

**SAFETY ASSESSMENT OF FREEWAY  
ACTIVE TRAFFIC MANAGEMENT BY  
EXPLORING THE RELATIONSHIP  
BETWEEN SAFETY AND CONGESTION**

**Final Report**

**PROJECT SPR 793**



Oregon Department of Transportation



**SAFETY ASSESSMENT OF FREEWAY ACTIVE TRAFFIC  
MANAGEMENT BY EXPLORING THE RELATIONSHIP  
BETWEEN SAFETY AND CONGESTION**

**Final Report**

**PROJECT 793**

by

Xiugang (Joe) Li, Ph.D.  
Research Coordinator

Tony Knudson  
Research Coordinator

for

Oregon Department of Transportation  
Research Section  
555 13<sup>th</sup> Street NE, Suite 1  
Salem OR 97301

and

Federal Highway Administration  
400 Seventh Street, SW  
Washington, DC 20590-0003

**November 2018**



1. Report No. FHWA-OR-RD-19-05	2. Government Accession No.	3. Recipient's Catalog No.	
4. Title and Subtitle Safety Assessment of Freeway Active Traffic Management by Exploring the Relationship between Safety and Congestion		5. Report Date November 2018	
		6. Performing Organization Code	
7. Author(s) Xiugang (Joe) Li, Ph.D., Tony Knudson		8. Performing Organization Report No.	
9. Performing Organization Name and Address  Oregon Department of Transportation Research Section 555 13 <sup>th</sup> Street NE, Suite 1 Salem, OR 97301		10. Work Unit No. (TRAIS)	
		11. Contract or Grant No.	
12. Sponsoring Agency Name and Address  Oregon Dept. of Transportation Research Section 555 13 <sup>th</sup> Street NE, Suite 1 Salem, OR 97301		13. Type of Report and Period Covered  Federal Highway Admin. 400 Seventh Street, SW Washington, DC 20590-0003  Final Report	
		14. Sponsoring Agency Code	
15. Supplementary Notes			
16. Abstract: This study has shown how Oregon crash incident data and PSU Portal traffic data can be combined to determine what factors lead to increased crash risk. Data related to Oregon highway 217 was used to conduct the analysis. The first analysis used Portal traffic data to determine Level of Service (LOS) and then determine the relationship between that and crash rate (Fatal and Injury) which was derived from ODOT's crash database. Similar to studies conducted in other states, there was a clear relationship between LOS and crash rate, with worse LOS associated with increased crash rate. The second part of the study used Portal data again, but this time the mean and variation of the variables speed, occupancy and volume were calculated 5-10 and 10-15 minutes before a crash incident on Oregon 217 on both the upstream and downstream directions. The crash incidents this time were derived from the Traffic Management Operations Center (TMOC) incident data which gave incident times to the nearest minute as opposed to the nearest hour in the LOS study. Given the number of correlated predictors in the data, logistic regression modeling may have led to regression estimates with large variances. Instead, logistic lasso regression was used to select a subset of significant predictor variables to predict the probability of a crash occurring given the traffic conditions at the time. Increasing upstream speed variation and occupancy, and downstream occupancy variation, volume and volume variation were associated with increased crash risk. Slower or decreasing downstream speed was associated with an increased crash risk.			
17. Key Words  Crash Data Analysis, Traffic Data, Crashes, Level of Service, Logistic Lasso Regression, Crash Rate		18. Distribution Statement  Copies available from NTIS, and online at <a href="http://www.oregon.gov/ODOT/TD/TP_RES/">www.oregon.gov/ODOT/TD/TP_RES/</a>	
19. Security Classification (of this report)  Unclassified	20. Security Classification (of this page)  Unclassified	21. No. of Pages  32	22. Price



## SI\* (MODERN METRIC) CONVERSION FACTORS

APPROXIMATE CONVERSIONS TO SI UNITS					APPROXIMATE CONVERSIONS FROM SI UNITS				
Symbol	When You Know	Multiply By	To Find	Symbol	Symbol	When You Know	Multiply By	To Find	Symbol
<b><u>LENGTH</u></b>					<b><u>LENGTH</u></b>				
in	inches	25.4	millimeters	mm	mm	millimeters	0.039	inches	in
ft	feet	0.305	meters	m	m	meters	3.28	Feet	ft
yd	yards	0.914	meters	m	m	meters	1.09	yards	yd
mi	miles	1.61	kilometers	km	km	kilometers	0.621	miles	mi
<b><u>AREA</u></b>					<b><u>AREA</u></b>				
in <sup>2</sup>	square inches	645.2	millimeters squared	mm <sup>2</sup>	mm <sup>2</sup>	millimeters squared	0.0016	square inches	in <sup>2</sup>
ft <sup>2</sup>	square feet	0.093	meters squared	m <sup>2</sup>	m <sup>2</sup>	meters squared	10.764	square feet	ft <sup>2</sup>
yd <sup>2</sup>	square yards	0.836	meters squared	m <sup>2</sup>	m <sup>2</sup>	meters squared	1.196	square yards	yd <sup>2</sup>
ac	acres	0.405	hectares	ha	ha	hectares	2.47	acres	ac
mi <sup>2</sup>	square miles	2.59	kilometers squared	km <sup>2</sup>	km <sup>2</sup>	kilometers squared	0.386	square miles	mi <sup>2</sup>
<b><u>VOLUME</u></b>					<b><u>VOLUME</u></b>				
fl oz	fluid ounces	29.57	milliliters	ml	ml	milliliters	0.034	fluid ounces	fl oz
gal	gallons	3.785	liters	L	L	liters	0.264	gallons	gal
ft <sup>3</sup>	cubic feet	0.028	meters cubed	m <sup>3</sup>	m <sup>3</sup>	meters cubed	35.315	cubic feet	ft <sup>3</sup>
yd <sup>3</sup>	cubic yards	0.765	meters cubed	m <sup>3</sup>	m <sup>3</sup>	meters cubed	1.308	cubic yards	yd <sup>3</sup>
*NOTE: Volumes greater than 1000 L shall be shown in m <sup>3</sup> .									
<b><u>MASS</u></b>					<b><u>MASS</u></b>				
oz	ounces	28.35	grams	g	g	grams	0.035	ounces	oz
lb	pounds	0.454	kilograms	kg	kg	kilograms	2.205	pounds	lb
T	short tons (2000 lb)	0.907	megagrams	Mg	Mg	megagrams	1.102	short tons (2000 lb)	T
<b><u>TEMPERATURE (exact)</u></b>					<b><u>TEMPERATURE (exact)</u></b>				
°F	Fahrenheit	(F-32)/1.8	Celsius	°C	°C	Celsius	1.8C+32	Fahrenheit	°F

\*SI is the symbol for the International System of Measurement





## **ACKNOWLEDGEMENTS**

The authors would like to acknowledge the valuable assistance in this project from members and friends of our Technical Advisory Committee (TAC), who helped us make key decisions at various points of the project and kept us focused on relevance and usability of project outcomes:

Zahidul Siddique, Highway Safety Engineer

Robin Ness, Crash Data Analysis

Dennis Mitchell, ODOT Region 1 Traffic Engineer

Brian Dunn, ODOT Transportation Planning Analysis Unit Manager

Galen McGill, ODOT ITS Unit Manager

We would also like to thank Michael Bufalino for his support and guidance as this research project was conducted.

## **DISCLAIMER**

This document is disseminated under the sponsorship of the Oregon Department of Transportation and the United States Department of Transportation in the interest of information exchange. The State of Oregon and the United States Government assume no liability of its contents or use thereof.

The contents of this report reflect the view of the authors who are solely responsible for the facts and accuracy of the material presented. The contents do not necessarily reflect the official views of the Oregon Department of Transportation or the United States Department of Transportation.

The State of Oregon and the United States Government do not endorse products of manufacturers. Trademarks or manufacturers' names appear herein only because they are considered essential to the object of this document.

This report does not constitute a standard, specification, or regulation.



## TABLE OF CONTENTS

<b>1.0</b>	<b>LITERATURE REVIEW .....</b>	<b>1</b>
1.1	OVERVIEW OF ACTIVE TRAFFIC MANAGEMENT .....	1
1.2	REAL-TIME CRASH-RISK ASSESSMENT .....	1
1.3	REMARKS .....	2
<b>2.0</b>	<b>CRASH RATE AND TRAFFIC DENSITY .....</b>	<b>5</b>
<b>3.0</b>	<b>CRASH PREDICTION .....</b>	<b>9</b>
<b>4.0</b>	<b>CONCLUSION .....</b>	<b>17</b>
<b>5.0</b>	<b>REFERENCES.....</b>	<b>19</b>

## LIST OF TABLES

Table 2.1. LOS Category and Traffic Density (source: (Potts, 2014)) .....	5
Table 3.1 - Traffic Variables and Descriptions.....	10

## LIST OF FIGURES

Figure 2.1 - Crash rate (fatal & injury) and traffic density .....	6
Figure 2.2 - Crash rate (PDO) and traffic density.....	7
Figure 2.3 - Crash rate (total) and traffic density.....	7
Figure 3.1 - Correlation between traffic metric variables.....	11
Figure 3.2 - Effect of upstream speed variation on crash probability.....	13
Figure 3.3 - Effect of upstream occupancy on crash probability .....	13
Figure 3.4 - Effect of downstream speed on crash probability .....	14
Figure 3.5 - Effect of downstream occupancy variation on Crash probability .....	14
Figure 3.6 - Effect of downstream volume on probability of a crash .....	15
Figure 3.7 - Effect of downstream volume variation on probability of a crash.....	15



## **1.0 LITERATURE REVIEW**

### **1.1 OVERVIEW OF ACTIVE TRAFFIC MANAGEMENT**

Active Traffic Management (ATM) improves the efficiency of freeways through the use of traffic operation strategies and automated system technologies to increase throughput and safety. Examples of ATM operation strategies are dynamic speed limits, dynamic lane use control, and weather responsive curve warning. As traffic congestion is a significant concern in many metropolitan areas, ATM may be cost-effective without having to widen freeways to mitigate congestion.

ATM has been deployed in several areas around the United States. Federal Highway Administration (FHWA) summarized some ATM deployments: <https://ops.fhwa.dot.gov/atdm/approaches/atm.htm>. In Oregon the ATM deployments include variable advisory speed system (VAS), adaptive ramp metering and weather responsive curve warning system along OR 217, I-5 and I-405.

In July 2014, Oregon Department of Transportation (ODOT) deployed the OR217 ATM project composed of six systems including Travel Time, Queue Warning, Congestion Responsive Variable Speed, Weather Responsive Variable Speed, Dynamic Ramp-Metering and Curve Warning (ODOT, 2015). The system utilizes sensors to collect traffic and weather conditions in real time, process the data, and distribute information to the motorist by means of variable message and variable speed signs.

ATM has the potential to improve traffic flow by increasing capacity, throughput, and travel time reliability, and decreasing accidents and accident severity (Levecq, 2011). It may harmonize the speeds during congested periods and delay the onset of freeway break down. ODOT (2015) collected peak-hour traffic and crash data for nine months both before and after the ATM project deployment, and found that OR217 ATM reduced total crashes by 21% and buffer time (an index of travel time reliability) by 10%. This evaluation suggests the positive safety benefits of ATM, although long-term safety assessment typically needs three to five years of historical crash data before and after the ATM project deployment.

Recently, exploratory studies on real-time crash-risk assessment have been performed for freeway traffic operations. The methods of real-time assessments may be applied to assess and improve the safety and effectiveness of ATM.

### **1.2 REAL-TIME CRASH-RISK ASSESSMENT**

Real-time crash-risk models are estimated from the real-time traffic data and/or weather data. The models predict the probability of crash occurrence on a specific highway segment to proactively improve safety and possibly prevent the crash occurrence.

Estimated predictive crash-risk models from crash data and real-time traffic data over 16 months on the U.S. 101 NB corridor in California (Pande, Nuworsoo, Shew, & Mineta Transportation Institute, 2012). The crash data were from the California Highway Patrol Incident section of Caltrans' Performance Measurement System (PeMS) database. Real-time traffic data at intervals of 5 minutes was generated by loop detectors. Logistic regression and classification trees were used to estimate the models. The results showed that speed variations were significantly associated with crash risk. They evaluated the models' transferability with data from three other freeways, and found that the models may be transferred to them, but the prediction accuracy was not as good as the freeway that the models were estimated on. The model transfer was improved if they were estimated using only one upstream or downstream vehicle-detection station rather than using two or three vehicle detection stations. They provided an extensive literature review on the previous studies. In addition, (Shew, Pande, & Nuworsoo, 2013) used the same data to estimate a multi-layer perceptron (MLP) neural network model which performed better.

Developed real-time crash risk models using data for 13 months (from October 2010 to October 2011) on a mountainous 15-mile segment of I-70 in Colorado (Ahmed, 2013). This study used Automatic Vehicle Identification (AVI) data for space mean speed, and used Remote Traffic Microwave Sensors (RTMS) data for volume, occupancy, and time mean speed. The authors claimed it was the first study using real-time weather data (precipitation and visibility) gathered by weather stations located on the roadway section. Stochastic Gradient Boosting (SGB), a machine-learning technique, was used to estimate models from the real-time data at intervals of 6 minutes. The models were not tested using data from other freeways for transferability.

Used the data from the 15-mile segment on I-70 in Colorado to estimate a real-time risk model by using Support Vector Machines (SVM) (Yu, 2013). In addition, they used traffic simulations to develop an algorithm of Variable Speed Limits (VSL) to reduce crash risk. The logistic regression model was estimated to measure the crash risk. The traffic simulation showed that the VSL reduced crash risk from 0.2% to 11.8% at different locations.

Published a follow-on report from their work a year earlier, where they looked at data from the Gardiner Expressway in Toronto Canada (Lee, Hellinga, & Saccomanno, January 2003). Initially they looked at the crash precursors of coefficient of variation in speed within and across lanes, density, weather, and proportion of peak period volume to create a log-linear model of exposure when a crash has occurred. They used an interesting approach to determine time of crash by looking for sudden drops in speed at detectors upstream of the reported crash.

In their follow-on study, they looked at rational methods to determine the correct crash precursors. The precursors found to have different distributions when crashes occurred, versus when there were no crashes should be the main criteria for their selection.

### **1.3 REMARKS**

Real-time crash-risk modeling has been studied in order to proactively prevent crash occurrence and improve traffic safety. The ATM may adopt the models to change the traffic flow, such as speed variation reduction, to reduce crash risk and prevent crash occurrence.

OR-217 ATM can change upstream traffic speed if an incident or crash happens (DKS, 2015). For example, after a minor incident happened, the inductive loops detect slow speed. Then the VAS sign display “SLOW” at the upstream message board. Similarly, the VAS can adjust the upstream advisory speed after a crash happened. But the OR-217 ATM cannot predict the times and locations of increased risk of crashes, and adjust the traffic speed to prevent the crashes proactively.

Previous studies predicted crash risk 5 to 6 minutes ahead of time. Longer prediction times, such as 10 minutes or 15 minutes as done in this study, can provide more time to respond. Previous studies of real-time crash-risk prediction were for freeways, whereas Oregon has ATM on two-lane highways. The models estimated from the two-lane highways in the future could be adopted by the ATM to reduce crashes.

The real-time crash-risk models can be used to assess the safety of highway operations in the short term even without the ATM. The application includes most segments of freeways (or some two-lane highways in our plan). Further, for segments with the ATM, the assessment helps the ATM change the traffic flow to prevent crash occurrences.

The potential use of this research project includes

- Short-term safety assessments of freeway or highway traffic operation including the corridor with the ATM, such as OR 217. Currently the long-term evaluation needs at least 5-year historical crash data.
- The research predicts times and locations with higher crash risk in real-time (5 minutes or 15 minutes ahead). The ATM may adopt the prediction to proactively prevent the crash occurrences.
- The research could help Traffic Incident Management (TIM) responders to reduce time needed for arrival on scene. Reduced lane clearance time results in reduction of secondary crash risk.





## 2.0 CRASH RATE AND TRAFFIC DENSITY

The Strategic Highway Research Program 2 (SHRP 2) project “Further Development of the Safety and Congestion Relationship for Urban Freeways” (Potts, 2014) produced the relationship between crash rate and traffic density by using data divided into 15-minute intervals. The results for different cities, such as Seattle and Minneapolis-St. Paul, produced unique results. The project did not use Oregon data, so it’s not clear if the results can be used in Oregon for applications such as the evaluation of ATM.

This study used crash and traffic data from OR 217 to produce a relationship between crash rate and traffic density. ODOT Crash Analysis and Reporting Unit (CARU) provided the crash data. The time of the crash was estimated to the nearest hour, so the analysis divided the time into one-hour intervals. The analysis may be less accurate but it did successfully produce a similar relationship to results from other studies.

The analysis aggregated the crash and traffic data at one-hour intervals and divided highway OR217 into segments. The start and end points of the segments are typically interchanges. The segment length is approximately 0.5 to 1.2 miles. Using ten years of data from 2005 to 2014 for each segment, traffic data (volume, speed and occupancy) was calculated from detectors installed on that segment. VMT was used to estimate traffic density which was then used to estimate the Level of Service (LOS).

LOS represents traffic conditions, such as congestion. The Highway Capacity Manual (National Research Council (U.S.), 2010) defines LOS using traffic density. This study adopted the LOS categories defined by (Potts, 2014), listed in Table 2.1.

**Table 2.1. LOS Category and Traffic Density (source: (Potts, 2014))**

LOS Category	Traffic Density (pc/mi/ln)	LOS Category	Traffic Density (pc/mi/ln)
A+	0-3	D+	26-29
A	3-7	D	29-32
A-	7-11	D-	32-35
B+	11-13	E+	35-38
B	13-15	E	38-41
B-	15-18	E-	41-45
C+	18-20	F+	45-50
C	20-23	F	50-55
C-	23-26	F-	55+

The average crash rate and traffic density were calculated for each category, then the relationship between traffic density and crash rate was estimated. The graphs shown as Figures 1 to 3 display this relationship for fatal and injury, Property Damage Only (PDO), and total respectively.

Each type of crash rate is much higher at LOS E and F, and increases with density. At LOS A, the crash rate decreases with density. This result is similar to Seattle, but different from Minneapolis-St. Paul. The crash rates are lower than those reported by (Potts, 2014) for Seattle.

The regression equations are as follows. Cr is the crash rate (crashes per million VMT) and d is traffic density, passenger cars per mile per lane (pc/mi/ln).

Fatal and Injury Crash:

$$Cr = 0.00003d^3 + 0.003d^2 - 0.0709d + 0.7922 \quad (2-1)$$

Property Damage Only Crash:

$$Cr = -0.00006d^3 + 0.006d^2 - 0.1453d + 1.3628 \quad (2-2)$$

Total Crash:

$$Cr = -0.00008d^3 + 0.009d^2 - 0.2162d + 2.155 \quad (2-3)$$

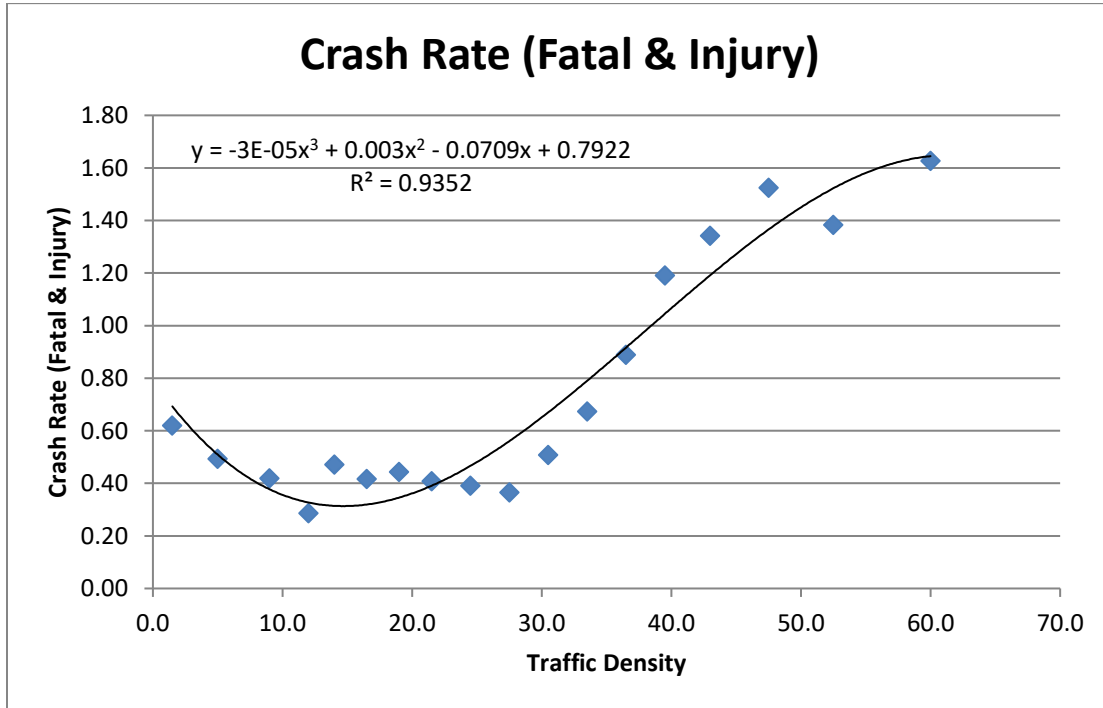
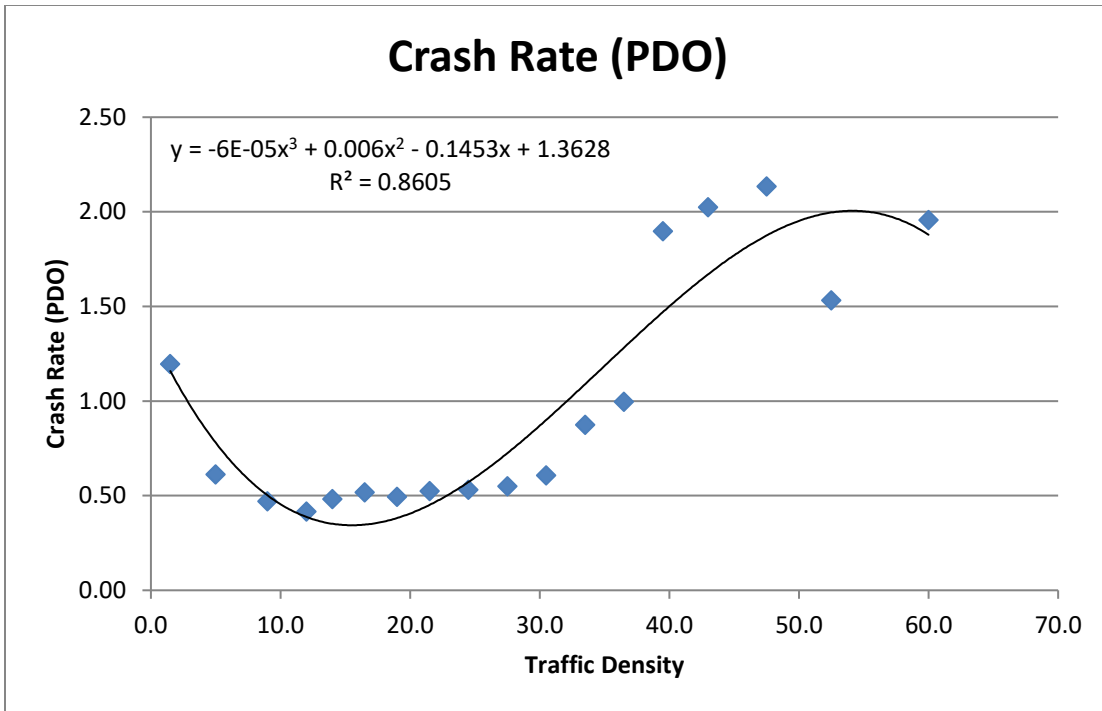
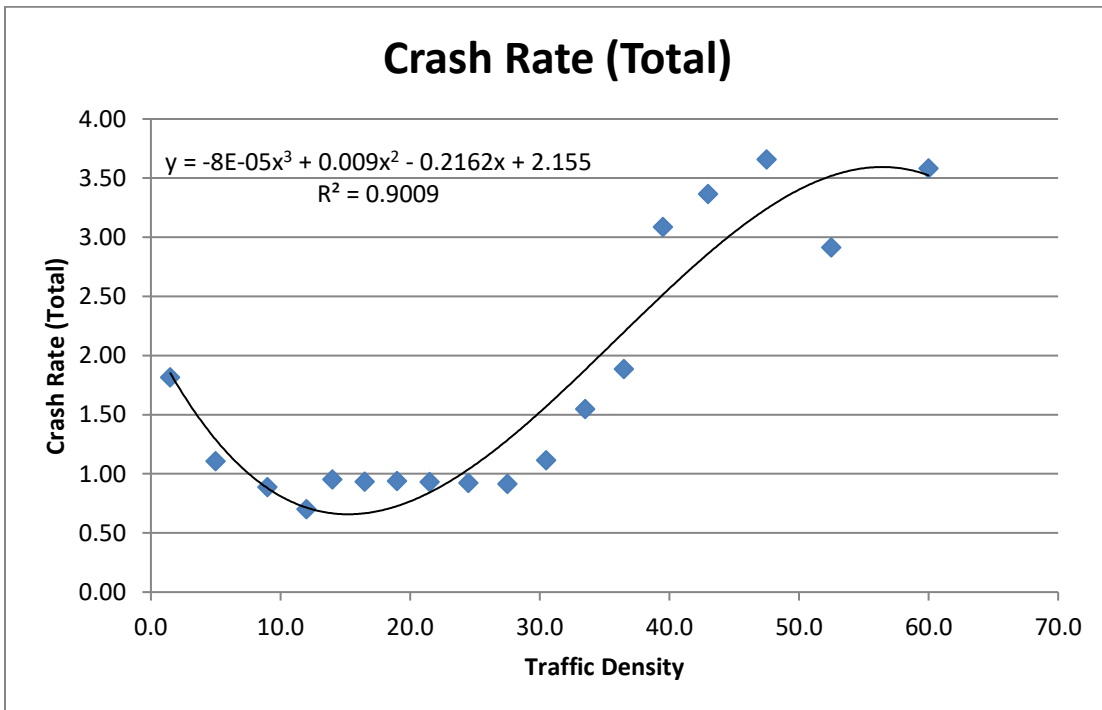


Figure 2.1 - Crash rate (fatal & injury) and traffic density



**Figure 2.1 - Crash rate (PDO) and traffic density**



**Figure 2.2 - Crash rate (total) and traffic density**



### 3.0 CRASH PREDICTION

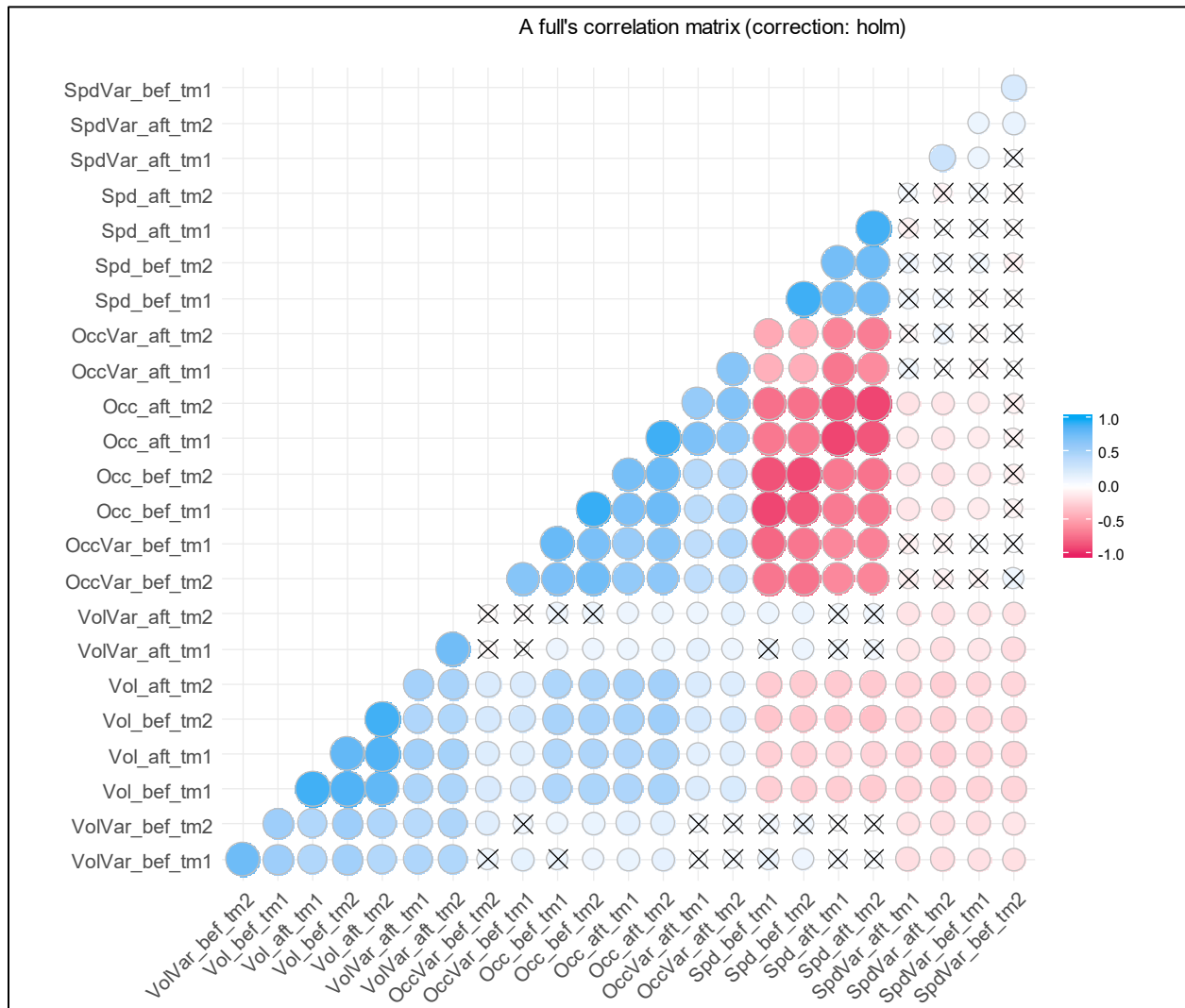
Using incident data from the Traffic Management Operations Center (TMOC) which provides time of incidents to the nearest minute instead of hour, a logistic regression model was fit to determine which variables were significant in predicting crashes. Incident data from OR-217 from July 2015 to August 2017 was used and only records containing Event\_type “Crash” or “Fatal Crash” were kept for further analysis. Both northbound and southbound roadways were segmented by milepost based on traffic detector locations. Random times with no crashes were also selected.

Data related to traffic volume, speed, and occupancy was downloaded for each detector site for the same time period as the incident data from the Portal site at Portland State University. Volume, occupancy, and speed were aggregated at five minute intervals based on start times from the Portal data set. Then for each of those variables by detector station located before and after the incident, the volume for five to ten minutes and ten to fifteen minutes before each crash or non-crash time period was summed. The same was done to calculate variation of volume, speed (calculated by summing up speed and then dividing by volume), variation of speed, mean occupancy and variation of occupancy. A table of the variables and their descriptions are in Table 3.1 below.

**Table 3.1 - Traffic Variables and Descriptions**

<b>Variable</b>	<b>Description</b>
<b>Vol_bef_tm1</b>	Sum of traffic volume from upstream station 5-10 minutes before incident.
<b>VolVar_bef_tm1</b>	Traffic volume variation from upstream station 5-10 minutes before incident.
<b>Spd_bef_tm1</b>	Mean traffic speeds from upstream station 5-10 minutes before incident.
<b>SpdVar_bef_tm1</b>	Traffic speed variation from upstream station 5-10 minutes before incident.
<b>Occ_bef_tm1</b>	Mean occupancy from upstream station 5-10 minutes before incident.
<b>OccVar_bef_tm1</b>	Mean occupancy variation from upstream station 5-10 minutes before incident.
<b>Vol_bef_tm2</b>	Sum of traffic volume from upstream station 10-15 minutes before incident.
<b>VolVar_bef_tm2</b>	Traffic volume variation from upstream station 10-15 minutes before incident.
<b>Spd_bef_tm2</b>	Mean speed from upstream station 10-15 minutes before incident.
<b>SpdVar_bef_tm2</b>	Speed variation from upstream station 10-15 minutes before incident.
<b>Occ_bef_tm2</b>	Mean occupancy from upstream station 10-15 minutes before incident.
<b>OccVar_bef_tm2</b>	Occupancy variation from upstream station 10-15 minutes before incident.
<b>Vol_aft_tm1</b>	Sum of traffic volume from downstream station 5-10 minutes before incident.
<b>VolVar_aft_tm1</b>	Traffic volume variation from downstream station 5-10 minutes before incident.
<b>Spd_aft_tm1</b>	Mean traffic speeds from downstream station 5-10 minutes before incident.
<b>SpdVar_aft_tm1</b>	Traffic speed variation from downstream station 5-10 minutes before incident.
<b>Occ_aft_tm1</b>	Mean occupancy from downstream station 5-10 minutes before incident.
<b>OccVar_aft_tm1</b>	Mean occupancy variation from downstream station 5-10 minutes before incident.
<b>Vol_aft_tm2</b>	Sum of traffic volume from downstream station 10-15 minutes before incident.
<b>VolVar_aft_tm2</b>	Traffic volume variation from downstream station 10-15 minutes before incident.
<b>Spd_aft_tm2</b>	Mean speed from downstream station 10-15 minutes before incident.
<b>SpdVar_aft_tm2</b>	Speed variation from downstream station 10-15 minutes before incident.
<b>Occ_aft_tm2</b>	Mean occupancy from downstream station 10-15 minutes before incident.
<b>OccVar_aft_tm2</b>	Occupancy variation from downstream station 10-15 minutes before incident.

Several of these variables are significantly correlated with each other as Figure 3.1 shows. There are negative correlations between speed and occupancy and positive correlations between volume and occupancy metrics and other volume and occupancy variables respectively.



**Figure 3.1 - Correlation between traffic metric variables**

Given the number of correlated predictors, methods such as step-wise least squares were not used given the number of issues with those methods, a good summary of those issues can be found in Warren R. Thompson's paper. (Thompson, 2009). Multicollinearity among predictors in logistic regression modeling is especially problematic since it can lead to regression estimates with large variances. Instead, logistic lasso regression was used to select a subset of significant predictor variables to predict the probability of a crash occurring given the traffic conditions at the time.

If  $Y$  is our dependent variable and our explanatory variables are  $x_1, \dots, x_k$ , then the logistic model is given by:

$$\text{Logit } \{Y = 1|x\} = \ln\left(\frac{P(Y=1|x)}{1-P(Y=1|x)}\right) = \beta_0 + \beta_1x_1 + \beta_2x_2 + \dots + \beta_kx_k \quad (3-1)$$

Using least squares to estimate  $\beta_0, \beta_1, \dots, \beta_k$  requires one to minimize Residual Sums of Squares (RSS), defined by

$$\text{RSS} = \sum_{i=1}^n (y_i - \beta_0 - \sum_{j=1}^k \beta_j x_{ij})^2 \quad (3-2)$$

Lasso regression is similar, except the coefficients are estimated by minimizing the following

$$\sum_{i=1}^n (y_i - \beta_0 - \sum_{j=1}^k \beta_j x_{ij})^2 + \lambda \sum_{j=1}^k |\beta_j| \quad (3-3)$$

This estimation attempts to shrink the coefficient estimates to zero and will set some of the estimates to exactly zero if the tuning parameter  $\lambda$  is sufficiently large enough. This gives a sparse model with only a subset of the variables selected and is easily interpreted. Using 10-fold cross validation `cv.glmnet` (Friedman, Hastie, & Tibshirani, 2010) in R, the procedure fit possible  $\lambda$  values on the data, chose the best  $\lambda$  and then trained the model with appropriate parameters.

Using the crash data, the odds of a crash using the sparse subset of predictors is shown in equation (3-4) below

$$\text{Odds} = e^{(-0.36398 + 0.0004*\text{SpdVar\_bef\_tm1} + 0.0428*\text{Occ\_bef\_tm1} + -0.0414*\text{Spd\_aft\_tm1} + 0.0009*\text{OccVar\_aft\_tm1} + 0.0005*\text{Vol\_aft\_tm2} + 0.0304*\text{VolVar\_aft\_tm2})} \quad (3-4)$$

As you can see, some selected parameters will have a smaller effect on the odds since they are closer to zero than others since the data was normalized before the cross validation procedure was done.

Finally, we can calculate the probability of a crash for various values of each subset parameter and graph the results by calculating

$$\mathbf{P} \{Y = 1|x\} = \frac{\text{Odds}}{1+\text{Odds}} \quad (3-5)$$

Graphing how the probability of a crash changes as you vary one of the model parameters over its range while holding the others at their mean is shown in Figures 3-2 through 3-7. Two of the parameters having the largest effects on the probability of a crash were upstream occupancy and downstream speed 5 – 10 minutes before a crash. Increasing upstream occupancy leads to



increasing probabilities for a crash, and decreasing downstream speeds also increase that probability.

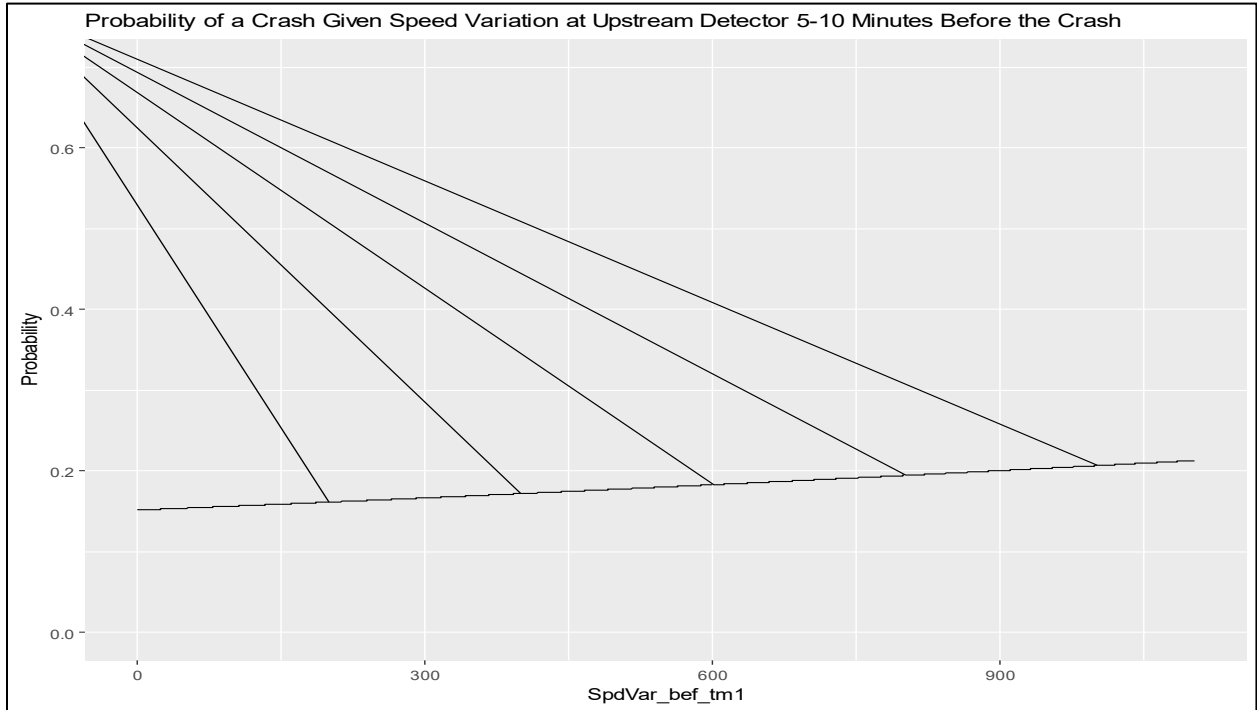


Figure 3.2 - Effect of upstream speed variation on crash probability

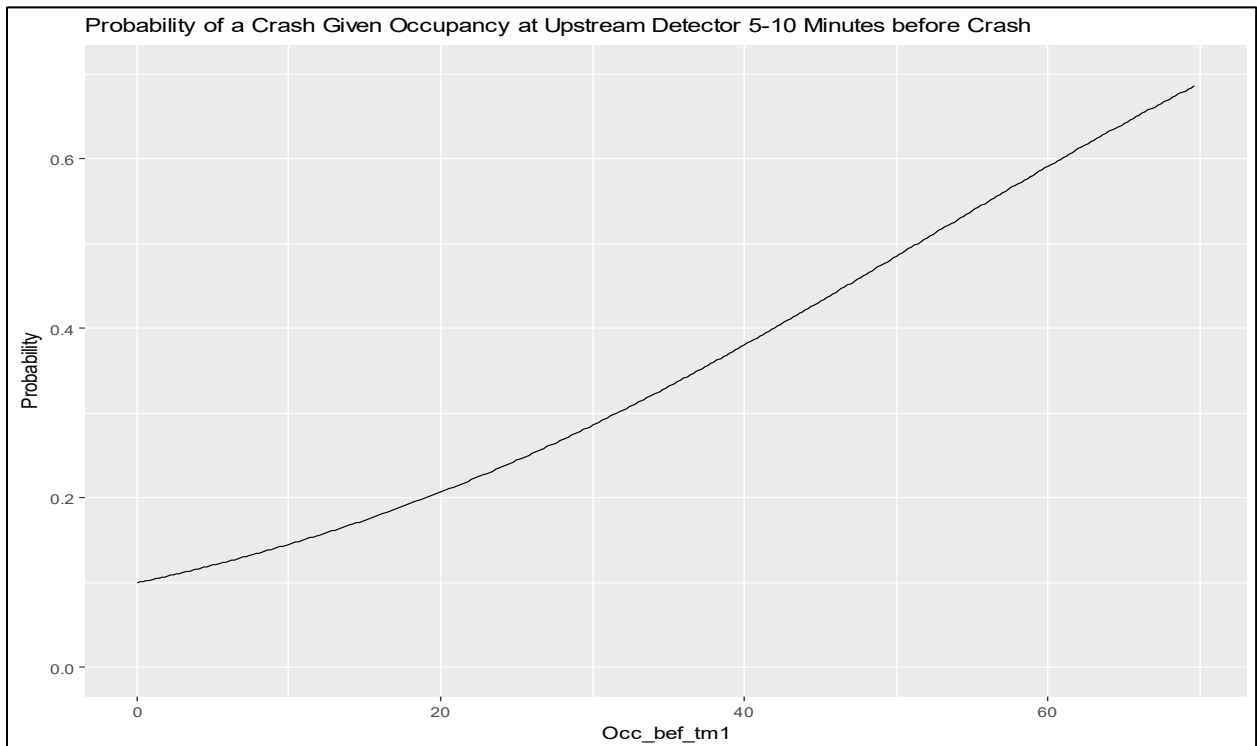
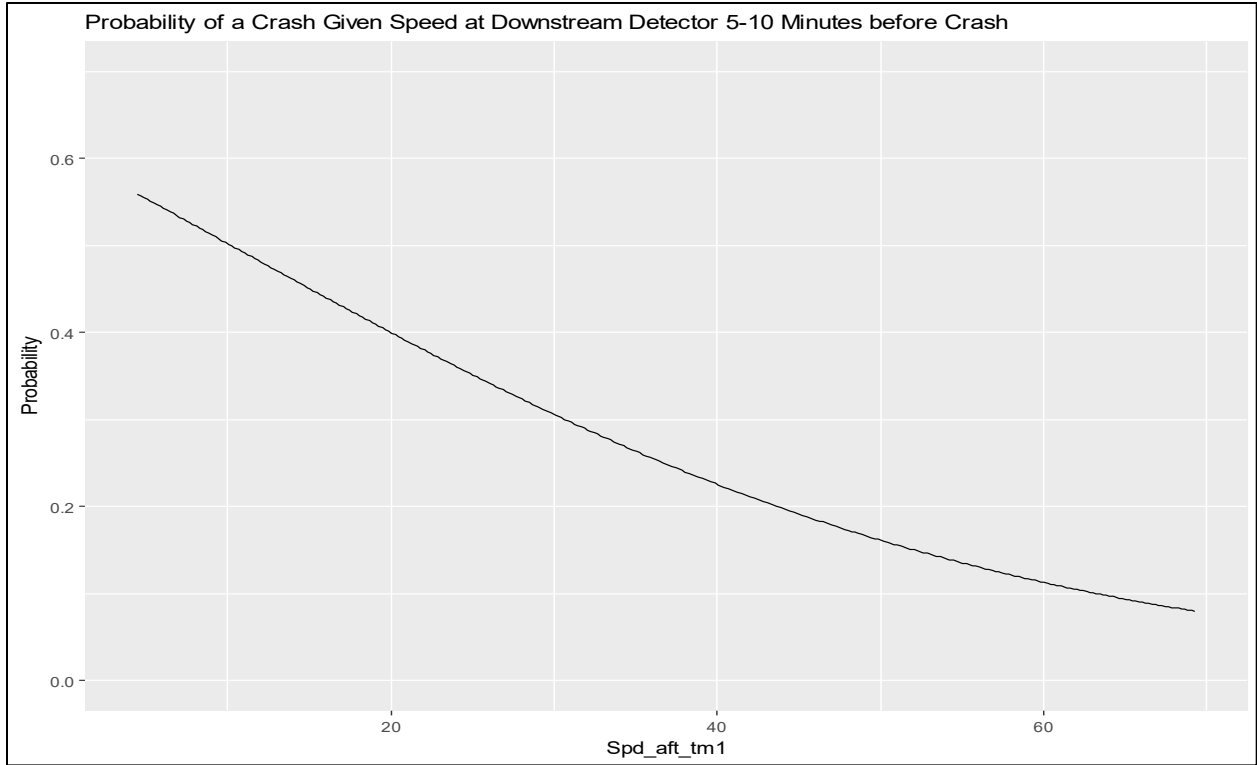
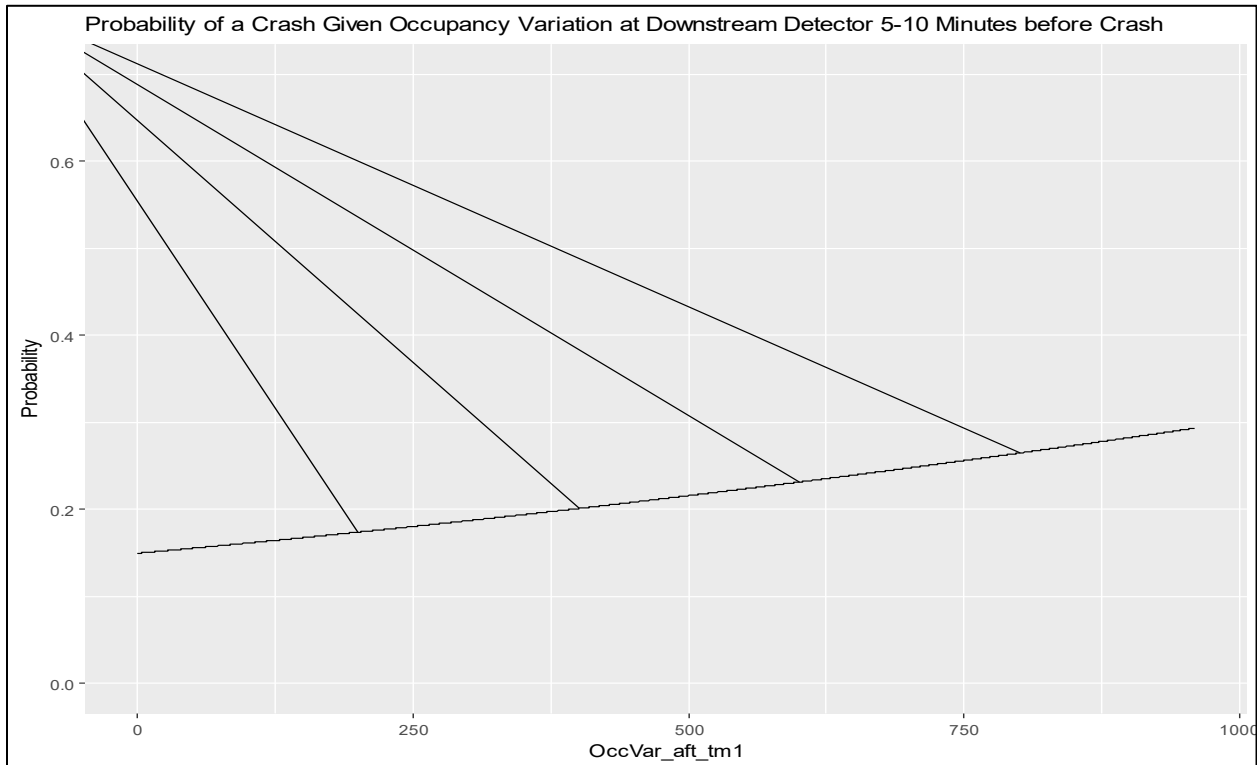


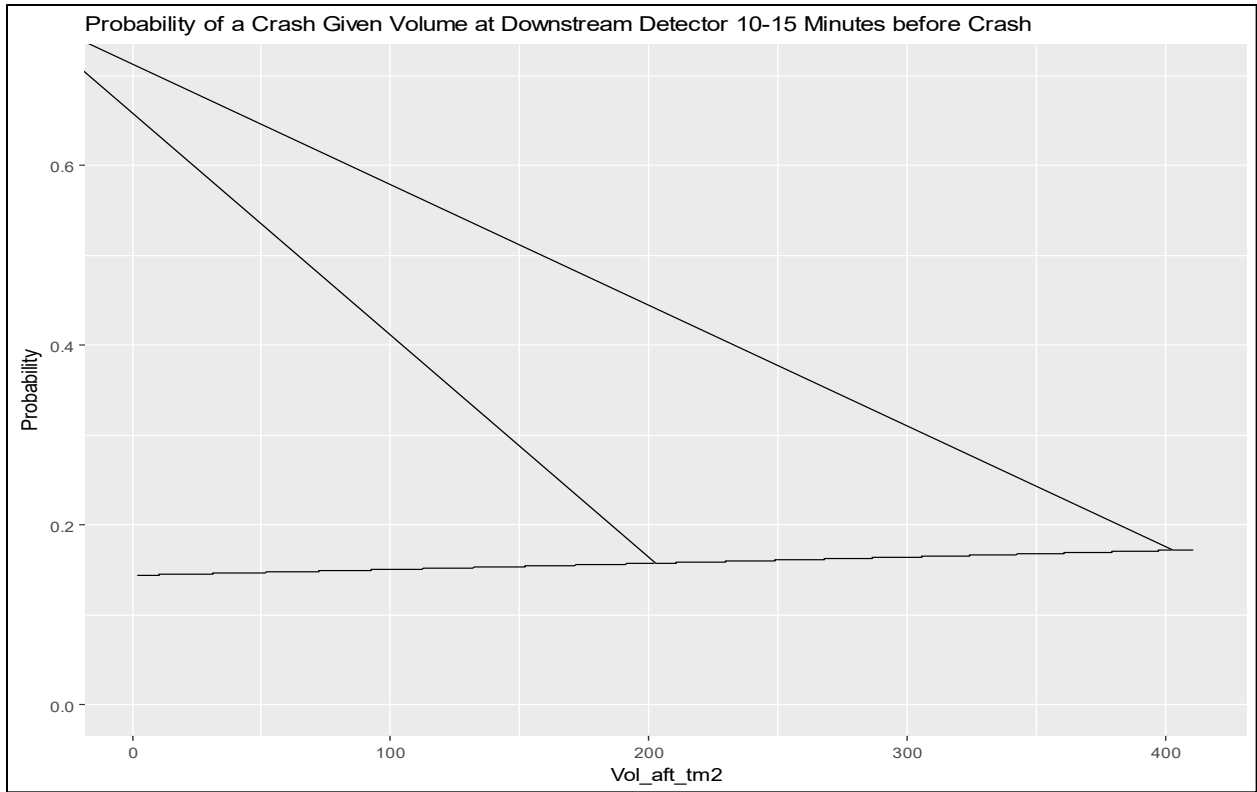
Figure 3.3 - Effect of upstream occupancy on crash probability



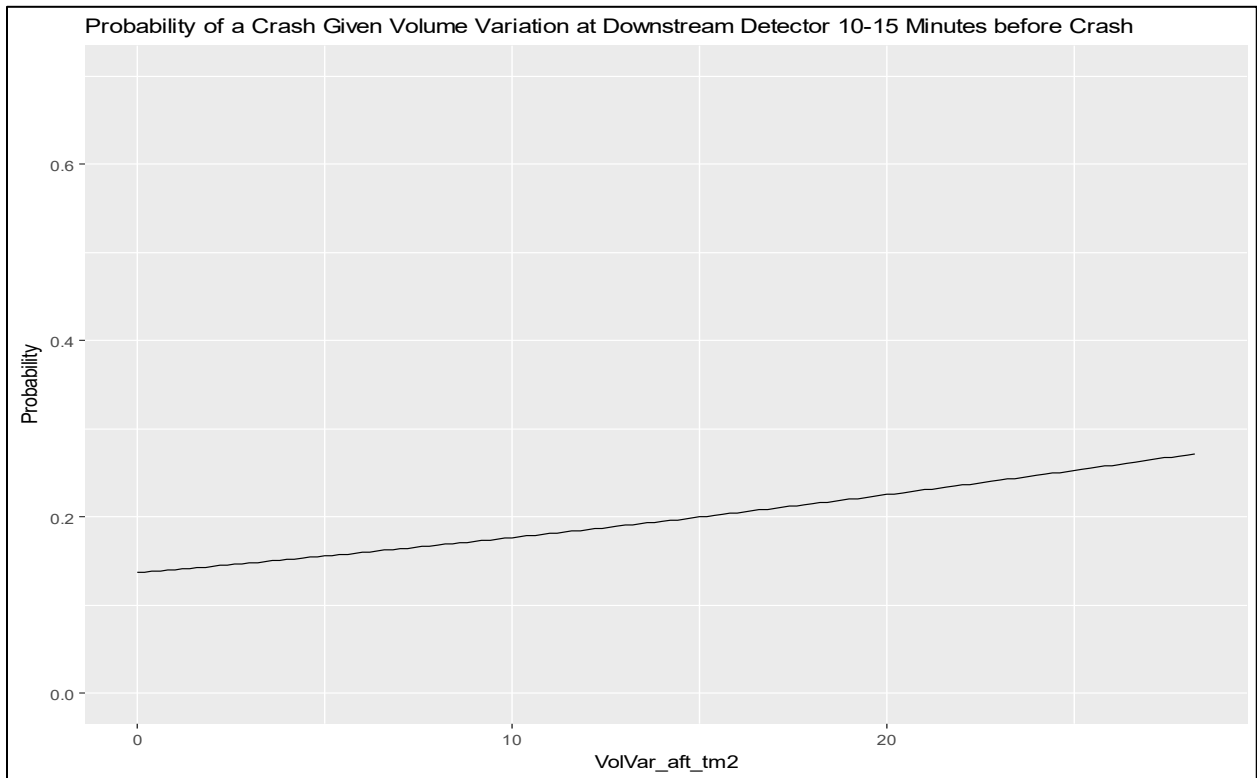
**Figure 3.4 - Effect of downstream speed on crash probability**



**Figure 3.5 - Effect of downstream occupancy variation on crash probability**



**Figure 3.6 - Effect of downstream volume on probability of a crash**



**Figure 3.7 - Effect of downstream volume variation on probability of a crash**



## **4.0 CONCLUSION**

The previous discussion has shown how Oregon crash incident data and PSU Portal traffic data can be combined to determine what factors lead to increased crash risk. A clear link was shown to exist between traffic density and crash risk, as well as between occupancy, speed, and volume and crash probability.

A follow-up to this work would be to see if using traffic speed data by itself could predict the probability of a crash, for example using HERE or Waze data. Utilizing that data could expand the method to highways which don't have access to loop or video detector technology.



## 5.0 REFERENCES

- Ahmed, M., & Abdel-Aty, M. (January 01, 2013). A data fusion framework for real-time risk assessment on freeways. *Transportation Research Part C*, 26, 203-213.
- Lee, C., Hellinga, B., & Saccomanno, F. (January 01, 2003). Real-time crash prediction model for application to crash prevention in freeway traffic. *Transportation Research Record*, 1840, 67-77.
- DKS Associates. (2015). Oregon Statewide Variable Speed System Concept of Operations (Publication No. 0-4). Salem, OR: Oregon Department of Transportation. Retrieved from [https://o10019.eos-intl.net/elibsql17\\_O10019\\_Documents/DKS\\_2015-03-02 State VSS ConOps v0-4 - Final.pdf](https://o10019.eos-intl.net/elibsql17_O10019_Documents/DKS_2015-03-02_State_VSS_ConOps_v0-4_Final.pdf)
- Federal Highway Administration. (n.d.). Active Traffic Management. Retrieved June, 2017, from <https://ops.fhwa.dot.gov/atdm/approaches/atm.htm>
- Friedman, J., Hastie, T., & Tibshirani, R.. (February 01, 2010). Regularization Paths for Generalized Linear Models via Coordinate Descent. *Journal of Statistical Software*, 33, 1. Retrieved June 2018 from <https://www.jstatsoft.org/v33/i01/>
- Levecq, C., Kuhn, B. T., Jasek, D., Texas Transportation Institute., United States., & University Transportation Centers Program (U.S.). (2011). *General guidelines for active traffic management deployment*. College Station, TX: University Transportation Center for Mobility, Texas Transportation Institute, Texas A & M University System.
- National Research Council (U.S.). (2005). *Highway capacity manual*. Washington, D.C: Transportation Research Board of the National Academies.
- Mitchell, D. (2015). *OR 217 Active Traffic Management (Rep.)*. Salem, OR: Oregon Department of Transportation. Retrieved from [https://transops.s3.amazonaws.com/uploaded\\_files/ODOT-ITS\\_2015-Award\\_Submittal-OR-217-Active-Traffic-Management.pdf](https://transops.s3.amazonaws.com/uploaded_files/ODOT-ITS_2015-Award_Submittal-OR-217-Active-Traffic-Management.pdf).
- Pande, A., Nuworsoo, C., Shew, C., & Mineta Transportation Institute. (2012). *Proactive assessment of accident risk to improve safety on a system of freeways*. San José, CA: Mineta Transportation Institute, College of Business, San José State University.
- Potts, I. B., National Research Council (U.S.), & Second Strategic Highway Research Program (U.S.). (2015). *Further development of the safety and congestion relationship for urban freeways*.
- Shew, C., Pande, A., & Nuworsoo, C. (January 01, 2013). Transferability and robustness of real-time freeway crash risk assessment. *Journal of Safety Research*, 46, 83-90.

Thompson, W. R. (2009). *Variable selection of correlated predictors in logistic regression: Investigating the diet-heart hypothesis*. Tallahassee, Florida: Florida State University.

Yu, R. (2013). *Real-time traffic safety evaluation models and their application for variable speed limits*. Orlando, Fla: University of Central Florida.