PREDICTING CASUALTY-ACCIDENT COUNT BY HIGHWAY DESIGN STANDARDS COMPLIANCE

Arief Rizaldi
Ministry of Public Works of Indonesia
Jln. Pattimura No.20, Kebayoran Baru,
Jakarta Selatan, 12110 Indonesia
Tlp. (+62 21) 7246654 Fax. (+62 21) 7246654
arief_rizaldi@yahoo.co.uk

Vinayak Dixit
University of New South Wales
UNSW Sydney NSW 2052 Australia
Tlp. (+61 2) 9385 5381 Fax.(+61 2) 9385 6139
v.dixit@unsw.edu.au

Abstract

Compliance to standard has been the main doctrine in highway design, but its relationship with accident count has not been widely scrutinized. One of the key programs in road safety in Indonesia is road-worthiness test which assesses the compliance of a road to national design specifications and criteria. In light of current improvements in the crash data system in Indonesia, this study is carried out to develop a model to predict the accident count per type of crashes and to identify significant road features based on their compliance to a national standard. About 272,200 kilometers of arterial road in East Java North Corridor is selected as case study and 2012 to 2014 crash data is analyzed. The Zero-Inflated Negative Binomial Model is preferred to develop crash prediction model with significant variables. This study has several findings. First, the number of median opening per unit length and disturbance level to pedestrian and road reserve area are the features that having positiverelationships with total accident count. Meanwhile, the ROW disturbance, conformance of intersection and of road marking also show significant value but negative relationship with total accident count. Second, significant variables for each type of crash may have different sign. For example, in right angle crash, median width has positive relation with the number of accident, while in run off and rear end crash, median width compliance is shown to have negative relation.

Keywords: highway standard, accident prediction, arterial highway

Abstrak

Kepatuhan terhadap standar merupakan doktrin utama dalam desain jalan, tapi hubungannya dengan jumlah kecelakaan belum banyak diteliti. Salah satu program utama dalam keselamatan jalan di Indonesia adalah pengujian laik fungsi jalan, yang dinilai dari pemenuhan jalan terhadap kiriteria dan spesifikasi desain nasional. Sejalan dengan perbaikan sistem pendataan kecelakaan di jalan di Indonesia saat ini, penelitian ini dilakukan untuk mengembangkan model untuk memprediksi jumlah kecelakaan per jenis kecelakaan dan mengidentifikasi fitur jalan yang signifikan berdasarkan pemenuhan terhadap standar nasional. Sekitar 272.200 kilometer jalan arteri di Koridor Utara Jawa Timur dipilih sebagai studi kasus dan data kecelakaan tahun 2012-2014 dianalisis. Model Binomial Negatif Zero-Inflated dipilih untuk mengembangkan model prediksi kecelakaan dengan variabel-variabel yang signifikan. Penelitian ini menghasilkan beberapa temuan. Pertama, jumlah bukaan median per satuan panjang dan tingkat gangguan untuk pejalan kaki dan ruang milik jalan merupakan fitur yang memiliki hubungan positif dengan jumlah total kecelakaan. Sementara itu, gangguan pada ruang milik jalan ROW, kesesuaian persimpangan dan marka jalan juga signifikan namun mempunyai hubungan negatif dengan jumlah total kecelakaan. Kedua, variabel yang signifikan untuk setiap jenis kecelakaan mungkin memiliki tanda yang berbeda. Misalnya, pada kecelakaan dari samping (tegak lurus), lebar median memiliki hubungan positif dengan jumlah kecelakaan, sementara kecelakaan di bagian jalan yang lurus, pemenuhan lebar median memiliki hubungan negatif.

Kata-kata kunci: standar jalan, prediksi kecelakaan, jalan arteri

INTRODUCTION

Many road designers believe that when a road is built to meet standards, it should be safe. Thus, adherence to standards has been the main doctrine in highway design practice. However, the level of safety cannot be measured only by the standards. Many highway standards are limited in this sense. They do not govern the safest design, rather, they control the permissible limit of safety for each element of a road.

The relationship between highway standards and the real measure of safety (i.e., accident rate and accident severity) is conjectural (Hauer, 2000). Standards are developed to avert failures. On the other hand, the concept of failure in highway design is defined by surrogates, such as a shortfall in lane width, or deficient centripetal force. Accordingly, a road built only to standards is built without a premeditated level of safety. Yet, there has not been much research conducted to understand the relationship between the governed standard with safety.

While a reliable crash data was unavailable, Bina Marga, Indonesian Highway Authority, was under heavy pressure to increase the level of safety on its national roads. In 2011, the road- worthiness program was introduced to rectify safety deficiencies based on nominal safety. The main goal of the program is to provide a sound level of safety for all road users. Since then, every existing road has been thoroughly assessed to meet the current Indonesian highway standard, starting with national roads on the Java Island. A new problem was then raised when the road- worthiness assessment found that almost all of our national roads in Java were unable to meet this minimum standard. A limited budget has made it difficult for Bina Marga to implement their safety program. Therefore, understanding the relationship between regulations and crash frequency is important to be understood at a national level, and evaluate whether these regulations are helping improve safety. Furthermore, it can serve as a guidance to a national road improvement program, especially for developing countries that are struggling to build roads in compliance with their national standard. This study uses crash data for Indonesian arterial road to study this relationship.

Indonesia Road-Worthiness Evaluation

In 2011, Bina Marga, Indonesian Highway Authority, developed an approach to evaluate the keyroad elements based on its technical attributes as a way to identify and rectify the safety deficiencies of Indonesian roads. Road-worthiness evaluation or Uji Laik Fungsi Jalan is regulated by the Road National Law No. 34 Year 2006. It is stated that every road must go through an assessment of technical standards in order to ensure safety. It is expected that safety can be achieved by the compliance of road elements to technical design standards.

National highway standards are the basis on which the road is evaluated in the road-worthiness program. In total, there are 158 road elements which become the focus of technical evaluation. The highway standard is differentiated by the road function (arterial, collector, regional or local road) and type of road classification (freeway, highway, medium road and small road). The standards may be assigned in a fixed number (e.g. number of lanes), minimum value (e.g. lane width, shoulder width, median width, overtaking distance, and stopping distance) or maximum value (e.g. traffic volume, critical length, and international roughness index), range value (e.g. median height and curb height), or discrete (e.g. shoulder drop and approach from minor road). Specification and criteria in Indonesian highway standards are mostly derived from the American Association of State Highway and Transportation Officials (AASHTO). They are regulated in the following legal documents:

1) Public Works Ministerial Decree No. 19/PRT/M/2011 regarding highway technical design specifications and criteria;
2) Public Works Ministerial Decree No. 20/PRT/M/2010 regarding road utilization guidelines;
3) Government Regulation No. 34/2006 regarding highways;
4) National Law No. 22/2009 regarding traffic and land transport;
5) National Law No. 38/2004 regarding roads.

METHODOLOGICAL APPROACH

The nature of this study is a cross-sectional study. It features a 'snap-shot' description of a situation at one specific point in time. In this study, the 2012-2014 casualty crash data was used as a predicted variable. East Java province was selected as a case study due to its higher level of accidents compared to other provinces in Indonesia. Forty-five key elements were analyzed as predictor variables using Stata software. The Zero-Inflated Negative Binomial Model (ZINB) regression was preferred as the best model due to the nature of the accident count.

Data Collection

About 100 meter segmentations of East Java North Corridor (EJNC) were based upon iRAP survey data which was carried out in 2011. Of 4116 segments available in the iRAP database, only 2722 segments were able to be used for analysis. The segments were sorted in 22 chainages according to National Road Nomenclature (Ministry of Public Works, 2009).

Crash data was extracted from the national IRSMS site which is available online for a limited access through the Traffic Police website. 2010-2013 crash data was extracted on August 2014, while 2014 crash data was extracted on April 2015. Only casualty crashes were included in the data analysis due to the suspiciously small percentage of Property Damage Only (PDO) crashes recorded (93 out of 2205). Moreover, PDO was rarely recorded in Indonesia due to the reluctance of road users to be involved in the tedious process of accident reporting. Chainages with bad crash data were identified based on the accuracy of crash coordinates. Chainages where most reported crashes were not within 15+5 meters from the road axis were omitted from the analysis. 15 meters is the standard width of the road reserve area, while 5 meters compensates for Google Earth accuracy (Paredes-Hernández, et al., 2013). Depending on the nature of the road environment, Police officers may record the accident coordinates from the house or building on the side of the road. Bad crash data was also indicated by highly agglomerated crashes in a location with accuracy of less than 5 meters. It was suspected that the Police inputted coordinate data of their stations, rather than of the crash locations.

Traffic volume and the mixture of traffic was developed from the Integrated Road Management System (IRMS) which is currently used in Bina Marga as the basis for planning, programming and budgeting. The 2014 data was acquired in February 2015.

Road-worthiness data was acquired from Balai Besar Pelaksanaan Jalan Nasional V Surabaya (BBPJN V), the executing agency for the road-worthiness program for Central Java and East Java Province. It was acquired in November 2014, when BBPJN V had not yet completed their assessment of all roads in the North Java Corridor network. The road-worthiness process recorded more than 150 properties for a single segment. For the purpose of this study, only elements that have a relationship with the accident (based on studies in the previous chapter) were included for analysis.

In road-worthiness evaluation, a road is segmented into several sections for the purpose of assessment. It is segmented by its specification and type of road, by distinctive physical and traffic volume similarities. Road key elements for each segment are then measured and compared to specifications. The percentage of deviation is recorded and assigned for each element. This figure will be revisited at a later stage of the road-worthiness program for budgeting purposes.

For the purpose of formal evaluation, the mark of 'pass' is awarded to categories in which all elements have 0% deviation ($\Sigma Dev = 0$), otherwise 'pass with rectifications' is awarded. The lowest mark in the category shall determine the overall mark for technical standard. The overall mark in both technical and administrative standards will determine the final status of the road (pass, pass with rectifications or fail). The status of 'fail' is given when all elements in a single technical category deviate from the standard (Ministry of Public Works, 2010).

Due to the different sources and measurement of data, a 'data mapping' process was undertaken. Coordinates (latitudes and longitudes) for segmentation were acquired from the iRAP database in the form of a Comma-Separated Value (*.csv) file. Google Earth was used to provide visual representation of segmentation. Those coordinates were converted to Keyhole Markup Language (*.kml) file using free online converter website http://www.convertcsv.com. Crash data was also converted to .kml, overlapped with iRAP segmentation and then verified. The total number of accidents and each type of accident were recorded per 100m. Each 100m segment was treated as a single entity and assigned with ID (i.e. segment 1 to 2722).

Analyzing Data

This research utilizes 3 years (2012-2014) of data. The dominant type of casualty-crash recorded was rear end crashes (37%), followed by right angle crash (25%) and head on crash (18%). Pedestrian and run off crash were the least likely to be a casualty-crash.

Overall, there are 48 road sections that are divided into 2722 segments, there were a total of 3128 casualty-crashes resulting in 709 deaths and 4046 injured. The average number of crash in each segment is 1.1492. The maximum number of accidents in one segment was 15 casualty crashes, while the 50% percentile of the data had 1 accident count. The variance (σ = 2.1078) was greater than the mean (μ = 1.1492). This indicated over-dispersion of data. Furthermore, the variance showed that the distribution of the accident count was not symmetrical. The high kurtosis value (9.65) indicated that the variability of the data was caused by the extreme differences from the mean. The following graph shows the distribution of the accident frequency which demonstrates that the distribution is skewed to the left.

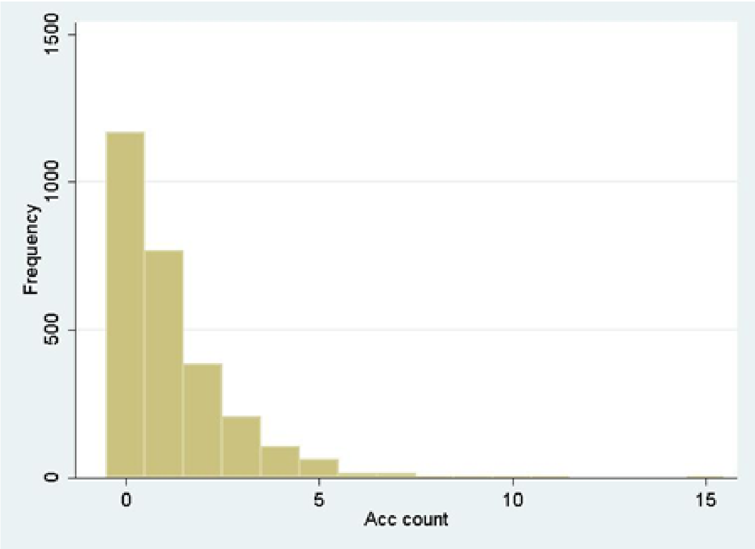

Figure 1 The Distribution of The Accident Frequency

To obtain a good predictive model, Zero-Inflated Negative Binomial (ZINB) regression was considered. Explanatory variables for the initial model were selected based on the theories and previous studies. Additional variables were included based on the practical knowledge in the field of road safety in Indonesia. This expert knowledge was considered very important and guided most of the model formulation in this study. Backward Elimination methods were employed to eliminate possible variables to acquire sound prediction models. Backward elimination started with a model which includes all possible variables. Variables were deleted one by one from the model until all remaining variables in the model significantly contribute to the explained variable. P-to-remove was decided to be as close as to 0.05. One of the drawbacks of this method was that the removal of one variable inclines to raise the significance of the remaining explanatory variables. One can easily overstate the importance of those variables.

A level of significance α less than 0.05 was selected for rejecting the null hypothesis with consideration to the consequences of the decision. This study is aimed to provide Bina Marga with a sound ground to develop a road safety program. The high probability of a variable in determining casualty-accident occurrences could decrease error in model interpretation. Thus, it increases the chance to save road users' lives.

A total of 32 explanatory variables were included in the initial model. Then the variables with low significant value were eliminated one by one until all the remaining variables had a significant level higher than P critical. This resulted in a stable negative binomial model. The next step was creating a ZINB model. Due to the nature of ZINB, a stepwise procedure was chosen to add and remove the variables in the logit part of the model based on the z value. The value of the Likelihood Ratio (LR) Chi-Square test was used as one of the considerations in building the final model. The probability of obtaining chi-square is compared to alpha level (willingness to accept a type I error). For prediction and test of the final models, the parameter estimation, and fit statistics were presented side-by-side for comparison. Three tests: (1) Bayesian Information Criterion, (2) Akaike Information Criteria, and (3) Vuong test were applied to compare the maximum statistical fit of the models.

RESULTS

Comparing the log-likelihood between models, accident prediction models per type of crash are better in predicting the crash counts than of total accidents. Of 34 possible variables, 12 variables are identified as significant factors in a model for total accident count, 5 variables are identified as significant factors for pedestrian crashes, 5 variables are identified as significant factors for run off crashes, 10 variables are identified as significant factors for right angle crashes, 12 variables are identified as significant factors for head on crashes, and 8 variables are identified as significant factors for rear end crashes.

The percentage of non-motorized vehicles in the traffic is found to be a very significant factor for head on and pedestrian crashes. An increase of 1% in the non-motorized vehicle volume will increase head on and pedestrian crashes by the factor of 1.087 and 1.042 respectively, holding the other factors constant in three years' time. A high percentage of non-motorized vehicles usually occurs in high density populated areas where a lot of local movement can be observed. This mixed traffic environment has a high potential for head on crashes when faster vehicles attempt to overtake the slower vehicles in front of them, especially when there is no extra lane for the slower vehicles. Head on crashes also has a significant relationship with the number of lanes.

Reducing 10% deviation from minimum spacing between median openings has been found to reduce the number of rear end and right angle crashes by the factor of 0.914 and 0.940 respectively, holding all variables constant in three years' time.

Disturbance to pedestrian paths also highly contributes to the number of pedestrian and rear end crashes. A 10% increase in effective pedestrian paths would decrease pedestrian and rear end crashes by the factor of 0.920 and 0.965 respectively, holding all variables constant in 3 years' time. Utilization of pedestrian paths for other functions, such as informal commercial activities, pushes pedestrians into the traffic lanes and increases the risk of pedestrian crashes. The pedestrian who disturbs the traffic, may cause a rear end crash as the vehicles stop abruptly to avoid hitting the pedestrians.

Interestingly, motorcycle percentage has a negative relationship with the number of rear end and head on crashes. An increase of 1% in motorcycle volume, decreases rear end and head on crashes by the factor of 0.990 and 0.978 respectively, holding all variables constant in 3 years' time. An increase of slower vehicle percentage such as motorcycles may reduce the overall operating speed on the road. Speed will even lessen when the proportion of motorcycles is higher (as shown by the motorcycle population mean of 35.346%).

ROW utilization is also an important predictor with negative relationship with the run off, head on and rear end crashes. The higher the percentage of ROW use for other purposes, the higher the disturbance to the traffic. It will also cause a slower speed environment. Clearly the lower the speed the safer the road, but it is counterproductive. If a ROW was made to be 100% effective (i.e. to eliminate side friction), a higher number of crashes can also be expected.

Another interesting finding is the negative relationship between the type of intersection (approach from minor road) and the number of total accidents, pedestrian and right angle crashes. A standard intersection (equipped with traffic lights and accelerating and decelerating lane) is found to be more dangerous than a non-standard intersection. This may be explained by other factors such as the phasing of the traffic light, or higher speed in the intersection due to provision of an acceleration lane.

SUMMARY AND CONCLUSIONS

Road features such as speed limits, median openings and pedestrian facilities are significant predictors for the accident count. The lower their deviations from the national standard, the lower the probability that accidents may occur. These features should be targeted for the national road improvement program to increase safety. Motorcycle percentage, ROW utilization, and the type of intersection are also significant variables which have a negative relationship with accidentoccurrences.

The accident prediction models in this study should be tested and compared with other national roads in Indonesia. They can also be useful in predicting the number of crashes in other national roads in Indonesia and to help identify the worst segments for treatments. For further study, the accident count of full-compliance-to-standard road segments can beinvestigated and the relationship of standard adjustment to the change in the accident count can be measured.

In terms of crash prediction model development, it would be interesting to see whether or not the shifting in modelling, has something to do with the state of road safety in one's country. Currently, Indonesia is having very significant problems with crashes, while in most developing countries, the number of accidents has decreased due to major efforts in improving safety on their roads. Fewer accidents means more zero value in a given segment of road.

This study is the first study in Indonesia that analyzes the newly developed Police crash data system. It started at the beginning of the data improvement process. The reliable data is still far from being achieved. Therefore the analysis done in this study must be revisited in the future to get more valid results.

ACKNOWLEDGEMENTS

This study would not have been possible without the help of Bina Marga, Indonesian Highway Authority, Indonesian Traffic Police, and iRAP. We thank them for making all the data available.

REFERENCES

Hauer, E. 2000. *Safety in Geometric Design Standards II: Rift, Roots and Reform*. In Krammes, R., Brilon, W. (ed) 2nd International Symposium on Highway Geometric Design, Mainz, Germany.

Ministry of PublicWorks. 2009. *Penetapan Ruas-Ruas Jalan Menurut Statusnya sebagaiJalan Nasional*. In Works, M.o.P. (ed) 631/KPTS/2009.

Ministry of Public Works. 2010. *Road-Worthiness Assessment, Specifications and Procedures*. In Works, M.o.P. (ed) 11/PRT/M/2010, Jakarta, Indonesia.

Paredes-Hernández, C.U., Salinas-Castillo, W.E., Guevara-Cortina, F., and Martínez-Becerra, X. 2013. *Horizontal Positional Accuracy of Google Earth's Imagery Over Rural Areas: A Study Case in Tamaulipas, Mexico*. Boletim de Ciências Geodésicas, 19(4): 588-601.