ELSEVIER

Accident Analysis and Prevention 40 (2008) 1123–1134

ACCIDENT ANALYSIS & PREVENTION

www.elsevier.com/locate/aap

Application of the Conway–Maxwell–Poisson generalized linear model for analyzing motor vehicle crashes

Dominique Lord *, Seth D. Guikema [1], Srinivas Reddy Geedipally [2]

Zachry Department of Civil Engineering, Texas A&M University, 3136 TAMU, College Station, TX 77843-3136, USA

Received 7 September 2007; received in revised form 30 November 2007; accepted 6 December 2007

Abstract

This paper documents the application of the Conway–Maxwell–Poisson (COM-Poisson) generalized linear model (GLM) for modeling motor vehicle crashes. The COM-Poisson distribution, originally developed in 1962, has recently been re-introduced by statisticians for analyzing count data subjected to over- and under-dispersion. This innovative distribution is an extension of the Poisson distribution. The objectives of this study were to evaluate the application of the COM-Poisson GLM for analyzing motor vehicle crashes and compare the results with the traditional negative binomial (NB) model. The comparison analysis was carried out using the most common functional forms employed by transportation safety analysts, which link crashes to the entering flows at intersections or on segments. To accomplish the objectives of the study, several NB and COM-Poisson GLMs were developed and compared using two datasets. The first dataset contained crash data collected at signalized four-legged intersections in Toronto, Ont. The second dataset included data collected for rural four-lane divided and undivided highways in Texas. Several methods were used to assess the statistical fit and predictive performance of the models. The results of this study show that COM-Poisson GLMs perform as well as NB models in terms of GOF statistics and predictive performance. Given the fact the COM-Poisson distribution can also handle under-dispersed data (while the NB distribution cannot or has difficulties converging), which have sometimes been observed in crash databases, the COM-Poisson GLM offers a better alternative over the NB model for modeling motor vehicle crashes, especially given the important limitations recently documented in the safety literature about the latter type of model.

Keywords: Statistical models; Conway–Maxwell–Poisson distribution; Negative binomial; Regression models; Bayesian models

1. Introduction

There has been considerable research conducted on the development of statistical models for analyzing crashes on highway facilities (Abbess et al., 1981; Hauer et al., 1988; Kulmala, 1995; Poch and Mannering, 1996; Lord, 2000; Miaou and Lord, 2003; Lord et al., 2005a; Miaou and Song, 2005; Xie et al., 2007). The most common probabilistic structure of the models used by transportation safety analysts for modeling motor vehicle crashes are the traditional Poisson and Poisson-gamma (or negative binomial or NB) distributions. Since crash data have often been shown to exhibit over-dispersion (see Lord et al., 2005b), meaning that the variance is greater than the mean, Poisson-gamma models are usually preferred over Poisson regression models. On occasion, crash data have sometimes shown characteristics of under-dispersion, especially in cases where the sample mean is very low (Oh et al., 2006). The Poisson-gamma distribution offers a simple way to accommodate the over-dispersion, since the final equation has a closed form and the mathematics to manipulate the relationship between the mean and the variance structures are relatively simple (Hauer, 1997).

Over the last few years, several studies have documented important limitations associated with Poisson and NB models. For instance, it has been shown that when the sample mean value becomes small, traditional methods used to assess the goodness-of-fit (GOF) of generalized linear models (GLMs) estimated using the maximum likelihood estimating (MLE) method (both for Poisson and NB) can be highly unreliable and provide a biased estimate of the fit (Maycock and Hall, 1984; Maher and Summersgill, 1996; Wood, 2002; Lord, 2006). Recent studies have also shown that the inverse dispersion parameter of NB models, $\phi = 1/\alpha$, can be significantly mis-estimated when the

* Corresponding author. Tel.: +1 979 458 3949; fax: +1 979 845 6481.
E-mail addresses: d-lord@tamu.edu (D. Lord), sguikema@civil.tamu.edu (S.D. Guikema), srinivas8@tamu.edu (S.R. Geedipally).
[1] Tel.: +1 979 458 3586; fax: +1 979 845 6554.
[2] Tel.: +1 979 862 3446; fax: +1 979 845 6481.

0001-4575/$ – see front matter © 2007 Elsevier Ltd. All rights reserved.
doi:10.1016/j.aap.2007.12.003

sample size becomes small and the sample mean value is low. This characteristic was observed both for MLE (Clark and Perry, 1989; Piegorsch, 1990; Lord, 2006; Lloyd-Smith, 2007) and full Bayesian (FB) NB models (Airoldi et al., 2006; Lord and Miranda-Moreno, 2008). A mis-estimated inverse dispersion parameter can affect analysis tools commonly used in highway safety analyses, such as the application of the empirical Bayes (EB) estimate for identifying hazardous locations (Hauer, 1997; Miranda-Moreno and Fu, 2007) and building confidence intervals for evaluating and screening highway projects (Wood, 2005; Geedipally and Lord, 2008). NB models have also been shown to be unable to handle data with extremely low mean values, which often produces sample data with many zeros. Given these important limitations, there is a need to evaluate whether other count data models could be used for modeling motor vehicle crashes.

The objectives of this study are to evaluate the application of the Conway–Maxwell–Poisson (COM-Poisson) GLM for analyzing motor vehicle crashes and compare the results with those produced from the NB model. The COM-Poisson distribution has very recently been re-introduced by statisticians for modeling count data that are characterized by either over- or under-dispersion (Shmueli et al., 2005; Kadane et al., 2006; Guikema and Coffelt, 2008). This distribution was first introduced in 1962 (Conway and Maxwell, 1962), but has been evaluated in the context of a GLM model only once (Guikema and Coffelt, 2008). Consequently, nobody has so far examined how the COM-Poisson GLM could be used for modeling crash data using common functional forms linking crash data to traffic flow variables (often referred to as general annual average daily traffic or AADT models). This type of functional form is the most popular type of models developed and used by transportation safety analysts (Hauer, 1997; Persaud et al., 2001). They are often preferred over models that include several covariates because they can be easily re-calibrated when they are developed in one jurisdiction and applied to another (Persaud et al., 2002; Lord and Bonneson, 2005). In fact, this type of model will be the kind of model used for estimating the safety performance of rural and urban highways as well as for intersections in the forthcoming highway safety manual (HSM) (Hughes et al., 2005). Although such models will suffer from an omitted variables bias (because many non-flow related factors are known to affect the frequency of crashes), the empirical assessment carried out in this work still provides valuable insight into the potential applicability of COM-Poisson GLM for modeling crash data.

To accomplish the objectives of the study, NB and COM-Poisson GLMs were developed and compared using two datasets. The first dataset contained crash data collected at four-legged signalized intersections in Toronto, Ont. The second dataset included data collected for rural four-lane divided and undivided highways in Texas. The study will show that COM-Poisson GLMs perform as well as NB models when AADT is used as the only covariate. The COM-Poisson GLM offers potential for safety analysis and modeling crash data, since it can also handle under-dispersed data (which the NB GLM cannot or has difficulties converging, see below) and datasets that contain intermingled over- and under-dispersed counts (for dual-link models only, since the dispersion characteristic is captured using the covariate-dependent shape parameter).

This paper is divided into six sections. Section 2 describes the parameterization characteristics of the COM-Poisson distribution and its GLM framework. Section 3 describes the methodology used for estimating and comparing the models. Section 4 presents the summary statistics of the two datasets. Section 5 summarizes the results of the comparison analysis. Section 6 provides a discussion about the application of COM-Poisson GLMs in highway safety, as well as ideas for further research. Section 7 presents a summary of the analysis carried out in this work.

2. Background

This section describes the characteristics of the COM-Poisson distribution and the GLM for modeling crash data. The first part covers the parameterization of the distribution. The second part describes its extension for modeling count data subjected to over- and under-dispersion.

2.1. Parameterization

The COM-Poisson distribution is a generalization of the Poisson distribution and was first introduced by Conway and Maxwell (1962) for modeling queues and service rates. Shmueli et al. (2005) further elucidated the statistical properties of the COM-Poisson distribution using the formulation given by Conway and Maxwell (1962), and Kadane et al. (2006) developed the conjugate distributions for the parameters of the COM-Poisson distribution. Its probability mass function (PMF) can be given by

$$P(Y = y) = \frac{1}{Z(\lambda, \nu)} \frac{\lambda^y}{(y!)^\nu} \tag{1}$$

$$Z(\lambda, \nu) = \sum_{n=0}^{\infty} \frac{\lambda^n}{(n!)^\nu} \tag{2}$$

where, Y is a discrete count; λ is a centering parameter that is approximately the mean of the observations in many cases; and, ν is defined as the shape parameter of the COM-Poisson distribution.

The COM-Poisson can model both under-dispersed ($\nu > 1$) and over-dispersed ($\nu < 1$) data, and several common PMFs are special cases of the COM with the original formulation. Specifically, setting $\nu = 0$ yields the geometric distribution; $\lambda < 1$ and $\nu \to \infty$ yields the Bernoulli distribution in the limit; and $\nu = 1$ yields the Poisson distribution. This flexibility greatly expands the types of problems for which the COM-Poisson distribution can be used to model count data.

With the original formulation, the first two central moments of the COM-Poisson distribution are given by

$$E[Y] = \frac{\partial \log Z}{\partial \log \lambda} \tag{3}$$

$$\text{var}[Y] = \frac{\partial^2 \log Z}{\partial \log^2 \lambda} \tag{4}$$

The COM-Poisson distribution does not have closed-form expressions for its moments in terms of the parameters λ and ν. However, the mean can be approximated through a few different approaches, including (i) using the mode, (ii) including only the first few terms of Z when ν is large, (iii) bounding $E[Y]$ when ν is small, and (iv) using an asymptotic expression for Z in Eq. (1). Shmueli et al. (2005) used the last approach to derive the approximation:

$$E[Y] \approx \lambda^{1/\nu} + \frac{1}{2\nu} - \frac{1}{2} \tag{5}$$

Using the same approximation for Z as in Shmueli et al. (2005), the variance can be defined as

$$\text{var}[Y] \approx \frac{1}{\nu}\lambda^{1/\nu} \tag{6}$$

Care should be taken in using these approximations. In particular, they may not be accurate for $\nu > 1$ or $\lambda^{1/\nu} < 10$ (Shmueli et al., 2005).

Despite its flexibility and attractiveness, the COM-Poisson has limitations in its usefulness as a basis for a GLM, as documented in Guikema and Coffelt (2008). In particular, neither λ nor ν provide a clear centering parameter. While λ is approximately the mean when ν is close to one, it differs substantially from the mean for small ν. Given that ν would be expected to be small for over-dispersed data, this would make a COM model based on the original COM formulation difficult to interpret and use for over-dispersed data.

To circumvent this problem, Guikema and Coffelt (2008) proposed a re-parameterization of the COM-Poisson distribution to provide a clear centering parameter. This new formulation of the COM-Poisson is summarized:

$$P(Y = y) = \frac{1}{S(\mu, \nu)}\left(\frac{\mu^y}{y!}\right)^\nu \tag{7}$$

$$S(\mu, \nu) = \sum_{n=0}^{\infty}\left(\frac{\mu^n}{n!}\right)^\nu \tag{8}$$

By substituting $\mu = \lambda^{1/\nu}$ in Eqs. (4), (5), and (41) of Shmueli et al. (2005), the mean and variance of Y are given in terms of the new formulation as

$$E[Y] = \frac{1}{\nu}\frac{\partial \log S}{\partial \log \mu}, \tag{9}$$

and

$$V[Y] = \frac{1}{\nu^2}\frac{\partial^2 \log S}{\partial \log^2 \mu} \tag{10}$$

with asymptotic approximations $E[Y] \approx \mu + 1/2\nu - 1/2$ and $\text{var}(Y) \approx \mu/\nu$, respectively. It is important to note that these approximations are especially accurate once $\mu > 10$. With this new parameterization, the integral part of μ is now the mode leaving μ as a reasonable approximation of the mean. The substitution $\mu = \lambda^{1/\nu}$ also allows ν to keep its role as a shape parameter. That is, if $\nu < 1$, the variance is greater than the mean while $\nu > 1$ leads to under-dispersion.

This new formulation provides a good basis for developing a COM-Poisson GLM. The clear centering parameter provides a basis on which the centering link function can be built, allowing ease of interpretation across a wide range of values of the shape parameter. Furthermore, the shape parameter ν provides a basis for using a second link function to allow the amount of over-dispersion or under-dispersion to vary across measurements (similar to the varying dispersion parameter of the NB model, as discussed by Miaou and Lord, 2003).

2.2. Generalized linear model

Guikema and Coffelt (2008) developed a COM-Poisson GLM framework for modeling discrete count data. Eqs. (11)–(14) describe this modeling framework. The framework is in effect a dual-link GLM in which both the mean and the variance depend on covariates. In Eqs. (11)–(14), Y is the count random variable being modeled, x_i and z_j are covariates, and there are assumed to be p covariates used in the centering link function and q covariates used in the shape link function. The sets of parameters used in the two link functions are not necessarily identical:

$$P(Y = y) = \frac{1}{S(\mu, \nu)}\left(\frac{\mu^y}{y!}\right)^\nu \tag{11}$$

$$S(\mu, \nu) = \sum_{n=0}^{\infty}\left(\frac{\mu^n}{n!}\right)^\nu \tag{12}$$

$$\ln(\mu) = \beta_0 + \sum_{i=1}^{p}\beta_i x_i \tag{13}$$

$$\ln(\nu) = \gamma_0 + \sum_{j=1}^{q}\gamma_j z_j \tag{14}$$

The GLM described above is highly flexible and readily interpreted. It can model under-dispersed data sets, over-dispersed data sets, and data sets that contain intermingled under-dispersed and over-dispersed counts (for dual-link models only, since the dispersion characteristic is captured using the covariate-dependent shape parameter). The variance is allowed to depend on the covariate values, which can be important if high (or low) values of some covariates tend to be variance decreasing while high (or low) values of other covariates tend to be variance increasing. The parameters have a direct link to either the mean or the variance, providing insight into the behavior and driving factors in the problem, and the mean and variance of the predicted counts are readily approximated based on the covariate values and regression parameter estimates. In this paper, the research team modified the formulation of Guikema and Coffelt (2008) by removing the second link and estimating a single shape parameter ν. This makes the model more directly comparable to the standard single-link NB model. In future work, the potential benefit of adding the second link function, the

shape link, could be explored in the context of modeling traffic crashes, as Guikema and Coffelt (2008) have done in other contexts.

The parameter estimation for the COM-Poisson GLM presented above is challenging. The likelihood equation for the COM-Poisson GLM is complex, making analytical and numerical maximum likelihood estimation difficult. Thus, Bayesian estimation provides an attractive alternative for estimating the coefficients of the model (e.g., Eqs. (13) and (14)). The next section describes the methodology used for estimating and comparing COM-Poisson GLMs and NB models.

3. Methodology

This section describes the methodology used for estimating and comparing the two types of model. For each dataset, COM-Poisson GLMs and NB models were initially estimated using the entire dataset. Then, five samples, which consisted of 80% of the original data, were randomly extracted. The models were developed using the subsets and were then applied to the remaining 20% to evaluate their predictive performance.

For demonstration purposes, the functional forms used for the models were the following (see Eq. (13) above):

Toronto intersection data:

$$\mu_i = \beta_0 F_{\text{Maj\_}i}^{\beta_1} F_{\text{Min\_}i}^{\beta_2} \tag{15}$$

Texas segment data:

$$\mu_j = \beta_0 L_j F_j^{\beta_1} \tag{16}$$

where, μ_i = the mean number of crashes per year for intersection i; μ_j = the mean number of crashes per year for segment j; $F_{\text{Maj\_}i}$ = entering flow for the major approach (average annual daily traffic or AADT) for intersection i; $F_{\text{Min\_}i}$ = entering flow for the minor approach (average annual daily traffic or AADT) for intersection i; F_j = flow traveling on segment j (average annual daily traffic or AADT) and time period t; L_j = length in miles for segment j; and β_0, β_1, β_2 = estimated coefficients.

The functional forms described above are very frequently used by transportation safety analysts. Although they are not considered the most adequate functional form for flow-only models (see Miaou and Lord, 2003), since they under-perform near the boundary conditions (at least for intersections), they are still relevant for this study, as they are considered established functional forms in the highway safety literature.

Several methods were used for estimating the GOF and predictive performance of the models. The methods used in this research include the following.

3.1. Deviance information criterion (DIC)

The DIC is defined as

$$\text{DIC} = \bar{D} + p_D \tag{17}$$

where $\bar{D} = -2 \log L$ represents the posterior mean of the deviance of the unstandardized model where L is the mean of the model log likelihood; $p_D = \bar{D} - \bar{D}(y|\bar{\theta})$ represents the penalty for the number of effective model parameters where $D(y|\bar{\theta})$ is the point estimate of deviance for the posterior means $\bar{\theta}$.

3.2. Mean absolute deviance (MAD)

MAD provides a measure of the average mis-prediction of the model (Oh et al., 2003). It is computed using the following equation:

$$\text{mean absolute deviance (MAD)} = \frac{1}{n}\sum_{i=1}^{n} |\hat{y}_i - y_i| \tag{18}$$

3.3. Mean-squared predictive error (MSPE)

MSPE is typically used to assess the error associated with a validation or external data set (Oh et al., 2003). It can be computed using the

$$\text{mean absolute predictive error (MSPE)} = \frac{1}{n}\sum_{i=1}^{n} (\hat{y}_i - y_i)^2 \tag{19}$$

The coefficients of the COM-Poisson GLMs and FB NB models were estimated using the software WinBUGS (Spiegelhalter et al., 2003) (*note*: to distinguish between the Bayesian and the MLE methods, FB NB refers to a model estimated using a Bayesian approach while MLE NB refers to a model estimated using the Frequentist approach). Vague or non-informative hyper-priors were utilized for the COM-Poisson GLMs and FB NB (described below). A total of three Markov chains were used in the model estimation process with 35,000 iterations per chain and the thinning parameter was set to 1. The first 20,000 iterations (burn-in samples) were discarded. Thus, the remaining 15,000 iterations were used for estimating the coefficients. The Gelman–Rubin (G–R) convergence statistic was used to verify that the simulation runs converged properly. In the analysis, the research team ensured that G–R statistic was less than 1.1. For comparison, Mitra and Washington (2007) suggested that convergence was achieved when the G–R statistic was less than 1.2.

Given the limitations described above and to confirm the results of the FB NB models, MLE NB models were also estimated using SAS (SAS Institute Inc., 2002). These models were not used for the comparison analysis, but to ensure the FB NB models were stable. It should be pointed out that the PMF of the NB models estimated in this work was the following:

$$f(y_i; \phi, \mu_i) = \binom{y_i + \phi - 1}{\phi - 1} \left(\frac{\phi}{\mu_i + \phi}\right)^{\phi} \left(\frac{\mu_i}{\mu_i + \phi}\right)^{y_i} \tag{20}$$

where, y_i = response variable for observation i; μ_i = mean response for observation i; and, ϕ = inverse dispersion parameter of the Poisson-gamma distribution.

The next section describes the characteristics of the two datasets.

4. Data description

This section describes the characteristics of the two datasets. The first part summarizes the characteristics of the Toronto data. The second part presents the summary statistics for the Texas data.

4.1. Toronto data

The first dataset contained crash data collected in 1995 at four-legged signalized intersections located in Toronto, Ont. The data have previously been used for several research projects and have been found to be of relatively good quality (Lord, 2000; Miaou and Lord, 2003; Miranda-Moreno and Fu, 2007). In total, 868 signalized intersections were used in this dataset. Table 1 presents the summary statistics for the full dataset and the five samples (both for the fitted data and predicted data).

4.2. Texas

The second dataset contained crash data collected at four-lane rural undivided and divided segments in Texas. To increase the sample size, divided and undivided segments were merged together. Since the study objective was not related to examining the safety performance of segments per se, putting the data together was deemed adequate. The data were provided by the Texas Department of Public Safety (DPS) and the Texas Department of Transportation (TxDOT) and were used for the project NCHRP 17–29 (methodology for estimating the safety performance of multilane rural highways). The final database included 3220 segments (≥0.1 mile) and 5 years of crash data. Table 2 presents the summary statistics for the full dataset and the five samples (both for fitted data and predicted data).

5. Modeling results

This section presents the modeling results for the COM-Poisson GLMs as well as for the FB and MLE NB models and is divided into three parts. The first part explains the modeling results for the Toronto data. The second part provides details about the modeling results for the Texas data. The last part documents the marginal analysis used for examining the regression coefficients.

5.1. Toronto data

Table 3 summarizes the results of the COM-Poisson GLMs for the Toronto data. This table shows that the coefficients for the flow parameters are below one, which indicates that the crash risk increases at a decreasing rate as traffic flow increases. It should be pointed out that the 95% marginal posterior credible intervals for each of the coefficients did not include the origin.

Table 1
Summary statistics for the Toronto data

	Fitting data				Predicting data			
	Min.	Max.	Average	Total	Min.	Max.	Average	Total
Full data[a]								
Crashes	0	54	11.56 (10.02)	10030	–	–	–	–
Major AADT	5469	72178	28044.81 (10660.4)	–	–	–	–	–
Minor AADT	53	42644	11010.18 (8599.40)	–	–	–	–	–
Sample 1[b]								
Crashes	0	54	11.49 (9.94)	7974	0	51	11.82 (10.33)	2056
Major AADT	5469	72178	28097.36 (10656.9)	–	9622	67214	27835.21 (10702.5)	–
Minor AADT	71	41288	10904.03 (8421.3)	–	53	42644	11433.55 (9289.40)	–
Sample 2								
Crashes	0	54	11.50 (9.96)	7983	0	48	11.76 (10.28)	2047
Major AADT	5469	67214	27946.05 (10490.5)	–	7361	72178	28438.69 (11335.8)	–
Minor AADT	53	42644	10862.04 (8532.00)	–	877	41029	11601.03 (8863.55)	–
Sample 3								
Crashes	0	53	11.67 (10.07)	8099	0	54	11.10 (9.84)	1931
Major AADT	5469	72178	28115.98 (10824.1)	–	5967	56623	27760.95 (10005.6)	–
Minor AADT	71	42644	11169.03 (8678.03)	–	53	36002	10376.61 (8272.24)	–
Sample 4								
Crashes	0	54	11.63 (10.02)	8074	0	50	11.24 (10.02)	1956
Major AADT	5469	68594	28072.28 (10652.6)	–	7361	72178	27935.25 (10721.7)	–
Minor AADT	465	42644	11119.84 (8749.1)	–	53	41288	10572.82 (7983.41)	–
Sample 5								
Crashes	0	54	11.69 (10.08)	8113	0	44	11.02 (9.79)	1917
Major AADT	5469	72178	28103.95 (10641.9)	–	5967	56697	27808.91 (10761.52)	–
Minor AADT	71	42644	11104.99 (8712.5)	–	53	34934	10632.03 (8145.70)	–

[a] 868 observations.
[b] 694 observations for the fitted data and 174 observations for the predicted data.

Table 2
Summary statistics for the Texas dataset (5 years)

	Fitting data				Predicting data			
	Min.	Max.	Average	Total	Min.	Max.	Average	Total
Full data[a]								
Crashes	0	108	4.89 (8.45)	15753	–	–	–	–
AADT	42	89264	8639.27 (6606.57)	–	–	–	–	–
Length (miles)	0.1	11.21	0.80 (1.02)	2576.18	–	–	–	–
Sample 1[b]								
Crashes	0	108	4.94 (8.75)	12722	0	52	4.71 (7.16)	3031
AADT	42	89264	8704.98 (6715.86)	–	420	52294	8376.39 (6147.98)	–
Length (miles)	0.1	11.21	0.80 (1.02)	2054.91	0.1	8.319	0.81 (1.04)	521.269
Sample 2								
Crashes	0	108	4.93 (8.72)	12703	0	48	4.74 (7.31)	3050
AADT	42	89264	8699.58 (6681.18)	–	266	52294	8398.03 (6298.54)	–
Length (miles)	0.1	11.21	0.80 (1.03)	2063.87	0.1	8.517	0.80 (0.96)	512.313
Sample 3								
Crashes	0	108	4.91 (8.21)	12655	0	97	4.81 (9.38)	3098
AADT	158	89264	8645.19 (6685.83)	–	42	53714	8615.58 (6284.50)	–
Length (miles)	0.1	8.548	0.81 (1.02)	2076.86	0.1	11.21	0.78 (1.01)	499.318
Sample 4								
Crashes	0	108	4.94 (8.44)	12731	0	97	4.69 (8.53)	3022
AADT	42	89264	8646.10 (6619.57)	–	158	53714	8611.93 (6559.38)	–
Length (miles)	0.1	11.21	0.80 (1.01)	2073.67	0.1	8.517	0.78 (1.04)	502.51
Sample 5								
Crashes	0	108	5.00 (8.69)	12891	0	53	4.44 (7.43)	2862
AADT	42	89264	8700.86 (6712.74)	–	264	52294	8392.89 (6162.47)	–
Length (miles)	0.1	8.548	0.81 (1.02)	2091.04	0.1	11.21	0.75 (1.02)	485.143

[a] 3220 observations.
[b] 2756 observations for the fitted data and 644 observations for the predicted data.

Table 3
Modeling results for the COM-Poisson GLMs using the Toronto data

Estimates[a]	Full data	Sample 1	Sample 2	Sample 3	Sample 4	Sample 5	Average
$\ln(\beta_0)$	−11.53 (0.4159)[b]	−11.7589 (0.742)	−11.7252 (0.560)	−11.2626 (0.487)	−11.2033 (0.729)	−11.1643 (0.709)	–
β_1	0.6350 (0.04742)	0.6527 (0.076)	0.6641 (0.055)	0.6078 (0.049)	0.6071 (0.072)	0.5949 (0.066)	–
β_2	0.7950 (0.03101)	0.7999 (0.029)	0.7854 (0.029)	0.7960 (0.031)	0.7912 (0.030)	0.8010 (0.027)	–
ν	0.3408 (0.02083)	0.3333 (0.023)	0.3454 (0.023)	0.3359 (0.023)	0.3497 (0.024)	0.3499 (0.024)	–
DIC	4953.7	3974.34	3953.69	3981.33	3953.66	3956.85	–
MAD_{fit}	4.129	4.141	4.075	4.156	4.132	4.074	4.118
$MSPE_{fit}$	33.664	34.433	33.102	34.108	33.508	33.176	33.665
MAD_{pred}	–	4.082	4.3003	4.034	4.106	4.316	4.168
$MSPE_{pred}$	–	30.529	34.695	32.339	34.059	34.663	33.257

[a] The coefficient estimates are based on the mode (posterior value) (see Section 6).
[b] Posterior credible standard error.

Table 4
Modeling results for the FB NB models using the Toronto data

Estimates	Full data	Sample 1	Sample 2	Sample 3	Sample 4	Sample 5	Average
$\ln(\beta_0)$	−10.11 (0.4794)	−9.862 (0.4018)	−10.48 (0.4359)	−9.799 (0.5509)	−10.14 (0.5256)	−9.87 (0.5272)	–
β_1	0.6071 (0.046)	0.5788 (0.039)	0.6459 (0.044)	0.5775 (0.052)	0.6057 (0.051)	0.5787 (0.055)	–
β_2	0.6852 (0.021)	0.6903 (0.022)	0.6840 (0.025)	0.6848 (0.023)	0.6902 (0.024)	0.6918 (0.023)	–
ϕ	7.120 (0.619)	6.898 (0.669)	7.256 (0.721)	7.045 (0.687)	7.388 (0.734)	7.567 (0.756)	–
DIC	4777.59	3821.9	3817.8	3836.35	3811.16	3824.74	–
MAD_{fit}	4.141	4.174	4.094	4.168	4.145	4.096	4.136
$MSPE_{fit}$	32.742	33.503	32.104	33.271	32.527	32.354	32.750
MAD_{pred}	–	4.024	4.379	4.058	4.121	4.346	4.186
$MSPE_{pred}$	–	29.594	35.091	30.855	33.331	33.989	32.572

Table 5
Modeling results for the MLE NB models using the Toronto data

Estimates	Full data	Sample 1	Sample 2	Sample 3	Sample 4	Sample 5	Average
$\ln(\beta_0)$	−10.2458 (0.465)	−10.1664 (0.525)	−10.4031 (0.520)	−9.7398 (0.513)	−10.2040 (0.513)	−9.8473 (0.511)	–
β_1	0.6207 (0.046)	0.6079 (0.051)	0.6393 (0.051)	0.5707 (0.051)	0.6119 (0.051)	0.5778 (0.051)	–
β_2	0.6853 (0.0211)	0.6910 (0.0241)	0.6826 (0.024)	0.6860 (0.024)	0.6905 (0.024)	0.6903 (0.0234)	–
α[a]	0.1398 (0.0122)	0.1443 (0.014)	0.1372 (0.014)	0.1410 (0.014)	0.1349 (0.0134)	0.1315 (0.0134)	–
AIC	5077.3	4068.8	4052.6	4080.5	4045.2	4054.3	–
MAD_{fit}	4.142	4.170	4.092	4.168	4.146	4.096	4.136
$MSPE_{fit}$	32.699	33.444	32.127	33.264	32.517	32.370	32.737
MAD_{pred}	–	4.026	4.374	4.062	4.117	4.348	4.185
$MSPE_{pred}$	–	29.547	34.973	30.898	33.271	34.002	32.538

[a] $\alpha = 1/\phi$.

Table 4 summarizes the results of the FB NB models for the Toronto data. This table exhibits similar characteristics as for the COM-Poisson GLMs in terms of GOF statistics and predictive performance despite the fact that the coefficients are a little bit different. Nonetheless, this difference did not affect the fit and predictive capabilities of the models. A comparison of the models' output is presented below.

Table 5 summarizes the results of the MLE NB models for the Toronto data. This table shows exactly the same results as for the FB NB. This is expected since a vague prior was used for the FB NB models. The results indicate that the FB NB models are relatively stable and can, therefore, be compared with the COM-Poisson GLM.

Fig. 1 compares the estimated number of crashes from the COM-Poisson and NB models for three minor AADT flows (F_{Min}). The figure illustrates that the estimated values are slightly different, especially when F_{Min} is equal to 500 vehicles/day, with the COM-Poisson output being always lower than the NB output. For $F_{Min} = 500$, the maximum absolute difference is about 0.9 crash per year. At the other end of the spectrum, the maximum absolute difference is about two crashes per year for $F_{Maj} = 70{,}000$ and $F_{Min} = 3000$. Although this difference appears to be large, the relative difference is about 17%. It should be pointed out that when the posterior mean value is used for the COM-Poisson model rather than the mode (e.g., assuming μ is the predicted mean), both curves get closer for all minor flow values. For $F_{Min} = 3000$, the absolute maximum difference becomes less than one crash per year. Thus, both models could be used for analyzing this dataset.

5.2. Texas

Table 6 summarizes the results of the COM-Poisson models for the Texas data. Similar to the first dataset, the 95% marginal posterior credible intervals for each of the coefficients did not include the origin. In addition, the coefficients do not vary significantly between the different samples.

Table 7 summarizes the results of the FB NB models for the Texas data. This table indicates that FB NB models estimate a slightly lower value for the coefficient for the traffic flow variable than for the COM-Poisson GLMs. Similar to the Toronto data, the COM-Poisson GLMs offer the same statistical performance as for FB NB models.

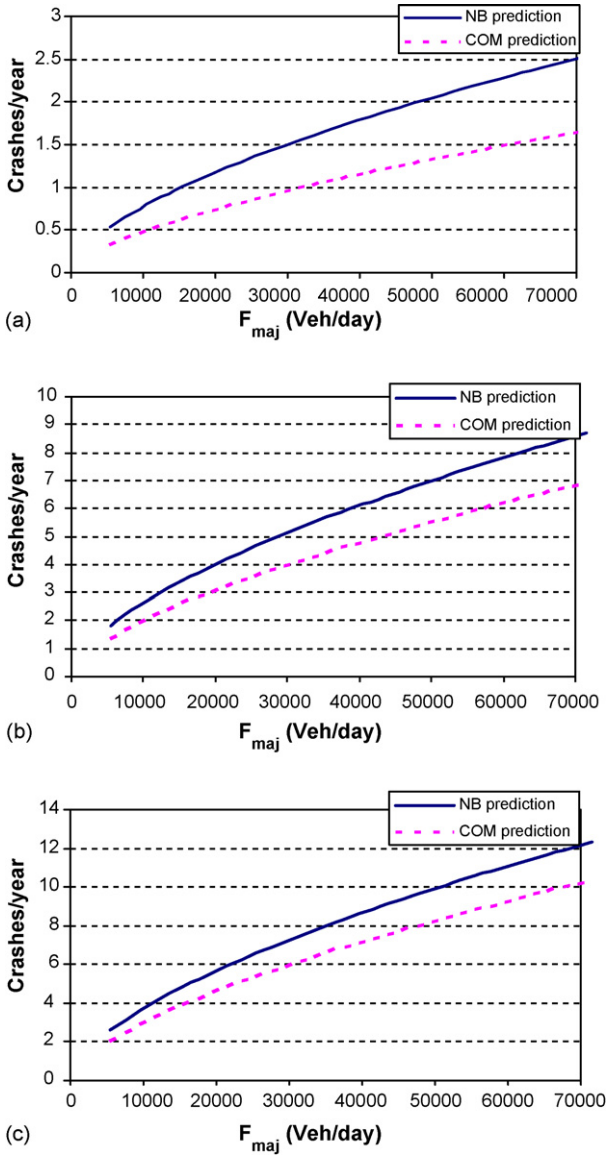

Fig. 1. Estimated values for the Toronto DATA. NB and COM-Poisson models: (a) minor AADT = 500 vehicles/day; (b) minor AADT = 3000 vehicles/day; and (c) minor AADT = 5000 vehicles/day.

Table 6
Modeling results for the COM-Poisson GLMs using the Texas data

Estimates[a]	Full data	Sample 1	Sample 2	Sample 3	Sample 4	Sample 5	Average
$\ln(\beta_0)$	−8.235 (0.242)[b]	−8.442 (0.267)	−8.333 (0.284)	−7.877 (0.223)	−8.155 (0.234)	−8.338 (0.249)	–
β_1	1.081 (0.025)	1.102 (0.028)	1.089 (0.030)	1.044 (0.023)	1.074 (0.025)	1.092 (0.026)	–
ν	0.3608 (0.012)	0.3504 (0.014)	0.3465 (0.013)	0.3699 (0.014)	0.3701 (0.015)	0.3650 (0.013)	–
DIC	13325.6	10688.8	10711.2	10673.9	10652.6	10710.6	–
MAD_{fit}	2.385	2.433	2.435	2.369	2.371	2.415	2.401
$MSPE_{fit}$	21.985	24.297	23.708	18.970	20.050	22.938	21.991
MAD_{pred}	–	2.240	2.242	2.388	2.413	2.283	2.313
$MSPE_{pred}$	–	14.462	16.650	31.748	28.745	18.835	22.088

[a] The coefficient estimates are based on the mode (posterior value) (see Section 6).
[b] Posterior credible standard error.

Table 7
Modeling results for the FB NB models using the Texas data

Estimates	Full data	Sample 1	Sample 2	Sample 3	Sample 4	Sample 5	Average
$\ln(\beta_0)$	−6.512 (0.227)	−6.701 (0.264)	−6.488 (0.211)	−6.356 (0.221)	−6.449 (0.209)	−6.618 (0.204)	–
β_1	0.9206 (0.025)	0.9403 (0.029)	0.9178 (0.023)	0.9046 (0.024)	0.9134 (0.023)	0.9320 (0.022)	–
ϕ	3.229 (0.177)	3.155 (0.189)	3.073 (0.184)	3.235 (0.198)	3.253 (0.198)	3.328 (0.200)	–
DIC	12408.7	9882.13	9908.99	9988.68	9964.02	9983.94	–
MAD_{fit}	2.437	2.466	2.508	2.430	2.424	2.453	2.453
$MSPE_{fit}$	20.699	22.758	22.487	18.070	18.284	21.651	20.658
MAD_{pred}	–	2.297	2.143	2.509	2.464	2.378	2.358
$MSPE_{pred}$	–	12.929	13.307	31.162	29.854	17.369	20.924

Table 8
Modeling results for the MLE NB models using the Texas data

Estimates	Full data	Sample 1	Sample 2	Sample 3	Sample 4	Sample 5	Average
$\ln(\beta_0)$	−6.5605 (0.199)	−6.6293 (0.224)	−6.4570 (0.224)	−6.4266 (0.222)	−6.4652 (0.220)	−6.6163 (0.220)	–
β_1	0.9260 (0.022)	0.9324 (0.025)	0.9143 (0.025)	0.9125 (0.025)	0.9151 (0.024)	0.9318 (0.024)	–
α[a]	0.3095 (0.017)	0.3172 (0.019)	0.3255 (0.020)	0.3094 (0.019)	0.3075 (0.019)	0.3009 (0.018)	–
AIC	13375	10674	10724	10773	10741	10749	–
MAD_{fit}	2.440	2.463	2.506	2.434	2.424	2.453	2.453
$MSPE_{fit}$	20.766	22.654	22.439	18.190	18.289	21.647	20.664
MAD_{pred}	–	2.293	2.141	2.510	2.463	2.378	2.357
$MSPE_{pred}$	–	12.856	13.282	31.146	29.861	17.365	20.902

[a] $\alpha = 1/\phi$.

Table 8 summarizes the results of the MLE NB models for the Texas data. This table shows exactly the same results as for the FB NB.

Fig. 2 shows the comparison results between the estimated number of crashes per mile per 5-year of the COM-Poisson and NB models for the Texas data (full dataset). This figure illustrates that both estimates are indeed very close.

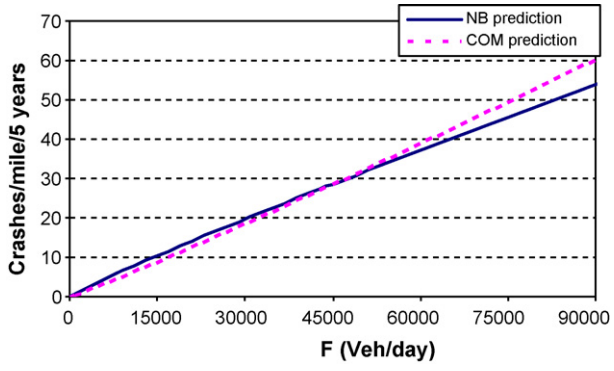

Fig. 2. Estimated values for the Texas data: NB and COM-Poisson models.

5.3. Marginal effects

An important issue in developing or using a regression model is the interpretation of the coefficients. Computing the marginal value of a particular variable can provide valuable information about how the regression coefficient related to that variable influence the expected mean value. For this exercise, the calculations of the marginal effect are slightly more complicated for the COM-Poisson distribution than for the NB distribution, which is usually straightforward. This is attributed to the fact that the parameter μ for the COM-Poisson is a centering parameter, as opposed to the expected mean value typically found in NB models.

The relative marginal effect of a particular variable or covariate X_i can be estimated using the following equation (Cameron

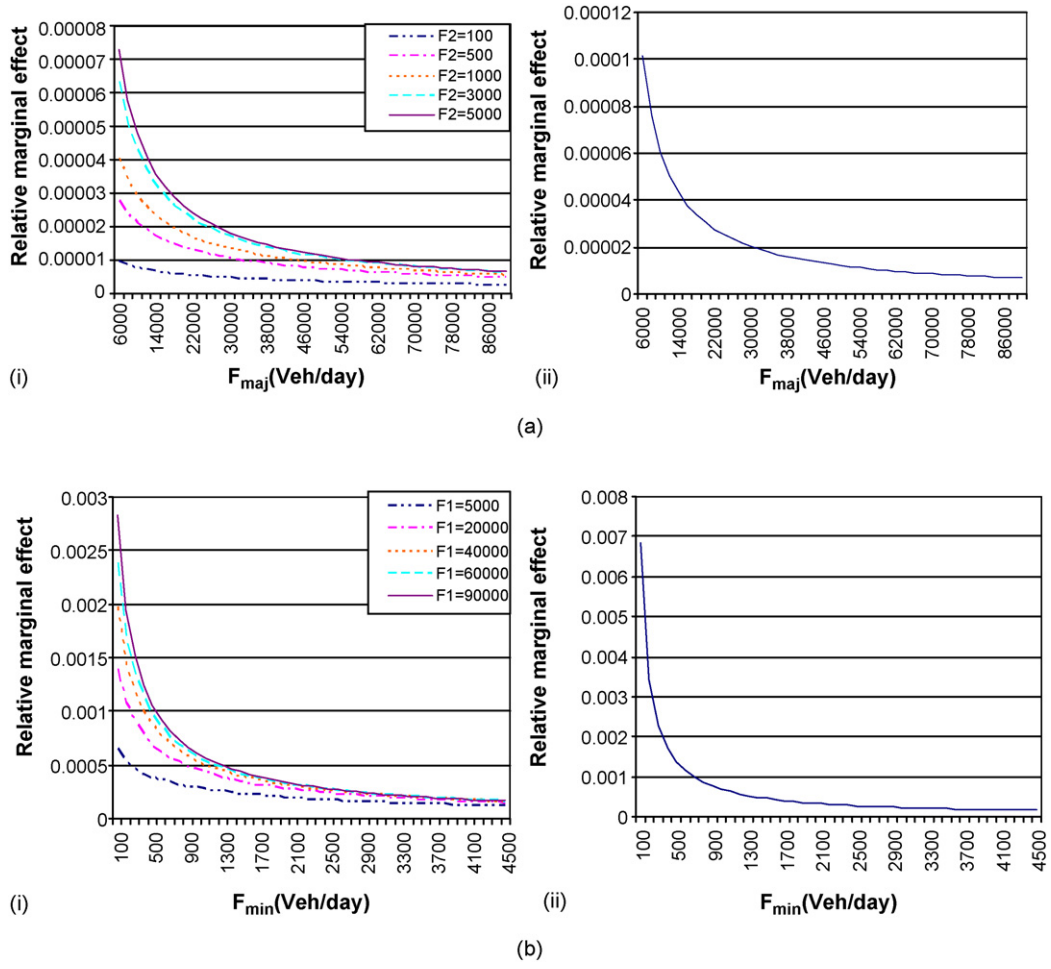

Fig. 3. Marginal effect of the traffic flows for the Toronto model. (a) Marginal effect of the major flow: (i) marginal effect of major flow with COM-Poisson and (ii) marginal effect of major flow with NB. (b) Marginal effect of the minor flow: (i) marginal effect of minor flow with COM-Poisson and (ii) marginal effect of minor flow with NB.

and Trivedi, 1998):

$$\frac{1}{E[Y|X_i]}\frac{\partial E[Y|X_i]}{\partial X_i} \tag{21}$$

In estimating the relative marginal effect of the variables, the mean approximation:

$$E[Y] \approx \mu + 1/2\nu - 1/2 \tag{22}$$

can be used for the COM-Poisson models.

As seen in Fig. 3a, the relative marginal effect of the major flow for the COM-Poisson model depends on the major and minor entering flows. This figure shows that the relative marginal effect on the expected mean for a unit increase in major flow remains nearly constant for lower minor flow volumes (e.g., 100 vehicles/day) and the rate of curvature increases with the increase in minor flows. On the other hand, the relative marginal effect of the major flow for the NB model is only independent of the minor flow. With a unit increase in major flow, the relative marginal effect on the expected mean value decreases. This decrease is smaller for higher major flows. Similar results can be seen for both COM-Poisson and NB models for the marginal effect related to the minor flow (Fig. 3b).

Fig. 4 shows that the relative marginal effect on the expected mean value decreases with a unit increase in flow for both COM-Poisson and NB models. In this figure, the *y*-axis is formatted under the logarithmic scale. In Fig. 4, it can be seen that the marginal effect of traffic flow is higher for the NB model than for the COM-Poisson GLM. For the NB model, there is a sharp decrease in the marginal effect at lower flows and the curve decreases slightly for flows above 10,000 vehicles/day.

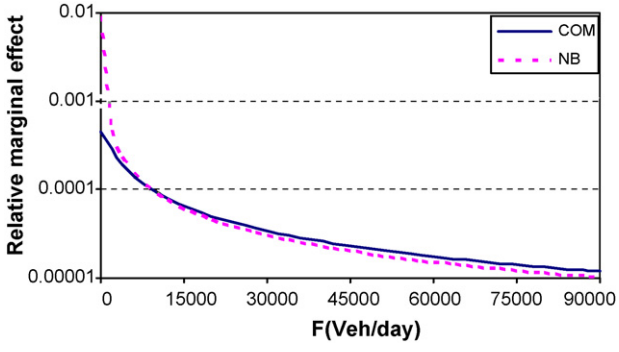

Fig. 4. Marginal effect of traffic flow for the Texas model (*note*: *y*-axis is formatted under a logarithmic scale).

6. Discussion

This paper has shown that the COM-Poisson GLM offers potential for modeling motor vehicle crashes. First, the model performs as well as the NB model (FB and MLE) for the functional form that only includes traffic flow as covariates. As detailed in the modeling results, both models provided similar GOF statistics and predictive performance. Guikema and Coffelt (2008) have reported similar comparison results between the COM-Poisson GLM and the FB NB model. The models used in Guikema and Coffelt (2008) included six covariates in both the centering and shape links. Hence, it is expected that COM-Poisson GLMs developed with several covariates, such as lane and shoulder widths, should work as well as the NB model.

Second, although almost all crash datasets have been shown to exhibit over-dispersion (see Lord et al., 2005b), it has been documented that some crash datasets can sometimes experience under-dispersion (Oh et al., 2006). The NB GLM could theoretically handle under-dispersion, since the dispersion parameter can be negative ($\text{var}(Y) = \mu + (-\alpha)\mu^2$)). However, in this case, the mean of the Poisson is no longer gamma distributed because this latter distribution cannot have negative parameters (i.e., gamma(ϕ, ϕ)). In addition, researchers who have worked on the characterization of the NB distribution and GLM have indicated that a negative dispersion parameter could lead to a mis-specification of the pmf (when $-1/(\text{max of counts}) < \alpha$) (Clark and Perry, 1989; Saha and Paul, 2005). On the other hand, the COM-Poisson distribution has been shown to easily handle such datasets (Shmueli et al., 2005; Kadane et al., 2006; Geedipally et al., 2007; Guikema and Coffelt, 2008). The fact that the model handles under-dispersed data makes it more useful than the NB model, which has difficulty coping with this kind of data (as described above). Although not a good analysis approach, a transportation safety analyst could theoretically not have to worry about the characteristics of the dispersion in the data, since the COM-Poisson GLM can handle both over- and under-dispersed data, and a combination of both, if the data are characterized as such.

Third, as discussed above, crash data can sometimes be subjected to very low sample mean values, which create data characterized by a large number of zeros (with the hypothesis that the space and time scales have been appropriately used, see Lord et al., 2005b). Consequently, NB models do not perform well with such datasets since they may tend to under-predict zero values (or over-estimate non-zero count values). To overcome this problem, some researchers have suggested the use of zero-inflated Poisson and NB models (Shankar et al., 1997). However, other researchers have argued against the use of such models for modeling crash data, since this kind of data does not exhibit two distinct generating processes, one of which is characterized by having a long-term mean equal to zero (which is not feasible for crash data) (Lord et al., 2005b, 2007; Warton, 2005; Wedagama et al., 2006; Kadane et al., 2006). Depending upon the specification of the parameters λ and ν, the COM-Poisson model can predict more zeros than the NB model for the same mean value. Nonetheless, both models should not be used as a direct substitute to zero-inflated models (when they are warranted) (see Kadane et al., 2006). In short, further work is needed on this topic.

Fourth, the COM-Poisson model is not significantly more difficult to implement than the FB NB model once the code for the likelihood is available. Guikema and Coffelt (2008) developed the code needed implement the COM-Poisson GLM in WinBUGS, and this code will be made available through the WinBUGS developer web page (www.winbugs-development.org.uk/). For the models produced in this work, non-informative or vague priors were used for the regression coefficients. For the β coefficients, normal(0, 100) priors were used, for ν, a gamma(0.03, 0.1) prior was used, and for ϕ a gamma(0.1, 0.1) prior was used. The research team also experimented with other non-informative priors, but found that the priors did not significantly affect either GOF of the models or the posterior parameter estimates. In addition, the difference in computational times for these models was not enormous. For example, for the full Toronto data set, a run of the COM-Poisson model with 35,000 replications took about 5 h while a run of the FB NB model with 35,000 replications took between 1 and 1.5 h; the absolute difference seems large, but some simulation runs can sometimes take up to 2 or 3 days in WinBUGS to converge depending on the complexity of the model hierarchical structure. Overall, implementing the COM-Poisson model is not significantly more difficult than implementing the FB NB model once the code for the COM-Poisson model is available.

Given the fact that the COM-Poisson GLM has so far only been applied once, there are many lines of research activities that could be investigated. First, similar to the work performed in this research, general AADT models should be evaluated for under-dispersed data using the COM-Poisson GLM. Second, further research should be conducted about data characterized by extremely low sample mean values (data with many zeros). Third, given the fact that crash data are often subjected to low sample mean values and small sample size, the stability of the COM-Poisson GLM should be investigated. Fourth, since the EB method is now used frequently in highway safety analyzes, an EB modeling framework should be developed for the COM-Poisson model. Fifth, methods about how to use COM-Poisson GLMs for identifying hazardous sites should be investigated. Sixth, although the analysis carried out in this research was conducted by assuming a fixed shape parameter (independent of covariates), further research should be done to examine the effects of a covariate-dependent shape parameter on COM-Poisson GLMs. Preliminary work conducted by the research team on some of these topics seems to offer positive results about using the COM-Poisson GLM for analyzing motor vehicle crashes (see, e.g., Geedipally et al., 2007).

7. Summary and conclusions

This paper has documented the application of the COM-Poisson GLM for analyzing motor vehicle crashes. The COM-Poisson distribution, originally developed in 1962, has recently been re-introduced by statisticians for analyzing count

data subjected to either over- or under-dispersion. This innovative distribution is an extension of the Poisson distribution. The objectives of this study were to evaluate and compare the COM-Poisson GLM with the NB model commonly used for analyzing motor vehicle crashes. The comparison analysis was carried out using the most common functional forms used by transportation safety analysts, which link crashes to the entering flows at intersections or on segments. To accomplish the study objectives, several FB NB and COM-Poisson GLMs were developed using two datasets. The first dataset contained crash data collected at four-legged signalized intersections in Toronto, Ont. The second dataset included data collected for rural four-lane divided and undivided highways in Texas. Several methods were used to assess the statistical fit and predictive performance of the models.

The results of this study show that COM-Poisson GLMs perform as well as FB NB models in terms of GOF statistics and predictive performance. This result is supported by another recent study on this topic (Guikema and Coffelt, 2008). Given the fact the COM-Poisson distribution can also handle under-dispersed data, which have sometimes been observed in crash databases, the COM-Poisson GLM offers a better alternative over the NB model for modeling motor vehicle crashes. The advantage may be even greater if the model is found to be more stable than its counterpart for small sample sizes and low sample mean values. Finally, it is hoped that the new type of model described in this paper will open the door for better and more reliable tools for estimating the safety performance of entities located on transportation networks.

Acknowledgments

The authors would like to thank two anonymous reviewers for providing useful comments to improve this paper.

References

Abbess, C., Jarett, D., Wright, C.C., 1981. Accidents at blackspots: estimating the effectiveness of remedial treatment, with special reference to the "regression-to-mean" effect. Traffic Engineering and Control 22 (10), 535–542.

Airoldi, E.M., Anderson, A.G., Fienberg, S.E., Skinner, K.K., 2006. Who Wrote Ronald Reagan's Radio Addresses? Bayesian Analysis 1 (2), 289–320.

Cameron, A.C., Trivedi, P.K., 1998. Regression analysis of count data. Econometric Society Monograph No. 30. Cambridge University Press, Boston.

Clark, S.J., Perry, J.N., 1989. Estimation of the negative binomial parameter κ by maximum quasi-likelihood. Biometrics 45, 309–316.

Conway, R.W., Maxwell, W.L., 1962. A queuing model with state dependent service rates. Journal of Industrial Engineering 12, 132–136.

Geedipally S., Guikema, S.D., Dhavala, S., Lord, D., 2007. Characterizing the Performance of a Bayesian Conway–Maxwell Poisson GLM. Working Paper, Zachry Department of Civil Engineering, Texas A&M University, College Station, TX.

Geedipally, S.R., Lord, D., 2008. Effects of the Varying Dispersion Parameter of Poisson-gamma models on the Estimation of Confidence Intervals of Crash Prediction models. In: Presented at the 87th Annual Meeting of the Transportation Research Board, Washington, D.C.

Guikema, S.D., Coffelt, J.P., 2008. A flexible count data regression model for risk analysis. Risk Analysis, in press.

Hauer, E., 1997. Observational Before–After Studies in Road Safety: Estimating the Effect of Highway and Traffic Engineering Measures on Road Safety. Elsevier Science Ltd., Oxford.

Hauer, E., Ng, J.C.N., Lovell, J., 1988. Estimation of safety at signalized intersections. Transportation Research Record 1185, 48–61.

Hughes, W., Eccles, K., Harwood, D., Potts, I., Hauer, E., 2005. Development of a Highway Safety Manual. Appendix C: Highway Safety Manual Prototype Chapter: Two-Lane Highways. NCHRP Web Document 62 (Project 17-18(4)). Washington, D.C. (http://onlinepubs.trb.org/onlinepubs/nchrp/nchrp_w62.pdf, accessed October 2007).

Kadane, J.B., Shmueli, G., Minka, T.P., Borle, S., Boatwright, P., 2006. Conjugate analysis of the Conway–Maxwell–Poisson distribution. Bayesian Analysis 1, 363–374.

Kulmala, R., 1995. Safety at Rural Three- and Four-Arm Junctions: Development and Applications of Accident Prediction Models. VTT Publications 233, Technical Research Centre of Finland, Espoo.

Lloyd-Smith, J.O., 2007. Maximum likelihood estimation of the negative binomial dispersion parameter for highly overdispersed data, with applications to infectious diseases. PLoS ONE 2 (2), e180 (http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=1791715, accessed July 2007.

Lord, D., 2000. The prediction of accidents on digital networks: characteristics and issues related to the application of accident prediction models. Ph.D. Dissertation. Department of Civil Engineering, University of Toronto, Toronto, Ontario.

Lord, D., 2006. Modeling motor vehicle crashes using Poisson-gamma models: examining the effects of low sample mean values and small sample size on the estimation of the fixed dispersion parameter. Accident Analysis & Prevention 38 (4), 751–766.

Lord, D., Bonneson, J.A., 2005. Calibration of predictive models for estimating the safety of ramp design configurations. Transportation Research Record 1908, 88–95.

Lord, D., Manar, A., Vizioli, A., 2005a. Modeling crash-flow-density and crash-flow-V/C ratio for rural and urban freeway segments. Accident Analysis & Prevention 37 (1), 185–199.

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. Safety Science, doi:10.1016/j.ssci.2007.03.005, in press.

Lord, D., Washington, S.P., Ivan, J.N., 2005b. Poisson, Poisson-gamma and zero inflated regression models of motor vehicle crashes: balancing statistical fit and theory. Accident Analysis & Prevention 37 (1), 35–46.

Lord, D., Washington, S.P., Ivan, J.N., 2007. Further notes on the application of zero inflated models in highway safety. Accident Analysis & Prevention 39 (1), 53–57.

Maher, M.J., Summersgill, I., 1996. A Comprehensive methodology for the fitting predictive accident models. Accident Analysis & Prevention 28 (3), 281–296.

Maycock, G., Hall, R.D., 1984. Accidents at 4-arm roundabouts. TRRL Laboratory Report 1120. Transportation and Road Research Laboratory, Crowthorne, Bershire.

Miaou, S.-P., Lord, D., 2003. Modeling traffic crash-flow relationships for intersections: dispersion parameter, functional form, and Bayes versus empirical Bayes. Transportation Research Record 1840, 31–40.

Miaou, S.-P., Song, J.J., 2005. Bayesian ranking of sites for engineering safety improvements: decision parameter, treatability concept, statistical criterion and spatial dependence. Accident Analysis & Prevention 37 (4), 699–720.

Mitra, S., Washington, S., 2007. On the nature of over-dispersion in motor vehicle crash prediction models. Accident Analysis & Prevention 39 (3), 459–568.

Miranda-Moreno, L.F., Fu, L., 2007. Traffic safety study: empirical Bayes or full Bayes? Paper 07-1680. In: Proceedings of the 84th Annual Meeting of the Transportation Research Board, Washington, D.C.

Oh, J., Lyon, C., Washington, S.P., Persaud, B.N., Bared, J., 2003. Validation of the FHWA crash models for rural intersections: lessons learned. Transportation Research Record 1840, 41–49.

Oh, J., Washington, S.P., Nam, D., 2006. Accident prediction model for railway-highway interfaces. Accident Analysis & Prevention 38 (2), 346–356.

Persaud, B.N., Lord, D., Palminaso, J., 2002. Issues of calibration and transferability in developing accident prediction models for urban intersections. Transportation Research Record 1784, 57–64.

Persaud B.N., Retting, R., Garder, P., Lord, D., 2001. Observational before-after study of U.S. roundabout conversions using the empirical Bayes method. Transportation Research Record 1751, 1–8.

Piegorsch, W.W., 1990. Maximum likelihood estimation for the negative binomial dispersion parameter. Biometrics 46, 863–867.

Poch, M., Mannering, F.L., 1996. Negative binomial analysis of intersection-accident frequency. Journal of Transportation Engineering 122 (2), 105–113.

Saha, K., Paul, S., 2005. Bias-corrected maximum likelihood estimator of the negative binomial dispersion parameter. Biometrics 61 (1), 179–185.

SAS Institute Inc., 2002. SAS System for Windows. Version 9. Cary, NC.

Shankar, V., Milton, J., Mannering, F.L., 1997. Modeling accident frequency as zero-altered probability processes: an empirical inquiry. Accident Analysis & Prevention 29 (6), 829–837.

Shmueli, G., Minka, T.P., Kadane, J.B., Borle, S., Boatwright, P., 2005. A useful distribution for fitting discrete data: revival of the Conway–Maxwell–Poisson distribution. Journal of the Royal Statistical Society, Part C 54, 127–142.

Spiegelhalter, D.J., Thomas, A., Best, N.G., Lun, D., 2003. WinBUGS Version 1.4.1 User Manual. MRC Biostatistics Unit, Cambridge. Available from: http://www.mrcbsu.cam.ac.uk/bugs/welcome.shtml.

Warton, D.I., 2005. Many zeros does not mean zero inflation: comparing the goodness-of-fit of parametric models to multivariate abundance data. Environmetrics 16, 275–289.

Wedagama, D.M., Bird, R.N., Metcalf, A.V., 2006. The influence of urban land-use on non-motorised transport casualties. Accident Analysis & Prevention 38 (6), 1049–1057.

Wood, G.R., 2002. Generalized linear accident models and goodness of fit testing. Accident Analysis & Prevention 34 (1), 417–427.

Wood, G.R., 2005. Confidence and prediction intervals for generalized linear accident models. Accident Analysis & Prevention 37 (2), 267–273.

Xie, Y., Lord, D., Zhang, Y., 2007. Predicting motor vehicle collisions using Bayesian neural networks: an empirical analysis. Accident Analysis & Prevention 39 (5), 922–933.