixed Lindley distribution. Similar to the NB-P, the NBP-L is the most general formulation of the NB-L models. Given $k_i=\phi\mu_i^{2-p}$, the hierarchical representation of the NBP-L model could be formulated as follows:

$$\begin{aligned}&P(Y=y,\mu_i,k_i|\epsilon)=\mathrm{NB}(y;k_i,\epsilon\mu_i)\\&\epsilon\sim\mathrm{Lindley}(\theta)\end{aligned} \tag{17}$$

2.5. Dispersion parameter functional form

So far, six NB models with different parameterizations and different variance structures have been discussed. Each functional form of the NB could be formulated with a fixed or varying dispersion parameter. In this study, four different parameterizations of the inverse dispersion parameter (ϕ) were evaluated. These functional forms were selected according to the best formulations proposed by Geedipally et al. (2009). In addition, Cafiso et al. (2010) mentioned that in shorter segments, variability of the dispersion parameter matters more. Also NFAS roads are basically characterized by short segment lengths. Consequently, this study ensured that segment length is included in all the functional forms selected from the ones proposed by Geedipally et al. (2009). For all the six models, a fixed, and the following functional forms of the inverse dispersion parameter, ϕ, were modeled and examined:

$$\phi=e^{\eta_0}*\mathrm{AADT}_i^{\eta_1}*L_i^{\eta_2} \tag{18}$$

$$\phi=e^{\eta_0}*\mathrm{AADT}_i^{\eta_1}*L_i \tag{19}$$

$$\phi=e^{\eta_0}*L_i^{\eta_2} \tag{20}$$

$$\phi=e^{\eta_0}*L_i \tag{21}$$

where the η's are the parameters needed to be estimated and L is the segment length.

It should be pointed out that the varying dispersion function may not be needed as the number of parameters used with the mean function increases. Mitra and Washington (2007) indicated that as the mean function gets better defined, the dispersion function becomes less structured or may even become fixed for well-defined mean functions. However, recent research on this topic by Zou et al. (2014) indicates that the varying dispersion function may be data dependent rather than dependent on the mean function. Even with a well-defined mean function, the variance was still structured and dependent on the covariates included in the model according to the dataset used in their study.

2.6. Parameter estimation

To generate valid posterior inferences, a full Bayesian approach was utilized. The FB method can incorporate all the information and prior knowledge into a single hierarchical model and yield robust estimates even when limited data is available. Since the Lindley distribution is not a standard distribution to draw samples from, a simpler formulation of the NB-L was used. According to this formulation, the Lindley distribution could be parameterized as a sum of two gamma distributions with the mixture components following the Bernoulli distribution. The equivalent hierarchical representation of the Lindley distribution can be shown as follows (Zamani and Ismail, 2010):

$$\begin{aligned}&\epsilon\sim\mathrm{Gamma}(1+z,\theta)\\&z\sim\mathrm{Bernoulli}\left(\frac{1}{1+\theta}\right)\end{aligned} \tag{22}$$

Moreover, the FB method requires to specify the prior distribution on all the unknown hyper-parameters to combine the data likelihood with the past evidence. This study assumed a non-informative normal prior for the regression coefficients of the mean and the varying dispersion function, and a gamma prior on the fixed dispersion parameter, $1/\phi$. Furthermore, priors should be chosen to preserve the identifiability of the model. In the NB-L model, as seen in Eqs. (15) and (16), both conditional mean and variance are adjusted by the Lindley parameter, θ. As a result, expectation of ϵ should be equal to one to preserve the

Table 1
Summary statistics of Virginia data.

Roadway functional class	Variable	Min	Max	Average (Std. dev.)	Skewness
All NFAS roads ($N = 2598$)	Number of crashes	0	33	2.25 (3.22)	2.83
	AADT over 5 years (vpd)	8	2347	589 (434)	
	Segment length (mi)	0.1	5.73	1.37 (1.20)	
Rural minor collector ($N = 1778$)	Number of Crashes	0	33	2.75 (3.55)	2.59
	AADT over 5 years (vpd)	21	2346	584 (395)	
	Segment length (miles)	0.1	5.73	1.59 (1.20)	
Rural local $N = 455$	Number of crashes	0	17	1.32(2.16)	2.85
	AADT over 5 years (vpd)	8	2093	379 (333)	
	Segment length (miles)	0.1	5.7	1.28 (1.18)	
Urban local ($N = 365$)	Number of crashes	0	14	0.98 (1.76)	3.31
	AADT over 5 years (vpd)	9	2347	874 (553)	
	Segment length (miles)	0.1	4.52	0.4 (0.48)	

conditional mean and ensure the identifiability issue. As suggested in Shaon et al. (2018) and Geedipally et al. (2012), a Beta $\left(\frac{N}{3}, \frac{N}{2}\right)$ would be a good prior distribution on $(1/1+\theta)$ since it guarantees $E(\epsilon) = 1$, and is also relevant to the likelihood function through the parameter N (number of observations).

Having specified the hierarchical joint model, we can draw random samples from the posterior distribution of the unknown parameters using the MCMC method. Depending on the availability of the full conditional distribution of the unknown parameter given the other parameters, Gibbs sampling method or otherwise, Metropolis-Hasting algorithm can be used to draw random samples from the posterior distribution. In this study, an open-source R package, called "rjags", was used to conduct MCMC analysis (Plummer et al., 2016). All the proposed models were implemented in the Bayesian framework. A total of three Markov chains, each containing 50,000 iterations, were run to make sure the convergence is achieved. The first 10,000 samples of each chain were considered as burn-in samples, and the remaining samples were used to estimate the unknown coefficients. Also, to mitigate the possible sample auto-correlation, out of three successive samples, only one sample was stored for estimation.

3. Data description

Virginia roadway information, traffic volume, and crash data from the Virginia Department of Transportation (VDOT) were gathered, processed, and integrated in order to develop SPFs for NFAS roads. Roadway inventory attributes such as lane width, shoulder width, number of lanes, etc., were, unfortunately, missing for a considerable number of segments. It is important to note that the data used in this study are collected from NFAS roadways which are mostly known as local roadways with low volume (2000 vpd or lower). Low-volume roadway inventory data are not usually well maintained and there are many missing geometric data such as horizontal and vertical curvature, shoulder width, etc. The other possible covariates available in the dataset were the percentage of trucks, and the percentage of buses in the roadways, which turned to be insignificant in all the regression models. The only reliable and available variables were segment length and AADT, which match with the basic variables used by the HSM. So, this study used flow-only models to make SPFs similar to the ones found in the HSM. Out of 92,834 reported crashes and 117,863 roadway information collected from 2014 to 2018 (latest dataset available), 3708 NFAS roadways, and corresponding 14,212 crashes were identified. Finally, after excluding the missing records, outliers, roadways with low-quality AADT counts (count estimates labeled as poor quality by the VDOT), and intersection related crashes, the final database, including a five-year period information on 2598 segments and 5856 crashes was obtained. Nearly 37% of the roadways did not experience any crash during the five years. Table 1 present the summary statistics of the input data that was used to develop SPFs. Descriptive statistics are summarized for both categorized data (based on the roadway functional classes), and all NFAS roads.

4. Modeling results

This section describes the details of the SPF modeling results. In total, six count models, including the NB-1, NB-2, NB-P, NB1-L, NB2-L, and NBP-L, each with five different dispersion structures, were developed and run. Segment length and AADT were included in the SPF models as the possible covariates. Segment length was considered as a separate covariate rather than an offset since its estimate was statistically different from one. It should be pointed out that, even though this study only used segment length and AADT, the omitted-variable bias is not critical in this study since the models were compared using the same dataset and functional form (i.e., the link between the dependent and independent variables). Also, to make the MCMC process faster and overcome the poor convergence resulting from the multicollinearity issue (Shaon et al., 2018), the standardized covariates were input for estimation and then transformed back to the original scale.

Tables 2–4 summarize the estimation results for each NB model with fixed, AADT and length dependent, and length-only dependent dispersion structure, respectively. The first and second part of each table provides estimates for the mean function coefficients, βs, and the dispersion function coefficients, ηs, respectively. For the models associated with a fixed dispersion structure, the inverse dispersion parameter is also reported only for those models that follow the original NB distribution structure without any adjustment to the dispersion structure (i.e., NB2 and NB2-L). The last part demonstrates the performance evaluation metrics for evaluation and comparison purposes.

This study used the Bayesian counterpart of the confidence interval, credible interval, to test the significance of the parameters. The coefficients, which their highest posterior density credible interval (HPD credible interval) included zero at 5% level, were underlined in Tables 2–4.

In regards to the mean function parameters, both AADT and segment length had a significant positive influence on the crash frequency, which confirms the previous findings regarding SPFs for low-volume roadways (Das et al., 2019; Cook, 2010; Stamatiadis et al., 1999; Dell'Acqua and Russo, 2011; Zegeer et al., 1994). As opposed to the coefficients of the mean function, $\beta' s$, some coefficients of the dispersion function, $\eta' s$, were neither significant nor similar across the modeling approaches. As illustrated in Table 3, the estimates for the intercept and AADT

Table 2
Model estimation results (fixed dispersion structure).

Variable	NB-1	NB-2	NB-P	NB1-L	NB2-L	NBP-L
Intercept (β_0)	−4.07 (0.18)	−4.41 (0.20)	−4.41 (0.20)	−4.46 (0.26)	−4.47 (0.26)	−4.46 (0.26)
Ln(AADT) (β_1)	0.63 (0.03)	0.65 (0.03)	0.65 (0.03)	0.64 (0.09)	0.65 (0.09)	0.64 (0.09)
Length (β_2)	0.56 (0.01)	0.65 (0.01)	0.65 (0.02)	0.68 (0.02)	0.68 (0.02)	0.68 (0.02)
ϕ	–	3.12 (0.25)	–	–	17.54 (3.10)	–
θ	–	–	–	1.41 (0.06)	1.41 (0.06)	1.41 (0.06)
P	–	–	1.94 (0.11)	–	–	0.57 (0.38)
WAIC	8296	8232	8234	7612	7640	7619
LOO	8296	8232	8234	8112	8174	8132
MAD	1.25	1.25	1.26	1.15	1.16	1.16
MASE	0.55	0.56	0.56	0.57	0.57	0.56
Log-likelihood	−4096	−4064	−4064	−3290	−3312	−3287

Table 3
Model estimation results (AADT and length dependent dispersion structure).

Variable	NB-1	NB-2	NB-P	NB1-L	NB2-L	NBP-L
Functional form (1): $\phi_i = e^{\eta_0} AADT_i^{\eta_1} L_i^{\eta_2}$						
Intercept (β_0)	−4.51 (0.19)	−4.51 (0.20)	−4.56 (0.20)	−4.41 (0.27)	−4.39 (0.27)	−4.50 (0.32)
Ln(AADT) (β_1)	0.72 (0.04)	0.70 (0.04)	0.70 (0.03)	0.64 (0.09)	0.65 (0.09)	0.69 (0.09)
Length (β_2)	0.51 (0.01)	0.55 (0.02)	0.60 (0.02)	0.66 (0.02)	0.63 (0.02)	0.58 (0.02)
θ	–	–	–	1.39 (0.05)	1.38 (0.05)	1.39 (0.06)
P	–	–	3.48 (0.20)	–	–	3.91 (0.09)
η_0	4.19 (0.71)	0.20 (0.68)	− 4.93 (0.96)	6.33 (4.57)	2.72 (3.07)	−2.61 (2.31)
η_1	−0.67 (0.10)	0.07 (0.10)	1.08 (0.17)	−0.02 (0.67)	0.40 (0.45)	1.22 (0.33)
η_2	0.41 (0.12)	0.85 (0.10)	1.92 (0.17)	3.65 (1.02)	3.60 (0.64)	3.65 (0.4)
WAIC	8150	8089	8043	7485	7497	7564
LOO	8150	8089	8043	7978	8018	7981
MAD	1.29	1.28	1.35	1.15	1.16	1.29
MASE	0.55	0.55	0.56	0.56	0.55	0.55
Log-likelihood	−4070	−4040	−4017	−3284	−3310	−3373
Functional form (2): $\phi_i = e^{\eta_0} AADT_i^{\eta_1} L_i$						
Intercept (β_0)	−4.54 (0.18)	−4.52 (0.19)	−4.52 (0.20)	−4.48 (0.26)	−4.51 (0.27)	−4.56 (0.28)
Ln(AADT) (β_1)	0.74 (0.03)	0.71 (0.03)	0.69 (0.03)	0.65 (0.09)	0.65 (0.09)	0.65 (0.09)
Length (β_2)	0.47 (0.01)	0.54 (0.01)	0.60 (0.02)	0.68 (0.02)	0.68 (0.02)	0.77 (0.03)
θ	–	–	–	1.39 (0.05)	1.41 (0.06)	1.42 (0.07)
P	–	–	2.56 (0.10)	–	–	3.93 (0.06)
η_0	3.87 (0.73)	−0.05 (0.64)	−1.21 (0.66)	6.48 (11.24)	8.36 (12.67)	−5.34 (2.06)
η_1	−0.62 (0.10)	0.11 (0.09)	0.38 (0.10)	−0.08 (1.75)	−0.29 (1.97)	1.77 (0.33)
η_2	–	–	–	–	–	–
WAIC	8170	8087	8066	7482	7489	7672
LOO	8170	8087	8067	7981	8030	8055
MAD	1.38	1.30	1.30	1.15	1.16	1.24
MASE	0.57	0.55	0.56	0.57	0.57	0.64
Log-likelihood	−4081	−4041	−4030	−3274	−3281	−3385

coefficient of the dispersion function, η_0 and η_1, were not statistically significant at 5% significance level when using the NB-L models with AADT and length dependent dispersion structure. On the other hand, as shown in Tables 2–4, the magnitude of the coefficients of the dispersion function vary markedly across the models. The differences in significance and magnitude of ηs could be partially attributed to the different variance structures in the models. The NB-1, NB-2, and NB-P models, each has a specific structure to capture the variation. Introducing the Lindley distribution to the NB models makes the variance structure even more complex since it provides the model with additional complexity and hence, more flexibility. Therefore, the source of variation in each model is different, which makes it difficult to compare the dispersion coefficients, individually. Moreover, as seen in Tables 3 and 4, the sign of the length coefficient is positive in the dispersion functions, indicating that the dispersion parameter, $1/\phi$, and therefore, the unobserved variation decreases as roadway length increases. These findings are in line with Hauer (2001) and Cafiso et al. (2010), indicating that shorter segments have higher crash frequency variances.

Models were evaluated based on a combination of different goodness of fit measures. Two fully Bayesian metrics, widely applicable information criteria (WAIC) and leave-one-out cross-validation (LOO), along with other commonly used metrics, were used for performance evaluation and comparison purposes. The superiority of the NB-L models over the traditional NB models is demonstrated through all the GOF metrics.

Among the traditional NB models, models with the less restricted mean-variance structure, i.e., NB-2 and NB-P showed better performance. However, the NB-L models performed better when formulated with less flexible mean-variance relationships, i.e., NB1-L and NB2-L.

All the NB parameterizations with varying dispersion parameters, regardless of the dispersion structure, showed superior fit compared to the NB parameterizations with fixed dispersion parameters. Moreover, as indicated in Tables 2–4, the performance measures vary when employing different dispersion functions. Given the results, it could be interpreted that the NB-L models perform better if the length-only

Table 4
Model estimation results (length-only dependent dispersion structure).

Variable	NB-1	NB-2	NB-P	NB1-L	NB2-L	NBP-L
Functional form (3): $\phi_i = L_i^{\eta_2} e^{\eta_0}$						
Intercept (β_0)	−4.09 (0.18)	−4.54 (0.20)	−4.52 (0.21)	−4.39 (0.26)	−4.47 (0.27)	−4.72 (0.30)
Ln(AADT) (β_1)	0.65 (0.04)	0.71 (0.04)	0.70 (0.04)	0.64 (0.09)	0.66 (0.09)	0.72 (0.09)
Length (β_2)	0.51 (0.01)	0.55 (0.01)	0.57 (0.02)	0.65 (0.02)	0.63 (0.02)	0.60 (0.02)
θ	–	–	–	1.39 (0.05)	1.39 (0.05)	1.39 (0.05)
P	–	–	2.41 (0.14)	–	–	0.13 (0.12)
η_0	−0.20 (0.07)	0.70 (0.07)	1.15 (0.17)	6.11 (1.34)	5.24 (1.06)	6.22 (0.63)
η_1	–	–	–	–	–	–
η_2	0.50 (0.12)	0.83 (0.10)	0.99 (0.11)	3.75 (0.90)	3.51 (0.62)	3.80 (0.17)
WAIC	8188	8087	8081	7482	7497	7535
LOO	8188	8087	8081	7971	8015	7963
MAD	1.24	1.29	1.29	1.15	1.18	1.27
MASE	0.55	0.55	0.55	0.56	0.55	0.55
Log-likelihood	−4090	−4040	−4037	−3285	−3313	−3345
Functional form (4): $\phi_i = L_i e^{\eta_0}$						
Intercept (β_0)	−4.14 (0.17)	−4.55 (0.20)	−4.52 (0.21)	−4.48 (0.26)	−4.45 (0.26)	−4.74 (0.31)
Ln(AADT) (β_1)	0.68 (0.03)	0.72 (0.03)	0.70 (0.03)	0.64 (0.09)	0.64 (0.08)	0.68 (0.09)
Length (β_2)	0.48 (0.01)	0.53 (0.01)	0.57 (0.02)	0.68 (0.02)	0.68 (0.02)	0.72 (0.03)
θ	–	–	–	1.38 (0.05)	1.39 (0.06)	1.39 (0.06)
P	–	–	2.41 (0.12)	–	–	3.44 (0.97)
η_0	−0.21 (0.08)	0.69 (0.06)	1.14 (0.16)	5.37 (1.48)	5.88 (1.43)	6.12 (0.81)
η_1	–	–	–	–	–	–
η_2	–	–	–	–	–	–
WAIC	8201	8087	8079	7482	7481	7625
LOO	8201	8087	8079	7981	8024	8075
MAD	1.32	1.30	1.30	1.13	1.15	1.22
MASE	0.57	0.55	0.55	0.56	0.57	0.60
Log-likelihood	−4097	−4041	−4036	−3275	−3277	−3348

dependent dispersion functions are used, whereas the traditional NB models favor the AADT and length dependent dispersion structures more.

Based on the combination of GOF criteria and also the significance of the model coefficients, NB1-L and NB2-L with length-only dispersion structure (dispersion structure (3) and (4)) ranked as the best models. As Hauer and Bamfo (1997) suggested, the cumulative residual (CURE) plot was utilized to assess the model performance by directly analyzing the residuals. The CURE plot of a well-fitted SPF should not include an upward or downward trend or a noticeable periodicity. It should

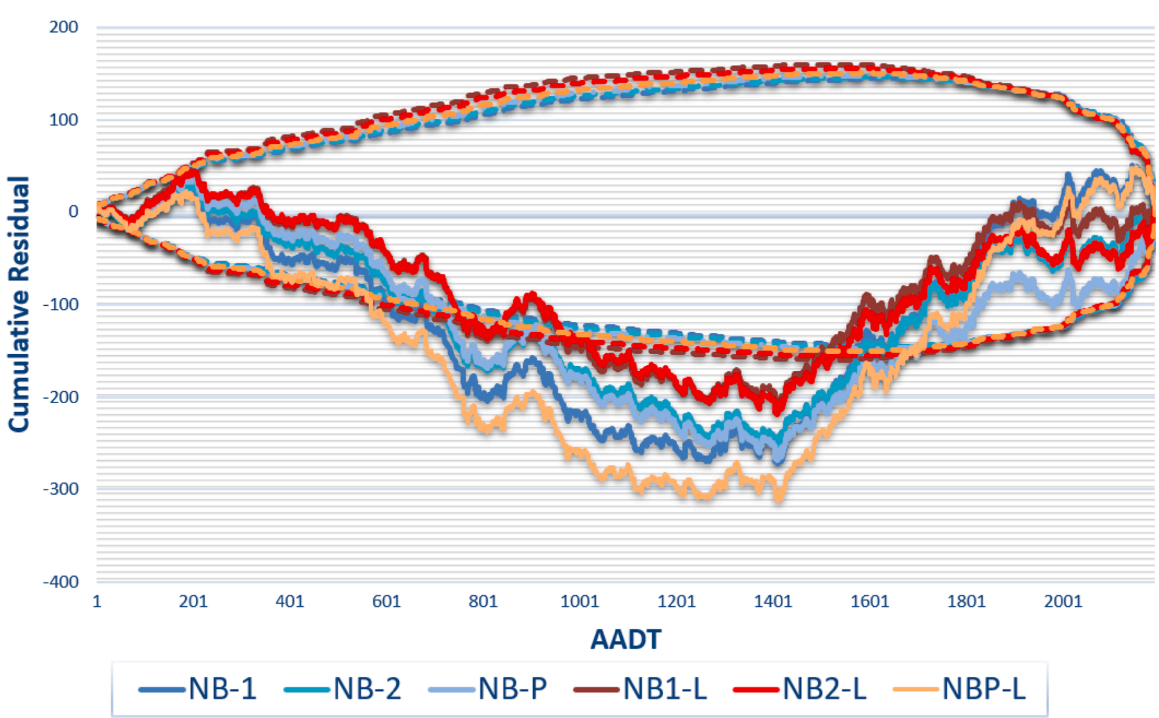

Fig. 1. CURE plots for AADT variable (dotted lines represent ±1.96 std. dev.).

fluctuate around zero while being in the boundary of two standard deviations (confidence interval). Adjusted CURE plot for the NB1-L model, as well as other models with the same dispersion function, are depicted in Fig. 1. All the CURE plots are adjusted to end at zero to make them comparable. In comparison to the traditional NB models, there are less sudden falls and rises in plots of the NB-L models. Also, the CURE plots of the NB-L models seem to be within the confidence intervals more often than their traditional counterparts. Aside from the CURE plots, the unadjusted cumulative residual itself could be a valid indicator of the predictive ability of the model. The last values of the cumulative residuals (sum of all the residuals) are equal to − 235, − 298, − 339, − 3, − 26, − 206 for the NB-1, NB-2, NB-P, NB1-L, NB2-L, and NBP-L, respectively. To put it differently, cumulative residual plot of the NB1-L and NB2-L models converges to zero naturally; whereas, that of the other NB parameterizations are far away from zero. These findings are also in line with Shirazi et al. (2017) findings that the maximum deviation of the NB-L models are smaller than the NB models.

5. Discussion

This study aimed to develop SPFs for NFAS roads. As these roads commonly have low volumes and short lengths, their crash statistics could be characterized by specific features that make it challenging to accurately quantify their safety performances. As there are not many studies done on these particular roadway classes, there is no commonly agreed predictive model that performs adequately. This study compared the application of different count models in three different levels; (1) between the traditional NB model and more flexible, zero-favored NB models (i.e., NB-L), (2) between different forms of the mean-variance association through the dispersion parameter, and (3) between different functional forms of the dispersion parameter.

As the results showed, regardless of the dispersion structure, all the GOF measures indicated that the NB-L models provide a better statistical fit. In the dataset analyzed in this study, 37% and 20% of the roadways had recorded zero and one crashes for a five year period, respectively. Also, dividing the crash counts by the number of years that data was collected, we observed that around 78% of the segments have crash frequencies below 0.6 crash per year. This information confirms Geedipally et al. (2012) findings that the NB-L models offer superior performance for datasets characterized by a large number of zeros. Also, the results are in line with Shirazi et al. (2017), which showed that for crash data with skewness higher than 1.92 (2.83 in this study) the NB-L model performs better.

For both groups of traditional NB and NB-L models, different mean-variance structures were examined as well. Similar to the results of Wang et al. (2019), the NB-P and NB-2 favored the NB-1 model, regardless of the dispersion structure. It was not unexpected since the NB-1 model introduces a less flexible variance structure (linear mean-variance relationship) to the model. The NB-P model performed slightly better than NB-2 due to the more flexibility through the parameter P. However, the improvements were negligible since the estimated P parameter in the NB-P model was close to 2 in the most cases, which made the NB-P model similar to the NB-2 model (e.g., NB-P with functional forms (2), (3), and (4)). A different pattern was observed in the NB-L models. Compared to the NBP-L, the NB1-L and NB2-L appeared to be better models in terms of almost all the performance measures. These findings showed that even though the NBP-L model offers a more flexible variance structure and more ability to fit to complex data, it might not always be the best choice. In other words, the choice of the variance structure is not only dependent upon the dataset being analyzed, but also depends on the formulation of the NB model. These results suggest that, even though the mixed NB-L distribution provides the model with another source of variation, the variance structure still matters and should be considered in the SPF development process.

Finally, the effect of using different dispersion structures was examined. All the models with varying dispersion parameter outperformed the models with fixed dispersion parameter. These findings support the results of Miranda-Moreno et al. (2005) and Lord and Park (2008) that the varying dispersion parameter can better capture the structured variations observed in the dataset. The traditional NB models showed better performance when formulated by the AADT and length dependent dispersion functions (e.g., dispersion functions (1) and (2)); however, the NB-L models favored the length-only dependent dispersion functions more. Having excluded models with insignificant coefficients, dispersion function (1), dependent upon both length and AADT, and dispersion function (3), dependent on length only, provided better statistical fit in the traditional NB models and the NB-L models, respectively. According to the results, the following conclusions could be obtained. First, the functional form of the dispersion parameter can significantly affect the model performance. This has been documented in the literature that applying different dispersion functions to the same model lead to different model performances (Geedipally et al., 2009; Cafiso et al., 2010). However, the improvements in the traditional NB models were more significant than the NB-L models. It was expected since the NB-L models are typically characterized by smaller over-dispersion parameter and hence, are less sensitive to the choice of the dispersion structure. Second, each NB parameterization calls for its own appropriate dispersion function. This means that if, for instance, a NB-2 model is enhanced by using a specific dispersion function, a NB-L model is not essentially improved by using the same dispersion function. This study also found that, within the traditional NB models, NB-1 shows more sensitivity to the dispersion function choice compared to the NB-2 and NB-P. This is probably due to the less flexible variance structure of the NB-1 compared to the other two models. So, providing it with an appropriate dispersion function could highly affect its ability to account for the data heterogeneity. These results support Wang et al. (2019) findings that more flexible variance structures in the NB models (e.g., NB2 and NB-P) can even capture the variation in the dispersion parameter. Therefore, these models are less sensitive to the dispersion structure. So, it can be concluded that researchers should choose the dispersion structure for each dataset (Geedipally et al., 2009), as well as for each parameterization of the crash-frequency model.

This study used a combination of metrics to evaluate and compare the models. Median absolute deviation (MAD) computes the average absolute difference between the observed and predicted values. Mean absolute scaled error (MASE) is a scale-free metric that normalizes the MAD by the average error. However, all these metrics are the measures of accuracy of the prediction. To consider the prediction accuracy and complexity of the model simultaneously, a cross-validation and an information criteria based method were used. These methods estimate the out of sample accuracy using within sample fits (Vehtari et al., 2017). Leave-one-out cross-validation assesses the predictive accuracy of the model by estimating the prediction error for the sample i without using it to train the model. However, it requires re-fitting the model N times (N is the sample size) to calculate the predictive accuracy. In the method proposed by Vehtari et al. (2017), they approximated the underlying process by using the sample draws from the full posterior distribution, $p(\theta|y)$, which is the typical outcome of any Bayesian analysis. Consequently, the leave-on-out cross-validation could be approximated by fitting the model once.

Also, WAIC appeared to be a more robust metric compared to the Deviance Information Criteria (DIC) in the Bayesian framework (Watanabe and Opper, 2010). As Geedipally et al. (2014) mentioned, regardless of the similarity of the estimates among the NB models, different parameterization of the model, especially different definitions of the likelihood functions in the hierarchical models, lead to different DIC values. Therefore, considering a fully Bayesian alternative metric seems essential, especially when both likelihood function and dispersion structure vary across the models. WAIC makes use of the entire posterior distribution and also is invariant to re-parameterization of the model (Vehtari et al., 2017). In general, all the information criteria approaches interpret the effective number of parameters as a measure of the model complexity. Although both DIC and WAIC use a variance-based computation to estimate the effective number of parameters (Gelman et al., 2014), WAIC produces more reliable results as it calculates the variance for each point separately (Vehtari et al., 2017). Both LOO and WAIC showed superiority over the traditional metrics such as AIC and DIC; however, they are computationally intensive and costly. In this study, we extracted $N \times S$ log-likelihood matrix, (where S is the number of simulation, and N is the number of observations) and then used the R package, called "loo" and set up by Vehtari et al. (2018), to compute LOO and WAIC.

Finally, after excluding models with insignificant estimates, the NB1-L and NB2-L models with length-only dependent dispersion function outranked the others in terms of almost all the metrics. However, within the NB-L models, there was some discrepancy. NB2-L showed better performance in terms of WAIC, and MASE, whereas NB1-L outperformed the others in terms of LOO and MAD. These findings justify the use of different metrics for evaluation and comparison purposes.

6. Summary and conclusion

NFAS roads comprise a significant part of the roadway network; however, not much research efforts have gone toward accurately evaluating the safety of these roadways. As these roadways are characterized by different features as major facilities (e.g., highways, arterials, etc.), there is a need to improve the currently used models to accurately quantify the safety issues associated with them. The primary objective of this study was to evaluate the application of different parameterizations of the negative binomial models in the SPF development of NFAS roadways. In the first level, both traditional NB and zero-favored NB (i.e., NB-L model) were considered to model the crash counts. Then, for each model, three different variance structures were considered, leading to six different NB parameterizations. Finally, for each of the six crash-frequency models, five different dispersion structures were employed. Using crash data, roadway inventory, and traffic volume data from 2014-2018 in Virginia, this study showed that the NB-L models fit better than the traditional NB models. Within the NB-L models also the NB1-L and NB2-L models showed better fits. This study also found that the variance and dispersion structure choices are highly dependent upon the NB parameterization. As opposed to the traditional NB models, the NB-L models performed better when using the length-only dependent dispersion function. All the models were evaluated using various GOF measures, including two recently documented fully Bayesian metrics, WAIC and LOO. This study provides additional insight into the choice of the predictive model to evaluate the safety performance of NFAS roadways. The advanced models developed in this study could contribute to the betterment of safety evaluation of these roadways and any other crash dataset that requires a more flexible modeling structure. This study showed that a reasonably chosen variance and dispersion structure can effectively enhance the count models (even the more advanced models which have been proven that outperform the traditional models) leading to better model performances, more accurate estimates, and hence more reliable decision making. Using a more detailed dataset and the inclusion of other traffic-related variables could further enhance the model performance. This study did not separate different categories of NFAS roadways. Further work needs to be done to develop SPF for each functional class individually. Also, separating the crash data by severity level (e.g., KABCO, KAB, etc.) and crash type would further improve the predictive accuracy of the model.

Authors' contribution

The authors confirm contribution to the paper as follows: study conception and design: Ali Khodadadi, Dominique Lord; data collection: Ali Khodadadi, Ioannis Tsapakis, and Yingfeng Li; analysis and interpretation of results: Ali Khodadadi, Ioannis Tsapakis, and Subasish Das; draft manuscript preparation: Ali Khodadadi, Ioannis Tsapakis, Dominique Lord, and Subasish Das. All authors reviewed the results and approved the final version of the manuscript.

Conflict of interest

The authors declare that they have no known competing financial interests or personal relationships that could have influenced the work reported in this paper.

Acknowledgements

The authors would like to thank In-Kyu (Q) Lim and Wenling Chen from the Virginia Department of Transportation (VDOT) who contributed time in processing the dataset. The authors are also pleased to thank Dr. Ali Shirazi at University of Maine for providing input on the analysis. The work was completed as part of SAFE-D University Transportation Center project *Use of Disruptive Technologies to Support Safety Analysis and Meet New Federal Requirements*. The study was partly funded by the A.P. and Florence Wiley Faculty Fellow.

References

Aguero-Valverde, J., Jovanis, P.P., 2009. Bayesian multivariate poisson lognormal models for crash severity modeling and site ranking. Transp. Res. Rec. 2136, 82–91.
Cafiso, S., Di Silvestro, G., Persaud, B., Begum, M.A., 2010. Revisiting variability of dispersion parameter of safety performance for two-lane rural roads. Transp. Res. Rec. 2148, 38–46.
Cameron, A.C., Trivedi, P.K., 2013. Regression Analysis of Count Data, Vol. 53. Cambridge University Press.
Cook, D.J., 2010. Secondary Low Volume Rural Road Safety: Segmentation, Crash Prediction, and Identification of High Crash Locations (Master's thesis). Iowa State University.
Das, S., Tsapakis, I., Datta, S., 2019. Safety performance functions of low-volume roadways. Transp. Res. Rec. 2673, 798–810.
Dell'Acqua, G., Russo, F., 2011. Safety performance functions for low-volume roads. Baltic J. Road Bridge Eng. 6, 225–234.
Farid, A., Abdel-Aty, M., Lee, J., Eluru, N., 2017. Application of bayesian informative priors to enhance the transferability of safety performance functions. J. Saf. Res. 62, 155–161.
Geedipally, S.R., Lord, D., Dhavala, S.S., 2012. The negative binomial-lindley generalized linear model: characteristics and application using crash data. Accid. Anal. Prev. 45, 258–265.
Geedipally, S.R., Lord, D., Dhavala, S.S., 2014. A caution about using deviance information criterion while modeling traffic crashes. Saf. Sci. 62, 495–498.
Geedipally, S.R., Lord, D., Park, B.J., 2009. Analyzing different parameterizations of the varying dispersion parameter as a function of segment length. Transp. Res. Rec. 2103, 108–118.
Gelman, A., Hwang, J., Vehtari, A., 2014. Understanding predictive information criteria for Bayesian models. Stat. Comput. 24, 997–1016.
Greene, W., 2008. Functional forms for the negative binomial model for count data. Econ. Lett. 99, 585–590.
Hauer, E., 2001. Overdispersion in modelling accidents on road sections and in empirical Bayes estimation. Accid. Anal. Prev. 33, 799–808.
Hauer, E., Bamfo, J., 1997. Two tools for finding what function links the dependent variable to the explanatory variables. In: Proceedings of the ICTCT 1997 Conference. Lund, Sweden.
Heydari, S., Miranda-Moreno, L.F., Lord, D., Fu, L., 2014. Bayesian methodology to estimate and update safety performance functions under limited data conditions: a sensitivity analysis. Accid. Anal. Prev. 64, 41–51.
HSM, 2010. Highway Safety Manual, 1st ed. AASHTO, Washington, DC.
Ismail, N., Zamani, H., 2013. Estimation of claim count data using negative binomial, generalized Poisson, zero-inflated negative binomial and zero-inflated generalized Poisson regression models. Casualty Actuarial Society E-Forum 1–18.
Lord, D., Geedipally, S.R., 2011. The negative binomial-Lindley distribution as a tool for analyzing crash data characterized by a large amount of zeros. Accid. Anal. Prev. 43, 1738–1742.
Lord, D., Geedipally, S.R., 2018. Safety prediction with datasets characterised with excess zero responses and long tails. Safe Mobility: Challenges, Methodology and Solutions.
Lord, D., Mannering, F., 2010. The statistical analysis of crash-frequency data: a review and assessment of methodological alternatives. Transp. Res. Part A: Policy Pract. 44, 291–305.
Lord, D., Miranda-Moreno, L.F., 2008. Effects of 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. Saf. Sci. 46, 751–770.
Lord, D., Park, B.J., Model, P.G., 2012. Negative binomial regression models and estimation methods. Probability Density and Likelihood Functions. Texas A&M University, Korea Transport Institute, pp. 1–15.

Lord, D., Park, P.Y.J., 2008. Investigating the effects of the fixed and varying dispersion parameters of Poisson-gamma models on empirical Bayes estimates. Accid. Anal. Prev. 40, 1441–1457.
Lord, D., Qin, X., Geedipally, S.R., 2021. Highway Safety Analytics and Modeling. Elsevier B.V., Amsterdam, The Netherlands.
Lord, D., Washington, S., Ivan, J.N., 2007. Further notes on the application of zero-inflated models in highway safety. Accid. Anal. Prev. 39, 53–57.
Lord, D., Washington, S.P., Ivan, J.N., 2005. Poisson, poisson-gamma and zero-inflated regression models of motor vehicle crashes: balancing statistical fit and theory. Accid. Anal. Prev. 37, 35–46.
Mehta, G., Lou, Y., 2013. Calibration and development of safety performance functions for alabama: two-lane, two-way rural roads and four-lane divided highways. Transp. Res. Rec. 2398, 75–82.
Meng, Y., Wu, L., Ma, C., Guo, X., Wang, X., 2020. A comparative analysis of intersection hotspot identification: fixed vs. varying dispersion parameters in negative binomial models. J. Transp. Saf. Secur. 1–18.
Miranda-Moreno, L.F., Fu, L., Saccomanno, F.F., Labbe, A., 2005. Alternative risk models for ranking locations for safety improvement. Transp. Res. Rec. 1908, 1–8.
Mitra, S., Washington, S., 2007. On the nature of over-dispersion in motor vehicle crash prediction models. Accid. Anal. Prev. 39, 459–468.
Pawlovich, M.D., Li, W., Carriquiry, A., Welch, T., 2006. Iowa's experience with road diet measures: use of Bayesian approach to assess impacts on crash frequencies and crash rates. Transp. Res. Rec. 1953, 163–171.
Pei, X., Wong, S., Sze, N.N., 2011. A joint-probability approach to crash prediction models. Accid. Anal. Prev. 43, 1160–1166.
Plummer, M., et al., 2016. rjags: Bayesian Graphical Models Using MCMC. R Package Version 4.
Shaon, M.R.R., Qin, X., Shirazi, M., Lord, D., Geedipally, S.R., 2018. Developing a random parameters negative binomial-lindley model to analyze highly over-dispersed crash count data. Anal. Methods Accid. Res. 18, 33–44.
Shirazi, M., Dhavala, S.S., Lord, D., Geedipally, S.R., 2017. A methodology to design heuristics for model selection based on the characteristics of data: Application to investigate when the negative binomial Lindley (nb-l) is preferred over the negative binomial (nb). Accid. Anal. Prev. 107, 186–194.
Stamatiadis, N., Jones, S., Aultman-Hall, L., 1999. Causal factors for accidents on southeastern low-volume rural roads. Transp. Res. Rec. 1652, 111–117.
Stapleton, S.Y., Gates, T.J., Avelar, R., Geedipally, S.R., Saedi, R., 2019. Safety performance functions for low-volume rural stop-controlled intersections. Transp. Res. Rec. 2673, 660–669.
Vehtari, A., Gelman, A., Gabry, J., 2017. Practical Bayesian model evaluation using leave-one-out cross-validation and waic. Stat. Comput. 27, 1413–1432.
Vehtari, A., Gelman, A., Gabry, J., 2018. loo: Efficient Leave-One-Out Cross-Validation and WAIC for Bayesian Models. R Package Version 2, p. 1003.
Wang, K., Zhao, S., Jackson, E., 2019. Functional forms of the negative binomial models in safety performance functions for rural two-lane intersections. Accid. Anal. Prev. 124, 193–201.
Watanabe, S., Opper, M., 2010. Asymptotic equivalence of Bayes cross validation and widely applicable information criterion in singular learning theory. J. Mach. Learn. Res. 11.
Zamani, H., Ismail, N., 2010. Negative binomial-Lindley distribution and its application. J. Math. Stat. 6, 4–9.
Zegeer, C.V., Stewart, R., Council, F., Neuman, T.R., 1994. Accident relationships of roadway width on low-volume roads. Transp. Res. Rec. 160.
Zou, Y., Zhang, Y., Lord, D., 2014. Analyzing different functional forms of the varying weight parameter for finite mixture of negative binomial regression models. Anal. Methods Accid. Res. 1, 39–52.