Effect of Pavement Condition on Traffic Crash Frequency and Severity in Virginia

Ali Mohagheghi

Thesis submitted to the faculty of the Virginia Polytechnic Institute and State University in partial fulfillment of the requirements for the degree of

Master of Science In

Civil Engineering

Montasir M. Abbas, Chair

Linbing Wang

Antonio A. Trani

August 8, 2020, Blacksburg, Virginia

Keywords: Traffic Safety Analysis, Pavement Management System, Pavement Condition Assessment, Asphalt Concrete Pavement, Critical Condition Index, International Roughness Index, Rutting Depth, Interstate Highway System, Virginia

Effect of Pavement Condition on Traffic Crash Frequency in Virginia

Ali Mohagheghi

ABSTRACT

Previous studies show that pavement condition properties are significant factors to enhance road safety and riding experience, and pavements with low quality might have inadequate performance in terms of safety and riding experience. Pavement Management System (PMS) databases include pavement properties for each segment of the road collected by the agencies. Understanding the impact of road characteristics on crash frequency is a key step to prevent crashes. Whereas other studies analyzed the effect of different characteristics such as International Roughness Index (IRI), Rutting Depth (RD), Annual Average Daily Traffic (AADT), this thesis analyzed the effect of Critical Condition Index (CCI) on crash frequency, in addition to the other factors identified in previous studies. Other characteristics such as Percentage of Heavy Vehicles, Road Surface Condition, Road Lighting Condition, and Driver Conditions are taken into the consideration. The scope of the study is the interstate highway system in Fairfax County, Virginia. Negative Binomial, Least Square and Nominal Logistic Models were developed, showing that the CCI value is a significant factor to predict the number of crashes, and that it has different effect for different values of AADT. The result of this study is a substantial step towards developing an integrated transportation control and infrastructure management framework.

Effect of Pavement Condition on Traffic Crash Frequency in Virginia

Ali Mohagheghi

GENERAL AUDIENCE ABSTRACT

Many factors cause crashes in the roads. Although there is a common sense that road characteristics such as asphalt quality are important in terms of road safety, there are few studies that scientifically prove that statement. In addition, asphalt maintenance decisions making process is mainly based on cost benefit optimization, and traffic safety is not considered at the process. The purpose of this study is to analyze crashes and road characteristics related to each crash to understand the effect of those characteristics on crash frequency, and eventually, to build a model to predict the number of crashes at each part of the road. The model can help transportation agencies to have a better understanding in terms of safety consequences of their infrastructure management plans. The scope of this study is the highway interstate system in Northern Virginia. Results suggest that pavement condition has a significant impact on crash frequency.

To my parents Parvaneh and Jafar,

And my sister Maryam

ACKNOWLEDGMENT

I would like to thank my advisor, teacher and mentor Dr. Montasir Abbas, not only for giving me the opportunity to be a Virginia Tech student, but to let me take advantage of his valuable knowledge and expertise. His endless desire to learn and solve problems has motivated me and made me the student that I am now.

I like to thank my committees Dr. Antonio Trani and Dr. Linbing Wang for their precious guidance through my research process. I want to appreciate Dr. Gerardo Flintsch and Dr. Kevin Heaslip for being brilliant teachers and help me to learn priceless lessons.

Finally, I would like to express my deepest gratitude towards my family. They always support me with their endless love and energy.

Table of Contents

LIST OF FIGURES

LIST OF TABLES

1 INTRODUCTION

Two main goal of traffic engineering is to improve safety and mobility. Safety has always been a priority in transportation design, transportation planning and management. One of the most practical safety measures is the number of crashes in a period of time at a place. Agencies use the safety performance functions (SPF) to predict the number of crashes in a part of the network. There are different SPFs for different parts of the transportation system, but they all use AADT as one of their inputs. An example of a highway safety performance function is;

$$Predicted\ Crashes = \exp[\alpha + \beta \times \ln(AADT) + \ln(Segment\ Length)], \quad (1)$$

where AADT is annual average daily traffic in vehicles/day, Segment Length is in miles, $\alpha\ and\ \beta$ are calibration parameters.

Whereas some studies show pavement quality has effect on number of crashes (summary of those studies is provided in the literature review section), but, as it is shown in the function above, there is no factor other than AADT and segment length in the equation. It means that number of crashes per segment has nothing to do with other factors such as quality of asphalt pavement, weather condition, lighting condition, etc. And when there is no input for those characteristics, they will not be considered in decision making process. But in reality, and based on previous studies, the number of crashes have relationship with road quality indices and road condition. By considering those factors in crash prediction models, not only the accuracy and reliability of the models will increase, but also there will be future safety measures for infrastructure and traffic planning. This study sheds light on infrastructure and traffic management by providing a model to predict crash frequency in highway system. The crash model was developed by analyzing the effect of road characteristics and traffic condition on number of crashes per mile per segment. To obtain a better perspective of factors, Negative Binomial, Least Square and Nominal Logistic Models were developed, the results were compared, and the best model is suggested. The result shows that some factors such as Critical Condition Index (CCI) and Percentage of Heavy Vehicles are significant in predicting the number of annual crashes.

2 LITERATURE REVIEW

2.1 PAVEMENT INDICES

The main parameters that define pavement condition are Rutting Depth (RD), Roughness, and Percentage of Fatigue Cracking Area. Each agency uses a different method to come up with a practical value to define the general quality of pavement such as Present Serviceability Index (PSI), Pavement Condition Index (PCI), Critical Condition Index (CCI), etc.

Rutting is defined as roadway deformation along the wheel path and the units for its measurement are inches or millimeters. If the rutting depth is deep, it creates an unsafe condition for the users when they try to get out of the rut path, especially at higher speeds. Rutting can cause more severe safety problems in case of wet weather when water accumulates in the rut paths. Previous studies indicate that 0.30 in. (7.6 mm) of Rutting Depth is the threshold that increases crashes significantly [Marc R. et al., 1998]. There are several models to predict the future rutting depth of the pavement. The equation below is a rutting prediction model by Ricardo et al.

$$RD_{it} = \beta_{i10} + a_i N_{it}^{b_i} , \tag{2}$$

where RD_{it} is rut depth for section i at time t in mm, N_{it} is cumulative number of load applied on section i up to time t, a_i and b_i function characteristics of pavement I such as thickness, gradation. etc.; β_{i10} is initial rut depth immediately after construction.

International Roughness Index (IRI) was developed by the World Bank in 1980s. Most of agencies around the world use IRI as an indicator for pavement condition. It indicates the characteristics of longitudinal profile for a traveled wheel path. “The IRI is a filtered ratio of a vehicle’s accumulated suspension motion in inches divided by the distance traveled by the vehicle in miles.” Common units of IRI are inch per mile and meter per kilometer. IRI is the most common pavement roughness measure in the United States as well as most countries all around the world. When the pavement roughness increases, the area of contact between wheels and pavement decreases, and the situation leads to a lower brake friction. Driving comfort level decrease when a pavement has a high IRI value rough pavement can cause damages to vehicle’s suspension system. “Pavement roughness is also the main cause of load loss crashes.” The model below is one of the IRI prediction models (in m/km) (Saleh et al 2000).

$$IRI = -1.415 + 2.923\sqrt{IRI_0} + 0.00129\sqrt{N} + 0.000113T - 5.485 \times 10^{-10}P^4 - 10^{-5}T\sqrt{N} + 5.777 \times 10^{-12}P^4\sqrt{N}, \tag{3}$$

where IRI is international roughness index in m/km, IRI_0 is initial IRI in m/km, N is number of axle repetitions, P is axle load in kN, T is asphalt layer thickness in mm.

Fatigue cracking is one of the major distresses that appears on asphalt pavement surface and it can cause serious damages to the pavement especially when accumulated water in cracks freezes. Fatigue cracking is measured as the ratio of accumulated cracking area divided by total area of the pavement surface. Hence, it is unitless and it is a percentage. It has always been a case of study for asphalt pavement designers. A study shows that fatigue cracking has a higher probability to occur on the asphalt pavement when the pavement is experiencing a large deflection and more load frequency. Several types of fatigue cracking models have been found, so different mechanistic–empirical (M-E) methods are required for pavement design. Pavement design methods vary for different roads, due to the different weather and traffic condition. Severe fatigue cracking can cause safety issues; when some parts separate from the road surface and when tires touch those small objects, they can be thrown towards the windshields of other vehicles. One of the models to predict future fatigue cracking is as follows (Sheng Hu et al. 2012);

$$FC = \left(\frac{6000}{1+e^{C_1 - C_2 \log(D)}}\right) \times \left(\frac{1}{60}\right), \tag{4}$$

$$D = \sum_{i=1}^{T} \frac{n_i}{N_i}, \tag{5}$$

$$C_1 = -2 \times C_2, \tag{6}$$

$$C_2 = -2.40874 - 39.748 \times (1 + \times\, h_{ac})^{-2.65609}, \tag{7}$$

where FC is percentage of fatigue cracking of total lane area, D is damage factor, n_i is actual traffic for period i, N_i is allowable failure repetitions under condition prevailing in period i, C_1 and C_2 are thickness factors, and h_{ac} is asphalt layer thickness in inch.

There are other asphalt pavement distresses such as bleeding, patching, raveling, stripping, etc., they are not as common as the three main distresses, but they are included in the calculation of some pavement quality indices.

2.2 Critical Condition Index Significance

There are several works related to the subject of this study, but the contribution of this study to the literature is that Critical Condition Index (CCI) is considered as a factor in the analysis, and also the location of study is a great model for all metropolitan areas in the US. Due to the variation of methods for crash analysis, this work provides different analysis methods and compare the results. “CCI is calculated as the minimum of Load-Related Distress Rating (LDR) and Non-Load-Related Distress Rating (NDR). These indices were first developed in 1998 based on the PAVER methodology developed by the 15 US Army Corps of Engineers and have been validated by using the Long-Term Pavement 16 Performance (LTPP) data collected through the Strategic Highway Research Program (SHRP).” (VDOT, 2018) CCI is a numerical index between 0 and 100 that indicates the quality of pavement. A healthy pavement that is just paved, has the CCI value of 100 and the lower the quality gets, the lower the CCI value gets. LDR is a function of loads (the traffic) that pass the road and NDR is a function of factors such as weather condition, chemical properties of the pavement and radiation. A model has been developed to predict CCI value and it has Modified Structural Index (MSI) as an input “The Akaike Information Criterion (AIC) suggesting that the model that includes the MSI is, at least, 50,000 times more likely to be more accurate to the true model than the model that does not include the MSI” [Katicha et al., 2016]. Accurate and reliable models for CCI makes the contribution of this work more significant in terms of future transportation design and infrastructure maintenance and repair planning. Equation 1 is the CCI model;

$$CCI_{Model} = 100 - T^{\beta_2} e^{\left(\beta_0 + \beta_1 \frac{1}{MSI^4} + \beta_3 T\right)}, \tag{8}$$

where CCI_{Model} is the predicted value of Critical Condition Index (CCI at T=0 is 100),

T is time in years,

MSI is Modified Structural Index value (Appendix)

$\beta_0, \beta_1, \beta_2\ and\ \beta_3$ are calibration factors

2.3 Similar Works

However, there are several studies about crash analysis, few of them are about the effect of road characteristics on crashes. All the related works have been reviewed and more relevant studies were identified. A key factor to contribute something valuable to the literature, is to carefully understand previous works, especially the methods that they used to analyze the data. Table 1 shows some highlights about previous works;

TABLE 1 - LITERATURE HIGHLIGHTS

Title		1. Tsubota et al., 2018 "Effect of Pavement Conditions on Accident Risk in Rural Expressways"	2. Lee et al., 2015 "Effects of Pavement Surface Conditions on Traffic Crash Severity"	3. Chan et al., 2010 "Investigating Effects of Asphalt Pavement Conditions on Traffic Accidents in Tennessee Based on The Pavement Management System"	4. Awad et al., 2018 "Investigation the Effect of Pavement Condition Characteristics on Bend Segments Accident Frequency"	5. Vinayakamurthy et al., 2017 "Effect of Pavement Condition on Accident Rate"	6. Hou et al., 2020 "A Correlated Random Parameters Tobit Model to Analyze the Safety Effects and Temporal Instability of Factors Affecting Crash Rates"	7. li et al., 2013 "Impact of Pavement Conditions on Crash Severity"
Factors Related to Pavement		Rutting Depth, IRI, Cracking Ratio	PSI	Rutting Depth, IRI, PSI	Skid Resistance, Geometric Design, Rutting Depth	Rutting Depth, IRI	Rutting Depth, Friction Coefficient, Distress Ratio	IRI, Skid Score, Distress Score, Condition Score
Y (Response)		Accident Risk	Crash Severity	Accident Frequency	Accident Frequency	Accident Rate	Crash Rate	Crash Severity
Method of Analysis		Poisson Regression	Bayesian ordered logistic regression	Negative Binomial	Negative Binomial	Sigmoidal function	Correlated Random Parameters Tobit	Chi-Square Statistics
Crash Rate Function						✓	✓	
Safety Performance Function		✓		✓	✓			
Categorizing Deteriorations		✓	✓	✓				
Other Factors	AADT	✓	✓	✓	✓	✓	✓	✓
	Surface Cond.	✓	✓	✓	✓			✓
	Crash Time		✓	✓				✓
	Lighting Cond.		✓	✓				✓
	Speed Limit		✓		✓			✓
	No. of Veh.		✓					
	Truck %		✓		✓		✓	✓
	Road Type	✓	✓	✓	✓	✓		
	Work Zone		✓					
	Peak Hour			✓				
	Crash Location							
Result (Pavement Significant Factors)		RD in *Rain, IRI (inv.)	Poor Pavement increases. Multi-vehicle crashes.	PSI, IRI in *Rain	Not Significant	Thresholds for IRI & RD	Rutting Depth in Rain	Poor Pavement More Severity

To better understand the similar works, a brief description for each one is given below;

1. The goal of study number 1 is to reveal the effects of road condition on traffic crash risk. The motivation is to demonstrate the safety respond of the road for infrastructure planners, so they can include safety in their pavement maintenance plans. The study uses empirical analysis to find the relationships between the pavement conditions and crash risk. The pavement condition factors in the analysis are rutting depth, IRI and cracking ratio. The regression method for modeling the crash frequency is Poisson regression. The results indicated that rutting depth could have a significant effect in increasing the risk of crashes, especially in wet surface condition. Results also showed that IRI has an inverse effect on crash risk. The analysis showed that cracking ratio does not have a conclusive effect on crash risk.
2. The other study (2) as unlike other crash analysis studies, investigated the effect of pavement condition on crash severity levels using a discrete model that can handle ordered data. The paper focuses on the effect of poor pavement condition on crash severity levels at low/medium/high speed roads and single/multiple collision cases, using Bayesian Ordered Logistic models. The models showed that the severity of single-vehicle crashes decrease at the poor pavement condition at low-speed roads, and they increase at high speed roads. In addition, the poor pavement condition increases the severity of multiple-vehicle crashes on all roads.
3. The study (3) looks into the relationship between crash frequency and pavement distress variables, ant it eventually integrated the Tennessee Pavement Management System (PMS) and Crash History Database (AHD). Study focused on urban interstates and it developed several Negative Binomial Regression models to predict number of crashes under different pavement condition factors, including rutting depth, Roughness (IRI), and

Present Serviceability Index (PSI). Results showed that the RD did not perform well, except for predicting crashes at night and crashes under rain weather conditions; but, IRI and PSI were significant prediction variables in all types of crash models. After Comparing the results, the paper concluded that PSI is more significant to predict crash frequency than IRI and RD.

4. The next paper (4) used fixed and random parameter negative binomial models to investigate the effect of pavement condition properties on crashes for a period of six years (2005–2010) in Norfolk County in the U.K. The study used random and fixed parameters to conduct the analysis, results demonstrated that log-likelihood is considerably more significant when using random parameters models. The study showed that the AADT, percentage of heavy vehicles, gradients and number of minor access, and rut depth have relationship with crashes. Other factors like Speed limit, radius of curvature, number of lanes, skid resistance, and texture depth are significant to decrease the numbers of crashes.

5. The last study (5) was conducted based on crash data of highways in the states of Arizona, North Carolina and Maryland for the years between 2013 and 2015 to investigate the relationship between crash rate and IRI and RD. Results showed that not only roughness and rutting are effective factors to predict number of crashes but combination of IRI and RD with AADT and human factors are possibly significant.

6. The next study (6) views the crash rates directly as a continuous variable left-censored at zero and explores the application of an alternate approach based on Tobit regression. The study also helped in developing more effective countermeasures for safety for better

understanding the effect of different factors associated with freeway design characteristics and pavement conditions.

7. The last related study (7) "analyzed the correlation between several key pavement condition ratings or scores and crash severity based on a large number of crashes in Texas between 2008 and 2009. The results showed that poor pavement condition scores and ratings were associated with proportionally more severe crashes, but very poor pavement conditions were actually associated with less severe crashes. Very good pavement conditions might induce speeding behaviors and therefore could have caused more severe crashes." (7)

Based on the literature, different studies use multiple methods for modeling the severity and the frequency of crashes. The variety of methods might be due to different datasets that different studies have used. Based on a study, Bayesian models normally have a better performance that their counterpart models. None of the studies in the literature has analyzed severity and frequency of crashes together. Therefore, both crash severity and crash frequency analysis are included in this study.

2.4 Scope of Study

Washington, DC metropolitan area is one of the most populated areas in the United States (ranked 6th in the US). It has also a growth rate of 11% which is moderately high. The growth rate is expected to be higher due to the opening of Amazon second headquarter (it will bring approximately 50,000 jobs to the area by 2023). Fairfax County is the most populated area in metropolitan Washington, DC, it has more population than the District of Colombia itself. Public transportation in the area is limited and the primary trip mode for users is driving. Hence, interstate highway system, especially in Fairfax County, is serving most of the travelers, consequently, it is very congested, number of crashes are considerable, and percentage of injury crashes are high. Interstate Highway System in Fairfax County (and the extension of I-95 in Prince William County) is the scope of this study due to the following reasons; Virginia is the only state in the US that uses CCI as an indicator to demonstrate pavement quality. Fairfax County is a congested area in the US and the number of crashes is considerable. There was no similar study found in Northern Virginia area, so the results of this study might be a helpful contribution to transportation engineering community. Table 2 shows statistics about Roads in Fairfax County Interstate Highway system in a three-year period from 2016 to 2018.

TABLE 2 - FAIRFAX INTERSTATE HIGHWAY SYSTEM HIGHLIGHTS

Road Name	Total Length	Number of Crash	AADT	Study Period (Years)	Crash/Mile/Year/ AADT (in 10000)
I-66	37.94	3945	71000	3	4.881688656
I-95	33.12	2386	102000	3	2.354283098
I-395	12.32	471	86000	3	1.481803081
I-495	29.18	3083	94000	3	3.746615529
Total	112.56	9885			12.46439036

2.5 DATA COLLECTION IN VIRGINIA

A key factor to understand the data, is to understand the data collection process In Virginia. “The Virginia Department of Transportation (VDOT) is responsible for more than 128,000 lane miles of roadway.” (VDOT 2018) Virginia Department of Transportation’s contractor, Furgo-Roadware Inc. is in charge of collecting pavement data by using continuous digital imaging and automated crack detection technology. The data is collected by vans equipped with cameras and sensors. Roughness and Rutting are captured with sensors and cracks are detected after computations on continuous captured images. Other type of distresses such as patching, delamination and bleeding, are identified by analyzing images.

As it was mentioned, CCI is the minimum of the Load-Related Distress Rating (LDR) and Non-Lod-Related Distress Rating (NDR). The distresses along wheel paths are considered as LDR and all other distresses in other locations but the wheel paths are considered as NDR.

TABLE 3 - PAVEMENT QUALITY DESCRIPTION BASED ON CCI VALUE FROM VA STATE OF THE PAVEMENT 2018

Pavement Condition	Index Scale (CCI)
Excellent	90 and above
Good	70-89
Fair	60-69
Poor	50-59
Very Poor	49 and below

2.6 DECISION MAKING IN VIRGINIA

Pavement maintenance can be defined as "a program employing a network level, long-term strategy that enhances functional pavement performance by using an integrated, cost-effective set of practices that extend pavement life, improve safety, and meet motorist expectations".

VDOT uses a pavement management software (PMS) system developed by Agile Assets. The system performs network-level and multi-constraint optimization, which provides infrastructure maintenance plans using single objectives and multiple constraints. The objective is either cost minimization while achieving performance targets or performance maximization while meeting budget constraints. The result of the optimization is to recommend one of the following maintenances and repair treatment strategy for each segment of the road: Do Nothing (DN), Preventive Maintenance (PM), Corrective Maintenance (CM), Restorative Maintenance (RM), or Rehabilitation and/or Reconstruction (RC).

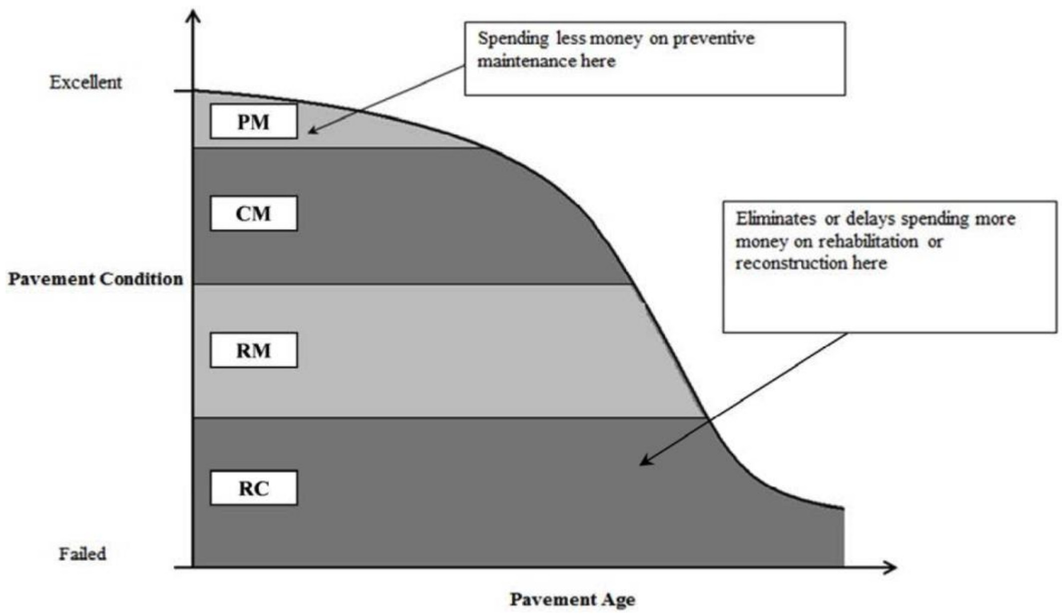

FIGURE 1 PAVEMENT MAINTENANCE BENEFITS

1. Preventive Maintenance (PM)

Preventive Maintenance is "a planned strategy of cost-effective treatments to an existing roadway system and its appurtenances that preserves the system, retards future deterioration, and maintains or improves the functional condition of the system (without significantly increasing the structural capacity)." Preventive maintenance maintains the performance condition of roadway systems, without improving their structural capacity. PM treatments are mainly used for pavements in relatively good condition, and they extend pavement lifetime by maintaining the pavement in an acceptable state.

2. Corrective Maintenance (CM),

Corrective Maintenance treatments are applied when there is one or multiple deficiencies that cause negative impact on safety and driving experience in a section of a pavement. Corrective maintenance strategies are mostly reactive, not proactive, and applied to restore a section of pavement to an acceptable level of service due to insufficient performance.

3. Restorative Maintenance (RM)

Restorative Maintenance treatments are necessary to return an infrastructure facility back to a minimum required level of service while a permanent restoration is being designed and scheduled. Some of the pavement distresses that need RM treatments are pavement blow-ups, road washouts, avalanches, or rockslides.

4. Rehabilitation and/or Reconstruction (RC)

Pavement Rehabilitation is described as "structural enhancements that extend the service life of an existing pavement and/or improve its load carrying capacity. Rehabilitation techniques include restoration treatments and structural overlays." Rehabilitation projects are applied to extend the life of pavement either by restoring existing structural capacity by eliminating the age-related distresses or increasing the pavement thickness. Pavement Reconstruction is the process of the replacement of the entire existing pavement structure to obtain a pavement with structural capacity equal to or more than the initial capacity of previous pavement. Reconstruction process requires the complete removal and replacement of the existing pavement structure. Table below shows different treatment costs for a lane-mile of primary road in Virginia (from E. de León Izeppi, A. Morrison, G. W. Flintsch, and K. K. McGhee, "Best practices and performance assessment for preventive maintenance treatments for Virginia pavements," Virginia Center for Transportation Innovation and Research, 2015)

TABLE 4 PAVEMENT PRESERVATION GUIDELINES

Activity Category	Activities	Cost per Lane-Mile for Primary Road
Do Nothing (DN) N/A		
Preventive Maintenance (PM)	Surface Treatment (Chip Seal, Slurry Seal, Microsurfacing, Ultra-thin bonded wearing course, etc.)	6977
	Crack Sealing	
	Minor Patching (< 5% Pavement Area) Surface Patching (Depth of ≤ 2”)	
Corrective Maintenance (CM)	Mill and AC Overlay (≤ 2”)	59686
	Partial Depth Patching and cover with surface treatment (< 10% Area and Depth of 4-6”)	
	Partial Depth Patching and cover with Thin AC Overlay (< 10% Area and Depth of 4-6”; overlay ≤ 2”)	
	Moderate Patching (< 10% Area and Depth of 6”)	
Restorative Maintenance (RM)	Mill and AC Overlay (≤ 4”)	153229
	Heavy Patching (< 20% Area, Full Depth Patch, Depth of 12”)	
	Full Depth Patching with AC Overlay (< 20% Area, Full Depth Patch, Depth of 9-12”; 4” overlay)	
Rehabilitation / Reconstruction (RC)	Mill, Break, and Seat and AC Overlay (9-12” overlay)	480494
	Reconstruction	

TABLE 5 TREATMENTS CONSIDERED

Treatment	PCI		PCI Reset after apply treatment	Cost per lane meter
	Min	Max		
Do Nothing (DN)	90	100		
Preventive Maintenance (PM)	60	90	PCI=PCI+10	4.33
Corrective Maintenance (CM)	40	60	PCI=PCI+20	37.10
Restorative Maintenance (RM)	25	40	100	95.23
Rehabilitation / Reconstruction (RC)	0	25	100	298.62

Maximum ‘‘Benefit” approach is typically considered for planning of capital investment projects. General benefits include better chances of making correct decisions, improved interagency coordination, and better use of technology; specific benefits, such as justification of programs, would accrue primarily to elected representatives and senior management. Several benefits include tort liability, decrease in travel time, improved motorist comfort and safety, decreased or deferred capital expenditures through capital preservation, vehicle operating and maintenance costs, and reduced pavement deterioration rate.

Cost-effectiveness process is a method of economic evaluation and the alternatives can be evaluated based on what is sacrificed (i.e. the cost) to what has been gained (effectiveness). The cost-effectiveness method might be more suitable for long-term projects. The cost-effectiveness evaluation is being performed in two ways based on the view of the economist: the maximum benefit approach and the minimum cost approach. The first way is applied more frequently in capital investment decision-making, while the second way is more appropriate for maintenance cost planning.

3 PROBLEM STATEMENT

Preventing crashes must be achieved in design and maintenance decision making process. At the moment, there is a gap between traffic management and infrastructure management. Crash prevention is a fundamental goal of transportation engineering and it can be an agent for integrating traffic control and infrastructure management. Developing a model for predicting crashes considering infrastructure and traffic characteristics of the road, is the first step to achieve this purpose. The goal of this study is to develop a crash model based on the field data and to fill the mentioned gap. A reliable crash model can be used in infrastructure decision making process and traffic control process. By having a new control strategy, we will have a different distribution of vehicles (loads) over the network and by having that data, infrastructure engineers have a different perspective for pavement maintenance and repair strategy. By using this model, in addition to predicting the severity and collision type of each crash, the number of crashes at each segment can be predicted and planning variables can vary to minimize the number of crashes as low as possible. In the first part of the study, different models and methods were used to obtain the best predictive model for crash severity and collision type based on the different road and traffic characteristics. The second paper focuses on predicting the number of crashes at each segment based on the road and traffic condition. The purpose of the studies is to initially find the relationship of different factor with severity and frequency of crashes and, potentially, provide predicting models for crash severity, crash type and crash frequency.

4 Effects of Pavement Condition and Road Characteristics on Crash Severity and Collision Type Using Different Regression Methods

4.1 Abstract

Enhancing the level of road safety as a consequence of having an appropriate pavement maintenance plan is a fundamental goal of infrastructure management systems. Based on previous studies, deteriorated pavement has an inadequate performance in terms of safety and riding comfort. Although many studies analyzed the factors effecting the crash occurrence and (few studies) the crash severity, the effect of pavement and traffic condition on collision type and crash severity has not been analyzed by considering Critical Condition Index (CCI) value of pavement as an input and using various regression methods such as Naive Bayes Classifier, Bootstrap Forest, and Neural Network. This study fills the gap by analyzing the effect of different road and traffic characteristics (including CCI) on collision type of crash and crash severity. Effective factors are; Critical Condition Index (CCI), Road Curvature, Pavement Surface Condition, Lighting Condition, AADT, IRI, Rutting Depth (RD), and Heavy Vehicle Percentage. The scope of study is the segment of I-95 in Northern Virginia between the year 2011 to 2016. Different modeling methods including Nominal Logistic Regression, Naive Bayes Classifier, Bootstrap Forest and Neural Network were used to obtain the best results. Results show that the Bootstrap models provide the best fit for the data and quality of pavement might be effective on crash severity and collision type.

Keywords: Transportation Safety, Crash Severity Analysis, Collision Type, International Roughness Index (IRI), Rutting Depth (RD), Critical Condition Index (CCI), Nominal Logistic Model, Naive Bayes Classifier, Bootstrap Forest, Neural Network, Interstate 95, Virginia.

INTRODUCTION

Increasing safety is the most important objective for traffic engineers and it has always been a trending subject for academic communities. Based on previous studies, human factors are the most effective component of each crash (more than 90% effective in crash occurrence), but pavement condition is also a considerable factor of a crash. Pavement condition often impacts human driving behavior, which can lead to significant changes in the number of crashes. One study investigated the effect of pavement condition on crash severity levels using a discrete model that can handle ordered data. The mentioned paper focuses on the effect of poor pavement condition on crash severity levels at low/medium/high speed roads and single/multiple collision cases, using Bayesian Ordered Logistic models. The results showed that the severity of single-vehicle crashes decrease at the poor pavement condition at low-speed roads, and they increase at high speed roads. In addition, the poor pavement condition increases the severity of multiple-vehicle crashes on all roads. This study focuses on the effect of different road and traffic characteristics in addition to Critical Condition Index (CCI) on crash severity. This is important for Department of Transportations (DOTs) and agencies that are working with CCI (e.g., Virginia DOT uses CCI to describe pavement quality of its network) so they can use the result of this work to define their safety thresholds based on CCI values. The result of this work will help decision makers to prioritize projects easier based on safety thresholds.

The main parameters used to measure pavement conditions are Rutting Depth (RD), International Roughness Index (IRI), and Percentage of Fatigue Cracking Area. Rutting is defined as roadway deformation along the wheel path and is measured in inches or millimeters. In a case that rutting depth is deep, it is hard and unsafe to get out of the rut path for the driver especially when the vehicle speed is high. Rutting can be more hazardous in wet weather when water accumulates in the rut paths. Previous studies show that 0.30 in. (7.6 mm) of Rutting Depth is the threshold that increases crashes significantly [Marc R. et al., 1998].

IRI was created by the World Bank in 1980s. IRI is the most common and standard roughness measurement. It is used to indicate the characteristics of longitudinal profile for a traveled wheel path. “The IRI is a filtered ratio of a vehicle’s accumulated suspension motion in inches divided by the distance traveled by the vehicle in miles.” Consequently, the unit of IRI is inch per mile or meter per kilometer. IRI is the most common pavement roughness measure in the United States. When the pavement roughness increases, the contact area between wheels and pavement will decrease, and it leads to a lower brake friction. Rough pavement reduces riding comfort and it can cause damages to vehicle’s suspension system. “Pavement roughness is also the main cause of load loss crashes.”

Fatigue cracking is one of the major distress kinds considered in asphalt-surfaced pavement. In 1955, Hveem demonstrated the concept that fatigue cracking has a higher probability to occur on an asphalt pavement that is experiencing a larger deflection and a higher loading frequency [Hveem et al., 1955]. Fatigue cracking is measured as a ratio of accumulated area with fatigue cracking divided by total area of the pavement surface. Thus, it is unitless and it is a percentage.

There are other asphalt pavement distress types such as bleeding, patching, raveling, stripping, etc., but they are not as common as rutting fatigue cracking and roughness or they are a subcategory of those main distresses.

4.3 CCI Significance

VDOT developed a set of Pavement Performance Prediction Models (PPPM) in units of CCI as a function of time and category of the last maintenance and repair (M&R) activity applied. CCI is an index, indicating asphalt pavement performance, ranging from 0 (complete failure) to 100 (perfect pavement) that represents the minimum of load-related or non-load-related distresses. VDOT classified M&R activities into five categories: (0) Do Nothing (DN), (1) Preventative Maintenance (PM), (2) Corrective Maintenance (CM), (3) Restorative Maintenance (RM), and (4) Reconstruction/Rehabilitation (RC). Using the base form corresponding to the equation below, VDOT defines PPPM for CM, RM and RC. The coefficient of VDOT's load related PPPM expressed through Equation I for asphalt pavements of Interstate highways are presented in Table 6.

CCI(t)=CCI0-e^(a+b*c^(ln(1/t)))

Table 6 - Calibrated Coefficients for CCI Equation

M&R activity	CCI0	a	b	c
CM	100	9.176	9.18	1.27295
RM	100	9.176	9.18	1.25062
RC	100	9.176	9.18	1.22777

Where CCI(t) is the CCI in year t since the last M&R activity (i.e., CM, RM or RC); CCI0 is the Critical Condition Index immediately after treatment; and a, b, and c are the load related PPPM coefficients.

As it was mentioned above, CCI is calculated as the minimum of the LDR and NDR. These indices were first developed in 1998 based on the PAVER methodology developed by the US Army Corps of Engineers and have been validated by using the Long-Term Pavement Performance (LTPP) data collected through the Strategic Highway Research Program (SHRP). Critical Condition Index is a numerical index between 0 and 100 that shows the state of a pavement it is the minimum of two different ratings; Load-Related Distresses Rating (LDR) and Non-Load-Related Distresses Rating (NDR). LDR is a function of loads (the traffic) that pass the road and NDR is a function of factors such as weather condition, chemical properties of the pavement and radiation. The minimum of those two different indexes is CCI. Virginia Transportation Research Council (VTRC) and Virginia Tech Transportation Institute (VTTI) have developed a model using Modified Structural Index (MSI) to calculate LDR. The MSI was found to be a significant input parameter that affects the rate of deterioration of a pavement section. “The Akaike Information Criterion (AIC) suggesting that the model that includes the MSI is, at least, 50,000 times more likely to be more accurate to the true model than the model that does not include the MSI” [Katicha et al., 2016]. Accurate pavement performance prediction can significantly help pavement managers to plan for their further projects and budget prioritization. Consequently, CCI is an interesting case of study specially for a state like Virginia that is using CCI as the main criteria for pavement quality.

4.4 Virginia Pavement Data Collection, Processing And Quality Control

Virginia Department of Transportation’s contractor, Furgo-Roadware Inc. collects pavement data by using continuous digital imaging and automated crack detection technology. They use vans equipped with cameras and sensors to collect data of the roads and their shoulders. Roughness and Rutting are captured with sensors and cracks are detected after computations on continuous images. Other type of distresses such as patching, delamination and bleeding, are identified by analyzing images.

As it was mentioned earlier, CCI is calculated as the minimum of the LDR and NDR. These indexes were derived based on the paver methodology developed by US Army Corps of Engineers in 1998. There is a category for CCI describing pavement condition that is shown below.

TABLE 7 - PAVEMENT QUALITY DESCRIPTION BASED ON CCI VALUE FROM VA STATE OF THE PAVEMENT 2018

Pavement Condition	Index Scale (CCI)
Excellent	90 and above
Good	70-89
Fair	60-69
Poor	50-59
Very Poor	49 and below

Despite that there is a category to define the quality of pavement based on the CCI, but there is no threshold for the CCI in terms of safety. Thus, project prioritization is mostly based on budget priorities rather than safety issues. It means that there is no limitation for CCI value, and if there was a threshold it would be easier for decision makers and stakeholders to make decisions of pavement rehabilitation.

4.5 METHODOLOGY

Considering the significance of safety in transportation and concerns of Virginia Department of Transportation (VDOT) to increase safety and reduce the severity of crashes as much as possible, our team decided to work on the relationship of different road characteristics, especially CCI value of the pavement, and the crash severity. The first step was to analyze Virginia Crash Data (from Virginia SmarterRoads.org Database) to choose an appropriate case for our framework. Almost 5% of all crashes that happened in all the roads in Virginia network, happened in Interstate 95, and around half of those crashes occurred in Northern Virginia District. Consequently, I-95 was selected as our case study.

TABLE 8 - TOP 5 ROUTES WITH MOST CRASHES IN VA (FROM SMARTERROADS DATABASE)

Route	Total Number of crashes from 2010 to 2017	% in VA network
I-95	36333	4.599975692
I-64	29474	3.731585158
I-81	18573	2.351453184
I-66	15947	2.018985835
I-495	8143	1.030952634
	Virginia Total Crashes from 2010 to 2017	789852

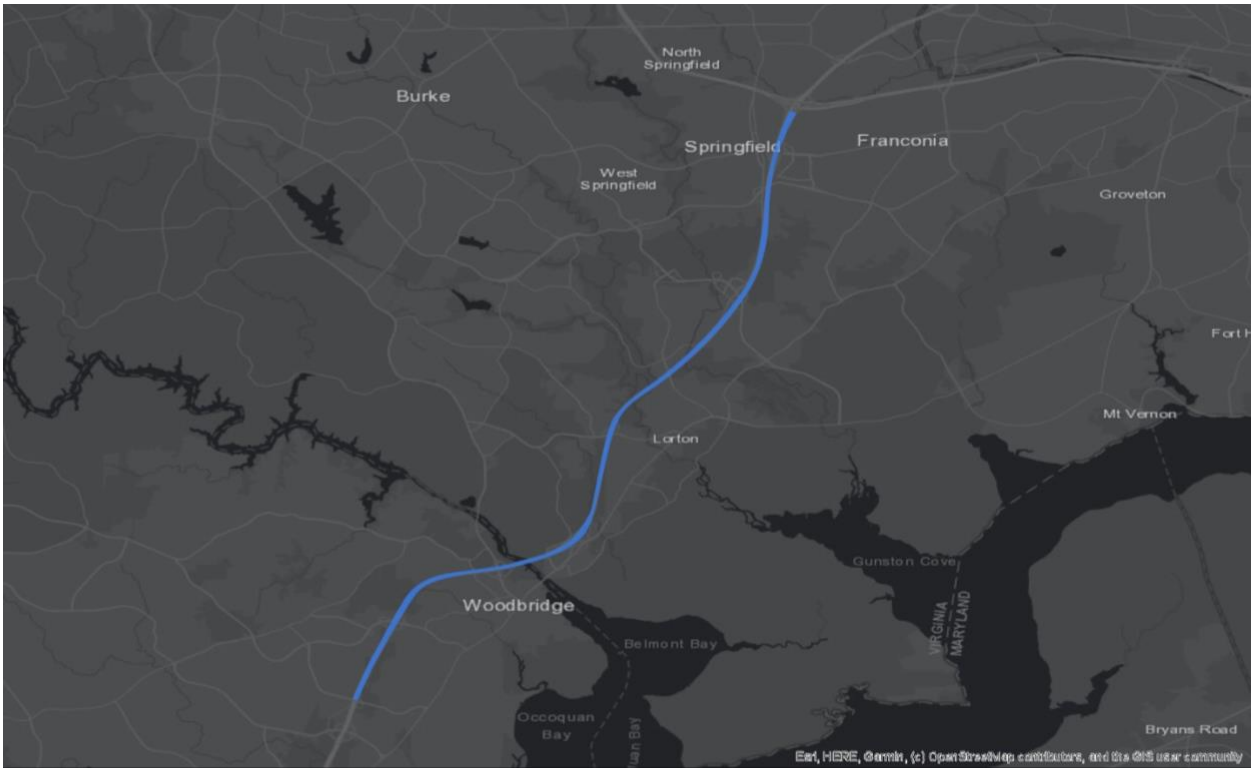

FIGURE 2 – THE SCOPE OF STUDY IN INTERSTATE 95 IN NORTHERN VIRGINIA DISTRICT (FROM ARCGIS)

The figure above shows the segment of I-95 in Northern Virginia District. The length of the road for the purpose of this study is 14 miles in each direction (28 miles total). The road is divided into smaller segments and each segment has its corresponding AADT, Heavy Vehicle Percentage, CCI, IRI and RD. The pavement data is from Pavement Management System (PMS) database and the traffic data is from VDOT traffic volume database.

Virginia Smarter Roads database has the crash data, and the crash data includes many attributes such as exact date and time of the crash, crash location, crash severity, collision type, number of vehicles involving the crash, road surface and lighting condition, and etc. The spatial pavement and traffic attributes of each crash (from PMS dataset and VDOT dataset) were assigned to the crash based on the geographical location of the crash. For understanding the effect of different characteristics of the road on crash severity and crash collision type, initially, Nominal Logistic Models were developed.

According to the description, “Nominal logistic regression models the relationship between a set of predictors and a nominal response variable. A nominal response has at least three groups which do not have a natural order, such as scratch, dent, and tear.” (J, Frost) In the case of crash severity or collision type analysis, the response is a nominal variable such as; “Property Damage Crash”, “Fatal Crash” or “Rear-End Collision”. Consequently, for the purpose of this analysis, nominal logistic models were developed. Table below shows the different categories of responses in the dataset;

TABLE 9 - CRASH SEVERITY CATEGORIES IN DATASET

Crash Severity
Fatal Crash
Injury Crash
Pedestrian Fatal Crash
Pedestrian Injury Crash
Property Damage Crash

TABLE 10 - COLLISION TYPE CATEGORIES IN DATASET

Collision Type
Rear End
Deer
Other Animals
Pedestrian
Backed Into
Other
Angle
Head On
Sideswipe - Same Direction
Sideswipe - Opposite Direction
Fixed Object in Road
Non-Collision
Fixed Object - Off Road

Other methods that were used to conduct the analysis are; Naive Bayes Classifier, Bootstrap Forest and Neural Network Modeling. Brief descriptions of the methods are provided below;

“Nominal logistic regression is a classification method that generalizes logistic regression to multiclass problems, i.e. with more than two possible discrete outcomes.” (Greene 2012)

Naive Bayes is described as “A simple technique for constructing classifiers: models that assign class labels to problem instances, represented as vectors of feature values, where the class labels are drawn from some finite set. There is not a single algorithm for training such classifiers, but a family of algorithms based on a common principle: all naive Bayes classifiers assume that the value of a particular feature is independent of the value of any other feature, given the class variable. For instance, a fruit may be considered to be an apple if it is red, round, and about 10 cm in diameter. A naive Bayes classifier considers each of these features to contribute independently to the probability that this fruit is an apple, regardless of any possible correlations between the color, roundness, and diameter features. For some types of probability models, naive Bayes classifiers can be trained very efficiently in a supervised learning setting. In many practical applications, parameter estimation for naive Bayes models uses the method of maximum likelihood.” (Zhang 2004)

The bootstrap is a computer-based, non-parametric method for assessing the accuracy of almost any statistical estimate. It is mostly used where analytic methods are impractical. In Bootstrap Forest method Maximum Parsimony (MP) is used to select the best solution (tree) instead of calculating the likelihood.

“A neural network is an interconnected assembly of simple processing elements, units or nodes, whose functionality is loosely based on the animal neuron. The processing ability of the network is stored in the interunit connection strengths, or weights, obtained by a process of adaptation to, or learning from, a set of training patterns.” (Gurney 1997)

For each method models try to predict the number of each crash severity and collision type categories as accurate as possible. The tables below show the numbers of crashes at each category through the years of analysis;

TABLE 11 - CRASH SEVERITY THROUGH THE YEARS OF ANALYSIS

Crash Severity	2011	2012	2013	2014	2015	2016	Total
Fatal Crash	2	1	5	3	0	2	13
Injury Crash	285	260	272	326	259	288	1690
Pedestrian Fatal Crash	0	1	0	0	0	2	3
Pedestrian Injury Crash	0	2	2	1	1	3	9
Property Damage Crash	525	521	687	817	789	813	4152
Total	812	785	966	1147	1049	1108	5867

TABLE 12 - COLLISION TYPE THROUGH THE YEARS OF ANALYSIS

Collision Type	2011	2012	2013	2014	2015	2016	Total
1. Rear End	494	470	592	678	658	757	3649
10. Deer	7	5	6	6	3	5	32
11. Other Animal	0	0	1	0	0	0	1
12. Ped	0	2	1	1	0	4	8
15. Backed Into	1	3	0	2	3	2	11
16. Other	4	5	8	9	7	5	38
2. Angle	45	60	29	62	38	29	263
3. Head On	4	3	2	1	6	0	16
4. Sideswipe - Same Direction	133	138	210	230	191	184	1086
5. Sideswipe - Opposite Direction	0	2	0	2	1	1	6
6. Fixed Object in Road	1	2	2	13	18	4	40
8. Non-Collision	6	2	5	8	3	4	28
9. Fixed Object - Off Road	117	93	110	135	121	113	689
Total	812	785	966	1147	1049	1108	5867

4.6 RESULTS

4.6.1 NOMINAL LOGISTIC MODELS

FOR CRASH SEVERITY ANALYSIS, RESULTS SHOW THAT THE MODEL IS STATISTICALLY SIGNIFICANT FOR PREDICTING THE SEVERITY OF CRASH, AND THE FOLLOWING FACTORS ARE STATISTICALLY SIGNIFICANT; COLLISION TYPE, WORK ZONE, CRASH HOUR, NUMBER OF VEHICLES (IN CRASH), SURFACE CONDITION, RD AND CCI. THE SECOND NOMINAL LOGIT MODEL WAS DEVELOPED TO PREDICT COLLISION TYPE OF EACH CRASH BASED ON THE DIFFERENT ATTRIBUTES. THE RESULTS SUGGEST THAT THE MODEL IS STATISTICALLY SIGNIFICANT AND FOLLOWING FACTORS ARE STATISTICALLY SIGNIFICANT TO PREDICT THE COLLISION TYPE OF CRASH; CRASH HOUR,

Number of Vehicles (in crash), Work Zone, Surface Condition, Light Condition, Crash Location (the position on the road), CCI, Percentage of Heavy Vehicle, and IRI.

Table 14 - Crash Severity and Collision Type Effect Likelihood Ratio Tests shows the effect summary of both crash severity and collision type nominal logistic models.

Table 14 - Crash Severity and Collision Type Effect Likelihood Ratio Tests shows the likelihood ChiSquare test of the input factors.

Table 13 - Crash Severity and Collision Type Effect Summary

Crash Severity			Collision Type		
Source	**LogWorth**	**PValue**	**Source**	**LogWorth**	**PValue**
COLLISION_TYPE	1473.652	0	HOUR	4503.865	0
WORK_ZONE	34.99	0	NO_OF_VEHI	141.652	0
HOUR	32.369	0	WORK_ZONE	130.726	0
NO_OF_VEHI	22.532	0	SURFACE_COND	127.458	0
SURFACE_COND	4.345	0.00005	LIGHT_COND	65.515	0
RD	1.71	0.01951	LOCATION_O	42.984	0
CCI	1.376	0.04203	CCI	2.635	0.00232
IRI	0.89	0.1287	HV%	2.26	0.00549
AADT	0.657	0.22018	IRI	1.142	0.07219
HV%	0.36	0.43603	RD	0.776	0.16737
IF_CURVE	0.116	0.76611	AADT	0.292	0.51065
LIGHT_COND	0.007	0.98363	IF_CURVE	0.272	0.53455

Table 14 - Crash Severity and Collision Type Effect Likelihood Ratio Tests

Crash Severity					Collision Type				
Source	**Nparm**	**DF**	**L-R ChiSquare**	**Prob>ChiSq**	**Source**	**Nparm**	**DF**	**Wald ChiSquare**	**Prob>ChiSq**
WORK_ZONE	20	14	203.5705	<.0001*	WORK_ZONE	60	18	673.974655	<.0001*
NO_OF_VEHI	4	4	111.8464	<.0001*	NO_OF_VEHI	12	12	701.386321	<.0001*
LIGHT_COND	24	14	5.149236	0.9836	LIGHT_COND	72	28	394.108813	<.0001*
SURFACE_COND	28	9	35.67942	<.0001*	SURFACE_COND	84	31	704.021362	<.0001*
HOUR	92	84	343.5188	<.0001*	LOCATION_O	72	15	245.521144	<.0001*
AADT	4	4	5.730766	0.2202	HOUR	276	148	21610.5598	<.0001*
HV%	4	3	2.724799	0.436	AADT	12	12	11.2143039	0.5106
CCI	4	4	9.906529	0.0420*	HV%	12	11	26.484654	0.0055*
IRI	4	4	7.139374	0.1287	CCI	12	12	30.537371	0.0023*
RD	4	3	9.891471	0.0195*	IRI	12	12	19.7388137	0.0722
COLLISION_TYPE	48	22	6919.193	<.0001*	RD	12	11	15.3418704	0.1674
IF_CURVE	4	4	1.83478	0.7661	IF_CURVE	12	12	10.9344006	0.5345

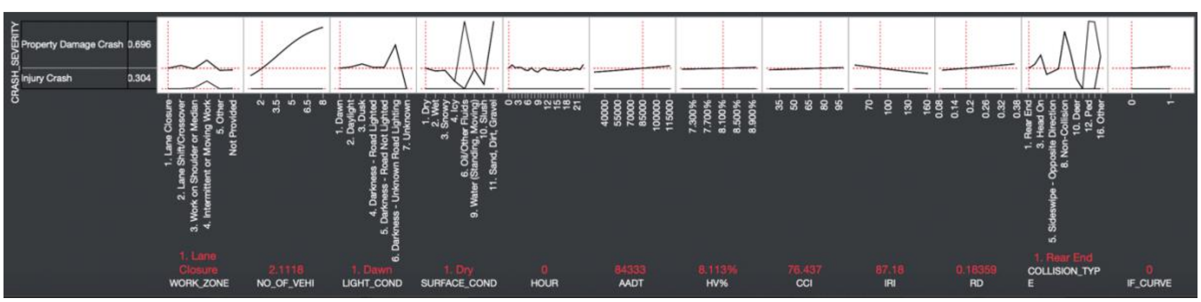

Figure 3 - Crash Severity Profiler for Nominal Logistic Model

As it is shown in the figure above, percentage of injury crashes increases by increasing the Rutting Depth. It can be concluded that; there are potentially more injury crashes in the roads with higher values of rutting depth and deep deformation of the pavement might cause the driver to lose the control over the vehicle in case of a crash (or a hazardous situation).

Based on the profiler graphs for crash severity, percentage of injury crashes considerably increased by increasing the number of vehicles in crash. It means that the more vehicles in a crash, the more chance of having injuries for the users.

The percentage of Injury Crashes increases in higher CCI values. It can be argued that when the quality of the road is higher, drivers tend to speed up and the chance of having injury in a crash increase.

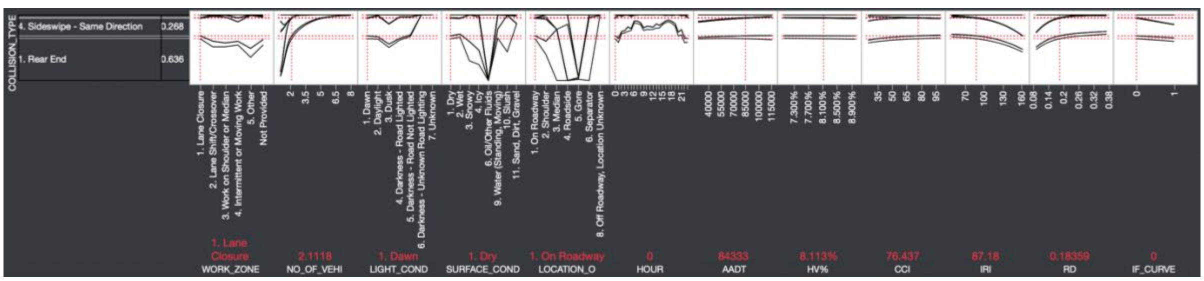

FIGURE 4 - COLLISION TYPE PROFILER FOR NOMINAL LOGISTIC MODEL

As it is shown in the collision type profiler graphs, and as it was expected, percentage of Rear-End crashes increases by increasing Number of Vehicles in crash. Percentage of Rear-End crashes increases by increasing the rutting depth and CCI which can be the cause of more injury crashes in crash severity analysis.

4.6.2 NAIVE BAYES

The tables below are the fit detail of the Naïve Bayes Models for Crash Severity and Collision Type. shows the fit details of Naive Bayes model for Crash severity. As it is shown in the table, RSquare value is low and this method is not the best method for predicting crash severity.

TABLE 15 - JMP FIT DETAILS OF NAIVE BAYES FOR CRASH SEVERITY

Measure	Training	Definition
Entropy RSquare	0.0517	1-Loglike(model)/Loglike(0)
Generalized RSquare	0.0880	(1-(L(0)/L(model))^(2/n))/(1-L(0)^(2/n))
Mean -Log p	0.5980	∑ -Log(ρ[j])/n
RASE	0.4505	√ ∑(y[j]-ρ[j])²/n
Mean Abs Dev	0.3953	∑ |y[j]-ρ[j]|/n
Misclassification Rate	0.2901	∑ (ρ[j]≠ρMax)/n
N	5867	n

TABLE 16 - JMP FIT DETAILS OF NAIVE BAYES FOR COLLISION TYPE

Measure	Training	Definition
Entropy RSquare	0.3592	1-Loglike(model)/Loglike(0)
Generalized RSquare	0.6279	(1-(L(0)/L(model))^(2/n))/(1-L(0)^(2/n))
Mean -Log p	0.7461	∑ -Log(ρ[j])/n
RASE	0.4815	√ ∑(y[j]-ρ[j])²/n
Mean Abs Dev	0.3829	∑ |y[j]-ρ[j]|/n
Misclassification Rate	0.2724	∑ (ρ[j]≠ρMax)/n
N	5867	n

4.6.3 BOOTSTRAP FOREST METHOD

As it is shown in the tables below, Bootstrap Forest method has a considerably better fitting performance for crash severity analysis in, and it is a great method for Collision Type analysis.

TABLE 17 - JMP FIT DETAILS OF BOOTSTRAP FOREST FOR CRASH SEVERITY

Measure	Training	Definition
Entropy RSquare	0.3698	1-Loglike(model)/Loglike(0)
Generalized RSquare	0.5201	(1-(L(0)/L(model))^(2/n))/(1-L(0)^(2/n))
Mean -Log p	0.3973	∑ -Log(ρ[j])/n
RASE	0.3501	√ ∑(y[j]-ρ[j])²/n
Mean Abs Dev	0.2971	∑ |y[j]-ρ[j]|/n
Misclassification Rate	0.1733	∑ (ρ[j]≠ρMax)/n
N	5867	n

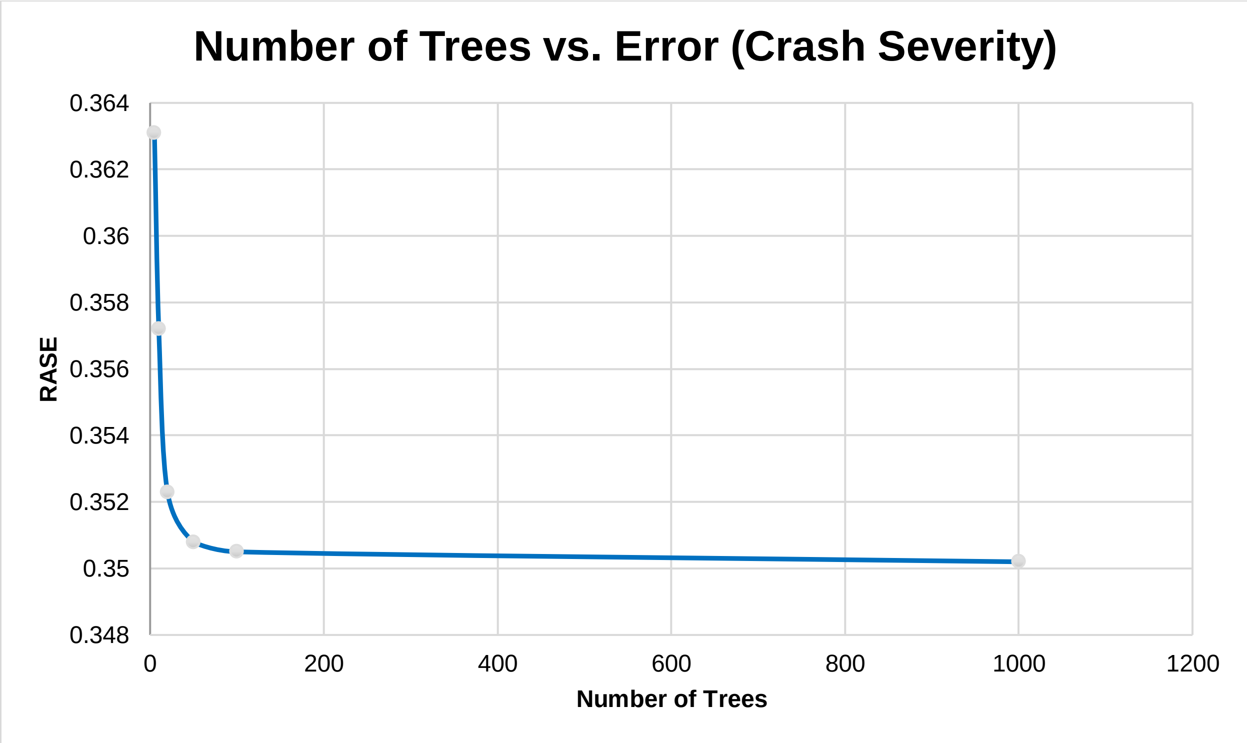

FIGURE 5 - NUMBER OF TREES IN FOREST VS. ERROR FOR CRASH SEVERITY

Figure above shows that after of about 100 trees in the forest, the model improves very slightly but it takes so much processing time for the system.

TABLE 18 - JMP FIT DETAILS OF BOOTSTRAP FOREST FOR COLLISION TYPE

Measure	Training	Definition
Entropy RSquare	0.5377	1-Loglike(model)/Loglike(0)
Generalized RSquare	0.7912	(1-(L(0)/L(model))^(2/n))/(1-L(0)^(2/n))
Mean -Log p	0.5383	∑ -Log(ρ[j])/n
RASE	0.4279	√ ∑(y[j]-ρ[j])²/n
Mean Abs Dev	0.3424	∑ |y[j]-ρ[j]|/n
Misclassification Rate	0.2340	∑ (ρ[j]≠ρMax)/n
N	5867	n

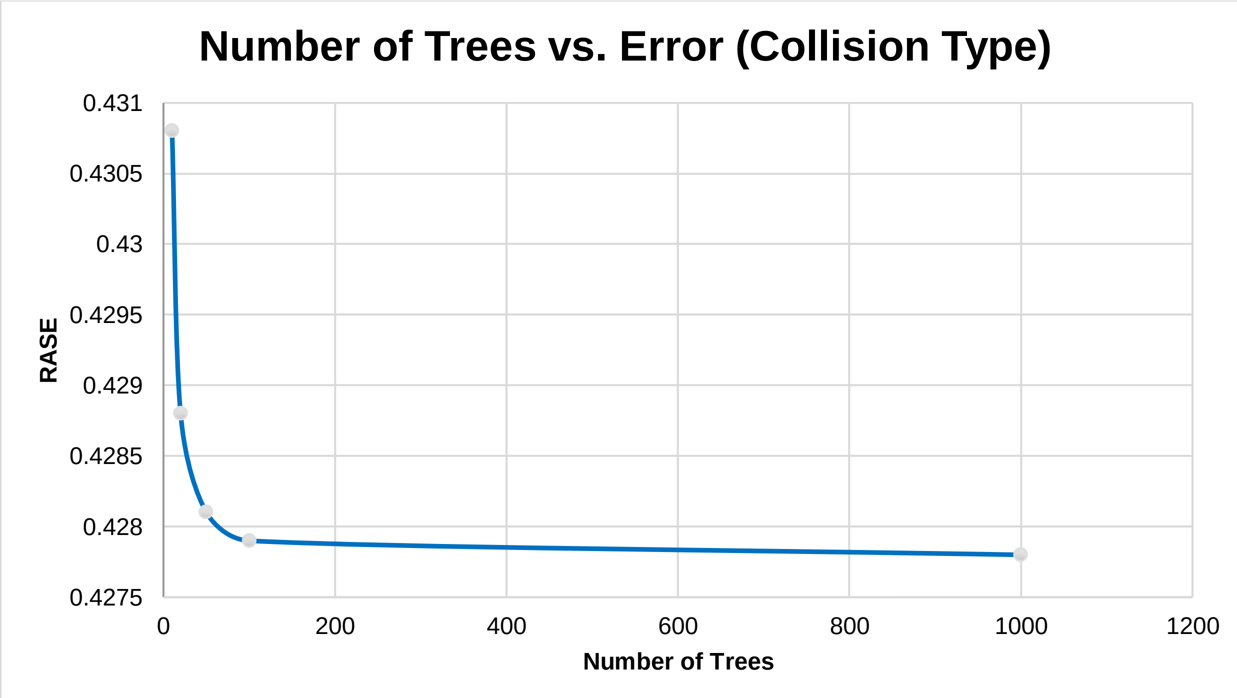

Figure 6 - Number of Trees in Forest vs. Error for Collision Type

As it is shown in figure above, after of about 100 trees in the forest, the model improves very slightly but it takes so much processing time for the system.

4.6.4 Neural Network

As it is shown in the fit detail tables below, Neural Network is not an appropriate method for crash severity analysis (RSquare=0.0767). And for the collision type has an acceptable RSquare, but it is not as good as the Bootstrap method.

Table 19 - JMP Fit Details of Neural Network for Crash Severity

Measures	Value
Generalized RSquare	0.0766924
Entropy RSquare	0.0448622
RMSE	0.4483004
Mean Abs Dev	0.4000326
Misclassification Rate	0.2910486
-LogLikelihood	2350.9652
Sum Freq	3910

TABLE 20 - JMP FIT DETAILS OF NEURAL NETWORK FOR COLLISION TYPE

Measures	Value
Generalized RSquare	0.6062005
Entropy RSquare	0.3407823
RMSE	0.4824401
Mean Abs Dev	0.3936856
Misclassification Rate	0.2646532
-LogLikelihood	2990.2668
Sum Freq	3907

4.6.5 MODEL COMPARISON

The table below shows the generalized RSquare of different models for crash severity and collision type. As it shown in the table, Bootstrap Forest method has the best RSquare values for both crash severity and collision type analysis.

TABLE 21 - GENERALIZED RSQUARE COMPARISON TABLE ALL MODELS

Generalized RSquare Table		Crash Severity	Collision Type
Method	Nominal Logistic	0.070	0.360
	Naive Bayes	0.088	0.628
	Neutral Network	0.077	0.606
	Bootstrap Forest	0.502	0.783

However different methods have different qualities in terms of fit to the data, the area under the receiver operating characteristic (ROC) curves have the same order for different crash severity and collision type categories within different methods. ROC curve is defined as; “A graphical plot that illustrates the diagnostic ability of a binary classifier system as its discrimination threshold is varied.” And the area under the curve “(Often referred to as simply the AUC) Is equal to the probability that a classifier will rank a randomly chosen positive instance higher than a randomly chosen negative one.” (Fawcett, 2005)

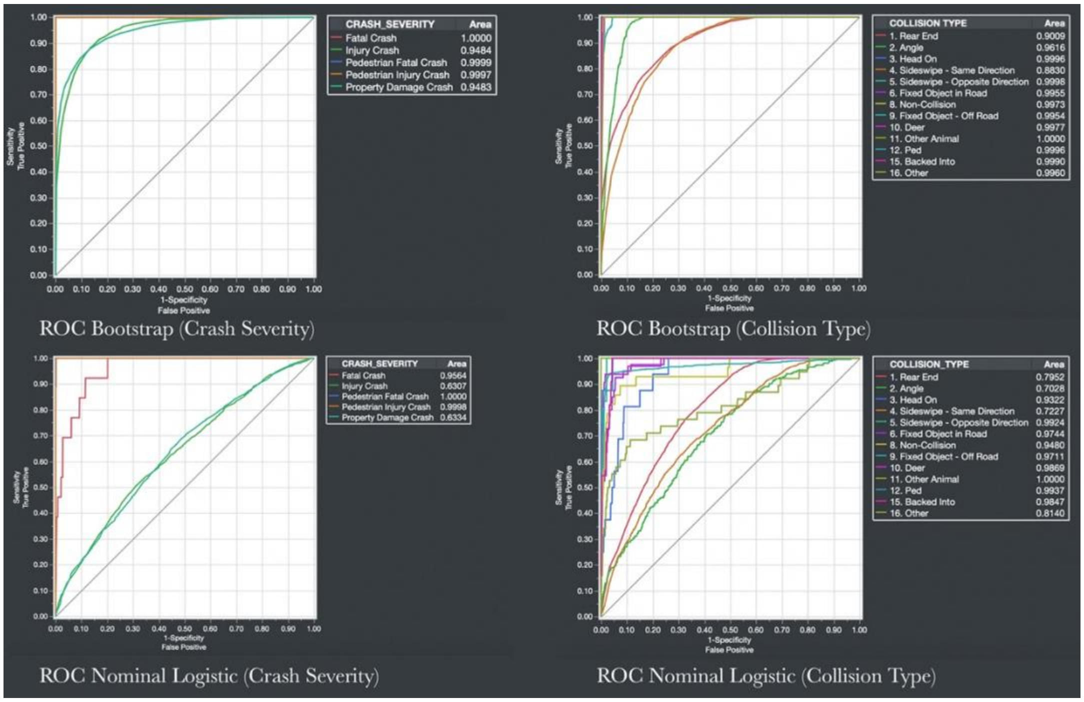

FIGURE 7 - ROC CURVES FOR CRASH SEVERITY AND COLLISION TYPE USING DIFFERENT METHODS

4.7 SUMMARY AND CONCLUSION

Interstate 95 is one of the roads with most crashes in Virginia. Crashes are denser in places that the road has curves. Based on this segment of the road, that there are more crashes in the curves that vehicles are turning to the right side (clockwise) than the curves with vehicles turning to the left side (counterclockwise). This can be a further case study.

Nominal Logistic Regression, Neural Network, Naive Bayes and Bootstrap Forest methods were used to analyze crash severity and collision type. The results suggest that;

- Nominal Logistic Regressions are statistically significant.
- Nominal Logistic Regression for crash severity suggests that Collision Type, Work Zone, Crash Hour, Number of Vehicles (in crash), Surface Condition, RD and CCI are statistically significant to predict the severity of crash.
- Nominal Logistic Regression for crash severity suggests that Crash Hour, Number of Vehicles (in crash), Work Zone, Surface Condition, Light Condition, Crash

Location (the position on the road), CCI, Percentage of Heavy Vehicle, and IRI are statistically significant to predict the collision type of crash.

- The Bootstrap Forest models provide the best fit for both crash severity and collision type analysis.

4.8 Further Works

For future works a model for number of crashes in each segment can be developed and based on the results of this work, percentage of each crash severity of each segment can be predicted. And by having the models, future safety condition of each road can be predicted. Travel demand management plans for the next couple of years have been developed and can be obtained from the DOT. Most of the travel demand models are designed to minimize overall delay of the network, but they are not designed to minimize the pavement deterioration. Investigation on the effect of different travel demand management scenarios on pavement deterioration rate and the impact of it on decreasing crash quantity is a great case of study for further works.

Acknowledgments

This work was funded by the Virginia Transportation Research Council (VTRC). The authors would like to acknowledge and thank Mr. Kevin McGhee, Dr. Mike Fitch, Ms. Elaine Thurman, and Mr. Ibrahim Abuawad for their support and help.

Author Contributions

The authors confirm contribution to the paper as follows: study conception and design: Montasir Abbas, Linbing Wang; analysis and interpretation of results: Ali Mohagheghi; draft manuscript preparation: Ali Mohagheghi and Montasir Abbas. All authors reviewed the results and approved the final version of the manuscript.

References

1. Huang, Y.H., 1993. Pavement analysis and design.

2. Start, M. R., Kim, J. and Berg, W. D. (1998) 'Potential Safety Cost-Effectiveness of Treating Rutted Pavements', *Transportation Research Record*, 1629(1), pp. 208–213.

3. Transportation Research and Board (T.R.B). A State-of-the-Art Report, the Influence of Roadway Surface Discontinuities on Safety, Transportation Research Board, National Research Council Washington D.C. 1984.
4. Horne WB, and Leland TJW. (1962) Influence of Tire Tread Pattern and Runway Surface Condition on Braking Friction and Rolling Resistance of a Modern Aircraft Tire. NASA TN D-1376. National Aeronautics and Space Administration.
5. Quinn BE, and Hilderbrand SE. Effect of road roughness on vehicle steering. Highway Research Record 1973; 47:62–75.
6. Wambold JC, Brickman AD, Park WH, and Ingram J. Effect of road roughness on vehicle braking. Highway Research Record 1973; 471:76–82.
7. T. Nakatsuji, T. Kaku, T. Fujiwara, T. Hagiwara, and Y. Onodera, "Dynamic Behavior of a Vehicle on a Rutted Road," in Surface Characteristics of Roadways: International Research and Technologies, ed. W. Meyer and J. Reichert (West Conshohocken, PA: ASTM I

8. Burns, J.C., 1981. Roughness and roadway safety. *Transportation research record*, *836*, pp.8-14.

9. Calibrating and Validating Overlay Tester–Based Fatigue Cracking Model with Data from National Center for Asphalt Technology, Sheng Hu, Fujie Zhou, Thomas Scullion, and Joe Leidy.
10. Hveem, F. N. *HRB Bulletin 114: Pavement Deflections and Fatigue Failures.* HRB, National Research Council, Washington, D.C., 1955, pp. 43–87.
11. DOT OHIO: PAVEMENT CONDITION RATING SYSTEM.
12. Gharaibeh, Nasir G.; Zou, Yajie; Saliminejad, Siamak (2010): Assessing the Agreement among Pavement Condition Indexes. In J. Transp. Eng. 136 (8), pp. 765–772.
13. Medlock, Ronald D.; Laffrey, David C. (Eds.) (1998): Structural Materials Technology III: An NDT Conference. Non-Destructive Evaluation Techniques for Aging Infrastructure & Manufacturing. San Antonio, TX, Tuesday 31 March 1998: SPIE (SPIE Proceedings).
14. Ram, Prashant V.; Smith, Kelly L.; Zimmerman, Kathryn A.; Ratliff, Tammy; Amenta, James (2016): Developing a Correlation Between the Pavement Condition Ratings Used by Five Federal Lands Management Agencies. In Transportation Research Record 2589 (1), pp.
15. Shah, Yogesh U.; Jain, S. S.; Tiwari, Devesh; Jain, M. K. (2013): Development of Overall Pavement Condition Index for Urban Road Network. In Procedia - Social and Behavioral Sciences 104, pp. 332–341.
16. Virginia State of the Pavement 2016.
17. Virginia State of the Pavement 2017.
18. State of the Pavement 2018.

19. Yin, Y., Wang, Y., Lu, J. and Wang, W., 2011, July. ICCTP 2011. American Society of Civil Engineers.
20. Z. Al-Qadi, Imad L.; Murrell, Scott (Eds.) (06182013): Airfield and Highway Pavement 2013. 2013 Airfield & Highway Pavement Conference. Los Angeles, California, United States, June 9-12, 2013. Reston, VA: American Society of Civil Engineers.
21. Gianluca Dell'Acqua, Fred Wegman, Transport Infrastructure and Systems.
22. Start, M. R., Kim, J. and Berg, W. D. (1998) 'Potential Safety Cost-Effectiveness of Treating Rutted Pavements', Transportation Research Record, 1629(1), pp. 208–213.
23. Samer W. Katicha, Ph.D., Safak Ercisli, Gerardo W. Flintsch, Ph.D., P.E., James M. Bryce, Ph.D., and Brian K. Diefenderfer, Ph.D., P.E. Development of Enhanced Pavement Deterioration Curves, 2016
24. Zhang, Harry. The Optimality of Naive Bayes (PDF). FLAIRS2004 conference.
25. Caruana, R.; Niculescu-Mizil, A. (2006). An empirical comparison of supervised learning algorithms. Proc. 23rd International Conference on Machine Learning.
26. Efron, B., Halloran, E. and Holmes, S., 1996. Bootstrap confidence levels for phylogenetic trees. *Proceedings of the National Academy of Sciences*, *93*(23), pp.13429-13429.
27. Hoang, D.T., Vinh, L.S., Flouri, T. *et al.* MPBoot: fast phylogenetic maximum parsimony tree inference and bootstrap approximation. *BMC Evol Biol* **18,** 11 (2018).
28. Greene, William H. (2012). Econometric Analysis (Seventh ed.). Boston: Pearson Education. pp. 803–806. ISBN 978-0-273-75356-8.
29. Gurney, K., 1997. *An introduction to neural networks*. CRC press.
30. Fawcett, Tom (2006); An introduction to ROC analysis, Pattern Recognition Letters, 27, 861–874.

5 Effects of Pavement Condition and Road Characteristics on Crash Frequency on Interstate System in Virginia

5.1 Abstract

Providing a safe environment for transportation users through a proper pavement maintenance plan is an essential goal of transportation engineers. Improving road safety is one of the most important objectives for pavement management systems. While many studies analyzed the effect of several factors (related to road condition) on crash occurrence, none of which considered Critical Condition Index (CCI) value of asphalt pavement as a factor for crash analysis. This paper analyzed the effect of different road characteristics including; Critical Condition Index (CCI), International Roughness Index (IRI), Rutting Depth (RD), Annual Average Daily Traffic (AADT), and Percentage of Heavy Vehicles (HV%) on crash frequency in the interstate highway system in Fairfax County, Virginia between the years 2016-2018. Negative Binomial, Least Square models were developed for different road lighting conditions and surface conditions. A Nominal Logistic Regressions were conducted to model the crash severity. Results show that CCI value is a significant factor to predict the number of crashes, and it has different behavior within different values of AADT. The number of crashes decrease by increasing the value of CCI (better pavement quality) when the traffic volume is high. This result can help decision makers to plan the pavement maintenance procedure for a potentially safer transportation network.

Keywords: Transportation Safety, Crash Prevention, Crash Pattern, Interstate Highway System, Fairfax County, Virginia, Infrastructure Maintenance, Critical Condition Index, International Roughness Index, Rutting Depth, Negative Binomial Regression, Nominal Logistic Model

5.2 INTRODUCTION

Elevating the level of safety is a fundamental goal of traffic engineering. Designing and maintaining a safe network for users, not only decrease the number of crashes but enhance the user experience by reducing the moments that users feel unsafe. By them, the level of mobility increases. Number of crashes per mile per year is a general expression of the level of safety in that segment. Different segments have different number of crashes per mile per year, even though they are in a same road. It means that there are some factors causing that variation. A large variety of past literature shows that human factors primarily and predominantly impact the crash occurrence. Although, road characteristics and traffic condition are not as effective as human factors in case a crash happens, but it does not mean that improving the road environment is not as necessary as improving technologies to alarm users to avoid any crash or unsafe behavior. Notwithstanding the fact that human factors have the main impact on crash occurrence, but road condition significantly impacts driving behavior, and driving behavior is a part of human factors. A previous study has analyzed the relationship of International Roughness Index (IRI), Rutting Depth (RD) and Present Serviceability Index (PSI) on crash frequency and it shows that PSI has the most significant relationship with crash frequency following by IRI and RD [Chan et al., 2010].

Preventing crashes must be achieved in design and maintenance decision making process. At the moment, there is a gap between traffic management and infrastructure management. Crash prevention is a fundamental goal of transportation engineering and it can be an agent for integrating traffic control and infrastructure management. Developing a model for predicting crashes considering infrastructure and traffic characteristics of the road, is the first step to achieve this purpose. The goal of this study is to develop a crash model based on the field data and to fill the mentioned gap. A reliable crash model can be used in infrastructure decision making process and traffic control process. By having a new control strategy, we will have a different distribution of vehicles (loads) over the network and by having that data, infrastructure engineers have a different perspective for pavement maintenance and repair strategy. By sing this model, the number of crashes in each segment can be predicted and planning variables can vary to minimize the number of crashes as low as possible. To have a better understanding of effective factors, each one of them is described briefly.

The main parameters which define pavement condition are Rutting Depth (RD), International Roughness Index (IRI), and Percentage of Fatigue Cracking Area. Each agency uses a different method to come up with a practical value to define the general quality of pavement such as Present Serviceability Index (PSI), Pavement Condition Index (PCI), Critical Condition Index (CCI), etc. Rutting is defined as roadway deformation along the wheel path and the units for its measurement are inches or millimeters. If the rutting depth is deep, it creates an unsafe condition for the users when they try to get out of the rut path, especially at higher speeds. Rutting can cause more severe safety problems in case of wet weather when water accumulates in the rut paths. Previous studies indicate that 0.30 in. (7.6 mm) of Rutting Depth is the threshold that increases crashes significantly [Marc R. et al., 1998]. International Roughness Index (IRI) was developed by the World Bank in 1980s. Most of agencies around the world use IRI as an indicator for pavement condition. It indicates the characteristics of longitudinal profile for a traveled wheel path. “The IRI is a filtered ratio of a vehicle’s accumulated suspension motion in inches divided by the distance traveled by the vehicle in miles.” Common units of IRI are inch per mile and meter per kilometer. IRI is the most common pavement roughness measure in the United States as well as most countries all around the world. When the pavement roughness increases, the area of contact between wheels and pavement decreases, and the situation leads to a lower brake friction. Driving comfort level decrease when a pavement has a high IRI value rough pavement can cause damages to vehicle's suspension system. "Pavement roughness is also the main cause of load loss crashes." Fatigue cracking is another common distress type that appears on the asphalt-surface. Fatigue cracking is measured as the ratio of accumulated cracking area divided by total area of the pavement surface. Hence, it is unitless and it is a percentage. There are other asphalt pavement distresses such as bleeding, patching, raveling, stripping, etc., they are not as common as the three main distresses.

5.3 BACKGROUND

There are several works related to the subject of this study, but the contribution of this study to the literature is that Critical Condition Index (CCI) is considered as a factor in the analysis, and also the location of study is a great model for all metropolitan areas in the US. Due to the variation of methods for crash analysis, this work provides different analysis methods and compare the results. CCI is calculated as the minimum of the Load-Related Distress Rating (LDR) and Non-Load-Related Distress Rating (NDR). These indices were first developed in 1998 based on the PAVER methodology developed by the 15 US Army Corps of Engineers and have been validated by using the Long-Term Pavement 16 Performance (LTPP) data collected through the Strategic Highway Research Program (SHRP). CCI is a numerical index between 0 and 100 that indicates the quality of pavement. A healthy pavement that is just paved, has the CCI value of 100 and the lover the quality gets, the lover the CCI value get. LDR is a function of loads (the traffic) that pass the road and NDR is a function of factors such as weather condition, chemical properties of the pavement and radiation. A model has been developed to predict CCI value and it has Modified Structural Index (MSI) as an input "The Akaike Information Criterion

(AIC) suggesting that the model that includes the MSI is, at least, 50,000 times more likely to be more accurate to the true model than the model that does not include the MSI" [Katicha et al., 2016]. Accurate and reliable models for CCI makes the contribution of this work more significant in terms of future transportation design and infrastructure maintenance and repair planning. Equation 1 is the CCI model;

$$CCI_{Model} = 100 - T^{\beta_2} e^{\left(\beta_0 + \beta_1 \frac{1}{MSI^4} + \beta_3 T\right)}, \tag{9}$$

where CCI_{Model} is the predicted value of Critical Condition Index (CCI at T=0 is 100),

T is time in years,

MSI is Modified Structural Index value (Appendix)

$\beta_0, \beta_1, \beta_2$ and β_3 are calibration factors

Washington, DC metropolitan area is one of the most congested areas in the United States. Fairfax County is the most populated area in metropolitan Washington, DC, it has more population than the District of Colombia itself. Public transportation in the area is limited and interstate highway system, especially in Fairfax County, is serving most of the travelers, hence, it is very congested, and number of crashes are considerable. Due to considerable number of traffic crashes and high percentage of injury crashes, the interstate system has been considered as the scope of this study. The period of study is from 2016 to 2018. Table below shows statistics about Roads in Fairfax County Interstate Highway system.

TABLE 22 - FAIRFAX INTERSTATE HIGHWAY SYSTEM HIGHLIGHTS

Road Name	Total Length	Number of Crash	AADT	Study Period (Years)	Crash/Mile/Year/ AADT (in 10000)
I-66	37.94	3945	71000	3	4.881688656
I-95	33.12	2386	102000	3	2.354283098
I-395	12.32	471	86000	3	1.481803081
I-495	29.18	3083	94000	3	3.746615529
Total	112.56	9885			12.46439036

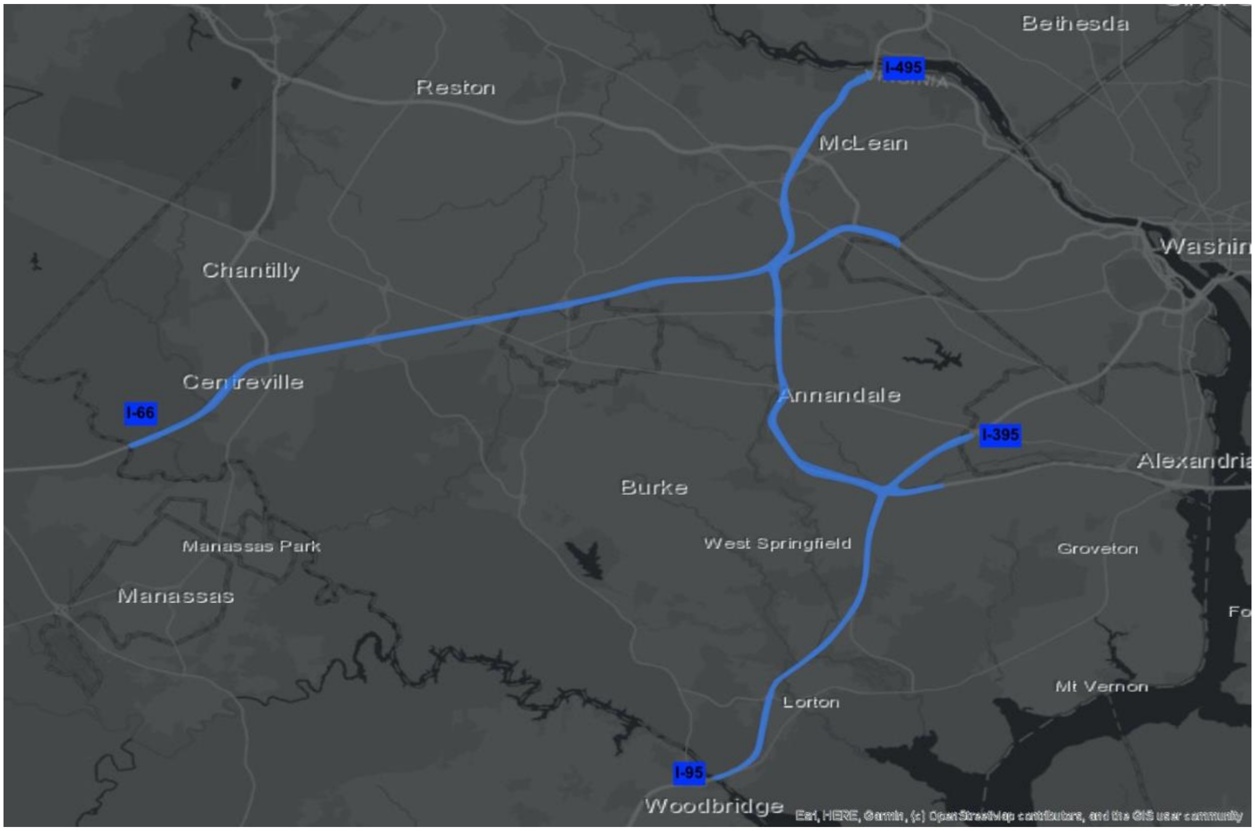

FIGURE 8 - INTERSTATE HIGHWAY SYSTEM IN FAIRFAX COUNTY, VIRGINIA

A key factor to understand the data, is to understand the data collection process. Virginia Department of Transportation's contractor, Furgo-Roadware Inc. is in charge of collecting pavement data by using continuous digital imaging and automated crack detection technology. The data is collected by vans equipped with cameras and sensors. Roughness and Rutting are captured with sensors and cracks are detected after computations on continuous captured images. Other type of distresses such as patching, delamination and bleeding, are identified by analyzing images.

As it was mentioned, CCI is the minimum of the Load-Related Distress Rating (LDR) and Non-Lod-Related Distress Rating (NDR). The distresses along wheel paths are considered as LDR and all other distresses in other locations but the wheel paths are considered as NDR.

TABLE 23 - PAVEMENT QUALITY DESCRIPTION BASED ON CCI VALUE FROM VA STATE OF THE PAVEMENT 2018

Pavement Condition	Index Scale (CCI)
Excellent	90 and above
Good	70-89
Fair	60-69
Poor	50-59
Very Poor	49 and below

5.4 METHODOLOGY

Crashes are random events that are dependent to many factors that some of which factors are impacted by one another. Human factors and vehicle performance have the main impact on a potential crash and most of the research and studies are related to them. The amount of work about the analysis of road environment itself (not road design) on traffic crashes are considerably lower than other studies related to the crash. Most of previous studies have used Negative Binomial Regression method to find the relation of road characteristics and number of crashes. In this study, Least Square Regression method, additional to the Negative binomial, is used to have a better understanding of the performances of each input. To study different attributes of each crash, a Nominal Logistic Model was developed.

Before developing the models, segments were highlighted, and crashes were mapped on segments in order to visualization of the site and check if all the points are on the highway system. This process was completed in ArcGIS (Figure 7). A geo-density map was developed in MATLAB to demonstrate the density of crashes along the roads (Figure 8).

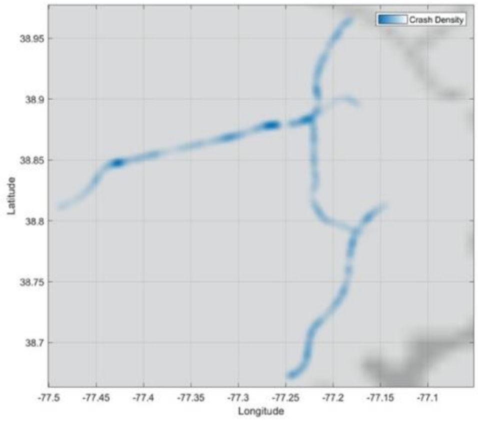

FIGURE 9 - CRASH GEO-DENSITY MAP

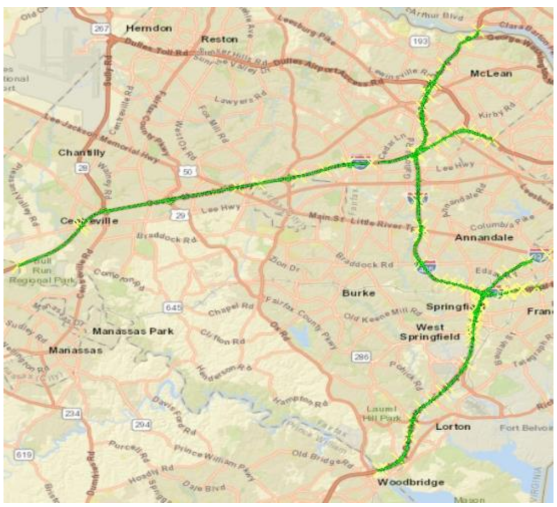

FIGURE 10 - CRASHES WITHIN THE SEGMENTS

As it is shown in the geo-density plot, events are denser in I-66, especially in curves and near exits and entrances. Figure 11 shows the frequency of segments within AADT vs. CCI graph (heatmap of the data along different AADT and CCI values). Figure 12 s the contour map of the crash frequency in the AADT vs. CCI graph.

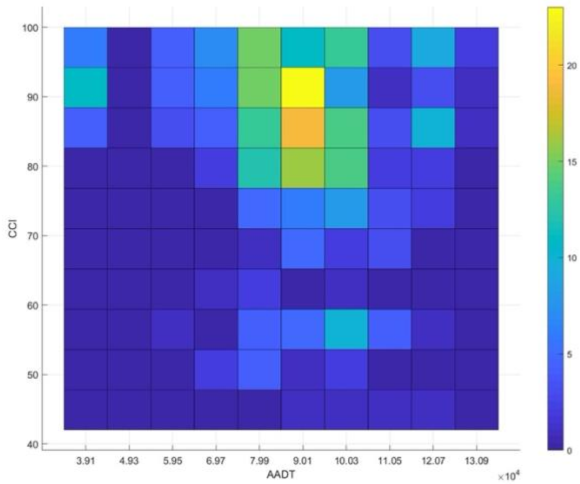

FIGURE 11 - AADT VS. CCI HEATMAP

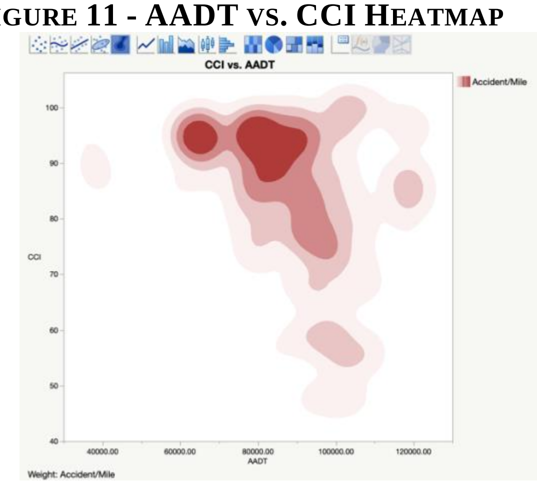

FIGURE 12 - AADT VS. CCI CONTOUR MAP (NUMBER OF CRASHES AS FREQUENCY)

Studies mainly use Poisson regression and Negative Binomial regression for modeling and studying traffic safety matters. Crashes are discrete and non-negative events; hence, Poisson regression is an appropriate method for crash analysis. Poisson regression can be used under the assumption of the conditional mean being equal to the conditional variance for the dependent variable. When the data is over dispersed and conditional variance exceeds the conditional mean, Negative Binomial Regression is an appropriate method to develop a model. In a case of overdispersion, Poisson regression cannot be used due to avoid any standard error in the model. The formulation of Poisson regression is provided below;

$$P(y_i) = \frac{\lambda_i^{y_i} e^{-\lambda_i}}{y_i!}, \tag{10}$$

where $P(y_i)$ is the probability of y number of crashes happen on road section i per year. In Poisson regression, λ_i is both mean and variance of y_i. So λ_i is estimated number of crashes and here is the formulation for λ_i;

$$\lambda_i = e^{\beta X_i}, \tag{11}$$

where X_i is independent variable I and β is coefficient of independent variable. To choose the best regression method, the data should be described properly. The crash record was obtained from Virginia Roads open dataset. It is a reliable and very detailed dataset for crashes in Virginia. The data includes exact time and location of each crash, collision type, lighting and weather condition, crash severity, and many other attributes of each crash. Pavement Management System (PMS) has all the pavement records for all Virginia network. Each road is divided into small segments and each segment has its own starting and ending location, a unique ID, and precise pavement information collected from the field. AADT is also available through Virginia Department of Transportation dataset and like pavement data road segments has specific attributes related to them. The total value of Crashes per mile for each segment in different lighting and weather condition were calculated and a new integrated dataset was developed. Table 21 is the descriptive statistics of crashes/mile in a period of a year for all segments.

TABLE 24 - DESCRIPTIVE STATISTICS OF CRASHES

Dependent Variables	Min	Max	Mean	SD	Variance
All Crash	0.394477318	335.483871	26.63912287	34.61346719	1198.092111
Daylight Crash	0	290.3225806	18.63644736	26.79075235	717.7444114

Nighttime Crashes	0	63.47804259	8.002675513	9.634536812	92.82429958
Dry Road Crash	0	306.4516129	22.90722311	30.95646022	958.3024291
Non-Dry Road Crash					
Independent Variables	0	29.03225806	3.731899756	4.532475191	20.54333136
CCI	42	100	83.23668639	13.45440086	181.0209026
IRI (In/Mile)	46	150	87.02071006	21.00611738	441.2569675
Rut (In)	0.08	0.24	0.136804734	0.031149055	0.000970264
AADT	34000	136000	88404.7619	19825.6502	393056405.9
Heavy Vehicle %	0.9037%	9.8782%	4.6979%	2.5171%	0.0634%

As it is shown in the Table 20, the variance is much higher than the mean for dependent variables. In this case, Negative Binomial regression is a great method for predicting the number of crashes per mile per year for each segment. Negative Binomial regression formulation is as below;

$$\lambda_i = e^{\beta X_i} + e^{\varepsilon_i}, \tag{12}$$

where ε_i is a gamma distribution error with mean of 1.0 and variance of α^2. Consequently, λ_i=Number of Crashes/mile/years at segment i and X_i are the characteristics of that segment such as IRI, CCI and etc., and β is the coefficient of those characteristics.

5.5 Results

A nominal logistic model was developed to analyze crash severity considering different aspects of the road, driver, and traffic situation at the time of the crash. The model shows that RD, Distraction of the Driver, Percentage of Heavy Vehicle, Drunkenness of the Driver and Driver's Age are significant in predicting Crash Severity. The result of crash severity is shown in table below;

Table 25 - Nominal Logistic Model Highlights for Crash Severity

Source	LogWorth	PValue
Rut (in)	4.435	0.00004
IF_DISTRACTED	3.657	0.00022
HV %	3.617	0.00024
IF_DRUNK	3.052	0.00089
IF_SENIOR	2.237	0.00580
CCI	1.312	0.04879
IRI	0.871	0.13447
Crash Hour	0.869	0.13535
IF_SPEEDING	0.825	0.14950
Roadway Surface Condition	0.014	0.96897
AADT	0.005	0.98794
IF_YOUNG	0.003	0.99198
Light Condition	0.002	0.99576

As it was expected, the Least Square method is not the best regression method for our dataset and the RSquare value for the regression is 0.16 which indicates that the model is not reliable, but it shows that CCI value is significant for predicting the number of crashes per mile. Least Square Model shows that number of crashes are higher at higher CCI values when the traffic is low, but by increasing traffic, there are less crashes at the segments with higher CCI value. A reasonable explanation of this finding in terms of human behaviors can be; drivers might drive with a higher speed and carelessly when the traffic is low and the quality of road is good, but in higher traffic they pay more attention to other vehicles, and any severe distress on the road (low quality of road) can cause a crash.

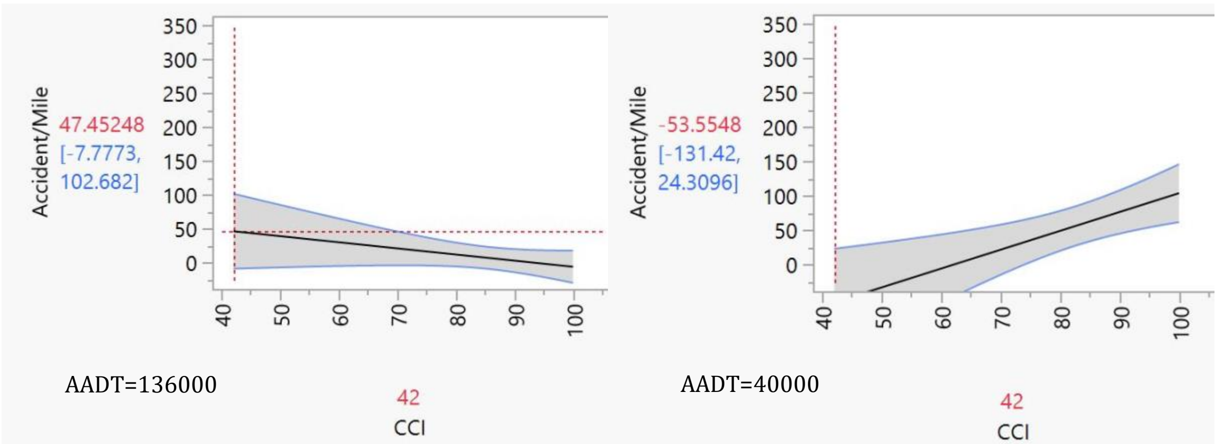

FIGURE 13 - CCI VS. CRASH/MILE LEAST SQUARE REGRESSION GRAPHS

5 Negative Binomial models were developed and as it is shown in Table 23, CCI value and Heavy Vehicle Percentage are significant factors and the models are significant. All models have RSquare value of around 0.25 except for the model for crashes in non-dry road condition, which is surprisingly high (RSquare=1), and it might be due to the low number of crashes at that category. Figure 7 shows the negative binomial solution path. Table 22 shows a summary of all models. In addition, just like Least Square Model, the effect of CCI on Crashes changes rapidly in different AADT values (Figure 14).

TABLE 26 - NEGATIVE BINOMIAL MODELS HIGHLIGHTS

	All Crashes	Daylight Crashes	Nighttime Crashes	Dry Surface Crashes	Non-Dry Surface Crashes
Negative Binomial Distribution Parameters	Dispersion	Dispersion	Dispersion	Dispersion	Dispersion
Estimate	0.7113606	0.7426411	0.6141701	0.7245924	2.44E-20
Std Error	0.0547676	0.0596753	0.0616463	0.0569351	0
Wald ChiSquare	168.70688	154.8704	99.257557	161.96757	.
Prob > ChiSquare	<.0001*	<.0001*	<.0001*	<.0001*	.

Lower 95%	0.6040181	0.6256797	0.4933456	0.6130017	2.44E-20
Upper 95%	0.818703	0.8596026	0.7349946	0.8361831	2.44E-20
Number of rows	338	338	338	338	338
-LogLikelihood	950.06498	903.89651	750.93319	950.97701	-5855.521
Number of Parameters	33	33	33	33	33
BIC	2092.2905	1999.9535	1694.0269	2094.1145	-11518.88
AICc	1973.5115	1881.1746	1575.248	1975.3356	-11637.66
Generalized RSquare	0.2454521	0.2594434	0.2128998	0.2621418	1

TABLE 27 - NEGATIVE BINOMIAL SIGNIFICANT FACTORS FOR ALL CRASHES

Term	Estimate	Std Error	Wald ChiSquare	Prob > ChiSquare	Lower 95%	Upper 95%
CCI	0.0215746	0.0056535	14.563179	0.0001*	0.010494	0.0326552
(AADT-88360.9)*(HV%-0.04698)	-0.00053	0.0001538	11.874119	0.0006*	-0.000831	-0.000229
(CCI-83.2367)*(AADT-88360.9)	-9.748e-7	4.0901e-7	5.6805887	0.0172*	-1.776e-6	-1.732e-7
(CCI-83.2367)*(IRI-87.0207)*(Rut (In)-0.1368)*(HV%-0.04698)	-0.857601	0.3651222	5.5168929	0.0188*	-1.573228	-0.141975
HV%	7.9619674	3.4014005	5.4793022	0.0192*	1.295345	14.62859
(CCI-83.2367)*(AADT-88360.9)*(HV%-0.04698)	-3.139e-5	0.0000145	4.6850642	0.0304*	-5.982e-5	-2.966e-6

As it is shown in the figure below, when the traffic volume is low, there are more crashes at the segments with higher quality of asphalt, and it can be argued that the drivers tend to speed up when the traffic is low and pavement quality is high and they might cause crashes due to the high speed. But in the segments with higher volume, quality of pavement is crucial to prevent crashes from happening.

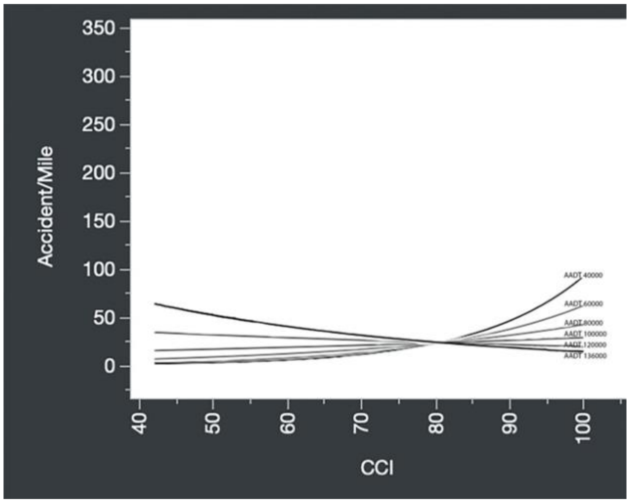

FIGURE 14 - CCI VS. CRASH/MILE NEGATIVE BINOMIAL MODEL GRAPH

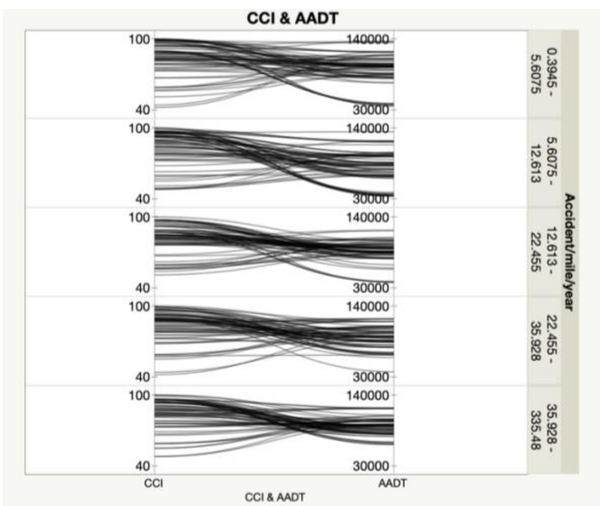

FIGURE 15 - CCI & AADT VS. CRASH/MILE PARALLEL LINE GRAPH

5.6 SUMMARY AND CONCLUSIONS

There is a gap between infrastructure planning and traffic control. This study used crash analysis as an integrating factor to shed light on the relationship between traffic control and infrastructure planning and the reflection of two subjects in number of crashes. Fairfax County is the most populated area in Washington, DC metropolitan area. Commuters mainly use Interstate Highway System as their choice of route. Number of crashes are considerable, and percentage of injury crashes is high in the Fairfax Interstate Highway System. This study analyzed crashes in the area by considering different factors such as International Roughness Index, Critical Condition Index, Rutting Depth, Annual Average Daily Traffic, Percentage of Heavy Vehicle, Road's Surface Condition and Road's Lighting Condition.

Initially a nominal logistic model was developed to understand all the factors. It shows that Rutting Depth, Driver Distraction, Percentage of Heavy Vehicles, Being Drunk and Being Senior are significant to predict the crash severity.

Crashes were categorized by Daylight and Nighttime crashes and Dry Surface and Non-Dry Surface Condition and Least Square Regression and Negative Binomial Regressions were developed by having Critical Condition Index, International Roughness Index, Rutting Depth, Annual Average Daily Traffic and Percentage of Heavy Vehicles as inputs and Number of Crashes Per Mile of each segment as model role variable. The result shows that CCI value and Percentage of Heavy Vehicles are significant to predict the number of crashes. Based on the findings, number of crashes have reverse relation with the CCI value when the road has high volume. By increasing pavement quality from CCI=60 to CCI=80 number of crashes drop from 50 crashes/mile/year to 25 crashes/mile/year which is a significant improvement.

This model predicts the number crashes per year per mile having CCI, IRI, RD, AADT, HV% as its inputs. There are many other inputs that can be taken into the consideration, but they make the model more complicated and some data might not be available for each road. Based on the M&R plans, the future road characteristics can be predicted and based on the traffic control strategies, AADT and HV% values can be predicted for each segment. Eventually, the number of crashes for that segment can be predicted for that segment using the equation below;

$$Accidents\ for\ Segment\ i = Accident\ Model\ for\ Segment\ i\ \times\ Segment\ Length \quad (13)$$

5.7 Further Works

Based on what have been discussed above, transportation decision making and is being processed in two different levels of traffic control and infrastructure management but there is no integrated system to merge these two layers. Traffic control plans can provide us with future traffic volume values in the network. Infrastructure plans provide future pavement information based on pavement models and current traffic situation (considering the growth factor). There should be a real-time accurate and dynamic transportation framework to demonstrate the effect of each transportation strategy in future transportation condition. There are some works about the infrastructure quality and traffic mobility. By using the result of those works and this study, an integrated framework can be developed. As a matter of fact, this is the next step of our lab to create the mentioned system. To give a more detailed description of the work, CCI value is predictable and it has an accurate model to predict CCI value based on future travel demand data. Travel demand management plans for the next 20-30 of years have been developed and can be obtained from the DOTs these plans include many future traffic information. Infrastructure plans provide future pavement information values. All of these parameters have effect on one another but there is no system including all of the parameters. A model that have all of the parameter as its inputs and provide an optimized solution for the future transportation project can be a great improvement in terms of transportation engineering, Most of the travel demand models are designed to minimize overall delay of the network, but they are not designed to minimize the pavement deterioration. The value of this study is to define a safety constraint for that framework

ACKNOWLEDGMENTS

This work was funded by the Virginia Transportation Research Council (VTRC). The authors would like to acknowledge and thank Mr. Kevin McGhee, Dr. Mike Fitch, Ms. Elaine Thurman, and Mr. Ibrahim Abuawad for their support and help.

AUTHOR CONTRIBUTIONS

The authors confirm contribution to the paper as follows: study conception and design: Montasir Abbas, Linbing Wang; analysis and interpretation of results: Ali Mohagheghi; draft manuscript preparation: Ali Mohagheghi and Montasir Abbas. All authors reviewed the results and approved the final version of the manuscript.

APPENDIX A

$$MSI = \frac{SN_{eff}}{SN_{req}} \frac{0.4728(D_0 - D_{1.5H_p})^{-0.4810} H_p^{0.7581}}{0.05716[\log(ESAL) - 2.32\log(M_r) + 9.07605]^{2.36777}}$$

where

D_0= the FWD deflection under the applied load

H_p = total pavement depth (i.e., measured from the top of the pavement to the top of the subgrade)

$D_{1.5H_p}$= FWD deflection at a distance equal to 1.5 times the total pavement depth

M_r = resilient modulus calculated using FWD measurements

ESAL = total accumulated truck traffic over a total design period of 20 years.

REFERENCE

1. Huang YH. Pavement Analysis and Design; 2nd edn, Pearson Prentice Hall: New Jersey, 2004.

2. Strat MR, Kim J, and Berg WD. Potential safety cost-effectiveness of treating rutted pavements. Transportation Research Record 1998; 1629:208–213.

3. Transportation Research and Board (T.R.B). A State-of-the-Art Report, the Influence of Roadway Surface Discontinuities on Safety, Transportation Research Board, National Research Council Washington D.C. 1984.

4. Horne WB, and Leland TJW. (1962) Influence of Tire Tread Pattern and Runway Surface Condition on Braking Friction and Rolling Resistance of a Modern Aircraft Tire. NASA TN D-1376. National Aeronautics and Space Administration.

5. Quinn BE, and Hilderbrand SE. Effect of road roughness on vehicle steering. Highway Research Record 1973; 47:62–75.

6. Wambold JC, Brickman AD, Park WH, and Ingram J. Effect of road roughness on vehicle braking. Highway Research Record 1973; 471:76–82.

7. Nakatsuji T, Kaku T, Fujiwara T, Hagiwara Y, and Onodera Y. Dynamic Behavior of Vehicle on a Rutted Road, Surface Charateristics of Roadways: International Research and Technologies,

8. ASTM STP 1031; Meyer WE, Reichert J. (Eds.), American Society for Testing and Materials: Philadelphia, 1990; 198–210.

9. Burns JC. Roughness and roadway safety. Transportation Research Record 1981; 836:8–14.

10. Calibrating and Validating Overlay Tester–Based Fatigue Cracking Model with Data from National Center for Asphalt Technology, Sheng Hu, Fujie Zhou, Thomas Scullion, and Joe Leidy

11. Hveem, F. N. HRB Bulletin 114: Pavement Deflections and Fatigue Failures. HRB, National Research Council, Washington, D.C., 1955, pp. 43–87.

12. DOT OH: PAVEMENT CONDITION RATING SYSTEM.

13. Gharaibeh, Nasir G.; Zou, Yajie; Saliminejad, Siamak (2010): Assessing the Agreement among Pavement Condition Indexes. In J. Transp. Eng. 136 (8), pp. 765–772. DOI: 10.1061/(ASCE)TE.1943-5436.0000141.

14. Shankar V, Mannering F, and Barfield W. Effect of roadway geometrics and environmental factors on rural crash frequencies. Crash Analysis & Prevention 1995; 27(3):371–389.

15. Abdel-Aty MA, and Radwan AE. Modeling traffic crash occurrence and involvement. Crash Analysis & Prevention 2000; 32(5):633–642.

16. Savolainen PT, and Tarko AP. Safety impacts at intersections on curved segments. Transportation Research Record 2005; 1908:130–140.

17. Medlock, Ronald D.; Laffrey, David C. (Eds.) (1998): Structural Materials Technology III: An NDT Conference. Non-Destructive Evaluation Techniques for Aging Infrastructure & Manufacturing. San Antonio, TX, Tuesday 31 March 1998: SPIE (SPIE Proceedings).

18. Ram, Prashant V.; Smith, Kelly L.; Zimmerman, Kathryn A.; Ratliff, Tammy; Amenta, James (2016): Developing a Correlation Between the Pavement Condition Ratings Used by Five Federal Lands Management Agencies. In Transportation Research Record 2589 (1), pp. 78–86. DOI: 10.3141/2589-09.

19. Shah, Yogesh U.; Jain, S. S.; Tiwari, Devesh; Jain, M. K. (2013): Development of Overall Pavement Condition Index for Urban Road Network. In Procedia - Social and Behavioral Sciences 104, pp. 332–341. DOI: 10.1016/j.sbspro.2013.11.126.

20. Shekharan, Raja: Microsoft Word - State_of_the_Pavement_2016.docx. State_of_the_Pavement_2018.

21. Vlacich, Branco (2017): State of the Pavement 2017. Virginia Department of Transportation. Richmond, VA.

22. Yin, Yafeng; Wang, Yinhai; Lu, Jian; Wang, Wei (Eds.) (07262011): ICCTP 2011: Towards Sustainable Transportation Systems. 11th International Conference of Chinese Transportation Professionals (ICCTP). Nanjing, China, August 14-17, 2011. Reston, VA: American Society of Civil Engineers.

23. Z. Al-Qadi, Imad L.; Murrell, Scott (Eds.) (06182013): Airfield and Highway Pavement 2013. 2013 Airfield & Highway Pavement Conference. Los Angeles, California, United States, June 9-12, 2013. Reston, VA: American Society of Civil Engineers.

24. (DOT OH; Gharaibeh et al. 2010; Ram et al. 2016; Shah et al. 2013; Medlock and Laffrey 1998; shekharan; Vlacich 2017; Yin et al. 07262011; Z. Al-Qadi and Murrell 06182013) Gianluca Dell'Acqua, Fred Wegman, Transport Infrastructure and Systems.

25. Strat MR, Kim J, and Berg WD. Potential safety cost-effectiveness of treating rutted pavements. Transportation Research Record 1998; 1629:208–213.

26. Samer W. Katicha, Ph.D., Safak Ercisli, Gerardo W. Flintsch, Ph.D., P.E., James M. Bryce, Ph.D., and Brian K. Diefenderfer, Ph.D., P.E. Development of Enhanced Pavement Deterioration Curves, 2016

27. Chang L, and Chen W. Data mining of tree-based models to analyze freeway crash frequency. Journal of Safety Research 2005; 36(4):365–375.

6 Conclusions

The first study demonstrates the effect of different road and traffic characteristics on crash severity and collision type. The second study analyze the effect of different road and traffic characteristics on number of crashes per each segment of the road. Despite the fact that the scope and period of two studies are different and various types of analysis were done to get the result, but both studies show that road characteristics such as pavement condition and surface and lighting condition are effective on number of crashes and severity of each crash in the highway system. By having the both models we can predict the number of crashes per segment and severity of each crash based on the time, location and some other attributes of each crash. First research shows that Collision Type, Number of Vehicles (in crash), Crash Hour, Being Wok Zone, Surface Condition, Rutting Depth, and Critical Condition Index are significant factors to predict the Crash Severity, and Crash Severity, Crash Hour, Being Work Zone, Crash Location (the position on the road), Surface Condition, Number of Vehicles in Crash, Light Condition, Critical Condition Index Value, and Percentage of Heavy Vehicle are significant to predict the collision type of each crash. Second study shows that CCI and Percentage of Heavy Vehicles are significant factors. The results show that lines in CCI vs. Predicted Crashes/mile/year have different slopes by changing AADT value. Both models are significant, and they can be supplementary for each other to have a brighter understanding of the future safety condition of the transportation system based on different factors. It can be concluded that no matter how accurate and reliable our models are, there are other effective factors that are not available in the dataset and they might cause the error in the models. It is recommended that every available factor be included in the future models based on this study. As the results show, models are reliable and significant, and it seems that they can be repeatable for the other highway roads, but they might need to be calibrated. The outcome of this study can be very helpful for the infrastructure decision making process of transportation agencies, especially agencies in Virginia and metropolitan DC area. This study has used various analysis methods and it can be a great reference for further studies and other educational activities. CCI prediction model is more reliable and accurate so a crash model that includes CCI as an input might be more accurate.

7 Future Works

The developed models are significant, and they can be used in an integrated transportation planning system that VT-SCORES lab is trying to create in the near future. A system that include future infrastructure maintenance plans and future traffic strategies for the transportation network and can predict the outcome of each traffic plan on the future infrastructure condition of the network and vice versa. The predicted number of crashes will be an important output for this system. Because it can be used as a constraint in the optimization process and decision-making process.

References

1. H. Sahin, P. Narciso, and N. Hariharan, "Developing a five-year maintenance and rehabilitation (M&R) plan for HMA and concrete pavement networks," APCBEE procedia, vol. 9, pp. 230-234, 2014.

2. K. A. Abaza and M. M. Murad, "Pavement rehabilitation project ranking approach using probabilistic long-term performance indicators," Transportation Research Record, vol. 2153, no. 1, pp. 3-12, 2010.
3. D. Geiger, "Pavement preservation definitions," Accessed on June, vol. 29, no. 2009, p. 18, 2005.
4. Assets, "Tutorial: PMS Users-Tutorial on Using the Agile Pavement Manager," ed: Virginia Department of Transportation, Richmond, 2012.
5. E. de León Izeppi, A. Morrison, G. W. Flintsch, and K. K. McGhee, "Best practices and performance assessment for preventive maintenance treatments for Virginia pavements," Virginia Center for Transportation Innovation and Research, 2015.
6. D. G. Peshkin and T. E. Hoerner, Pavement preservation: Practices, research plans, and initiatives. Citeseer, 2005.
7. L. C. Falls, S. Khalil, W. R. Hudson, and R. Haas, "Long-term cost-benefit analysis of pavement management system implementation," Transportation Research Board, vol. 2, pp. 133-138, 1994.
8. P. Babashamsi, N. I. M. Yusoff, H. Ceylan, N. G. M. Nor, and H. S. Jenatabadi, "Evaluation of pavement life cycle cost analysis: Review and analysis," International Journal of Pavement Research and Technology, vol. 9, no. 4, pp. 241-254, 2016.
9. D. N. Geoffroy, Cost-effective preventive pavement maintenance (no. Project 20-5 FY 1993). 1996.
10. Huang YH. Pavement Analysis and Design; 2nd edn, Pearson Prentice Hall: New Jersey, 2004.
11. Strat MR, Kim J, and Berg WD. Potential safety cost-effectiveness of treating rutted pavements. Transportation Research Record 1998; 1629:208–213.
12. Transportation Research and Board (T.R.B). A State-of-the-Art Report, the Influence of Roadway Surface Discontinuities on Safety, Transportation Research Board, National Research Council Washington D.C. 1984.
13. Horne WB, and Leland TJW. (1962) Influence of Tire Tread Pattern and Runway Surface Condition on Braking Friction and Rolling Resistance of a Modern Aircraft Tire. NASA TN D-1376. National Aeronautics and Space Administration.
14. Quinn BE, and Hilderbrand SE. Effect of road roughness on vehicle steering. Highway Research Record 1973; 47:62–75.
15. Wambold JC, Brickman AD, Park WH, and Ingram J. Effect of road roughness on vehicle braking. Highway Research Record 1973; 471:76–82.
16. Nakatsuji T, Kaku T, Fujiwara T, Hagiwara Y, and Onodera Y. Dynamic Behavior of Vehicle on a Rutted Road, Surface Charateristics of Roadways: International Research and Technologies, ASTM STP 1031; Meyer WE, Reichert J. (Eds.), American Society for Testing and Materials: Philadelphia, 1990; 198–210.
17. Burns JC. Roughness and roadway safety. Transportation Research Record 1981; 836:8–14.
18. Calibrating and Validating Overlay Tester–Based Fatigue Cracking Model with Data from National Center for Asphalt Technology, Sheng Hu, Fujie Zhou, Thomas Scullion, and Joe Leidy
19. Hveem, F. N. HRB Bulletin 114: Pavement Deflections and Fatigue Failures. HRB, National Research Council, Washington, D.C., 1955, pp. 43–87.
20. DOT OH: PAVEMENT CONDITION RATING SYSTEM.

21. Gharaibeh, Nasir G.; Zou, Yajie; Saliminejad, Siamak (2010): Assessing the Agreement among Pavement Condition Indexes. In J. Transp. Eng. 136 (8), pp. 765–772. DOI: 10.1061/(ASCE)TE.1943-5436.0000141.
22. Shankar V, Mannering F, and Barfield W. Effect of roadway geometrics and environmental factors on rural accident frequencies. Accident Analysis & Prevention 1995; 27(3):371–389.
23. Abdel-Aty MA, and Radwan AE. Modeling traffic accident occurrence and involvement. Accident Analysis & Prevention 2000; 32(5):633–642.
24. Savolainen PT, and Tarko AP. Safety impacts at intersections on curved segments. Transportation Research Record 2005; 1908:130–140.
25. Medlock, Ronald D.; Laffrey, David C. (Eds.) (1998): Structural Materials Technology III: An NDT Conference. Non-Destructive Evaluation Techniques for Aging Infrastructure & Manufacturing. San Antonio, TX, Tuesday 31 March 1998: SPIE (SPIE Proceedings).
26. Ram, Prashant V.; Smith, Kelly L.; Zimmerman, Kathryn A.; Ratliff, Tammy; Amenta, James (2016): Developing a Correlation Between the Pavement Condition Ratings Used by Five Federal Lands Management Agencies. In Transportation Research Record 2589 (1), pp. 78–86. DOI: 10.3141/2589-09.
27. Shah, Yogesh U.; Jain, S. S.; Tiwari, Devesh; Jain, M. K. (2013): Development of Overall Pavement Condition Index for Urban Road Network. In Procedia - Social and Behavioral Sciences 104, pp. 332–341. DOI: 10.1016/j.sbspro.2013.11.126.
28. Shekharan, Raja: Microsoft Word - State_of_the_Pavement_2016
29. State_of_the_Pavement_2018.
30. Vlacich, Branco (2017): State of the Pavement 2017. Virginia Department of Transportation. Richmond, VA.
31. Yin, Yafeng; Wang, Yinhai; Lu, Jian; Wang, Wei (Eds.) (07262011): ICCTP 2011: Towards Sustainable Transportation Systems. 11th International Conference of Chinese Transportation Professionals (ICCTP). Nanjing, China, August 14-17, 2011. Reston, VA: American Society of Civil Engineers.
32. Z. Al-Qadi, Imad L.; Murrell, Scott (Eds.) (06182013): Airfield and Highway Pavement 2013. 2013 Airfield & Highway Pavement Conference. Los Angeles, California, United States, June 9-12, 2013. Reston, VA: American Society of Civil Engineers.
33. Gianluca Dell'Acqua, Fred Wegman, Transport Infrastructure and Systems.
34. Strat MR, Kim J, and Berg WD. Potential safety cost-effectiveness of treating rutted pavements. Transportation Research Record 1998; 1629:208–213.
35. Samer W. Katicha, Ph.D., Safak Ercisli, Gerardo W. Flintsch, Ph.D., P.E., James M. Bryce, Ph.D., and Brian K. Diefenderfer, Ph.D., P.E. Development of Enhanced Pavement Deterioration Curves, 2016
36. Chang L, and Chen W. Data mining of tree-based models to analyze freeway accident frequency. Journal of Safety Research 2005; 36(4):365–375.
37. Sheng Hu, Fujie Zhou, Thomas Scullion, and Joe Leidy, Calibrating and Validating Overlay Tester–Based Fatigue Cracking Model with Data from National Center for Asphalt Technology
38. Mofreh F. Saleh, Michael S. Mamlouk, and Emmanuel B. Owusu-Antwi, Mechanistic Roughness Model Based on Vehicle-Pavement Interaction
39. Zhang, Harry. The Optimality of Naive Bayes (PDF). FLAIRS2004 conference.

40. Caruana, R.; Niculescu-Mizil, A. (2006). An empirical comparison of supervised learning algorithms. Proc. 23rd International Conference on Machine Learning.
41. Efron B, Tibshirani R(1993) An Introduction to the Bootstrap (Chapman & Hall, London)
42. Felsenstein J(1985) Evolution 39:783–791
43. Hillis D, Bull J(1993) Syst Biol 42:182–192
44. Full description for Bootstrap method: https://doi.org/10.1186/s12862-018-1131-3
45. Greene, William H. (2012). Econometric Analysis (Seventh ed.). Boston: Pearson Education. pp. 803–806. ISBN 978-0-273-75356-8.
46. Kevin Gurney; An introduction to neural networks 1997