MODELING TRAFFIC ACCIDENTS AT SIGNALIZED INTERSECTIONS IN THE CITY OF NORFOLK, VA

Sharad Maheshwari, Hampton University, Hampton, VA.
Kelwyn A. D'Souza, Hampton University, Hampton, VA.

ABSTRACT

This study was an attempt to apply a proactive approach using traffic pattern and signalized intersection characteristics to predict accident rates at signalized intersections in a city's arterial network. An earlier analysis of accident data at selected intersections within the City of Norfolk indicated that in addition to traffic volume, other controllable factors contributed to traffic accidents at specific intersections. These factors included area topography, lane patterns, type of road signs, turning lanes, etc. It is also known that administrative factors such as signal types, signal polices, road closures, etc., and maintenance factors such as road conditions, condition of the signals, condition of road signs, etc. also impact road accidents.

The objective of this study was to relate these variables to accident rate and delineate variables that are statistically more significant for accident rate. Data on several topographical variables was collected in the City of Norfolk. These variables included number of lanes, turn lanes, pedestrian crossing, restricted lanes, etc. A linear regression model was used to establish relationship between these variables and the accident rate. The resulting regression model explained 60% of the variability. It also showed that four topographical variables are more important than other variables. These variables include number of lanes, number of turn lanes, presence of median and presence of permanent hazard like railway crossing. However, validation of model showed higher than expected variation. The model developed, in this study, overestimates the accident rate by 33%, thus, limiting its practical application.

INTRODUCTION

The main objective of this research was to study the signalized intersection in a city to delineate intersection geometry and design factors which may be contributing to traffic accidents. The City of Norfolk was selected for this study since it is one of the largest and oldest cities in the Hampton Roads region; and is home to roughly quarter million people. In 2006 the Hampton Roads had the highest crash incidents in the state based on the millions of VMT (vehicle mile traveled) (Nichols, 2007). The City of Norfolk contributed roughly 17% of those crashes with annual traffic accident count of approximately 5,400.

The literature review shows that road design factors could impact traffic safety. Several highway engineering factors like lane widths, shoulder widths, horizontal curvature, vertical curvature, super-elevation rate, median, auxiliary lane, etc. are estimated based on some traffic safety considerations. Additional factors like road signage, vegetation, line sight of signal especially on horizontal and vertical curvature, and number of driveways have also been reported to have impact on the traffic safety. To study the impact of these factors along with traffic control rules, researchers have utilized variety of statistical models. The most often used model is multivariate regression where the dependent variable is generally based on traffic accidents and a set of independent variables including roadway design, traffic control, demographic variables, etc. The negative binomial model is used to account for large variability among the accident rates on different intersections. Research results show relationship exists between the various roadway design and control factors and traffic accidents. Research also indicates divergences on the importance of individual factor on the traffic safety. There is reported difference based on the regional demographic factors indicating regional accident rate differences due to interactions between design/control factors and local driving population. This study was designed to understand the impact of the road design factors on the traffic accident rate in a local area.

This study was preceded by a pilot study conducted in the City of Norfolk for signalized intersections (Maheshwari and D'Souza, 2008). An intersection accident was defined as any accident occurring within the 250' of the intersection. The pilot study results showed that intersection topography/design factors and traffic control rules have positive relationship with the traffic accident rate. These factors included number of driveways, pedestrian crossing, and presence of physical median. Despite indicating number of positive relationships, the pilot study results could not be generalized as the sample size was very small. A sample of ten intersections was selected based solely on the accident rate. Also, the pilot study model was not validated for other intersections in the City. Hence, it was logical to further investigate the impact of topographical and other controllable factors on the traffic accidents with an expanded sample size and validate the model using other intersections within the City.

LITERATURE REVIEW

Automobile accidents contribute to staggering amount of property damage and large number of deaths in United States. According to the Insurance Information Institute, New York (Hot Topic and Issues Update: Auto Crashes, 2006), 42,636 people died in motor vehicle crashes in 2004 alone and an additional 2,788,000 people were injured. There were over 6 million police reported auto accidents in 2004. It is reported that about 50% of crashes occur at the intersections (Hakkert & Mahalel, 1978; National Highway R&T Partnership, 2002). It has been reported extensively in the literature that traffic volume is the major explanatory factor for traffic accidents (Vogt, 1999). However, studies have been carried out showing that design and other related factors contribute towards 2% - 14% of accidents. Ogden, et al., 1994 reported that about 21% of the variation in accidents was explained by variations in traffic flow volume, while the remaining majority of the variation was explained by other related factors. Vogt (1999) provides an extensive review of the factors, which have been considered in past research studies. These factors include channelization (right and left turn lane), sight distance, intersection angle, median width, surface width, shoulder width, signal characteristics, lighting, roadside condition, truck percentage in the traffic volume, posted speed, weather, etc. Beside these factors, researchers have also considered other minor details such as surface bumps, potholes, pavement roughness, pavement edge drop-off, etc. (Graves, et al., 2005).

The relationship between the accidents and pertinent factors is usually established using multivariate analysis (Corben & Foong, 1990; Hakkert & Mahalel, 1978; Ogden, et al., 1994; Ogden and Newstead, 1994; Vogt, 1999). A study by Corben and Foong, 1990 led to development of a seven-variable linear regression model for predicting right-turn crashes at signalized intersections. This model explained 85% of the variance of accident occurrence. In a FHWA study by Harwood, et al., (2000), quantitative data on accidents and other factors were combined with the expert's judgment about design factors as well as expected impact of these design factors on the accident rate. Mountain, Fawaz & Jarrett (1996) showed in a British study that the road design features- link length relates to accident rate, especially in dual carriageway. Retting, et at. (2001) studied the affect of roundabout on the traffic accidents; and found that replacing signals or stop signs with roundabouts could reduce traffic accidents. Road design factors like, the curve radii, spiral lengths, lane width, shoulder width, and tangent lengths are shown to be related the collusion frequency (Easa and Mehmood, 2008). It was exhibited through a comprehensive study of Korean road accident data that three categories of factors influence the accident rate -- road geometric condition, driver characteristic and vehicle type (Lee, Chung & Son, 2008). Wang, Quddus and Ison (2009) studied roadways based on congestion and reported that beside traffic volume, segment length, number or lanes, curvature and gradient also influence the accident rates.

Malyshkina and Mannering (2010) studied the impact of design exceptions allowed in the highway construction on the traffic accident rate (design exception: safety deviation in roadway design factors). They found exceptions don't necessarily increase accidents in their dataset. In another analysis of the data of 10 Canadian cities, Andrey (2010) related weather and accident rates and found that accident rates drop under severe weather conditions.

It is clear from the research that variety of statistical models are used for traffic accident analsysis, however, it is evident from the literature that negative binominal or Poisson distribution is often employed in relating the frequency of accidents to design factors (Lord, Guikema, Geedipally, 2008; Malyshkina and Mannering ,2010; Shankar, Manning and Barfield 1995; and Wang and Abdel-Aty, 2008). The technique is largely used to account for the higher variability in the frequency of accidents at different intersection. For example, Shankar, et al. (1995) used negative binomial distribution to show interaction between roadway geometry factors and weather accidents. They showed that certain geometry elements are more critical during the severe weather conditions. Milton, Shankar, & Mannering (2008) used logit model to include several parameters like weather, type of traffic, and road geometry.

Recent studies have applied data mining techniques along with statistical modeling to determine the impact of major factors like traffic volume and road design characteristics along with minor factors such as potholes and surface roughness. Graves, et al., (2005) reported about the impact of potholes and surface roughness on the accident rate. However, due to the paucity of data, a clear link could not be established between these surface factors (pot holes, roughness, etc.). Washington, et al., (2005) performed an extensive study to validate previously reported accident prediction models and methods. Validation was performed using recalculation of original model coefficients using additional year's data as well as using data from a different state. The study reported that beside traffic volume other factors should be considered on a case-by-case basis for a given site.

The literature discusses the variety of factors affecting traffic accidents including road geometry, layout and traffic control factors. However, there is divergence of opinion on what factors have more influence on safety. Also, there are regional differences in the importance of factors which influence safety. Studies on rural highways are not directly applicable to urban settings as the traffic pattern and other factors differ at rural and city intersections. Furthermore, before and after studies may be less valuable in rural settings as road design changes are not made as often as in a city with growing traffic volume. Moreover, literature shows that traffic accident analyses are commonly conducted in a larger geographical area (one or more states). This research was build upon past research and evidence from the literature to apply a systematic approach of identifying factors in accident-prone intersections in a city such as Norfolk, VA and analyzing factors which could significantly influence the accident rates in that specific area.

METHODOLOGY

The approach in this research was to collect and analyze data from the intersections with higher accident rates in the City. Restricting data set to higher accident intersections allows to reduce the variability in the data set. Therefore, it makes generalized linear model (GLM) applicable for the analysis of the data set. The study was conducted for the signalized intersections within the City of Norfolk, VA. The study concentrated on 65 signalized intersections that experienced high accident rates during the period 2001-2004. Thirty of these intersections were selected for the analysis and 10 were used for validation. Rest of the intersections could not be used because of traffic count data for those intersections were not available.

The City of Norfolk has stored traffic accident data in an electronic format for the past 11 years from 1994 to 2004. Only accidents related to single vehicles were considered in the study due to technical limitations of importing multi-vehicle into the available database. The City's accident database was developed from individual police accident reports that included type of accident, road conditions, traffic signs and signal, drivers' actions, vehicle(s) condition, demographic data, nature of injury, and other related information, all of which are subsequently entered in the City's accident database. The traffic accidents without a police report were not included in this database hence those accident were not part of this study. The traffic volume data, Annual Average Daily Traffic (AADT), was obtained from the Department of Transportation, Commonwealth of Virginia. Some of the local and feeder roads traffic count were not available hence those intersections were eliminated from the study.

The physical attributes included number of lanes, type of lanes, type of turn signals, existence of median and shoulder, pedestrian crossing, number of driveways within 250' of the intersection, and other safety features. A schematic of the intersection is shown in Figure 1. For each intersection, 56 different physical attributes were collected. The AADT data was collected from the Department of Transportation, Commonwealth of Virginia. A review of data revealed that the certain variables could be eliminated as they were rarely present in the data collected, this included shoulder variables and no right turn signal. This reduced the variable set to 44 independent variables.

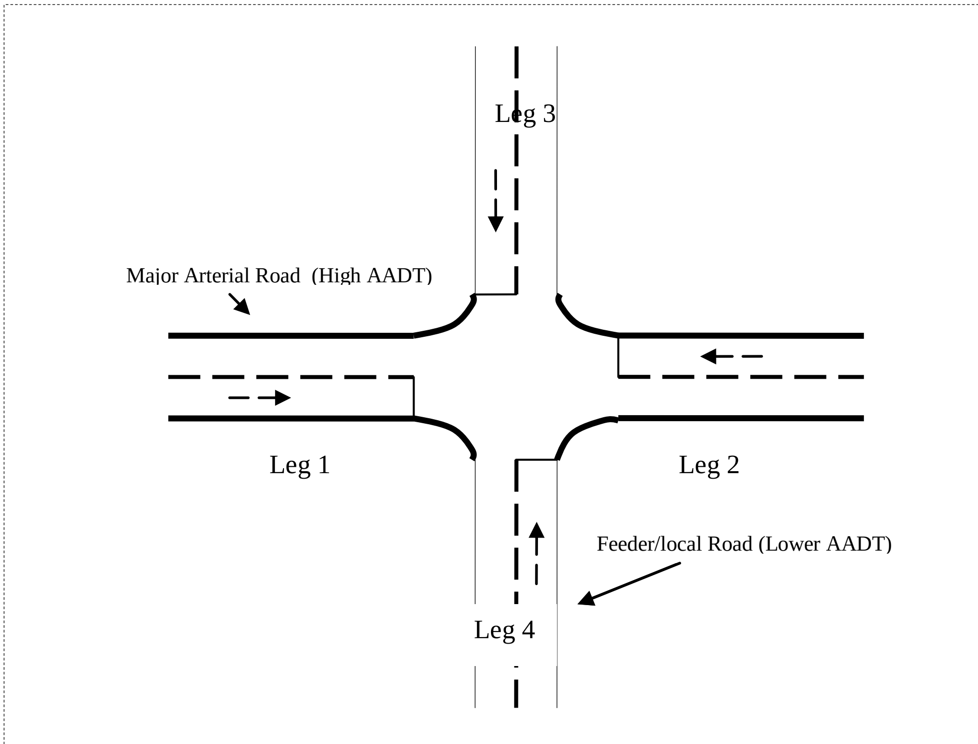
The traffic volume for the 40 intersections was computed using the Annual Average Daily Traffic (AADT) data published by the Commonwealth of Virginia. The total AADT for each intersection was calculated by adding traffic (AADT) coming into and leaving the intersection for both highways. The total AADT at an intersection is the sum of the average of AADT for the each highway as follows:

> Intersection total AADT = {[(Traffic Volume Approaching the Intersection from Direction 1 + Traffic Volume Leaving Intersection from Direction 1)/2]+[(Traffic Volume Approaching the Intersection from Direction 2 + Traffic Volume Leaving Intersection from Direction 2)/2]}

RESULT AND ANALYSIS

Although topographical data for each leg of the intersection was collected, accident data was not available for each leg. Therefore, a composite variable was created for the number of lanes, turn lanes etc. These composite variables were input into the regression model as the independent variables. A list of all independent variables is provided below in Table 1.

Figure 1. Schematic of an Intersection



The linear regression technique was used in this analysis in which total accident count was used as the dependent variable. Pearson correlation coefficients calculated are shown in Table 2.

Table 1. Independent Variables for the at Intersection Accident Model

Variable	Definition
AADT	Annual average daily traffic at the intersection
LANE	Total number of lanes at the intersection
TURN	Total number of turn lanes at the intersection
MEDN	Total number of physical median at the intersection (MEDN1+MEDN2+MEDN3+MEDN4)
PEDN	Total number of pedestrian crossing at the intersection (PEDN1+PEDN2+PEDN3+PEDN4)
DRWY	Total number of driveways at the intersection
HZRD	Number of legs with extra hazards at the intersection
EXSF	Number of legs with extra safety features at the intersection
RLFL	Number of legs with restricted left turn signal at the intersection.

The linear regression technique was used in this analysis in which total accident count was used as the dependent variable. Pearson correlation coefficients calculated are shown in Table 2.

Table 2. Correlation Coefficient			
Variable	**ACCT**	**p-Value**	**Sig at 10%**
LANE	.502	0.00475261	Yes
TURN	.559	0.00133763	Yes
DRW	-.006	0.97315647	No
MEAD	.330	0.07532854	Yes
PEDN	-.083	0.66431599	No
EXSF	.024	0.89843575	No
HZRD	.578	0.00081547	Yes
RLTL	-.030	0.87507706	No
AADT	.416*	0.02214094	Yes

Above table shows that five variables: number of lanes, number of turn lanes, presence of medians, presence of hazards and AADT, are significantly (at alpha of 10%) correlated to number of accidents. A linear regression model was developed using these variables. Coefficients of the model are presented below in the Table 3.

The linear model analysis showed that regression accounted for 60% of variability in the accident rate (R-square = .602). The analysis of variance of the regression model shows that the variability explained by the model was significant at less than 1% level.

Table 3. Linear Model Coefficient	
Constant	7.246
LANE	.438
TURN	3.225
MEAD	.596
HZRD	13.751
AADT	.001

The regression model can be written as:

ACCTOL = 7.246 + 0.438*LANE + 3.225* TURN + 0.596*MEAD + 0.001*AADT + 13.751*HZRD

Where ACCTOL—Total number of accidents at different intersections.

This result was significantly different than the pilot study result where R-square was 97%. The pilot study indicated that factors like number of driveways and pedestrian crossing were significant whereas presence of extra hazard (railway line, another traffic light with 250’, etc.) factors were not significant.

To validate results, the current model was used to predict the total number of accidents in a different set of ten intersections. It was found that the model was predicting higher than the actual number of accidents. This difference between actual and predicted values was on an average more than 33% higher. A t-test was conducted and difference between actual and predicted values was found to be significant with p-value of .003. Table 4 shows the results.

Table 4. Difference between Predicted and Actual Accidents			
Street	**Actual Accident**	**Predicted Accidents**	**Diff**
1	46	49.07	3.07
2	49	54.69	5.69
3	42	63.67	21.67
4	17	29.74	12.74
5	35	53.16	18.16
6	37	54.98	17.98
7	21	43.40	22.40
8	12	24.78	12.78
9	36	36.33	0.33
10	38	32.12	-5.88
Total	333	441.95	108.95

DISCUSSION

This study was an attempt to replicate the result of the earlier pilot study. However, the model developed with a larger sample size could not confirm the results of the pilot study. The R-square dropped from 97% in the pilot study to 60%. This was a significant change in the explained variation of the accident rates.

Furthermore, the variables which were found to be significant at the pilot study were not the same in the current model. It was encouraging to see that the presence of pedestrian crossings and number of driveways was significant in the pilot study model, indicating that certain policy decisions can be made based on the results of that model. However, those variables are no longer significant in the current study. It could be due to the fact that pilot study sample size was more homogeneous both in terms of the number of accidents as well as in terms of traffic volume.

Stepwise regression technique was also used to eliminate the affect of multi-collinearity. The model resulting from the stepwise regression included only two variables in the model; those factors were presence of hazards and number of turn lanes. As expected it gave lower R-square than model using simple regression. The R-square from the stepwise regression model was just 52%.

CONCLUSIONS AND LIMITATIONS

This study attempted to replicate the earlier pilot study to relate the traffic accidents and controllable factors such as road design, signal policies, and other data prevalent at signalized intersections. However, it was not able to fully replicate the pilot study. A larger sample of 65 intersections was chosen out of which only 40 data points were usable. A regression model was developed but it could only explain 60% of variability in the accident rate. Based on this research and literature, it is clear that there is some relationship between the topographical, design, and other controllable factors (60% of variation explained), however, more factors must be included to improve the results. Nevertheless, the analysis shows that traffic accidents and certain factors like presence of hazards should be further evaluated to see if these factors could be included to mitigate accident rates in future.

This study has the following limitations:

i. The accident data set is 5 years old compared to data collection on the roadways.
ii. The current model was unable to account for 40% of accident variations.
iii. The model was tested in a different set of intersections and showed that it was over estimating the number of accident by approximately 33%. Hence, its predictive capabilities were limited.
iv. Study of the impact of controllable factors could be improved if data was collected over time to reflect the changes made in the roadways.
v. Data on other relevant factors such as signal policy, road closure, etc. were not available. These factors could have an impact on the accident rates.
vi. This study excluded low accident rate intersections that could result in a biased model and may not be applicable to lower accident rate intersections.
vii. Records of the current data on the intersection geometry were not available; hence, onsite observations were conducted. These observations could have been affected by observers' skill level, fatigue, distraction, etc.

ACKNOWLEDGEMENT

The authors thank the City of Norfolk, Division of Transportation for providing data and inputs during the conduct of the study.

REFERENCES

Andrey, J. (2010). Long-Term Trends in Weather-Related Crash Risks. *Journal of Transport Geography*. 18 (2), 247-258.

Corben, B. F. and Foong Chee Wai (1990). Pro-Active Traffic Engineering Safety Study: Final Report Part 2 - Right-Turn-Against Crashes at Traffic Signals. *Monash University Accident Research Center Report*, Number 13, December, 1990, Victoria 3800, Australia.

Easa, S. M . and Mehmood, A. (2008). Optimizing Design of Highway Horizontal Alignments: New Substantive Safety Approach. *Computer-Aided Civil and Infrastructure Engineering*. 23 (7), 560-573.

Graves, S. J., Rochowiak, D. and Anderson, M. D. (2005). Mining and Analysis of Traffic Safety and Roadway Condition Data. University Transportation Center for Alabama , The University of Alabama, Tuscaloosa, AL. *UTCA Report* # 04310.

Hakkert, A.S. and Mahalel, D. (1978). Estimating the Number of Accidents at the Intersections Form a Knowledge of the Traffic Flows on the Approaches. *Accident Analysis and Prevention*. 10 (1), 69-79.

Harwood, D.W., Council, F.M., Hauer, E., Hughes, W.E., and Vogt A. (2000). Prediction of the Expected Safety Performance of Rural Two-Lane Highways Report. US Department of Transportation, Federal Highway Administration. *Report # FHWA-RD-99-207*.

Hot Topic and Issues Update: Auto Crashes (2006). Retrieved July 25, 2006 from http://www.iii.org/media/hottopics/insurance/test5./?table_sort_748377=6.

Lee, J., Chung, J. and Son, B. (2008). Analysis of Traffic Accident Size For Korean Highway Using Structural Equation Models. *Accident Analysis and Prevention*. 40(6), 1955-1963.

Lord, D., Guikema, S.D., Geedipally, S.R. (2008). Application of the Conway–Maxwell–Poisson generalized linear model for analyzing motor vehicle crashes. *Accident Analysis and Prevention.* 40 (3), 1123–1134.

Maheshwari, S.K. & D'Souza, K.A. (2008). Modeling Signalized Intersection Traffic Accidents: Empirical Study in City of Norfolk, VA. *International Academy of Business and Economics Journal of Business Research*, 8 (1) 219-238.

Malyshkina, N.V. and Mannering, F. L (2010). Empirical Assessment of the Impact Of Highway Design Exceptions On The Frequency And Severity of Vehicle Accidents. *Accident Analysis and Prevention*. 42 (1), 131–139.

Milton, J.C., Shankar, V.N. and Mannering, F.L. (2008). Highway Accident Severities and the Mixed Logit Model: An Exploratory Empirical Analysis. *Accident Analysis and Prevention*. 40 (1), 260-266.

Mountain. L., Fawaz, B. and Jarrett D. (1996). Accident Prediction Models For Roads With Minor Junctions. *Accident Analysis and Prevention*. 28 6) 695-707.

National Highway R&T Partnership. (2002). Highway Research and Technology: the Need for Greater Investment. U. S. Department of Transportation, Washington, D.C.

Nichols, K.M. (2007). Hampton Roads Regional Safety Study: General Crash Data and Trend 2007 Update. http://www.hrpdc.org/Documents/Transportation/Gen_Crash_ Data _ Trends_07_Final.pdf

Ogden, K. W., Newstead, S. V., Ryan, P. K., and Gantzer, S. (1994). Factors Affecting Crashes at Signalized Intersections. *Monash University Accident Research Center Report,* Number 62, October, 1994, ISBN 0732600618, Victoria 3800, Australia.

Ogden, K. W. and Newstead, S. V. (1994). Analysis of Crash Patterns at Victorian Signalized Intersections. *Monash University Accident Research Center Report,* Number 60, February, 1994, ISBN 0732600596, Victoria 3800, Australia.

Retting, R.A., Persaud, B. N, Garden, P.E. and Lord, D. (2001). Crash and Injury Reduction Following Installation of Roundabouts in the United States. *American Journal of Public Health*. V 9, N4, 628-631.

Shankar, V., Mannering, F., and Barfield,W. (1995). Effect of Roadway Geometrics and Environmental Factors on Rural Freeway Accident Frequencies. *Accident Analysis and Prevention.* 27 (3), 371–389.

Vogt, A. (1999). Crash Models For Rural Intersections: Four-Lane by Two-Lane Stop-Controlled and Two-Lane by Two-Lane Signalized. *US Department of Transportation, Federal Highway Administration Report* #FHWA-RD-99-128.

Wang, X., Abdel-Aty, M. (2008). Modeling left-turn crash occurrence at signalized intersections by conflicting patterns. *Accident Analysis and Prevention.* 40 (1) 76–88.

Wang, C., Quddus, M.A. and Ison, S.G. (2009). Impact of Traffic Congestion on Road Accidents: A Spatial Analysis of the M25 Motorway in England. *Accident Analysis and Prevention*. 41(4), Pages 798-808.

Washington, S., Persaud, B,, Lyon, C. and Oh, J. (2005). Validation of Accident Models for Intersections. US Department of Transportation, *Federal Highway Administration Report* # FHWA-RD-03-*037*.