International Journal of Injury Control and Safety Promotion

ISSN: 1745-7300 (Print) 1745-7319 (Online) Journal homepage: https://www.tandfonline.com/loi/nics20

Crash severity modelling using ordinal logistic regression approach

Isaac Ofori Asare & Alice Constance Mensah

To cite this article: Isaac Ofori Asare & Alice Constance Mensah (2020) Crash severity modelling using ordinal logistic regression approach, International Journal of Injury Control and Safety Promotion, 27:4, 412-419, DOI: 10.1080/17457300.2020.1790615

To link to this article: https://doi.org/10.1080/17457300.2020.1790615

Published online: 05 Jul 2020.

INTERNATIONAL JOURNAL OF INJURY CONTROL AND SAFETY PROMOTION
2020, VOL. 27, NO. 4, 412–419
https://doi.org/10.1080/17457300.2020.1790615

Crash severity modelling using ordinal logistic regression approach

Isaac Ofori Asare[a] and Alice Constance Mensah[b]

[a]MSc Applied Statistics, Vita Verde Consult, Accra, Ghana; [b]Mathematics and Statistics, Accra Technical University, Accra, Ghana

ABSTRACT
Road traffic accident is one of the major problems facing the world. The carnage on Ghana's roads has raised road accidents to the status of a 'public health' threat. The objective of the study is to identify factors that contribute to accident severity using an ordinal regression model to fit a suitable model using the dataset extracted from the database of Motor Traffic and Transport Department, from 1989 to 2019. The results of the ordinal logistic regression analyses show that the nature of cars, National roads, over speeding, and location (urban or rural) are significant indicators of crash severity. Strategies to reduce crash injuries should physical enforcement through greater Police presence on our roads as well as technology. There is also the need to train drivers to be more vigilant in their travels especially on the national roads and in the urban areas. The Recommendation is, a well thought out and contextualised written laws and sanctioned schemes to monitor and enforce strict compliance with road traffic rules should be put in place.

ARTICLE HISTORY
Received 16 April 2020
Revised 19 June 2020
Accepted 29 June 2020

KEYWORDS
Ordinal logistic regression; proportional odds; road traffic accidents; injury; chi-square test

Introduction

Road traffic accident is one of the major problems facing the world. It is any vehicle accident occurring on a public highway and is a major public health concern. World Health Organization (2018) reports that: "between 20 and 50million more people suffer non-fatal injuries, with many incurring a disability as a result of their injury. Road crashes are ranked as the ninth most serious cause of death in the world and it is anticipated to surge to the third place by the year 2020 if initiative to improve road safety is disregarded. Ensuing current trends, WHO reports that, about 2million people could be expected to be killed in motor vehicle crashes each year by 2030. Traffic fatalities have increased in developing countries from 1990 to 2014, 44 percent to 243 percent. Again, 93 percent of the world's fatalities on the road occur in low – and middle-income countries.

The report further states that, more than 90 percent of road traffic deaths occur in low- and middle-income countries, which have 59 percent of the world's registered vehicle population. Road traffic injury deaths rates are highest in the African region, 26.6 percent per 100,000 population". From Global Road Safety Partnership (2014) report, it is estimated that 1.24million people are killed each year taking an immeasurable toll on families and communities. Up to 50million people suffer injuries, life altering injuries which in many low- and middle-income countries directly contribute to the poverty cycle, (GRSP, 2014). The report further states that road fatalities per 100,000 of the population is 24.1 in Africa, the highest in the world with the least being 10.3 in Europe.

The major system of transportation in Ghana is road transportation on which most people rely for their daily commuting and the conveyance of raw materials and foodstuffs. This constitutes about 96 percent of the country's freight. Nevertheless, the rate at which accidents occur on our roads is very alarming and it is seen as the deadliest disease (Osei – Kyei & Chan, 2016). Studies have showed that pedestrians fatalities accounted for 42 percent of all road traffic fatalities with children constituting almost a third (Ackaah & Adonteng, 2011) and ranged between 21 percent and 41 percent among pedestrians in northern Ghana based to the collision type (Damsere-Derry et al., 2017). Occupants of goods vehicles accounted for 12 percent of all road traffic fatalities although goods vehicle constitute about 9% of the total motor vehicle population in Ghana (Ackaah & Adonteng, 2011). Females were more likely to die as pedestrians and males as drivers or riders while children under 10 years and those above 60 years being twice more likely to die in traffic collisions as compared to casualties aged between 30 and 59 years, (Damsere-Derry et al., 2017). Many institutions including the Motor Traffic and Transport Department (MTTD), the national Road Safety Commission (NRSC) and the Drivers and Vehicle Licensing Authority (DVLA) have been charged with the responsibility of managing road traffic in Ghana. There has been a steady rise in the occurrence of road traffic accidents for the past ten years, despite every effort made by the relevant institutions to manage and curb the incidence of road traffic accidents. Statistics has shown that, the number of fatal accidents in Ghana keeps increasing on yearly basis with its negative impact on families, government and insurers (Albert, 2014). National Road Traffic Crash Statistics of

CONTACT Isaac Ofori Asare asare.re3@gmail.com

2016 for Ghana, indicated an increase of 15.6 and 6.77 percent in fatalities and serious injuries respectively.

Traffic accident mortality rate in Ghana was 24.7 per 100,000 population in 2016 and ranked 18th in Africa, (World Health Organization, 2018). A release by the MTTD (2018), indicates that, a total of 2,076 people died in road traffic accidents in 2017 as compared to 2,084 in 2016. In the same period 20,444 vehicles- 8,080 commercial vehicles, 8,877 private vehicles and 3,487 motorcycles were involved in accidents, with pedestrian knockdowns and killed at 3,300 and 879 respectively. According to the data, 12,166 travellers were injured in 2017 compared to 12,500 in 2016. The latest statistics released by the MTTD for the period January to June 2019, revealed that four people die daily on Ghanaian roads and estimates that Ghana loses over 230 million USD yearly due to road accidents with more than 1,600 deaths. Also, there is a 3.3 percent rise in the number of casualties as compared to the same period in 2018.

According to the World Health Organization (2015), road traffic injuries have considerable economic losses on individuals, families and the nation at large. These said losses arise from the cost of treatment as well as cost of productivity for those killed or disabled by injuries and their family members who need to take time off work or school to care for the injured. Atubi (2012) testified that, road traffic injuries are a leading cause of death in children which accounted for 55 percent of all accidental death in healthy children who might have been expected to have had productive lives. This causes immeasurable distress and guilt to the parents and other parties involved.

Available literature on accident severity have attributed it to a number of factors: road user characteristics, vehicle types involved, the speed of travel, manner of vehicle collision, road environment, crash location, time and crash season. Haadi (2014) investigated the factors that contributed to road crash severity in Ghana's Northern Region using binary logistic regression. The study found that overloading and obstruction were the two most significant factors contributing to road crash severity in Ghana. A similar study was undertaken to explore the factors affecting motorcycle crash severity in Ghana by Wahab and Jiang (2019a). The result indicated the following factors increase the probability of fatal injuries: at a junction, weekend, signage, poor road shoulder, village settlement among others.

In a research by Wang et al. (2017), blacktop road surface was associated with a significant increase in fatality risk compared with gravel/stone and concrete surface. They further reported that road traffic crashes occurring at expressway and secondary road had a significant increment of fatality risk, which is supported by (Haleem & Gan, 2011). Crash location, time and crash season are significant variables influencing injury severity. According to World Health Organization (2018) reported that, approximately 60 percent of road traffic fatalities occur on the non-urban sections of the road networks.

This is corroborated by (Afukaar et al., 2003; Damsere-Derry et al., 2008) . In a study by Odero et al. (1997), distribution cases showed Saturday as the overall as well as the fatal cases peak. Several researches have used various models and techniques in predicting road traffic accidents such as; Poisson lognormal model (Miaou & Lord, 2003; Lord & Miranda-Moreno,2008); Zero-Inflated Poisson and negative binomial (Washington et al., 2003, 2010); Negative multinomial models(Caliendo et al., 2007); Neural, Bayesian Neutral Network, Support vector machine models(Liang, 2005; Xie et al., 2007). A strong association has been proven between speed and crash severity (Elvik & Vaa, 2004; Taylor et al., 2002).

Crash severity models generally explore the relationship between crash severity injury and the contributing factors: driver or human error, vehicle characteristics, roadway geometry and road-environment conditions. In Ghana, Wahab and Jiang (2019a), used machine learning methods to predict the severity of motorcycle crashes and revealed the most significant factors associated with motorcycle crash injury to be location type, settlement type time of the crash, collision type and collision partner. A study by Rezapour et al. (2019) provided insights into the factors contributing to downgrade crashes in mountainous areas, used selected ordered logistic models on crash injury severity of downgrade crashes. Yuan et al. (2020) investigated injury severity in of expressways in Beijing using Bayesian ordered logistic regression model. The outcome of the study was that crash location, time and crash season are significant variables influencing injury severity. In another study, Park et al. (2012) assessed the influencing factors contributing to the degree of injury severity sustained in traffic crashes on Korea expressways. Examining ordered probit, ordered logit and multinomial logit, sixteen variables were identified as major contributing factors to severity of injuries.

Using the generalized ordered logistic regression model, Michalaki et al. (2015) explored the factors affecting motorway accident severity in England. The findings suggested that the factors positively affecting the severity include the number of vehicles involved, peak-hour traffic time and low visibility. On their, part Yoon et al. (2017) fitted a hierarchical ordered model and revealed that vehicle speed, vehicle age, road alignment, surface status, road class and traffic light installation were significant at the lower level. The relationship between vehicle age and risk of a car crash with injury has also been investigated (Ablin Consult, 2007; Blow et al., 2003; Peden, 2004).

A cursory look of the above statistics indicates the severity of the safety issue on our roads (Wahab & Jiang, 2019b; Damsere-Derry et al., 2017). As posited by Olmu and Erbas (2012), a greater emphasis should be placed on the issue of traffic accidents because the problems caused by them continue to increase in seriousness. Thus, there is an urgent need to fashion out a predictive model on road accident or crash severity to help manage this menace efficiently. Prediction models in traffic accidents are very useful in highway safety, since it has the potential for determining both the frequency of accident occurrence and the degree of crash severity. Although many studies have taken place, they were carried out prevalently in countries where traffic characteristics, driver behaviour and geometry of roads differ

Table 1. Variables included in the model.

Variable	χ^2	df	Sig
Average age of vehicles	207.55	120	0.000
Nature of vehicle involve	1759.81	6	0.000
Day of the week	255.11	12	0.000
Type of road	510.66	6	0.000
Average speed limit	1820.64	10	0.000
Urban or rural area	1171.33	2	0.000
Number of vehicles involve	1896.69	28	0.000
Average age of drivers involved in accidents	252.44	192	0.002
Average years of driving by drivers involved in accidents	68.51	20	0.000

Significance value < 0.05 indicates significance.

from those of Ghana. This paper aims to fit an ordered logistic regression model to determine road accident severity level using the dataset from MTTD of the Ghana Police Service database. To develop measures aimed at improving road safety, it is necessary to understand how, why and where crashes occur, who are involved and what the main contributory factors in crashes are.

Methods

Data description

The study utilized the MTTD of the Ghana Police Service database on road accidents from the whole country. The reported road traffic accidents data from January 1989 to December 2019 were collected from the MTTD. The MTTD among other duties, record all accidents or crash information, which includes the severity of injuries and property damaged, and publishes the statistics on quarterly basis in print and electronic media. The data collection form was prepared with excel using variables as; date, age, gender, location, injury severity outcome, pedestrian victim, type of road, type of vehicle, etc. The MTTD is organized into 11 units in the police regions of Ghana, divisional and district MTTU as well as accident squads. Personnel of the MTTD are stationed in all the six hundred and fifty-one (651) police stations in the country.

Data collected were then exported into R, for the statistical analysis. The dependent variable in this study is accident severity levels of crashes categorized into three groups; 1-fatal, 2-Serious and 3- Minor. It represents an ordered outcome and therefore, an ordered-response model is appropriate to analyse the data (Pour-Rouholamin & Zhou, 2016). The independent variables considered as contributory factors of road crashes involving human injuries include; Average age of vehicles; Nature of Vehicle type involve (1-Commercial;2-Private;3-Motor bike); Day of the Week (1-Monday to 7-Sunday); Type of road (1-National, 2-Inter Regional, 3-Regional and 4-Rural); Average years of driving by drivers involve; Average Speed limit; Location (Urban- 1 and Rural − 2); Number of Vehicles involve; Average age of drivers.

Statistical analysis

The Chi-square test (X^2) of independence was used to determine the inclusion of variables in the model. The ordinal logistic regression model was to assess road crash variables which are proximately associated with road crash severity level. The variables that have statistical and significant association with road crash severity level at 5% significance level (Sig ≤ 0.05) were used for further analysis using the Ordinal Logistic Regression model.

Inclusion of road accidents crashes variables

Table 1 depicts the result when the chi-square test was applied on the dataset. The variables that have statistical and significant association with road crash severity level at 5% significance level (Sig ≤ 0.05) were used for further analysis using the Ordinal Logistic Regression model. The results show that nine (9) variables associate well with road crash severity level at 5% significance level. These variables have significance value less than ($p < 0.05$).

Point-biserial correlation

Table 2 shows the point -biserial correlation and collinearity diagnostics analysis for the variables. The multicollinearity among the contributory factors was checked via the calculation of point -biserial correlations among the variables. The collinearity analysis was performed to check for the independence of the study variables using (correlation coefficient $|r| \geq 0.7$) which occurs when correlations exist among the variables (Hair et al., 2010). A Variance Inflation Factor (VIF) was performed to check if multi-collinearity exists among the independent variables. The general rule of thumbs for VIF test is that if the VIF value is greater than 5, then there is multi-collinearity (Hair et al., 2010; Ringle et al., 2015). Since none of the VIF values is greater than 5.00 as seen in Table 1, it can be concluded that there is no multi-collinearity in the dataset (Byrne, 2010).

Model description

The statistical technique, logistic regression is a special case of the generalized linear model which generalizes the ordinary linear regression by allowing the linear model to be related with a response variable that follows the exponential family via an appropriate link function (Jeong et al., 2018). When the response variable is binomially distributed with parameter *pi*, the logit function is used for the link function. The logistic regression establishes the relationship between the response variable $y = (y_1, y_n)$ given $p = (p_1, \dots p_n)$ and a set of k predictors, $X = (x_1, \dots x_k)$: Ordered logistic regression usually known as proportional odds logistic regression was used to analyse the data gathered statistically. The technique uses log-odds of cumulative probabilities. According to (Fox & Hong, 2009) the proportional-odds model, perhaps the regression model most commonly used for an ordinal response, is often derived by assuming a linear regression for a latent response variable γ.

$$\gamma_i = \alpha + \acute{x}_i\beta + \varepsilon_i = \alpha + \eta_i + \varepsilon_i \quad (1)$$

where $\acute{x}_i$ i is the i^{th} row of the model matrix, β is a vector

Table 2. Point-biserial correlation and collinearity diagnostic.

Variable	*1*	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*	*Tol*	*VIF*
Average age of vehicle (1)	1										0.997	1.003
Nature of vehicle involve (2)	−0.002	1									0.330	3.033
Day of the week (3)	−0.002	0.002	1								1.000	1.000
Road type (4)	−0.001	−0.039**	0.003	1							0.996	1.004
Average speed limit (5)	0.005	−0.013**	−0.002	−0.044**	1						0.959	1.043
Location (6)	0.019**	0.048**	−0.019**	−0.009**	0.195**	1					0.957	1.045
Number of vehicles involve (7)	−0.002	0.818**	0.001	−0.035**	0.006*	0.062**	1				0.329	3.035
Average age of drives involved in the accident (8)	0.011**	0.006*	0.001	0.002	−0.006*	0.015**	0.007**	1			0.431	2.318
Average years of driving by the drivers (9)	−0.026**	0.005	0.000	0.006*	−0.005	0.023**	0.008**	0.753**	1		0.431	2.320
Accident severity level (10)	0.000	0.092**	0.001	−0.036**	−0.032**	0.001	0.074**	0.005	0.003	1		

of regression coefficients, and α is an intercept parameter (which will prove redundant). Although the latent response cannot be observed directly, a binned version of it, y, with m levels is available as follows:

$$y_i = \begin{cases} 1 \; for \; \gamma_{i<}\alpha_i \\ 2 \; for \; \alpha_i < \gamma_i \leq \alpha_2 \\ . \\ . \\ . \\ m-1 \; for \; \alpha_{m-2} < \gamma_i \leq a_{m-1} \\ m \; for \; a_{m-1} < \gamma_i \end{cases}$$

where the thresholds given $\alpha_1 < \alpha_2$ - - - $< \alpha_{m-1}$, are parameters to be estimated from the dataalong with the regression coefficients. The cumulative distribution function of the observed response is given as follows;

$$\begin{gathered} P_r(y_i \leq j) = \; P_r(\gamma_i \leq \alpha_j) \\ P_r(\alpha + \eta_i + \varepsilon_i \leq \alpha_j) \\ P_r(\varepsilon_i \leq \alpha_j - \alpha - \eta_j \end{gathered}$$

for $j = 1, 2, \ldots m-1$. Assuming that the errors ε_i are normally distributed produces an ordered probit model; assuming that the errors are logistically distributed, with distribution function

$$\Lambda(\varepsilon_i) = \frac{1}{1 + e^{-\varepsilon_i}}$$

produces an ordered logit model:

$$\begin{aligned} logit = [P_r(y_i > j)] &= log_e \frac{P_r(y_i > j)}{P_r(y_i \leq j)} \\ &= -\alpha_j + x\acute{}\beta \quad 1; 2; \ldots m-1 \end{aligned} \tag{2}$$

The intercept α in equation 1 is set to 0 to identify the model, in effect fixing the origin of the latent response. This model is called the proportional-odds model because different cumulative log-odds, logit [Pr ($y_i > $ j)] and logit [Pr ($y_i > \acute{j}$)], differ by the constant α $\acute{j}$-αj, and therefore the odds themselves are proportional regardless of the value of $\acute{x}_i$ i.

Each level of y save the last has its own intercept, which is the negative of the corresponding threshold, but there is a common coefficient vector β. One of the key assumptions in an ordered logistic regression model is the assumption of parallel curves. In this assumption, the regression parameter obtained in the model is the same in all categories of the dependent variable (Akin & Şentürk, 2012). The interpretation of the parameters in the ordered logistic regression is difficult to interpret or explain. Methods such as the use of standardized coefficients, calculation of estimated probabilities, calculation of factor change in estimated probabilities and percentage change in estimated probabilities are used for interpreting parameters. Also, the odds ratio is used to interpret the parameters in the model. In the event that all other variables are held constant,exp (β_x) is an odd ratio for a dummy variable. To standardize odds ratios, sk showing standard deviation,exp $(\beta_k * sk)$ is calculated provided that all other variables are held constant. For continuous variables; the percentage is found by $\left[expexp \; (\beta - 1) \; *100\right]$.

We employ the polr () function in the MASS package associated with (Venables & Ripley, 2013), which is one of the recommended packages in R-software. The vglm () function in the VGAM package was used to fit a variety of threshold-based models to the ordinal data (Yee, 2015). The poTest function in R implements tests proposed by Brant (1990) for proportional odds for logistic models fit by the polr function in the MASS package was used to fit the model. It is used to test on separate binary logit models fit to the various cumulative logits, taking into account not just the variances of the estimated coefficients but also their covariances (Brant, 1990). The test function provides both an overall test of the proportional-odds assumption and separate tests for each regressor, all of which produce small p-values for our model, casting doubt on the assumption of proportional odds.

Table 3 represents the result of Brant Test of parallel regression assumption for the ordinal logistic regression developed for the dataset. From the result, it can be concluded that the parallel assumption holds since the probability (*p-values*) for all variables are greater than alpha = 0.05. The output also contains an Omnibus variable, which stands for the whole model, and it is still greater than 0.05. Therefore, the proportional odds assumption is not violated and the model is a valid model for this dataset (Long & Freese, 2014).

Result

Descriptive statistics of study variables

Data from the various regions were put together for a period of 30 years from 1989 to 2019 by the MTTD of the

Ghana Police Service. It was a retrospective analysis of 127,739 road accidents cases in Ghana. The dataset was split into two, 70% for training (89345) and 30% for testing (38394) randomly for the study. In Ghana, accident severity is recorded as fatal, serious and minor. This study employed the discrete choice modelling, in which a decision maker chooses an alternative from a set of exhaustive and mutually exclusive alternatives (Train, 2009). It is chosen in order to explore the most important factors affecting the severity of accidents on roads in Ghana.

Table 4 shows the descriptive statistics of road crashes variables. From the results, the average age of the vehicles involved in road crashes is estimated to be approximately (M = 4.2; SD = 1.4). Day of the week at which accidents occurred most was Saturday (16.3%), followed by Friday (15.3%). This is the period where people travel for gatherings such as funerals, weddings, parties etc in Ghana. The result shows that most of the road accident crashes occur on the rural roads (72.2%) which is along the national highways, accidents on regional roads accounted for 15.6%; national roads (7.6%) and least accidents on inter-regional roads accounting for 4.6%.

Also, most of the roads accidents statistics were recorded in the urban areas of the country accounting for 67.3% as compared with accidents in the rural areas forming about 32.7%. The average speed recorded of the vehicles involved in the road accident crashes was estimated to be (M = 44.7; SD = 25). Average number of vehicles involved in the road crashes is estimated as (M = 1.8; SD = 0.7) and most of the crashes recorded were minor (78.3%) with few considered as fatal (4.4%).

Driving experience by drivers involved in the accident was 6.2 years on the average. Statistics on the estimated mean number of people involved in road accidents under nature of road accident severity level is given as follows; Fatal (M = 16; SD = 8.2); Serious (M = 74; SD = 10.6); Minor (M = 211.4; SD = 28.6). Accidents as a result of commercial vehicles, private vehicles and motorbike accounted for 30.8%,60.3% and 8.9% respectively. The skewness and kurtosis of majority of the variables concerned were within the recommended values of ±2.00 (George & Mallery,George and Mallery, 2003) to indicate normality with few of the variables having skewness and kurtosis

Table 3. Brand test of parallel regression assumption.

Variable	X^2	*df*	*probability*
Omnibus	11.419	15.000	0.176
Nature of vehicles involved (Private)	6.492	1.000	0.058
Nature of vehicles involved (Motor bike)	1.983	1.000	0.159
Day of Week (Tuesday)	0.002	1.000	0.968
Day of Week (Wednesday)	2.182	1.000	0.140
Day of Week (Thursday)	0.425	1.000	0.514
Day of Week (Friday)	0.204	1.000	0.651
Day of Week (Saturday	0.501	1.000	0.479
Day of Week (Sunday)	0.283	1.000	0.595
Road type (Inter-Regional)	25.409	1.000	0.156
Road type (Regional)	20.589	1.000	0.755
Road type (Rural)	0.65	1.000	0.420
Average Speed Limit (ASL)	50.910	1.000	0.421
Road type (Rural)	437.873	1.000	0.199

Table 4. Descriptive statistics of road accidents variables.

Variable	%	*M*	*SD*	*Skew*	*Kurt*	*Min*	*Max*
Average age of vehicles+		4.2	1.4	0.8	1.8	1	6
Average years of driving by drivers involve+		6.2	2.3	−0.3	−0.4	1	11
Average Speed limit+		44.7	25	2	3.3	30	120
Number of Vehicles involve+		1.8	0.7	1.8	15.9	1	23
Average age of drivers+		39	14.7	0.8	0.4	3	99
Average age of casualty		36.5	18.5	0.6	−0.1	0	98
Number of Casualties		1.3	0.8	6.1	146.2	1	42
Pedestrians knocked down (PKD)		56.6	15	1.8	2.5	0	287
Number of people Injured		146.9	105.8	1.6	1.8	3	699
Nature of vehicle involves							
Commercial	30.8	193.7	21.6	1.3	0.6	4	872
Private	60.3	250.3	26.1	1.3	0.6	20	1049
Motorbike	8.9	39.6	19.8	1.6	2.4	0	504
Day of the week							
Monday (1)	11.5						
Tuesday (2)	13.8						
Wednesday (3)	15.0						
Thursday (4)	14.9						
Friday (5)	15.3						
Saturday (6)	16.3						
Sunday (7)	13.3						
Road type							
National	7.6						
Inter -regional	4.6						
Regional	15.6						
Rural	72.2						
Location where the accident occurred							
Urban	67.3						
Rural	32.7						
Accident severity level							
Fatal	4.4						
Serious	17.3						
Minor	78.3						

Note: Variables with + are proximate associated variables with road accidents severity level.

Table 5. Results of ordered logistic regression analysis of variables affecting road accident severity level.

Variable	*Beta(β)*	*Std. Error*	*z value*	*Pr(>|z|)*
Nature of vehicles involved (Private)	0.53[1]	0.02	30.61	< 2e-16***
Nature of vehicles involved (Motor bike)	0.52[2]	0.03	16.27	< 2e-16***
Day of Week (Tuesday)	0.21	0.03	6.73	1.73E-11***
Day of Week (Wednesday)	0.26[3]	0.03	8.15	3.57E-16***
Day of Week (Thursday)	0.22	0.03	6.93	4.24E-12***
Day of Week (Friday)	0.24	0.03	7.54	4.72E-14***
Day of Week (Saturday	0.17	0.03	5.68	1.37E-08***
Day of Week (Sunday)	0.06	0.03	2.05	0.040581*
Road type (Inter-Regional)	−0.18	0.05	−3.44	0.000592***
Road type (Regional)	−0.22	0.04	−5.60	2.20E-08***
Road type (Rural)	−0.37	0.03	−10.61	< 2e-16***
Average Speed Limit (ASL)	0.00	0.00	−7.76	8.49E-15***
Location (Rural)	−0.08	0.02	−4.79	1.71E-06***
Intercepts (Thresholds):				
Fatal |Serious	−3.04	0.05	−65.53	<2e-16 ***
Serious |Minor	1.23	0.04	27.72	-<2e-16 ***
LogLik=-55497.38; df = 15				
AIC = 11024.8; BIC = 111165.7				

Significance: Codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1.

Table 6. Odd ratio coefficient and confidence interval.

Variable	Odds Ratio (OR)	2.50%	97.50%
Nature of vehicles involved (Private)	1.705	1.649	1.761
Nature of vehicles involved (Motor bike)	1.674	1.574	1.781
Day of Week (Tuesday)	1.238	1.164	1.317
Day of Week (Wednesday)	1.291	1.215	1.372
Day of Week (Thursday)	1.242	1.168	1.320
Day of Week (Friday)	1.265	1.190	1.345
Day of Week (Saturday	1.188	1.126	1.255
Day of Week (Sunday)	1.066	1.003	1.133
Road type (Inter-Regional)	0.836	0.756	0.924
Road type (Regional)	0.800	0.740	0.864
Road type (Rural)	0.691	0.645	0.739
Average Speed Limit (ASL)	0.997	0.997	0.998
Location (Rural)	0.919	0.888	0.951

above the recommended values as proposed indicating non-normality.

Ordered logistic regression outcome

The Ordered Logistic technique was applied to the dataset and Table 5 depicts the results of the impact of the associated variables on road accidents severity level. Out of the total of nine (9) variables associated well with road accident severity level, the study reveals that only five (5) were statistically significant. The result presents the estimated β coefficients of ordered logistic regression analysis. Interpretation of the coefficients as indicated in Table 5 is however difficult; therefore, it is recommended to convert the log of odds into odds ratio for easier comprehension.

Table 6 shows the result of the odd ratio for 25% and 95% confidence interval of each coefficient. In Tables 5 and 6, the coefficient for the nature of vehicle is 0.53 for private and 0.52 for motorbike and the odds ratios were 1.705 and 1.674 respectively. The odds ratio means that a crash resulting into fata/serious/minor injury is 1.705 higher for private vehicles and 1.674 times higher for motorbike as compared to commercial cars. The odds of a crash being fatal/serious/minor is highest for Wednesday (1.29 times) referencing Monday. From the study, type of road is seen to be statistically significant with accident severity. Road type, inter-regional, regional and rural are about 0.836, 0.800 and 0.691 times less likely to result in fatal/serious/minor injuries as compared to national road type. With the increase in speed limit of the vehicles the odds of a crash being fatal/serious/minor will decrease (0.997 times). Locations of accident occurrence in the rural areas has lower odds of 0.919 times resulting in fatal/serious/minor injury as compared with urban areas.

Discussion

During the period studied (1989–2019), there was a steady upsurge in the number of road traffic accidents, death and injuries. This is as a result of high number of registered vehicles and the long stretches of relatively good highways that pass-through towns and villages, (Afukaar et al., 2003). Preliminary analysis discovered the significance of these variables – average age of the vehicle, average age of the driver, number of vehicles involved and driver experience - to road traffic crashes.

Ordered logistic regression analysis revealed that nature of vehicles involved, Day of the Week, Road Type, Average Speed Limit and Location were statistically significant predictors of road accident severity level in Ghana. The nature of vehicles has a significant influence on the road severity level in Ghana, the odds of accident crashes causing fatal/serious/minor injury is higher in private cars and motorbike as compared with commercial vehicles. This result is consistent with the study by ERSO (The European Road Safety Observatory) (2012). In this study with reference to Monday, the odds of causing fatal/serious/minor injury from Tuesday through to Sunday were higher. This finding is supported by other studies, (Odero et al., 1997; Adeloye & Ssembatya-Lule, 1997; Karacasu et al., 2014).

Speed at which a vehicle travels on the roads directly influences the risk of a crash as well as the severity of the accident, injuries and likelihood of death resulting from the incidence (Vadeby & Forsman, 2018).

The study result revealed that, the speed limit of a vehicle has a significant impact on the severity level of road crashes. This finding seemed to be in accordance with (Bachani et al., 2017). According to Bachani et al. (2017), 1% increase in average speed of a vehicle is likely to produce a 4% increase in the fatal crash risk and a three-percent increase in the serious crash risk. However, a five-percent reduction in the average speed limit of a vehicle can reduce fatalities of road crashes by approximately 30% (WHO, 2018). Previous studies have indicated that there is a significant relationship between location and severity level of accidents (Al-Ghamdi, 2002; Wahab & Jiang, 2019b). The odds of traffic accidents being fatal/serious/minor injury is higher in urban areas as compared to rural areas. This is not consistent with the study conducted by (Damsere-Derry et al., 2008; Ackaah & Adonteng, 2011).

One disadvantage of the study is that, in Ghana, crash records reported by the police are the main source of traffic crash data which is used to make the official estimates of fatalities in road traffic crashes. This may lead to underreporting which is foreseeable. Also, a more detailed list of variables could enhance the liability of the study. Based on the results obtained, accident severity level as a result of set of independent variables as indicated in the previous sections, it could be stated that to the best of our knowledge this is the first time such a macroscopic analysis of accident severity for different vehicle types is carried out in Ghana taking data for a period of 30 years.

Conclusion

Statistics indicates the severity of safety issue on our roads, and as such a greater emphasis must be placed on the issue of traffic accidents due to the harm and the ever-increasing road accidents cases in the country. Although many studies have taken place, they were carried out prevalently in countries where traffic characteristics, driver behaviour and geometry of roads differ from those of Ghana. This paper proposes an ordered logistic regression model to define the ordinal feature of road injury severity.

The estimated model indicates that, nature of vehicles involved, Day of the Week, Road Type, Average Speed Limit and Location have significant impact on crash severity level in Ghana. Certainly, these findings have policy and practical implication for stakeholders in road safety. To implement strategies to reduce crash injuries, there should be physical enforcement through the country police force on our roads and the application of technologies such as radars, camera to aid their work. It is believed that the results implicate the driving behaviour and as such these strategies will go a long way to reduce fatalities (Barua & Tay, 2010).

In addition, there will be identification of best practices in order to achieve a complete safer system to help raise awareness more effectively. There is also the need to train drivers to be more vigilant in their travels especially on the national roads and in the urban areas. In the long run we believe the identification of the road injury severity, will inform the health system as a whole in Ghana, the need to prioritize injury management. The recommendation is a well thought out and contextualised written laws and sanctioned schemes to monitor and enforce strict compliance with road traffic rules. These should be based on proving leading practice for effective enforcement. Further research could focus on examining additional parameters such as weather conditions, traffic conditions and other factors that influence road accidents.

Acknowledgement

We wish to thank the management of the Ghana police service for the provision of the dataset for the study.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding details

This study received no funding.

ORCID

Isaac Ofori Asare http://orcid.org/0000-0002-1503-1470
Alice Constance Mensah http://orcid.org/0000-0002-0435-010X

References

Ablin Consult. (2007). *Conditions of service of commercial vehicle drivers and the impact on road safety*. National Road Safety Commissions Report.

Ackaah, W., & Adonteng, D. O. (2011). Analysis of fatal road traffic crashes in Ghana. *International Journal of Injury Control and Safety Promotion*, *18*(1), 21–27. https://doi.org/10.1080/17457300.2010.487157

Adeloye, A., & Ssembatya-Lule, G. C. (1997). Aetiological and epidemiological aspects of acute head injury in Malawi. *East African Medical Journal*, *74*(12), 822–828.

Afukaar, F. K., Antwi, P., & Ofosu-Amaah, S. (2003). Pattern of road traffic injuries in Ghana: Implications for control. *Injury Control and Safety Promotion*, *10*(1–2), 69–76. https://doi.org/10.1076/icsp.10.1.69.14107

Akin, H. B., & Şentürk, E. (2012). Bireylerin mutluluk düzeylerinin ordinal lojistik regresyon analizi ile incelenmesi-analysing levels of happiness of individuals with ordinal logistic analysis. *Öneri Dergisi*, *10*(37), 183–193.

Albert, C. (2014). Road traffic accidents in Ghana: A public health concern, and a call for action in Ghana (and the sub- region). *Open Journal of Preventive Medicine*, *4*, 822–828. https://doi.org/10.4236/ojpm.2014.411092

Al-Ghamdi, A. S. (2002). Using logistic regression to estimate the influence of accident factors on accident severity. *Accident; Analysis and Prevention*, *34*(6), 729–741. https://doi.org/10.1016/S0001-4575(01)00073-2

Atubi, A. O. (2012). Determinants of road traffic accident occurrences in Lagos State: Some lessons for Nigeria. *International Journal of Humanities and Social Science*, *2*(6), 252–259.

Bachani, A. M., Zia, N., Hung, Y. W., Adetunji, R., Cuong, V., Faried, A., & Hyder, A. A. (2017). Speeding in urban South East Asia: Results from a multi-site observational study. *Journal of the Australasian College of Road Safety*, *28*(2), 27.

Barua, U., & Tay, R. (2010). Severity of urban transit bus crashes in Bangladesh. *Journal of Advanced Transportation*, *44*(1), 34–41. https://doi.org/10.1002/atr.104

Blows, S., Ivers, R. Q., Woodward, M., Connor, J., Ameratunga, S., & Norton, R. (2003). Vehicle year and the risk of car crash injury.

Injury Prevention: Journal of the International Society for Child and Adolescent Injury Prevention, *9*(4), 353–356. https://doi.org/10.1136/ip.9.4.353

Brant, R. (1990). Assessing proportionality in the proportional odds model for ordinal logistic regression. *Biometrics*, *46*(4), 1171–1178. https://doi.org/10.2307/2532457

Byrne, B. M. (2010). *Structural equation modeling with AMOS: Basic concepts, applications, and programming (multivariate applications series)* (Vol. 396, p. 7384). Taylor & Francis Group.

Caliendo, C., Guida, M., & Parisi, A. (2007). A crash-prediction model for multilane roads. *Accident; Analysis and Prevention*, *39*(4), 657–670. https://doi.org/10.1016/j.aap.2006.10.012

Damsere-Derry, J., Afukaar, F. K., Donkor, P., & Mock, C. (2008). Assessment of vehicle speeds on different categories of roadways in Ghana. *International Journal of Injury Control and Safety Promotion*, *15*(2), 83–91. https://doi.org/10.1080/17457300802048096

Damsere-Derry, J., Palk, G., & King, M. (2017). Road accident fatality risks for "vulnerable" versus "protected" road users in northern Ghana . *Traffic Injury Prevention*, *18*(7), 736–743. https://doi.org/10.1080/15389588.2017.1302083

Elvik, R., & Vaa, T. (2004). *The handbook of road safety measures*. Elsevier Science.

ERSO (The European Road Safety Observatory). (2012). *Country overviews*. http://ec.europa.eu/transport/road_safety/specialist/erso/pdf/country_overviews/dacota-country-overview-el_en.pdf.

Fox, J., & Hong, J. (2009). Effect displays in R for multinomial and proportional-odds logit models: Extensions to the effects package. *Journal of Statistical Software*, *32*(1), 1–24. https://doi.org/10.18637/jss.v032.i01

George, D., & Mallery, M. (2003). *Using SPSS for Windows step by step: A simple guide and reference*. Report 2014.

Global Road Safety Partnership (GRSP). (2014). *Annual Report 2014*. Global Road Safety Partnership.

Haadi, A. R. (2014). Identification of factors that cause severity of road accidents in Ghana: A case study of the Northern Region. *International Journal of Applied Science and Technology, 4*(3), 242–249.

Hair, J. F., Anderson, R. E., Babin, B. J., & Black, W. C. (2010). *Multivariate data analysis: A global perspective*. Pearson Education.

Haleem, K., & Gan, A. (2011). Identifying traditional and nontraditional predictors of crash injury severity on major urban roadways. *Traffic Injury Prevention*, *12*(3), 223–234.

Jeong, H., Jang, Y., Bowman, P. J., & Masoud, N. (2018). Classification of motor vehicle crash injury severity: A hybrid approach for imbalanced data. *Accident; Analysis and Prevention*, *120*, 250–261. https://doi.org/10.1016/j.aap.2018.08.025

Karacasu, M., Ergül, B., & Altin Yavuz, A. (2014). Estimating the causes of traffic accidents using logistic regression and discriminant analysis. *International Journal of Injury Control and Safety Promotion*, *21*(4), 305–313. https://doi.org/10.1080/17457300.2013.815632

Liang, F. (2005). Bayesian neural networks for nonlinear time series forecasting. *Statistics and Computing*, *15*(1), 13–29. https://doi.org/10.1007/s11222-005-4786-8

Long, J. S., & Freese, J. (2014). *Regression models for categorical dependent variables using stata* (3rd ed.). Stata Press.

Lord, D., & Miranda-Moreno, L. F. (2008). Effects of low sample mean values and small sample size on the estimation of the fixed dispersion parameter of Poisson-gamma models for modeling motor vehicle crashes: A Bayesian perspective. *Safety Science*, *46*(5), 751–770. https://doi.org/10.1016/j.ssci.2007.03.005

Miaou, S. P., & Lord, D. (2003). Modeling traffic crash-flow relationships for intersections: Dispersion parameter, functional form, and Bayes versus empirical Bayes methods. *Transportation Research Record: Journal of the Transportation Research Board*, *1840*(1), 31–40. https://doi.org/10.3141/1840-04

Michalaki, P., Quddus, M. A., Pitfield, D., & Huetson, A. (2015). Exploring the factors affecting motorway accident severity in England using the generalised ordered logistic regression model. *Journal of Safety Research*, *55*, 89–97. https://doi.org/10.1016/j.jsr.2015.09.004

Motor Traffic and Transport Department (MTTD) of the Ghana Police Service. (2018). *Road accident statistics*.

Odero, W., Garner, P., & Zwi, A. (1997). Road traffic injuries in developing countries: A comprehensive review of epidemiological studies. *Tropical Medicine & International Health : TM & Ih*, *2*(5), 445–460. https://doi.org/10.1046/j.1365-3156.1997.d01-296.x

Olmuş, H., & Erbaş, S. (2012). Analysis of traffic accidents caused by drivers by using Log-linear models. *Promet-Traffic & Transportation*, *24*(6), 495–504.

Osei–Kyei, R., & Chan, A. P. (2016). Developing transport infrastructure in Sub-Saharan Africa through public–private partnerships: Policy practice and implications. *Transport Reviews*, *36*(2), 170–186. https://doi.org/10.1080/01441647.2015.1077288

Park, S., Jang, K., Park, S. H., Kim, D. K., & Chon, K. S. (2012). Analysis of injury severity in traffic crashes: A case study of Korean expressways. *KSCE Journal of Civil Engineering*, *16*(7), 1280–1288. https://doi.org/10.1007/s12205-012-1527-3

Peden, M. (2004). World report on road traffic. *Injury Prevention*, *10*(2), 67–67. https://doi.org/10.1136/ip.2004.005405

Pour-Rouholamin, M., & Zhou, H. (2016). Investigating the risk factors associated with pedestrian injury severity in Illinois. *Journal of Safety Research*, *57*, 9–17. https://doi.org/10.1016/j.jsr.2016.03.004

Rezapour, M., Moomen, M., & Ksaibati, K. (2019). Ordered logistic models of influencing factors on crash injury severity of single and multiple-vehicle downgrade crashes: A case study in Wyoming . *Journal of Safety Research*, *68*, 107–118. https://doi.org/10.1016/j.jsr.2018.12.006

Ringle, C. M., Wende, S., & Becker, J. M. (2015). *SmartPLS 3*. SmartPLS GmbH. http://www.smartpls.com

Taylor, M. C., Baruya, A., & Kennedy, J. V. (2002). *The relationship between speed and accidents on rural single-carriageway roads* (Vol. 511). TRL.

Train, K. E. (2009). *Discrete choice methods with simulation*. Cambridge University Press.

Vadeby, A., & Forsman, Å. (2018). Traffic safety effects of new speed limits in Sweden. *Accident; Analysis and Prevention*, *114*, 34–39. https://doi.org/10.1016/j.aap.2017.02.003

Venables, W. N., & Ripley, B. D. (2013). *Modern applied statistics with S-PLUS*. Springer Science & Business Media.

Wahab, L., & Jiang, H. (2019a). A multinomial logit analysis of factors associated with Severity of motorcycle crashes in Ghana. *Traffic Injury Prevention*, *20*(5), 521–527. https://doi.org/10.1080/15389588.2019.1616699

Wahab, L., & Jiang, H. (2019b). Severity prediction of motorcycle crashes with machine learning methods. *International Journal of Crashworthiness*, 1–8. https://doi.org/10.1080/13588265.2019.1616885

Wang, Z., Yue, Y., Li, Q., Nie, K., Tu, W., & Liang, S. (2017). Analyzing risk factors for fatality in urban traffic crashes: A case study of Wuhan. *Sustainability*, *9*(6), 897. https://doi.org/10.3390/su9060897

Washington, S. P., Karlaftis, M. G., & Mannering, F. L. (2003). *Statistical and econometric methods for transportation data analysis*. Chapman Hall/CRC.

Washington, S. P., Karlaftis, M. G., & Mannering, F. L. (2010). *Statistical and econometric methods for transportation data analysis* (2nd ed.). Chapman Hall/CRC.

World Health Organization. (2015). *Managing speed*. http://www/who/int/violence_injury_prevention/publication/road_traffic/managing-speed/en/

World Health Organization. (2018). *Global status report on road safety*.

Xie, Y., Lord, D., & Zhang, Y. (2007). Predicting motor vehicle collisions using Bayesian neural Network models: An empirical analysis. *Accident; Analysis and Prevention*, *39*(5), 922–933. https://doi.org/10.1016/j.aap.2006.12.014

Yee, T. W. (2015). *Vector generalized linear and additive models: With an implementation in R*. Springer.

Yoon, S., Kho, S. Y., & Kim, D. K. (2017). Effect of regional characteristics on injury severity in local bus crashes: Use of hierarchical ordered model. *Transportation Research Record: Journal of the Transportation Research Board*, *2647*(1), 1–8. https://doi.org/10.3141/2647-01

Yuan, Q., Xu, X., Zhao, J., & Zeng, Q. (2020). Investigation of injury severity in urban expressway crashes: A case study from Beijing. *PLoS One*, *15*(1), e0227869. https://doi.org/10.1371/journal.pone.0227869