Accident Analysis and Prevention 41 (2009) 1118–1123

Contents lists available at ScienceDirect

Accident Analysis and Prevention

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

Accident prediction models with random corridor parameters

Karim El-Basyouny*, Tarek Sayed[1]

Dept. of Civil Engineering, University of British Columbia, Vancouver, BC, Canada V6T 1Z4

ARTICLE INFO

Article history:
Received 4 May 2009
Received in revised form 16 June 2009
Accepted 26 June 2009

Keywords:
Collision prediction models
Full Bayes estimation
Markov Chain Monte Carlo
Random parameters
Corridor variation

ABSTRACT

Recent research advocates the use of count models with random parameters as an alternative method for analyzing accident frequencies. In this paper a dataset composed of urban arterials in Vancouver, British Columbia, is considered where the 392 segments were clustered into 58 corridors. The main objective is to assess the corridor effects with alternate specifications. The proposed models were estimated in a Full Bayes context via Markov Chain Monte Carlo (MCMC) simulation and were compared in terms of their goodness of fit and inference. A variety of covariates were found to significantly influence accident frequencies. However, these covariates resulted in random parameters and thereby their effects on accident frequency were found to vary significantly across corridors. Further, a Poisson-lognormal (PLN) model with random parameters for each corridor provided the best fit. Apart from the improvement in goodness of fit, such an approach is useful in gaining new insights into how accident frequencies are influenced by the covariates, and in accounting for heterogeneity due to unobserved road geometrics, traffic characteristics, environmental factors and driver behavior. The inclusion of corridor effects in the mean function could also explain enough variation that some of the model covariates would be rendered non-significant and thereby affecting model inference.

1. Introduction

Accident prediction models (APM) currently constitute the main tool for estimating the safety performance of road locations. Traditionally, these models are developed using the Poisson-gamma hierarchy, which leads to the negative binomial regression model (e.g., Poch and Mannering, 1996; Hauer, 1997; Hinde and Demetrio, 1998; Lord, 2000; Miaou and Lord, 2003; Sawalha and Sayed, 2006). Another common approach uses the Poisson-lognormal (PLN) hierarchy. The PLN is a good candidate for modeling accident occurrence in the presence of outliers, since its tails are known to be asymptotically heavier than those of the Gamma distribution (Kim et al., 2002; Lord and Miranda-Moreno, 2007). Other empirical evidence and advantages of the PLN model are discussed in Winkelmann (2003).

Recently, several advances in the development of APMs have been made. These include (i) the application of the multivariate Poisson-lognormal (MVPLN) to model accident count data at different levels of severity (Tunaru, 2002; Park and Lord, 2007; Ma et al., 2008; Ye et al., 2009; Aguero-Valverde and Jovanis, 2009; El-Basyouny and Sayed, 2009b), (ii) the use of a two-state Markov Switching model (Malyshkina et al., 2009) to analyze collision frequencies, (iii) the use of variable dispersion parameters to model accident count data (Heydecker and Wu, 2001; Miaou and Lord, 2003; Miranda-Moreno et al., 2005; El-Basyouny and Sayed, 2006; Mitra and Washington, 2007; Lord and Park, 2008) and (iv) the application of different modeling techniques to overcome the excessive zeroes observed in collision data (e.g., Lee and Mannering, 2002; Shankar et al., 2004; Qin et al., 2004; Warton, 2005; Kumara and Chin, 2006; Lord et al., 2005, 2007).

Although Lord and Park (2008) focused on the structure of the variance function, they emphasized that transportation safety analysts should concentrate on the structure of the mean function. Ideally, a good structure along with a proper selection of the covariates for the mean function would simplify the structure of the variance function or at least significantly minimize the magnitude of the variance (Miaou and Song, 2005; Mitra and Washington, 2007).

A recent approach for modeling the mean function advocates the use of random parameters (Anastasopoulos and Mannering, 2009). In contrast to the traditional APM, which fits one regression model to the dataset, the random parameters approach develops different regression models for individual sites. Li et al. (2008) have also considered various models ranging from mixed models (models with random intercepts) to hierarchical models (models where all parameters are allowed to vary from site to site). Random intercept

* Corresponding author at: Dept. of Civil Engineering, University of British Columbia, 2002 - 6250 Applied Science Lane, Vancouver, BC, Canada V6T 1Z4. Tel.: +1 604 716 4470.
E-mail addresses: basyouny@civil.ubc.ca (K. El-Basyouny), tsayed@civil.ubc.ca (T. Sayed).
[1] Tel.: +1 604 822 4379.

0001-4575/$ – see front matter.
doi:10.1016/j.aap.2009.06.025

models were considered earlier by Shankar et al. (1998) and random parameters logit models were considered by Milton et al. (2008).

To benefit from the flexibility of the random parameters approach, it is proposed to cluster the road entities into rather homogeneous groups (e.g., districts, municipalities, zones) and fit a different regression curve for each group. In this paper a dataset composed of urban arterials in Vancouver, British Columbia, is considered where the 392 segments were clustered into 58 corridors. The proposed model fits a different regression curve for each corridor. This model focuses on explaining part of the extra-variation through improvements in the mean function. Alternatively, the inclusion of an additional variance component, corresponding to corridor variation, in a spatial PLN model has been found to alleviate both spatial and extra Poisson variations (El-Basyouny and Sayed, 2009a). This latter model focuses on explaining part of the extra-variation through improvements in the variance function.

To investigate the consequences of incorporating corridor effects in APMs, two models that account for the corridor variation through the mean and variance of an extended PLN model are compared to the traditional PLN model. The models are estimated in a Full Bayes context via MCMC simulation and are compared in terms of their goodness of fit and inference.

2. The models

2.1. Poisson-lognormal (PLN) model

Let Y_i denote the number of accidents at road segment i ($i = 1, \ldots, n$). It is assumed that accidents at the n segments are independent and that:

$$Y_i|\theta_i \sim Poisson(\theta_i). \quad (1)$$

To address over-dispersion for unobserved or unmeasured heterogeneity, it is assumed that:

$$\theta_i = \mu_i \exp(u_i), \quad (2)$$

where

$$\ln(\mu_i) = \beta_0 + \beta_1 X_{i1} + \cdots + \beta_m X_{im}, \quad (3)$$

the X_{ij} are covariates representing certain traffic and geometric characteristics, the $\beta_0, \beta_1, \ldots, \beta_m$ are model parameters and the term $exp(u_i)$ represents a multiplicative random effect. The PLN regression model is obtained by the assumption:

$$\exp(u_i) \sim Lognormal(0, \sigma_u^2) \quad \text{or} \quad u_i \sim N(0, \sigma_u^2), \quad (4)$$

where σ_u^2 denotes the extra Poisson variance. Under the PLN model, we have:

$$E(Y_i) = \mu_i \exp(0.5\sigma_u^2), \qquad Var(Y_i) = E(Y_i) + [E(Y_i)]^2(\exp(\sigma_u^2) - 1). \quad (5)$$

2.2. Accounting for Corridor variation through the variance (M1)

Typically, the n segments under consideration belong to K mutually exclusive corridors. In such cases, an additional variance component can be included in the model to allow for the possibility that different corridors have different accidents risks because traffic, geometric and environmental conditions vary across corridors. Suppose that the ith segment belongs to corridor $c(i) \in \{1, 2, \ldots, K\}$. This leads to the extended model M1:

$$\theta_i = \mu_i \exp(u_i) \exp(w_{c(i)}), \quad (6)$$

where $w_{c(i)} \sim N(0, \sigma_c^2)$ and σ_c^2 denotes the additional variance component representing the variation among different corridors. Under M1, we have:

$$E(Y_i) = \mu_i \exp(0.5[\sigma_u^2 + \sigma_c^2]),$$
$$Var(Y_i) = E(Y_i) + [E(Y_i)]^2(\exp(\sigma_u^2 + \sigma_c^2) - 1). \quad (7)$$

Let $\mu_i^+ = \mu_i \exp(w_{c(i)})$ and note that $\ln(\mu_i^+) = \beta_{c(i),0} + \beta_1 X_{i1} + \cdots + \beta_m X_{im}$, where $\beta_{c(i),0} = \beta_0 + w_{c(i)}$. The model M1 is thus equivalent to a PLN with a random intercept that varies across corridors.

2.3. Accounting for Corridor variation through the mean (M2)

Alternatively, the corridor variation can be represented by allowing all regression coefficients (not just the intercept) to vary randomly from one corridor to another leading to the second extended model (M2):

$$\ln(\mu_i) = \beta_{c(i),0} + \beta_{c(i),1} X_{i1} + \cdots + \beta_{c(i),m} X_{im}, \quad (8)$$

where

$$\beta_{c(i),j} \sim N(\beta_j, \sigma_j^2), \quad j = 0, 1, \ldots, m. \quad (9)$$

Several alternative distributions were considered in (9) but the normal distribution was found to provide the best statistical fit (Milton et al., 2008; Anastasopoulos and Mannering, 2009; see also Li et al., 2008). Under M2, we have:

$$E(Y_i) = \mu_i^* \exp(0.5\sigma_i^2),$$
$$Var(Y_i) = E(Y_i) + [E(Y_i)]^2(\exp(\sigma_i^2) - 1). \quad (10)$$

where

$$\ln(\mu_i^*) = \beta_0 + \beta_1 X_{i1} + \cdots + \beta_m X_{im},$$
$$\sigma_i^2 = \sigma_0^2 + X_{i1}^2\sigma_1^2 + \cdots + X_{im}^2\sigma_m^2 + \sigma_u^2. \quad (11)$$

In practice, a random parameter $\beta_{c(i),j}$ is used in (8) whenever the posterior estimate $\hat{\sigma}_j$ is significantly greater than 0, otherwise the parameter β_j is fixed across segments (Anastasopoulos and Mannering, 2009).

Although PLN is nested within M1, which in turn is nested within M2, the models have quite different characteristics. The PLN and M1 models have the same form (3) for their mean functions and have fixed dispersion parameters σ_u^2 and $\sigma_u^2 + \sigma_c^2$, respectively. In contrast, M2 has a variable dispersion function (11), involving the covariates X_{ij}, and permits both corridor-level inference, using (8), and overall-level inference, using (11).

3. Methodology

3.1. Prior distributions

To obtain the full Bayes estimates it is required to specify prior distributions for the parameters. Prior distributions are meant to reflect to some extent prior knowledge about the parameters of interest. If such prior information is available, it should be used to formulate the so-called informative priors. The specification of informative priors for generalized linear models was dealt with by Bedrick et al. (1996), who considered conditional means priors as well as data augmentation priors of the same form as the likelihood and showed that such priors result in tractable posteriors. A good discussion on the elicitation of priors in accident data analysis can be found in Schluter et al. (1997). In contrast, uninformative (vague) priors are usually used to reflect the lack of prior information (Miaou and Lord, 2003; Qin et al., 2005; Miaou and Song, 2005; Mitra and Washington, 2007; Lord et al., 2008; Li et al., 2008).

The most commonly used priors are diffused normal distributions (with zero mean and large variance) for the regression parameters, $Gamma(\varepsilon,\varepsilon)$ or $Gamma(1,\varepsilon)$ for σ_u^{-2} and σ_c^{-2}, where ε is a small number, e.g., 0.01 or 0.001. It should be noted that introducing vague priors on all unknown parameters can be risky under some conditions such as the combination of low mean and small sample size (Miranda-Moreno and Lord, 2007). In such cases, better results can be obtained by using semi-informative priors with a small mean and variance for the dispersion parameter. The priors reviewed in this section have been successfully used in various applications including road safety.

3.2. Full Bayes estimation

The posterior distributions needed in the full Bayes approach are obtained using MCMC sampling available in the statistical software WinBUGS (Lunn et al., 2000). MCMC methods are used to repeatedly sample from the joint posterior distribution. The techniques generate sequences (chains) of random points, whose distributions converge to the target posterior distributions. A sub-sample is used to monitor convergence and then excluded as a burn-in sample. The remaining iterations are used for parameter estimation, performance evaluation and inference.

Monitoring convergence is important because it ensures that the posterior distribution has been "found". Thereby indicating when parameters sampling should begin. To check convergence, two or more parallel chains with diverse starting values are tracked to ensure full coverage of the sample space. Convergence of multiple chains is assessed using the Brooks–Gelman–Rubin (BGR) statistic (Brooks and Gelman, 1998). A value under 1.2 of the BGR statistic indicates convergence. Convergence is also assessed by visual inspection of the MCMC trace plots for the model parameters as well as by monitoring the ratios of the Monte Carlo errors relative to the respective standard deviations of the estimates; as a rule of thumb these ratios should be less than 0.05.

3.3. Model comparisons and goodness-of-fit

Let D denote the un-standardized deviance of the postulated model. Spiegelhalter et al. (2002) proposed the Deviance Information Criteria (DIC) as a measure of model complexity and fit:

$$DIC = \bar{D} + p_D; \qquad p_D = \bar{D} - \hat{D}, \tag{12}$$

where $\bar{D}$ is the posterior mean of D, $\hat{D}$ is the point estimate obtained by substituting the posterior means of the model's parameters in D and p_D is a measure of model complexity estimating the effective number of parameters. As a goodness of fit measure, DIC is a Bayesian generalization of Akaike's information criteria (AIC) that penalizes larger parameter models.

According to Spiegelhalter et al. (2005), it is difficult to determine what would constitute an important difference in DIC. Very roughly, differences of more than 10 might definitely rule out the model with the higher DIC. Differences between 5 and 10 are substantial. If the difference in DIC is less than 5, and the models make very different inferences, then it could be misleading just to report the model with the lowest DIC.

While the DIC is used for model comparisons, a posterior predictive approach (Gelman et al., 1996; Stern and Cressie, 2000; Li et al., 2008) can be used to assess the goodness-of-fit (adequacy) of the model with the lowest DIC. Such procedures involve generating replicates under the postulated model and comparing the distribution of a certain discrepancy measure such as the chi-square statistic to the value of chi-square obtained using observed data. A model does not fit the data if the observed value of chi-square is far from the predictive distribution; the discrepancy cannot reasonably be explained by chance if the p-values are close to 0 or 1 (Gelman et al., 1996).

The replicates are best obtained simultaneously with model estimation in WinBUGS in order to account for all uncertainties in model parameters as reflected by the estimated distributions.

The chi-square statistic is computed from:

$$\chi^2 = \sum_{i=1}^{n} \left[\frac{[y_i - E(Y_i)]^2}{Var(Y_i)} \right], \tag{13}$$

where the y_i denotes either the observed or the replicated accident frequencies.

4. Data and model description

A total of $n = 392$ urban road segments in the city of Vancouver were investigated for the purpose of developing accident prediction models relating the safety of these road segments to their traffic and geometric characteristics (Sawalha and Sayed, 2001). The data were obtained from the urban city of Vancouver and covered the period from 1994 to 1996. The 392 segments belong to $K = 58$ corridors. Table 1 shows a list of the explanatory variables with their corresponding abbreviations and units.

In this application, Eqs. (3) and (8) are used with $m = 6$, $X_1 = ln(L)$, $X_2 = ln(V)$ and X_3–X_6 as described in Table 1. Although other functional forms are available in the literature (Miaou and Lord, 2003; Qin et al., 2004), only the above functional form is investigated to demonstrate the effect of including corridor variation on the development of APMs. The form allows for no accidents if the exposure is zero and it has been investigated extensively in the literature since the early 1950s (Lord, 2000).

Average values for the traffic volumes (over the 3 years period) were used to build APMs for predicting the total (aggregate) number of accidents. The aggregation is justified on several grounds. For instance, an aggregate APM was found to perform well compared with APMs developed to handle temporal correlation (Lord and Persaud, 2000; see also Anastasopoulos and Mannering, 2009). Moreover, the aggregation of accidents over a period of reasonable length helps to avoid confounding effects and such a phenomenon like regression-to-the-mean (Cheng and Washington, 2005). A statistical summary of the data is shown in Table 2.

Table 1
Model covariates and their corresponding description.

Covariates	Acronym	Symbol	Units
Segment length	L	L	Kilometers (km)
Annual average daily traffic	AADT	V	Vehicles per day (veh/day)
Crosswalks density	CROD	X_3	Crosswalks per km
Business land use	IBUS	X_4	Indicator variable yes = 1, no = 0
Unsignalized intersection density	UNSD	X_5	Intersections per km
Between signal number of lanes	NL	X_6	
Number of accidents	Y	Y	Accidents per 3 years

Table 2
Statistical summary of data set (n = 392 road segments).

Variable	Minimum	Maximum	Mean	Standard deviation
L	0.11	3.61	0.82	0.40
V	4232	62,931	24,412	13,151
X_3	0	10	1.94	2.11
X_4	0	1	0.26	0.44
X_5	0	21.16	5.90	3.40
X_6	2	7	3.89	1.36
Y	1	311	51.45	50.12
Corridor size[a]	1	17	6.76	3.83

[a] The number of road segments within the corridor.

5. Results and discussion

The posterior summaries were obtained via WinBUGS using two chains with 20,000 iterations each, 10,000 of which were excluded as a burn-in sample. Examination of the BGR statistics, ratios of the Monte Carlo errors relative to the standard deviations of the estimates and trace plots for all model parameters indicated convergence.

The DIC statistics were 2840, 2832 and 2829 under the PLN, M1 and M2 models, respectively. Thus, the two extended PLN models fitted the data better than PLN, with the random corridor parameters PLN model providing a significantly better fit than PLN, according to the DIC guidelines.

Under M2, the observed value of chi-square was 290.8. To assess goodness-of-fit, the distribution of the chi-square discrepancy measure in replicated datasets was generated. The observed value of chi-square was located near the center of the replicated distribution, with an associated p-value of 0.485. As a result, M2 seem to perform well in terms of accommodating the variation in accident frequency across segments. In contrast, the observed value of chi-square under M1 was larger, 347.9, with a smaller p-value of 0.156.

Table 3 summarizes the parameter estimates and their 95% credible intervals for the PLN, M1 and M2 models. The table shows that the parameter estimates are significant as the 95% credible intervals were bounded away from zero, except for X_6 (NL) under M2. Apart from the intercepts, the regression coefficients are all positive, indicating that factors such as segment length, AADT, crosswalks density, business land use, unsignalized intersection density and number of lanes are positively associated with the number of accidents.

The incorporation of corridor variation affected the estimation of parameters. In particular, the variable X_6 representing the number of lanes (NL) is significant under PLN and M1, but is not significant under M2. In general, the regression coefficients have changed under the three models but the respective 95% credible intervals still overlapped.

The estimates of σ_u^2 are significant under all three models, demonstrating the presence of over-dispersion in the data. However, the PLN estimate of 0.388 has been reduced by 29% and 44% under M1 and M2, respectively. As expected, accounting for corridor variation reduces the estimates of extra Poisson variation.

The results of M2 in Table 3 indicate that random parameters were associated with the intercept as well as all six covariates. Since the M2 intercept is normally distributed with mean −3.090 and standard deviation 2.294, then it is expected to be less than the PLN fixed intercept of −3.177 on 48.5% of the segments and greater than the PLN intercept on 51.5% of the segments. This variability is capturing the unobserved heterogeneity across corridors.

The road segment length also resulted in a random parameter that is normally distributed, with a mean 0.712 and standard deviation 0.163. Thus, for almost all corridors, accident counts are expected to increase with segment length although by varying magnitudes. A similar result was obtained for AADT, where accident counts are expected to increase with AADT, for the vast majority (99.6%) of the corridors, and are expected to decrease for only a small proportion (0.4%) of the corridors. As noted by Anastasopoulos and Mannering (2009), this AADT finding is likely picking up a complex interaction among traffic volume, driver behavior and accident frequency. It may be capturing, among other factors, the response and adaptation of drivers to various levels of traffic volume.

The estimates of the distributional parameters for CROD, IBUS and UNSD indicate that accident frequency is expected to increase with crosswalks density, business land use and the density of unsignalized intersections on the majority of the corridors; the per-

Table 3
Parameter estimates and 95% credible intervals for the PLN, M1 and M2 models.

	PLN	M1	M2
β_0	−3.177 (−4.395, −2.055)	−2.777 (−4.062, −1.496)	−3.090 (−4.681, −1.587)
β_1	0.683 (0.542, 0.836)	0.711 (0.555, 0.871)	0.712 (0.558, 0.875)
β_2	0.561 (0.435, 0.698)	0.539 (0.400, 0.681)	0.571 (0.413, 0.738)
β_3	0.115 (0.075, 0.149)	0.112 (0.084, 0.143)	0.123 (0.081, 0.163)
β_4	0.249 (0.107, 0.397)	0.290 (0.113, 0.456)	0.286 (0.105, 0.483)
β_5	0.096 (0.075, 0.115)	0.081 (0.060, 0.102)	0.098 (0.066, 0.133)
β_6	0.118 (0.051, 0.186)	0.089 (0.023, 0.163)	0.063 (−0.019, 0.151)
σ_u^2	0.338 (0.286, 0.400)	0.240 (0.199, 0.288)	0.188 (0.150, 0.233)
σ_c^2		0.123 (0.064, 0.211)	
σ_0			2.294 (1.129, 4.404)
σ_1			0.163 (0.067, 0.308)
σ_2			0.214 (0.081, 0.436)
σ_3			0.074 (0.047, 0.116)
σ_4			0.224 (0.078, 0.445)
σ_5			0.077 (0.049, 0.115)
σ_6			0.135 (0.069, 0.224)
DIC	2840	2832	2829

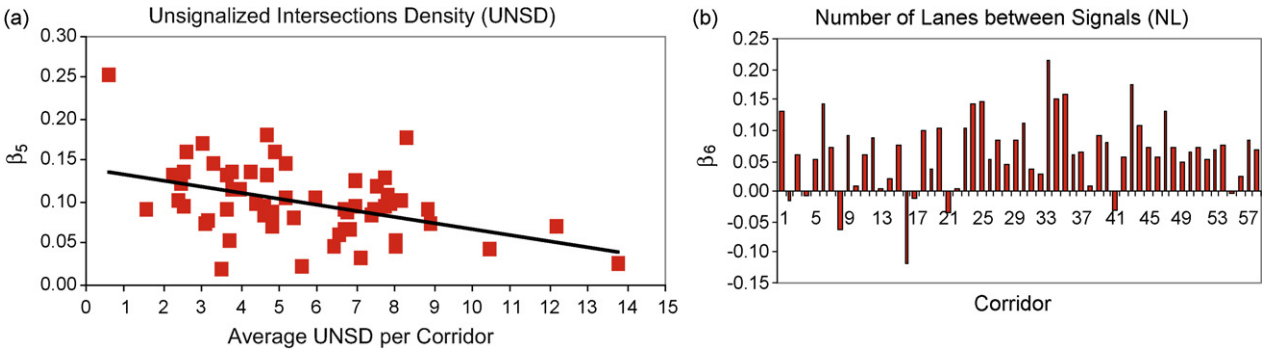

Fig. 1. (a) Scatter plot of β_5 vs. average UNSD per corridor and (b) bar chart of β_6 by corridor.

centages were 95%, 89.9% and 89.9%, respectively. On the remaining corridors, accident frequency is expected to decrease; reflecting heterogeneity across corridors.

As mentioned above, the number of lanes was found to be significant under PLN and M1, but not under M2. This finding maybe due to the fact that the corridor effects associated with traffic volumes, crosswalks density and business land use seem to be related to the number of lanes. Such multicollinearity appears to negate the need to include NL in M2.

In addition to the overall estimates of $\beta = (\beta_0, \beta_1, \ldots, \beta_6)$, provided in Table 3, there were 58 estimates of β corresponding to the 58 corridors. Such estimates can be used for corridor-level inference using (8). The nature of these regression coefficients is illustrated in Fig. 1.

For UNSD, the average values were computed for each of the 58 corridors and the 58 regression coefficients β_5 are plotted against these averages in Fig. 1(a), which shows that: (i) the corridors' estimates of β_5 varied between 0.02 and 0.25; (ii) the corridors varied in terms of UNSD as the average per corridor ranged from 0.6 to 13.8; and (iii) there is a tendency for the corridors' regression coefficients to decrease as the average UNSD increases. The tradeoff suggested by (iii) is logical in terms of goodness-of-fit. The same pattern has occurred for $\beta_1, \ldots, \beta_4$. However, while the correlation coefficient between β_5 and average UNSD per corridor was -0.44, which is significant at $p = 0.001$, the corresponding correlations for the other regression coefficients were weaker and not significant.

Fig. 1(b) displays the regression coefficients β_6 of the number of lanes by corridor. Not only there were 8 corridors with negative estimates, but also there were considerable variability as the coefficients ranged from -0.120 to 0.215.

6. Conclusions and future research

This paper compared the traditional PLN model with two extended PLN models using a data set for 392 urban arterials in the city of Vancouver, BC, that were clustered into 58 corridors. To assess the effects of incorporating corridor variation with alternate specifications, two models that account for corridor variation through the variance function (M1) and the mean function (M2) were considered.

A variety of covariates representing segment length, AADT, crosswalks density, business land use, unsignalized intersection density and the number of lanes between signals were found to significantly influence accident frequencies. The M2 model provided the best fit, representing a considerable improvement over PLN, while fitting the data a little better than M1. Under M2, the covariates resulted in random parameters and thereby their effects on accident frequency were found to vary significantly across corridors.

The results of this paper provide some strong evidence for the benefit of clustering road segments into rather homogeneous groups (e.g., corridors) and incorporating random corridor parameters in accident prediction models. Apart from the improvement in the goodness of fit, such an approach can be used to gain new insights into how the covariates affect accident frequencies, and to account for heterogeneity due to unobserved road geometrics, traffic characteristics, environmental factors and driver behavior. As well, the inclusion of corridor effects in the mean function could explain enough variation that some of the model covariates would be rendered non-significant and thereby the inference would be affected as well.

The results presented in this paper are based on a single dataset. Further research with different datasets is required to confirm the paper's findings. The work in this paper could be extended to compare other ways of clustering road segments, e.g., districts, municipalities, zones.

References

Aguero-Valverde, J., Jovanis, P.P., 2009. Bayesian multivariate Poisson log-normal models for crash severity modeling and site ranking. In: Presented at the 88th Annual Meeting of the Transportation Research Board.

Anastasopoulos, P.Ch., Mannering, F., 2009. A note on modeling vehicle accident frequencies with random-parameters count models. Accident Analysis and Prevention 41 (1), 153–159.

Bedrick, E.J., Christensen, R., Johnson, W., 1996. A new perspective on priors for generalized linear models. Journal of the American Statistical Association 91, 1450–1460.

Brooks, S.P., Gelman, A., 1998. Alternative methods for monitoring convergence of iterative simulations. Journal of Computational and Graphical Statistics 7, 434–455.

Cheng, W., Washington, S.P., 2005. Experimental evaluation of hotspot identification methods. Accident Analysis and Prevention 37, 870–881.

El-Basyouny, K., Sayed, T., 2006. Comparison of two negative binomial regression techniques in developing accident prediction models. Transportation Research Record 1950, 9–16.

El-Basyouny, K., Sayed, T., 2009a. Urban arterial accident prediction models with spatial effects. Transportation Research Record, in press.

El-Basyouny, K., Sayed, T., 2009b. Collision prediction models using multivariate Poisson-lognormal regression. Accident Analysis and Prevention 41 (4), 820–828.

Gelman, A., Meng, X., Stern, H., 1996. Posterior predictive assessment of model fitness via realized discrepancies. Statistica Sinica 6, 733–807.

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.

Heydecker, B.G., Wu, J., 2001. Identification of sites for accident remedial work by Bayesian statistical methods: an example of uncertain inference. Advances in Engineering Software 32, 859–869.

Hinde, J., Demetrio, C.G.B., 1998. Over-dispersion: model and estimation. Computational Statistics and Data Analysis 27 (2), 151–170.

Kim, H., Sun, D., Tsutakawa, R.K., 2002. Lognormal vs. gamma: extra variations. Biometrical Journal 44 (3), 305–323.

Kumara, S.S.P., Chin, H.C., 2006. Disaggregate models to examine signalized intersection crash frequencies. In: Presented at the 85th Annual Meeting of the Transportation Research Board, Washington, DC.

Lee, J., Mannering, F., 2002. Impact of roadside features on the frequency and severity of run-off-road accidents: an empirical analysis. Accident Analysis and Prevention 34 (2), 349–361.

Li, W., Carriquiry, A., Pawlovich, M., Welch, T., 2008. The choice of statistical models in road safety countermeasures effectiveness studies in Iowa. Accident Analysis and Prevention 40 (4), 1531–1542.
Lord, D., Miranda-Moreno, L.F., 2007. 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. In: Presented at 86th Annual Meeting of the Transportation Research Board, Washington, DC.
Lord, D., Persaud, B., 2000. Accident prediction models with and without trend: application of the generalized estimating equation. Transportation Research Record 1717, 102–108.
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., Guikema, S.D., Geedipally, S.R., 2008. Application of the Conway–Maxwell–Poisson generalized linear model for analyzing motor vehicle crashes. Accident Analysis and Prevention 40, 1123–1134.
Lord, D., Park, Y.-J., 2008. Investigating the effects of the fixed and varying dispersion parameters of Poisson-gamma models on empirical Bayes estimates. Accident Analysis and Prevention 40 (4), 1441–1457.
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. Accident Analysis and 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 and Prevention 39 (1), 53–57.
Lunn, D.J., Thomas, A., Best, N., Spiegelhalter, D., 2000. WinBUGS—a Bayesian modelling framework: concepts, structure, and extensibility. Statistics and Computing 10, 325–337.
Ma, J., Kockelman, K.M., Damien, P., 2008. A multivariate Poisson-lognormal regression model for prediction of crash counts by severity, using Bayesian methods. Accident Analysis and Prevention 40, 964–975.
Malyshkina, N.V., Mannering, F.L., Tarko, A.P., 2009. Markov switching negative binomial models: an application to vehicle accident frequencies. Accident Analysis and Prevention 41 (2), 217–226.
Miaou, S., Song, J., 2005. Bayesian ranking of sites for engineering safety improvements: decision parameter, treatability concept, statistical criterion, and spatial dependence. Accident Analysis and Prevention 37 (4), 699–720.
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.
Milton, J., Shankar, V., Mannering, F., 2008. Highway accident severities and the mixed logit model: an exploratory empirical analysis. Accident Analysis and Prevention 40 (1), 260–266.
Miranda-Moreno, L.F., Lord, D., 2007. Evaluation of alternative hyper-priors for Bayesian road safety analysis. In: Presented at the 87th Annual Meeting of the Transportation Research Board, Washington, DC.
Miranda-Moreno, L.F., Fu, L., Saccomanno, F.F., Labbe, A., 2005. Alternative risk models for ranking locations for safety improvement. Transportation Research Record 1908, 1–8.
Mitra, S., Washington, S., 2007. On the nature of over-dispersion in motor vehicle crash prediction models. Accident Analysis and Prevention 39, 459–468.
Park, E.S., Lord, D., 2007. Multivariate Poisson-lognormal models for jointly modeling crash frequency by severity. Transportation Research Record 2019, 1–6.
Poch, M., Mannering, F.L., 1996. Negative binomial analysis of intersection-accident frequency. Journal of Transportation Engineering 122 (2), 105–113.
Qin, X., Ivan, J.N., Ravishankar, N., 2004. Selecting exposure measures in crash rate prediction for two-lane highway segments. Accident Analysis and Prevention 36 (2), 183–191.
Qin, X., Ivan, J.N., Ravishanker, N., Liu, J., 2005. Hierarchical Bayesian estimation of safety performance functions for two-lane highways using Markov Chain Monte Carlo modeling. Journal of Transportation Engineering 131 (5), 345–351.
Sawalha, Z., Sayed, T., 2006. Traffic accidents modeling: some statistical issues. Canadian Journal of Civil Engineering 33 (9), 1115–1124.
Sawalha, Z., Sayed, T., 2001. Evaluating safety of urban arterial roadways. Journal of Transportation Engineering 127 (2), 151–158.
Schluter, P.J., Deely, J.J., Nicholson, A.J., 1997. Ranking and selecting motor vehicle accident sites by using a hierarchical Bayesian model. The Statistician 46 (3), 293–316.
Shankar, V.N., Albin, R.B., Milton, J.C., Mannering, F.L., 1998. Evaluating median crossover likelihoods with clustered accident counts: an empirical inquiry using the random effects negative binomial model. Transportation Research Record 1635, 44–48.
Shankar, V.N., Chayanan, S., Sittikariya, Shyu, M.-B., Juvva, N.K., Milton, J.C., 2004. Marginal impacts of design, traffic, weather, and related interactions on roadside crashes. Transportation Research Record 1897, 156–163.
Spiegelhalter, D., Thomas, A., Best, N., Lunn, D., 2005. WinBUGS User Manual. MRC Biostatistics Unit, Cambridge. Available from http://www.mrc-cam.ac.uk/bugs.
Spiegelhalter, D.J., Best, N.G., Carlin, B.P., Van der Linde, A., 2002. Bayesian measures of model complexity and fit. Journal of the Royal Statistical Society B 64, 1–34.
Stern, H.S., Cressie, N., 2000. Posterior predictive model checks for disease mapping models. Statistics in Medicine 19, 2377–2397.
Tunaru, R., 2002. Hierarchical Bayesian models for multiple count data. Austrian Journal of Statistics 31 (3), 221–229.
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 (2), 275–289.
Winkelmann, R., 2003. Econometric Analysis of Count Data. Springer, Germany.
Ye, X., Pendyala, R.M., Washington, S.P., Konduri, K., Oh, J., 2009. A simultaneous equations model of crash frequency by collision type for rural intersections. Safety Science 47 (3), 443–452.