Available online at www.sciencedirect.com

ScienceDirect

Transportation Research Procedia 17 (2016) 626 – 635

Transportation Research Procedia

www.elsevier.com/locate/procedia

11th Transportation Planning and Implementation Methodologies for Developing Countries, TPMDC 2014, 10-12 December 2014, Mumbai, India

Development of comprehensive crash models for four-lane divided highways in heterogeneous traffic condition

Praveen Vayalamkuzhi[a] and Veeraragavan Amirthalingam[b]*

[a]*Doctoral Research Scholar, Department of Civil Engineering, Indian Institute of Technology Madras, Chennai 600 036, INDIA*
[b]*Professor, Department of Civil Engineering, Indian Institute of Technology Madras, Chennai 600 036, INDIA*

Abstract

Traffic safety is of prime concern worldwide. Highway geometry should be designed for vehicle safety and efficiency. Several researches have been carried out to identify the factors contributing to road crashes and for finding measures to reduce the crash rate. One of the critical gaps in the management of highway safety is the lack of a reliable method for estimating the safety of an existing roadway with, widely varying road geometrics and vehicle mix. The focus of this work is mainly to quantify the relationship between geometric design characteristics and level of safety of intercity highways under heterogeneous traffic conditions. Study was carried out in a four-lane divided rural highway in Tamil Nadu, India and a relationship was established using statistical modeling technique. Crash Prediction Models (CPM) were developed by Poisson regression, Poisson-gamma and negative binomial modeling approach for three categories, namely, current (i^{th}) segment, with preceding (i-1^{th})segment and with succeeding and preceding (i+1&-1^{th}) segments. Results showed the significance of identified variables and the effect of preceding and succeeding segments on the current segment in the case of CPM. Attempts were also made to develop operating speed models for curve and tangent elements. From the developed models, the effect of contiguous element on the operating speed could be understood.

Peer-review under responsibility of the Department of Civil Engineering, Indian Institute of Technology Bombay

*Keywords:*Crashes; Geometric characteristics;Heterogeneous traffic; Statistical modeling; Poisson regression; Negative binomial

1. Introduction

Road crashes are complex events and are influenced by many factors such as road geometric design, traffic volume and composition, speed, weather, motivation for travelling, driver's physical and mental conditions

* Corresponding author. Tel.: +91-44-2257-4272; fax: +91-44-2257-4252
E-mail address: av@iitm.ac.in

Peer-review under responsibility of the Department of Civil Engineering, Indian Institute of Technology Bombay
doi:10.1016/j.trpro.2016.11.117

(*American Association of State Highway Transport Officials, 2004*). For identifying the accident prone location or black spot along a roadway for the purpose of detailed engineering studies and to prioritize the road networks for implementation of safety measures, consistency of the design needs to be examined. Highway geometry should be designed for traffic safety and efficiency. As concerns traffic on Indian roads, it is highly heterogeneous in character and is composed of slow and fast-moving vehicles. As per World Health Organization (WHO) report 2013, India loses more than 100,000 lives due to road traffic crashes every year with a fatality rate of 18.9 deaths per 1,00,000 populations. Approximately half of all deaths on the country's roads are among vulnerable road users - motorcyclists, pedestrians and cyclists (*Global Status Report on Road Safety – Time for Action, 2009*). Recent statistics show that, there were 4,86,476 reported accidents during 2013, out of which 1,37,572 people were killed and 4,94,893 people were injured (*Ministry of Road Transport and Highways, Govt. of India, 2014*). The accident severity viz., persons killed per 100 vehicles is 28.2. According to MoRTH 2012, Tamil Nadu stands at second position with a share of 11.7% in total number killed in road crashes in the year 2012. Road crash statistics in 2013 shows that, 66,238 road accidents were occurred, which resulted in 15,563 fatalities (*Transport department, Government of Tamil Nadu, 2014*).

Several studies were conducted to evaluate the influence of speed on the safety of roadways. According to AASHTO - Geometric Design of Highways and streets, "the safest speed for any highway depends on design features, road conditions, traffic volumes, weather conditions, roadside development, spacing of intersecting roads, cross-traffic volumes, and other factors". Most of the studies considered the operating speed as the 85^{th} percentile speed of those vehicles travelling on the roadway and found speed-related crashes are more likely to occur at mid-block than an intersection (*Solomon, 1964; Lamm et al., 1990; Liu and Chen, 2006; Lu, 2006*). Models were also developed for studying the dependence of crash rates on speed and geometric characteristics, which showed that they are not linearly related (*Garber and Ehrhart, 2000*). Vehicle speed could be related to traffic safety in two ways: (1) greater a vehicle's velocity, the lesser will be the time available for the driver to react to a hazard under the presence of other motorists, bicyclists, or pedestrians. If this relationship exists, it would be expressed in relative incidence of crashes at different speeds and (2) due to the physical relationship of mass and speed to energy, it would be possible to express the relative severity of crashes at different speeds (*Federal Highway Administration, 2000*). Most of the reported studies were carried out on undivided rural highways and only a few on divided highways. Research related to geometric characteristics showed that variables like, radius of the curve and super elevation have a significant effect on the safety of roadways. Radius of curves was identified as one of the significant variables while defining the effect of horizontal and vertical curves on road crashes and also while estimating the speeds on rural highways. Crash rate was found to increase significantly when radii are below 200 m. Super elevation, degree of curve, shoulder width and average daily traffic also contributes to road crashes (*FHWA, 2000; Aram, 2010*). Researchers also concluded that roads with heavier traffic volume, more road lanes and higher speed limits tend to have more severe crashes (*Ma, 2010*). Studies related to divided roadway showed that the presence of raised medians reduces crash rates up to 15 % and also reduces the severity of crashes when modelled accidents using Poisson and negative binomial regression (*Frawley et al., 2005; Sawalha and Sayed, 2003*).

As far as Indian scenario is concerned, data collection and their availability are major issues, as systematic data collection and road crash database are not in place. Few studies were carried out which focused on the development of Accident Prediction Models (APM) for a particular type of road, mainly undivided rural roads. Regression models were found appropriate for evaluating road crashes (*Bhagat, 2008; Robert, 2006; Ramesh and Kumar, 2011; Dinu, 2012*). Studies were carried out to develop APMs with the stepwise introduction of identifying predictor variables based on their magnitude of the coefficient of determination. This resulted in developing multiple linear regression, Poisson, Negative Binomial, exponential and logistic regression models for evaluating total accidents, injuries, fatalities and property damage, casualties and accident rate (*Robert, 2006*). APMs were developed using a Poisson model with random coefficients to predict single vehicle and multi-vehicle crashes for evaluating the safety performance of two-lane undivided rural roads (*Dinu, 2012*). From the literature, it was found that most of the research works are based on the homogeneity and lane disciplined traffic conditions. Hence, the results of the studies may not be directly applicable to the heterogeneous traffic such as the one prevailing in developing countries like India. As discussed earlier, the traffic conditions in India are highly heterogeneous in nature, where the traffic comprises vehicles with diverse static and dynamic characteristics and all such vehicles use the same right-of-way without any physical segregation. Thus, the vehicles of the said heterogeneous traffic, under high-volume conditions, move on Indian roads by sharing the available road space without sufficient lateral as well as longitudinal clearances. The lane-less movement further adds to the complexity of analyzing/modeling mixed traffic.

There are only a limited number of studies reported on the subject matter under Indian conditions mainly because of non-availability of required data and inapplicability of the available models for the analyses.

The present study attempted to investigate the safety effects of four-lane divided rural intercity highway in heterogeneous traffic condition by developing comprehensive crash models. Study mainly focused on analysing the data pertaining to crash frequency, geometric and traffic characteristics of the study corridor by developing crash prediction models (CPMs) after classifying the highway into segments. Further the study also explores to establish relationship between the contiguous elements (curve and tangent) of a horizontal alignment by developing the operating speed models.

2. Data description

The study was carried out on a four-lane divided rural highway in Tamil Nadu, India with heterogeneous traffic condition. In order to establish the relationship between geometric and traffic characteristics with crash occurrence, field studies were carried out on the highway to identify the factors influencing crash occurrence. Data pertaining to road geometry, traffic characteristics, accident and environmental condition were collected. Roadway characteristics were collected using Hawkeye series 2000, automated data collection equipment and were extracted using the Hawkeye processing toolkit by standard rating form. Hawkeye, an automated data collection vehicle developed by the Australian Road Research Board (ARRB), is a system fully integrated and the outputs are linked to both spatial (GPS) and linear references. It has three high-resolution digital cameras by which digital images were collected with a 150 to 180 degree field of view (centered in the travel lane). Geo-reference data**,** including distance along the road (from an established start point) plus latitude and longitude coordinates were also recorded. In addition to these, it will provide automated measurements of radius of curvature for horizontal curves and gradient for vertical alignment. Pavement condition data were also collected via a digital laser based profile beam fixed to the front of the vehicle. Data extracted includes gradient, cross slope, horizontal curvature, vertical curvature, pavement condition, width and type, shoulder condition, width and type (as shown in Figure 1).

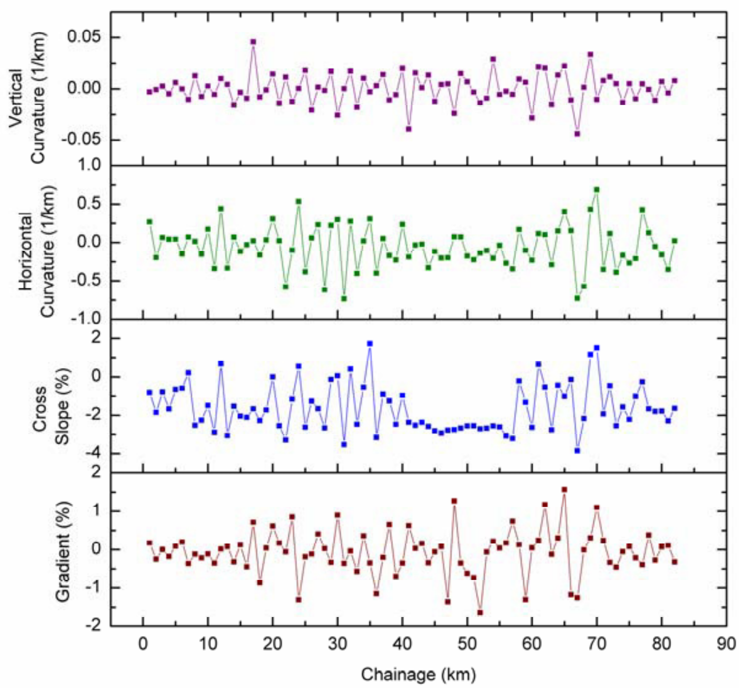

FIGURE 1. Geometric profile of the study corridor

The other important data required for the study are the historical crash data. This data was collected from records for a period of four years, i.e., 2009 to 2012, which includes a number of crashes per year, location, type, cause, severity and time of the day. In order to evaluate the relationship between crash and exposure factor (traffic volume), it is really essential to collect the traffic characteristics such as classified volume count, traffic composition, operating speed along the roadway and also at contiguous elements. Figure 2 shows the crash frequency along the study corridor with the operating speed. To have a clear picture of the heterogeneous traffic scenario that exists on the study corridor, it has been shown in Figure 3. Other features of the roadway such as access density, presence of median, type and width, were also examined during field studies.

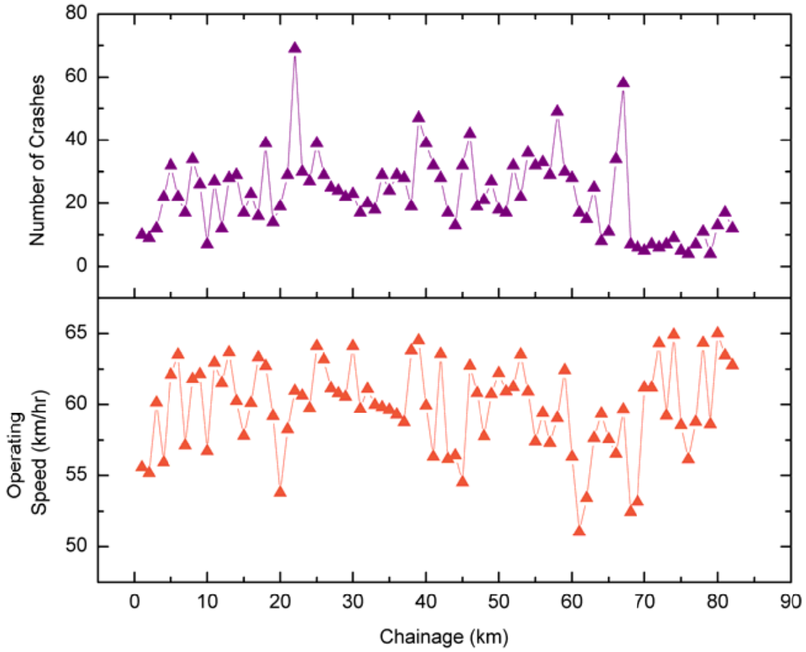

FIGURE 2. Number of crashes along the study corridor with operating speed

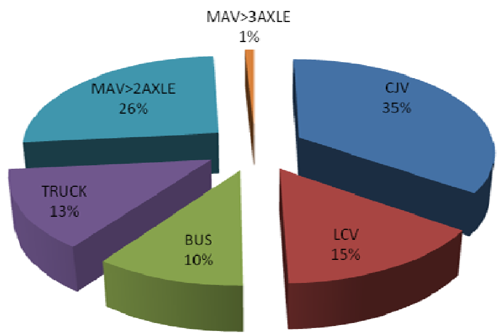

FIGURE 3. Vehicle composition in the study corridor

3. Analysis of data

The present work attempts to develop comprehensive crash models for divided rural highways operating under heterogeneous traffic condition, by quantifying the influence of geometric design and traffic characteristics. To examine the influence of the parameters collected, preliminary analysis was carried out in the form of correlation analysis and scatter plot. These analyses help in identifying the input parameters to the model as well as the model form to be developed. Predictive models were developed for the study period using statistical modeling approach. In order to develop models, the study corridor was segmented as it plays an important role in safety analysis. Generally, segmentation is carried out in two fundamental ways; with respect to length of highway and composition (*Lord et al., 2005; Cook et al., 2011; Cafiso et al., 2013*). In the present study also, an attempt was made to segment the study corridor based on the above two fundamental ways; (1) fixed length - one kilometer and (2) exposure - Annual Average Daily Traffic (AADT)

3.1. Model development

Based on the preliminary analysis, parameters identified were gradient (G), cross slope (CS), horizontal curvature (HC), vertical curvature (VC), operating speed (OS), average daily traffic (AADT) and presence of access point (AP) and median opening (MO) from three major factors (exposure, geometry, and context factors), for i^{th}, $i-1^{th}$ and $i+1$&-1^{th}segment. To explore the relationship between the number of road crashes as dependent variable and identified parameters as independent variables, models were developed using Generalized Linear Modeling (GLM) approach with Poisson regression, Poisson-gamma and a Negative binomial error structure, because of the non-negative integer nature of crash occurrence. These models account for a unique distribution of count data by preserving the validity of statistical analysis (*Miller and Freund, 1977*). Fitness of the regression models to the observed data was assessed by likelihood ratio and Akaike's Information Criteria (AIC) (*Naderan and Babaei, 2011*).

3.1.1. Poisson regression

Poisson regression is a fundamental method for analyzing data that have non-negative integer values, which is having wide range of application on rare event count data (*Miller and Freund, 1977*). The link function for Poisson regression is usually a log transformation of the mean. It assumes that the mean equals the variance, referred to as equi-dispersion property of Poisson. Eqn. 1 and 2 shows the general form of Poisson regression.

$$P\left(y_i\right) = \frac{exp\left(-\lambda_i\right)\lambda\, y_i}{y_i!} \tag{1}$$

Where λ_i *is* the Poisson parameter of the number of accidents, which is equal to E[y_i]. Poisson regression models are estimated by specifying the Poisson parameter $\lambda_i$as a function of explanatory variables.

$$Y = \exp(\, Intercept + \beta_1\ x_1 + \beta_2\ x_2 + \beta_3\ x_3\,) \tag{2}$$

The log-likelihood function *LL(β)* (as shown in Eqn. 3) can be used for estimating the model and is given by

$$LL(\beta) = \sum_{i=1}^{n}\left[-exp\left(\beta X_i\right) + y_i \beta X_i - L_n\left(y_i!\right)\right] \tag{3}$$

To assess the fit of the Poisson regression model to the observed data, likelihood ratio, Akaike's Information Criteria (AIC) and so on are generally used.

3.1.2. Poisson gamma regression

Poisson gamma is one other model form for count data, which overcome the drawback of over-dispersion in the Poisson regression with the introduction of dispersion parameter and the probability model is shown in Eqn. 4.

$$\text{Prob}[y_i = m] = G(\alpha_m, \lambda_i) - G(\alpha_m + \alpha, \lambda_i) \tag{4}$$

Where,

$\lambda_i = exp(\beta' \mathbf{x}_i)$ and $G(\alpha_m, \lambda_i) = 1$, if $m = 0$

'α' is the dispersion parameter; under-dispersion if $\alpha > 1$, over-dispersion if $\alpha < 1$, and equi-dispersion if $\alpha = 1$, which reduces the gamma probability to the Poisson model.

3.1.3. Negative Binomial (NB) regression

Negative binomial regression is also used for modeling over-dispersed count data. It does not assume an equal mean and variance and is particularly suitable when the variance is greater than the mean (Greene (2008), Cummings (2000)). The form of the model is same as that for Poisson regression, with an additional parameter 'α' (dispersion parameter) to model the over-dispersion. The selection between these two models is dependent on the value of the parameter 'α'. The likelihood function $L(\lambda_i)$ is given in Eqn. 5.

$$L\left(\lambda_i\right) = \Pi_i \frac{\Gamma\left((1/\alpha) + y_i\right)}{\Gamma(1/\alpha)\, y_i!} \left[\frac{1/\alpha}{\left(1/\alpha\right) + \lambda_i}\right]^{1/\alpha} \left[\frac{\lambda_i}{\left(1/\alpha\right) + \lambda_i}\right]^{y_i} \tag{5}$$

3.2. Goodness of fit of model

To access the quality of the models developed, goodness of fit is evaluated. Likelihood ratio chi^2statistic is chi^2 distributed with degrees of freedom equal to the difference in the numbers of parameters in the restricted (β_R) and unrestricted (β_u) model is a goodness of test. Akaike's Information Criteria (AIC) also helps in identifying the best model and is shown in Eqn. 6.

$$\text{AIC} = \frac{\left[-2L_n\left(\beta_u\right) + 2b\right]}{N} \tag{6}$$

Where $L_n(\beta_u)$ is the log-likelihood value of the model, b is the number of variables and N is the number of observations. The model which has the minimum AIC value is considered as the best model (*Miller and Freund, 1977; Naderan and Babaei, 2011; Greene, 2008: Cummings, 2000*).

3.3. Operating speed model of contiguous element

To establish a relationship between the operating speed (85th percentile) and the geometric design parameters, models were developed with an operating speed at mid-curve and at a tangent as dependent variables and geometric design parameters as independent variables using multiple linear regression. The general form is shown in Eqn. 7.

$$Y = \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_4 X_4 + \ldots + \beta_0 \tag{7}$$

Operating speed at the mid of the curve and the tangent were taken as dependent variables and the independent variables considered after preliminary analysis were radius of curve (R), tangent length (TL), cross slope (CS), gradient (G), inverse radius (IR), deflection angle (DA), and curvature change rate (CCR). These models will help in calculating operating speeds on a curve and tangent which in turn will help in forecasting crashes.

4. Results and summary

4.1. Crash Prediction Models

In order to understand the effect of identified variables, models were developed by count data regression approach. Models were developed based on three approaches, namely Poisson regression, Poisson gamma and Negative binomial. In the case of fixed length segmentation approach, models were developed considering three categories *viz.***,** i^{th} (current segment), $i\text{-}1^{th}$ (preceding segment) and $i+1\&\text{-}1^{th}$ (current, succeeding and preceding segments), whereas in the case of AADT based segmentation, only i^{th} (current segment) is considered. Results showed that most of the identified independent variables are statistically significant at 1%, 5% and 10% level, when the different modeling approach is concerned. For each three categories, fifteen models were developed using three modeling approaches. The log-likelihood (LL), restricted log likelihood, LL ratio, AIC, over-dispersion and dispersion parameter (α) values for fixed length and AADT based segmentations are shown in Table 1 and 2 respectively.

Table 1. Model results from fixed length based segmentation approach

Description		Poisson Regression			Poisson-Gamma			Negative Binomial		
		i^{th}	$i\text{-}1^{th}$	$i+1\&\text{-}1^{th}$	i^{th}	$i\text{-}1^{th}$	$i+1\&\text{-}1^{th}$	i^{th}	$i\text{-}1^{th}$	$i+1\&\text{-}1^{th}$
Log likelihood	Upper	769.08	718.82	640.48	566.39	551.78	529.35	559.08	544.19	522.71
	Lower	323.12	285.67	274.38	313.65	285.35	274.35	308.63	284.48	274.25
Restricted Log likelihood	Upper	1107.5	1097.4	1087.5	769.08	718.82	640.48	769.08	718.82	640.48
	Lower	347.8	343.23	338.91	323.12	285.66	274.38	323.12	285.67	274.38
Log likelihood Ratio	Upper	0.31	0.34	0.41	0.26	0.23	0.17	0.27	0.24	0.18
	Lower	0.05	0.09	0.13	0.03	0.01	0.01	0.04	0.01	0.01
AIC	Upper	9.17	8.78	8.05	6.79	6.81	6.72	6.71	6.72	6.647
	Lower	3.92	3.62	3.64	3.82	3.63	3.65	3.76	3.62	3.65
Over dispersion	Upper	4.66	6.05	6.09	-	-	-	-	-	-
	Lower	1.89	0.52	0.60	-	-	-	-	-	-
Dispersion parameter	Upper	-	-	-	0.54	0.61	1.03	0.32	0.62	0.23
	Lower	-	-	-	0.15	0.30	0.23	0.18	0.13	0.02

Table 2. Model results from AADT based segmentation approach

Description		Poisson Regression	Poisson-Gamma	Negative Binomial
Log likelihood	Upper	-173.90	-150.19	-117.84
	Lower	-66.32	-66.24	-66.25
Restricted Log likelihood	Upper	-871.38	-173.90	-173.90
	Lower	-93.35	-66.32	-66.32
Log likelihood Ratio	Upper	0.80	0.39	0.38
	Lower	0.29	0.01	0.01
AIC	Upper	14.14	12.40	9.91
	Lower	6.10	6.17	6.18
Over dispersion	Upper	4.51	-	-
	Lower	0.51	-	-
Dispersion parameter	Upper	-	0.90	0.0002
	Lower	-	0.08	0.65

Out of the total models developed, one model from each category and modeling approach has been identified has the best model from fixed length segmentation based on goodness of fit values (section 3.2) and is shown in Tables 3, 4 and 5.

Table 3. Best models based on fixed length based segmentation while considering i^{th} segment

Description	Models
Poisson Regression	$Y = exp\left(23.543 - 0.289G - 0.041CS + 0.037OS + 0.095AP - 3.454LnAADT\right)$
Poisson Gamma	$Y = exp\left(28.189 - 0.339G + 0.048OS - 4.292LnAADT\right)$
Negative Binomial	$Y = exp\left(22.639 - 0.183G + 0.036OS + 0.142AP - 3.315LnAADT\right)$

Table 4. Best models based on fixed length based segmentation while considering i^{th} and $i\text{-}1^{th}$ segment

Description	Models
Poisson Regression	$Y = exp\begin{pmatrix} 26.937 - 0.288G - 0.051CS + 0.032OS + 0.067AP - 1.758LnAADT \\ - 0.045CSpr - 0.221MOpr - 2.207LnAADTpr \end{pmatrix}$
Poisson Gamma	$Y = exp\left(31.232 - 0.331G + 0.039OS - 0.275MOpr\right)$
Negative Binomial	$Y = exp\left(26.354 - 0.184G + 0.036OS - 0.246MOpr\right)$

Table 5. Best models based on fixed length based segmentation while considering i^{th}, i-1 &i+1 th segment

Description	Models
Poisson Regression	$Y = exp\begin{pmatrix} 20.164 - 0.218G - 0.217HC - 4.626VC + 0.028OS + 0.109AP - 0.043Gpr \\ - 0.331HCpr - 0.260MOpr - 1.233LnAADTpr - 0.163Gsu + 0.052CSsu \\ - 0.341HCsu - 3.911VCsu + 0.231APsu - 0.338MOsu \end{pmatrix}$
Poisson Gamma	$Y = exp\begin{pmatrix} 22.585 - 0.241G - 5.530VC + 0.033OS - 0.386HCpr - 0.309MOpr - 0.178Gsu \\ + 0.264APsu - 0.391MOsu \end{pmatrix}$
Negative Binomial	$Y = exp\left(24.211 - 0.151G + 0.031OS - 0.261MOpr - 0.107Gsu + 0.219APsu - 0.342MOsu\right)$

Crash prediction models were also developed by the AADT based segmentation approach as explained in Section 3. In this method, best model while considering i^{th} segment alone was identified for the three count data modeling approach as shown in Table 6.

Table 6. Best models based on AADT based segmentation for i^{th} segment

Description	Models
Poisson Regression	$Y = exp\begin{pmatrix} 12.645 - 0.277G - 0.084CS - 0.674HC - 0.028OS + 0.083AP - 0.056MO \\ - 1.151LnAADT + 0.135LENG \end{pmatrix}$
Poisson-Gamma	$Y = exp\left(4.671 + 0.189AP\right)$
Negative Binomial	$Y = exp\left(9.604 + 0.127AP\right)$

Different methods of count data modeling approach were used for model development in order to eliminate the drawbacks of one method and to find the best model to predict the crash for a four-lane divided highway having similar roadway and traffic scenario. The explanatory variable which has a positive coefficient in the model reflects that each one unit increase results in an increase in the number of crashes and vice versa. It was observed that the operating speed in all the models is positively significant. Other variables which have positive effects were access point and cross slope. Few other explanatory variables considered for modeling such as gradient, curvature (inverse of the radius), median opening, and AADT had a negative effect on crash occurrence. Model evaluation was carried out using the likelihood (LL) ratio and AIC values. These results were found to be within the limit, *i.e.*, for LL ratio, a model having rho square value close to 1 and the lower AIC value is identified.

4.2. Operating speed models for curve section

In the case of model development of curve section, operating speed at the midpoint of the curve ($OS_{(MC)}$)is considered as the dependent variable and the independent variables used are radius (R), tangent length (TL), operating speed at preceding tangent ($OS_{(T)}$), length of curve (LC), degree of curvature (DC), deflection angle (DA), cross slope (CS) and gradient (G). For modeling, total of 202 curves were considered. Different combinations were developed and the best one identified based on the R square value is as shown in Eqn. 8.

$$OS_{(MC)} = 15.138 – 0.684G – 0.285CS + 0.718\ OS_{(T)}\ (R^2 - 0.64) \tag{8}$$

4.3. Operating speed models of tangent section

Models were developed for tangent section ($OS_{(T)}$) by considering operating speed at the preceding tangent of the curve as the dependent variable and the independent variables used are tangent length (TL), radius of succeeding curve (R), curve length (L_C), operating speed at mid-curve($OS_{(MC)}$), cross slope (CS) and gradient (G). A total of 226 tangent sections are used for model development. Eqn. 9 shows the best model based on the R-square value.

$$OS_{(T)} = 13.327 + 0.786OS_{(MC)}\ (R^2 - 0.61) \tag{9}$$

From the models developed, it was observed that the operating speed on the preceding tangent is influenced by the operating speed of curve section and vice versa with an R^2 value of 0.64 and 0.61 respectively, because of the influence of geometric variables at the tangent. Other variables which are statistically significant are tangent length, inverse radius and cross slope.

Operating speed models developed using MLR to evaluate the influence of contiguous elements explained clearly the relationship, i.e., operating speed at preceding tangent have a significant effect on the operating speed at curve and vice versa in addition to other geometric characteristics as shown in Eqns. 8 and 9.

5. Conclusions

The influence of geometric design characteristics and traffic characteristics on the level of safety on a four-lane divided rural highway in India, operated under heterogeneous traffic condition was studied. Count data modelling approach was used for developing crash models, because the crash occurrences are rare and random in nature. Geometric characteristics of highway were observed to be varied, which in turn affect the operating speed on the highway, thereby affecting the level of safety. Operating speed (85^{th} percentile speed) of highway, one of the explanatory variables, which has a positive effect on the crash occurrence, was found to be the most influential factor in all the models developed. It was also found that geometric characteristics of the roadway such as gradient, median opening, access point density, curvature, and traffic volume of preceding and succeeding segments influence the crash occurrence, in addition to the effect of the segment under study. The relationship between contiguous elements was established by developing operating speed models for curve and tangent section and was observed that operating speed of one element is associated with the other element.

Developed crash prediction models (CPMs) also gave insight that not only the segment under study have an influence on crash, but also preceding and succeeding segments have a significant role in crash occurrence. In this way, developed models will be really helpful for the transportation engineers in the design of highways operating under heterogeneous traffic.

Acknowledgements

The authors would like to thank the Ministry of Human Resource Development (MHRD) for funding and Indian Road Survey and Management Pvt. Ltd. (IRSM), Chennai, India, for the help in data collection.

References

American Association of State Highway Transport Officials, *A Policy on Geometric Design of Highways and Streets,* PE Exam Edition, ISBN Number: 1-56051-263-6, 2004.

Aram, A., 2010, Effective safety factors on horizontal curves of two-lane highways, *Journal of Applied Sciences*, Vol. 10, No. 2, pp.2814 – 2822.

Bhagat, P. K., 2008, Road accident analysis of Patna – A case study, *Indian Highways*, Indian Road Congress, Vol. 36, No. 8, pp.13 - 20.

Cafiso, S., Agostino, D., and Persaud,B., 2013, Investigating the influence of segmentation in estimating safety performance functions for roadway sections, Presented at the 92ndTransportation Research Board Annual Meeting, CD-ROM, 2013.

Cook, D., Souleyrette, R., and Jackson, J., 2011, Effect of road segmentation on highway safety analysis, Presented at the 90thTransportation Research Board Annual Meeting, CD-ROM, 2011.

Cummings, P., 2009, Methods for estimating adjusted risk ratios, *The Stata Journal*, Vol. 9, No. 2, 2009, pp. 175–196.

Department of Transport, Government of Tamil Nadu, www.tn.gov.in/departments/ transport.html, Accessed on 20-07-2014

Dinu, R. R., *Studies on safety performance of two-lane rural highways in heterogeneous traffic*, Thesis Dissertation, submitted to Indian Institute of Technology Madras, Chennai, 2012.

Federal Highway Administration, *Report on synthesis of safety research related to speed and speed management*, Report Number FHWA-RD-98-154, U.S. Department of Transportation, Washington, D.C, 1998.

Federal Highway Administration, *Speed prediction for two lane rural highways*, Report Number FHWA-RD 99-171, U.S. Department of Transportation, Washington, D.C., 2000.

Frawley, W. E., and Eisele, W. L., *Investigation of access point density and raised medians: Crash analysis and micro-simulation*. Report Number FHWA/TX-05/0-4221-P1. Texas Transportation Institute, Texas A&M University System, College Station, Texas, 2005

Garber, N. J., and Ehrhart, A. A., 2000, The effect of speed, flow, and geometric characteristics on crash frequency for two-lane highways, Presented at the 79th Transportation Research Board Annual Meeting, CD-ROM, 2000.

Greene, W., 2008, Functional forms for the negative binomial model for count data, *Economics Letters*, Vol. 99, pp. 585–590.

Lamm, R., Choueiri, E. M., and Mailaender, T., 1990, Comparison of operating speed on dry and wet pavements of two-lane rural highways, *Transportation Research Record*, Vol. 1280, pp. 199-207.

Liu, C., and Chen, C. L., *An analysis of speeding-related crashes: definitions and the effects of road environments*. Report Number DOT HS 811 090, National Highway Traffic Safety Administration, 2009.

Lord, D., Simon P. Washington, S. P., and Ivanc, J. N., 2005, Poisson, Poisson-gamma and zero-inflated regression models of motor vehicle crashes: balancing statistical fit and theory, Accident Analysis and Prevention, Vol. no. 37, pp. 35–46

Lu, M., 2006, Modelling the effects of road traffic safety measures, *Journal of Accident Analysis and Prevention,* Vol. – 38, pp. 507–517

Ma, M., Yan, X., Abdel-Aty, M., Huang,H., and Wang, X., 2010, Safety analysis of urbanarterials under heterogeneous traffic in Beijing, China, Presented at the 89th Transportation Research Board Annual Meeting, CD-ROM, 2010.

Miller, I., and Freund, J.E., *Probability* and Statistics for *Engineers*, 2nd *ed*. Prentice-Hall, Englewood Cliffs, N.J., 1977

Ministry of Road Transport and Highways (MoRTH) Transport Research Wing, Government of India, *Road Accidents in India 2013*,http://morth.nic.in/, Accessed in October 2014.

Naderan, A., and Babaei, M., 2011, Assessment of statistical models for aggregate crash prediction, Presented at the 90thTransportation Research Board Annual Meeting, CD-ROM, 2011.

Ramesh, A., and Kumar.M..2011, Road accident models for Hyderabad metropolitan city of India, *Indian Highways*, Indian Road Congress, Vol. 39, No. 7, pp. 17 - 28.

Robert V., R., *Studies on road safety problems under heterogeneous traffic flow*, Thesis Dissertation, Submitted to Bangalore University, 2006.

Sawalha, Z., and Sayed, T., 2003, Statistical issues in traffic accident modeling, Presented at the 82ndTransportation Research Board Annual Meeting, CD-ROM, 2003.

Solomon, D., 1964, Accidents on main rural highways related to speed, driver, and vehicle, Bureau of Public Roads (precursor to FHWA), Washington, DC, (Reprinted 1974)

World Health Organization, *Global Status Report on Road Safety- Time of action*, 2009

World Health Organization, *Global Status Report on Road Safety*, Switzerland, 2013.