

Modeling Motorcycle Accident on Rural Highway

A.K.Sharma, V.S. Landge, and N.V.Deshpande

Abstract— Motorcyclists are the most crash-prone road-user group in many Asian countries including India. Statistics of accident on Indian roads reveals that motorcycles accounted for the highest share in total road accidents in 2011. Earlier studies have authenticated the use of Generalized Linear models with Poisson or Negative binomial error structure, for accident modeling. This study examines the effect of road geometry and traffic variables on motorcycle crashes using a statistical technique called as Zero inflated negative binomial regression. The independent variables selected for this study includes access density (AD), annual average daily traffic(AADT), heavy vehicle percentage(HVPER), speed variation from modal speed (VARMSP), standard deviation of speed(STDSP), and shoulder width deficiency(SWDEF). Accident per year per km(ACCR) is taken as dependent variable. Accident data collected between 2005-09 over a stretch of 100 km of road length are used for modeling. It is observed that shoulder width deficiency, percentage of heavy vehicles in traffic and speed variations have significant impact on safety of motorcyclist.

Keywords— Crash rate, access density, shoulder width deficiency, zero inflated negative binomial

I. INTRODUCTION

ROAD accidents are a human tragedy, which involve high human suffering. They impose a huge socio-economic cost in terms of untimely deaths, injuries and loss of potential income. The total number of road accident in India in 2011 was 497,686, out of which 24.4% were fatal. The age profile of accident victims reveals that more than half were in the wage earning age group (25-65) [1].

Motorcyclists are the most crash-prone road-user group in many Asian countries including India. Statistics of accident on Indian roads reveals that motorcycles accounted for the highest share in total road accidents (23.7%) in 2011 [1]. Motorcyclists are the most vulnerable road users with maximum fatality (26.6%) in total road accident.

Rural roads were reported to have more number of accidents (53.5%) than urban areas (46.5%). Also rural roads had more fatalities (63.4%) than urban areas (36.6%) and number of

persons injured was more in rural areas (59.4%) as compared to urban areas (40.6%) [1].

National Highways (one of the rural roads in India, abbreviated as NH) accounted for 30.1 % in total road accidents and 37.1% in total number of person killed in 2011[1]. Present study, on motorized two-wheelers, is based on data collected from National Highway No.6 (NH-6) for a road length of 100km between Amravati (a city on NH-6) and Lakhni (a township on NH-6).

Earlier studies have authenticated the use of Generalized Linear models with Poisson or Negative binomial error structure, for accident modeling. This study presents development of predictive models for motorcycle accidents using Zero Inflated Negative Binomial.

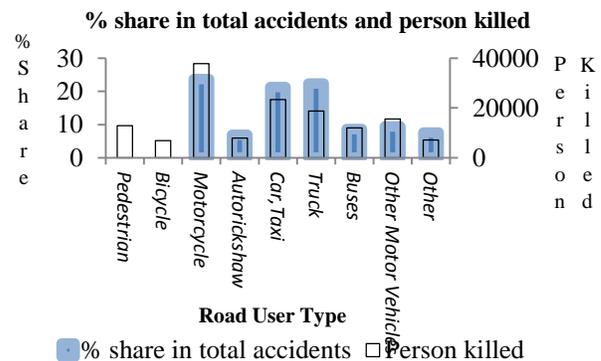


Fig.1 Percentage share of different vehicles in total road accidents and person killed (India-2011)

II. OBJECTIVE

The specific objectives of studies are:

- Development of Correlation between road accidents and geometric design parameters of highway along with traffic operating characteristics for pedestrian accident.
- Evolving engineering remedial measures for improving safety on the selected stretch.
- Practical recommendations for improving traffic safety on the said highway.

A.K.Sharma is with Civil Engineering Department of Shri Ramdeobaba College of Engineering and Technology, Nagpur(India) Pin-440013, phone-0712-2582844, mobile-09823126311.(email:sharmaakrkn@yahoo.co.in)

Dr. V.S.Landge is with Civil Engineering Department of VNIT, Nagpur (email:vslandge@rediffmail.com)

Dr.N.V.Deshpande is with Gurunanak Institute of Technology, Nagpur as a Principal. (email:narendravdeshpande@gmail.com)

III. STUDY AREA AND SCOPE

The study area chosen is National Highway No.6 commonly refer to as NH-6 or G.E. Road (Great Eastern Road), which is a one of the most busy national highways in India. It's a connecting corridor to major states of India namely Gujarat, Maharashtra, Chhatisgarh, Orissa, Jharkhand and West Bengal. NH-6 eventually will be one of the important links of Asian Highway network (AH-46).

The scope of present study is limited to road section passing through central Indian states of Maharashtra. Maharashtra, one of the most advanced states of India, has high traffic accident rate. Most of these accidents occur on the National highways.

IV. METHODOLOGY

Methodology adopted for the study is as specified below:

- Identification of study area
- Crash data collection from the law enforcement agency and insurance companies.
- Road geometric and Traffic parameter data from field studies
- Selection of variables and modeling methods
- Development of accident prediction models
- Testing of models, interpretation of results and remedial measures suggestions.

V. DATA COLLECTION AND ANALYSIS

Accident data was collected from the police stations and insurance companies. Road geometry and traffic data was collected through field studies and traffic count survey for a road length of 100km between Amravati City and Nagpur City of Maharashtra State in India. For the purpose of collecting road geometry data, the road was divided into segments of similar characteristics ranging from 0.2 to 0.6km length for curve portion and 0.7 to 1.6 km for straight portion.

The preliminary analysis of the data are given in fig.2 and fig.3. Fig.2 shows a positive relation between dependent variable and shoulder width deficiency, which suggest that more deficient shoulders give rise to more accidents. Fig.3 which gives relation of heavy vehicle percentage with the dependent variable suggests that accident rates may increase with percentage of heavy vehicles in the traffic stream. Similarly relationships of other variables can be presented graphically.

VI. VARIABLES USED IN THE MODEL

A thorough scrutiny and the preliminary analysis of the data led to selection of following highly influential independent parameters.

Shoulder Width Deficiency: Shoulder provides an area along the highway for vehicle to stop, particularly during

emergency. A report by Zegeer et al.[2] indicated that a paved shoulder widening of 2 feet per side reduces accidents by 16%. Shoulder width deficiency from a standard minimum (5m in this study inclusive of both sides) has been a variable with significant influence on safe operations of traffic and hence selected as a variable.

Access Density: The availability of access point is necessary to commercial or residential developments, usually at the expense of traffic operations and the safety of local highway systems. To achieve a good coordination of these two aspects, compromises are often required to be made between accessibility and mobility or capacity and safety. On the study area access density ranging from 0 to as much as 6 per km is observed on the said highway.

Traffic Volume: Traffic volume is believed to have considerable impact on the crash rate [3]. For this study annual average daily traffic (AADT) is used as a parameter to indicate traffic volume.

Heavy Vehicle Percentage: There are two main traffic related issues associated with commercial vehicles, namely: delays that they may cause to other vehicles and safety related impacts [4]. It has been suggested by a number of authors that the presence of a truck in front of any other vehicle may result in the driver being more cautious due to the large size of the vehicle and the diminished sight distances[5] [6].

Speed Variations: Spot speed is one of the major parameter that is used as an indicator of traffic performance. The data collected shows a wide variation in the spot speed from 25kmph to 60kmph. The data collected showed positive impacts of speed variations on traffic safety. The speed variations of the vehicles are incorporated in terms of variation in speed from modal speed and standard deviation of speed.

VII. MODELING OPTION AVAILABLE

Various modeling techniques have been tried to model accidents aiming for accuracy. However the suitability of the model depends on the data quality and is location specific. The model building methodology is selected based on the availability of data and accuracy of the data. Various modeling techniques popularly employed are as discussed below :

Stochastic Models : Early in 1989; Okamoto et al.[7] suggested that the occurrence of traffic crashes follows stochastic distribution. In 1990, Garber et al.[8] developed several models to describe the occurrence of crashes using stochastic modeling techniques, like Poisson regression(PR) [9] and negative binomial regression(NBR)[10][11]. More stochastic models were also proposed other than Poisson regression models and Negative Binomial Regression Model, which included Zero Inflated Poisson Regression Model and Zero Inflated Negative Binomial Regression Models. This paper presents models build using Zero Inflated Negative Binomial Regression Zero Inflated Negative Binomial

Regression Models: Transportation safety analysts have typically justified the use of Zero Inflated (ZI) models because of the improved statistical fit compared to traditional Poisson and NB models. Zero Inflated regression models are two regime models. First probability model governs whether a count number is zero or positive number, called as inflate model. Then the positive part of the distribution is described by suitable stochastic distribution, called as base model. In Miaou’s[12] study, the negative binomial regression model was of the form;

$$y_i = 0,1,2 \dots \dots \text{with probability} \frac{\Gamma\left(\frac{1}{\alpha} + y_i\right)}{\Gamma\left(\frac{1}{\alpha}\right) \Gamma(y_i + 1)} \left(\frac{1}{1 + \alpha * \lambda_i}\right)^{\frac{1}{\alpha}} \left(\frac{\alpha * \lambda_i}{1 + \alpha * \lambda_i}\right)^{y_i} \quad (1)$$

$\lambda_i = e^{\beta_i x_i}$ where x_i is i^{th} covariate and β_i is the regression coefficient

For Zero Inflated Negative Binomial regression

$$y_i = 0 \text{ with probability } p_0 + (1 - p_0) \left(\frac{1}{1 + \alpha * \lambda_i}\right)^{\frac{1}{\alpha}} \quad (2)$$

$$y_i = 1,2 \dots \dots \text{with probability} (1 - p_0) \frac{\Gamma\left(\frac{1}{\alpha} + y_i\right)}{\Gamma\left(\frac{1}{\alpha}\right) \Gamma(y_i + 1)} \left(\frac{1}{1 + \alpha * \lambda_i}\right)^{\frac{1}{\alpha}} \left(\frac{\alpha * \lambda_i}{1 + \alpha * \lambda_i}\right)^{y_i} \quad (3)$$

Where p_0 can be represented by probability model incorporating the effects of covariates, such as logit model.

$$p_0 = \frac{e^{r'w_i}}{1 + e^{r'w_i}} \quad (4)$$

r' is the coefficient matrix and w_i is the i^{th} covariate. $\Gamma(\cdot)$ is Gamma function; and α is the rate of over dispersion.

VIII. MODEL SELECTION CRITERIA

Maximum likelihood estimation method has been employed widely in estimating Poisson, negative binomial [13] and zero inflated regression models. Akaike Information Criterion (AIC) [14] and Bayesian Information Criteria(BIC) was used to judge the performance of the model. Smaller the AIC and BIC values, the better is the model.

$$AIC = -2 \text{Log } L + 2K \dots \dots \quad (5)$$

Where Log L is the log likelihood; K is the number of estimated parameters.

IX. VARIABLE SELECTION

Various combinations of the variables selected were tried the best combinations are given in Table I. Accident rate per year per km was chosen as dependent variable. Crash rate is defined as $CR_i = TA_i / L_i / NY$, where CR_i = Crash Rate on segment , TA_i = Total accidents on segment i, L_i = Length in Km. of segment i, NY = Number of Years

X. RESULTS

The results obtained after analysis of data, using SPSS software, are shown in TABLE II and TABLE III.

In order to see the performance of the model the coefficient of variables has to be examined. Models with logical algebraic signs of the variables were selected. Along with the strong statistical tools, proper engineering judgments are required to decide upon the selection of final model. For example, a positive sign of shoulder width deficiency suggests that more shoulder width deficiency will be responsible for more accidents and it is logically acceptable.

A. Base Model

$$\text{Acc/Yr/Km} = -2.211 + 0.133 * \text{SWDEF} + 0.079 * \text{HVPER} + 0.044 * \text{VARMSP} + 0.0000288 * \text{AADT} \quad (6)$$

$$\lambda = e^{-2.211 + 0.133 * \text{SWDEF} + 0.079 * \text{HVPER} + 0.044 * \text{VARMSP} + 0.0000288 * \text{AADT}} \quad (7)$$

Substituting average values for variables , we get

$$\lambda = e^{0.562}$$

B. Inflate Model

$$p_0 = \frac{e^{-13.512 + 1.688 * \text{SWDEF} + 0.411 * \text{HVPER}}}{1 + e^{-13.512 + 1.688 * \text{SWDEF} + 0.411 * \text{HVPER}}} \quad (8)$$

Substituting average values for variables, we get

$$p_0 = 0.571$$

C. Final Models

Model to predict frequency of zero accidents

$$P(y = 0) = \varepsilon * [p_0 + (1 - p_0) * \left(\frac{1}{1 + \alpha * \lambda_i}\right)^{\frac{1}{\alpha}}] \quad (9)$$

Model to predict frequency of accidents other than zero

$$P(y) = \varepsilon * \left[(1 - p_0) * \frac{\Gamma\left(\frac{1}{\alpha} + y_i\right)}{\Gamma\left(\frac{1}{\alpha}\right) \Gamma(y_i + 1)} \left(\frac{1}{1 + \alpha * \lambda_i}\right)^{\frac{1}{\alpha}} \left(\frac{\alpha * \lambda_i}{1 + \alpha * \lambda_i}\right)^{y_i} \right] \quad (10)$$

Where ε is exposure term = total no. of road sections(128 in present study) α is overdispersion parameter =0.62 (for present data collected)

TABLE I
VARIABLE IN THE MODEL

Model No.	Independent Variables	Dependent Variable
1	SWDEF, HVPER, VARMSP, AADT	Accidents/Yr/Km (Crash Rate)
2	SWDEF, HVPER, VARMSP	
3	SWDEF, HVPER, STDSP	
4	SWDEF, HVPER, VARMSP, AD	

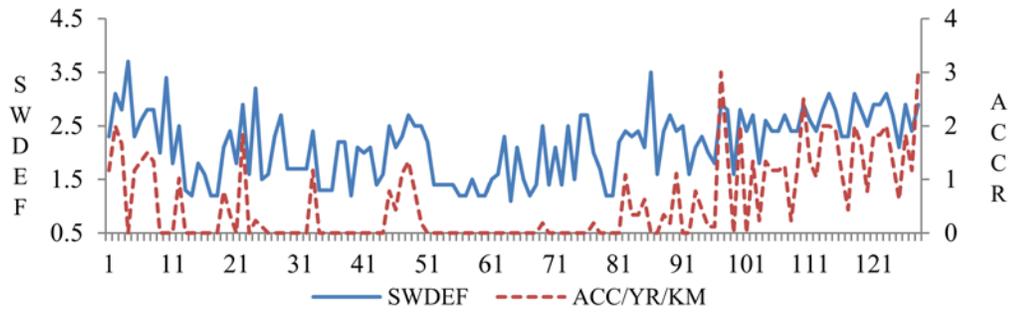


Fig. 2 Variation of accident rate with shoulder width deficiency

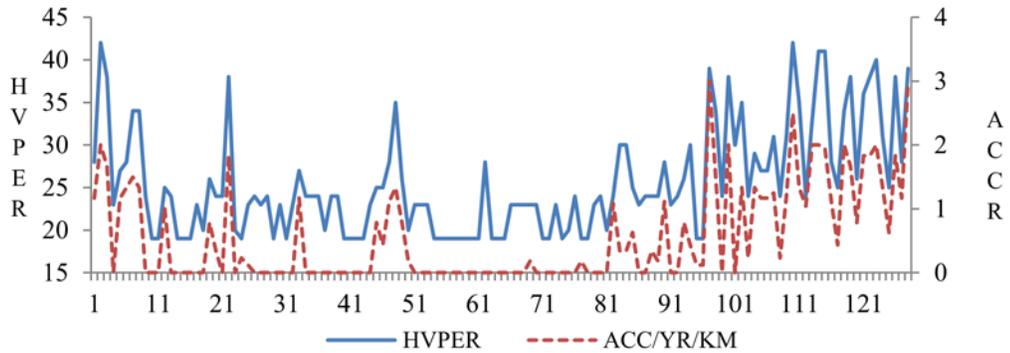


Fig. 3 Variation of accident rate with percentage of heavy vehicles

TABLE II
PARAMETER ESTIMATES FOR BASE MODEL (USING SPSS)

Variable	Model 1	Model 2	Model 3	Model 4
Intercept	-2.211	-2.028	-2.164	-2.025
SWDEF	0.133	0.142	0.146	0.121
HUPER	0.079	0.082	0.089	0.082
VARMSP	0.044	0.041		0.040
AD	-----	-----	-----	0.043
AADT	0.0000288	-----	-----	-----
STDSP	-----	-----	0.035	-----
AIC	61.52	66.49	77.55	65.46
BIC	77.93	79.26	89.32	81.88

TABLE III
PARAMETER ESTIMATE FOR INFLATE MODEL

Inflate Variable	SWDEF	HUPER	Intercept
Coefficient	+1.688	+0.411	-13.512

TABLE IV
PREDICTED AND OBSERVED FREQUENCY

Accident Frequency	Predicted Frequency	Observed Frequency	R-Square
0	89.8	71	0.959
1	14.06	24	
2	9.55	4	
3	5.99	2	
4	3.61	1	
5	2.11	1	
6	1.21	1	
7	0.69	0	
8	0.38	0	
9	0.21	0	
10	0.12	0	

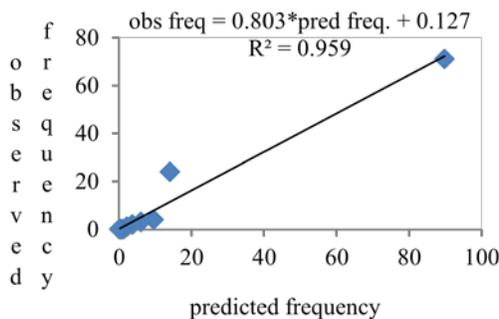


Fig.4 Observed vs predicted frequency

XI. CONCLUSION

Zero inflated negative binomial regression was proposed to establish an empirical relationship between motorcycle accidents and highway geometric and traffic parameters. The result of this study can eventually be employed to identify the locations with certain numbers of accident frequencies under different geometric and traffic conditions for national highway number-6 (NH-6) in India. The results obtained here provide valuable insight into the underlying relationship between risk factors and vehicle accidents. The model selected may be used to develop strategies for enhancing safety with optimum use of resources. The effects of various variables can also be quantified from the parameter estimates of the models. Motorcycle accident is greatly influenced by shoulder width deficiency on a particular road section. Reducing the shoulder width deficiency by 1m may reduce the accidents by about 24%. Every 2% increase in heavy vehicle traffic may increase the accidents by about 28%. Similarly speed variation is also a influential factor affecting accidents.

XII. SCOPE FOR FUTURE WORK

The regression method may be used to model traffic accident for different set of crash data collected from different sections of the same or some other highway. The method may be compared with other methods like ANN. Future work might also focus on improving the prediction performance of the ZINB models.

REFERENCES

- [1] Road accident in India (2011)" (Ministry of Road Transport and Highway).
- [2] Zegeer, C. V., Deen, R. C. & Mayes, J. G. (1981) "Effect of lane and shoulder widths on accident reduction on rural two-lane roads". TRR 806 .pp 33-43.
- [3] Karlaftis M.G. & Golias, I.(2002) , "Effect of road geometry and traffic volumes on rural roadway accident rates", Accident Analysis and Prevention vol.34, Issue 3, pp. 357-365.
- [4] Krammes, R. and Crowley, K. Passenger Car Equivalents for Trucks on Level Freeway Segments. *Transportation Research Record*, **1091**, pp. 10-17, 1987.
- [5] Molina, C. Development of Passenger Car Equivalencies for Large Trucks at Signalised Intersections. *ITE Journal*, November, 1987.
- [6] Kockelman, K. and Shabih, R. Effect of Light-Duty Trucks on the Capacity of Signalised Intersections. *Journal of Transportation Engineering*, **126(6)**,pp. 506-512, 2000.
- [7] Okamoto (1989), "A Method to cope up with random errors of observed accident rates" Safety literature, Vol.4, page 317-332.
- [8] Garber, N.J., Wu, L(1989), "Stochastic models relating crash probabilities with geometric and corresponding traffic characteristics data", Research report No UVACTS-5-15- 74 Center for transportation studies at the University of Virginia.
- [9] Jovanis, P. & Chang, H(1986)., "Modeling the relationship of accidents to miles traveled", Transportation Research Record 1068, pp 42-51.
- [10] Poach, M. & Mannering, F.(1996), "Negative binomial analysis of intersection accident frequencies" ASCE's Journal of Transportation Engineering Vol.122 No. 2, pp. 105-113.
- [11] Ziad Sawalha & Tarek Sayed (2003) "Statistical Issues In Traffic Accident Modeling" TRB annual meeting (CD-ROM).
- [12] Miaou, S.P.(1994) "The relationship between truck accidents and geometric design of road sections: Poisson versus negative binomial regressions", Accident Analysis and Prevention vol. 26, issue 4, pp. 471-482

- [13] Landge, V.S., Jain S.S. & Parida, .(2006), "Modeling traffic accidents on two lane rural highways under mixed traffic conditions", 87th Annual Meeting of Transportation Research Board.(CD- Rom).
- [14] Sharma A.K., Landge V.S.(2012), "Pedestrian Accident Prediction Model For Rural Road", International Journal of Science and Advanced Technology Volume 2, No 8 .

A.K.Sharma ,born on 28-12-64 ,graduated in civil Engineering from NIT Raipur(Chattisgarh,India) in 1987. He completed his post-graduate in Highway and Traffic Engineering from IIT Kharagpur (India) in 1989. Presently working at Shri Ramdeobaba College of Engineering and Management, Nagpur , as Associate professor in Civil Engineering Department. His main field of work has been pavement and traffic engineering. He has guided many undergraduate and post graduate dissertations. Presently he is persuing his Doctoral research in the field of traffic safety.

V.S.Landge, born on 02-10-68, graduated from RKNEC, Nagpur (India) in 1991.He completed his post graduate studies in Transportation Engineering from BITS Pillani(Rajasthan, India) in 1993. He completed his Doctoral degree from IIT Roorkee(India) in 2006. Presently working at VNIT, Nagpur as Associate professor in Civil Engineering Department. His main field of work has been traffic engineering particularly in the field of traffic Safety. He is a member of the State Technical Agency for Pradhan Mantri Gram Sadak Yojna for Vidarbha Region in India He has guided many undergraduate and post graduate dissertations.

N.V.Deshpande, born on 18-07-1956, graduated from VNIT, Nagpur (India).He completed his post graduate studies in Structural Engineering from VJTI, Mumbai (Maharashtra, India). He completed his Doctoral degree from VNIT Nagpur (India). Presently working at GNIET Nagpur as Professor and Principal. His main field of work has been Structural Engineering. He has guided many undergraduate and post graduate and Doctoral Thesis.