A conceptual framework for development of safety performance functions for rural roads in Indonesia

*W B Dermawan1 and T Tjahjono1

1 Department of Civil Engineering, University of Indonesia, Kampus Baru UI Depok
Jawa Barat – 16424, Indonesia
*Corresponding author: widodo.budi@ui.ac.id

Abstract. Road safety policies provide important guidance for actions to enhance road safety, but often absence an extensive conceptual framework, system theory, and evidence-based study and practice. One effort to enhance road safety is to decrease the number of road accidents. By knowing the accident data at a location, transportation planners can conduct analysis so that they can provide recommendations for improving road safety. However, the prediction model in Indonesia in some studies is more local purposes. Meanwhile, the prediction model used in various countries needs to be adapted to the conditions of Indonesia. The objective of this paper is to integrate all components that contribute to accidents. The work identifies tools that are most and least widely used for components. This framework also provides guidance for developing safety performance functions, what the model is like, what the variables are and what conditions need to be considered according to Indonesian conditions. It is expected that the developed framework can be applied to improve the development of safety performance functions for rural roads in Indonesia.

1. Introduction

Indonesia is the fourth most populous nation in the world with more than 260 million populations. The country's economic growth has led to an rise in demand for transport and car ownership. Road accidents in Indonesia are arising on a growing trend in the previous decade. Out of 105,374 cases of traffic accidents recorded in 2016, 25,859 people died, 22,939 people severely injured and 120,913 people minorly injured [1]. There are 135,883 motorcyclists that are involved in a traffic accident. The motorcycle is the most popular mode of transportation in Indonesia. Compared to other transportation modes, the number of serious motorcycle crashes is relatively high [2]. Fatality risk among motorcyclists is about 30 times higher and serious injuries are about 8 times higher compared to other groups of drivers [3].

Road accidents are often caused by several factors such as human factors, vehicle factors, and external factors (including road conditions). Road damage is a factor that can prevent accidents if the damage is handled properly. According to data from the Ministry of Public Works (PUPR) in 2017, the length of national roads reached 47,017 km with good conditions 51.76%, moderate 38.59%, lightly damaged 5.76% and severely damaged 3.89%. Moreover, compared to urban roads, rural roads are more prone to accidents [4]. Rural roads are usually characterized by low traffic volumes, generally high speeds, fewer conflict turns, and fewer conflicts with pedestrians. Accidents that occur on rural roads generally occur at high speeds and if an accident occurs, the emergency response will be longer due to the distance to the nearest hospital [5]. Road accidents have always been the subject of extensive research in the world. Traffic accidents have been identified as a very important factor in identifying
discrepancies in traffic management and the entire transportation system [6]. Studies show the identification of factors affecting road crashes obtained from the crash investigation can be used to prevent accidents on a road section in a certain condition in the future. Historically different types of statistical methods have been used to model road accidents. The early models developed were based on Multiple Linear Regression while the models currently used are based on the Generalized Linear Model (GLM) technique.

Today, the Highway Safety Manual (HSM) Prediction Model [7] is the most widely used to predict road accidents. Furthermore, previous researches suggest that the use of HSM outside the United States and Canada should be performed with caution. In this case, Indonesia's number of motorcycle users should be properly considered as motorcycle crashes account for a very high proportion of the total vehicle fatalities. Another concern that needs to be considered is the site condition before the implementation of treatment, also known as the base conditions. A specified set of Indonesian base conditions will be used to determine the values of Crash Modification Factors (CMF). The CMFs are an important factor of the HSM safety management process used to estimate the change in the expected number of accidents at the site where a specific counter-measure is being implemented.

The purpose of this study is to develop transferable safety performance functions (SPFs) to estimate the amount of road accidents that can be used to estimate the amount of road accidents in rural national road sections by looking at the characteristics of road conditions in Indonesia using the Generalized Linear Model (GLM) method and the HSM model.

2. Literature review

Early accident prediction models were generated from simple multiple linear regression techniques. [8-9]. This model is based on the assumption that errors are normally distributed with zero mean and constant variance. Because the number of accidents cannot be negative and errors from accident rates have a variable value that is not constant, Jovanis and Chang [10] state that the linear regression model has several disadvantages. Maycock and Hall [11] concluded that a linear regression model is not suitable for predicting accident rates since the accident rate is discrete, so it does not follow a normal distribution. The other reason is the variance in the number of accidents is not constant but tends to increase when the traffic flow increases.

To overcome the limitations in the linear regression model, Miaou [12], recommend Poisson regression models in analyzing accident rates. The Poisson regression model is quite interesting to use because it only has one parameter to estimate. However, despite widespread use, the Poisson regression model has one significant constraint, which is the chance that standard errors, which are estimated by the maximum likelihood technique, may be biased if the mean value is not equal to the variance. If this occurs, the test statistics obtained from the model will be inaccurate. This can lead to coefficient bias and standard error of model in the case of crash data that are usually over-dispersed with variance exceeding the mean value.

To overcome the issue of over-dispersion in crash data counts, the Negative Binomial (NB) distribution models have been proposed [11]. NB distribution is more appropriate to describe discrete and non-negative events such as accident rate data. The NB distribution has advantages over the Poisson distribution in that the mean value does not need to be the same as the variance value. The suitability of using the NB regression model compared to the Poisson regression model is determined from the significance of the \( \alpha \) parameter, which is a parameter associated with the degree of over-dispersion. If \( \alpha \) is significantly greater than zero (as calculated using t-statistics), then the NB regression model is more suitable to use than the Poisson regression model [13]. The Poisson regression model can be considered a limited case of the NB regression model that is if \( \alpha \) is close to zero. However, when the sample size is very small, the NB models have some constraints when the mean of the crash counts is greater than the variance to manage the under-dispersion of crash counts. As a result, scientists have begun to apply Poisson-log normal models that are deemed more suitable to solve the under-dispersed problem of crash counts.

Another commonly used form of accident prediction model is the zero-inflated Poisson and zero-inflated negative binomial models that were primarily implemented to address the issue of over-dispersion induced by excessive zeroes (i.e. locations where no accidents can be observed) in traffic data.
counts [14]. This condition can lead to inaccurate conclusions that entities that have zero-accident data are safe entities, whereas in reality only a portion of these entities are truly safe (in the absence of an accident) while other entities are actually not safe, but happens to have zero accidents during the observation period [15-16].

In addition, another approach to estimate road accidents is the use of Bayesian methods, consisting of an empirical Bayes (EB) method and a hierarchical Bayes (HB) method. The name Bayes theorem is taken from the name of the inventor of the theorem, namely an Englishman named Reverend Thomas Bayes (1702-1761). The empirical Bayes (EB) method combines the local accident history and expected number of crashes obtained from the prediction model for similar sites [17]. This technique increases estimation accuracy and corrects for regression-to-mean bias, and is therefore recommended for use in road safety studies.

3. Method

3.1. A meta-analysis of Methodological Method
Meta-analysis may be defined as a research study. This refers to a statistical study of a large set of results from particular research in order to integrate the findings [18]. This method links the outcome from several studies to make a general conclusion. It was initially applied to studies conducted in medicine and psychology, and eventually it was expanded to other natural sciences including transportation.

Two main methods have been used to identify the relevant studies. First, all studies published were identified by searching in Science Direct and the Transportation Research Board online library of publications. The search term were “road safety”, “crash prediction model”, “safety performance function “, “GLM” and “highway safety manual”. Second, to classify additional studies, the reference segments of available studies were reviewed. These articles are then reviewed whether they met the criteria.

3.2. A Conceptual Framework of Prediction Model
A conceptual framework is defined as a network of linked concepts [19]. The conceptual framework for the development of crash prediction model on national rural roads provides a detailed overview of the factors that contribute to accidents. It also offers the opportunity to consider all potential methods that can be extended to any of the relevant variables in the term of decreasing the number of accidents as part of the road safety strategies. These strategies had been introduced by Haddon [20] such as a component of the initial logical framework; vehicle, human, infrastructure, equipment, physical and socio-economic environment. Models can define crashes as a sequence of events over time or a process, such as the Haddon's framework; pre-crash, crash and post-crash phases. The proposed framework for accident prediction models could be implemented by road safety professionals and ultimately tested for use. It could also be used to analyze past strategies to identify which techniques and sub-components were not previously applied. The framework could be used by researchers as a guide for defining other variables. Hughes [21] described all the prospective contributors to the crash as shown in Table 1.

Tjahjono [22] provides information that the factors that cause accidents written in police reports tend to emphasize errors in human factors or it can be said that human factors are the main factors in the majority of accidents. However, it is not easy to study the influence of human factors in the event of an accident considering that human characteristics and behavior vary naturally so that accident management strategies are generally directed at applying technical corrective actions to road infrastructure to decrease the possibility of drivers making mistakes [23]. This study emphasis more on the road infrastructure factors that are suitable for conditions in Indonesia as described in Table 2. No every variable in the road infrastructure components become explanatory variables in the model. The selection of explanatory variables should be driven by the earlier documented evidence of crash and injury risk factors available from research literature or by the availability of data.
Table 1. Road Safety Components

| Components                        | References                                      |
|----------------------------------|------------------------------------------------|
| Transport and Land Use Context   | Elvik [24]                                     |
| Economic Context                 | Chen [25], Rangel [26]                         |
| Social Context                   | Backe S [27]                                   |
| Natural Environment              | Road Research Laboratory [28], Limpert [29],   |
|                                  | Baker and Fricke [30], Elvik [24]              |
| Human                            | Road Research Laboratory [28], Limpert [29],   |
|                                  | Baker and Fricke [30]                          |
| Vehicles                         | Road Research Laboratory [28], Limpert [29],   |
|                                  | Baker and Fricke [30], Elvik [24]              |
| Road Infrastructure              | Road Research Laboratory [28], Limpert [29],   |
|                                  | Baker and Fricke [30]                          |
| Crash Response                   | Tournier [31]                                  |
| Traffic Volume                   | Baker and Fricke [30]                          |

Source: Hughes [21]

Table 2. Road Infrastructure Components

| Sub Components | Variables                                      |
|----------------|-----------------------------------------------|
| Surface        | Roughness, friction cracks                    |
| Geometry       | Alignment geometry, crest, curve, gradient,   |
|                | lanes, passing lane, median, shoulder, physical dimensions, dual carriageway, |
| Safety devices | Fence, guardrail, barrier, landslide protection, escape ramp |
| Signs          | Pavement marking, delineator, signal, advisory |
| Lighting       | Roadway, reflectors                           |
| Road type      | Highway, urban road, rural road               |
| Miscellaneous  | Work zones, stopped buses, parked cars, road debris, landslides, crossing railway |
| Traffic        | Composition (motorcycle, heavy vehicle)       |

Source: Hughes [21]

4. Data collection

In this study, the data is obtained from the Eastern Indonesia National Roads Improvement Project (EINRIP) which collected the data for 8 years (2008-2016) within identical locations in 8 provinces of Indonesia namely: Bali, NTB (Sumbawa), NTT (Flores), Central Sulawesi, South Sulawesi, Southeast Sulawesi, West Kalimantan and South Kalimantan. The objective of the Eastern Indonesian National Road Improvement Project (EINRIP) is to support economic and social development in Eastern Indonesia through a program of improvements to approximately 400 km of national roads in the region. The project involved reconstruction and resurfacing, widening and minor realignments, which were expected to reduce vehicle operating costs and journey times. The improved quality of construction was expected to result both in better quality roads and reduced maintenance costs. Hence, from this project, the data of traffic, speeds, road condition, and maintenance activity can be obtained.

The accident data collection is supported with the established Integrated Road Safety Management System (IRSMS) developed by Indonesian National Traffic Police Corps (Korlantas Polri) in collaboration with funded by the World Bank under the Strategic Road Infrastructure Project (SRIP). It aims to acquire specific and accurate data of traffic accidents throughout Indonesia with a web-based program. Thus, report of traffic accident can be directly received through IRSMS programmed tablet.
device. The system included a facility for validation of accident information. Only the validated data are included in statistics produced by the system. The system also includes a wide variety of output possibilities such as Google Maps or Open Street Map. However, in this study, if accident data is inadequate, it can be obtained from the local police station.

5. Discussion and conclusion
Crash prediction models had been widely used in identifying and evaluating road safety problems. Internationally recognized, countries all over the world are advancing their crash prediction models of GLM, using the HSM method by exploring the modification factors (CMF) and calibration factors. However, most of the crash prediction models developed by these researchers are limited to one area within a period of time. Whereas a country like Indonesia, which still faces challenges to develop its appropriate crash prediction models due to difficulties acquiring accurate crash data from poor police records and reports needs a transferable model. This will save time, effort and cost if the estimated accident model developed for one area in a certain time period can be transferred to be used in a different time and region periods (transferability model). Therefore, in this study, the crash prediction model will be developed with a safety performance function-HSM approach with modification factors and calibration factors.

In addition, the more variables included in the model, the better its predictive features will perform. However, variable numbers depend on data availability and access (data collection). Such in previous studies of developed countries, where variables or factors of the accident had not been included due to some major differences with Indonesia such as the composition of vehicles on the road and road surface conditions. The research in Indonesia is mostly focused on urban roads, highways/toll roads or locally based on one particular type of vehicle. Therefore, in this study, the traffic accident prediction model will be developed by looking at new variables and those that are in accordance with the conditions in Indonesia. The framework is expected to lead the comprehensive research approach in the development of safety performance functions for rural roads in Indonesia.

6. References
[1] Badan Pusat Statistik Indonesia 2017. Statistik Transportasi Darat 2016 (BPS-Statistics Indonesia)
[2] Haworth N, Greig K and Nielson A 2009 Comparison of risk taking in moped and motorcycle crashes Transportation research record 2140(1) 182-187
[3] Huang B, and Preston J 2004 A literature review on motorcycle collisions: Final report. Transport Studies Unit (Oxford University)
[4] Zhong C, Sisiopiku V, Ksaibati K and Zhong T 2011 Crash prediction on rural roads. In Proceeding of the 3rd International Conference on Road Safety and Simulation, Indianapolis, Indiana
[5] The Road Information Program (TRIP) 2005 “Growing Traffic in Rural America: Safety, Mobility and Economic Challenges in America’s Heartland”. The Road Information Program, Washington, DC
[6] Soehodho S 2017 Public Transportation Development and Traffic Accident Prevention in Indonesia JATSS Research 40 (2) 76-80
[7] American Association of State Highway and Transportation Officials 2010. Highway safety manual vol 1, Washington, DC
[8] Solomon D 1964 Accidents on main rural highways related to speed, driver, and vehicle Report (Columbus, GPO/Bureau of Public Roads)
[9] Lau M Y K and May A D 1988 Injury Accident Prediction Models for Signalized Intersections Transportation Research Record 1172 58-67
[10] Jovanis P P and Chang H 1986 Modelling the Relationship of Accidents to Miles Traveled Transportation Research Record 1068 42-51
[11] Maycock G and Hall R 1984 Accidents at 4-Arm Roundabouts. TRRL Laboratory Report 1120 (Crowthorne U.K: Transportation and Road Research Laboratory)
[12] Miaou S P 1994 The Relationship Between Truck Accidents and Geometric Design of Road Sections: Poisson Versus Negative Binomial Regressions *Accident Analysis and Prevention* 26(4) 471-482

[13] Poch M and Manering F 1996 Negative binomial analysis of intersection-accident frequencies *Journal of transportation engineering* 122(2) 105-113

[14] Miaou S P and Lum H 1993 Modeling vehicle accidents and highway geometric design relationships *Accident Analysis & Prevention* 25(6) 689-709

[15] Shankar V, Milton J and Manering F 1997 Modeling accident frequency as zero-altered probability processes: an empirical inquiry. *Accident Analysis and Prevention* 29(6) pp 829–837

[16] Kumara S and Chin H 2003 Modeling accident occurrence at signalized tee intersections with special emphasis on excess zeros *Traffic Injury Prevention* 3(4) 53–57

[17] Hauer E, Harwood D W, Council F M and Griffith M S 2002 Estimating safety by the empirical Bayes method: a tutorial *Transportation Research Record* 1784(1) 126-131

[18] Glass G V, McGaw B and Smith M L 1981 *Meta-Analysis in Social Research. Beverly Hills*, (London : Sage).

[19] Schepers P, Hagenzieker M, Methorst R, Van Wee B and Wegman F 2014 A conceptual framework for road safety and mobility applied to cycling safety *Accident Analysis & Prevention* 62 331-340

[20] Haddon Jr W 1972 A logical framework for categorizing highway safety phenomena and activity. *J. of Trauma and Acute Care Surgery* 12(3) 193-207

[21] Hughes B P, Anund A and Falkmer T 2016 A comprehensive conceptual framework for road safety strategies *Accident Analysis & Prevention* 90 13-28

[22] Tjahjono T, Iqbal M and Marino A 2012 *Pemanfaatan Highway Safety Manual (HSM) di Indonesia dalam memprediksi kecelakaan pada segmen jalan. Jurnal Transportasi* 1299(3) 175-184

[23] Kusumawati A, Upahita D P and Hendarto S 2013 Development of Accident Prediction Model for Rural Toll Road Sections. In *Proceedings of the Eastern Asia Society for Transportation Studies* 9

[24] Elvik R 2008 The predictive validity of empirical Bayes estimates of road safety *Accident Analysis and Prevention* 40 1964–1969

[25] Chen G 2005 Safety and economic impacts of photo radar program. *Traffic injury prevention* vol 6(4) 299-307

[26] Rangel T, Vassallo J M and Herraiz I 2013 The influence of economic incentives linked to road safety indicators on accidents: the case of toll concessions in Spain *Accident Analysis and Prevention* 59 529–536

[27] Backe S, Ericson L, Janson S and Timpka T 2009 Rock climbing injury rates and associated risk factors in a general climbing population *Scandinavian journal of medicine & science in sports* 19(6) 850-856

[28] Road Research Laboratory 1963 *Research on Road Safety* (London: H.M.S.O).

[29] Limpert R 1984 *Motor vehicle accident reconstruction and cause analysis* (Charlottesville: Michie Co)

[30] Baker J S and Fricke L B 1986 *The Traffic-accident Investigation Manual: At-scene Investigation and Technical Follow-up* (Evanstone III: Northwestern University Traffic Institute Evanston).

[31] Tournier C, Charnay P, Tardy H and Chosssegros L 2014. A few seconds to have an accident, a long time to recover: consequences for road accident victims from the ESPARR cohort 2 years after the accident *Accident Analysis and Prevention* 72 422–432