TRAFFIC ACCIDENT STUDY IN MIXED TRAFFIC CONDITION

Rizky JUNIRMAN[1] and Hiroyuki ONEYAMA[2]

[1]Doctoral Student, Dept. of Civil Eng., Tokyo Metropolitan University
(Minamiosawa 1-1 Hachioji-shi, Tokyo 192-0397, Japan)
E-mail: rizky-adelwinjunirman@ed.tmu.ac.jp
[2]Member of JSCE, Professor, Dept. of Civil Eng., Tokyo Metropolitan University
(Minamiosawa 1-1 Hachioji-shi, Tokyo 192-0397, Japan)
E-mail: oneyama.hiroyuki@tmu.ac.jp (Corresponding Author)

Improving road safety in low- and middle-income countries; local road characteristics must be understood. Among many local road characteristics, mixed traffic is infamous to be accident-prone, especially for vulnerable road users, but research on such issues is still rare. This study attempts to clarify how motorcycle composition contributes to accident rates in urban mixed traffic conditions in Indonesia; hence, other contributing road elements that represent such conditions is essential. Two accident models for motorcycle and motor vehicle analyzed using negative binomial regression, that suitable for count data and account overdispersion of the dependent variable. Negative binomial regression revealed that motorcycle and motor vehicles have similar contributing factors to their respective accident rate. However, not all the independent variables were statistically significant. This study finds mixed traffic with a large proportion of motorcycle contributes to accident rate, but does not necessarily warrant exclusively motorcycle lane.

Key Words : *Accident rate; motorcycle accident; negative-binomial regression.*

1. INTRODUCTION

The disproportion of road deaths between low- and middle-income countries (LMIC) and high-income countries (HIC) is tremendous, based on the data from 185 countries (reference). About 90% of road deaths occurred in LMIC, while only 10% in HIC (Burton, 2013). Actions to improve such conditions are often ineffective because they are often not based on local road traffic characteristics (Hills & Downing, 1980).

The local road traffic characteristics include road user behavior, traffic composition, types of motor vehicle, and road environment (Jacobs & Bardsley, 1977). Among the types of motor vehicle, motorcycles dominate the traffic stream in most LMIC, especially in Southeast Asia, which are a characteristic that easily observes and reasonably considers as hazardous situation. However, research on road safety on mixed traffic conditions that are dominated by motorcycles remains limited. Thus, this research attempts to explore the impacts of motorcycle composition to accidental rates in an urban road network environment with the mixed-traffic condition. Furthermore, the encouragement to study traffic volume composition and the accident rate stems from the following theory: "accident risk further is dependent on the composition of various road users" (Elvik et al., 2009).

As a case study, this research focuses on a set of an LMIC, particularly the city of Bandung, West Java, Indonesia. It attempts to clarify how motorcycle composition contributes to accident rates in urban road networks with mixed-traffic conditions. Hence, the use of other contributing variables to represent the typical mixed-traffic situation and urban road network environment. The urban road network consists of roads with different hierarchy. In this study, arterial and collector roads selected because of the varying degree of varying traffic compositions between different road users are more prevalent compared to roads with lower functions. The roads, then divided into a smaller unit of analysis concerning their road type.

The dependent and independent variables are allocated based on the pre-determined units of analysis of the type of multivariate analysis decided by considering the accident rate dispersion parameter of the dependent variable. Accident rates of motorcycles and motor vehicles selected for analysis. A motor vehicle is a four-wheeled or more vehicles includes car occupant vehicle, bus, and truck. The method of analysis is using multivariate analysis to understand and compare how the selected explanatory variables contribute to each of their accident risks (Lord and Mannering, 2010, pp. 291-305).

The result of this study aims to clarify the effect of high motorcycle composition in a mixed-traffic situation in an urban road network environment in Bandung, West Java, Indonesia. Furthermore, this new insight can be one of many considerations for the Road Authority to decide road safety countermeasures, for instance, if an exclusive motorcycle lane is necessary to improve safety.

2. LITERATURE REVIEW

Mixed traffic often thought to be dangerous for road users, especially vulnerable road users such as motorcyclists and bicyclists. Accident studies that consider mixed traffic conditions including vehicle types, road variations across and within countries, a different mode of travel, and heterogeneous traffic, are still deficient in the Southeast Asia region (World Health Organization (WHO), 2015). The term mixed traffic itself has been used loosely in different researches. However, in general, we can conclude that mixed traffic is a situation in the traffic stream when two or more road user types with different static and dynamic characteristics interact with each other at the same time and space. Furthermore, a discussion on mixed traffic conditions is a discussion on road user composition.

As mentioned before, it is challenging to improve road safety in LMIC without considering the local road characteristics. In this study, a mixed traffic condition that constitutes a high proportion of motorcycle composition compared to other road users leads to the decision to study such characteristics. **Table 1** provides traffic composition data in some LMIC.

As a literature review of this study goes, researches that studied the mixed traffic and the proportion of road users are as follows. First, a study conducted by Tjahjono (2009) in inter-urban roads that takes account of motorcycle proportion, but unfortunately, the mixed-traffic condition is not as prevalent as in an urban road. Second, a study that involved different volume compositions of road users. However, the focus of the study was pedestrians and bicyclists in road junctions without focusing on other road traffic conditions (Brüde and Larsson, 1993, pp. 499-509).

Table 1 Sample of traffic composition data

Vehicle Type	Traffic composition (unit)			
	Indonesia	India	Philip-pines	Vietnam
Cars and 4-wheeled light vehicles	8,148,330	15,313,000	2,770,591	556,945
Motorized 2 and 3-wheelers	60,152,752	82,402,000	3,482,149	31,452,503
Heavy trucks	3,296,315	6,041,000	347,182	552,244
Buses	1,095,554	1,486,000	34,933	97,468
Others	0	9,710,000	0	67,607
Total	72,692,951	114,952,000	6,634,855	33,166,411

Source: Global Status Report on Road Safety (2013)

The roads and traffic in urban environments are different from those in inter-urban environments. Traffic flow is often interrupted owing to the existence of junctions and road access. Roads located in a developed urban area often cannot be effectively utilized due to illegal or even legal use of the roadside, which is common in LMIC. Furthermore, roads in urban environments form a vast road network compared to inter-urban roads. It is crucial to discern such characteristics of the above situations when studying traffic composition and accident studies on urban roads.

Owing to frequent traffic flow interruption, the traffic state in an urban road environment can be expressed by moving and stopping states. The two-fluid model can depict both states (Prigogine & Herman, 1971). Various researches showed that the two-fluid model correlates with the main traffic elements of road, driver, and vehicle. Road network features and road user characteristics correlate with the two-fluid model (Herman and Ardekani, 2008, pp. 101-140). Furthermore, it correlates with road user characteristics (Williams, Mahmassani and Herman, 1995, pp. 121-140) and accident occurrences, respectively (Dixit et al., 2011, pp 1610-1616).

Roads in urban environments form a network that affects its operational characteristics and accessibility, consequently affecting road safety (Li and Wang, 2017, pp. 100-111). Centrality parameters of betweenness centrality and closeness centrality can represent the road network topological characteristics. Betweenness centrality describes how central a road is located among other roads in the network; hence, it captures pass by movements, whereas closeness centrality describes how other roads are near located to a road, representing the attraction and production of movements (Jayasinghe, 2017).

Most accident studies treated intersections and road segments as separate units of analysis. However, in an urban road network, which may have a high density of intersections located close to each other, such methods may not be appropriate (Li and Wang, 2017, pp. 100-111). The argument is that the interaction between a segment and an intersection cannot be adequately studied; hence, the proposal of a meso-level model that enables analysis of both segments and intersections as one single unit.

Typically, accident studies use count models to consider the nature of accident data with non-negative integer values, such as the Poisson model, negative binomial model, zero-inflated models, and logit models. In choosing among many models, an understanding of available data with the assumptions embedded in the models is essential. Considerations of choosing a model should involve analysis of its dispersion parameter and whether the data has a preponderance of zero values, along with the possibility of spatial or temporal correlation (Lord and Mannering, 2010, pp. 291-305).

As provided in **Table 1**, mixed traffic with a large proportion of motorcycles is a traffic phenomenon prevalent not only in Indonesia but also in some other LMIC, that consider as a dangerous situation, especially for motorcyclists, and the ideal solution is to separate them from other road users. However, quantitative research on such matters, particularly in the urban context, is deficient. Furthermore, lane separation of motorcyclists from other road users is not always applicable because of space or even budget constraints. This study aims to contribute to clarifying the relationship between a high proportion of motorcycle and safety scientifically.

3. DATA DESCRIPTION

In consideration of the technical and practical aspects, we chose Bandung as the study location. From a technical point of view, Bandung is the third-largest metropolitan city in Indonesia that has vast urban road networks while from a practical point of view, access to resources in obtaining secondary data, and conducting surveys. An arterial and a collector road in the road network ideally should have no or limited connection with the lower-class road; however, in reality, it does not. The road configuration in the road network has variations between divided and undivided roads of two-way and one-way roads.

There are several types of variables in this study. Land use variables typically exist in traffic accident studies excluded in this study due to the uniformity of land use type in the study location. Similarly, transit facility-related variables are excluded since such a facility is considerably scarce in the study location. In general, the source of data for such variables comes from accident data, traffic data, and road data collected with various methods. Most of these variables require estimation before being used in the analysis. In total, the study has 479 data observations, and the process of collating the data explained in the next section of the paper. A sample of the data is available in **Table 2**, and the explanation of the fields is available in **Table 3**.

Table 2 Sample of the data

ID	MAC	MVAC	Expo	MCAR	MVAR	MCP	MVP	TmCar	nCar	TmMc	nMc	RSDI	RT	AD	NACHr10000	NAINr10000	NACHr1000m	NAINr1000m	SI
1	1	4	1.404	0.712	2.848	52	48	1.525	1.148	1.76	0.752	0.042	Type3	4	1.375	0.849	1.391	1.145	TypeC
2	3	6	1.404	2.136	4.273	52	48	1.525	1.148	1.76	0.752	0.027	Type3	8	1.357	0.817	1.368	1.143	TypeA
3	0	5	1.404	0	3.56	52	48	1.525	1.148	1.76	0.752	0.008	Type3	7	1.37	0.852	1.34	1.128	TypeA
4	1	4	1.404	0.712	2.848	52	48	1.525	1.148	1.76	0.752	0	Type3	6	1.371	0.857	1.336	1.146	TypeA
5	6	12	1.404	4.273	8.545	52	48	1.525	1.148	1.76	0.752	0.07	Type3	3	1.372	0.864	1.335	1.173	TypeA
6	5	7	1.404	3.56	4.985	52	48	1.525	1.148	1.76	0.752	0.047	Type3	4	1.372	0.864	1.337	1.179	TypeA
7	5	6	1.404	3.56	4.273	52	48	1.525	1.148	1.76	0.752	0.044	Type3	5	1.371	0.871	1.377	1.258	TypeA
8	5	6	1.404	3.56	4.273	52	48	1.525	1.148	1.76	0.752	0.044	Type3	6	1.379	0.865	1.444	1.305	TypeA
9	5	8	1.404	3.56	5.697	52	48	1.525	1.148	1.76	0.752	0.041	Type3	4	1.369	0.811	1.449	1.371	TypeA
10	6	9	1.404	4.273	6.409	52	48	1.525	1.148	1.76	0.752	0.007	Type3	10	1.369	0.807	1.421	1.344	TypeA
11	6	8	1.404	4.273	5.697	52	48	1.201	3.618	1.21	10.46	0.051	Type3	7	1.369	0.815	1.395	1.338	TypeA
12	2	2	1.404	1.424	1.424	52	48	1.201	3.618	1.21	10.46	0.011	Type3	15	1.369	0.814	1.395	1.337	TypeA
13	2	4	1.404	1.424	2.848	52	48	1.201	3.618	1.21	10.46	0.029	Type3	10	1.372	0.823	1.358	1.335	TypeA
14	0	0	1.404	0	0	52	48	1.201	3.618	1.21	10.46	0.029	Type3	17	1.372	0.823	1.355	1.325	TypeA
15	5	5	1.404	5.003	5.003	52	48	1.201	3.618	1.21	10.46	0.036	Type3	14	1.378	0.865	1.318	1.401	TypeA
16	7	10	1.404	7.005	10.007	52	48	1.201	3.618	1.21	10.46	0	Type3	11	1.38	0.865	1.342	1.479	TypeA
17	1	3	0.999	1.001	3.002	62	38	1.201	3.618	1.21	10.46	0	Type3	0	1.304	0.778	1.196	1.115	TypeA
18	0	0	0.999	0	0	62	38	1.201	3.618	1.21	10.46	0	Type3	0	1.304	0.778	1.196	1.115	TypeA
19	3	4	0.999	3.002	4.003	62	38	1.201	3.618	1.21	10.46	0	Type3	0	1.307	0.782	1.137	0.862	TypeA
20	0	0	0.999	0	0	62	38	1.201	3.618	1.21	10.46	0	Type3	0	1.316	0.806	1.059	0.817	TypeA

Table 3 Fields in the data

Fields	Annotation
MAC	Motorcycle accident count
MVAC	Motor vehicle accident count
Exposure	The traffic exposure value of 100-million-kilometer travel
MCAR	Motorcycle accident rate
MVAR	Motor-vehicle accident rate
MCP	Motorcycle percentage
MVP	Motor-vehicle percentage
TmCar	Tm of two-fluid parameters for car occupants
nCar	n of two-fluid parameters for car occupants
TmMc	Tm of two-fluid parameters for motorcycle
nMc	n of two-fluid parameters for motorcycle
RSDI	Roadside disturbance index to represent disturbance to traffic from roadside
RT	Road type base on median and traffic flow directions
AD	Number of access in a unit of analysis
NACHr 10000m	Betweenness centrality for a radius of 10 km
NAINr 10000m	Closeness centrality for a radius of 10 km
NACHr 1000m	Betweenness centrality for a radius of 1 km
NAINr 1000m	Closeness centrality for a radius of 1 km

(1) Study location

Overall, 14 road segments vary on its lengths from the Bandung road network (e.g. **Figure 1**) used in the study, each of them divided into 100 m units of analysis using a meso-level approach that combines road segments and intersections (Li and Wang, 2017, pp. 100-111). Li and Wang's division method based on speed characteristics and geometry. However, in this research, the application of the same approach is not practicable owing to the limitation of total road length, which would result in a small unit of analysis that cannot be adequately analyzed.

The division into 100 m unit of analysis uses data collected from a mobile road survey vehicle belong to the Institute of Road Engineering (IRE) in Bandung. The vehicle has a GPS feature to tag location every 100 m while collecting visual data and geometric road data in a designated route. The location data from this vehicle was exported to GIS software as a reference location of the units and reference to input other data into each unit of analysis, resulting in 479 observations from this division.

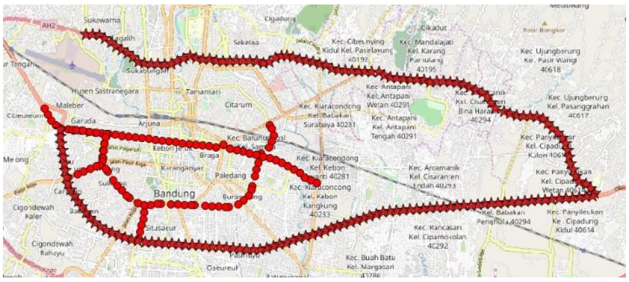

Figure 1 Study Location in Bandung, West Java, Indonesia

(2) Accident data

Accident data management in Indonesia is under the Indonesia Road Safety Management System (IRSMS). We obtained accident data through the Ministry of Public Works that has accessed to the system. To prevent the effect of regression to mean in the analysis, five year period of was accident data (2013-2018) used in this research.

The motorcycle accident count (MAC) and the motor vehicle accident count (MVAC) converted into accident rates with a unit of 100 million vehicle kilometer travel (VKT). The conversion requires additional data such as the length of the unit of analysis and traffic volume. Equation (1) described the accident rate equation (1):

$$Accident\ rate = \frac{Accident\ count\ x\ 10^8}{Segment\ length\ x\ period\ of\ observation\ x\ ADT\ x\ 365} \quad (1)$$

(3) Traffic data

This study uses traffic volume data from 25 observations across the selected study location to obtained traffic composition and calculates accident rate. The traffic composition is allocated into 479 units of analysis following the road name, where the data is collected. In the analysis, initially, the traffic composition is in motorcycle percentage (MCP) and motor vehicle for occupant percent (MVP). However, due to the high correlation, only MCP was used as one of the explanatory variables.

(4) Two-fluid parameters with GPS data

The logic of using the two-fluid parameters, which developed based on Bose and Einstein's theory, is due to its ability to represents the traffic as moving and stopping movement (Prigogine & Herman, 1971). Moving and stopping movement in urban road networks caused by road characteristics, road user maneuver characteristics, and unrestricted access to the main road. The parameters of the two-fluid model are Tm and n. Tm is the minimum free-flow travel time when there is no stopping, and n is a parameter describing network operational resistance to operational degradation with increasing demand. Both parameters are estimated using equation (2):

$$Tr = Tm^{1/n+1}\, T^{n/n+1} \quad (2)$$

where
Tr = running time
T = total travel time

The parameter estimation comes from the moving and stopping time of a probe survey vehicle. In this case, sampling is essential. Previous research has used different sampling strategies. For example, observations of 20-25 samples using one vehicle were sufficient (Dixit *et al.*, 2011), whereas others have found that sampling requirements for the two-fluid model use 10 to 20 units of vehicles and a survey period of 10 to 15 min (Williams, Mahmassani and Herman, 1995). In this research, a total of 64 observations of data collected using eight vehicles (four-unit of cars and motorcycles) from four days of observation. The method of data collection is the chase car method using GPS to collect speed and travel time. This method requires the selected survey vehicles to mimic random vehicles along with the designated study location. The two-fluid model estimation procedure consists of acquiring data from GPS devices, dividing the data into two kilometre micro-trips, identifying moving and stopping times and summarizing each of these times for each micro-trip, and estimation (Yeon, 2005, pp. 560-572).

In the analysis, only Tm of the car occupants and Tm of the motorcycle used in the model. The reason is that having a variable to represent the characteristics of motorcycle and motor vehicles of car occupants when they freely move in the traffic stream should affect the accident rate.

(5) Road data

The road segments in this study consist of road types with three different configurations concerning the median and traffic direction. Therefore, to capture the effect of these different road conditions on the accident rate, a categorical variable was proposed. Type 1 is undivided and one-way roads, while type 2 is undivided and two-way roads, and type 3 divided and two-way roads.

(6) Roadside disturbance (RSD) data

The formulation of RSD in this study is because there are many legal or illegal uses of the roadside (including the outer lane) that disturb the traffic flow in the road network. The disturbance causes stop and go movement and encourage dangerous overtaking maneuver. The RSD is somewhat new in traffic accident research. RSD calculation utilized data from the survey vehicle of the IRE; the video data contained visual conditions of the location and coordinates. Furthermore, the dimension of objects in the video to be calculated and the coordinate to be tagged. The length of the disturbance is calculated through the coordinate location. The formula for RSD is available in formula three:

$$RSDI = \frac{\sum_{i=1,2,\ldots}^{n} (li\ x\ wi)}{Ls\ x\ Ws} \tag{3}$$

where

li = length of object i
wi = width of object i
Ls = segment length
Ws = segment width

To ensure consistent identification of RSD, a simple rule of thumb is necessary. First, we calculated the length for each disturbance in the meso unit analysis. If the distance between disturbances is less than 1 m, they are considered as a single continuous disturbance. Second, we calculated the width of the disturbance from the edge of the traffic lane to the outer edge of the position of disturbance. Third, the area of disturbance is considered to be rectangular irrespective of the disturbance type of form, and different types of disturbance may be identified as a single disturbance if located near each other.

(7) Access density (AD) data

The location of the arterial and collector roads understudy should have limited direct access to lower class roads or property. However, owing to uncontrolled urban development, many roads and property access built directly connected to the road under study. AD is a variable that represents the number of access points to a property in a unit of analysis. The property in this variable can be a residential complex and business properties. The data collection method was observation through Google Maps and Google Street View.

(8) Signalized intersection (SI) data

The intersection is well known as part of the road that, owing to the possibility of conflicts between road users, consider have high accident risk, thus separately analyzed from a straight road segment. This study uses a meso-approach that defines the unit of analysis that may combine between the road segment and intersection. To account for the presence of signalized intersections in the analysis, and thus, the introduction of the SI variable. SI represents a signalized intersection and intersecting road in terms of the number of lanes. SI type A is straight segments of the road; SI type B is signalized intersections that the intersecting road has two lanes; finally, SI type C is signalized intersections that the intersecting road has three lanes or more.

(9) Signalized intersection (SI) data

The concept of NC allows us to understand the importance of road links in a road network. There are several measures of network centrality, among them closeness centrality and betweenness centrality. Closeness centrality measures how far a link is to all others links along the shortest path, whereas betweenness centrality refers to the extent to which a link belongs to the shortest paths between any pairs of nodes in a network; in other words, many shortest paths traversed that link (Jayasinghe, 2017). The formulas of closeness centrality and betweenness centrality presented in equations (4) and (5), respectively:

$$CC_{i[r]} = \frac{(N-1)}{\sum_{j \in N, j \neq i} d_{ij}} \tag{4}$$

$$BC_{i[r]} = \frac{1}{(N-1)(N-2)} \sum_{j,k \in N, j \neq k, k \neq i} \frac{p_{jk_{(i)}}}{p_{jk}} \tag{5}$$

where

CCi = closeness centrality of link i
N = the total number of links in a network.
d_{ij} = distance link i and link j
r = radius of influence system considered.
BCi = betweenness centrality of link "i"
pjk = number of geodesics between link "j" and "k"
$p_{jk(i)}$ = number of geodesics between link "j" and "k" that passing through i.

The calculation of centrality also accounts for the size of the influence area. However, in determining the suitable size of influence are for accident studies, no previous study has established a specific size (Li and Wang, 2017, pp. 110-111). A large radius was set as 10 km because it is the closest round number of distances to the maximum distance of the study location radio network that reached a radius of approximately 15 km. A small radius of 1 km is considered to be the smallest radius of travel distance when using a motor vehicle. The calculation of network centrality conducted using the Space Syntax Toolkit in QGIS (Gil, Varoudis, and Bartlett, 2014). The network centrality variable coded as NACHr 10 km (betweenness centrality 10 km), NAINr 10 km (closedness centrality 10 km), NACHr 1 km (betweenness centrality 1 km), and NAINr 1 km (closedness centrality 1 km).

4. METHOD

This research focuses on accidents for motorcycles and four-wheeled or more motor vehicles, thus using two different accident rates. Therefore, motorcycle and motor vehicles were separately analyzed using the same explanatory variables to compare how the formulated variables affect different road users. The reason is that the proportion of motorcycles is dominant (average of 69.897%); therefore, other road user proportions such as car occupants and heavy vehicles would have a small proportion of each of them were categorized separately.

At the beginning of the analysis, descriptive statistics and correlation analysis were conducted for all the dependent and independent variables. The result of descriptive statistics is available in **Table 4**. Both dependent variables of the motorcycle accident rate (MCAR) and motor vehicle accident rate (MVAR) had signs of overdispersion in which negative binomial regression was deemed suitable for this analysis. The mean and variance of MCAR are 5.317 and 41.965, respectively, whereas the mean and variance of MVAR 7.533 and 69.995, respectively.

All variables except for categorical variables were analyzed for correlation (result in **Table 5**). The correlation test was performed to evaluate the possibility of collinearity, gain insight into the correlation sign, and evaluate the strength of the linear relationship among variables. The correlation between MCP and MVP (percentage of the motor vehicle) is high, and the two variables cannot enter the model together. The inclusion is unfortunate since, initially, through the study, we aimed to determine how the different compositions of road users affect the accident rate. Also, there is a high correlation between TmCar and nCar (two-fluid parameters of car occupants), with a correlation coefficient above 0.5; thus, both variables need careful consideration when estimating the accident rate. In general, there is no other high correlation among the variables; correlation of accident rate and accident count variables with other variables is low. However, this does not indicate that there will be insignificant results.

Negative binomial regression is a count model, whereas the accident rate is non-integer. Hence, using accident rates directly should cause a warning from the R software. Thus, the exposure part of the accident rate input with the explanatory variables with an offset function. As a result, the mathematical model changed to equation (6). Both model development for motorcycle and motor vehicle conducted using backward elimination. Model development by backward elimination does not produce a conclusion error either a variable is significant and acceptance of the variable that explains the higher significant deviance when removed from the models (Tjahjono, 2009, pp. 1-13). When conducting backward elimination, the consistency of the coefficient sign of the independent variables in the model to the correlation analysis result is check throughout the process.

$$log\left(\frac{accident\ count}{exposure}\right) = intercept + a_1\ variable_1 + \cdots + a_n\ variable_n \quad (6)$$

Table 4 Descriptive Statistics

	MAC	MVAC	MCAR	MVAR	MCP	MVP
Minimum	0	0	0	0	52	18
1st quartile	1	2	1.468	2.001	65	27
Median	2	3	3.56	5.341	69	31
Mean	3.138	4.559	5.317	7.553	69.55	30.6
Variance	9.06	18.113	41.962	69.995	47.624	44.453
3rd quartile	5	6.5	6.784	9.781	73	35
Maximum	16	25	69.928	69.928	83	48
	TmCar	**nCar**	**TmMc**	**nMc**	**AD**	**RSDI**
Minimum	0.042	0.32	0.556	0.008	0	0
1st quartile	1.366	0.736	1.416	0.794	5	0
Median	1.644	1.728	1.494	1.258	8	0.012
Mean	1.699	3.021	1.539	2.847	8.374	0.027
Variance	0.233	49.978	0.085	18.256	23.825	0.003
3rd quartile	2.046	2.804	1.666	2.244	12	0.034
Maximum	2.617	50.37	2.108	23.329	25	0.578
	NACHr 10 km	**NAINr 10 km**	**NACHr 1 km**	**NAINr 1 km**		
Minimum	1.18	0.745	0.95	0.806		
1st quartile	1.372	0.902	1.274	1.144		
Median	1.413	0.979	1.316	1.274		
Mean	1.403	0.9837	1.303	1.259		
Variance	0.003	0.012	0.007	0.028		
3rd quartile	1.445	1.055	1.356	1.37		
Maximum	1.483	1.177	1.456	1.66		

Table 5 Variables correlation result

	MAC	MVAC	MCAR	MVAR	MCP	MVP	TmCar	nCar	TmMc	nMc	AD	RSDI	NACHr 10 km	NAINr 10 km	NACHr 1 km	NAINr 1 km
MAC	1															
MVAC	0.943	1														
MCAR	0.723	0.657	1													
MVAR	0.744	0.762	0.956	1												
MCP	0.043	-0.001	0.075	0.064	1											
MVP	-0.044	0.001	-0.086	-0.076	-0.991	1										
TmCar	0.149	0.095	0.150	0.120	0.330	-0.304	1									
nCar	-0.050	-0.038	-0.058	-0.053	0.016	-0.025	-0.529	1								
TmMc	0.105	0.104	0.052	0.058	0.130	-0.165	0.346	0.107	1							
nMc	-0.027	-0.030	-0.017	-0.023	-0.287	0.290	-0.248	0.167	-0.515	1						
AD	0.008	-0.042	0.055	0.017	0.345	-0.345	0.238	0.046	0.093	-0.275	1					
RSDI	0.030	0.021	-0.076	-0.082	0.143	-0.152	-0.106	-0.104	-0.068	-0.158	0.126	1				
NACHr 10 km	0.100	0.144	-0.059	-0.002	-0.098	0.081	-0.205	0.069	0.082	-0.224	-0.168	0.093	1			
NAINr 10 km	-0.137	-0.126	-0.092	-0.089	-0.142	0.142	-0.283	-0.184	-0.206	-0.334	-0.099	0.081	0.445	1		
NACHr 1 km	0.043	0.051	0.042	0.065	0.019	-0.032	-0.035	0.034	0.128	-0.417	0.145	0.127	0.311	0.241	1	
NAINr 1 km	-0.101	-0.089	0.016	0.020	-0.071	0.060	-0.162	-0.020	-0.049	-0.208	-0.064	0.031	0.132	0.335	0.577	1

5. RESULTS AND DISCUSSIONS

The multivariate analysis using a negative binomial regression result is available in **Table 6**. The result showed that not all proposed explanatory variables are significant. The significant variables were the same for motorcycle and motor vehicle models. They are MCP, NACHr 1 km, AD, SI Type B, SI Type C, and RT Type 3. Based on AIC and standard error, both models may not be the best model for prediction purposes. **Figure 2** provides a graph that showed the predicted accident rate vs. observed accident rate for both motorcycle and motor vehicle models, further explained the result. The plotted data points unbalanced in the y-axis shows a sign that the model is not perfect, which can be further explained based on the histogram frequency of accident rates (**Figure 3**). The accident rate most of the time is less than 40 VKT for motorcycle and 15 VKT for the motor vehicle, but on specific situation accident rate increases until 80 VKT and 30 VKT for motorcycle and motor vehicle, respectively. The goal if this study is to clarify the effect of motorcycle proportion and typical urban road characteristics to accident rate rather than for building model prediction. Hence, to that extent, we decided the result is still useful. Adding new roads is ideal for improving the model. However, due to the restriction of data availability and time, unfortunately, it is impossible to expand the studied road network.

Table 6 Analyses result

Variables	Motorcycle		Motor vehicle	
	Coefficient	P>z	Coefficient	P>z
Intercept	-1.289	0.151	-0.881	0.312
Motorcycle percentage (MCP)	0.014	0.043*	0.013	0.067·
Tm Car occupants	0.145	0.182	0.087	0.418
NACHr 1 km	1.911	0.008*	2.022	0.004**
NAINr 1 km	-0.277	0.442	-0.307	0.384
RSDI	-0.655	0.470	-0.934	0.295
AD	-0.023	0.033·	-0.029	0.005**
SI Type B	-1.146	0.004**	-1.082	0.003**
SI Type C	-0.408	0.076·	-0.465	0.038*
RT Type 2	-0.011	0.504	-0.011	0.954
RT Type 3	-0.475	0.010*	-0.319	0.078·
AIC	2197		2514.7	
Standard error	0.162		0.130	

Significant codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

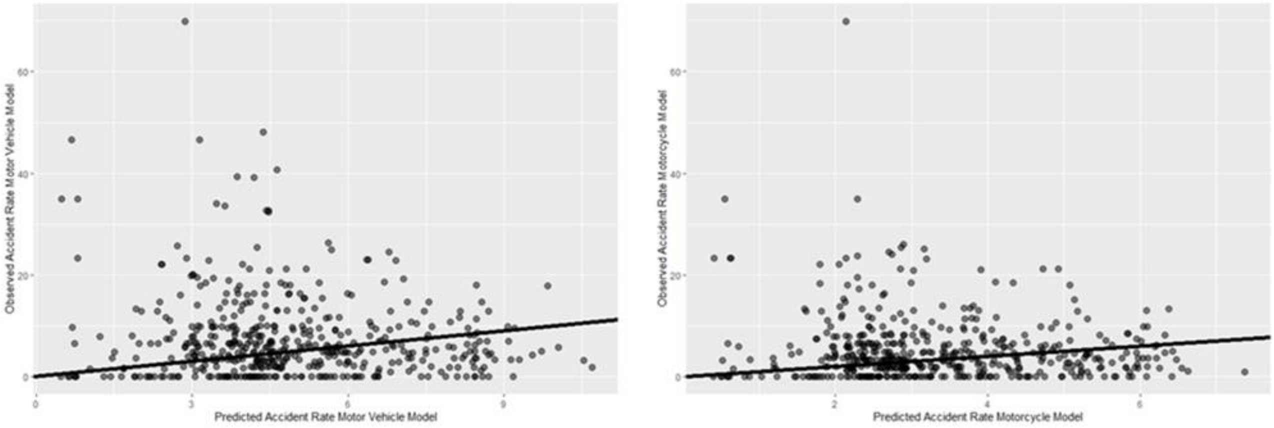

Figure 2 Predicted Vs. Observed Accident Rates

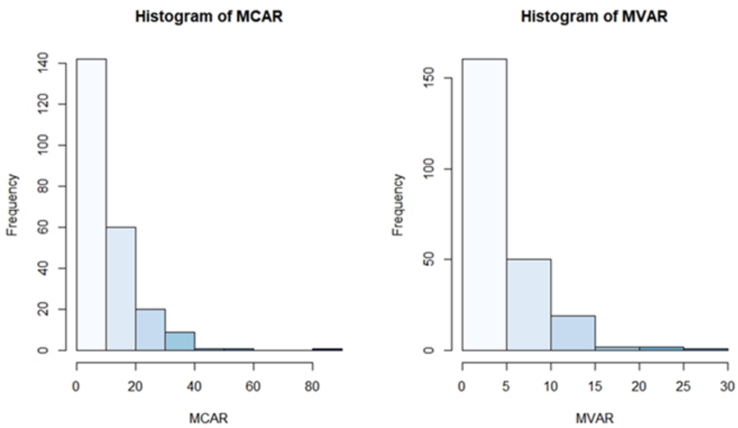

Figure 3 Histogram Frequency of Observed Accident Rates

The increasing motorcycle proportion contributes to the increase in accident rates for both motorcycles and motor vehicles. Further investigation on the motorcycle (%) and accident rates show that accident rate increases as motorcycle (%) increase, but when the motorcycle (%) is 73%, and above, there is a decrease in accident rate, as shown by **Figure 4**. Overall, this result supports the general opinion that a vulnerable road user where the motorcycle is one of them better have a dedicated lane. However, it does not necessarily oppose the opinion that both can share the same space and time under a particular condition, such as low speed and appropriate composition (Asian Development Bank, 2003). This study uses aggregated data where some detailed information may be lost, and traffic composition is dynamic and may change throughout the day. A micro-level analysis on the modeling scale or analysis using high-resolution data would provide more information on the effect of traffic composition.

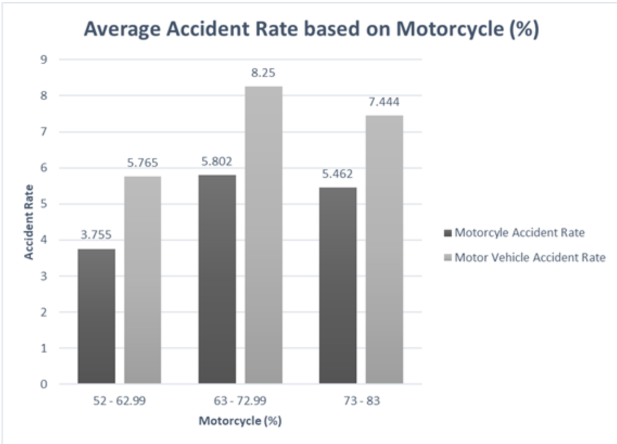

Figure 4 Histogram Frequency of Observed Accident Rates

TmMc and nMc (two-fluid parameter for motorcycle) were not included in the model because when they entered in the model, other related variables coefficient sign changes compared to the sign from the correlation analysis. The two-fluid parameter for both models for TmCar occupants was not significant. From previous research, Tm has a negative correlation with n and contributes to the increase in rear crash accident rates. A lower Tm correlates with aggressive driving behavior, which quickly deteriorates traffic flow, represented by higher n. As a result, the accident risk is higher (Dixit et al., 2011, pp. 1610-1616). However, the correlation analysis and model result showed in this research showed that Tm positively correlated with accident rates, which indicates that the longer time spent in the traffic stream with less aggressive driving of car contributes to a higher accident rate.

Further checking on the observed data showed that Tm increases as motorcycle proportion increases **(Figure 5)**; in a larger population of motorcycles, aggressive driving becomes problematic. In such a situation, dangerous movements or the slightest mistake would cause accident occurrence. In regards to the statistical significance, the two-fluid parameters are estimated for two km of micro-trips for each road segment by using probe data. Even though the number of observations is set based on previous studies, there is a possibility that the ratio of micro trips and road segment length, especially for short road segments, need additional observations.

Other than the TmCar occupants, RSDI is not significant for both models. The reason RSDI is not significant to accident rate because most of the recorded road disturbances are disturbances that already exist for an extended period; hence most road users already aware of such conditions and able to adapt to the situation accordingly. The road disturbance may disrupt the traffic flow but not necessarily contributes to the accident rate. **Figure 6** depicts a typical situation of roadside disturbance in the study location. AD is also significant for both motorcycle and motor vehicle models. A negative sign of AD is not as expected. An explanation for this is that road users drive more careful on roads that have high access to a property, which is mostly a commercial area.

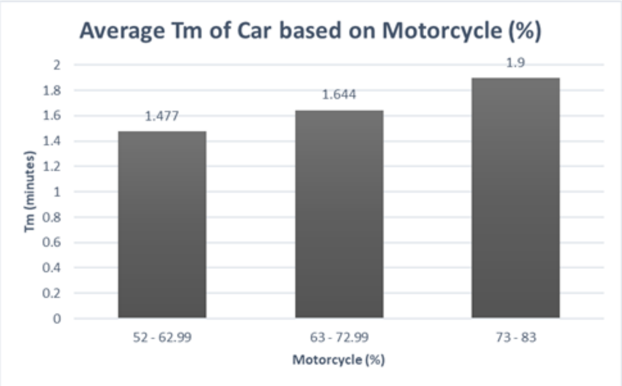

Figure 5 Tm and Motorcycle (%)

Figure 6 Typical of Roadside Disturbance

NACHr 10 km and NAINr 10 km not included in the model due to its interaction with other variables in the model cause change of variables coefficient sign, which makes comparing betweenness centrality and closeness centrality to accident rate difficult. NACHr 1 km significantly contributes to accident rate rather than in NAINr 1 km. The result indicates that through movement analyzed within an area of 1 km radii significantly affect the accident rate of motorcycle and motor vehicle, while local movement on 1 km radii is not significant. **Figure 7** showed that as the value of NACHr 1 km is getting higher, which indicates higher through traffic, the average accident rate also increases. The through traffic represented by NACH is often considered to be accident-prone owing to higher exposure and speed.

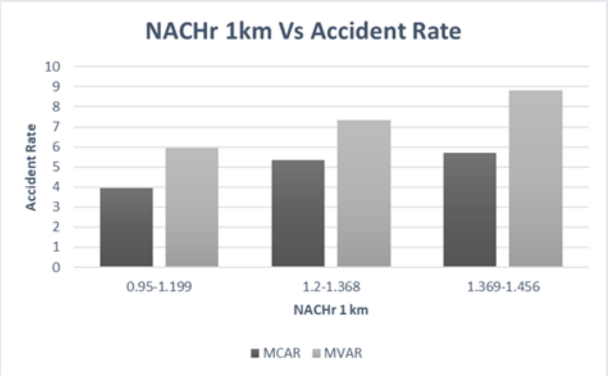

Figure 7 NACHr 1 km and Accident Rate

For both motorcycle and motor vehicle, SI type B and SI type C are significant and have a negative contribution to the accident rate. This finding was counterintuitive since it indicates that the intersection is safer than a straight section. However, the comparison was to signalized intersection instead of the non-signalized intersection. This result may show that having traffic control at an intersection contributes to safety. Also, the application of an advanced stopping lane for motorcycles in many signalized intersections in the study location improves safety.

Similarly, counterintuitive findings also exist for the RT variable, which is significant for the motorcycle and motor vehicle model. Roads with a median to separate different traffic streams are inherently safer. In this case, it is significant for motorcycles since it determines dangerous maneuvers such as overtaking using the opposing traffic lane, but is it safer than a one-way road? The answer should depend on the accident characteristics of the study location because the one- and two-way roads have different accident characteristics.

6. CONCLUSIONS

In general, the results show that motorcycle and motor vehicles have the same significant contributing factors to their respective accident rates. The main findings from this study are that motorcycle proportions affect the accident rate for both motorcycles and motor vehicles. The increasing proportion of motorcycles not only affects the accident rate for accidents between the same type of road users and different types of road users. Furthermore, a study using smaller aggregate data or micro-level analysis with high-resolution data should provide more information on understanding the effect of traffic composition on the accident rate in mixed traffic conditions.

Aggressive driving is represented by Tm, which is prevalent in Indonesia and thought to have an impact on accident rate; in this case, it was found to be insignificant and has a different coefficient sign from previous research. Whether this result is due to the need for more samples or indicates that aggressiveness is not a significant contributor is something that requires further exploration.

Through the betweenness centrality, it confirmed that in general, through traffic contributes to accident rates. The coefficient of NACHr 1 km for both motorcycle and motor vehicle models is significant than NAINr 1 km, which indicates that exposure from through traffic travel is more dangerous for road users, mainly for mixed traffic conditions.

Regarding road features, in this case, signalized intersection and road type have a counterintuitive result. The study found that a signalized intersection has a more significant impact on the accident rate for both motorcycle and motor vehicle compared to a straight road segment. The effect of proper signal control and installation of an advanced stopping lane for motorcycles may have some contribution.

The existing road type appears to provide some degree of safer road infrastructure to road users, especially on divided roads. However, it is interesting to explore further this situation concerning accident type.

ACKNOWLEDGMENT: We wish to thank colleagues at IRE, Department of Highway, Maranatha Christian University, Parahyangan Christian University, Bandung Department of Transport, and Indonesian National Police for sharing the data. We hope that the results of this research will clarify traffic composition and mixed traffic conditions.

REFERENCES

1. Burton, A. (2013) 'Global status on road report 2015', *World Health Organization*, 19(3 ed), p. 150. doi: 10.1136/injuryprev-2013-040775.
2. Jacobs, G. D. & Bardsley, M. N., 1977. Road Accidents As a Cause of Death in Developing Countries, Crowthorne: TRRL Supplementary Report 277.
3. Elvik, R. (Institute of T. E. *et al.* (2009) *The Handbook of Road Safety Measures*. Second Edi. Bingley: Emerald.
4. Lord, D. and Mannering, F. (2010) 'The statistical analysis of crash-frequency data: A review and assessment of methodological alternatives', *Transportation Research Part A: Policy and Practice*. Elsevier Ltd, 44(5), pp. 291–305. doi: 10.1016/j.tra.2010.02.001.
5. World Health Organization (WHO) (2015) 'Global Status Report on Road safety in the South-East Asia', pp. 1–28. doi: WA 275.
6. Tjahjono, T. (2009) 'The Effect of Traffic and Road Conditions to the Fatality Rates on Rural Roads in Eastern Indonesia', *Proceedings of the Eastern Asia Society for Transportation Studies*, 7, pp. 1–13.
7. Brüde, U. and Larsson, J. (1993) 'Models for predicting accidents at junctions where pedestrians and cyclists are involved. How well do they fit?', *Accident Analysis and Prevention*, 25(5), pp. 499–509. doi: 10.1016/0001-4575(93)90001-D.
8. Prigogine, I. & Herman, J. C., 1971. Kinetic Theory of Vehicular Traffic. s.l.:Elsevier.
9. Herman, R. and Ardekani, S. (2008) 'Characterizing Traffic Conditions in Urban Areas', *Transportation Science*, pp. 101–140. doi: 10.1287/trsc.18.2.101.
10. Williams, J. C., Mahmassani, H. S. and Herman, R. (1995) 'Sampling strategies for two-fluid model parameter estimation in urban networks', *Transportation Research Part A*, 29(3), pp. 229–244. doi: 10.1016/0965-8564(94)00031-5.
11. Dixit, V. V. *et al.* (2011) 'Quality of traffic flow on urban arterial streets and its relationship with safety', *Accident Analysis and Prevention*. Elsevier Ltd, 43(5), pp. 1610–1616. doi: 10.1016/j.aap.2011.01.006.
12. Li, J. and Wang, X. (2017) 'Safety analysis of urban arterials at the meso level', *Accident Analysis and Prevention*. Elsevier, 108(September), pp. 100–111. doi: 10.1016/j.aap.2017.08.023.
13. Yeon, S. (2005) 'Analysis of Two-Fluid Model Using Gps Data', 6, pp. 560–572.
14. Jayasinghe, A. B. (2017) *A network centrality-based simulation approach to model traffic volume*. Nagaoka University of Technology.
15. Gil, J., Varoudis, T. and Bartlett, T. (2014) 'Space Syntax Toolkit: Extending QGIS with exploratory spatial analysis for'.
16. Asian Development Bank (2003) 'Vulnerable Road Users in the Asian and Pacific Region'. Available at: https://think-asia.org/bitstream/handle/11540/3007/vulnerable-road-users.pdf?sequence=1.

(Received October 3, 2020)
(Accepted October 3, 2020)