ISSN 2250-3153

Development of model for road accidents based on intersection parameters using regression models

Dr. NSSR Murthy[1] and R. Srinivasa Rao

[1]Professor, Mumtaz College of Engg & Tech, Malakpet, Hyderabad

Abstract- The issue of road accidents is an increasing problem in developing countries. This could be due to increasing road traffic/vehicle occupancy. This has been increasing over years. Regulating traffic on roads is an important task. There by reducing accidents in accident prone zones. The accident was drastically increased over a decade from 4% to 31%. This is a alarming issue. The analysis and identification of such road accident prone zones is essential to reduce the accidents. A model was developed based on intersection parameters and no. of accidents by regression analysis.

Index Terms- intersection parameters, major traffic, minor traffic, correlation coefficient, regression analysis, regression model

INTRODUCTION

The issue of road accidents is an increasing problem in developing countries. This could be due to increasing road traffic/vehicle occupancy. This has been increasing over years. Regulating traffic on roads is an important task. There by reducing accidents in accident prone zones. The accident was drastically increased over a decade from 4% to 31%. This is a alarming issue. The analysis and identification of such road accident prone zones is essential to reduce the accidents. The factors influencing such accidents are to be analyses for remedies. Using regression Models, factors influencing road accidents have analyzed using SPSS model. The study area selected is part of Hyderabad. Telangana (GHMC) with heavy traffic.The objective of the paper is to review relation between accident per year and intersection Parameters. To develop prediction models and test their validity. To suggest improvement measures to prevent road accidents and to derive a model for accident parameters.

Factors responsible for road accidents:

The paper considers Identifying factors influencing road accidents, identifying intersection parameters and collection of data. Using statistical methods, regression models have to be developed

Factors influencing road accidents:

These can be grouped into following.

1. Vehicle related factors: this may be due to inherent design limitations or defects to lack of maintenance, failure of components like brakes, tires and lighting. Visibility, speed and vehicle lighting are also important.
2. Road related factors: this includes pavement design and conditions, horizontal curves, insufficient lane and shoulder width, vertical curves.
3. Road user related factors : psychological factors of the users, alertness and intelligence, patience of driver, drivers experience and age
4. Environmental related factors: rain, reduced visibility, bad weather etc. heavy fog and mist and heavy rain also plays important role.

Analysis considerations:

1. A data of 1-3 year was considered for moderate to heavy Traffic locations
2. Segments of minimum 0.15 km in region of accident prone locations were taken for analysis.

Methods for identifying high accident locations:

a) Spot map method: simplest way is to examine accident spot map which gives spots of max no. of accidents.
b) Accident frequency method: it is based on accidents frequency; locations are classified from high to low accident areas.
c) Accident rate method it is comparing no. of accidents at location with no. of vehicles travel resulting accident rate.

Earlier models survey:

ISSN 2250-3153

Neumann and Glenn on (1982) described a theoretical model that relates accident on crest curves to available sight distance. The development of this model was not based on accident data, rather the model

relied on intuitively logical relationship and engineering judgment. Highway designers to systematically evaluate the cost effectiveness spot improvement of location with deficient SSD's can use the model.

The model is as follows:

$$N = ARH(L)(V) + ARh\ (Lr)(V)(F_{ar})$$

Where,

N = Number of accident on a segment of highways containing a crest curve.
ARH = Average accident rate for specific highways
L = Length of highway segment in miles.
V = Traffic volumes in millions of vehicles.
L = Length of restricted sight distance in miles
F_{ar} = A hypothetical accident rate factor that varies according to both the severity of the sight restriction and the nature of the hidden hazard.

Glen on (1983) developed model in the basis of the available literature, sufficient evidence appears to indicate that in general, horizontal curves experience higher accident rates than tangents and that accident rates and that accidents are generally increase as a function of increasing degree of curvature. Glen on horizontal Curve Model reported in the FHWA report accident relationship is presented below:

$$A = AR_S\,(L)\,(V) + 0.0336(D)\,(V)$$

Where,

A = total number of accidents on the segment
ARs = accident rate on comparable straight segment in accident
L = length of highway segment in miles
V = Traffic volume in millions of vehicles
D = curvature in degrees
Lc = length of curved component in miles

Zegeer er al (1992) developed the following accident prediction model for the 1991 FHWA study cost- effective improvements for horizontal curves.

$$A = [\ (1.552)(L)(V) + (0.014(D)(V) – (0.012)(S)(V)](0.978)^{W}\ )$$

Where,

A = number of total accidents on the curve in 3 years period
L = length of curves
V = volume of vehicles in million vehicles passing through (both direction) in a 5 Year time
D = degree of curve
S = presence of spiral
W = width of roadway (twice the lane plus shoulder width) on the curve.

To evaluate the combined effect of road elements, Transportation and Research Laboratory (TRRL) carried out research work in Kenya and Jamaica. The equations developed in this study are:

$$Y = 1.45 + 1.02X_5 + .017X_3 \text{ (KENYA)}$$
$$Y = 1.09 + 0.031X_3 + 0.62X_5 + 0.0003X_4 + 0.062X_2 \text{ (JAMICA)}$$
$$Y = 5.77 + 0.755X_1 + 0.275X_5 \text{(JAMICA)}$$
$$Y = 5.77 + 0.755X_1 + 0.275X_5 \text{ (JAMICA)}$$

ISSN 2250-3153

Where,

Y = rate per million vehicle kilometer per year
X_1 = road width (m)
X_2 = vertical curvature (m/Km)
X_3 = horizontal curvature (degree/Km)
X_4 = surface irregularity (mm/Km)
X_5 = junctions per Km.

Charles V zegeer et al (1995) developed Accident Relationships of Roadway Width on low-VolumeRoads Horizontal curves represent a considerable safety problem on rural two lane highways. A 1980 study estimated that there are more than 10 million curves on the two-lane highway system in the U.S. Accident studies further indicated that curves experience a higher accident rate than do tangents, with rates that range from one and a half to four times higher than similar perhaps be more important in light of improvements being made related to resurfacing, restoration, and rehabilitation projects, commonly known as the 3R program. Using this model form, the variable indicating spiral transitions was also found to be statistically significant, and the basic estimated model for total accident rate was given by:
With all coefficients significant at the p = .0001 level (standard errors are shown in parentheses).
Total accidents per curve per year could be estimated by the model

$$\text{Accident} = (\text{ADT})(\text{Length})(1.94+0.24\text{Deg}-0.026\text{Width}-0.25\text{Spirals})$$

Luis F. Miranda-Moreno et al (2005) developed Alternative Risk Models for Ranking Locations for Safety Improvement. The authors compared the performance and practical implications of these models and ranking criteria when they are used for identifying hazardous locations. This research investigates the relative performance of three alternative models: the traditional negative binomial model, the heterogeneous negative binomial model, and the Poisson lognormal model. In particular, this work focuses on the impact of the choice of two alternative prior distributions (i.e., gamma versus lognormal) and the effect of allowing variability in the dispersion parameter on the outcome of the analysis.

Henry C. Brown et al (2006) considered the effects of access control on safety on urban arterial streets. Access control techniques are used to improve traffic performance and safety on highways. One important benefits of access control is improved safety. For a quantitative assessment of the benefits of access control on safety, impact models are needed to predict crash frequencies based on the geometric and access control characteristics of the segments. The objective of this research was to develop regression models to predict crash frequencies on urban multilane arterial segments. To develop these models, data were collected on geometric and access control characteristics of the segments and the number of crashes on the segment by severity type. Negative binomial regression models were developed to predict the total number of crashes, number of property-damage0only crashes, and number of fatal and injury crashes.

Fajaruddin Mustakim etal (2007) did his study on block spot study and accident prediction model using multiple regression linear models. The study area was federal Route (FT50) Batu Pahat – Ayer Hitam. The regression model was

$$\text{In (APW)}^{0.5} = 0.0212(\text{AP}) + 0.0007\,(\text{HTV}^{0.75} + \text{GAP}^{1.25}) + 0.0210(85^{th}\ \text{PS})$$

Where,

APW = accident point weight age.
AP = number of access points per kilometer
HTV = hourly traffic volume
Gap = amount of time, between the end of one vehicle and the beginning of the next in second.
85^{th} PS = 85^{th} percentile speed

The model has R-square of 0.9987

The results of the paper have shown that the existence of a larger major junction density, an increase in traffic volume and vehicle speed in federal Route 50 is the contributors to traffic accident. Reduction of vehicle speed, access point, traffic volume and gap are likely to have influential effect on the road traffic accidents.

Fajaruddin Mustakimetal (2008) developed a model for Black Spot Study and Accident prediction Model using Multiple Linear Regression. This approach is considered to be helpful for strengthening freeway management and reducing the likelihood of incidents. The freeway incident model developed in this study can constitute a useful decision support tool for analysis of traffic accident and the implementation of freeway patrol systems. Further, it is useful for selecting optimal location of shoulder rumble strips and sonic napping alert system.In case of knowing the trip length or travel time by the AVIS, the applicability of developed model will be higher. Future research needs the more specific accident model considering vehicle types such as passenger car, bus, and freight vehicle.

Ali ,Osman Yielder etal (2009) used fractional factorial method for the sensitivity analysis of accident prediction model. The evaluation of sensitivity analysis indicated that average daily traffic (ADT), lane width (W), width of paved shoulder (PA), median(H) and their interactions (i.e., ADT-W, ADT-PA and ADT-H) has significant effects on numbers of accidents. The effects due to each parameter and parameter interactions are estimated using the following equation:

$$E_j = S \quad E_j = S_{ij}(X_i)/N_j$$

In which Ej represent the effect of the jth factor (i.e., in jth column), n the total number of experimental runs (i.e., n=8), Sij represents the sign in row i and column j, Xi represents the value of the prediction variable obtained from the ith experimental run and Njis the number of "+" signs in column j. Accident prediction model developed is given below.

$$A= 0.0019(ADT)^{0.882}(0.879)^{w}(0.919)^{PA}(0.932)^{up}(1.232)^{H}*(0.882)^{T1}(1.322)^{T2}$$

In which A is the number of run-off-road, head on, opposite direction sideswipe, and same-direction sideswipe accidents per mile per year, ADT the two-directional average daily traffic, W the lane width in feet, PA the width of paved shoulder in feet, UP the width of unpaved (gravel, turf, earth)shoulder in feet, H the median roadside hazard rating for the highway segment, measured subjectively on a scale from 1 (least hazardous) to 7 (most hazardous), T1 = 1 for flat terrain, 0 otherwise. T2 = 1 for mountainous terrain, 0 otherwise.

Av rascleetal (2010) A camera system has been installed on a stretch of 20 km length at a construction site on this highway. Since then traffic data have been recorded. The purpose of the camera supervisor is not on accidents reduction the cameras do not fully capture the traffic and therefore so far data of only one accidents could be retrieved this is clearly insufficient to calibrate the model never the less we present the available data. The reported accidents happen at 328 km on the highway A1 with cars travelling in the direction. We present density & velocity data at a 383 km to the accidents at the side of the accidents 383 km and at kilometer 381 after the accidents referring to condition.

Donatus lygas, Velmajasiniene etal (2011) analysis of accidents prediction feasibility on the roads of Lithuania. Modeling of road accidents carried out based on 1997-2011. Data on fatal & injury accidents on the main roads enabled to construct mathematical models for the optimum selection of road safety improvement measures one of the constructed mathematical models enables to optimally select road safety improvement measures to mathematically reduce the no. of people killed under unrestricted amount of trends.

$$YS = 193.52+0.2t-67.7t1-69.54t2-14.48t3$$

Where Y = forecasted average value

T = trend variable Ti variable taking the value in the quarter 1 of the year & value 0 in the other quarter the fourth quarter at the year is corresponding by the value of variable.

Lazim Abdullah nurnadian zamri (2012) developed road accidents models with two threshold linear fuzzy linear regression. The research was design to determine the model road accident a fuzzy regression model was developed in describing road accidents in Malaysia over the period of 1974 to 2007 using three prediction thresholds. Level h=0.5 were accounted in this paper the model structure was developed after consider linearity assumption of model variable the model for road accidents uses registered vehicle population and road length as variable based on one response variable of road accident. The model shows that the variable of registered vehicle & population provide higher impact to the number of road accident.

$$Ra= (35436.69312.458.141052)+(0.022814)Rv+(0.000220)Po+(0.023)R$$

Multiple Linear Regression Model is if Y is the number of accident per year as dependent variable and X_1,X_2 and X_3 are the independent variable representing the traffic volume (AADT), length pf the segment (km) and number of median openings per km respectively. Then the Multi Linear Regression model is in the forms as follows

$$\boxed{Y = D + AX_1 + BX_2 + CX_3}$$

D is constant, A, B, C are proportional coefficients.

Methodology:

The proposed methodology is as shown in fig1.

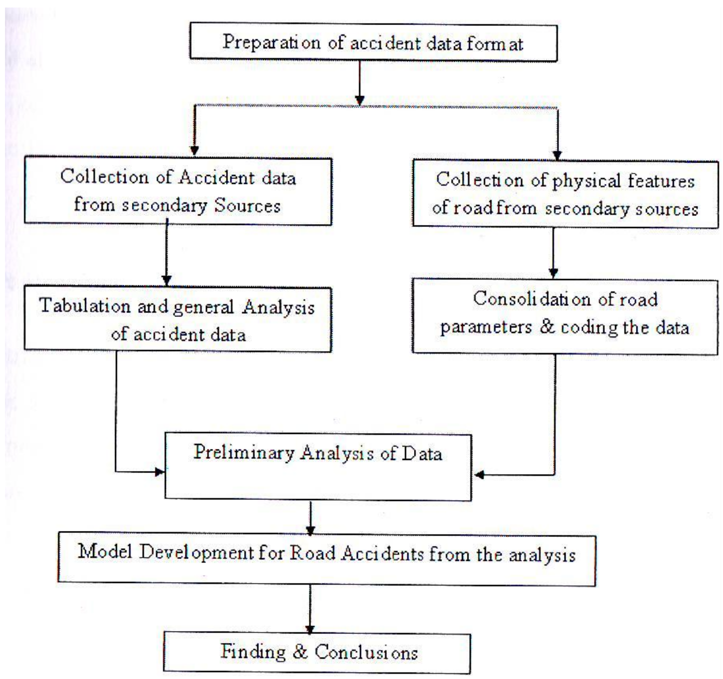

Fig.1 flow chart for methodology

Detailed analysis will be carried out through monthly, annual and hourly data and analyzed. Data collection: There are 6 selected intersections in selected segments. Intersection details are as given in table1. The data for above areas and intersections are collected from police records in accidents/year, traffic volume (major and minor road), turning traffic volume in vehicles/day, pedestrian volume etc.. Data is to be extracted from Accident data of whole record and also some data from HMDA and R&B department. Primary and secondary data was collected for Cybarabad city, from state police department and national crime records bureau. It includes no. of vehicles involved, Time, severity etc., the data was from 2007 to 2011. Data was also collected from cybarabad police stations of Balanagar Zone

Table 1 INTERSECTION Details

CHAINAGE(KM)	INTERSECTION PLACE	TYPE
1	Chandanagar	T- legged
6.6	Miyapur	T- legged

9.6	Jntuh	4- legged
16.4	Balanagar	T- legged
17.6	Sanathnagar	T- legged

Analysis and results:

Using curve fitting technique & SPSS software package, curve fitting is done to test relation between road intersection parameters and accident rate by Microsoft excel. SPSS software takes data from excel file and generates tabulated reports, charts and distribution trends. (Figs. 2-10). For each model, regression coefficients, variances were found. Based on statistical analysis of secondary data of road accidents, relation between accidents and intersection parameters was found .like Accidents /year vs major road volume, minor road volume, turning traffic volume, pedestrian volume, approach width, turning radius , speed and no of legs.

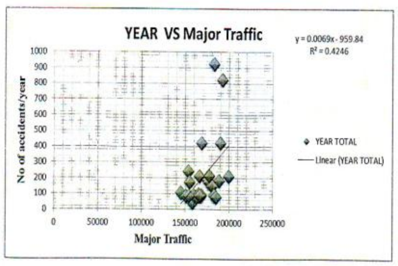

Fig 2 Major Traffic

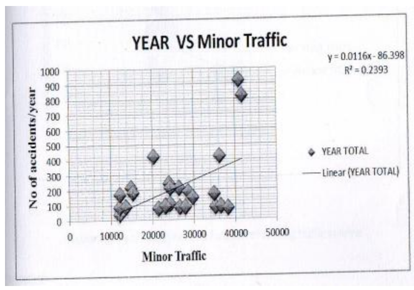

Fig 3 Minor traffic

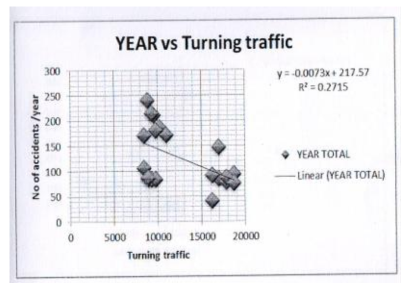

Fig 4 Turning Traffic

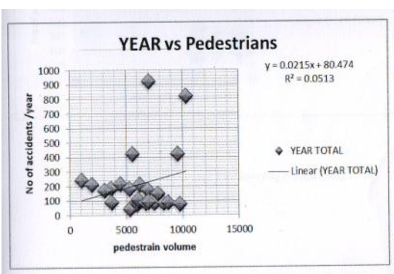

Fig 5 Pedestrians

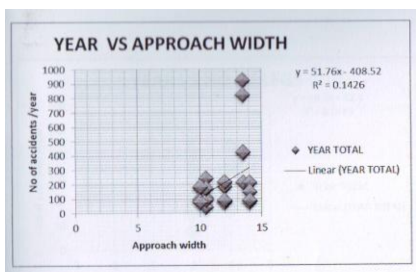

Fig 6 Approach Width

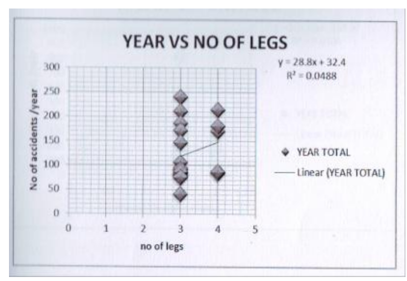

Fig 7 No.of Legs

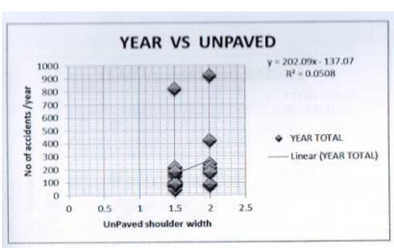

Fig 8 Unpaved

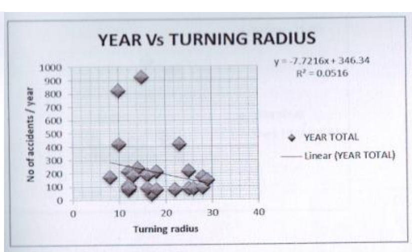

Fig 9 Turning Radius

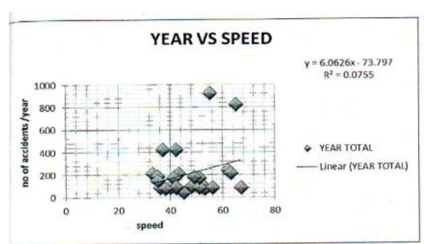

Fig 10 Speed

A model was developed between accident rate and intersection parameters using SPSS regression analysis. Scatter plots were drawn between accidents /year as a function of parameters major road volume, minor Volume, turning traffic volume, pedestrian volume, approach width, no. of legs, unpaved shoulder width and turning radius. Regression correlation was found. From the above trend line approach, variation of accident rate with respect to intersection parameters is estimated and a model was developed. Based on regression model of each intersection parameter prediction model was developed from the data as shown below in table 2

Table 2: regression coefficients

VARIABLES COFFICENTS	MODEL 1	MODEL 2	MODEL 3	MODEL 4
CONSTANT	67.5711	324.0681	-65.1354	187.2354
MAJOR TRAFFIC	…….	0.0000	-0.0001	0.0005
MINOR TRAFFIC	0.0077	……	0.0081	-0.0039
TURNING TRAFFIC	-0.0025	-0.0059	…….	-0.0069
PEDESTRIAN VOLUME	-0.0340	-0.0207	0.0346	………
APPROACH WIDTH	-1.6887	-0.8866	-2.2525	-0.2387
NO OF LEGS	84.8679	-7.8870	105.8817	-21.6706
UNPAVED SHOULDER	-80.4882	-57.8540	-47.0821	23.7664
TURNING RADIUS	2.8751	2.9388	2.7246	1.0413
SPEED	-1.5654	-0.2140	-1.6650	0.6787
Correlation, r	0.897	0.844	0.894	0.668
Co-efficientof determination,r^2	0.805	0.712	0.8	0.447

A model form was developed. Among these models from table 2, model 4 all parameters shows maximum relation with accident rate and hence adopted.

Model form
Accident rate, Y = 5.0 E -04 MRV -3.86 E -03mrv – 6.91 E -01AW -2.17 E+01NL + 2.38 E+01UP + 1.04 E+0.00TR +6.79 E-01V + 187.235

Conclusions:

Relation between accidents/year and various intersection parameters were found and a model was developed from regression analysis. As number of intersections increase, accident rate increases, major traffic, unpaved shoulder, speed and turning radius have positive relation with accident rate. Minor traffic shows negative relation with accident rate.

References:

1.**Charles V Zegeer (1995)** " Accident Relationships of Roadway Width on Low-Volume Roads Horizontal curves".TRB.
2. **Andrew Vogt and Joe Bared (1998),** "Accident Models for Two-Lane Rural Segment and Intersection " , TRB Paper No. 98-0294.
3. **Arun Chatterjee (2005),** " Planning Level Regression Models For Crash Prediction on Interchange and Non-Interchange Segment of Urban Freeways" , Southeastern Transportation Center, Center for Transportation Research, The University of Tennessee, Knoxville, TN 37996-4133 August ,2005.
4. **Andrew P. Tarko (2006)** " Calibration of Safety Prediction Models for Planning Transportation Networks" , Transportation Research Record 1950, pp. 83-91.
5**. Dr. B.N.Nagaraj, M.V.L .R . Anjaneyulu & K.S.Vipin (2002)**," Systematic Segmentation Approach for Prediction of accidents at intersections un urban areas" , Journal of the Indian Roads Congress, Aug 2002, vol63-3, 625-646.
6.**Dominique Lord and Bhagwant N.Persaud (2003),** " Accident Prediction Models With and Without Trend: Application of the Generalized Estimating Equations (GEE) Procedure", Paper No. 00-0496.
7**. El- Basyouny and Sayed T (2006),** " Comparison of Two Negative Binomial Regression Techniques in Developing Accident Prediction Models" , Transportation Research Record 1950 pp. 9-16.
8. **Fajaruddin Mustakim (2008),** " Black Spot Study and Accident Prediction Model Using Multiplication Linear Regression".Advancing and intergrating construction education, research and Practice, August 4-5, 2008.
9**. Henry C. Brown And Andrzej P. Tarko (2005),** " Effects of Access Control on Safety on Urban Arterial Streets" ,Transportation Research Record 1665 paper No. 99-1051.
10**. Hauer, E., J.C.N Ng, and J.Lovell (1989),** " Estimation of Safety at Signalized Intersections" ,In Transportation Research Record 1185, TRB, National Research Council, Washington, D.C., 1989, pp.48-61.
11. **Jianming Ma Bayesia,"** Multivariate Poisson Regression for Models of Injury count, by Severity".Transportation Research Record 1782.
12**. L.R Kadilayi,** "Traffic and Transportation Planning", text book, pp. No 428-438.
13. **Ma,.J. Kockelmen, K. M**, "Bayesian Multivariate Poisson Regression for Models of Injury Count, by Severity" , Transportation Research Record 1950, pp. no 24-34.

14. Ministry of Road Transportation and Highways, Govt. of India, New Delhi, “Manual for safety in Road Design”, 1998.
15. **Miranda-Moreno LF FU L Saccomanno FF Labbe A,**” Alternative risk models for ranking locations for safety improvement”, Transportation Research Record 1908, pp.1-8.

16. **Neuma and Glenn on (1982)**,” A Theoretical model that relates accident on crest curves to available sight distance “, Transportation Research Record 923.
17. **Paul Gutsokie** , “ SPECIAL REPORT” , IRTAD , Transport Canada August 2003.
18. **Sandip Chakraborty and Sudip K. Roy (2005),**” traffic accident characteristics of Kolkata”,Transportation and Communications Bulletin for Asia and the Pacific No. 74, 2005.
19. **Seungkirl Baek, Hyunho Chang, Byoungjo yoon(2005),**”An effect of trip length on freeway accident rates “, journal of the Eastern Asia Society for Transportation Studies, Vol. 6, pp. 3467-3481, 2005.
20. **Shane Turner, (2002),** “Accident Prediction Models for Signalized Intersection’s”.Transportaion engineer, Beca Infrastructure.
21. **T.Sayed (2005),** “Development od an Accident Prediction Model Using GLIM (Generalized Log-Linear Model) and EB method: A case of seoul”, Journal of the Eastern Asia Society for Transportation Studies (2005), Vol.6, pp. 3669-3682, 2005
22. **Tom Brijs, Dimitris Karlis, Filip Van den Bossche and Geert Wets(2004) ,** “ A Model for Identifying ans Ranking Dangerous Accident Locations”, A Case-Study in Flanders, Transportation Research Institute.
23. **Gardner W. Mulvey E.P and Shaw E.C (1995).** “Regression Analyses of Counts and Rates: Poisson Over dispersed Poisson, and Negative Binomial Models, “ Psyhological Bulletin, 118, 392-404.

AUTHORS

First Author – Dr. NSSR Murthy MTech Ph.D Mumtaz college of Engg & Technology Malakpet Hyderabad murtynssr@gmail.com

Second Author – R Srinivasa rao MTech . Mumtaz college of Engg & Technology Malakpet Hyderabad rsrao123 @gmail.com

Correspondence Author – Dr. NSSR Murthy MTech Ph.D Mumtaz college of engg & Technology Malakpet Hyderabad murtynssr@gmail.com