Journal of Infrastructure, Policy and Development 2024, 8(6), 6692.
https://doi.org/10.24294/jipd.v8i6.6692

Article

Prediction model of factors causing traffic accidents on rural arterial roads: A binary logistic regression approach

Novita Sari[1,2], Siti Malkhamah[1,*], Latif Budi Suparma[1]

[1] Faculty of Engineering, Gadjah Mada University, Yogyakarta 55281, Indonesia
[2] Polytechnic of Indonesian Land Transport, Bekasi 17510, Indonesia
*** Corresponding author:** Siti Malkhamah, malkhamah@ugm.ac.id

CITATION

Sari N, Malkhamah S, Suparma LB. (2024). Prediction model of factors causing traffic accidents on rural arterial roads: A binary logistic regression approach. Journal of Infrastructure, Policy and Development. 8(6): 6692. https://doi.org/10.24294/jipd.v8i6.6692

ARTICLE INFO

Received: 27 May 2024
Accepted: 19 June 2024
Available online: 25 June 2024

Abstract: In rural areas, land use activities around primary arterial roads influence the road section's traffic characteristics. Regulations dictate the design of primary arterial roads to accommodate high speeds. Hence, there is a mix of traffic between high-speed vehicles and vulnerable road users (pedestrians, bicycles, and motorcycles) around the land. As a result, researchers have identified several arterial roads in Indonesia as accident-prone areas. Therefore, to improve the road user's safety on primary arterial roads, it is necessary to develop models of the influence of various factors on road traffic accidents. This research uses binary logistic regression analysis. The independent variables are carelessness, disorderliness, high speed, horizontal alignment, road width, clear zone, road shoulder width, signs, markings, and land use. Meanwhile, the dependent variable is the frequency of accidents, where the frequency of accidents consists of multi-accident vehicles (MAV) and single-accident vehicles (SAV). This study collects data for a traffic accident prediction model based on collision frequency in accident-prone areas. The results, road shoulder width, and road sign factor all have an impact on the frequency of traffic accidents. According to a realistic risk analysis, MAV and SAV have no risk difference. After validation, this model shows a confidence level of 92%. This demonstrates that the model generates estimations that accurately reflect reality and are applicable to a wider population. This research has the potential to assist engineers in improving road safety on primary arterial roads. In addition, the model can help the government measure the impact of implemented policies and engage the public in traffic accident prevention efforts.

Keywords: arterial road; binary logistic regression; road safety; traffic accident factor

1. Introduction

Traffic safety relies on the integration and complex relationships between various components. The components are road users, vehicles, roads, and the environment (Cantisani et al., 2023). These components have different characteristics for each road section and could contribute to road traffic accidents (Bosurgi et al., 2024; C. Lee and Li, 2014; Wu et al., 2014; Yakupova et al., 2020). The causes of road traffic accidents are multi-factors (Chand et al., 2021; C. Lee and Li, 2014; Ma et al., 2021; Wang and Zhang, 2017). Therefore, it is necessary to carry out a comprehensive analysis to find out what factors influence the occurrence of traffic accidents.

In Indonesia, primary arterial roads in rural areas still have unrestricted access along roads that cross various types of land use. The land use activities around the road influence the traffic characteristics of the road section. Based on regulations, primary arterial roads are designed for high speeds. Hence, there is a mix of traffic between high-speed vehicles and vulnerable road users (pedestrians, bicycles, and motorcycles) around the land. As a result, several arterial roads in Indonesia have been identified as accident-prone areas (Aswardi et al., 2017; Mayuni et al., 2017; Pamungkas et al., 2017; Sujanto and Mulyono, 2010; Tjahjono and Nugroho, 2018).

To improve road safety, predicting the frequency of collisions has been carried out by researchers previously (Awadalla and De Albuquerque, 2021; Eboli et al., 2020; J. Lee et al., 2015; Umair et al., 2022). In the development of traffic accident models, they use statistical methods. Previous researchers used a logistic regression model to find factors causing the frequency and severity of collisions (Ariani, 2019; Champahom et al., 2022; Chen et al., 2016; Sze et al., 2014).

To identify safety targets, Sze et al. (2014) link vehicle ownership and national income levels to traffic accident fatality rates. Their findings suggest that setting lower safety achievement targets may more effectively enhance road safety. Similarly, Chen et al. (2016) utilized a hierarchical Bayesian logistic model to explore the factors significantly impacting crash rates, which varied in their effects on driver injury severity. Key influences on the severity of driver injuries include road curves, vehicle damage, the number of vehicles involved, road surface wetness, vehicle type, driver age and gender, safety belt usage, and alcohol or drug impairment. Moomen et al. (2019) found that factors affecting truck accidents encompass driver gender and age, weather conditions, lighting, road conditions, the number of inclines, accident type, the number of driveways, days of the week, and speed limits. Additionally, Champahom et al. (2022) identified critical factors influencing the severity of motorcycle accidents on Thailand’s arterial roads, including the driver’s age, the specific road route, and helmet use. Ariani (2019) noted that while vehicle factors have a nonsignificant impact, road and environmental factors, as well as human behavior, significantly affect traffic discipline among parents and children.

Traffic accidents at intersections are the most significant hazard worldwide. Due to unsafe crossing behavior, most pedestrians are victims of intersection crashes. These unsafe behaviors often result in pedestrian fatalities, especially if they violate traffic signals. A binary logistic regression model was also used to study and model pedestrian behavior at intersections. Proactive policies are necessary to design pedestrian-friendly facilities and reduce traffic violations, thereby improving pedestrian safety (Hapsari and Malkamah, 2012; Haque et al., 2024; Haque and Ahmad Kidwai, 2023). Traffic accidents also have a significant economic impact. The study by Pereira et al. (2020), using a negative binomial model, showed the effect of reduced income due to road accidents.

To enhance road safety in high-risk areas, it is essential to base recommendations on the factors contributing to accidents. Therefore, developing models that consider various factors influencing road traffic accidents is crucial for improving safety on primary arterial roads. The model of the influence of various factors on road traffic crashes has the main objective of analyzing the factors that cause crashes and formulating effective government policies to improve road safety. By understanding the causes of accidents, the government can design more targeted prevention programs, improve law enforcement, improve road infrastructure, and increase public awareness to reduce the risk of accidents. Furthermore, the model measures the impact of implemented policies and engages the community in traffic accident prevention efforts.

2. Materials and methods

The prediction model of traffic accident factors based on collision frequency uses traffic accident data from accident-prone areas as its data source. The primary arterial road 2/2UD (undivided), located in the accident-prone area of Ciamis Regency, West Java, Indonesia, serves as the research object. Around accident-prone areas, the land uses are residential, commercial, and educational.

Data for analysis was sourced from both secondary and primary data. The Ciamis Regency Police provided traffic accident data spanning the last five years (2016–2020) in Ciamis Regency. The Directorate of Road and Bridge Engineering Development, Ministry of Public Works, and Public Housing provided information on the road network and land use. The road network data utilized in this research encompassed road length, width, function, road equipment, and land use data across the study area.

The goal of the road inventory survey is to gather information or variables about the current state of the road section under study. The survey directly collects road inventory data from the field, which includes factors such as road length, road width, roadside hazards, traffic signs, road markings, intersection conditions and accessibility, complementary road facilities and direction systems, as well as land use around sections and intersections. The results of this survey can be used as a basis for determining road capacity. The V/C ratio of the road segment can then be determined by dividing the volume (pcu/hour) by the road capacity.

Vehicle speed surveys use a speed gun. The survey data shows the vehicle's average speed as it passes a specific point on the studied section. Traffic volume is the number of vehicles passing a certain point or section. Based on the Indonesian road capacity manual, types of vehicles are divided into two categories: motorized vehicles and non-motorized vehicles. Motorized vehicles include motorcycles (MC), light vehicles (LV), and heavy vehicles (HV).

2.1. Determination of variables from the factors causing accidents

The factors contributing to traffic accidents are assessed through an analysis of accident chronology, an evaluation of vehicle roadworthiness, an inspection of road geometry, and an examination of road equipment and roadside conditions.

The analysis was performed using SPSS 22 software with the Chi-square statistical test. The confidence level for this Chi-square test is 95%, or $\alpha = 5\%$. The hypotheses used to determine the relationship between accident frequency and the factors causing traffic accidents are as follows:

- If the test result shows a p-value ≥ 0.05, then Ho is rejected, indicating that there is no statistically significant correlation between the independent variable and the dependent variable.
- If the test result shows a p-value < 0.05, then Ha is accepted, indicating that there is a statistically significant correlation between the independent variable.

2.2. Binary logistic regression

In logistic regression, the model equation can be written as follows:

$$g(x) = \ln\left[\frac{p}{1-p}\right] = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \cdots + \beta_k x_k \quad (1)$$

where:

g(*x*): Dependent variable to nominal scale;

ρ: Probability of occurrence of an event;

x: Independent variable;

β: Parameter coefficients.

2.3. Probability

The probability of factors causing traffic accidents based on the frequency of traffic accidents is calculated using the equation:

$$p = \frac{1}{1 + e^{-y}} \quad (2)$$

where:

ρ = probability for the occurrence of an event

e = The number of natural (2.72)

$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \cdots + \beta_k$

x = independent variable

β = parameter coefficients

2.4. Relative Risk (RR)

The RR value is used to measure the relationship between factors that cause traffic accidents and the frequency of accidents. The RR value shows how high the risk of traffic accidents involving other vehicles is. Therefore, the probability of factors causing traffic accidents, based on the frequency of traffic accidents, can be calculated as follows:

$$RR = \frac{\rho_1}{\rho_2} \quad (3)$$

where:

RR = Relative Risk;

ρ_1 and ρ_2 = probability for the occurrence of an event.

3. Results

3.1. Traffic accident characteristics

Based on traffic accident data from the Ciamis Regency Police, there have been 153 accidents in accident-prone areas over the last five years (2016–2020). The highest frequency of incidents in the last five years occurred on Jalan Jendral Sudirman, amounting to 36% of the total incidents over the past five years. The fewest accidents occurred in Jalan Raya Cikoneng. **Table 1** describes the frequency of traffic accidents in accident-prone areas.

Table 1. Frequency of traffic accidents in accident-prone areas (2016–2020).

No	Location	Frequency of traffic accidents	Year				
			2016	2017	2018	2019	2020
1	Jalan Jend.Sudirman	55	13	10	8	16	8
2	Jalan Raya Imbanagara	44	8	9	10	6	11
3	Jalan Ciamis-Banjar	30	4	7	11	8	0
4	Jajalan Raya Cikoneng	24	6	6	8	3	1
Total		153	31	32	37	33	20

Generally, the severity of accident victims can be classified into several levels, ranging from minor to fatal injuries. The percentage of fatalities is 30%, serious injuries are 36%, and minor injuries are 35%. **Table 2** describes the severity of traffic accident victims in accident-prone areas.

Table 2. Accident severity of traffic accidents in accident-prone areas (2016–2020).

No	Location	Frequency of traffic accidents	Accident severity		
			Minor	Major	Fatal
1	Jalan Jend.Sudirman	55	27	29	19
2	Jalan Raya Imbanagara	44	22	31	14
3	Jalan Ciamis-Banjar	30	15	20	14
4	Jajalan Raya Cikoneng	24	9	13	16
Total		153	73	75	63

Traffic accidents are classified based on the road users involved, such as vehicle types and pedestrians. Traffic accidents based on vehicle type were the highest (52% for motorcycles, 16% for light vehicles, and 14% for heavy vehicles), while traffic accidents involving pedestrians amounted to 18%. This shows that traffic accidents involve 70% of vulnerable road users. **Figure 1** shows the percentage of road users involved in traffic accidents.

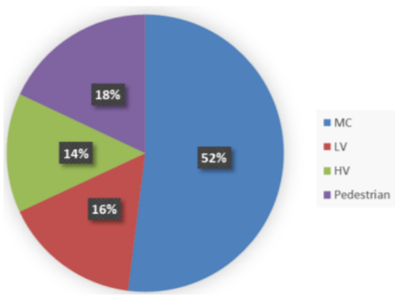

Figure 1. Road users involved in traffic accidents.

According to traffic accident characteristics, traffic accidents are caused by humans, roads, and environmental factors. Vehicle factors did not yield any data on the causes of accidents.

3.2. Human factor

In accident-prone areas, the frequency of traffic accidents over the past five years

(2016–2020) recorded only two causes: careless and disorderly drivers. Careless driving is a condition in which the driver temporarily loses concentration and alertness, thus becoming less aware of the surrounding situation. Careless driving, which accounts for 56% of total accidents, refers to a situation where a driver temporarily loses concentration or alertness, becoming less aware of their surroundings. This can encompass actions like using a mobile phone while driving, experiencing drowsiness, or becoming distracted by other elements in the vehicle.

Meanwhile, a disorderly driver is someone who drives their vehicle without obeying traffic rules. Disorderly drivers, accounting for 44% of the total accidents, are those who do not obey traffic rules such as breaking the speed limit, not using seat belts, and disobeying traffic signs. Driving a vehicle at high speeds is one of the factors causing traffic accidents. Vehicle speed surveys provide information on the speed of vehicles in accident-prone areas. The vehicle speed data was analyzed using the 85th percentile method. **Table 3** describes the results of the 85th percentile vehicle speed analysis.

Table 3. 85th percentile speed of vehicles in accident-prone areas.

No	Location	85 th percentile vehicle speed (KM/hour)					
		MC		LV		HV	
		Lane 1	Lane 2	Lane 1	Lane 2	Lane 1	Lane 2
1	Jalan jend Sudirman	51.00	47.45	48.05	48.20	40.35	49.30
2	Jl.Raya Imbanagara	43.00	51.90	41.00	46.00	40.30	41.50
3	Jl. Ciamis-Banjar						
	- Segment 1	66.90	52.00	47.90	45.70	49.80	41.90
	- Segment 2	56.60	49.80	65.80	57.40	59.00	58.40
4	Jalan Raya Cikoneng	56.00	41.75	44.75	44.50	46.50	40.25

Regulations govern vehicle speed based on land use (Peraturan Menteri Perhubungan Republik Indonesia, 2015). Based on the land use around the road, it appears that 85% of vehicle speeds exceed the maximum speed limit.

3.3. Road and environment factor

Determining the causes of traffic accidents based on environmental factors is carried out by analyzing data from Hawkeyes 2000 vehicle surveys and direct observations in the field. The results from road and environmental data are compared with standards and regulations (Kementerian Pekerjaan Umum dan Perumahan Rakyat, 2011; Kementerian Pekerjaan Umum dan Perumahan Rakyat Direktorat Jendral Bina Marga, 2021; Kementerian Perhubungan Republik Indonesia, 2014; Kementerian Perhubungan RI, 2018). **Table 4** describes the results of the analysis of the road and environmental factors causing traffic accidents.

Table 4. Road and environmental conditions in accident-prone areas.

No	Road and environment factor	Element	Standard	Result	Summary
1	Road Geometric	Horizontal alignment	Min. radius 148 m	89–178 m	53% of traffic accidents occur on curves with non-standard radius.
		Vertical alignment	Flat terrain max. 5%	± 3%	100% of the maximum grades meet the requirements.
		Road width	11 m	± 7 m	84% of the road width is not up to standard.
		Road Shoulder width	Min. 1 m	± 0.5 m	47% of road shoulder width is not up to standard.
2	Road surface	IRI	IRI ≤ 4 good 4 ≤ IRI ≤ 8 medium 8 ≤ IRI ≤ 12 light damage IRI > 12 heavily damage	-	37% of road surface conditions fall into the good category, 61% fall into the medium category, and 2% fall into the lightly damaged category. Road stability is above 90%.
3	Roadside	Clearzone	3–3.5 m (Speed < 60 KM/hour)	± 1–2 m	82% did not meet the standard.
4	Road Equipment	Traffic signs	Technical standards, function, location and reflective	-	22% have no traffic signs. In 78% of the locations where traffic signs are present, 68% of them do not meet the required standards.
		Road markings	Technical standards, function, location and reflective	-	6% have no road markings, and 94% have road markings. However, of the 94% of existing road markings, there are 37% that are no longer in accordance with standards.
		Public street lighting	Lights up at night	-	17% of accident sites have no public street lighting, and 83% have public street lighting.
5	Land use	Residential area Commercial areas (shops and offices) Educational areas	-	-	43% occur in residential, 39% of accidents occur in commercial areas (shops and offices), and 18% occur in educational areas.

The analysis of road and environmental factors causing traffic accidents reveals that vertical alignment satisfies standards, excluding it from the list of factors causing traffic accidents.

Based on Indonesian regulation (Kementerian Pekerjaan Umum dan Perumahan Rakyat Direktorat Jendral Bina Marga, 2021), the cumulative length of surface rise and fall per unit length yields the International Roughness Index (IRI), a measure of the unevenness of the road surface. Mathematically, IRI is the ratio between the cumulative length of damaged or potholed roads (in m units) and the total road length (in km units). So, the higher the IRI value (in m/km), the worse the road surface's condition. The IRI value provides information about a road section's pavement surface condition. The Directorate of Road and Bridge Engineering, Directorate General of Highways, and Ministry of Public Works and Public Housing conducted a survey using Hawkeyes 2000 vehicles, yielding the average IRI value for each road section under study. Meanwhile, for the last five years at the research location, the vehicle factor was not present in the data.

After identifying the factors that cause accidents in accident-prone areas, an analysis of the relationship between the factors that cause accidents and the frequency of accidents is carried out. The goal is to extract the model's accident-causing factors.

Chi-square analysis is used to determine the relationship between two variables.

In chi-square analysis, the dependent variable is the frequency of accidents, and the dependent variable is the factors causing traffic accidents (human, road, and environmental factors). If the chi-square significance value is less than 0.05, it indicates a relationship. According to **Table 5**, nine independent variables have a p value less than 0.05, indicating their potential use as independent variables in a multivariate model.

Table 5. Chi-square test results for each accident-causing variable.

No	Independent Variable	*p*-value
1	Carelessness	0.028
2	Disorderliness	0.112
3	High speed	0.000
4	Horizontal alignment	0.000
5	Road width	0.000
6	Clearzone	0.000
7	Road shoulder width	0.000
8	Signs	0.029
9	Markings	0.000
10	Land use	0.008

The *p* value requirement of <0.05 is not an absolute rule and does not have to be fulfilled if there is a variable that has a significance of >0.25. Since the disorderliness factor dominates accidents in Indonesia and significantly influences the dependent variable, it can still be included in the logistic regression analysis to create a model of factors causing traffic accidents.

3.4. Model of factors causing traffic accidents

Multivariate analysis seeks to determine which factors have the strongest influence on traffic accidents. This analysis is a binary logistic regression analysis using SPSS 22 software. This analysis simultaneously tests all factors causing accidents, also known as independent variables, against the dependent variable.

The independent variables are carelessness, disorderliness, high speed, horizontal alignment, road width, clearzone, road shoulder width, signs, markings, and land use. Meanwhile, the dependent variable is the frequency of accidents, which is composed of multi-accident vehicles (MAV) and single-accident vehicles (SAV). **Table 6** describes the results of the logistic regression test between factors causing traffic accidents and the frequency of traffic accidents.

Table 6. Results of binary logistic regression testing.

Variables	coefficients	Sig,	OR	95% C,I,for EXP(B)	
				Lower	Upper
carelessness	21.230	0.999	0.000	0.000	,
disorderliness	17.538	0.999	0.000	0.000	,
high speed	20.493	0.998	0.000	0.000	,

Table 6. (*Continued*).

Variables	coefficients	Sig,	OR	95% C,I,for EXP(B)	
				Lower	Upper
horizontal alignment	-1.078	0.407	0.340	0.027	4.346
road width	1.906	0.069	6.729	0.861	52.589
Clearzone	-2.383	0.066	0.092	0.007	1.170
road shoulder width	-2.044	0.023	0.130	0.022	.753
signs	2.248	0.017	9.469	1.503	59.664
markings	0.747	0.313	2.110	0.494	9.007
land use	-0.167	0.742	0.846	0.313	2.286
Constant	-44.416	0.999			
Omnibus Tests	0.000	<0.05			
Nagelkerke R Square	0.465				
Hosmer and Lemeshow Test	0.298	>0.05			
Classification Table	90.8				

The results of the analysis of model feasibility testing on binary logistic regression are as follows.

3.4.1. Omnibus test

Omnibus tests of model coefficients are simultaneous statistical tests. This study tested whether the independent variable simultaneously influenced the dependent variable. The significance level is 5%, so the decision-making criteria are as follows:

- If the *p*-value is <0.05, then H0 is rejected and H1 is accepted. This means that the independent variable simultaneously influences the dependent variable.
- If the *p*-value >0.05, then H0 is accepted and H1 is rejected. The interpretation suggests that the independent variables do not simultaneously influence the dependent variable.

The results of the Omnibus Test output in **Table 5** show that α 0.000 <0.05, which means a confidence level of 95%. To conclude that the first model is suitable for further analysis, at least one independent variable must significantly influence the dependent variable in this case.

3.4.2. Nagelkerke's r square

Binary logistic regression uses the Nagelkerke *R* square value to determine the coefficient of determination. Because the Nagelkerke *R* square value can be interpreted like the *R* square value in multiple regression, a Nagelkerke *R* square value close to zero indicates that the ability of the variables to explain the dependent variable is very limited. Conversely, a Nagelkerke *R* square value near one signifies that the independent variable can offer all the necessary information to forecast the variability of the dependent variable.

The multivariate analysis results show a value of 0.465 for Nagelkerke's R square. This suggests that independent variables (carelessness, high speed, horizontal level, road width, clear zone, road shoulder width, signs, markings, and land use) can explain 46.5% of the frequency of accidents (the dependent variable). Meanwhile, 53.5% was explained by other variables outside the independent variable in the model.

3.4.3. Hosmer and Lemeshow test

The feasibility test of the regression model was assessed using the Hosmer and Lemeshow tests, based on Chi-square values. The purpose of this model is to test the null hypothesis, which states that the empirical data fits the model. The following are the hypotheses formulated for this test:

- If the probability value (p-value) ≤ 0.05 (significance value), then H0 is rejected. This shows that there is a significant difference between the model and its observation value. As a result, the goodness-of-fit test cannot predict the value of an observation.
- If the probability value (p-value) is ≥ 0.05 (significance value), then H0 is accepted. This shows that the model matches its observation value. This indicates that the goodness of fit test can accurately predict the observation value.

The Hosmer and Lemeshow Test results yielded a Chi Square value of 9.548, accompanied by a probability value of 0.298 > 0.05. This indicates that the model has adequately explained the data. The Hosmer and Lemeshow test is used to determine the suitability of the model (goodness of fit) and whether it is in accordance with the empirical data or not.

3.4.4. Classification table

A value of 90.8% in the classification table indicates that the logistic regression model has a high level of accuracy, meaning that it is able to correctly classify 90.8% of the observations. This demonstrates excellent model performance.

3.4.5. The odds ratio (OR)

Based on the results of the significance value, there are two independent variables whose significance value is less than 0.05. These factors are road shoulder width, with a significance value of 0.023 <0.05, and road signs, with a significance value of 0.017 <0.05. This shows that the road shoulder width and road signs have the most influence on the frequency of accidents.

The odds ratio (OR) value reveals the strength of the correlation between the causative factor and the crash frequency. The OR value is a measure of comparison between the odds of two different groups or conditions. In binary logistic regression analysis, the OR describes the correlation between independent variables such as road shoulder width and road signs and the odds of crash frequency in traffic accidents. An OR greater than 1 indicates that the predictor variable increases the likelihood of the target class occurring (positive), whereas an OR smaller than 1 indicates that the predictor variable decreases the likelihood of the target class occurring.

The coefficient value for road shoulder width is −2.044, with an OR value of 0.130. This indicates a negative correlation between the variable road shoulder width and MAV. A negative coefficient value indicates that road shoulder width has a significant effect on SAV. The more standard the road shoulder width, the greater the chance that the frequency of single-traffic accidents will decrease. It applies to logistic regression analysis, where log-odds or logits of the frequency of road traffic accidents decreased by −2.044 units with every one unit increase in the variable "road shoulder width". The coefficient values reveal that a road shoulder width factor with a probability of 1/0.130 times, or 7.69 times, increases the chance of SAV occurrence.

Meanwhile, the sign factor has a coefficient value of 2.248 and an OR value of 9.469, indicating that it has the most influence on MAV occurrence. The OR value indicates that the less standard the sign is, the greater the likelihood that it becomes the cause of traffic accidents (9.469 times).

The logistic regression equation is modeled as follows:

$$g(x) = \ln\left[\frac{p}{1-p}\right] = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \cdots + \beta_k x_k \quad (4)$$

$g(x)$ is a variable bound to a nominal scale, thus $g(x) = y$, then:

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \cdots + \beta_k x_k \quad (5)$$

y = 44.416+ 2.248 (road signs) + (−2.044) (road shoulder width).

Each variable in the model has two values: 1 for "MAV" and 0 for "SAV." Thus, y's values are as follows:

$$y = 44.416 + 2.248\ (1) + (-2.044) \quad (6)$$

y = 45.

So, the probability of factors causing traffic accidents to the frequency of traffic accidents involving other vehicles (MAV) is:

$$pMAV = \frac{1}{1 + 2.72^{-(45)}} \quad (7)$$

$pMAV = 1.00$

Therefore, the analysis of the probability of factors causing traffic accidents shows that the frequency of accidents involving multiple vehicles (MAV) is 100%. On the other hand, the probability of factors causing single-vehicle accidents (SAV) is as follows:

$$pSAV = \frac{1}{1 + 2.72^{-(44.416)}} \quad (8)$$

ρSAV = 1.00

The results of the probability analysis for traffic accident factors on the frequency of single accidents (SAV) are 100%. If the model expresses the probability as 1, in logistic regression analysis or other statistics, it confidently predicts the event or outcome with no uncertainty. Probability 1 indicates a very high level of confidence in the event.

The result of the probability values of MAV and SAV can be calculated as the relative risk (RR) value. Relative risk is a statistical term that describes the risk of a particular event occurring in one group compared to another group. The RR value indicates how much risk the frequency of traffic accidents involving other vehicles poses. Therefore, the RR value, which represents the probability of factors causing traffic accidents relative to their frequency, is as follows:

$$\text{Relative Risk (RR)} = \frac{\rho MAV}{\rho SAV} = \frac{1.00}{1.00} = 1.00 \quad (9)$$

The relative risk value between MAV and SAV is 1. If the relative risk (RR) value is 1.00, it indicates that there is no significant difference in the risk or possibility of an accident occurring between the two groups of events, MAV and SAV. The risk of accident frequency in the MAV and SAV incident groups was equal or equivalent, and there was no significant difference between the two based on the model or analysis performed.

3.4.6. Model validation

This model validation is crucial to ensure that the developed model is reliable and relevant in identifying the contributing factors to accidents, thus enabling the design of effective prevention strategies.

The validation of this model was carried out in three distinct locations: West Bandung Regency, Purwakarta Regency, and Serang Regency. These locations consist of primary arterial roads (2/2TT), known for being accident-prone and sharing similar characteristics with the road sections used in the model. The accident data utilized for this validation spans a minimum of the past three years and was sourced from the district police departments in these regions.

Table 7 presents the details based on the analysis of traffic accident causation factors, which include road sign and road shoulder width, for Purwakarta Regency (Jl. Sukatani), West Bandung Regency (Jl. Purwakarta), and Serang Regency (Jl. Anyer).

Table 7. Number of accident incidents on validation road sections.

Causative factor	**Number of traffic accidents**			
	Model	**Validation area 1**	**Validation area 2**	**Validation area 3**
Road shoulder width	72	0	0	48
Road Sign	83	19	90	48

The model equation for the impact of causative factors on traffic accident frequency is as follows:

$$y_{\text{main model}} = 44.416 + 2.248\ (\text{signs}) + (-2.044)\ (\text{shoulder width}) = 44.416 + 2.248\ (83) + (-2.044)\ (72) = 83.83$$

Therefore, the validation value of y is as follows:

$y_{\text{area 1}} = 87.13$;

$y_{\text{area 2}} = 90$;

$y_{\text{area 3}} = 54.21$;

$y_{\text{average}} = 77.11$.

The average y-value for the validation areas is 77.11. Thus, the confidence level of the model for traffic accident causation factors in predicting the frequency of traffic accidents after validation is 92%.

4. Discussion

The model's confidence level results demonstrate the mathematical proof, logical consistency, and sufficiently close-to-reality estimations of the influence of traffic accident causative factors on the frequency of traffic accidents. Therefore, this model can be generalized to a broader population.

According to the results of the binary logistic regression model analysis, traffic signs and road shoulder width affect the frequency of traffic accidents. Previous studies have examined the effect of traffic signs and road shoulder width on traffic accident frequency.

4.1. The effect of traffic signs on traffic accidents frequency

One of the main principles of a safer road is the self-explaining road. A self-explaining road enables drivers to understand road conditions through traffic signs, thereby lowering the likelihood of accidents. Signs, as a means of traffic control, are mostly found on road sections. The existence of signs plays an important role in ensuring safety and providing information to road users. Therefore, the presence of traffic signs significantly contributes to enhancing safety in areas that are prone to accidents (Arianto and Heriwibowo, 2017; Kurniawan et al., 2015; Lestari and Anjarsari, 2020; Oktopianto and Pangesty, 2021; Pradana et al., 2019; Putri, 2014).

Traffic signs have a significant role in improving road safety (Mayuni et al., 2017; Sujanto and Mulyono, 2010). Traffic signs are elements of road equipment that can be symbols, letters, numbers, sentences, and/or a combination of both. Traffic signs provide directions, provide information, play a role in traffic management, and aid driver navigation. Installing repeated traffic signs at the right distance can reduce the frequency of accidents and even have an effectiveness of up to 75% (Pramesti and Budiharjo, 2020; Zheng et al., 2018).

The placement of traffic signs is an integral part of a road's design, based on its geometric characteristics and road functions. In this case, signs become one of the basic aspects of improving road safety (Arianto and Heriwibowo, 2017; Mayuni et al., 2017; Perdana et al., 2019; Sujanto and Mulyono, 2010). The performance of signs as road safety facilities must be known to suppress the increase in the frequency of traffic accidents (Aswardi et al., 2017; Kusuma et al., 2019; Pane et al., 2021). Traffic signs must be up to standard in order to reduce the risk of accidents (Priambodo and Siregar, 2018). If exposed to light at night, conventional types of signs must be clearly visible and illuminated. Therefore, the driver must be able to see and read the sign by using retro-reflective material (Arianto and Heriwibowo, 2017; Kusuma et al., 2019).

Traffic signs can influence drivers' behavior to reduce vehicle speed, making them one of the most effective devices for reducing vehicle speed (Adisthi, 2015; Andini et al., 2020; Sirait et al., 2014). Studies have shown that installing speed limit signs and warning signs in curve areas can decrease the frequency of accidents (Arianto and Heriwibowo, 2017; Guan et al., 2014; Manggala et al., 2016; Sumarsono et al., 2010). A delineator has a significant effect on decreasing speed in the curve segment (Zhao et al., 2015), and it can maintain the vehicle's position in the lane on the curved segment (Charlton, 2007).

Vehicle speed management aims to reduce the possibility of collisions and, if collisions do occur, prevent fatalities. Turner et al. (2017) propose how to manage speed on rural and urban arterial roads, locations where most accidents are fatal and seriously injured. Speed limit signs demonstrate their effectiveness in reducing speed. Despite the use of speed limit signs, the speed reduction remains temporary and has not reached the speed limit. Therefore, there is a need for road users to understand speed limits and safety awareness (Andini et al., 2020). To enhance pedestrian safety, road crossings, such as zebra crossings, can be supplemented with signs to indicate their presence (Sari and Hidayat, 2020).

4.2. The effect of road shoulder width on traffic accidents frequency

The shoulder width of the road for 2/2UD primary arterial roads with flat terrain type is 0.5–2 m (Kementerian Pekerjaan Umum dan Perumahan Rakyat, 2011). The road shoulder width can affect road capacity and traffic density. If the road shoulder width is too small for high traffic volumes, the likelihood of accidents may increase.

The shoulder of the road provides additional space in the side area of the main lane of the road. This can contribute to improving safety by providing a place for drivers who need additional areas to stop the vehicle or avoid emergencies, such as sudden stops or responding to traffic obstacles. In the bend area, the wider shoulder of the road provides greater free side space, which can reduce the risk of collisions (Ben-Bassat and Shinar, 2011; Zheng et al., 2018).

The influence of shoulder width, guardrails, and highway geometry on vehicle speed suggests that shoulder width has a significant influence on actual speed, lane position, and driving speeds that are considered safe, but when guardrails are absent, shoulder width loses many benefits and effects on driving behavior (Ben-Bassat and Shinar, 2011). Highway geometry, while useful in reducing driving speed, can also negatively impact the stability of lane position in sharp corners. Thus, controlling the width of the road shoulder and the placement of guardrails seems to be a safer approach to controlling speed and position in the lane.

5. Conclusions

Overall, improving road safety relies heavily on implementing preventive measures and initiatives aimed at reducing both the occurrence and severity of traffic accidents. Road geometry encompasses several components, including horizontal alignment, vertical alignment, lane width, road shoulder width, speed design, road equipment such as road signs and markings, and other road safety amenities. Based on the frequency of occurrence, the analysis of elements contributing to road traffic accidents indicates that traffic signs and road shoulder width are the primary factors influencing traffic accidents. The engineering method suggests the following strategies for improving road safety:

- It is necessary to install traffic signs that are according to standards, clearly visible, and understandable to road users.
- Make road shoulders wider in accordance with road technical standards, road status, and class.
- The road shoulder must have a steady structure and be usable in an emergency, so that the road shoulder surface is flush and almost the same height as the road body.
- In the bend area, install road safety equipment in the form of warning signs, delineators, and traffic mirrors. To prevent vehicles from exiting the road lane when passing curves, install road safety fences and add side-free areas to bends to ensure visibility at corners.
- Conduct vehicle speed management by installing speed limit signs based on road characteristics and land use around the road, such as in educational areas, residential areas, and commercial areas such as industrial areas, offices, and shopping center areas.

- It is important to harmonize road geometry and equipment, especially signs and markings.

The government plays a crucial role in formulating policies aimed at enhancing road safety. The government has the authority to take concrete steps to reduce the frequency of road traffic accidents, starting with building institutional systems, developing comprehensive policies, designing and implementing integrated and sustainable safety programs, and allocating budgets in an effort to improve road safety.

The Ministry of Transportation has primary responsibility for regulating the transportation sector, including roads. One of their key roles is to design and implement policies aimed at improving road user safety. This includes the development of road traffic regulations, vehicle safety standards, the improvement of safe road infrastructure, the supervision of transport operators, and public awareness programs on the importance of traffic safety. The Ministry of Transportation is also responsible for research and analysis related to traffic accident statistics to understand trends and causal factors, which form the basis for effective policy formulation. Through this role, the Ministry of Transportation is instrumental in creating a safer transportation environment for all road users.

Author contributions: Conceptualization, SM, LBS and NS; methodology, SM, LBS and NS; software, NS, SM and LBS; validation, NS, SM and LBS; formal analysis, SM, LBS and NS; investigation, NS, SM and LBS; resources, NS, SM and LBS; data curation, NS, SM and LBS; writing—original draft preparation, SM, LBS and NS. All authors have read and agreed to the published version of the manuscript.

Data availability statement: Some of the data used in this study may be available from the corresponding author upon reasonable request.

Funding: This research was funded by the Human Resource Development on Transportation Agency (Ministry of Transportation), grant number No. SK.164/BPSDMP-2020 and The APC was funded by the Center for Human Resources Development of Transportation Apparatus (Ministry of Transportation).

Acknowledgments: The author gives proper thanks to Gadjah Mada University, Polytechnic of Indonesian Land Transport, Ciamis Regency Police, and Directorate of Road and Bridge Engineering for administrative support and research data. They also thank Aulia, Candi, Iwan and Isti for their significant contribution given in surveying and data analysis.

Conflict of interest: The authors declare no conflict of interest.

References

Adisthi, M, Tjahjono, T. (2015). Development of a Method for Determining Safe Speed Signs on National Roads (Case Study: Trans Sulawesi Road) (Indonesian). Fakultas Teknik Universitas Indonesia.

Andini, F. R., Kusumastutie, N. S., Purwanto, E., et al. (2020). The Effectiveness of Speed Limit Sign and Marking as the Speed Management Devices. In: Proceedings of the 2nd International Symposium on Transportation Studies in Developing Countries (ISTSDC 2019).

Ariani, M. (2019). Analysis of Risk Factors for Traffic Accidents among Elementary School-Age Children in Yogyakarta City (Indonesian). Perpustakaan Universitas Gadjah Mada.

Arianto, S. B., & Heriwibowo, D. (2017). Evaluation of Traffic Signage Needs on the Border Road Section between Bantul and Gunungkidul Districts, Yogyakarta (Indonesian). Jurnal Penelitian Transportasi Darat, 18(1), 1. https://doi.org/10.25104/jptd.v18i1.107

Aswardi, T., Saleh, S. M., & Isya, M. (2017). Evaluation of Traffic Accidents from the Aspects of Road Geometric Visibility and Road Equipment Facilities at Sibreh Intersection (Indonesian). Jurnal Teknik Sipil Universitas Syiah Kuala, 1(1), 263–270.

Awadalla, D. M., & de Albuquerque, F. D. B. (2021). Fatal Road Crashes in the Emirate of Abu Dhabi: Contributing Factors and Data-Driven Safety Recommendations. Transportation Research Procedia, 52, 260–267. https://doi.org/10.1016/j.trpro.2021.01.030

Ben-Bassat, T., & Shinar, D. (2011). Effect of shoulder width, guardrail and roadway geometry on driver perception and behavior. Accident Analysis & Prevention, 43(6), 2142–2152. https://doi.org/10.1016/j.aap.2011.06.004

Bosurgi, G., Di Perna, M., Pellegrino, O., et al. (2024). Drivers' Steering Behavior in Curve by Means of New Indicators. Infrastructures, 9(3), 43. https://doi.org/10.3390/infrastructures9030043

Cantisani, G., Borrelli, C. C., Del Serrone, G., et al. (2023). Optimizing Road Safety Inspections on Rural Roads. Infrastructures, 8(2), 30. https://doi.org/10.3390/infrastructures8020030

Champahom, T., Wisutwattanasak, P., Chanpariyavatevong, K., et al. (2022). Factors affecting severity of motorcycle accidents on Thailand's arterial roads: Multiple correspondence analysis and ordered logistics regression approaches. IATSS Research, 46(1), 101–111. https://doi.org/10.1016/j.iatssr.2021.10.006

Chand, A., Jayesh, S., & Bhasi, A. B. (2021). Road traffic accidents: An overview of data sources, analysis techniques and contributing factors. Materials Today: Proceedings, 47, 5135–5141. https://doi.org/10.1016/j.matpr.2021.05.415

Charlton, S. G. (2007). The role of attention in horizontal curves: A comparison of advance warning, delineation, and road marking treatments. Accident Analysis & Prevention, 39(5), 873–885. https://doi.org/10.1016/j.aap.2006.12.007

Chen, C., Zhang, G., Liu, X. C., et al. (2016). Driver injury severity outcome analysis in rural interstate highway crashes: a two-level Bayesian logistic regression interpretation. Accident Analysis & Prevention, 97, 69–78. https://doi.org/10.1016/j.aap.2016.07.031

Eboli, L., Forciniti, C., & Mazzulla, G. (2020). Factors influencing accident severity: an analysis by road accident type. Transportation Research Procedia, 47, 449–456. https://doi.org/10.1016/j.trpro.2020.03.120

Guan, W., Zhao, X., Qin, Y., et al. (2014). An explanation of how the placement of traffic signs affects drivers' deceleration on curves. Safety Science, 68, 243–249. https://doi.org/10.1016/j.ssci.2014.04.007

Kementerian Pekerjaan Umum dan Perumahan Rakyat. (2011). Minister of Public Works Regulation Number: 19/PRT/M/2011 on Road Technical Requirements and Road Technical Planning Criteria (Indonesian). Available online: https://keselamatanjalan.wordpress.com/wp-content/uploads/2016/10/permen-pu-19-2011-persyaratan-teknis-jalan-dan-kriteria-perencanaan-teknis-jalan.pdf (accessed on 20 June 2024).

Kementerian Pekerjaan Umum dan Perumahan Rakyat Direktorat Jendral Bina Marga. (2021). Road Geometric Design Guidelines (No.13/P/BM/2021) (Indonesian). Kementerian Pekerjaan Umum dan Perumahan Rakyat Direktorat Jendral Bina Marga

Kementerian Perhubungan Republik Indonesia. (2014). Regulation of the Minister of Transportation of the Republic of Indonesia No. PM 13 of 2014 concerning Traffic Signs (Indonesian). Available online: https://www.regulasip.id/book/7395/read (accessed on 20 June 2024).

Kementerian Perhubungan RI. (2018). Regulation of the Minister of Transportation of the Republic of Indonesia No. 67 of 2018 concerning Amendments to Minister of Transportation Regulation No. 34 of 2014 concerning Road Markings (Indonesian). Kementerian Perhubungan RI.

Kurniawan, A., Septiana, D. A., Basuki, K. H., & Indriastuti, A. K. (2015). Traffic Accident Analysis on Primary Arterial Road Segments (Case Study of Maospati - Solo Road, Segment 28.029, STA 11+020-18+020) (Indonesian). Jurnal Karya Teknik Sipil, 4(4), 538–545.

Kusuma, Y., Sumarna, T., Mustika, D., et al. (2019). Traffic Sign Performance and Its Impact on Accidents (Case Study: Jalan Soekarno Hatta Bandung) (Indonesian). Potensi: Jurnal Sipil Politeknik, 21(2), 61–64. https://doi.org/10.35313/potensi.v21i2.1597

Lee, C., & Li, X. (2014). Analysis of injury severity of drivers involved in single- and two-vehicle crashes on highways in Ontario. Accident Analysis and Prevention, 71, 286–295. https://doi.org/10.1016/j.aap.2014.06.008

Lee, J., Abdel-Aty, M., & Jiang, X. (2015). Multivariate crash modeling for motor vehicle and non-motorized modesat the macroscopic level. Accident Analysis and Prevention, 78, 146–154. https://doi.org/10.1016/j.aap.2015.03.003

Lestari, U. S., & Anjarsari, R. I. (2020). Traffic Accident Analysis and Handling of Accident Prone Areas of Jalan Ahmad Yani (KM 17-KM36 Section) Banjarbaru City (Indonesian). Jurnal Teknologi Berkelanjutan (Sustainable Technology Journal), 9(2), 110–117.

Ma, Z., Mei, G., & Cuomo, S. (2021). An analytic framework using deep learning for prediction of traffic accident injury severity based on contributing factors. Accident Analysis and Prevention, 160(July), 106322. https://doi.org/10.1016/j.aap.2021.106322

Manggala, R., J., J. A., Purwanto, D., & Indriastuti, A. K. (2016). Case Study of Causal Factors of Traffic Accidents on Sharp Bends (Indonesian). Jurnal Karya Teknik Sipil, 4(4), 462–470.

Mayuni, S., Widodo, S., & Sulandari, E. (2017). Safety Evaluation of Road Infrastructure (Trans Kalimantan Road Case Study) (Indonesian). In: Proceedings of the National Conference on Civil Engineering and Planning (KN-TSP) 2017.

Moomen, M., Rezapour, M., & Ksaibati, K. (2019). An investigation of influential factors of downgrade truck crashes: A logistic regression approach. Journal of Traffic and Transportation Engineering (English Edition), 6(2), 185–195. https://doi.org/10.1016/j.jtte.2018.03.005

Oktopianto, Y., & Pangesty, S. (2021). Analysis of Accident Prone Areas on the Tangerang-Merak Toll Road (Indonesian). Jurnal Keselamatan Transportasi Jalan (Indonesian Journal of Road Safety), 8(1), 26–37. https://doi.org/10.46447/ktj.v8i1.301

Pamungkas, S. B., Amirotul, M., & Setiono. (2017). Analysis of accident-prone locations on the primary arterial roads of Surakarta city (Indonesian). E-Jurnal Matriks Teknik Sipil.

Pane, R. R., Lubis, M., & Batubara, H. (2021). Study of Road Safety Facility Needs in the Kisaran City Area of Asahan Regency (Indonesian). Buletin Utama Teknik, 16(3), 224–234.

Peraturan Menteri Perhubungan Republik Indonesia. (2015). Regulation of the minister of environment of the republic of Indonesia Number PM.111 of 2015 concerning Procedures for Determining Speed Limits. Peraturan Menteri Perhubungan Republik Indonesia. pp. 2004–2006.

Perdana, R. R., Permata, Y. K. A., Latifah, S., et al. (2019). Analysis of the Effect of Geometrics and Completeness of Traffic Signs on Accidents (Case Study: Uphill Kethekan Kec. Jambu, Road Section Ambarawa - Magelang Km. 46+000 TO 46+750) (Indonesian). Wahana Teknik Sipil: Jurnal Pengembangan Teknik Sipil, 24(1), 60. https://doi.org/10.32497/wahanats.v24i1.1607

Pradana, M. F., Intari, D. E., & Pratidina D, D. (2019). Analysis of Traffic Accidents and Their Causal Factors on Cilegon Highway (Indonesian). Kajian Teknik Sipil, 04(2), 165–175.

Pramesti, Y. T., & Budiharjo, A. (2020). Effectiveness of Speed Limit Signs on Collector Roads (Indonesian). Jurnal Teknik, 18(2), 73–81. https://doi.org/10.37031/jt.v18i2.98

Priambodo, M. I., & Siregar, M. L. (2018). Road Sign Detection Distance and Reading Distance at an Uncontrolled Intersection. E3S Web of Conferences, 65, 1–8. https://doi.org/10.1051/e3sconf/20186509004

Putri, C. E. (2014). Analysis of Accident Characteristics and Causal Factors at Blackspot Locations in Kayu Agung City (Indonesian). Teknik Sipil Dan Lingkungan, 2(1), 154–161.

Sari, N., & Hidayat, B. (2020). Transport Management for Sustainable Urban Development at Tambun Market Area. IOP Conference Series: Earth and Environmental Science, 501(1). https://doi.org/10.1088/1755-1315/501/1/012034

Sirait, W. J., Priyanto, S., & Jamson, S. (2014). The efficacy of motorcycle training to reduce self-reported negative behaviour and attitude (case study in UK and Indonesia). Civil Engineering Forum.

Sujanto, S., & Mulyono, A. T. (2010). Road Safety Inspection on Yogyakarta's South Ring Road (Indonesian). Jurnal Transportasi, 10(1), 13–22.

Sumarsono, A., Pramesti, F. P., & Sarwono, D. (2010). Modeling traffic accidents on curves due to the influence of horizontal alignment consistency in highway geometric design (Indonesian). Media Teknik Sipil, X(Juli), 85–92.

Sze, N. N., Wong, S. C., & Lee, C. Y. (2014). The likelihood of achieving quantified road safety targets: A binary logistic regression model for possible factors. Accident Analysis and Prevention, 73, 242–251. https://doi.org/10.1016/j.aap.2014.09.012

Tjahjono, N., & Nugroho, I. (2018). Ornamental plants as noise absorbers (Indonesian). In: Proceedings of the Conference on Innovation and Application of Science and Technology (CIASTECH 2018).

Turner, B., Makwasha, T., & Hillier, P. (2017). Infrastructure treatments for managing speeds on rural and urban arterial roads. Journal of the Australian College of Road Safety, 28(2), 13–20.

Umair, M., Rana, I. A., & Lodhi, R. H. (2022). The impact of urban design and the built environment on road traffic crashes: A case study of Rawalpindi, Pakistan. Case Studies on Transport Policy, 10(1), 417–426. https://doi.org/10.1016/j.cstp.2022.01.002

Wang, Y., & Zhang, W. (2017). Analysis of Roadway and Environmental Factors Affecting Traffic Crash Severities. Transportation Research Procedia, 25, 2119–2125. https://doi.org/10.1016/j.trpro.2017.05.407

Wu, Q., Chen, F., Zhang, G., et al. (2014). Mixed logit model-based driver injury severity investigations in single- and multi-vehicle crashes on rural two-lane highways. Accident Analysis and Prevention, 72, 105–115. https://doi.org/10.1016/j.aap.2014.06.014

Yakupova, G., Buyvol, P., & Shepelev, V. (2020). Identification of factors affecting the road traffic injury rate. Transportation Research Procedia, 50(2019), 735–742. https://doi.org/10.1016/j.trpro.2020.10.086

Zhao, X., Wu, Y., Rong, J., & Ma, J. (2015). The effect of chevron alignment signs on driver performance on horizontal curves with different roadway geometries. Accident Analysis and Prevention, 75, 226–235. https://doi.org/10.1016/j.aap.2014.12.004

Zheng, L., Sun, J., & Meng, X. (2018). Crash Prediction Model for Basic Freeway Segments Incorporating Influence of Road Geometrics and Traffic Signs. Journal of Transportation Engineering, Part A: Systems, 144(7), 04018030. https://doi.org/10.1061/jtepbs.0000155