Investigating a method for rating Sri Lankan roads through identifying the factors affecting road safety

D.M.Shirantha Ramanayake, Niranga Amarasingha
Faculty of Engineering, Sri Lanka Institute of Information Technology,
New Kandy Road, Malabe, 10115, Sri Lanka
sramanayake98@gmail.com, niranga.a@sliit.lk

ABSTRACT

Road safety assessment is imminent to reduce road accidents in Sri Lanka. The existing road safety assessments in other countries are inapplicable to Sri Lankan Road conditions as there are significant differences between the road conditions of Sri Lanka and other countries. This study aims to identify the governing roadside elements which are influential to the cause of accidents and can be used for road star rating. The study was conducted in three districts of Sri Lanka: Anuradhapura, Polonnaruwa and Kurunegala. The data for road characteristics were obtained through the Google Earth Web engine whereas accident data were collected from Sri Lanka Police. The variables such as road condition, road length, road width, delineation, shoulder condition, footpath, vehicle parking length, road divide status, roadside objects length, number of lanes, number of intersections, number of pedestrian crossings, and number of bus bays were analyzed to develop a Negative Binomial regression model considering the number of accidents as the dependent variable. The results reveal that six variables: number of lanes, road condition, number of intersections, road divide status, road section length and width of lane are significant towards the occurrence of accidents. Moreover, the results demonstrate the relationship between the road characteristics and the accident number which is crucial in road designing in order to reduce road accidents. The findings affirm the possibility in developing a road safety rating mechanism for Sri Lankan streets to standardize the road network with the international standards while enhancing the road conditions with reflecting to the required safety levels.

KEYWORDS: *Road safety, Road star rating, Sri Lankan roads, Negative binomial regression*

1 INTRODUCTION

Safety is a crucial component in road transportation due to the increase in road accidents as a result of the fast-growing infrastructure and the use of vehicles in Sri Lanka. Road safety assessment helps to analyze the potential causes for road accidents which leads to reduction of accidents through the effective adoption of the safety assessment methodology. Road safety assessment benefit in several perspectives such as produce designs that reduce the number and severity of crashes, reduce costs by identifying safety issues and correcting them before construction, promote awareness of safe road design practices, integrate multinomial road safety concerns, and identify key human factor considerations (Belcher et al. 2008).Therefore, this research study explores the potential of the implementation of a road safety assessment with considering the number of accidents and other governing road side elements to develop a generalized road safety assessment for Sri Lanka where heterogeneous traffic conditions are imminent.

2 LITERATURE REVIEW

2.1 Governing Factors in Road Safety Assessments

Belcher et al. (2008) assessed the road safety in Australia with respect to three core components which are road users at risk, scheme features and junction types. The road users at risk consisted of people with disabilities, pedestrians, pedal cyclists, and motorcyclists while scheme features consisted of road traffic signs, road markings, street lighting, road surfacing, surface water drainage, vehicle restraints and pedestrian guardrails. The junction types consisted of traffic signals, roundabouts, priority junctions and pedestrian or cycle openings. AUSTROADS (2009) provides a clear guideline in the risk assessment process and potential suggestions. Such that, the risk assessment is done in four main steps which are 1) estimate possible crash frequency (frequency is categorized as frequent, probable, occasional, and remote); 2) estimate the possible crash severity (severity is categorized as catastrophic, serious, minor and negligible); 3) determine the level of risk (risk level is categorized as low, medium, high and intolerable for frequency and severity), 4) determine a course of action (suggestion are provided for the four risk levels separately).

Demasi et al. 2018 analyzed the road safety while proposing a methodology to enhance risk management in urban roads and to promote better decision-making processes to design safer roads. The study developed risk factors for nine categories which are geometry, cross section, private access, pavement, lightning, road signs, intersection, urban furniture and stopping. The study further emphasized that these parameters contribute towards the Section Index Risk and the Branch Index Risk which are incorporated in evaluating the risk. Furthermore, the study acknowledged the correlations between the risk factors and the accident factors. Perera and Amarasingha (2021) did a similar analysis in Sri Lanka using risk metrics method and binomial negative regression for road risk assessment. Grasskopf (1998) developed a method to evaluate the existing road network safety performance. Such that, road safety index was calculated as a combination of accident index and the safety index. The accident index only had one element which is accidents. In this, the number of accidents and the accident severity were considered. The safety index was divided into four elements as operating conditions, land use, pedestrians and road elements. The operating elements was further divided into speed and weather conditions. Land use was divided into area characteristics and the road elements were divided into road hierarchy and characteristics, road lighting, access management, roadside hazard management, traffic control and geometry and pavement.

2.2 Road Safety and Accident Modelling

Road safety and accident prediction models have been developed since late 20th century where prediction models were mainly based on statistical models. Group et al. (1997) identified the road safety principles with predictive risk and accident consequence models. The study identified the factors influencing consequences such as traffic elements, accident type, speed, vehicle mass and age of injured. Ezra (2004) also modelled the road safety statistically with considering variables such as average annual daily traffic, percentage of trucks, speed limit, curvature, and gradient of the road.

Silva et al. (2020) developed a crash prediction model with using machine learning methods. The study identified three core models. Which are crash frequency, crash clarification by severity and crash frequency and severity. Bhasvar et al. (2021) modelled the safety of multi-lane rural highways in India as a statistical generalized linear model with having the accidents per km as the dependent variable and junction density, village settlement nearby, average daily traffic and average speed spots as independent variables. The authors identified accidents spots in the heterogeneous nature of accidents with assessing the severity of the accident as fatal and non-fatal.

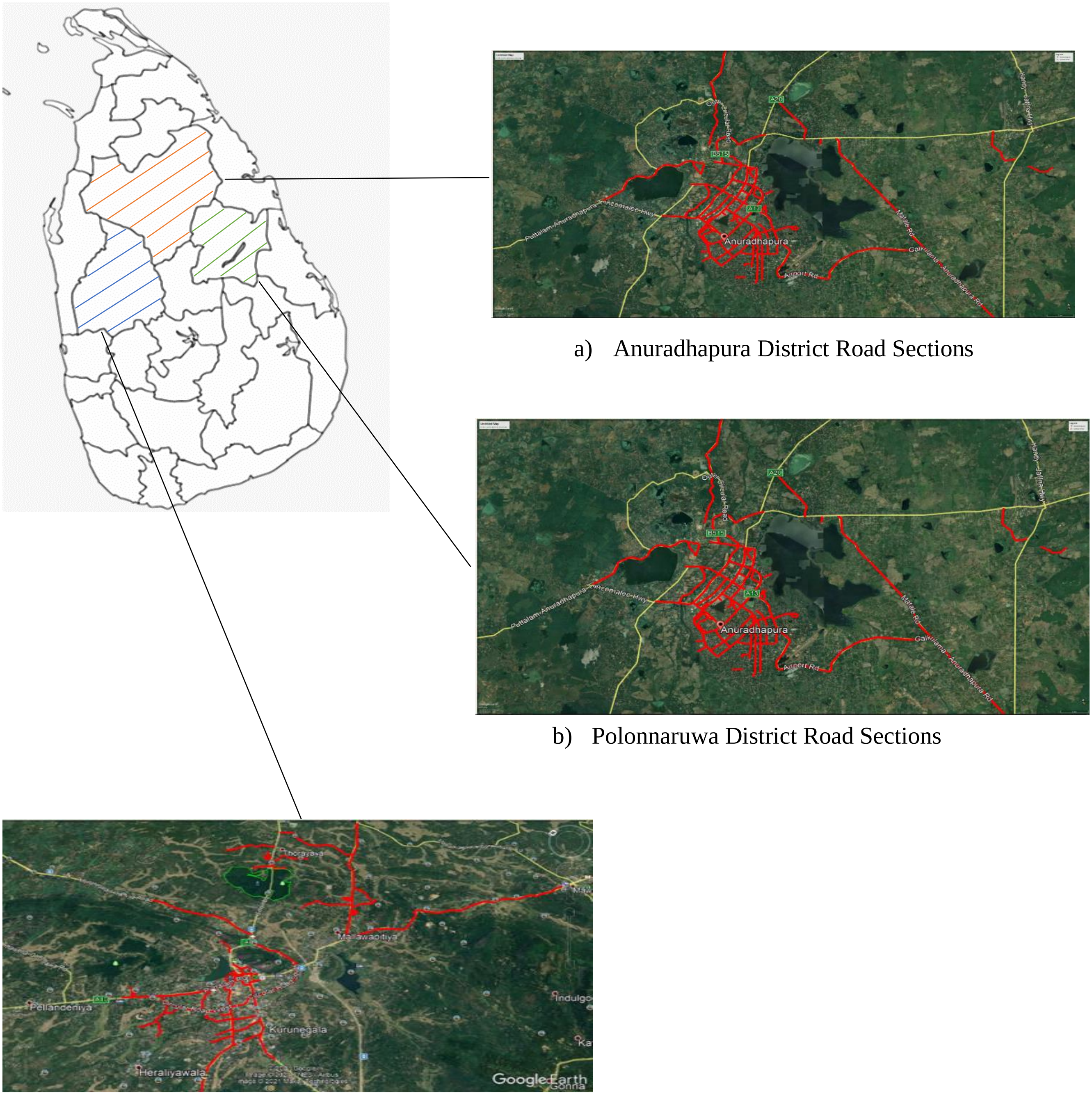
The methodologies of the regional Road Assessment Programs demonstrate the importance of a road star rating to a country to increase the safety performance of the roads and minimize the number of road accidents. Sri Lanka has been unable to develop such a road safety rating system thus making it more essential to implement such a system. However, there are various discrepancies among the regional methods in terms of the road demands, user desires, discipline, culture etc. Therefore, an expository analysis is required on developing such methodology for developing south Asian country like Sri Lanka to identify the governing parameters and to implement a complex road safety system.

3 METHODOLOGY

3.1 Study Area

The location selection was the initial step in the research and was considered as one of the vital steps as the selected locations provides an insight of the safety features in that respective road. The overall locations were selected in the Anuradhapura, Polonnaruwa and Kurunegala districts of Sri Lanka. The study areas are further illustrated in Figure 1. According to the Road Development Authority

(RDA) of Sri Lanka, the road classess in this study area are A Class, AB Class, B Class. The research study was carried out with 380 considered road sections which is seen in red in Figure 1.



a) Anuradhapura District Road Sections

b) Polonnaruwa District Road Sections

c) Kurunegala District Road Sections

Figure 1. The Selected Road Segments in the Study Area

3.2 Data Collection

The data collection was done separately for each of the road sections with adhering to each of the governing roadside elements. The road sections were selected based on the physical identification of local people from one intersection to other. Fourteen variables were collected in this process. Which are road length, road width, number of lanes, delineation, road condition, shoulder condition, number of intersections, footpath, vehicle parking length, number of pedestrian crossings, road divide status, roadside objects length, number of bus bays and number of accidents. The considered variables, variable type and the respective unit of measurement are summarized in Table 2. Road condition and shoulder condition were categorized based on the visual observation. When taking the road condition and shoulder condition both ways of the road were considered. The data for road characteristics were obtained by visual observation through the Google Earth Web engine whereas accident data were collected from Sri Lanka Police in the same year. The accidents included both vehicle – pedestrian accidents and vehicle – vehicle accidents.

Table 2. Definition of the Variable in the Study

Variable	Variable type	Unit Measurement/Category
Road Length	Continuous	Kilometers
Road Width	Continuous	Meters
Number of Lanes	Nominal	1 – One lane 2 – Two lanes 3 – Three lanes 4 – Four lanes
Delineation	Categorical	0 = No Delineation 1 = Adequate Delineation
Road Condition	Categorical	0 = Weak Condition 1 = Intermediate Condition 2 = Good Condition
Shoulder Condition	Categorical	0 = Weak Condition 1 = Intermediate Condition 2 = Medium Sealed Condition
Number of Intersections	Nominal	0 = 0 to 10 Intersections 1 = 11 to 20 Intersections 2 = 21 to 30 Intersections 3 = 31 to 40 Intersections
Footpath	Continuous	Kilometers
Vehicle Parking Length	Continuous	Kilometers
Number of Pedestrian Crossings	Nominal	0 = 0 to 5 Crossings 1 = 6 to 10 Crossings 2 = 11 to 15 Crossings
Road Divide Status	Categorical	0 = Undivided 1 = Divided
Roadside Objects Length	Continuous	Kilometers
Number of Bus Bays	Nominal	0 = 0 to 5 Bus Bays 1 = 6 to 10 Crossings 2 = 11 to 15 Crossings 3 = 16 to 20 Crossings
Accident Number	Continuous	Accidents in Last 5 Years

3.3 Data Analysis

The data analysis was started in considering three modeling approaches, which are generalized linear model, Poisson regression and Negative Binomial regression. The Poisson regression was initially applied followed by the assumption check to fit the model. The assumption check was performed with the goodness of fit and this was measured using the residual deviance as shown in Equation 1.

$$D(y,\hat{\mu}) = 2\left(log\left(p\left(y/\widehat{\theta_s}\right)\right) - log\left(p\left(y/\widehat{\theta_o}\right)\right)\right) \quad (1)$$

where;
y =the outcome,
$\hat{\mu}$ = the estimate of the model,

$\widehat{\theta_s}$ = parameters of the fitted saturated.
$\widehat{\theta_o}$ = parameters of the fitted proposed models,
p(y|θ) is the likelihood of data given the model.

A saturated model is perfect fit that has as many parameters as it has training points, that is, p=n. The proposed model can be the any other model. The residual deviance was checked to be small as possible where for the Poisson/Negative binomial regression, the residual deviance is calculated as shown in Equation 2.

Residual deviance = number of observations – number of parameters (2)

The residual deviance was also considered as the residual degrees of freedom of the model. However, the higher residual deviance which is greater than the residual degrees of freedom is not considered as a good fit model where Poisson regression model is invalid and thereby, negative binomial regression is applied to overcome the less model fit (Shingleton, 2012). However, the Poisson regression analysis was disregarded due to the assumption violation. Thus, the Negative Binomial regression was performed to obtain the relationships with adhering to the statistical viability. The negative binomial distribution is closely related to the Poisson distribution in that the negative binomial is a measure of instances of an event until reaching a concluding event (Shingleton, 2012). The formula for the negative binomial distribution is shown in Equation 3.

$$f(y) = \frac{T\,(r+y)}{y!\,T(r)}(1-p)^r p^y \quad y = 0,1,2.. \tag{3}$$

where r>0 and, when it is an integer, r can be interpreted as the number of failures and y is the number of successes required to get exactly r failures and p is the probability of a success. The function T(x) is the gamma function and is a continuous version of the choose function (Shingleton, 2012).

The overall analysis was conducted with using the IBM SPSS 23 software. The input variables for the dependent and the independent variables were as follows.

- Dependent variable – Accident number; this is a measure of safety which could be used easily for this study other safety measures such number of conflicts, Level of Service, etc may need excessive data collections.
- Independent variables – Road length, Road width, Number of lanes, Delineation, Road condition, Shoulder conditions, Number of intersections, Footpath, Vehicle parking Length, Number of pedestrian crossing, Road divide status, Roadside objects length, Number of bus bays.

4 RESULTS AND DISCUSSION

4.1 Negative Binomial Regression

The categorical variable information is shown in Table 3. The table represents the total data amount and the percentage of each categorical variable. Eight categorical variables are shown along with the respective category of each variable. Even though number of lanes was defined by four categories, the sample comprise of road sections of three categories which consisting with 1,2 and 4 lanes. Where the most frequent number of lanes were 2 with accounting to 365 sections which is over 96% of the sample. The four-lane category comprise of 3.2% sections whilst the minimum of 0.8% accounting for single lanes. The category of delineation comprises of two as road sections without delineation and with delineation. Road sections without delineation are 280 data which is 73.7% and adequate delineation comprise of 100 data which is 26.3% of the dataset. Road Condition is categorized into three as the weak condition, intermediate condition, and good condition. The category of weak condition has 33 data which is 8.7% of the dataset, intermediate comprise of 156 data which is 41.1% of the dataset and good condition comprise of 191 data which is 50.3% of the dataset of the category.

Table 3. Sample Size Distribution of Categorical Variables

Parameter		**N**	**Percent**
Number of lanes	1	3	0.8%
	2	365	96.1%
	4	12	3.2%
	Total	380	100.0%
Delineation	No	280	73.7%
	Yes	100	26.3%
	Total	380	100.0%
Road condition	Weak	33	8.7%
	Intermediate	156	41.1%
	Good	191	50.3%
	Total	380	100.0%
Shoulder condition	Week	207	54.5%
	Intermediate	62	16.3%
	Good	111	29.2%
	Total	380	100.0%
No of Intersections	0 to 10	313	82.4%
	11 to 20	53	13.9%
	21 to 30	12	3.2%
	31 to 40	2	0.5%
	Total	380	100.0%
No of Pedestrian crossing	0 to 5	346	91.1%
	6 to 10	31	8.2%
	11 to 15	3	0.8%
	Total	380	100.0%
Divided / Undivided	No	373	98.2%
	Yes	7	1.8%
	Total	380	100.0%
No of bus bays	0 to 5	339	89.2%
	6 to 10	35	9.2%
	11 to 15	5	1.3%
	16 to 20	1	0.3%
	Total	380	100.0%

The variable ‘shoulder condition’ in category of weak condition has 207 sections which is 54.5% of the dataset, intermediate comprise of 62 data which is 16.3% of the dataset and good condition comprise of 111 data which is 29.2% of the dataset of the category. The number of intersections was categorized to four according to the number of intersections. Thus, the majority of the road sections encompass 0 – 10 intersections with accounting to 313 road sections. The second category has 53 road sections, representing that 13.9% of the sample has 11 – 20 intersections. There are 12 road sections with having 21 - 30 number of intersections and only 2 road sections with having 31 – 40 number of intersections. The number of pedestrian crossing constitute to three categories. Thus, the majority of the road sections in the sample have 0 – 5 pedestrian crossings with accounting to 91.1% of the sample. While 8.2% of the road sections comprise of 6 – 10 pedestrians crossing and only 0.8% comprise of 11 – 15 pedestrian crossings. The Road Divide dataset comprised of 373 data for undivided which is 98.2% of the dataset while only 1.8% of the roads were divided which is 7 roads. The number of bus bays were also categorized into four categories. The majority of the 89.2% of the road sections comprise of 0 – 5 bus bays whereas 9.2% account to 6 – 10 bus bays. However, 1.3% of the road section have 11 – 15 bus bays and only 0.3% have 16 – 20 bus bays. In some categorical variables, some groups show few observations which can also be combined with some other categories and similar analysis can be done. However, in this study the model developed with combined categorical variable did not improve the model therefore the model presented in this paper used the variable definitions shown in Tables 3 and 4. The dependent variable and the continuous variable information are demonstrated in Table 4.

Table 4. Continuous Variable Information

Variable		**N**	**Minimum**	**Maximum**	**Mean**	**Std. Deviation**
Dependent Variable	Number of accidents	380	0	241	9.83	24.519
Covariate	Road section length (km)	380	.08	6.85	1.2373	1.08488
	Width of the lane (m)	380	.00	3.00	1.7057	.57394
	Shoulder width (inches)	380	.0	4.0	1.780	1.9542
	Foot path (km)	380	.000	4.900	.60826	.695778
	Vehicle parking length (km)	380	.000	3.750	.36603	.547762
	Road side objects length (km)	380	.000	3.650	.65009	.534341

The results show the minimum, maximum mean and the standard deviation of the dependent variable, accident number and six covariates which are road section length, width of lane, shoulder width, footpath, vehicle parking length, and roadside objects length. The number of accidents in a road section is a count variable which range from zero to 241 with a mean of 9.83. The road section length ranges from 0.08 km to 6.85 km with a mean of 1.2373 km. The width of lane ranges from 1 m to 3 m with a mean of 1.7057 m while the shoulder width has a minimum of 0 m and a maximum of 4 m with the mean of 1.78 m. The footpath length is between 0 km to 4.9 km with an average of 0.60826 km. The vehicle parking length ranges from 0 km to 3.75 km with a mean of 0.36603 km. Whereas the roadside objects length ranges between 0 km and 3.65 km and have a mean of 0.65009 km.

When developing the negative binomial model, the descriptive analysis results reveal that the mean of the dependent variable is 9.83 and the variance is 601.175. If the variance is higher than the mean the distribution should be negative binomial (Shingleton, 2012). The results of the current study show that the variance is more than the mean. Therefore, the results ascertain that the data set comprise of a negative binomial distribution. Thereby, the descriptive statistics of the dependent variable, and the negative binomial distribution is applicable due to the adherence and fulfillment to the assumptions checks with the applicability requirements of the negative binomial distribution. The likelihood ratio chi-square value is 771.892 along with a high significance of 0.000 which is less than 0.05. The goodness of fit results further affirm that the fitness of the developed model is appreciable. According to the results of the test of models, there are five variables which are significant, and they are number of lanes, road condition, number of intersections, divide status, number of bus bays, road section length and width of lane.

The parameter estimates for each variable along with the coefficient, standard error, significance and the exponential coefficient are shown in Table 6. The intercept has a significance of 0.46 which is higher than 0.05. Thus, the intercept of the model is not statistically significant towards the accident number. As per the results, the variable, 'one number of lanes' is significant as the p value is 0.012 and the variable 'two number of lanes' is also significant with the lower p value of 0.005. The variable 'one number of lanes' has a positive coefficient of 1.825 which reveals that it is positively affecting to the 'number of accidents' compared to that of 'four number of lanes'. However, the variable 'two number of lanes has a negative coefficient of – 1.093 which is negatively affecting towards the number of accidents compared to that of 'four number of lanes'. Furthermore, the coefficient of 'four number of lanes' is zero which signifies because it is the reference category. None of the levels of delineation variable were significant in the developed model. Even though the road condition encompasses of three levels, the results revealed that 'weak road condition' is insignificant. Whilst the intermediate road condition was statistically significant with a negative coefficient of -0.627, signifying that the intermediate level of road condition reduces the 'number of accidents' compared to 'good road condition'. According to the obtained results, 'shoulder condition' is insignificant towards the number of accidents. The variable, number of intersections is highly significant towards the dependent variable as that has lower p value obtained for all levels. Therefore, '0 – 10 number of intersections' affects the ' number of accidents' positively due to the positive coefficient of 3.324 while the 11 – 20 intersections a coefficient of 3.162 and 21 – 30 intersections with a coefficient of 3.161. However, 31 - 40 'number of intersections' has a zero coefficient which signifies that if the 'number of intersections' are between 31 – 40, there is no potential effect towards the accident number.

Table 6. Parameter Estimates of Accident Frequency Model for Road Segments

Parameter	B	Std. Error	Sig.	Exp(B)
Intercept	-1.975	2.6747	.460	.139
One number of lanes	1.825	.7265	.012	6.204
Two number of lanes	-1.093	.3868	.005	.335
Four number of lanes	0[a]	.	.	1
No delineation	.242	.1890	.200	1.274
Adequate delineation	0[a]	.	.	1
Weak road condition	-.627	.3574	.079	.534
Intermediate road condition	-.670	.2010	.001	.512
Good road condition	0[a]	.	.	1
Week shoulders	-3.443	2.0827	.098	.032
Intermediate shoulders	.027	.2389	.911	1.027
Medium sealed shoulders	0[a]	.	.	1
0 – 10 number of Intersections	3.324	.9792	.001	27.759
11 – 20 number of Intersections	3.162	.9542	.001	23.628
21 – 30 number of Intersections	3.161	.9824	.001	23.605
31 – 40 number of Intersections	0[a]	.	.	1
0 – 5 number of Pedestrian crossing	.205	.8385	.807	1.228
6 – 10 number of Pedestrian crossing	.299	.8161	.714	1.348
11–15 number of Pedestrian crossing	0[a]	.	.	1
Undivided	-1.051	.4583	.022	.349
Divided	0[a]	.	.	1
0 – 5 number of bus bays	1.428	1.3644	.295	4.168
6 – 10 number of bus bays	2.194	1.3397	.101	8.974
11 – 15 number of bus bays	.909	1.4333	.526	2.482
16 – 20 number of bus bays	0[a]	.	.	1
Road section length (km)	.729	.1860	.000	2.074
Width of the lane (m)	1.652	.1665	.000	5.216
Shoulder width (inches)	-.887	.5238	.090	.412
Road side objects length (km)	-.062	.3288	.851	.940

The overall coefficients reveal that, with the increase in the number of coefficients, the 'number of accidents in a section' is reduced. Thus, the fascinating results of reduced impact on the 'number of accidents' with the increasing 'number of intersections' has a major influence from the rural road conditions of the selected sample. The results further reveal that there is no significance from 'the number of pedestrian crossings' towards the 'number of accidents' as it has higher p value of above 0.05. However, 11 - 15 number of pedestrian crossings comprise of a coefficient of zero, signifying that if the 'number of pedestrian crossings' is between 10 – 15, there is no potential effect towards the number of accidents. 'Undivided road' is statistically significant variable decreasing the 'number of accidents' compared to that of 'divided road'. The 'number of bus bays' has no potential impact towards the 'accident number' as it is statistically insignificance in the obtained model. The 'road section length' variable is significant with the significance less than 0.05, whilst the coefficient is 0.729 which shows that when the 'road section length' is increased by 0.729 units, then the 'accident number' is increased by $e^{0.729}$ units. Moreover, the width of the lane variable is significant as the significance obtained as 0.000. The coefficient for the width of the road section is 1.652 and it explicates that when the width of the road section is increased by 1.652 units, the 'accident number' that road section is increased by $e^{1.652}$ units. However, the 'shoulder width' and the 'roadside objects length' is insignificant as p value is more than 0.05.

The road accidents are influenced by multiple factors such as road, traffic, driver, environment and several others. The comprehensive study conducted by Ashraf et al. (2019) proposed that a variety of road factors such as allowable speed, road structure, junctions and sections lengths are extremely influential towards the accident number. As Eboli et al. (2020) alluded, the influential road characteristics such as location, street type, paving condition, intersection type, surface conditions have significance on the road accidents and can be statistically developed through a binary logistic analysis. Therefore, the present study was mainly oriented around the potential road characteristic that are influential towards the increase of the accident numbers in the form of a Poisson and negative binomial regression analysis. The considered variables of the current study assessed fourteen variables in total to identify the influence on the number of accidents in road sections. As per the results, only the variables such as number of lanes, road condition, number of intersections, divide status, road section length and width of lane shows statistical significance towards the number of accidents.

The road section length is a substantial variable which also critically analyzed by Ashraf et al. (2019). Thus, the authors stated that number of accidents is dependent on the road length. The findings of the present study further affirm these results as the considered road section length is positively related to the accident number. Thus, the shorter road lengths results in less accidents. The road surface condition is a critical factor governing towards the accident number. Haque et al. (2009) highlighted the importance of a good condition of the road to reduce number of accidents while Shaheed and Gkritza (2014) found that dry and good road surfaces reduce the accident severity. The present study results also revealed that road condition is significant towards the number of accidents where intermediate road conditions have less impact and good road conditions further reduce the number of accidents. Therefore, the findings of the present study further entangle with Shaheed et al. (2013) as they found that the roadway surface has a greater influence on the accident number.

The geometrical parameters such as delineation and shoulder were obtained as statistically insignificant in the present study. These parameters have mixed results in the previous findings where Manan et al. (2018) affirmed that the road shoulder type is statistically insignificant towards the crash frequency, yet certain road geometrical parameters show significance. Moreover, the results for divide status in the current study explicate that there is a negative correlation between undivided roads and the accident number while no potential impact from divided roads. According to the results of alluded Manan et al. (2018) the road dividing with the markings has a higher statistical significance where unmarked and undivided roads have a substantially high impact on the crash amount.

The selected road sections of the present study are of rural areas where the number of lanes is ranging from 1 to 4. Therefore, the road types of the present study have a broader range from single lane to two way two lanes. However, the number of lanes was significant where the single and dual lanes were obtained in the final model. According to the results, the positive coefficient of single lane proves that single lane road sections impact positively towards the number of accidents while the negative coefficient of the dual lanes impact negatively towards the number of accidents. On the contrary, Ashraf et al. (2019) proved in a study conducted in Korea that road types have the minimal effect towards the accident rates. Moreover, Eboli et al. (2020) also found that roadways with multiple lanes have high accident severity and crash probability.

The significance of the number of intersections was statistically proven by Eboli et al. (2020) and Haque (2009). The results of the present study further back these findings with having all the considered categories for number of intersections significant towards the number of accidents. The obtained results of the present study show that the roadside objects length is insignificant. Yet Shankar (1996) found that fixed object interaction has an extensive influence towards single vehicle motorcycle crashes while Maistros (2014) stated that the rigid roadside objects contributes heavily towards single vehicle motorcycles and car crashes and further increases the severity.

The road width study conducted by Chen et al. (2020) analyzed the lane width and the footpath width based on the crash analysis. Thus, the results proposed that increased lane width reduces the risk of crash and the severity. However, the increment in the footpath width has disparate effects on the pedestrian safety. Contrastingly, the results of the present study showed no statistical significance for the shoulder width, which signifies that there is no potential impact towards the number of accidents from the shoulder width. However, the lane width is statistically significant with a positive coefficient. Therefore, it reveals that increase in the lane width further enhances the number of accidents. Furthermore, the rural road condition is a core factor on this result where with higher lane widths the vehicle speed increases further increasing the crash risk (Davoudi-Kiakalayeh & Yousefzade-Chabok, 2014).

The increase in the number of pedestrian crossing enhances the risk of pedestrian crash due to the higher exposure of the pedestrians (Gobalarajah, 2016). However, the current study results exposed that the number of pedestrian crossings shows no statistical significance towards the number of accidents. Moreover, Porcu et al. (2020) found that the number of bus bays increase the accident risks due to the increased frequency and the severity of the bus accidents. Therefore, the increase in the bus bays has a significant influence in enhancing the number of accidents. On the contrary, the results of the present study revealed that number of bus bays shows no potential significance towards the number of accidents in the developed negative binomial regression.

Therefore, it is illuminated that, variables such as number of lanes, road condition, number of intersections, divide status, road section length and width of lane are only statistically significant in the negative binomial regression model where other variables are unacceptable due to the insignificance. Furthermore, the significant variables depend on the accident number of the road section where the road star rating can be developed. Therefore, lower accident numbers denoting a higher road star rating, whilst higher accident numbers denoting lower road star rating. These results can be used by the road agencies when developing and testing the road star rating system.

5 CONCLUSIONS

This study is important in identifying the variables for a road safety rating in Sri Lanka. thus, the study explored the relationship between the road accidents other influential road geometric aspects with specified to rural road sections of Kurunegala, Polonnaruwa and Anuradhapura districts of Sri Lanka. The road sections in these districts were selected as a pilot study and there may be minor differences on driver behavior, traffic composition, and road geometry in the other areas in Sri Lanka. This study shows that the number of lanes, road conditions, number of intersections, road divide status, road section length and width of lane are parameters which affect the number of accidents in a particular road section. These findings can be used to evaluate road safety in Sri Lanka and developing a star rating system for Sri Lankan roads. Therefore, the research objectives were achieved with the identification of the governing parameters for the implementation of the rad star rating system. Moreover, the developed negative binomial equation represents the relationship between the governing road parameters and the number of accidents in the road sections.

However, the data collections were limited to rural road sections where more applications of the developed methodology is required in urban road sections. Thus, a generic road safety rating can be developed through the comprehensive analysis of different road conditions with different road and traffic characteristics. It is recommended to conduct a similar study in other areas in Sri Lanka including the urban road sections and thereby to develop a general road star rating for the Sri Lankan roads.

ACKNOWLEDGEMENT

The authors would like to acknowledge the data provided by Sri Lanka Police for conducting this study. This research was supported by the Accelerating Higher Education Expansion and Development (AHEAD) Operation of the Ministry of Higher Education, Sri Lanka, funded by the World Bank.

REFERENCECS

Ashraf, I., Hur, S., & Shafiq, M. (2019). Catastrophic factors involved in road accidents: Underlying causes and descriptive analysis. *PLoS ONE, 14*(10), 29.

AUSTROADS. (2009). *Guide to Road Safety: Part 6: Road Safety Audits.*

Belcher, S., Proctor, S., & Cook, P. (2008). *Practical Road Safety Auditing.*

Bhasvar, R., Amim, A., & Zala, L. (2021). Development of model for road crashes and identification of accident spots. *International Journal of Intelligent Transportation Systems Research*, 99-111.

Chen, T., Sze, N., Chen, S., & Labi, S. (2020). Urban road space allocation incorporating the safety and construction cost impacts of lane and footpath widths. *Journal of Safety Research*, *75*(1), 31.

Davoudi-Kiakalayeh, A., & Yousefzade-Chabok, S. (2014). Road traffi c crashes in rural setting: an experience of a middle-income country. *Chinese Journal of Traumatology*, *17*(6), 327 -330.

Demasi, F., Loprencipe, G., & Moretti, L. (2018). Road safety analysis of urban roads: case study of italian municipality. *safety*, 58.

Eboli, L., Forciniti, C., & MazzullaG. (2020). Factors influencing accident severity: an analysis by road accident type. *Transportation Research Procedia*, *47*, 449-456.

Gobalarajah, R. (2016). *STUDY OF ACCIDENTS AT MID-BLOCK PEDESTRIAN CROSSINGS IN SRI LANKA.* Moratuwa, Sri Lanka: University of Moratuwa.

Grasskopf, S. (1998). *South african road safety manual; road safety audits.* Road traffic management cooperation.

Group, O., Mikkonen, V., & Pelola, H. (1997). *Road safety principles and models review of descriptive, risk and accident consequence models.*

Haque, M., Chin, H., & Huang, H. (2009). Modeling fault among motorcyclists involved in crashes. *Accid. Anal. Prev.*, *41*(2), 327-335.

Maistros, A., Schneider, W., & Savolainen, P. (2014). A comparison of contributing factors between alcohol related single vehicle motorcycle and car crashes. *J. Saf. Res.*, *49*, 129-135.

Manan, M., Várhelyi, A., & Çelik, A. (2018). Road characteristics and environment factors associated with motorcycle fatal crashes in Malaysia. *IATSS Research*, *42*(4), 207-220.

Perera, B.D.S.P., Amarasingha, N., (2021), Risk Assessment of Road Intersections in Colombo Suburban, Proceedings of Eastern Asia Society for Transportation Studies, 12-15 Sep., Hiroshima, Japan.

Porcu, F., Olivo, A., & Maternini, G. (2020). Evaluating bus accident risks in public transport. *Transportation Research Procedia*, *45*, 443-450.

Shaheed, M., & Gkritza, K. (2014). A latent class analysis of single-vehicle motorcycle crash severity outcomes. *Anal. Methods Accident Res.*, *2*(0), 30-38.

Shaheed, M., Gkritza, K., Zhan, W., & Hans, Z. (2013). A mixed logit analysis of two-vehicle crash severities involving a motorcycle. *Accid. Anal. Prev.*, *61*, 119-128.

Shankar, V., & Mannering, F. (1996). An exploratory multinomial logit analysis of single-vehicle motorcycle accident severity. *J. Saf. Res.*, *27*(3), 183-194.

Shingleton, J. (2012). *CRIME TREND PREDICTION USING REGRESSION MODELS FOR SALINAS, CALIFORNIA* . MONTEREY, CALIFORNIA: NAVAL POSTGRADUATE SCHOOL .

Silva, P., Andrade, M., & Ferreira, S. (2020). Machine learning applied to road saefty modelling: a systematic literature review. *Journal fo Traffic and Transportation Engineering*, 775-790.