RESEARCHERID
THOMSON REUTERS

[Reddy, **10(1): January-March 2020]** **ISSN 2277 – 5528**
Impact Factor- 5.085

INTERNATIONAL JOURNAL OF ENGINEERING SCIENCES & MANAGEMENT

TRAFFIC ACCIDENTS USING LOGISTIC REGRESSION IN FORMATIVE EFFECTIVE VARIABLES AND PREVENTIVE MEASURES

Kandukuru Jagan Mohan Reddy [*1] Asst professor, M.Durga[2] Asst Professor & V.Seerisha[3] Asst Professor
Dept. Of Civil Engineering, Noida International University- U.P., INDIA

ABSTRACT

This research was conducted to determine the important influential variables upon the deaths from road traffic accidents and effect of each of those upon the studied phenomenon through applying logistic regression model. The maximum likelihood method was used to estimate parameters to determine the explanatory variables effect. Wald test was used to determine the significance of the explanatory variables. The data set used in this research consists of a sample of (212) observations and was obtained from the records of the directorate traffic- Garmian. The accident victims is response variable in this study and it is a dichotomous variable with two categories. The study led to a number of conclusions, among them; logistic regression models fit such data, three explanatory variables were found most significantly associated to accident victims response variable namely; high speed, car type, and location. Road accidents are one of the main causes of deaths worldwide. About half a million people are killed in road related crashes every year throughout the world. The probability of an accident occurring is influenced by numerous factors like roadway geometric characteristics, vehicle characteristics, pavement conditions and weather conditions each of these factors contribute its own share towards occurrence of accidents. This paper discusses the influence of various factors on accident caution based on statistical package Regression Analysis collected from the most accident prone stretch, Ayalurmetta to Thammarajupalli (30KM) in Andhra Pradesh on NH18 and from the results of the analysis, it can be concluded that this National Highway section needs improvement from safety point of view. A large number of accidents have been occurring over such a small section of 15 km length. Proper traffic guidance and control system to guide road users ensuring safe movement of vehicles has been recommended and some of the facilities such as pedestrian crossings and median openings, acceleration and deceleration lanes were re-designed in order to improve the safety of the road and minimize the accidents.

Keywords: Traffic accident, logistic regression model, maximum likelihood estimation, Wald's test.

I. INTRODUCTION

The problem of deaths and injury as a result of road accidents is now acknowledged to be a global phenomenon with authorities in virtually all countries of the world concerned about the growth in the number of people killed and seriously injured on the road. The word report on road traffic accident prevention has indicated the worldwide, an estimated 1.2 million people died in road traffic accident each year and as many as 50 million are being injured [3][6]. The logistic growth function was first proposed as a tool for use in demographic studies by Verhulst (1838, 1845) and was given its present name by reed and Berkson (1929). The function was also applied as a growth model in biology by Pearl and Reed (1924) [7]. Logistic regression is firstly developed by statistician.

D. R. Cox in 1958 as a statistical method, and after that it is used widely in many fields, including the medical and social sciences [12]. Logistic regression is used for prediction by fitting data to the logistic curve. It requires the fitted model to be compatible with the data. In logistic regression, the variables are binary or multinomial. Gordon (1974) pointed out that logistic regression models have plays a major role in biological and medical applications where cross- classified tables with large numbers of cells are typically replaced by a logistic or log- linear relationship among the variables, thus obviating the need for the table[14]. Logistic regression was first proposed in the 1970s as an alternative technique to overcome limitations of ordinary least square regression in handling dichotomous outcomes. It became available in statistical packages in the early 1980s [10]. For logistic regression, least squares estimation is not capable of producing minimum variance unbiased estimators for the actual parameters. In its place, maximum likelihood estimation is used to solve for the parameters that best fit the data [15].

ISSN 2277 – 5528
Impact Factor- 5.085

Bedard (2002) used the multivariate logistic regression model to determine the independent contribution of crash, driver and vehicular characteristics that lead to increasing driver's fatality risk. Reducing speed, increasing the use of seatbelts and reducing severity incidences attributed to driver side impacts was found to be preventing fatalities [8]. Mhamad (2011) used the logistic regression model in traffic problems in Sulaimani, the purpose of the study was that the modeling of traffic accidents linking the accident fatality and the various factors that cause it [2]. Odhiambo (2015) used artificial neural network to model the monthly number of road traffic injuries and the negative binomial regression model as our baseline model and noted that accident data are non-negative integers.

Figure 1.1: shows the key map of the section

II. TECHNICAL APPROACH AND ANALYSIS

1. Methodology of the Study

It is a known fact that the width of the road plays an important role in accident causation. The alignment of the road is also an important parameter which affects the accident rate. Alignment of the road affects sight distance. Adequate sight distance is essential for the safe traffic operations. The alignment is classified into straight alignment, curve alignment. For a straight aligned road there will be adequate sight distance. The sight distance goes on reducing as the degree of curvature increases. Increase in the number of side roads interrupts is an important parameter which affects the accident rate. When side roads increases, accident rate also increases.

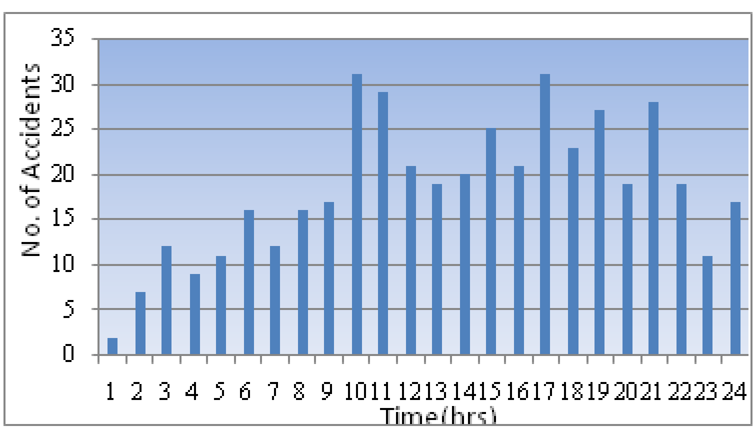


Figure : Accident Distribution by Time

Regression Analysis

ISSN 2277 – 5528
Impact Factor- 5.085

1. Assumptions

The technique of predicting the values of one variable (called the dependent variable) from measurements of the other (called the independent variable)is called "REGRESSION ANALYSIS". If the relation between the dependent and independent variables is linear, the analysis is known as "LINEAR REGRESSION". If the independent variables are two (or) more in number, the analysis is known as MULTIPLE LINEAR REGRESSION ANALYSIS.

If the independent variable is denoted by Y and the independent variable by X .The linear relationship between the two can be of the following form:

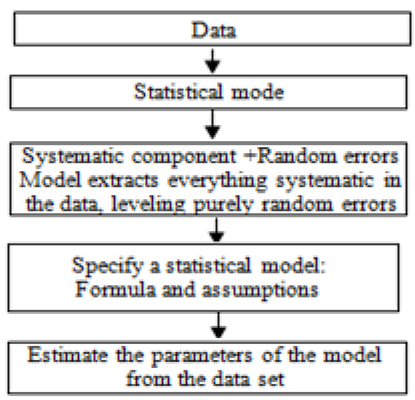
$$Y=a+bX+Є$$

Where a and b are constants and Є it is a term denoting the random error. The random error term Є will be small if the two variables are closely related. Thus the observed values of X and Y are:
$(x_1,y_1),(x_2,y_2),\ldots\ldots\ldots..(x_n,y_n)$,then

$$Y_{i=}\ a_i+bx_i+Є$$

The problem is to estimate the values of a and b such that the sum of the $Є_i$'s ($Є_1+Є_2+Є_{3\ldots\ldots\ldots}+Є_n$) is as small as possible and thus minimize the likely error in prediction , this is done by method of least squares, thus a and b are chosen to make a minimum.

The following are some the condition which must be satisfied if a multiple linear regression analysis is to be used:



- All the independent variables must be independent of each other and there should be no correlation between them.
- All the variables are normally distributed.
- All the variables are continuous.
- A linear relationship exists between the dependent variable and the independent variables.
- The influence of the independent variables is additive.

i.e. the inclusion of each variable in the equation contributes a distinct portion in the estimation of the dependent variable and all the variables together contribute additively in the estimation.

2. Model Building

$$R=\sum_{i=1}^{n} Є_i^{2}$$

where n is the total number of points observed

3. Multiple Linear Regressions

The statistical technique which will be most frequently encountered by a traffic engineer and traffic planner is the multiple linear regression analysis. The problem concerns with the establishment of relationship between a variable which is known to respond to changes in two or more other variables. The variable which is known to respond, Y variable, is commonly called the independent variable, and the other variable influencing it is called the independent variables, i.e X variable. The function will be the following form:
$Y=a_0+a_1x+a_2x_2+a_3x_3+...+a_mx_m$ Where Y=true estimate of the dependent variable,y $x_1,x_2,x_3........x_m$=m independent variables a_0=regression constant
$a_1,a_2,a_3.......a_m$=regression coefficients of the respective m independent variables.

The collected data were statistically analyzed to evaluate the effect of the selected parameters on accidents. The relationship between the accidents and various factors were also obtained. The data was analyzed using SPSS (Statistical Package for Social Sciences) software. Here the number of accidents was taken as the dependent variable and width of the road, alignment of the road, number of side roads and traffic volume were taken as independent variables. The table shows the ten points which we had taken in the stretch. This ten places shows more no of accidents happened from the past five years data which we have collected. This accident data is collected form our nearest police stations. And the traffic volume is collected for a week(day and night) The table shows the width and alignment and no of side roads in the ten selected points. The points are mentioned in the below Table 3.9

Name of The Stretch	*Width (M)*	*Alignment*	*No of Side Roads*	*Volume (PCU/ Hr)*
KMC to Thammarajupalli	6.8	Straight	1	818.35
Thammarajupalli to BSNL tower	6.5	Curve	1	896.35
BSNL to Sugalimitta	6.5	Curve	0	893
Sugalimitta to Panyam	7.9	Straight	4	965.52
Panyam to RGM college	8.5	Curve	3	1253
RGM college to Balapanuru	7.9	Straight	2	978.3
Balapanuru to Venkateshwarapuram	7.4	Straight	1	965.73
Venkateshwarapuram to K.C canal	6.9	Straight	3	997
K.C canal to Nandyal Public School	7.5	Curve	2	974.29
Nandyal Public School To Ayalurmetta	7	Curve	2	962.44

Where speed is in km/h and UI is unevenness index in mm/km. The values given in parentheses are the 't' values of coefficients, which are significant at 5 per cent level.

Effect of roughness on passenger car unit (PCU)
The PCU of a vehicle type was calculated by Equation (2). These are given in Table. Figs show the variation in PCU for different types of vehicles with road roughness at different sections. As may be seen, the PCU for a vehicle type decreases linearly with roughness, the slope of linearity depends on the type of vehicle. Non-motorized vehicles are not much influenced by road roughness while speed of motorized vehicles is greatly influenced. The fast moving vehicles are substantially influenced by road roughness while its effect on all other vehicle types is relatively low.

Speed±Volume Relationship

Speed, density, and volume are the most important components of a traf®c stream for estimating the traf®c carrying capacity of a road. Since the measurement of traf®c density in a mixed traf®c situation is dif®cult, attempts have always been to concentrate on the speed±volume relationship for different type of roads. For the determination of a speed±volume relationship in heterogeneous traf®c conditions, the total vehicles recorded for each counting period were converted into an equivalent number of the PCU's using values given in Table 3.

In a mixed traf®c situation, a large variation in speeds among slow and fast moving vehicles exists. Therefore, the spot speed or the space mean speed, as normally calculated for the homoge-neous traf®c, cannot be considered for the mixed traf®c. It needs to be modi®ed to suit the heterogeneous traf®c conditions. For this purpose, many researchers suggest the use of a weighted

Table . Calculation of Passenger Car Unit Factors

Vehicle type	Relation between passenger car unit and carriageway width ~w!	R^2 value
Bus	PCU[b]_0.1114*w* _3.073	0.92
Truck	PCU_0.146*w* _4.40	0.95
LCV	PCU_0.097*w* _1.956	0.99
Tractor Trailer	PCU_0.103*w* _4.95	0.99
Three-Wheeler	PCU_0.168*w* _0.327	0.95
Two-Wheeler	PCU_0.017*w* _0.158	0.97
Cycle	PCU_0.034*w* _0.225	0.99
Rickshaw	PCU_0.054*w* _1.132	0.97

[a]LCV indicates light commercial vehicle.
[b]PCU indicates passenger car unit

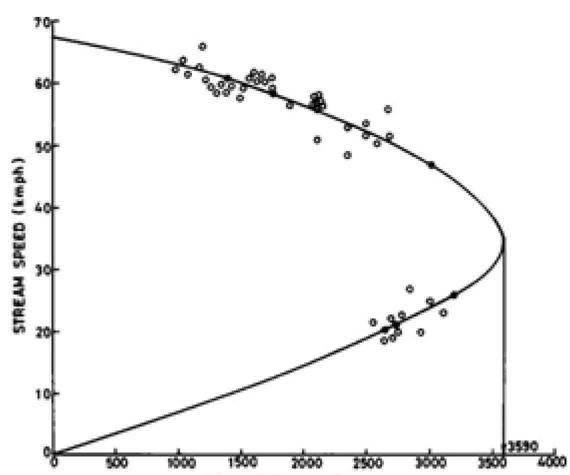


Fig. . Speed±volume relationship at section I

space mean speed or, simply, mean stream speed. To arrive at the mean stream speed, a trap of suitable length is made on the road and the speed of each category of vehicles considered for the count is calculated. The mean stream speed or weighted space mean speed is then given by

$$V_m = \frac{\sum_{i=1}^{k} n_i v_i}{\quad} \qquad (3)$$

k

$^{(}i\ \_1\ ^{n}\ i$

where k_total number of vehicle categories present in stream,
v_m _mean stream speed ~km/h!, v_i _speed of vehicle of category
i ~km/h!, and n_i _number of vehicles of category i.

The average stream speed calculated by Eq. ~3! was plotted against the traf®c volume. Typical curves showing speed±volume relationships are given in Figs. 4 and 5.

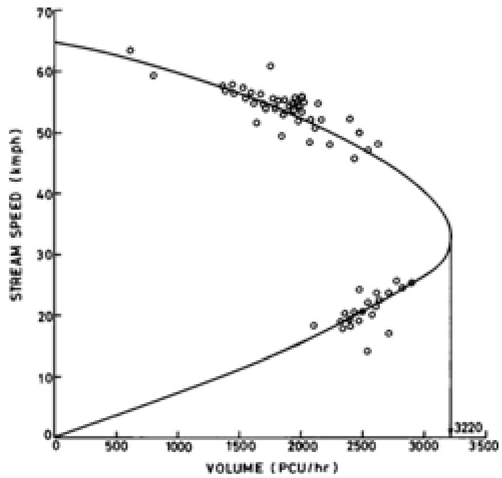


Fig. Speed±volume relationship at section IX

III. TEST RESULTS AND DISCUSSIONS

General

The stretch is taken from Ayalurmitta to Tammarajupalli. By taking the accident data from the near by police stations we conclude that the total no of accidents happened for the 5 years.And we had selected 10 main points where mostly accidents happened. This ten main points is shown in table no 2 .9.and by this a latest technique called Regression statistics we can find the no of accidents to be happened in future days.

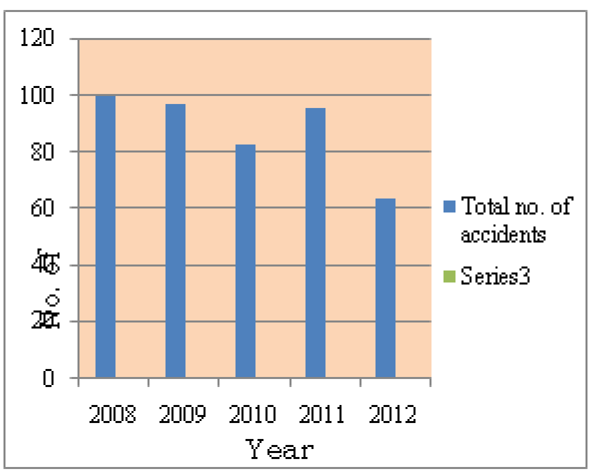


Figure Showing Year vs Number of Accidents

Table : No of accidents held from 2008-2012.

ISSN 2277 – 5528
Impact Factor- 5.085

Year	Total No. of Accidents
2008	100
2009	97
2010	83
2011	96
2012	64
Total	440

- **Vehicle Analysis**

For this analysis, vehicles involved in all types of accident have been studied. We found that out of all the vehicles involved in crashes reported on the NH18 during the study span, 123 of them were trucks. The total number of vehicles involved in crashes were shown in the Table 3.4.

Table .Vehicle Distribution

Year	Auto	Car	Buses	Trucks	Other Vehicles	Tractor	Two Wheeler
2008	36	16	5	21	1	6	10
2009	18	23	11	26	1	6	12
2010	15	20	6	29	2	9	11
2011	14	26	8	27	0	5	15
2012	16	9	5	20	0	6	7
Total	99	94	35	123	4	32	55

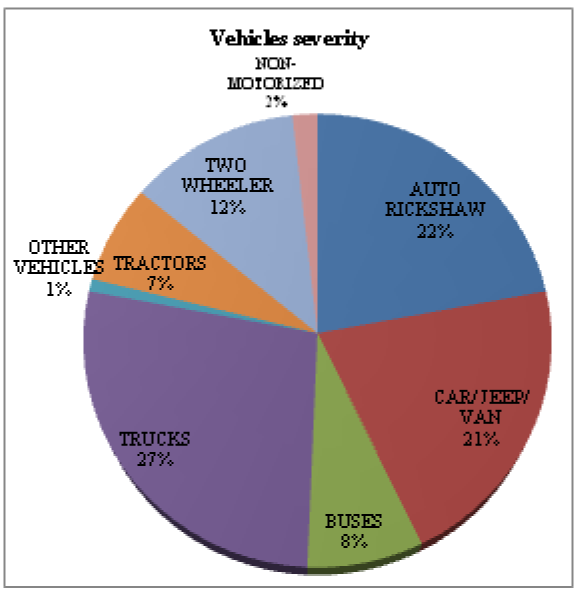


Figure Vehicle Distribution

- **Accident Distribution by Time**

The distribution of accidents by time is show in figure 5.3Maximum number of accident recorded between that is 31 accidents took place between 09:00 to 10:00 AM and 04:00 to 05:00 PM.

Table . : Accident Distribution By Time

S.No.	Time	Total No. of Accidents
1	0-1	2
2	1-2	7
3	2-3	12
4	3-4	9
5	4-5	11
6	5-6	16
7	6-7	12
8	7-8	16
9	8-9	17
10	9-10	31
11	10-11	29
12	11-12	21
13	12-13	19
14	13-14	20
15	14-15	25
16	15-16	21
17	16-17	31
18	17-18	23
19	18-19	27
20	19-20	19
21	20-21	28
22	21-22	19
23	22-23	11
24	23-24	17

IV.METHODS OF IDENTIFYING BLACKSPOTS

Blackspots on the roads are those places, where accidents often appear to cluster or concrete the techniques used to identify the blackspots may broadly be categorized as a

a) Statistical methods
b) Engineering Methods
c) Bio-medical methods
d) Subjective assessment techniques.

Legislative measures those are possible

A variety of legislative measures are possible and different countries have adopted different measures. Some of them are listed below:

- To stipulate age limits for drivers.
- To introduce penalties of fine, imprisonment, disqualification, or endorsements on licenses for careless driving.
- To enable police to check the drivers for their drunkenness and to impose suitable penalties.
- To prescribe maximum hours of work for drivers of commercial vehicles and buses to prevent them from fatigue.
- To prescribe uniform road signs throughout the country and provide for penalties for the non-observance of the same.
- To lay down rules for pedestrians when crossing streets and to impose penalties for their non-observance.
- To prescribe rules for cyclist.
- To prescribe rules for motor-cycle and scooter riders.
- To prescribe rules for the maximum size and weights (axle loads) of vehicles.

- To prescribe minimum standards for the design of vehicle.

V. CONCLUSION

The model developed and calibrated using regression analysis is used to predict the accidents depending upon the factors considered**.** Once the proposed measures are implemented, the number of accidents can be decreased. From data stimulation, it found that Road Markings, Condition, Traffic Volume, Median Opening and Carriageway condition were main parameters for causing accidents. All undeveloped major and minor intersections must be developed with adequate lighting provisions. Pedestrian guardrail should be provided all along the footpath of service road and at bus stops. All the horizontal curves and bridges approaches are to be provided with guardrails so as to guide the drivers through the right path.

REFERENCES

1. Traffic Engineering And Transportation Planning by Dr.
2. L.R. Kadilayi.
3. Transportation Engineering & Planning by Papacos C.S. Papacostas & P.D.Prevedourous.
4. M. Abdel-Aty, (2003). "Analysis of driver injury severity levels atmultiple locations using ordered probit models."
5. J. Safety Res., 34(5),PP597–603
6. M. Bedard, G.H. Guyatt, M.J. Stones, J.P. Hirdes, (2002). "The Independent Contribution of Driver, Crash, and Vehicle Characteristics to Driver Fatalities." Accident analysis and Prevention, Vol. 34, PP717-727
7. S.K. Khanna, and C.E.G. Justo, (2001). "Highway Engineering", Nemchand and Bros, Roorkee.