

Establishment of a Composite Safety Index for Pavement Management

Stephen A Arhin* and Asteway Ribbiso

Department of Civil and Environmental Engineering, Howard University, Washington, DC, 20059, USA

Abstract

One of the goals of local transportation agencies is to improve the quality of life for citizens and visitors by ensuring the efficient and safe movement of people and goods through the roadway system. Maintenance and rehabilitation of pavements are necessary to ensure that roadway networks continue to perform at their optimum. Currently, maintenance and rehabilitation of roadway networks depend on several factors including pavement condition indices, funding availability, among others. Previous studies have established relationships between crash frequency and pavement condition indices. However, the combined influence of speed, volume, and crash frequency on pavement indices, and thereby pavement management efforts has not been thoroughly examined. In this paper, a multinomial logistic regression was employed for 193 arterial segments to establish a new categorical variable: Composite Safety Index (CSI). The CSI values or ratings were based on pavement indices, crash frequency, traffic volumes and vehicular speeds to help categorize pavement sections for either maintenance or rehabilitation. The results indicated that the selected independent variables were statistically reliable in ranking pavement sections for rehabilitation or maintenance based on their CSI values.

Keywords: Composite safety index; Pavement management; Traffic safety data; Traffic operations

Introduction

A vibrant transportation infrastructure is essential for sustaining the social and economic growth of a city. When roadway networks are efficient, they provide economic and social opportunities and benefits which result in better access to the markets, and employment for citizens. Surface transportation infrastructure in the United States of America (USA) is aging according to a report prepared by Department of Homeland Security in 2014 [1]. The USA was ranked 16th in the overall quality of roads in a report developed by World Economic Forum despite a 3rd place ranking in an overall competitiveness of the economy among the rest of the world [2]. The American Society of Civil Engineers (ASCE) indicated in another study that it would take up to \$225 billion to upgrade the United States' aging transportation infrastructure to meet current demands and future growth. There are two significant concerns regarding the aging transportation infrastructure: the failing existing conditions, and how to prioritize and allocate funding for their Maintenance and Rehabilitation (M&R). Hence, the need for the United States Department of Transportation (USDOT) and local jurisdictions to develop programs and efficient tools to improve and rehabilitate roadway systems. These programs and tools can help local agencies to allocate fund to enhance and preserve transportation networks appropriately.

Annual pavement condition assessments are conducted by transportation agencies to determine the physical states of roadway networks. Maintenance and rehabilitation (M&R) of road segments are scheduled based on certain factors and funding availability in most jurisdictions. The notable factors include International Roughness Index (IRI) and Pavement Condition Index (PCI). However, most of these M&R programs do not include safety variables and other traffic operations measurements in the decision-making processes. The goal of this research is to incorporate safety and operational characteristics in the prioritizing of M&R of roadway networks.

Safety on roadway networks is an essential component for both economic success and quality of life for all stakeholders. Maintenance and rehabilitation is necessary to ensure that roadway networks may operate safely and efficiently. Even though the primary objective of local transportation agencies is to improve the roadway system within their jurisdictions, there are limited studies that focused on investigating the impact of safety and traffic operational issues on pavement conditions

and the utilization of additional variables to improve pavement M&R strategies. Currently, several states only consider pavement condition indices such as IRI and PCI, along with budget constraint as the primary factors for selecting roadway segments for M&R. Even though this is a methodological way of selecting road segments for M&R, it does not consider the safety component and traffic operational factors of the roadway segment selected.

Objectives

The goal of this research is to develop a safety-incorporated pavement logistic regression model that would aid engineers and local agencies in prioritizing and selecting roadways for M&R. This is achieved by developing a composite safety index (CSI) based on traffic and safety data, and pavement condition by employing multinomial logistic regression (MLNR) analysis. This evaluates a segment based on its CSI value on the basis of which the road segment may be selected for M&R.

The objectives of this research are:

1. To define a composite safety index based on crash frequency, crash rate, fatalities, injuries, and PDOs on each segment.
2. Use a multinomial logistic regression to predict the impact(s) of crash frequency, IRI, PCI, average speed, and AADT on CSI.

Ultimately, this research is aimed at developing a CSI classification model that can be used to aid in the selection of roadway segments for M&R while taking critical safety and traffic operational measures into consideration.

***Corresponding author:** Stephen A. Arhin, Assistant Professor, Department of Civil and Environmental Engineering, Howard University, Washington, DC, 20059, USA, Tel: +1 202-806-6100; E-mail: saarhin@howard.edu

Received April 04, 2017; **Accepted** April 14, 2017; **Published** April 18, 2017

Citation: Arhin SA, Ribbiso A (2017) Establishment of a Composite Safety Index for Pavement Management. J Civil Environ Eng 7: 273. doi: 10.4172/2165-784X.1000273

Copyright: © 2017 Arhin SA, et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Literature Review

The Asset Management Primer, published in 1999 by the Federal Highway Administration (FHWA), establishes the process and decision-making framework which utilizes engineering and economic principles as a base to measure asset performance [3]. The Primer was developed to improve the allocation of funds and plan repairs as well as rehabilitation of pavements. Pavement management is different from all other infrastructure maintenance. A Pavement Management System (PMS) is a tool that is used to collect and monitor the physical state of pavement sections, and forecast future conditions. This system only focuses on the distress and irregularities on the roadway systems.

Pavement distress can be expressed in terms of surface irregularities that can potentially affect the ride quality and safety of road users [4]. Many models have been established to predict the effect that pavement distress has on ride quality, but the potential of crash occurrence is only limited to road characteristics, environmental factors, and traffic conditions. Several factors can influence crashes on roadway networks such as pavement distress or roughness, speed, and volume. Pavement roughness is typically measured by automation and is quantified by an International Roughness Index (IRI). The National Cooperative Highway Research Program (NCHRP) Report 228 (Transportation Record Board) first defined IRI in the late 70's. IRI is commonly obtained from the longitudinal profile of a roadway segment measured in/mi or m/km [5]. It has become an accepted measure of smoothness/roughness since its introduction in 1986 [6]. Data obtained from the profile of a road is used in computer algorithms to calculate the pavement smoothness or roughness; the resultant value is called the IRI. IRI values range from 50 to 1,000. Lower scores are an indication of good quality roads (i.e., smooth), and values greater than 170 indicate lower quality (i.e., rough) road segments [4]. Figure 1 presents an image of a typical survey van used for IRI data collection.

In addition, distress photographs obtained from roadway profiles are reviewed and scored by trained pavement engineers and technicians for classification by severity type. The resulting score is known as the Pavement Condition Index (PCI). The United States Army Corps of Engineers [7] created PCI based on visual survey review of the number and types of distresses in the pavement segment. The outcome of the evaluation, PCI, is a value between 0 and 100, with 100 representing pavements with no observed distress and 0 representing pavement with extensive distresses.



Figure 1: Survey van used for IRI data Collection Florida DOT.

Speed is an important variable used for design and as a performance measure for the evaluation of roadway networks [8]. In an article published by the World Health Organization (2004), speed was identified as a primary risk factor in traffic crashes influencing both potential of accident and severity of injuries. Excessive speeding is defined as speed exceeding the 85th percentile speed of the segment. Each jurisdiction is responsible for setting speed limits based on roadway characteristics and geometry. Speed has been a major factor contributing to roadway crashes [9]. According to the World Health Organization's road safety facts, an increase in average operating speed of 1 km/hr typically results in a 3% higher risk of injury crashes, and a 4 - 5% increase in fatalities on all roadways. Speeding characteristics of a roadway network may be impacted by pavement conditions.

Traffic volumes could potentially contribute to the congestion and crashes on road segments. Average annual daily traffic of a roadway segment is defined as the average number of vehicles that pass a specific point on the road for 365 days expressed in vehicles per day (vpd). Average annual daily traffic is also a variable used in the classification of a roadway segments. Furthermore, engineers typically examine crash rates when comparing roadways with different traffic volumes and crashes using the following expression:

$$Crashrate(R) = \frac{100,000,000 * C}{365 * N * V * L}$$

where C is number of crashes, N is number of years, V is volume of segment, and L is length of the roadway (Federal Highway Administration 2015).

Shan et al. [10] conducted a study to explore the possibility of using a predictive model to improve pavement management in Indiana. The study, conducted in 2001, used average annual daily traffic, age, and IRI in a simple regression to evaluate the performance of arterials and collectors road segments over time. IRI was used as the dependent variable, with age of pavement and average annual daily traffic representing the independent variables. Roadway segments were randomly selected. The analysis resulted in a statistically significant prediction model. Three of the six regression models had a R^2 values higher than 0.50 which were recommended for use to the Indiana Department of Transportation (IDOT). The models are as follows:

$$IRI = 43 + 1.8 \times AGE + 0.0004 \times AADT, R^2 = 0.70$$

$$IRI = 65 + 1.9 \times AGE + 0.0003 \times AADT, R^2 = 0.50$$

$$IRI = 65 + 8.1 \times AGE + 0.0009 \times AADT, R^2 = 0.90$$

In another study, Zeng et al. [11] explored the influence of pavement condition on safety to prioritize highway maintenance. Crash and pavement condition indices (IRI, PCI, and rut depth) data were collected from Virginia Department of Transportation (VDOT) for 2007 through 2011. The study evaluated the performance of good pavement conditions versus deficient pavement conditions on rural, two-lane undivided highways in Virginia. The research employed the empirical Bayes method to explore the relationship between pavement conditions and crash frequency and crash severity. The results indicated that good pavement segments could reduce fatal and injury crashes by 26% compared to deficient pavements. According to the author, improvement of pavement conditions from deficient to good, offers a significant safety benefit in reducing crash severity, but not crash frequency. The resulting model was significant with an R^2 value of 0.75.

Currently, for the District Department of Transportation (DDOT) in the District of Columbia, M&R of pavements are typically scoped

during the winter season for implementation during spring and summer. In addition to budget constraints, DDOT considers pavement PCI data for its M&R programs. IRI data is submitted to FHWA, but is not utilized for pavement M&R. DDOT uses a software program called PAVER to manage its pavement M&R projects. After field evaluation of roadway segments, the resulting PCI is used as an input variable in PAVER which prioritizes the selection of roadway segments based on acceptable PCI value and/or financial constraints. This allows the agency to allocate funds for the most critical roadway segments without taking into consideration traffic operations or safety characteristics of the segments.

According to crash statistics for urban cities in the United States in 2013, major arterials experience the most crashes, followed by local roads as shown in Figure 2.

Other factors including traffic volume and speed may potentially affect the reliability and safety of a roadway segment, which may ultimately influence the need for maintenance and rehabilitation. Previous studies have established relationships between crash frequency and pavement conditions. However, the combined effect of pavement conditions, speed, volume, and crash frequency on the deterioration of pavement has not been examined thoroughly. This paper is aimed at developing a composite safety index that can be used to make decisions in selecting of roadway segments, while taking safety into consideration.

Research Methodology

In this research, a composite safety index (CSI_1 , CSI_2 , and CSI_3) was established using the following variables: crash rate (CR), fatalities (F), property damage only ($PDOs$) and injuries (I). The index ranged from one (1) to three (3), with three (3) representing a poor safety condition, and one (1) representing an excellent safety condition. These four crash characteristics were combined to establish the composite safety index using the following equation:

$$CSI = 0.25*CR + 0.30*I + 0.35*F + 0.10*PDOs$$

where CR is crash rate, I is injuries, F is fatalities, and $PDOs$ is property damage only. The weights were computed using average percentage cost of impact of each variable. Each roadway segment information was entered in the equation to determine the crash composite safety index. The four factors were weighted using values of 0.25 for crash rate, 0.30 for injuries, 0.35 for fatalities, and 0.10 for property damage only, as shown in the equation. The segments were then sorted in descending order based on the CSI values. The composite safety indices were classified as either 1, 2, or 3 based cumulative curve. The first 50% of the segments were assigned to CSI_1 , 51% to 75% to CSI_2 , and those greater than 75% were assigned CSI_3 .

Five independent variables were identified for the regression analysis based on data availability and the pertinent variables from the literature review. Table 1 presents the variables and their descriptions.

International Roughness Index (IRI), Pavement Condition Index (PCI), and the average speed data for 2014 were provided by the District Department of Transportation (DDOT). The International Roughness Index (IRI) and Pavement Condition Index (PCI) data sets included various types of pavement information such as pavement type, segment length, and functional classification. Crash data for the District of Columbia was accessed through DDOT's Traffic Accident Reporting and Analysis System (TARAS). AADT data was obtained from the Metropolitan Washington Council of Government, which oversees regional transportation data, and from DDOT. The most recent AADT data available was for 2013. The data was projected to 2014 using a 5% growth rate. The growth factor was based on a three-year moving average of AADT data from 2006 to 2013.

Multinomial Logistic Regression (MNL) is typically used to model relationships between dichotomous dependent or response variable,

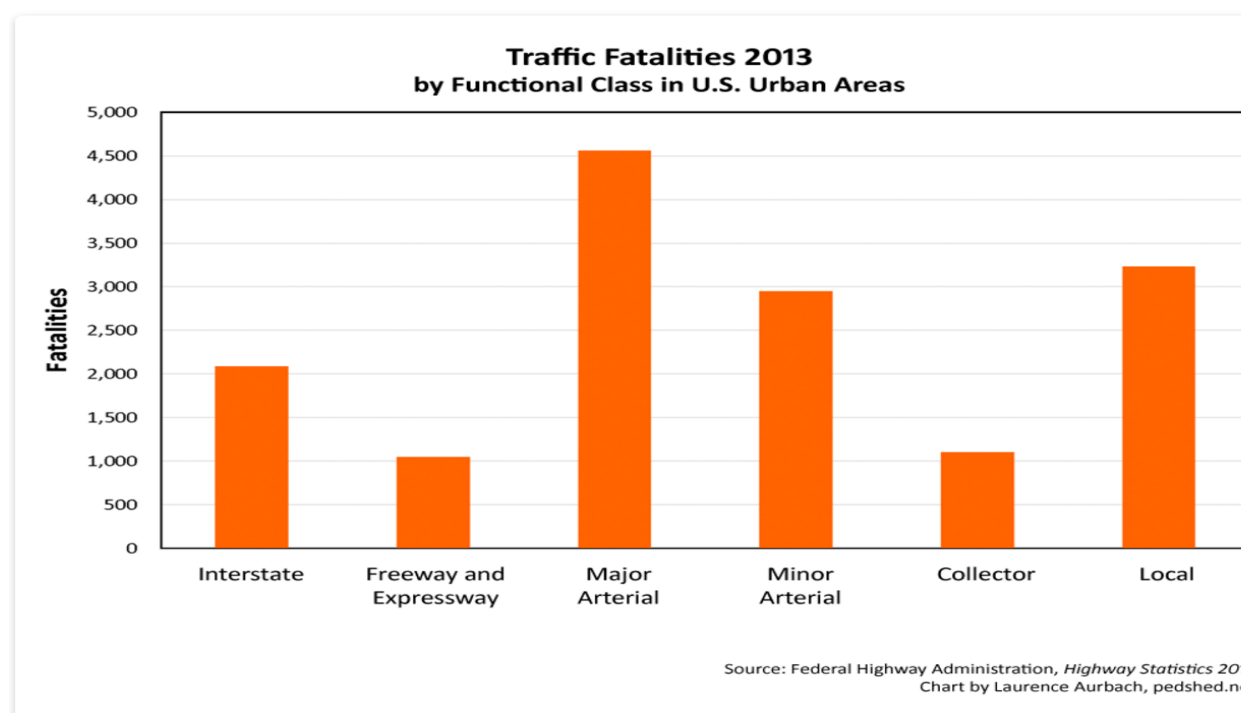


Figure 2: Fatal crashes based on functional classification [3].

Independent Variables	Variables Description	Value of Measurement
International Roughness Index (IRI)	An index for condition of pavement	Index (in/mi)
Pavement Condition Index (PCI)	An index which evaluates distress on pavement surface	Index value 0 to 100
Crash Frequency (CF)	Number of crashes from 2012-2014	Number of Crashes
Average Vehicular Speed (AVS)	Average segment speed	Average vehicular speed (mph)
Average Annual Daily Traffic (AADT)	Annual average daily traffic	Number of Vehicles

Table 1: Composite safety index rating.

and a set of independent variables [12]. Multinomial logistic regression is applicable since the dependent variable is categorical and not in any order. For multinomial logistic regression analysis, it is required that a comparison between two categories be made with one category used as a reference. The reference category for this analysis was CSI₃, which represents a poor safety condition, was compared to CSI₁ and CSI₂. The general equation is as follows:

$$\ln\left(\frac{P}{1-P}\right) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_p x_p$$

In this analysis, the dependent variables have three levels: CSI₁, CSI₂, and CSI₃. The general equation can be rearranged to express dependent variable with three levels. As such, the dependent variable is divided into j categories and one category can be selected as the reference. In this case, the probability equation can be denoted as:

$$P_j(Y_j) = \frac{\exp(\beta_0 + \beta_{1j} X_{1i} + \dots + \beta_{pj} X_{pi})}{1 + \sum_{j=1}^{k-1} \exp(\beta_0 + \beta_{1j} X_{1i} + \dots + \beta_{pj} X_{pi})}$$

where j = 1, 2, ..., k-1 and i = 1, 2, ..., n. The regression analysis compared CSI₁ relative to CSI₃ as well as CSI₂ relative to CSI₃. The logistic equation for each comparison group of categories can denote by these two logistic models:

$$\ln\left(\frac{P(Y=1)}{P(Y=3)}\right) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_p x_p$$

$$\ln\left(\frac{P(Y=2)}{P(Y=3)}\right) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_p x_p$$

where P denotes the number of dependent variable for the binary response Y. β is the coefficient of the equations and X represents the value of each independent variable associated with each coefficient of the equation.

For most regression analysis, it is assumed that the observed data are independent and unbiased. The dependent variable in a MLNR is assumed to follow a multinomial distribution. The probabilities are linked to the independent variable by a logit link function which is assumed to be linear. The logistic regression analyzed 193 segments in the District.

This regression employed the maximum likelihood and the likelihood ratio tests to evaluate the “goodness of fit.” For MLNR, the overall fit is tested using the difference between the constant and final model. The constant model uses the marginal probabilities to predict categories of the dependent variable, while the final model incorporates the influence of all the independent variables to predict the resulting categories. The chi-square test was used to determine the resulting model's significance and is defined as the difference between the

final and constant models. The difference between these two values is referred to as chi-square (χ^2) for goodness of fit and is computed by:

$$X^2 = -2LL_{\text{Constant}} - 2LL_{\text{Final}}$$

The regression output from SPSS shows how well the multinomial regression predicts each CSI category. The “enter” method in SPSS was used to evaluate all the independent variables in the regression analysis and included all the variables at the same time.

In linear regression analysis, the R^2 explains the proportion of variance that can be accounted for by the predictor variable. A high R^2 indicates that the variation in the model and ranges between 0 to 1. However, for MLNR, computing R^2 is not the same as that for a linear regression. The output of the multinomial regression provides a pseudo R^2 values as an approximation of the regression. One of the pseudo R^2 values in a multinomial regression is called the Nagelkerke R^2 . The Nagelkerke R^2 is a modified Cox and Snell R^2 . It ranges between 0 to 1. A high Nagelkerke R^2 value indicates a strong relationship in the logistic regression model.

Results

In the first step of the analysis, pearson correlation analyses between the dependent and each independent variable were performed to identify high relationships between two different groups of data sets. Due to multicollinearity, only four of the five variables were used to predict CSI categories and each of the independent variables.

A MLNR analysis was conducted to predict CSI categories for 193 arterial segments using IRI, AVS, CF, and AADT as the independent variables. A test of the final model against a constant only model was statistically significant, indicating that the independent variables as a set are reliably distinguished between CSI categories ($\chi^2_{(8,193)} = 113.42$, $p < .000$ with $df = 8$) as presented in Table 2.

Nagelkerke's R^2 of 0.537 indicates a moderately good relationship between the prediction and classification of roadway segments. The overall prediction rate was approximately 75%. The Wald criterion confirmed that only CF, IRI, and AADT were significant in contributing to the model while the intercept and AVS were not. The independent variables, AADT and CF, were statistically significant with $\chi^2_{(2,193)} = 34.7$, $p = 0.000$, and $\chi^2_{(2,193)} = 17.38$, $p = 0.000$, respectively when comparing CSI₁ to CSI₃. IRI was statistically significant with $\chi^2_{(2,193)} = 4.007$, $p = 0.000$ when comparing CSI₂ and CSI₃. However, AVS was determined not to be statistically significant as presented on Table 3.

The odds ratio is less than 1 for CF when comparing CSI₁ to CSI₃, which means that the outcome is less likely to be in the CSI₁ category when the CF values are high. On the other hand, when comparing CSI₂ to CSI₃, only IRI has the most statistically significance influence on the model ($p < 0.05$). Thus, when IRI values are high, the category will more likely be in CSI₃.

The reference category is CSI₃

Figure 3 presents a comparison of the actual CSI category versus the predicted CSI category from the resulting MLNR analyzed in the study.

From the graph, it can be concluded that the predicted CSI

Model	Model Fitting Criteria	Likelihood Ratio Tests		
	-2 Log Likelihood	Chi-Square	df	Sig.
Intercept Only	339.375			
Final	225.934	113.441	8	.00

Table 2: Regression “goodness of fit”.

categories are closer to the actual. The CSI model was able to predict categories 1, 2, and 3 at respective rates of 95%, 27%, and 47%, resulting in an overall prediction rate of approximately 75%.

Results and Discussion

A composite safety index (CSI) with a range from 1 to 3 was established using crash rate, fatalities, injuries, and property damage only. This paper used the CSI as the dependent variable with independent variables such as crash frequency (CF), international roughness index (IRI), average vehicular speed (AVS), and average annual daily traffic (AADT) in a multinomial logistic regression (MLNR) at 5% significance level.

The resulting model used arterial segments in a dense urban environment. The final sample size for the regression was 193. The analysis was performed using SPSS to predict CSI using four independent variables. Multinomial logistic regression analysis yielded statistically significant results. From the MLNR analysis, CR, IRI, AVS, and AADT were used as independent variables for the model. Due to the presence of multicollinearity, PCI was excluded from this analysis. The logistic regression model yielded an R^2 value of 0.537 and a statistically significant $\chi^2_{(8,193)} = 113.42$, $p < .000$ with $df = 8$ at 5% level of significance.

The Nagelkerke's R^2 of 0.537 (53.7%) represents a moderate relationship between prediction and grouping, indicating that a good prediction of each sample into the appropriate CSI category. The prediction success for the overall model was 74.6%. Thus, the CSI

categories can help identify segments with different safety attributes for M&R considerations, in addition to funding or budget constraints.

Conclusion

Over the past few years, there has been a significant effort for state agencies to develop to new methods to manage and improve their roadway network. Funding from the Federal government can be effectively applied for the implementation of pavement management programs based on data-driven models. The concept of adding safety and traffic operational factors to pavement condition indices and budget constraints may potentially improve the evaluation, selection, and prioritization process for roadway maintenance and/or rehabilitation. This research evaluated the potential of including four variables in a statistically significant logistic regression model in addition to pavement indices for pavement M&R for arterial roadway segments. The models were developed using data for an urban area (Washington DC) therefore they are not necessarily recommended for use in other jurisdictions.

Future research needs would include the following:

(i) Validate the model using the list of arterial roadway segments selected for maintenance or rehabilitation in 2014.

(ii) Develop and validate similar models for other roadway functional classifications for the District of Columbia.

References

- Department of Homeland Security (2010) Infrastructure: Issues, research, and technology. Buildings and Infrastructure Protection Series BIPS 01.
- Schwab K (2015) The Global Competitiveness Report 2015–2016, World Economic Forum, Geneva.
- Federal Highway Administration Office of Asset Management (1999) Asset Management Primer.
- Arhin S, Williams L, Ribbiso A, Anderson M (2015) Predicting pavement condition index using in a dense urban area. Journal of Civil Engineering Research. Volume 5-1:10-17.
- Sayers MW, Karamihas SM (1998) Little book of profiling. University of Michigan Transportation Research Institute.
- Sayers MW, Gillespie T, Paterson W (1986) Guidelines for the conduct and calibration of road roughness measurements. World Bank Technical Paper No. 46.
- ASTM International (2007) Standard practice for roads and parking lots pavement condition index surveys. ASTM International.
- Fitzpatrick K, Carlson P, Brewer M, Wooldridge M (2003) Design speed, operating speed, and posted speed practices. Report submitted to Washington, DC. NCHRP 504.
- Peden, M, Richard S, David S (2004) World Report on Road Traffic Injury Prevention. World Health Organization Geneva.
- Gulen S, Zhu K, Weaver J, Shan J, Flora W (2001) Development of improved pavement performance prediction models for the Indiana Pavement Management System. Publication FHWA/IN/JTRP-2001,1:69.
- Zeng H, Smith BL (2014) Estimation of the safety effect of pavement condition on rural, two-lane highways transportation research record: Journal of the Transportation Research Board, No. 2435: 45–52.
- Kutner MH, Nachtsheim C, Neter J, Li W (2005) Applied linear statistical model. (5th edn). New York, The McGraw-Hill Companies, Inc. NY.

CSI		B	Std. Error	Wald	Sig.
1*	Intercept	1.914	2.457	0.607	0.436
	CF	-0.269	0.046	34.714	0.000
	AADT	0.000	0.000	17.384	0.000
	IRI	-0.004	0.002	3.813	0.051
	AVS	0.066	0.093	0.494	0.482
2*	Intercept	3.174	2.447	1.683	0.195
	CF	-0.041	0.033	1.533	0.216
	AADT	0.000	0.000	2.263	0.133
	IRI	-0.005	0.002	4.007	0.045
	AVS	-0.066	0.092	0.518	0.472

Table 3: Results of regression analysis.

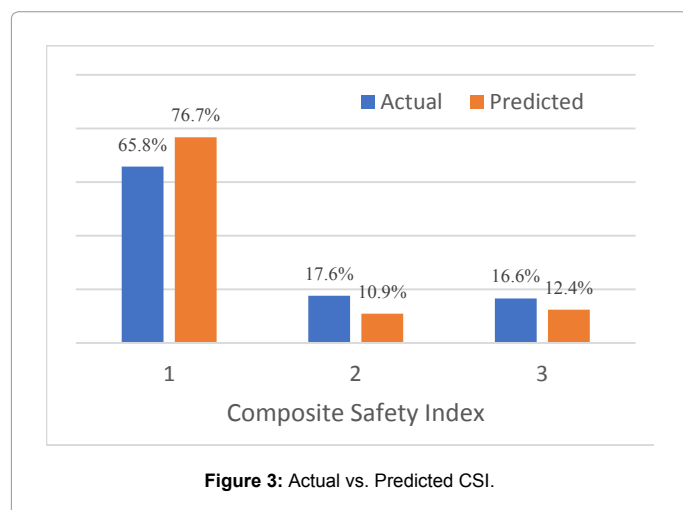


Figure 3: Actual vs. Predicted CSI.