Pavement condition and crashes

Toshihiro Yokoo[a] David Levinson[b] Mihai Marasteanu[c]

[a] Email: yokoo002@umn.edu

[b]Professor of Transport, University of Sydney. Email: david.levinson@sydney.edu.au

[c]Professor, University of Minnesota

Abstract

Change in weather state (such as the freeze-thaw cycle) leads to distresses in pavement materials. It has been hypothesized that poor pavement quality reduces the ability of roads to drain and reduces the ability of vehicles to resist skidding, and is thus associated with more crashes. This paper combines GIS data on crashes with a separate GIS database to test the hypothesis. Poor road quality is associated with more property damage and injury crashes. The interaction of road quality and curves was surprising, indicating that good pavement quality on curves increased the fatal, injury, and property-damage crash rate.

1 Introduction

In a place like Minnesota well known for its severe weather, maintaining road pavements to meet high standards remains a challenge. Change in weather state (such as the freeze-thaw cycle) leads to distresses in pavement materials. Some have found that crash rate depends on the pavement type and pavement condition (Buddhavarapu et al., 2013, Al-Masaeid, 1997, Abdel-Aty and Abdalla, 2004, Lee et al., 2015, Najafi et al., 2017, Merritt et al., 2015, Elghriany, 2016, Vinayakamurthy et al., 2017, Chan et al., 2009). One posited reason is that poor pavement quality reduces the ability of roads to drain and reduces the ability of vehicles to resist skidding, and is thus associated with more crashes. In order to improve road safety, several pavement maintenance treatments are carried out, such as "rout and seal cracks" and "hot-mix patching" for improving pavement roughness and distress (Tighe et al., 2000).

Crash rate of tined pavement sites is larger than the rate of ground pavement sites. When the pavement condition is wet or icy, crashes are more likely than under dry conditions (Drakopoulos et al., 1998). When the pavement condition is poor, severe crashes are more likely, but when the pavement condition is very poor, severe crashes are less likely to occur than poor pavement conditions (Li et al., 2013). In crash rate estimation models, the results indicate that most important independent variable is amount of traffic, while geometric design (lane width and access control) and pavement condition (friction, serviceability index, and pavement type) are also important variables (Karlaftis and Golias, 2002). Our research proposes to statistically test the relationship between incident number and road quality, while controlling for traffic data (annual average daily traffic (AADT) and percent truck), segment length, crash conditions (date, road characteristics, and road surface), and pavement type.

To investigate the relationship, we combine data from various sources. We then conduct a statistical analysis to ascertain the effects of good road quality on incident number and severity. This paper describes the data, methods, hypotheses, and results in turn.

2 Data

This research uses pavement quality data and crash data from the Minnesota Department of Transportation (MnDOT). Pavement quality data is available from 2000 to 2015, the crash data from 2003 to 2014. Therefore, we use the data from 2003 to 2014 in order to analyze the relationship between incident number and road quality. While MnDOT's crash data is recorded for all road sections in Minnesota, pavement quality data is only available for highway road segments.

The crash data is a GIS shapefile and contains information about each crash including: location, crash date, severity of crash, crash type, road characteristics, road design, and weather condition.

The pavement quality data records pavement roughness and surface distress information for each year and it is recorded on a mile-by-mile basis. AADT and percent truck on each segment are also collected. We also received an electronic highway map from MnDOT, which has highway segment information.

Several standard indicators of pavement quality (Surface Rating (SR), International Roughness Index (IRI), Pavement Quality Index (PQI)) are provided, but we focus on the Ride Quality Index (RQI). RQI ranges from 0-5 and indicates the smoothness of the pavement, with 5 indicating smoother. The correlation between the alternative pavement quality indices are high (RQI and SR: 0.55, PQI and SR: 0.89, PQI and RQI: 0.85, RQI and IRI: -0.97), so we use only RQI as an independent variable describing pavement quality.

To manage the data, we use QGIS version 2 (Sutton and Dassau, 2015), an open source geographic information system.

3 Methodology

The crash data is recorded as points and the pavement quality data is recorded on mile-by-mile basis. We match these two data by a function in QGIS.

In brief, we select for crashes by year. There are around 15,900 crashes per year, 190,918 in total. We aim to select crashes only on state highways for which pavement quality data is available (Figure 1). We count the crashes on each segment by severity level (1: Incapacitating Injury, 2: Non-incapacitating Injury, 3: Possible Injury, 4: Fatal, 5: Property Damage, 6: No Value). The count depends on the GIS buffer around the road, tighter buffers remove crashes from the data set, ultimately we use a buffer of 0.00001m (i.e. only accepting crashes that were accurately geocoded). Then we merge the crash data with the pavement quality data.

This paper tests the hypothesis that good road quality is negatively correlated number of crashes. We analyze RQI for each year on a mile-by-mile basis, and control for traffic, share of trucks, pavement type, highway geometry, weather conditions, day-of-week, month-of-year, and time-of-day. The dependent variable is the number of crashes (distinguished for each severity level). Number crashes by severity is given by ($Crash_S$) where (S=Fatal, Injury, or Property damage). Many segments had no crashes in a given year. Injury is the sum of Incapacitating Injury, Non-incapacitating Injury, and Possible Injury. Negative binomial regressions are used.

Table 2 shows the list of independent variables.

In order to avoid the dummy variable trap, we drop one category from the model for variables which would otherwise be determined, in this case, Year: 2014 and Pavement type: concrete.

We also add several independent variables about crash conditions (date, road characteristics, and road surface) to the model. To illustrate the coding, as shown in Figure 2, horizontal alignment of crash location is both 'straight' and 'curve' in this segment, and vertical alignment of crash location is 'level', 'grade' and 'hillcrest'. In this case, the value of 'curve', 'grade' and 'hillcrest' are 1 while the value of 'sag' is 0. We code for 'Spring Load Restrictions' (SLR) (March to May) when roads are weak during spring due to the spring thaw, therefore the local authority has begun Spring Load Weight Restrictions (SLR) to reduce road damage (MnDOT). We code for peak travel periods, 'Rush hour' is defined as 6 a.m. to 9 a.m. and 3 p.m. to 7 p.m (Brown, 2013)

4 Results

We analyze the relationship between several variables (pavement data, traffic data, crash conditions) and crash statistically with a Negative Binomial Regression. Table 3 showing regression results is calculated using the statistical package *R* version 3.

In all cases segment length is positive, longer segments have more opportunities for crashes.

For all cases, percentage trucks is negative, indicating number of crashes drop on facilities with a higher share of trucks. Roads serving a higher share of trucks may be built to a higher standard than other roads, so the causality might not be that trucks reduce crashes.

The relationship between traffic and crashes is more complex. We modeled this parabolically, including both *Traffic* and $Traffic^2$. For fatal crashes, at lower levels of traffic, crashes decline with increasing traffic, but beyond a threshold they increase. In contrast for property damage crashes, the relationship is the reverse, and for injury crashes, crashes increase with number of vehicles on the road. We hypothesize that congestion increases minor crashes but decreases fatal crashes (because traffic is slower), but in the data the opposite pattern is revealed. Perhaps not surprisingly, during rush hour periods, crashes of all types increase due to the increased opportunity for vehicular interaction.

Pavement material (bituminous rather than concrete) is associated with a higher number of injury and property damage crashes. Again, the causality might not be that bituminous causes crashes, rather it could be that concrete roads, which tend to serve higher levels of traffic, are built to a different or more modern standard.

Number of injury and property damage crashes is generally decreasing over time (compared with 2014).

Crashes of all types increase on weekends, on grades, and in snow.

Injury and property damage crashes also increase on hillcrests, sags, wet conditions, and during the Spring Load Restrictions period.

Good pavement quality is associated with lower crash rates in several conditions: for $RQI : Snow$ for fatal crashes, as well as $Bituminous_2 : RQI$ and $RQI : Sag$ for Injury and Property Damage Crashes, and $RQI : Wet$ for Injury and $RQI : Hillcrest$, and $RQI : SLR$ for Property Damage. The z-value from the model indicates that there are significant differences across pavement type.

However, counter-intuitively perhaps, for all three crash types, good pavement quality on curves ($RQI : Curve$) increases number of crashes compared with curves in general or good pavement quality in general. Perhaps poor pavement quality on curves positively affects driver alertness. Similarly for property damage crashes, $RQI : Snow$ is positive.

5 Discussion and Conclusion

This paper investigates the relationship between pavement quality and crashes in Minnesota from 2003 to 2014. The most pertinent findings from the results are that good road quality is negatively and significantly correlated with property damage crashes (both $Bituminous_1 : RQI$ and $Bituminous_2 : RQI$) and with injury crashes (for $Bituminous_2 : RQI$), representing 3 of 6 cases, the other 3 were statistically insignificant. RQI is related to driver's perception of smoothness, so it is assumed that the same road conditions that lead to uncomfortable driving are correlated with an increase in the number of crashes, after controlling for traffic levels, number of trucks, and geometric conditions.

Future studies should aim to replicate (or refute) this result.

One challenge with the analysis is that crash conditions differ within the pavement database segments, which are assumed homogeneous. For example, even though one fatal crash occurs at the segment in Figure 2, the value of both 'Grade' and 'Hillcrest' becomes 1 in the fatal model in Table 3. The data is stored on a mile-by-mile basis and it is referenced by mile posts along the highway. Therefore, if the length of each segment were shorter, the reliability of the regression model would improve.

This paper focuses on only roads managed by MnDOT (state highways), although many of them are 2 lane undivided roadways, they tend to be more important and designed to a higher standard than lower level roads. Future research should aim to analyze this relationship on non-highway road sections as well.

References

Prasad Buddhavarapu, Ambarish Banerjee, and Jorge A Prozzi. Influence of pavement condition on horizontal curve safety. *Accident Analysis & Prevention*, 52:9–18, 2013.

Hashem R Al-Masaeid. Impact of pavement condition on rural road accidents. *Canadian Journal of Civil Engineering*, 24(4):523–531, 1997.

Mohamed Abdel-Aty and Fathy Abdalla. Linking roadway geometrics and real-time traffic characteristics to model daytime freeway crashes: generalized estimating equations for correlated data. *Transportation Research Record: Journal of the Transportation Research Board*, (1897):106–115, 2004.

Jaeyoung Lee, BooHyun Nam, and Mohamed Abdel-Aty. Effects of pavement surface conditions on traffic crash severity. *Journal of Transportation Engineering*, 141(10):04015020, 2015.

Shahriar Najafi, Gerardo W Flintsch, and Alejandra Medina. Linking roadway crashes and tire–pavement friction: a case study. *International Journal of Pavement Engineering*, 18(2):119–127, 2017.

David K Merritt, Craig Lyon, Bhagwant Persaud, et al. Evaluation of pavement safety performance. Technical report, United States. Federal Highway Administration, 2015.

Ahmed F Elghriany. *Investigating Correlations of Pavement Conditions with Crash Rates on In-Service US Highways*. PhD thesis, University of Akron, 2016.

Mounica Vinayakamurthy, Michael Mamlouk, and PE Shane Underwood. Effect of pavement condition on accident rate. *Master of Science, Arizona State*, 452, 2017.

Chun Yip Chan, Baoshan Huang, Xuedong Yan, and Stephen Richards. Relationship between highway pavement condition, crash frequency, and crash type. *Journal of Transportation Safety & Security*, 1(4): 268–281, 2009.

Susan Tighe, Ningyuan Li, Lynne Falls, and Ralph Haas. Incorporating road safety into pavement management. *Transportation Research Record: Journal of the Transportation Research Board*, (1699):1–10, 2000.

Aris Drakopoulos, Thomas Wenzel, Stephen Shober, and Robert Schmiedlin. Crash experience on tined and continuously ground portland cement concrete pavements. *Transportation Research Record: Journal of the Transportation Research Board*, (1639):140–146, 1998.

Yingfeng Li, Chunxiao Liu, and Liang Ding. Impact of pavement conditions on crash severity. *Accident Analysis & Prevention*, 59:399–406, 2013.

Matthew G Karlaftis and Ioannis Golias. Effects of road geometry and traffic volumes on rural roadway accident rates. *Accident Analysis & Prevention*, 34(3):357–365, 2002.

Tim Sutton and Otto Dassau. Qgis. http://www.qgis.org/en/site/index.html, 2015. Accessed Nov. 24, 2015.

MnDOT. Spring load restrictions. http://dotapp7.dot.state.mn.us/research/seasonal_load_limits/thawindex/slr_introduction.pdf. Accessed May. 3, 2016.

Heather Brown. Good question: When is rush hour? http://minnesota.cbslocal.com/2013/07/29/good-question-when-is-rush-hour/, 2013. Accessed May. 3, 2016.

Figure 1: Crash data (a) original and (b) after processing of Select by location

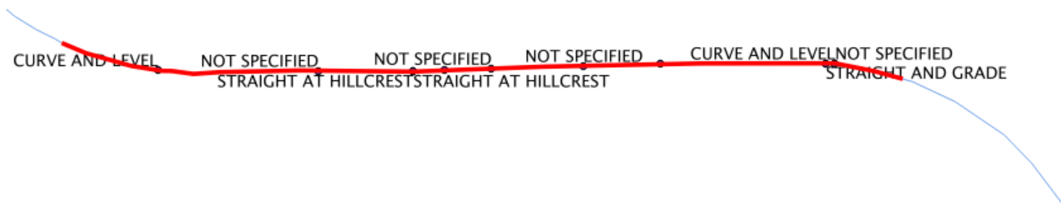

Figure 2: Road characteristics of crash location (Red line is one segment)

Table 1: Accuracy of buffer size

Year	Buffer size	Total crash (A)	Points in polygon (B)	# of Error (B-A)	Error rate
2004	10m	39,010	44,993	5,983	15.3%
2004	5m	38,465	43,224	4,759	12.4%
2004	3m	38,292	42,700	4,408	11.5%
2004	1m	38,129	41,900	3,771	9.9%
2004	0.1m	37,378	39,251	1,873	5.0%
2004	0.001m	36,388	37,143	755	2.1%
2004	0.0001m	32,303	32,343	40	0.1%
2004	0.00001m	29,581	29,581	0	0.0%

Table 2: List of Independent variables

Variables	**Definition**
$Trucks$	Percentage of truck volume among total traffic volume
$Traffic$	Annual average daily traffic (AADT)
$Length$	Segment length (miles)
$Bituminous_1$	Indicator, 1= pavement type is BAB, BFD, or BOB, (BAB: Bituminous Aggregate Base, BFD: Bituminous Full Depth, BOB: Bituminous Over Bituminous) 0 = otherwise
$Bituminous_2$	Indicator, 1= pavement type is BOC, (BOC: Bituminous Over Concrete) 0 = otherwise
$Concrete$	Indicator, 1= pavement type is Concrete, (CD: Concrete Doweled, CRC: Continuously Reinforced Concrete, CU: Concrete Undoweled) 0 = otherwise
$Year_{2003}$ - $Year_{2014}$	Indicator, 1= crash year is each year (2003 to 2014), 0 = otherwise
$Weekend$	Indicator, 1= crash date is Saturday or Sunday, 0 = otherwise
$Curve$	Indicator, 1= horizontal alignment of crash location is curve, 0 = otherwise
$Grade$	Indicator, 1= vertical alignment of crash location is grade, 0 = otherwise
$Hillcrest$	Indicator, 1= vertical alignment of crash location is hillcrest, 0 = otherwise
Sag	Indicator, 1= vertical alignment of crash location is sag, 0 = otherwise
Wet	Indicator, 1= road surface of crash location is wet, 0 = otherwise
$Snow$	Indicator, 1= road surface of crash location is snow, 0 = otherwise
SLR	Indicator, 1= crash date is during Spring Load Restrictions, 0 = otherwise
$Rushhour$	Indicator, 1= crash date is during rush hour, 0 = otherwise
$XX : RQI$	Interaction term, RQI: Ride quality index

Table 3: Negative Binomial Regression: Number of crashes by type

	Fatal			Injury			Property damage		
	Estimate	z value		Estimate	z value		Estimate	z value	
(Intercept)	-6.398E+00	-39.368	***	-3.024E+00	-104.461	***	-2.336E+00	-105.923	***
$Trucks$	-1.836E-02	-3.492	***	-2.692E-02	-25.887	***	-1.721E-02	-22.090	***
$Traffic$	-1.636E-05	-5.938	***	4.788E-06	11.105	***	1.100E-05	32.570	***
$Traffic^2$	8.180E-11	4.106	***	2.979E-12	0.985		-2.516E-11	-10.391	***
$Length$	7.312E-01	6.918	***	1.927E-01	10.644	***	5.968E-02	4.251	***
$Bituminous_1$	6.024E-02	0.268		1.710E-01	4.036	***	1.951E-01	5.946	***
$Bituminous_2$	-3.668E-01	-1.122		4.083E-01	7.687	***	6.599E-01	16.297	***
$Year_{2003}$	2.462E-01	2.216	*	1.396E-01	6.566	***	3.620E-02	2.208	*
$Year_{2004}$	1.746E-01	1.551		1.772E-01	8.407	***	1.171E-01	7.231	***
$Year_{2005}$	2.221E-01	1.987	*	1.255E-01	5.893	***	6.115E-02	3.744	***
$Year_{2006}$	1.050E-01	0.904		1.679E-01	7.764	***	1.040E-01	6.260	***
$Year_{2007}$	1.785E-01	1.561		1.355E-01	6.297	***	-1.400E-02	-0.838	
$Year_{2008}$	8.069E-02	0.691		9.398E-02	4.369	***	-3.147E-02	-1.901	.
$Year_{2009}$	-2.950E-02	-0.242		7.359E-02	3.367	***	-4.565E-02	-2.716	**
$Year_{2010}$	-1.403E-01	-1.125		1.107E-01	5.102	***	-1.350E-02	-0.808	
$Year_{2011}$	-1.817E-01	-1.444		5.851E-02	2.682	**	-4.176E-02	-2.497	*
$Year_{2012}$	-1.850E-01	-1.455		1.264E-01	5.738	***	3.719E-02	2.194	*
$Year_{2013}$	-1.907E-01	-1.541		7.143E-03	0.330		-2.421E-02	-1.475	
$Weekend$	6.966E-01	2.017	*	6.883E-01	11.784	***	6.545E-01	15.061	***
$Curve$	-2.234E-01	-0.734		-1.131E-02	-0.221		-4.729E-02	-1.185	
$Grade$	7.971E-01	2.451	*	1.809E-01	3.394	***	3.014E-01	7.356	***
$Hillcrest$	5.427E-01	1.219		2.991E-01	4.064	***	3.445E-01	5.838	***
Sag	-8.215E-02	-0.170		3.390E-01	4.396	***	3.498E-01	5.635	***
Wet	-3.373E-02	-0.104		7.134E-01	13.210	***	6.567E-01	15.924	***
$Snow$	6.920E-01	2.181	*	3.621E-01	6.799	***	4.471E-01	10.944	***
SLR	3.394E-02	0.102		7.174E-01	12.400	***	7.906E-01	18.232	***
$Rushhour$	8.613E-01	2.298	*	1.598E+00	23.750	***	1.759E+00	35.188	***
$Bituminous_1 : RQI$	8.621E-03	0.125		-1.959E-02	-1.464		-5.980E-02	-5.780	***
$Bituminous_2 : RQI$	1.199E-01	1.185		-8.553E-02	-5.023	***	-1.597E-01	-12.326	***
$RQI : Weekend$	5.779E-02	0.547		3.565E-02	1.939	.	2.473E-02	1.800	.
$RQI : Curve$	2.784E-01	2.951	**	6.980E-02	4.268	***	6.846E-02	5.368	***
$RQI : Grade$	-1.873E-01	-1.861	.	6.308E-03	0.371		-6.614E-03	-0.505	
$RQI : Hillcrest$	-1.118E-01	-0.777		-4.066E-02	-1.672	.	-5.272E-02	-2.704	**
$RQI : Sag$	9.596E-02	0.623		-5.401E-02	-2.126	*	-5.011E-02	-2.448	*
$RQI : Wet$	5.531E-02	0.554		-3.886E-02	-2.270	*	-1.769E-02	-1.346	
$RQI : Snow$	-2.395E-01	-2.449	*	2.036E-02	1.202		8.254E-02	6.331	***
$RQI : SLR$	1.436E-01	1.406		-2.290E-02	-1.260		-5.142E-02	-3.758	***
$RQI : Rushhour$	1.082E-01	0.948		-5.638E-02	-2.698	**	-1.114E-01	-7.139	***
AIC	17,311			191,321			260,910		

Legend: . p<0.1; * p<0.05; **p<0.01; *** p<0.001