European Transport \ Trasporti Europei (2016) Issue 62, Paper n° 1, ISSN 1825-3997

Safety Performance Evaluation of Multilane Highways Under Heterogeneous Traffic

Praveen Vayalamkuzhi [1], Veeraragavan Amirthalingam [2*]

[1] *Doctoral Research Scholar, Department of Civil Engineering,*
Indian Institute of Technology Madras, Chennai - 36, INDIA

[2] *Professor, Department of Civil Engineering,*
Indian Institute of Technology Madras, Chennai - 36, INDIA

Abstract

The traffic on the Indian highways is heterogeneous in nature. Vehicles with widely varying static and dynamic characteristics occupy the common carriageway space. The speed differentials between the vehicles of the same class and different classes significantly contribute to road crashes on the highways. The operating speeds of the different vehicle types vary widely. The design speed is different from the operating speeds observed on the highways. Design of the road geometrics for safety considering vehicles with widely varying static and dynamic characteristics is a challenging task to the designers. The present research work attempts to develop safety prediction models for inter-city highways catering to heterogeneous traffic considering the operating speeds of vehicles and relating it to crashes duly considering the roadway geometric elements. A typical four lane highway is considered in the research study. The data on roadway geometrics at different sections along the highway, the traffic details and other elements contributing to the road crashes are collected along with the road accident data for a period of four years. Statistical analysis was carried out to develop safety prediction models. The models may be used in assessment of safety performance of four lane rural highways under heterogeneous traffic conditions and in the calibration of Interactive Highway Safety Design Models for mixed traffic conditions.

Keywords: Traffic safety; Road crash; Heterogeneous traffic; Highway design; Regression modelling.

---

1. Introduction

According to World Health Organization (WHO 2013), over 1.2 million person are killed and 20 to 50 million person are injured annually in road crashes worldwide. As far as developing countries are concerned, 20 to 200 deaths occur per 10,000 motor vehicles, whereas only 2 to 5 deaths happen in the case of developed countries. India is no exception with over 1.3 million people gets killed in over 4.89 million road crashes every year. The increase in the number of vehicles has however, far outweighed the

---

[*] Corresponding author: Veeraragavan Amirthalingam (av@iitm.ac.in)

projected capacities and adaptive capabilities of the existing road infrastructure leading to deterioration in traffic performance and increasing difficulties in maintaining an acceptable and sustainable traffic-safety standards. Fatality rate, which is defined as the number of persons killed per 100,000 inhabitants per year, is 16.6 in 2014 in the country. Statistical data from India reveals that highways are facing a disparate rate of crashes, i.e., almost 53.7% of road crashes occurred on inter-city highways in the year 2014 when compared to urban roads (Transport Research Wing (TRW), Ministry of Road Transport & Highways, 2014). The heterogeneous traffic under high-volume conditions, share the available road space without sufficient safe lateral as well as longitudinal clearances. The lane-less movement further adds to the complexity of analyzing/modeling heterogeneous traffic. This motivates the need for a level of safety analysis under heterogeneous traffic conditions.

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, drivers' physical and mental conditions etc. Highway geometry should be designed for traffic safety and efficiency. Traffic-safety analysis was initially recognized in 1960's. Later several research works related to traffic safety analysis were performed by different research groups in different parts of the world by considering its seriousness to the public health, which resulted in different ways for evaluating safety performance. Safety Performance Function (SPF), is defined as the relationship between crash and the factors influencing crashes. SPFs are established in order to evaluate safety performance of highway or roadway under study. Recent development by Highway Safety Manual (HSM 2010) also describes SPF's for different classes of highways based on the data availability. In the case of design consistency evaluation methodology, each element of the road geometry and also the major traffic factors are considered individually in the analysis. Glennon *et* al. (1978) initially adopted this concept as the underlying principle of highway design and studied that the design inconsistency will affect the operating speed or leads to unsafe manoeuvre resulting in collision risk. Hence, emphasis has to be given to safety while designing highway geometrics.

The review of literature on traffic-safety analysis shows that the research attempts made are concerned with relating the causal factors to the crash frequency by adopting different approaches and analytical techniques. Research groups have developed Safety Performance Functions (SPF's). Poisson and Negative Binomial Regression models have been used to study the traffic safety characteristics of two-lane highways by Kloeden, *et* al. (1997), Garber and Ehrhart (2000), Sawalha and Sayed (2003) and Tie-jun and Dan (2010). The statistical relationship between crashes and hourly traffic flow characteristics such as traffic volume, vehicle density and volume/capacity ratios, for both rural and urban freeway segments was also developed (Lord, *et* al. 2004). Persaud and Look (2000), also developed relationship between safety performance and level of safety or capacity. Several studies related to geometric design consistency methods for rural two-lane highways have been reported (Lamm, *et* al. (1990), Anderson *et* al. (1999), Watters (2007) and Aram (2010). Consistency evaluation methods were developed by Cafiso, *et* al. (2006, 2009) in the form of safety index. Habib, *et* al. (2004, 2008) developed consistency models with thresholds for good, acceptable and poor design consistency of any section of highway. In order to study the effect of operating speed on crash occurrence, speed profile models were also developed by research groups, Fitzpatrick and Collins (2000), Federal Highway Administration (2000), Krammes (2000). In the case of heterogeneous traffic scenario, only very few studies were reported on the analysis of safety performance of highway.

1.1 Heterogeneous Traffic Condition

The methodology for prioritisation of accident black spots as well as for identification of accident causative parameters for rural highway based on computation of hazard rating scores was developed (Robert 2006). Accident prediction factors were developed by varying the roadway width and traffic volume. Models were also developed in the study for evaluating crash rate, casualties and injuries using Poisson, Negative Binomial and multiple regression models respectively. For evaluating the safety performance of two-lane undivided rural roads, Accident Prediction Models (APM's) were developed using Poisson model with random coefficients for predicting single vehicle and multi-vehicle accident (Dinu 2012). In the case of APM for single vehicle accident, parameters such as length of segment, average daily traffic, number of vertical curves have positive effect where as percentage of motorised two-wheelers and heavy goods vehicles have negative effect. But in the case of APM for multi-vehicle accidents model parameters such as length of segment, average daily traffic, number of vertical curves, percentage of motorised two-wheelers and heavy goods vehicles have positive effect. This indicates that accidents increase with exposure.

Most of the research works found in literature considered homogeneous and lane disciplined traffic conditions and hence, the results of the above studies may not be directly applicable to heterogeneous traffic such as the one prevailing in developing countries like India. Hence the present study, a methodology for the evaluation of safety performance of multi-lane rural highways operating under heterogeneous traffic is presented.

2. Data description

After reconnaissance survey, 85 km long four-lane divided inter-city roadway in Tamil Nadu state, in India, was identified as the study stretch. In order to develop models for evaluation of safety performance, data related to road geometry, traffic, volume, composition, speed etc and crash characteristics is essential. Road inventory surveys were carried out to collect the roadway geometric data.

The geometric data were collected using Hawkeye, automated data collection equipment developed by Australian Road Research Board and the outputs are linked to both spatial (GPS) and linear references. The equipment has three high-resolution digital cameras by which digital images were collected with a 150 to 180 degree field of view (centered on the travel lane) and the data were extracted using Hawkeye processing toolkit by standard rating form. The extracted data includes gradient, cross slope, horizontal curve radius, vertical curve radius, pavement width, shoulder width, access points, median opening and all particulars related to geometry of a highway. In order to collect traffic characteristics, video capturing technique was used and also data were collected from official records. Traffic composition data was extracted manually from videos and it comprises of five major categories of vehicles from city journey vehicles (two-wheelers, three-wheelers) to multi-axle trucks plying on the roadway (as shown in Figure 1). The presence of different categories of these motorised vehicles operating at different speeds and varied physical dimensions were affecting the average operating speed along the roadway, thereby playing a vital role on the safety. Traffic data collection includes composition of vehicles, operating speed along the highway for different categories of vehicles, and volume count. City journey vehicles like two-wheelers are the more vulnerable road users among the vehicle category plying on the study corridor.

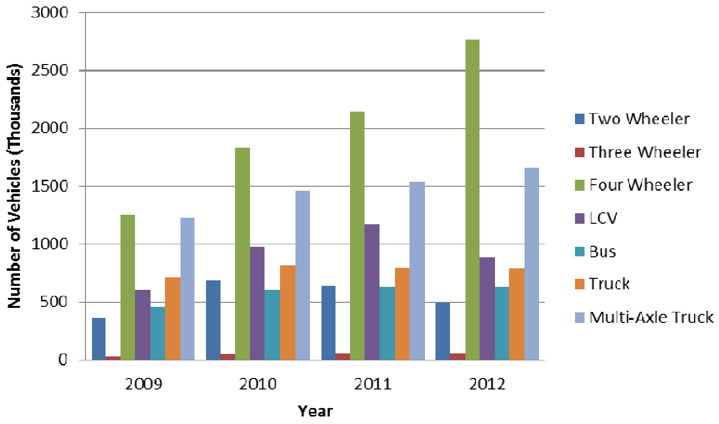

Figure 1: Traffic composition along the study stretch

Crash history is the major input for the model development. The data on crashes was collected from records for the study period (2009 to 2012), which includes the number, type and cause of crash, vehicles involved, severity, location and so on. Figure 2 shows the dependence of the number of crashes at each location to the operating speed and radius of horizontal curve. From the figure it can be observed that the variation in the geometric profile of highway under study has an influence on the crashes.

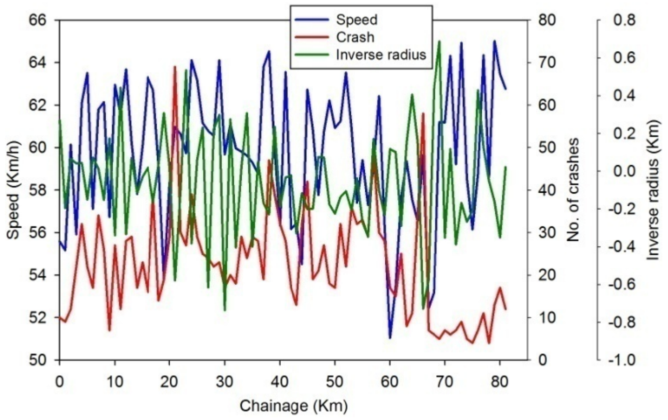

Figure 2: Plot of crashes with operating speed and radius along the study stretch

3. Data Analysis

The data collected for four years (2009-2012), were cleaned for developing models to evaluate safety performance under heterogeneous traffic condition for four-lane inter-city highways.. The data were segmented based on fundamental segmentation approaches i.e., length of roadway and annual average daily traffic (AADT). In the present work, segmentation based on fixed length of 500m length is adopted. A total of 170 segmental data are used in the crash analysis. To explore the relationship between the number of road crashes as dependent variable and geometric design characteristics and AADT (exposure variable) as independent variables, Poisson and/or Negative Binomial (NB) regression and Zero-inflated models were considered. Goodness of fit of the developed models are evaluated using likelihood ratio and Akaike's Information Criteria (AIC). The model which has the minimum AIC value is considered as the best model. The summary statistics is shown in Table 1.

Table 1:Summary Statistics of Parameters.

Variables	*Notation*	*Min.*	*Max.*	*Average*	*Std. Deviation*
No. of curves	C	0	250	-	-
No of tangents	T	0	282	-	-
Radius (m)	R	360	3000	1201.14	693.25
Degree of Curvature (deg)	DC	0.58	4.85	2.09	1.23
Length of Horizontal Curve (m)	LH	56.45	948.67	161.20	165.19
Deflection Angle (deg)	DA	0.50	78.30	9.63	10.82
Tangent Length (m)	TL	2.0	1101	253.85	239.86
Cross slope (%)	CS	-8.07	8.21	-1.31	3.79
Gradient (%)	G	-4.12	5.91	0.16	1.69
Median Opening (presence of median)	MO	0	1	-	-
Access Point	AP	0	1	-	-
Crash Frequency (Numbers/yr)	CF	0	14	-	-
Annual Average Daily Traffic (No. of 10,000 Vehicles)	AADT	1216	4749	-	-
Horizontal Curvature (1/km)	HC	-2.7	7.6	0.02	1.08
Vertical Curvature (1/km)	VC	-0.5	0.5	0.0005	0.12
Operating Speed (km/hr)	OS	37.1	69.8	57.6	7.3

3.1 Correlation Analysis

Analysis was carried out to study the correlation among the variables identified in the development of models. The correlation matrix for modeling the number of crashes as dependent variable and other geometric and traffic parameters as independent variables is shown in Table 2.

Table 2:Correlation Matrix between Variables Influencing Crashes.

	G	*CS*	*HC*	*VC*	*OS*	*AP*	*MO*	*AADT*
G	1							
CS	0.192	1						
HC	0.002	0.734	1					
VC	0.029	0.157	0.093	1				
OS	-0.183	-0.067	-0.063	0.153	1			
AP	-0.305	-0.147	-0.062	-0.086	0.103	1		
MO	-0.146	-0.194	-0.259	-0.058	0.030	0.303	1	
AADT	-0.224	-0.102	0.009	-0.037	0.033	0.136	-0.166	1

From the analysis (Table 2), it can be seen that many variables are not significantly correlated, except Cross Slope (CS) and Horizontal Curvature (HC).

Table 3 shows the correlation analysis carried out for modeling considering operating speed as dependent variable and geometric parameters as independent variables.

Table 3:Correlation Matrix between Variables Influencing Operating Speed

	OS_C	*G*	*CS*	*DC*	*LH*	*CCR*	*DA*	OS_T	*IR*
OS_C	1								
G	-0.390	1							
CS	-0.145	0.046	1						
DC	-0.012	0.128	0.221	1					
LH	0.083	-0.121	-0.087	-0.215	1				
CCR	-0.009	0.106	0.149	**0.766**	-0.198	1			
DA	0.070	-0.036	0.031	0.404	**0.648**	0.404	1		
OS_T	**0.780**	-0.307	-0.110	-0.001	0.056	-0.011	0.062	1	
IR	-0.108	0.004	**0.817**	0.019	-0.063	-0.057	-0.085	-0.117	1

From Table 3, following variables are found to be correlated.

- Curvature Change Rate (CCR) and Degree of Curvature (DC)
- Deflection Angle (DA) and Length of Horizontal curve (LH)
- Operating Speed at Tangent (OS_T) and Operating Speed at Curve (OS_C)
- Inverse Radius (IR) and Cross Slope (CS)

Models were developed by considering the interaction between significantly correlated variables.

4. Model Development

The present work attempts to analyse safety performance by developing models and to predict the number of crashes under heterogeneous traffic condition on multi-lane rural highways using NLogit 5 statistical software.

4.1 Poisson Regression

This method is referred to as a benchmark for modelling count data. Poisson regression model is the simplest model and the response variable is assumed to be independent and follows a Poisson distribution.

In this model each observed data y_i is from Poisson distribution with conditional mean, given vector X_i for case i. The probability mass function is shown in Eqn. 1.

$$P(y_i)=exp(-\lambda_i)\lambda_i\, y_i / y_i\,! \tag{1}$$

Where λ_i *is* the Poisson parameter of the number of crashes, which is equal to E[y_i].

Poisson regression models are estimated by specifying the Poisson parameter λ_i as a function of explanatory variables.

Limitation of using Poisson regression model is over-dispersion, i.e., the variance is greater than the mean. Over-dispersed data results in small p-values and standard errors.

4.2 Poisson-Gamma

Poisson-gamma is one other model form for count data and the probability model is shown in Eqn. 2.

$$\text{Prob}[y_i=j] = G(\alpha_j, \lambda_i) - G(\alpha_j + \alpha, \lambda_i) \tag{2}$$

Where, $\lambda_i= \exp(\beta'\mathbf{x}_i)$ (similar to Poisson), $G(\alpha_j, \lambda_i) = 1$, if $j = 0$

'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.

4.3 Negative Binomial (NB) Regression

Negative binomial regression is used for modeling over-dispersed count data. The Poisson regression model can be considered as a limiting model of the NB regression model, as the value of ‘α’ approaches zero. It does not assume an equal mean and variance and is particularly suitable when the variance is greater than the mean, (Sawalha and Sayed, (2003), Miller and Freund, (1977)). The form of the model is same as that for Poisson regression, with an additional parameter ‘α’ to model the over-dispersion. The selection between these two models is dependent on the value of the parameter ‘α’. The general form of NB regression is given in Eqn. 3.

$$P\left(y_i\right)=\frac{\Gamma\left(\left(1/\alpha\right)+y_i\right)}{\Gamma\left(1/\alpha\right)}\left[\frac{1/\alpha}{\left(1/\alpha\right)+\lambda_i}\right]^{1/\alpha}\left[\frac{\lambda_i}{\left(1/\alpha\right)+\lambda_i}\right]^{y_i} \qquad (3)$$

4.4 Zero-Inflated Models

Zero-Inflated models are used when the data frequently exhibits over dispersion and excess zeros. This technique was first introduced by Lambert (1992). Zero-Inflated of Poisson (ZIP) or Zero-Inflated of Negative Binomial model (ZINB) can be classified as a mixture of two data generation process. In this type of model, for each observation, there are two possible data generation processes. Process 1 is chosen with probability ϕi (generate always zero counts) and process 2 with probability 1−ϕi (generate counts from Poisson or NB model).

4.5 Operating Speed Models

Operating speed (85th percentile speed) models were developed to establish relationship between the operating speed (OS) and the geometric design parameters. OS models were developed considering operating speed at mid-curve and at tangent as dependent variables and geometric design parameters as independent variables using Multiple Linear Regression (MLR) approach. Summary statistics of data used for developing operating speed models are shown in Table 1.
General form of MLR is shown in Eqn. 4.

$$Y = \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_4 X_4 + \ldots + \beta_0 \qquad (4)$$

The developed models can be used as tool/criteria for evaluating the road design parameters from safety point of view.

5. Model Selection

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. Likelihood ratio *chi*^{2}statistic 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. Akaike’s Information Criteria (AIC) helps in identifying the best model and is defined as shown in Eqn. 5.

$$\text{AIC}=\left[-2L_n\left(\beta_u\right)+2b\right]/N \qquad (5)$$

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.

5.1 Vuong Test

This is a non-nested test that is based on a comparison of the predicted probabilities of two models that do not nest (Vuong, 1989). This test is used for comparison between Zero-inflated Poisson (ZIP) count models with Poisson regression, or Zero-inflated negative binomial (ZINB) against negative binomial model. The result of the test provides evidence of the superiority of one model over the other. If the value of the Vuong is greater than 1.96, the first model is preferred. But if it is less than 1.96, the second model is preferred (Miller and Freund, 1977).

6. Results and Summary

The analysis on the safety performance along a four lane divided rural highway was analysed and it is found that the design variables like gradient, cross slope, curvature, operating speed and AADT contribute to safety.

6.1 Safety Performance Function

The influence of geometric characteristics on the crash frequency was studied by relating the number of crashes to operating speed and to the radius of curve. The model results are shown in Tables 4 and 5, and it can be seen that the variation in the geometric parameters significantly affects crash frequency.

TABLE 4: SPF's developed using Poisson, Poisson-Gamma and NB Regression

Variables	*Poisson Regression*			*Poisson Gamma*			*Negative Binomial*		
	Coefficient	*Standard Error*	*p-value*	*Coefficient*	*Standard Error*	*p-value*	*Coefficient*	*Standard Error*	*p-value*
Constant	20.93***	1.91	0.00	24.25***	5.56	0.00	18.89***	5.30	0.00
Upward Grade	0.06*	0.03	0.09	0.06	0.10	0.55	0.04	0.09	0.66
Downward Grade	-0.40***	0.02	0.00	-0.45***	0.05	0.00	-0.32***	0.05	0.00
Cross slope	-0.09***	0.02	0.00	-0.12*	0.06	0.08	-0.10*	0.06	0.09
Hori curvature	0.18**	0.08	0.03	0.21	0.22	0.35	0.17	0.20	0.41
Vert curvature	-0.73	1.03	0.48	-0.81	2.98	0.79	-1.72	2.78	0.54
Operating Speed	0.04***	0.01	0.00	0.05***	0.02	0.10	0.04***	0.01	0.01
Access pt	0.10*	0.05	0.07	0.14	0.16	0.38	0.17	0.13	0.16
Median Open	0.22***	0.05	0.00	0.26*	0.13	0.05	0.11*	0.12	0.05
LnAADT	-3.13***	0.29	0.00	-3.76***	0.87	0.00	-2.82***	0.81	0.00
Log likelihood	-718.65			-557.79			-554.15		
Restricted LL	-1107.50			-718.65			-718.65		
AIC	1459.30			1139.60			1132.30		
Chi squared (d.f.)	777.69			321.72			328.99		
Overdispersionn		7.32		-			-		
Dispersion Parameter	-			0.18***			0.27***		

Note: ***, **, * - Significance at 1%, 5% & 10% level

Poisson regression model was developed and it was found that most of the variables identified for the model development are statistically significant. But greater variability is observed (i.e., 7.32), which clearly indicates the over-dispersion of data. Hence attempts were made to develop Poisson-gamma and NB regression models with number of crashes as dependent variable and upward gradient(%), downward gradient(%), cross slope (%), horizontal curvature (1/km), vertical curvature (1/km), operating speed (km/hr), presence of median opening (1 – if present; otherwise 0), presence of access points (1 – if present; otherwise 0), and annual average daily traffic as independent variables. The models developed using Zero-Inflated regression (count data parts, i.e., ZIP and ZINB) (due to the presence of zeros and over-dispersion) are shown in Table 5.

In the case of NB regression models few variables such as downward gradient, operating speed, and AADT are found to be statistically significant for most of the models at 1%, 5% and 10% significance levels. The value of dispersion parameter, 'α' is 0.274. The log-likelihood value was found to be -554.153, which is lowest of the three values. The mean AIC value was estimated to be 1132.3, which is also the lowest value among the other values. Log-likelihood ratio and Akaike's Information Criteria values obtained using NB regression model were lower than that of Poisson regression and Poisson-gamma models, which shows the ability of NB model to predicting crashes as shown in eqn. 8. Since the value of vuong statistics is less than 1.96, it is evident that negative binomial model is superior to ZINB models.

TABLE 5: Results of the Count Data Model parts of Zero-Inflated Regression

Variables	*Zero-Inflated Poisson*			*Zero-Inflated Negative Binomial*		
	Coefficient	*Standard*	*p-value*	*Coefficient*	*Standard*	*p-value*
Constant	20.69***	1.02	0.00	18.89***	5.30	0.00
Upward Grade	0.05***	0.02	0.00	0.04	0.08	0.66
Downward Grade	-0.40***	0.01	0.00	-0.32***	0.05	0.00
Cross slope	-0.09***	0.01	0.00	-0.10*	0.06	0.09
Hori curvature	0.17***	0.04	0.00	0.17	0.20	0.41
Vert curvature	-0.81	0.52	0.12	-1.72	2.76	0.53
Operated Speed	0.04***	0.00	0.00	0.04***	0.01	0.01
Access pt	0.10***	0.03	0.00	0.17	0.12	0.16
Median Open	0.23***	0.02	0.00	0.11	0.11	0.34
LnAADT	-3.08***	0.16	0.00	-3.13***	0.30	0.00
Log likelihood	-716.30			-554.15		
Vuong Statistics	0.71			0.00		
Dispersion	-			0.27***		
Tau	-2.29***			-23.49		

Models developed in the study using NLogit 5 software are as follows (eqn. 6 – 10):

Poisson Regression (eqn. 6)

$$Y = exp\left(20.931 + 0.057UG - 0.403DG - 0.094CS + 0.175HC + 0.037OS + 0.097AP + 0.222MO - 3.131LnAADT\right)$$

Poisson Gamma (eqn. 7)

$$Y = exp\left(24.245 - 0.446DG - 0.115CS + 0.045OS + 0.256MO - 3.758LnAADT\right)$$

Negative Binomial (eqn. 8)

$$Y = exp\left(18.886 - 0.315DG - 0.095CS + 0.036OS + 0.107MO - 2.817LnAADT\right)$$

Zero-Inflated Poisson (eqn. 9)

$$Y = exp\left(20.685 + 0.054UG + 0.402DG - 0.089CS + 0.170HC + 0.035OS + 0.099AP + 0.226MO - 3.079LnAADT\right)$$

Zero-Inflated Negative Binomial (eqn. 10)

$$Y = exp\left(18.886 - 0.315DG - 0.095CS + 0.036OS - 3.131LnAADT\right)$$

From the models developed, it can be seen that most of the parameters considered in analysis are significant in the case of Poisson regression and Zero-Inflated Poisson regression analysis. But due to the presence of overdispersion parameter, negative binomial and due to the existence of excess zeros, Zero-Inflated regressions were adopted. Based on the goodness of fit values (as mentioned above), it is evident that the model developed using negative binomial regression (eqn. 8) is considered for safety evaluation of multi-lane highways under heterogeneous traffic.

6.2 Operating speed model

Operating speed along the mid-point of the curve ($V_{85(MC)}$) is considered as the dependent variable and the independent variables considered are radius (R), inverse radius (IR), tangent length (TL), operating speed at preceding tangent(OS_T), length of curve (LH), curvature change rate (CCR), degree of curvature (DC), deflection angle (DA), cross slope (CS) and gradient (G). The final model identified is shown in Eqn. 11.

$$V_{85(MC)} = 9.071 - 0.447\ CS + 1.368\ IR + 0.\ 847\ OS_T \qquad (11)$$
$$(R^2 - 0.64)$$

Tangent length, length of curve, degree of curvature and deflection angle are not statistically significant whereas operating speed at preceding tangent, cross slope and inverse of radius are found to be the influencing variables on the operating speed along mid-curve.

In the case of tangent section, operating speed at the tangent preceding the curve is considered as the dependent variable and the independent variables considered are tangent length (TL), radius of curve (R), length of curve (LH), operating speed along mid-curve (OS_C), cross slope (CS) and gradient (G). Best fit model out of developing models is shown in Eqn. 12.

$$V_{85(T)} = 18.401 - 1.118\ IR + 0.678\ OS_C \qquad (12)$$
$$(R^2 - 0.63)$$

From the models developed, the identified is shown in eqn. 12, and was observed that the operating speed on the preceding tangent is influenced by the operating speed of curve section with a R^2 value of 0.63. Other variables that are statistically significant are tangent length, inverse radius and cross slope.

7. Best Fit Models For Safety Analysis

Out of the models developed, the best model was identified based on the goodness of fit test values. The identified models (shown in Table 6), will enables the transportation engineers to evaluate the safety performance of the existing highways and suggest counter measures to obtain a particular level of safety. More over this application can also be used by engineers and planners for the design/re-design of new/existing highways.

TABLE 6: Best Fit Models Identified from the Developed Models

Category	*Models*
Number of Crashes	$Y = exp\left(18.886 - 0.315DG - 0.095CS + 0.036OS + 0.107MO - 2.817LnAADT\right)$ Y- Number of crashes/year/segment DG – Downward Gradient (%) CS – Cross Slope (%) OS – Operating Speed (km/hr) MO – Presence of Median Opening (0/1) AADT – Annual average daily traffic (Number of vehicles)
Operating Speed – Mid of curve	$V_{85(MC)} = 9.071 - 0.447\ CS + 1.368\ IR + 0.847\ OS_T$ $V_{85(MC)}$ – Operating speed at mid of curve (km/hr) CS – Cross slope (%) IR - Inverse of curve radius (m) OS_T – Operating speed at previous tangent (km/hr)
Operating Speed – Mid of tangent	$V_{85(T)} = 18.401 - 1.118\ IR + 0.678\ OS_c$ $V_{85(T)}$ – Operating speed at tangent (km/hr) IR – Inverse of curve radius (m) OS_C – Operating speed at curve (km/hr)

The first part of the study enables the user to evaluate safety directly from the available geometric design characteristics, traffic characteristics and other roadway features. If all these data are available, then safety level of the highway under study can be evaluated using the model developed. The second part is applicable when the speed data of the study corridor is not available. In this case, the operating speed along curve and/or tangents can be estimated from the models developed. The safety in terms of number of crashes is then estimated. Hence, the resulted information enables the traffic planners/highways engineers to analyse safety of multi-lane highways under consideration based on the data availability.

8. Conclusions

The study was carried out to analyse safety performance of a four-lane divided highway in India, operated under heterogeneous traffic condition. The study was carried out in two sections, *viz*, development of safety performance functions and also operating speed models. Safety Performance functions (SPF’s) were developed using GLM approach likes Poisson, Poisson-gamma regression, negative binomial and zero-inflated models where as operating speed models for the mid of curve as well as tangent were developed using multiple linear regression approach. As far as SPF’s were considered, geometric design parameters such as gradient, cross slope, operating speed, median opening and annual average daily traffic have significant effect on crashes and this variation in turn affects the traffic characteristics thereby decreasing the level of safety, leading to more number of crashes. From the models developed based on the goodness of fit test, NB regression model was found to have lower value than that of Poisson regression and Poisson-gamma models, which shows the ability of NB model for predicting crashes.

The operating speed models gave insight that the operating speed on curve is influenced by operating speed on preceding tangent segment and vice versa. This explains the role of geometric design parameters in predicting crash at curve as well as tangent section in a multi-lane rural highway. The design consistency of the highway plays an important role on road crashes. The study resulted in a better understanding of the geometric and traffic characteristics on crash frequency of multi-lane highways. The study is a novel approach to predict safety level of highway operated under heterogeneous traffic scenario.

References

Alexander, G. J., and Lunenfeld, H., (1986) "Driver expectancy in highway design and traffic operations", Final report, Federal Highway Administration, Washington, DC.

Aljanahi, A. A. M., Rhodes, A. H., and Metcalfe, A.V., (1999) "Speed, speed limits and road traffic accidents under free flow conditions", *Accident Analysis & Prevention*, 31, pp: 161-168.

Anderson, I. B., Bauer, K. M., Harwood, D.W., and Fitzpatrick, K., (1999) "Relationship to safety of geometric design consistency measures for rural two-lane highways", *Transportation Research Record: Journal of the Transportation Research Board*, 1658, pp: 43-51.

Aram, A. (2010) "Effective safety factors on horizontal curves of two-lane highways", Journal of Applied Sciences, 10(2), pp: 2814 –2822.

Cafiso, S., Graziano, A. D., Silvestro, G. D., Cava, G. L., and Persaud, B., (2010) "Development of comprehensive accident models for two-lane rural highways using exposure, geometry, consistency and context variables", Accident Analysis & Prevention, 42(4), pp:1072–1079.

Dinu, R. R. (2012) Studies on safety performance of two-lane rural highways in mixed traffic, *Ph.D. thesis,* Indian Institute of Technology Madras, Chennai, India

Garber, N. J., Ehrhart, A. A., (2000) "The effect of speed, flow, and geometric characteristics on crash frequency for two-lane highways", 79th Annual Meeting of the Transportation Research Board, Washington D.C.

Habib, C. M., Polus, M., and Farah, H., (2008) "Further evaluation of the relationship between enhanced consistency model and safety of two-lane rural roads in Israel and Germany", EJTIR, 8(4), pp: 320 –332.

Krammes, R. A. (1995) "Horizontal Alignment Design Consistency for Rural Two-Lane Highways", Rep. No. FHWA-RD-94-034, Federal Highway Administration, Washington, DC.

Krammes R. A., (2000) "Design speed and operating speed in rural highway alignment design", *Transportation Research Record,* 1701, pp: 68-75.

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

Ministry of Road Transport and Highways (MORTH), Transport Research Wing, Government of India, *Road Accidents in India 2013, 2014.* , Naderan, A., Babaei, M., (2011) "Assessment of Statistical Models for Aggregate Crash Prediction", Presented at the 90th Transportation Research Board Annual Meeting, Washington D.C.

NCHRP Report 502 (2003) "Geometric design consistency on high – speed rural two lane roadways", Transportation Research Board, Washington, USA.

Persaud, B. N., (1994) "Accident prediction models for rural roads", *Canadian journal of civil engineering*, 21(4), pp: 547–554

Polus, M., Habib, C. M., (2004) "New consistency model for rural highways and its relationship to safety", *Journal of Transportation Engineering*, ASCE, 130(3), pp: 286–293

Robert, V. R. (2006) Studies on road safety problems under mixed traffic flow, Ph*.D. thesis*, Bangalore University, India

Sawalha, Z., Sayed, T., (2003) "Statistical Issues in Traffic Accident Modeling", Presented at the 82nd Transportation Research Board Annual Meeting, Washington D.C.

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 in rural highways*, Switzerland, 2013, 2015.

Acknowledgements

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