Development of Accident Prediction Model for Rural Toll Road Sections

Aine KUSUMAWATI[a], Dwi Phalita UPAHITA[b], Sri HENDARTO[c]

[a] *Lecturer at Faculty of Civil and Environmental Engineering, Institut Teknologi Bandung, Jl. Ganesha No. 10, Bandung 40132; E-mail : aine@ftsl.itb.ac.id*
[b] *Civil Engineering Department Graduate Student, Faculty of Civil and Environmental Engineering, Institut Teknologi Bandung, Jl. Ganesha No. 10, Bandung 40132; E-mail : dwi.phalita@yahoo.com*
[c] *Same as the first author; E-mail : hendarto_sh2@yahoo.co.id*

Abstract: Traffic safety is commonly defined as a condition without accident. The risk of traffic accident occurrences at toll roads is high; one way to increase safety of the toll roads is by developing accident prediction model, which is an important tool for estimating and predicting road safety, identifying hazardous road locations, and also for evaluating treatment effectiveness. This research aims to develop a general accident prediction model for rural toll road sections in Indonesia by considering relation between accident frequencies and traffic flows, and also various roadway geometric and environment characteristics. Several types of model were developed using negative binomial regression model. The resulting general accident prediction model shows that accidents at rural toll road sections are positively correlated with annual average daily traffic and the number of horisontal curvature, but negatively correlated with the lane width.

Keywords: traffic safety, toll road, accident prediction model, negative binomial regression model.

1. INTRODUCTION

Traffic safety is commonly defined as a condition with no accident. A condition of safety is a result of many actions made by people to get that condition. According to Hauer (1997), 'road safety is a manifest in the occurrences of accident and their harm'. A broader definition of road safety is 'the number of traffic accidents (crashes), or accident consequences, by kind and severity, expected to occur on the entity during a specified period'.

WHO (2009) wrote that 91% of road traffic accident deaths occurred in low- and middle-income countries (such as Indonesia), an increase of 6% compared to similar report published in 2004. The condition was not good, remembering that majority of world's vehicle population were in high income countries. According to the report from Indonesian Police Department, there were 31,234 deaths in the road on 2010, which means in one hour 3-4 people died due to road traffic accident. The result of accident data analysis in 2010 showed that road accident in Indonesia has cause dead of 86 people everyday and 67% of the victims are in working ages. Loss of productivity of the victim and loss of material caused by the accidents were estimated to be 2,9 – 3,1 % of Indonesia's gross domestic product (GDP), or equal to Rp 205 – 220 trillion on 2010 with total GDP of Rp 7,000 trillion (Repubic of Indonesia - Ministry of Transportation, 2011). Thus, it needs serious and real efforts to make the road traffic accident occurrences and deaths due to road traffic accident reduce.

Until now, a comprehensive program for road traffic accident reduction has not been applied thoroughly in Indonesia due to limited road traffic accident data availability. The program includes stages of identification of hazardous road locations, diagnosis of safety problems occurred in the hazardous locations, selection of countermeasure actions, and evaluation on effectiveness of the selected countermeasures in reducing accidents. The condition makes accident reduction initiatives/programs set up by the government are not based on proper statistical analysis of the accident data, thus no assessment on the effectiveness of the programs can be made.

At toll roads, the risk of traffic accident occurrences is high. One way to increase safety of the toll roads is by developing accident prediction model, which is an important tool for estimating and predicting road safety, identifying hazardous road locations, and also for evaluating treatment effectiveness. The model relates accident frequencies with traffic flows and various roadway geometric and environment factors contributing to accident occurrences. In previous research by Rakhmat et al (2013), a local accident prediction model for Purbaleunyi toll road has been developed. This research aims to make a more general accident prediction model that will be able to be used at other toll roads in Indonesia. The function of this developed model is to produce estimations and predictions of accident frequencies that are free from regression-to-the-mean bias. The results of the estimation or prediction will be used as a base for development of various methods to identify hazardous road locations (blackspots) at toll roads which is more reliable statistically compared to traditional methods (based on historical accident number only) commonly used in Indonesia.

2. ACCIDENT PREDICTION MODEL

Accident prediction model is usually developed using multivariate modeling technique. The model may include ordinary linear regression models and nonlinear regression models such as Poisson and negative binomial regression model. However, the use of ordinary linear regression models is not favoured anymore. Maycock and Hall (1984), Jovanis and Chang (1986), Miao and Lum (1993), and also Turner and Nicholson (1998) concluded that the ordinary linear model was not appropriate for predicting accident frequency. There were various reasons supporting this assertion. First of all, the accident frequency is discrete, and hence does not follow the normal distribution. Furthermore, the variance in the accident frequency is not constant, but tends to increase as the flow increases. The number of accidents also cannot be negative while a normally-distributed error structure implies a substantial probability of a negative number of accidents, especially when the flow is small and the expected number of accidents is also small.

Maycock and Hall (1984), Bonneson and McCoy (1993), Miaou (1994), Poch and Mannering (1996), Turner and Nicholson (1998), Berhanu (2004) and Hiselius (2004) have empirically demonstrated the superiority of the Poisson or negative binomial regression over the ordinary linear regression models in analysing accident frequencies. The Poisson regression model is attractive in that there is only a single parameter to be estimated but it does have its limitations, especially when the variance of the accident data is constrained to be equal to the mean. Accident data were often found to have a variance greater than the mean (a case of over-dispersion), which can result in biased model coefficients and erroneous standard errors. To overcome this problem of over-dispersion, the negative binomial regression model is recommended.

The description of the negative binomial regression is as follow. If Y_i is an independent random variable that follows a negative binomial distribution with expected value μ_i, then the probability function of Y_i is given by:

$$f(Y_i = y_i) = \frac{\Gamma\left(y_i + \frac{1}{\alpha}\right)}{\Gamma(y_i + 1)\,\Gamma\left(\frac{1}{\alpha}\right)} \left(\frac{1}{1+\alpha\mu_i}\right)^{\frac{1}{\alpha}} \left(\frac{\alpha\mu_i}{1+\alpha\mu_i}\right)^{y_i} \quad ; \quad i = 1, \ldots, n \tag{1}$$

with α being the dispersion parameter, and $\alpha \geq 0$. The mean and variance of this negative binomial regession model are:

$$E(Y_i) = \mu_i = \exp\left(X_i^{'}\beta\right) = \exp\left(\sum_{1}^{p} x_{ij}\beta_j\right) \tag{2}$$

and

$$Var(Y_i) = \mu_i + \alpha\mu_i^2 \tag{3}$$

The Poisson regression model can be regarded as a limiting model of the NB regression model as α approaches zero.

3. METHODOLOGY

3.1 Research Plan

The objective of this research is to develop a general accident prediction model for rural toll road sections, which consider the relation between accident frequencies and traffic flows, roadway geometric, and environment characteristics. This research coversthree rural toll road sections, which arePurbaleunyi, Palimanan-Kanci, and Tangerang-Meraktollroads. In this research, the development of the accident prediction model was conducted by re-calibrating the previously developed accident prediction model of Purbaleunyi toll road by Rakhmat et al (2013), so that the model can be used for other rural toll road sections, using additional data from Palikanci and Tangerang-Merak toll roads.

The needed data consists of traffic data, accident data, and roadway geometric and environment conditions. After all of the data were gathered, the transferability of accident prediction model from the previous study was tested using data from Palimanan-Kanci and Tangerang-Merak toll roads. If the model passed the transferability test, it means that the model can be used directly for the estimation and prediction of accident number in Palimanan-Kanci and Tangerang-Merak toll road. But, if the model did not pass the test, the re-calibration process of the model must be conducted using additional data from the Palikanci and Tangerang-Merak toll road sections.

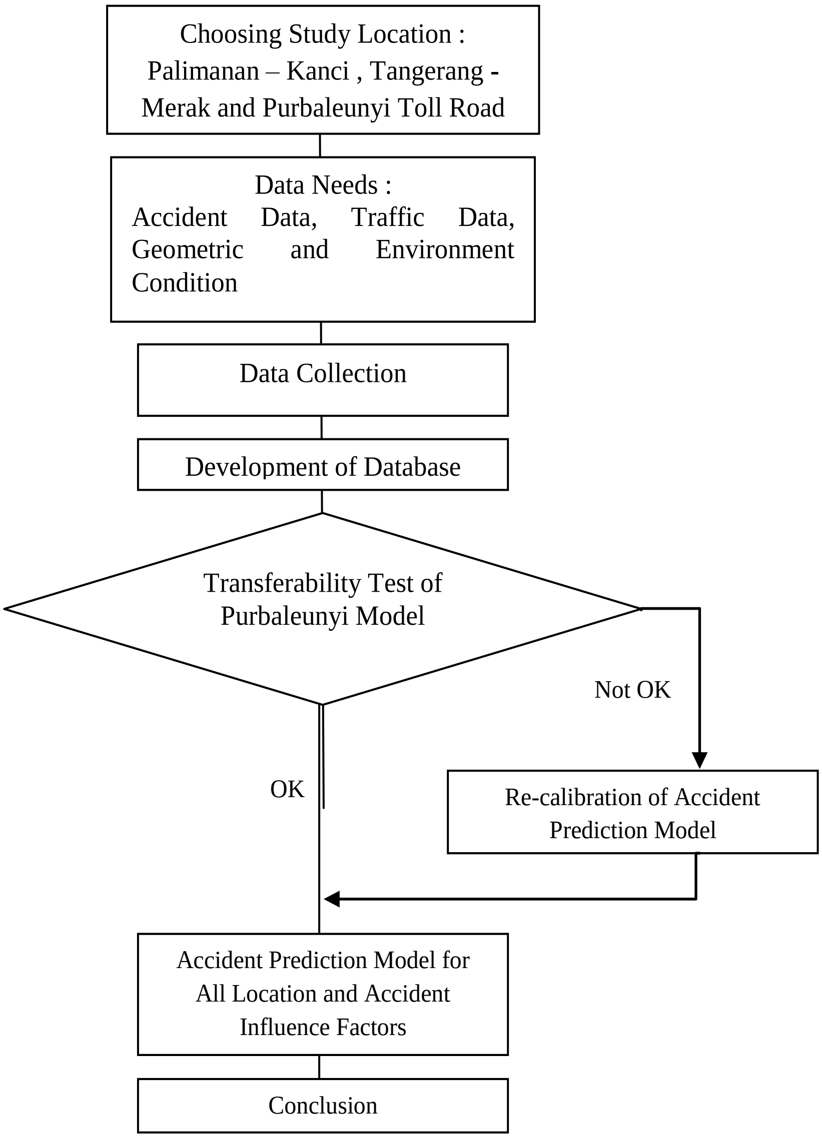

Figure 1. Research Steps

3.2 Research Object

The object of this research is Purbaleunyi, Palimanan-Kanci, and Tangerang-Merak toll roads. Palimanan-Kanci (Palikanci) toll road is a toll road that lays 26,2 km at Cirebon, West Java. This toll road lies from Plumbon Toll Gate until Kanci Toll Gate. The Palikanci toll road is a four-lane and two-way toll road. This toll road not only connects Palimanan with Kanci, but is also connected to Kanci-Pejagan toll road. The toll road is operated by PT Jasa Marga (Persero) Tbk.

Tangerang-Merak toll road is a toll road that connects Tangerang and Merak Port in Banten. This toll road is started from Cikupa Toll Gate and ended at Merak Toll Gate, lies about 72 km. It is operated by PT Marga Mandalasakti. Tangerang-Merak toll road is a four-lane and two-way toll road, with nine access gates. The toll road is connected with Jakarta-Tangerang toll road.

Purbaleunyi toll road is an 81 km toll road located at West Java. Purbaleunyi toll road is started at Kalihurip Interchange at Purwakarta and ended at Cileunyi Toll Gate, with 19 access gates located between them. It is operated by PT Jasa Marga (Persero) Tbk. Purbaleunyi toll road is divided into two segments, Cipularang segment and Padaleunyi segments.

3.3 Data

The accident prediction model was developed by taking the aggregated accident frequency of years 2007, 2008, and 2009 as dependent variable while traffic data (in the form of average Annual Average Daily Traffic - AADT of year 2007, 2008, and 2009) and various road geometric/environment data were model explanatory variables. Based on the accident data availability, the model was developed in the basis of constant 1-km length segments. Accident data for 2007-2010 period, traffic data for the same period, and road geometric/condition data were collected from PT Jasa Marga. Pavement condition data were obtained from Agency for Road Research, Indonesian Ministry of Public Works. The collected road geometric/condition data included type of median, median width, vertical grade, horizontal curve, pavement type, skid resistance, roughness, ramp and bridge locations. A video recording of road condition along the road was also carried out to support the data obtained from PT Jasa Marga.

3.4 Model Development Scenario

In this research, several accident prediction modelswere developed. According to the number of independent variables that were used in the development of the model, the accident prediction models were divided into two model types: single variable model and multivariate model. The single variable model uses only AADT variable as the independent variable. On the other hand, the multivariate model uses not only AADT but also geometric and environment variable as the independent variables.

According to the data that were used in the development process of the models and also the coverage of the model's location, the accident prediction models were divided into two types: general model and specific model. The general model was developed using the combination database from the three study objects; this model was intended to be used for all of the three study objects. The specific models were developed by using every toll roads own database and intended to be used only for the specific toll road section. In this research there were three specific models developed each for Purbaleunyi toll road, Palikanci toll road, and Tangerang-Merak toll road. It has to be noted that the specific model for Purbaleunyi toll has been developed on the previous research by Rakhmat et al (2013). Thus, this research developed specific models for Palikanci and Tangerang-Merak toll roads only. Both general and specific modelswere developed using single independent variable (single variabe models) and multi independent variable (multivariate models).

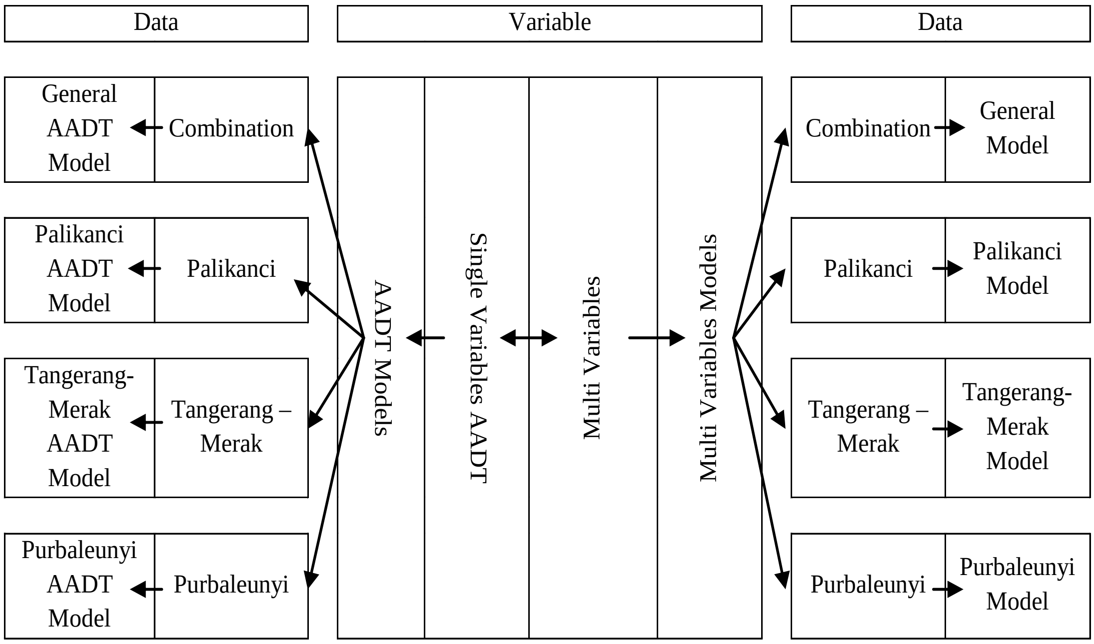

Figure 2. Model Development Scenario

3.5 Model Calibration and Development

The accident prediction for Purbaleunyi toll road that has been developed by Rakhmat et al (2013) is:

$$\mu = AADT^{0,484} \times \exp(0.242D+0.243M1 -3.238) \quad (4)$$

where μ is the expected accident frequencies (accidents in three years), D is the degree of curvature (in radian) and M1 is the existence of median with specification of width less than 2.5 m. The goodness-of-the-fit for the model is presented in Table 1, as follow.

Table 1. Goodness-of-The-Fit Results for General Multivariate Accident Prediction Model

Number of data	162
Number of parameter	4
Degree of freedom	158
Pearson χ^2	167.22
Scaled Deviance G^2	127.70
Critical $\chi2_{(0,95;158)}$	188.33
Result	**Accepted**

Firstly, a transferability test was carried out to the model shown in Equation (4). This test was intended to examine the suitability of this previously developed model to be applied in Tangerang – Merak and Palikanci toll roads. If this model did not pass the test, it will be re-calibrated using the additional data from the other toll road.

In the recalibration process, the new accident prediction models were developed by adding additional data from Palimanan-Kanci and Tangerang-Merak toll road sections into the Purbaleunyi toll road section database. Basic form of model that was used in the recalibration process is shown in Equation (5), as follow:

$$\mu = k \times (AADT)^{\beta} \times \exp(\gamma_1 X_1 + \gamma_2 X_2 + \cdots) \quad (5)$$

where μ is the expected accident frequencies (accidents in three years); k is a constant; β and γ are coefficients to be estimated in the model; AADT is the annual average daily traffic (in vehicles/day); X_1, X_2,.. are roadway geometric and environment variables.

During the model development/recalibration process, correlation test, regression analysis, and goodness-of-the-fit test were conducted. Correlation test was conducted to evaluate the correlation between variablesthose were used in the model development process. This process was conducted using Microsoft Excel. LIMDEP/NLOGIT version 4.0 software was used to estimate the model parameters; in this case, the negative binomial regression model was used. During the modeling process, variable selection was carried out using a backward elimination procedure whereby the most insignificant variable (based on the t-ratio or P-value of its estimated parameter) at 95% confidence level was progressively eliminated one by one from the model. This process was continued until all remaining variables were statistically significant to be retained in the model. The resulting models were considered appropriate when the estimated parameters of all model explanatory variables were statistically significant at 95% confidence level, as shown by the t-value of the estimated parameters that should be greater than or equal to 1.96, or by P-values of the estimated coefficients that should be less than or equal to 0.05. Other than that, the model shall pass goodness-of-fit assessment, in the form of Pearson X^2 and Scaled Deviance G^2 statistics.

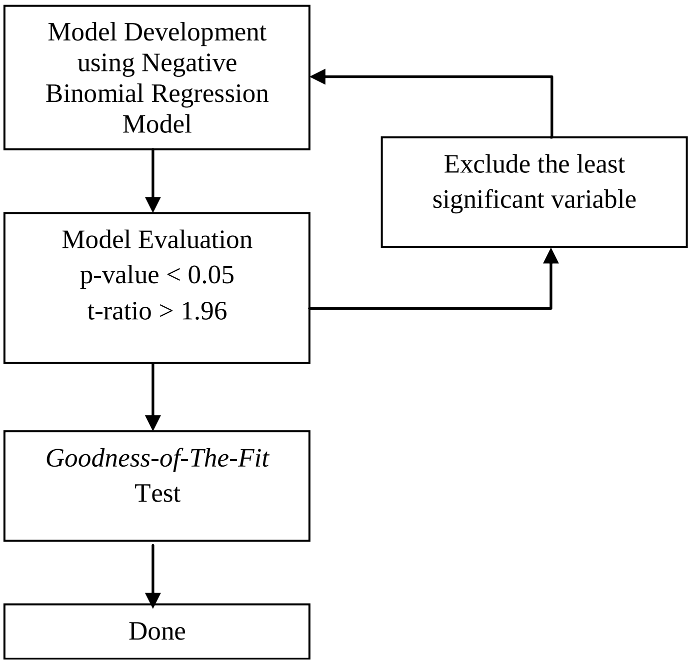

Figure 3. Model Development/Recalibration Process

4. MODELLING AND ANALYSIS

4.1 Accident Database

This accident database contains data tabulation for all variables that will be used in the modelingprocess, both independent and dependent variables. In this case, the dependent variable is the number of accident (LAKA), and the independent variables which are considered in the model development/recalibration are average annual daily traffic (AADT), existence of ramp junction (RAMP), number of bridge (J), radius of curvature (R), degree of curvature (D), number of horisontal curve (NTIK), curvature (B), vertical gradient (L), percentage of guardrail coverage area (G), lane width (W), type of pavement (P), number of bridge pillars (TJ), and number of roadsign pillars (TR).

Table 2 describes all variables that are included in the accident database and the range of those variables.

Table 2. Variables Description

Variable	Description	Value	
		Min	Max
LAKA	Number of accident in 3 years	0	78
AADT	Average annual daily traffic in 3 years	5186	29113
RAMP	Existence of ramp junction, 1 if exist and 0 if none	0	1
J	Number of Bridge	0	3
R	Radius of curvature (meter)	0	10000
D	Degree of curvature (rad)	0	4.584
NTIK	Number of horisontal curve	0	3
B	Curvature (rad/km)	0	3.841
L	Vertical gradient (%)	-4.86	4.86
M1	Existence of median with specification width < 2,5 m, 1 if it exists and zero if it not exist	0	1
G	Percentage of guardrail coverage area	0	1
W	Lane width (m)	3.5	4
P	Type of pavement, one for rigid pavement and zero for flexible pavement	0	1
TJ	Number of bridge pillars	0	18
TM	Number of road sign pillars	0	13

4.2 Transferability Test

Transferability test was conducted to examine the fitness of Purbaleunyi accident prediction model (Equation 1) when applied to Palikanci and Tangerang-Merak toll road data. The parameter that was used to check the transferability of the modelwas based on the Pearson χ^2. The expected acident number for Palikanci and Tangerang-Merak toll roads were firstly calculated using the Purbaleunyi accident prediction model before the Pearson χ^2 value can be calculated.

Transferability test calculation process, as explained in previous research by Sawalha (2002), was conducted as explained below. The first step is the calculation of Pearson χ^2 value using Equation (6) as follow:

$$\chi_p^2 = \sum \frac{\left[y_i - E(y_i)\right]^2}{E(y_i)\left[1 + E(y_i)\alpha\right]} \tag{6}$$

In this case, the overdispersion value (α) of the Purbaleunyi accident prediction model was 0.160. It was obtained that $\chi_p^2 = 1302$.

The next step is the calculation of the expected value, $E(\chi^2_p)$ and standard deviation, $\sigma(\chi^2_p)$ of Pearson χ^2 parameter using Equation (7) as follow:

$$\sigma\left(\chi_p^2\right)=\sqrt{2N\left(1+3\alpha\right)+\sum_{i=1}^{N}\frac{1}{E\left(y_i\right)\left[1+E\left(y_i\right)\alpha\right]}} \qquad (7)$$

where $E\left(\chi_p^2\right)=N=200$. By inputting all the required values into the equation, the value for $\sigma\left(\chi_p^2\right)$ was obtained to be 24.85.

Finally, the parameter to evaluate the transferability of the Purbaleunyi accident prediction model was calculated using the following formula:

$$z=\frac{\chi_p^2-E\left(\chi_p^2\right)}{\sigma\left(\chi_p^2\right)} \qquad (8)$$

The value of the z was found to be 44,43, which was out of the acceptable range of z. It indicates that Purbaleunyi Model is not suitable to be applied in other locations, especially in Palikanci and Tangerang – Merak toll roads.

4.3 Accident Prediction Model Development Process

The transferability test result shows that the Purbaleunyi accident prediction model is not suitable to be applied in Palikanci and Tangerang – Merak toll roads. Therefore, it was necessary to develop/calibrate a new accident prediction model that can be applied for all of the three toll roads. The calibration of the new model was carried out using the combined data from all of the three toll roads: Purbaleunyi, Palikanci, and Tangerang – Merak toll roads.

4.3.1 Correlation test

After the accident database for the three toll roads were combined, the process was continued with correlation test for all variables that would be considered in the modeling process. This correlation test was conducted to see the correlation within variables, not only between independent variable and dependent variable but also within the independent variables itself. Independent variables shall have high correlation values with dependent variable, which indicates that there are strong correlations between the independent and dependent variables. On the contrary, the correlation between independent variables shall be low. Table 3 shows the result of correlation test.

In this correlation test, when there was strong correlation between two independent variables, then only one of the two variables which have a higher correlation value with the dependent variable would be included in the modeling process later.

Table 3. Correlation Test Result

	D	G	J	L	LAKA	LHRT	M1	NTIK	P	R	RAMP	TJ	TM	W	B
D	1.000	0.046	-0.154	-0.003	**0.138**	-0.037	-0.050	0.681	0.169	-0.084	0.003	-0.214	0.048	-0.248	0.494
G		1.000	0.850	0.000	**-0.237**	-0.080	0.209	0.053	-0.120	-0.064	0.058	-0.115	-0.013	0.237	0.063
J			1.000	-0.002	**-0.139**	0.033	0.349	-0.110	0.078	-0.039	-0.020	0.177	0.099	0.031	0.124
L				1.000	**-0.002**	0.019	0.028	0.000	-0.001	0.000	0.002	0.000	-0.055	0.000	0.000
LAKA					**1.000**	**0.352**	**-0.193**	**0.190**	**0.056**	**-0.083**	**0.010**	**-0.083**	**-0.052**	**-0.258**	**0.005**
LHRT						1.000	-0.012	0.085	0.268	-0.094	-0.079	-0.157	0.046	-0.398	-0.031
M1							1.000	-0.021	0.289	-0.091	-0.042	-0.030	0.244	-0.047	0.242
NTIK								1.000	0.055	0.394	-0.004	-0.171	0.064	-0.130	0.608
P									1.000	-0.178	-0.135	-0.140	0.162	-0.539	0.179
R										1.000	0.013	0.128	-0.032	0.274	0.117
RAMP											1.000	0.082	0.120	0.119	-0.074
TJ												1.000	-0.053	0.519	-0.099
TM													1.000	-0.148	0.105
W														1.000	-0.136
B															1.000

4.3.2 Development of the multivariate models

The remaining variables that pass the correlation test were included in modelling process. During the modeling process, the selection of model' variables was carried out using a backward elimination procedure whereby the most insignificant variable (based on the t-ratio or P-value of its estimated parameter) at 95% confidence level was progressively eliminated one by one from the model. Significant variables is indicated by the t-value of the estimated parameters which should be greater than or equal to 1.96, or by P-values of the estimated coefficients which should be less than or equal to 0.05.

The resulting general accident prediction model is shown in Table 4. The estimated dispersion parameter α is found to be greater than zero, with t-value that is greater than 1.96 and P-value that is approaching zero, thus the appropriateness of the Negative Binomial model is confirmed. It has to be noted that the resulting model contain no constant parameter, since it has been found to be not significant to be included in the model.

Table 4. Final Result of General Multivariate Accident Prediction Model

Variable	Coefficient	Standard Error	t-value	p-value
NTIK	0.103	0.048	2.136	0.0327
LN(AADT)	0.521	0.053	9.811	0
W	-0.911	0.139	-6.532	0
α (dispersion parameter)	0.202	0.0282	7.163	0

Table 5 below shows the result of the goodness-of-the-fit for the general multivariate accident prediction model. It can be seen from the table that the Pearson χ^2 and Scaled Deviance G^2 values of the resulting model are lower than the critical χ^2 value, which means that the general multivariate accident prediction model pass the goodness-of-the-fit test.

Table 5. Goodness-of-The-Fit Results for General Multivariate Accident Prediction Model

Number of data	367
Number of parameter	3
Degree of freedom	364
Pearson χ^2	364.42
Scaled Deviance G^2	325.17
Critical $\chi 2_{(0,95;364)}$	409.47
Result	**Accepted**

As stated before, this research also developed specific models for individual toll road locations. Using similar methodology, the specific models for individual toll road locations were developed using data from its own location. The resulting multivariate general model and specific models (for Palikanci toll road and Tangerang – Merak toll road) are:

General model:

$$\mu = AADT^{0.521} \exp(0.103NTIK - 0.911W) \quad (9)$$

Palikanci model:

$$\mu = AADT^{0.165} \exp(-(4.19 \times 10^{-5})R) \quad (10)$$

Tangerang-Merak model:

$$\mu = AADT^{0.696} \exp(0.090TJ - 0.244P - 4.447) \quad (11)$$

where,

μ : accident number (accident/3 years)
AADT : average annual daily traffic (vehicle/day)
TJ : number of bridge pillar
P : pavement type (1 for rigid pavement, 0 for flexible pavement)
R : radius of curvature (m)
NTIK : number of horizontal curvature
W : lane width (m)

The goodness-of-the-fit test results for the Palikanci and Tangerang-Merak model are presented in Tables 6 and 7.

Table 6. Goodness-of-The-Fit Results for Palikanci - Multivariate Accident Prediction Model

Number of data	54
Number of parameter	2
Degree of freedom	52
Pearson χ^2	57.14
Scaled Deviance G^2	49.71
Critical $\chi2_{(0,95;52)}$	69.83
Result	**Accepted**

Table 7. Goodness-of-The-Fit Results for Tangerang-Merak - Multivariate Accident Prediction Model

Number of data	143
Number of parameter	4
Degree of freedom	139
Pearson χ^2	146.36
Scaled Deviance G^2	144.72
Critical $\chi2_{(0,95;139)}$	168.61
Result	**Accepted**

4.3.3 Development of the single variable models

Tables 8 and 9 show the results of modelling and goodness-of-the-fit test for the general single variable model (General – AADT model).

Table 8. Final Result of General – AADT Model

Variable	Coefficient	Standard Error	t-value	p-value
LNLHRT	0.188	0.003	52.915	0
Alpha	0.258	0.032	8.030	0

Table 9. Goodness-of-The-Fit Test for General – AADT Model

Number of data	367
Number of parameter	1
Degree of freedom	366
Pearson χ^2	371.72
Scaled Deviance G^2	334.24
Critical $\chi2_{(0,95;366)}$	411.59
Result	**Accepted**

The resulting model shows that traffic flow has positive correlation with accident number. It means that the more traffic flow in a location the more accident will occur on that location. The estimated dispersion parameter α is found to be greater than zero, with t-value that is greater than 1.96 and P-value that is approaching zero, thus the appropriateness of the Negative Binomial model is confirmed. It has to be noted that the resulting model contain no constant parameter, since it has been found to be not significant to be included in the model.

The result of the goodness-of-fit test of the model shows that the Pearson χ^2 and Scaled Deviance G^2 values of the resulting model are lower than the critical χ^2 value, which means that the general multivariate accident prediction model pass the goodness-of-the-fit test.

Similarly, this research also developed specific models for individual toll road locations. The resulting general-AADT model and specific-AADT models (for Palikanci toll road, Tangerang – Merak toll road, and Purbaleunyi toll road) are:

General - AADT model :

$$\mu = AADT^{0.188} \quad (12)$$

Palikanci - AADT model :

$$\mu = AADT^{0.129} \quad (13)$$

Tangerang-Merak - AADT model :

$$\mu = AADT^{0.227} \quad (14)$$

Purbaleunyi - AADT model :

$$\mu = AADT^{0.161} \quad (15)$$

The goodness-of-the-fit test results for the Palikanci, Tangerang-Merak, and Purbaleunyi - AADT models are presented in Tables 10, 11, and 12, respectively.

Table 10. Goodness-of-The-Fit Results for Palikanci - AADT Accident Prediction Model

Number of data	54
Number of parameter	1
Degree of freedom	53
Pearson χ^2	65.61
Scaled Deviance G^2	16.29
Critical $\chi 2_{(0,95;53)}$	70.99
Result	**Accepted**

Table 11. Goodness-of-The-Fit Results for Tangerang-Merak - AADT Accident Prediction Model

Number of data	143
Number of parameter	1
Degree of freedom	142
Pearson χ^2	139.28
Scaled Deviance G^2	71.53
Critical $\chi 2_{(0,95;142)}$	170.81
Result	**Accepted**

Table 12. Goodness-of-The-Fit Results for Purbaleunyi - AADT Accident Prediction Model

Number of data	179
Number of parameter	1
Degree of freedom	178
Pearson χ^2	186.12
Scaled Deviance G^2	178.52
Critical $\chi 2_{(0,95;178)}$	209.52
Result	**Accepted**

4.4 Analysis

The general multivariate accident prediction model, as presented in Equation (9), is formed by several parameters, which are average annual daily traffic, number of horisontal curvatures, and lane width. Average annual daily traffic is the exposure of accident, the higher the traffic flow in a location will result in the higher possibility of accident occurrences in that location. The number of horisontal curvatures has a positive correlation with accident number, which can be explain as follow: the number of horisontal curvatures in a road segment relates with the radius of curvature; the more number of horisontal curvatures in a road segment implies that the radius of curvatures in that particular segment are small, which means that the curvatures are sharper and this sharper curvatures may cause a safety problem for some lack-of-concentrate drivers. Lane width has a negative correlation with accident number which means that the wider the lane, the lower the accident number will be. A narrower lane width gives less space for vehicle movements, so the driver must drive very carefully in order to stay on their lane and to not disturb other vehicles.

The specific multivariate accident prediction model for Palikanci toll road, as presented in Equation (10), shows that accident occurrences in Palikanci toll road is influenced by average annual daily traffic and radius of curvature only. The radius of curvature has a negative correlation with accident number, it means that larger radius results in lower number of accident. A smaller radius of curvature will form a sharper curve, if it is not countered by proper driving skill of the driver, it may lead to an accident.

The specific multivariate accident prediction model for Tangerang – Merak toll road, as presented in Equation (11) shows that accident occurrences in Tangerang – Merak toll road is influenced by average annual daily traffic, number of bridge pillars, and type of pavement. Number of bridge pillars and average annual daily traffic has positive correlation with accident number. The existence of bridge pillars in a road segment may lead to the occurrence of accident for vehicles that are out of travelled lane unpurposely. The type of pavement gives a negative correlation to accident number. Considering there are only two quantitative values representing the type of pavement (i.e. zero for flexible pavement and one for rigid pavement) means that more traffic accidents at Tangerang-Merak toll roads occurr on the flexible pavement sections than on the rigid pavement sections.

In general, different characteristics of a toll road result in different variables influencing accident occurrences. Variables that are found to influence accident occurrences in rural toll road segments may include average annual daily traffic, radius of curvature, number of curve, lane width, number of bridge pillars, and type of pavement. It turns out that there are other important roadway geometric variable, such as vertical grade, which was not found to be statistically significant to be included in the accident prediction model. This situation may happen as an occurrence of accidents in a particular segment may be caused not by the adverse geometric condition of the segment itself, but by the adverse geometric condition of the previous segment. For example, at steep downgrade condition in a segment, a driver may loss control of his vehicle. However, if the incident ends with accident, the accident may not happen at the particular downgrade segment but in the subsequent segments when the downgrade ends. Therefore, it may be necessary to consider the influence of roadway geometric condition in previous segment in developing the accident prediction models.

5. CONCLUSIONS AND RECOMMENDATIONS

The research has developed a general accident prediction model for rural toll roads in Indonesia based on Purbalenyi, Palikanci, and Tangerang – Merak toll roads data. The models were developed using a negative binomial regression model. The model shows that accident occurrences in rural toll road in Indonesia is influenced by average annual daily traffic, number of horisontal curvatures, and lane width.

The research has also developed specific accident prediction models for Palikanci and Tangerang – Merak toll roads. Theoretically, the specific accident prediction model shall give better estimation/prediction than the general accident prediction model. Nevertheless, it is expected that the general accident prediction model can be applied for other rural toll road locations in Indonesia. It is envisaged that before the general model is applied for other toll road locations, the transferability of the model to the other toll road locations shall be examined first.

The research found that different characteristics of a toll road result in different variables influencing accident occurrences. One important finding from this research is that vertical grade, which is believed as an important roadway geometric variable influencing accident occurrences at toll road segments, was not found to be statistically significant to be included in the accident prediction model. This situation may happen as an occurrence of accidents in a particular segment may be caused not by the adverse geometric condition of the segment itself, but by the adverse geometric condition of the previous segment. Therefore, it is suggested to consider the influence of roadway geometric condition in previous segment when developing the accident prediction models in future research.

REFERENCES

Berhanu, G. (2004), Models Relating Traffic Safety with Road Environment and Traffic Flows on Arterial Roads in Addis Ababa, *Accident Analysis & Prevention*, Vol.36, No.5, pp. 697-704.

Bonneson, J.A. and McCoy, P.T. (1993), Estimation of Safety at Two-Way Stop-Controlled Intersections on Rural Highways, *Transportation Research Record*, Vol.1401, pp. 23-29.

Hauer, E. (1997), Observational Before-After Studies in Road Safety: Estimating the Effect of Highway and Traffic Engineering Measures on Road Safety, Pergamon.

Hiselius, L.W. (2004), "Estimating the Relationship between Accident Frequency and Homogeneous and Inhomogeneous Traffic Flows", Accident Analysis & Prevention, Vol.36, No.6, pp. 985-992.

Jovanis, P.P. and Chang, H. (1986), Modelling the Relationship of Accidents to Miles Traveled, *Transportation Research Record*, Vol.1068, pp. 42-51.

Kusumawati, A.and Wong, Y.D. (2010), The Application of Empirical Bayes Approach for Identifying and Ranking Hazardous Junctions, *Journal of the Eastern Asia Society for Transportation Studies*, Vol. 8.

Maycock, G. and Hall, R.D. (1984), Accidents at 4-arm Roundabout, *TRRL Laboratory Report*, Transport and Road Research Laboratory, UK.

Miaou, S.P. (1994), The Relationship between Truck Accidents and Geometric Design of Road Sections: Poisson versus Negative Binomial Regressions, *Accident Analysis & Prevention*, Vol.26, No.4, pp. 471-482.

Pickering, D., Hall, R.D., and Grimmer, R. (1986), Accidents at Rural T-junctions, *TRRL Research Report*, Transport and Road Research Laboratory, U.K.

Poch, M. and Mannering, F.L. (1996), Negative Binomial Analysis of Intersection-accident Frequencies, *ASCE Journal of Transportation Engineering*, Vol.122, No.2, pp. 105-113.

Rakhmat, L.A., Kusumawati, A., Frazilla, R.B., and Hendarto, S. (2013), The Development of Accident Prediction Model for Purbaleunyi Toll Road, *Jurnal Teknik Sipil*, Vol. 19, No. 3. (in Indonesian)

Republic of Indonesia - Ministry of Transportation (2011), National General Plan for Road Safety. (in Indonesian)

Sawalha, Z. (2002), Traffic Accident Modeling: Statistical Issues and Safety Applications, University of British Columbia, United States.

Shankar, V., Mannering, F, and Barfield, W.(1995), Effect of Roadway Geometrics And Environmental Factors on Rural Freeway Accident Frequencies, *Accident Analysis & Prevention*, Vol.27, No.3, pp. 371-389.

Turner, S. and Nicholson, A. (1998), Intersection Accident Estimation: The Role of Intersection Location and Non-Collision Flows, *Accident Analysis & Prevention*, Vol.30, No.4, pp. 505-517.

WHO (2009), Global Status Report on Road Safety, World Health Organization.