Statistical models for right-turn related crashes at high crash locations in the Las Vegas Valley

Paul John Villaluz

University of Nevada, Las Vegas

Follow this and additional works at: https://digitalscholarship.unlv.edu/rtds

Repository Citation
https://digitalscholarship.unlv.edu/rtds/1995

This Thesis is brought to you for free and open access by Digital Scholarship@UNLV. It has been accepted for inclusion in UNLV Retrospective Theses & Dissertations by an authorized administrator of Digital Scholarship@UNLV. For more information, please contact digitalscholarship@unlv.edu.
STATISTICAL MODELS FOR RIGHT-TURN RELATED CRASHES AT HIGH CRASH LOCATIONS IN THE LAS VEGAS VALLEY

by

Paul John Villaluz, PE, PTOE

Bachelor of Science
University of Nevada, Las Vegas
1997

A thesis submitted in partial fulfillment of the requirements for the

Master of Science Degree in Engineering Department of Civil and Environmental Engineering Howard R. Hughes College of Engineering

Graduate College University of Nevada, Las Vegas May 2006
INFORMATION TO USERS

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleed-through, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.
The Thesis prepared by
PAUL JOHN VILLALUZ

Entitled
STATISTICAL MODELS FOR RIGHT-TURN RELATED CRASHES AT HIGH CRASH LOCATIONS IN THE LAS VEGAS VALLEY

is approved in partial fulfillment of the requirements for the degree of
MASTER OF SCIENCE IN ENGINEERING

Examination Committee Chair

Dean of the Graduate College

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
ABSTRACT

Statistical Models for Right-Turn Related Crashes
at High Crash Locations in the Las Vegas Valley

by

Paul John Villaluz, PE, PTOE

Dr. Mohamed Kaseko, Examination Committee Chair
Professor of Civil Engineering
University of Nevada, Las Vegas

The research objective was to develop a statistical model that related right-turn related crashes (RTRC) to volumetric factors at High Crash Locations (HCL) in the Las Vegas Valley. Information from selected HCL was analyzed with simple bivariate regression analysis and multiple regression analysis.

Response variables included the number of RTRC, the ratio of RTRC / Million Entering Vehicles (MEV), and the ratio of RTRC/ Total Intersection Crashes. Predictor variables included Right-Turn Volume, Right-Turn-On-Red Volume, Red Time / Cycle Time Percentage, Cross Product (per 1000 Vehicles) of Right-Turn Volumes and Opposing Through Volumes [Cross Product], Frequency of Gaps greater than 6.5 seconds [Gaps > 6.5 s], and the Frequency of Gaps less than 6.5 seconds [Gaps < 6.5 s].

Regression models of the relationships between these particular responses and predictor variables were found to explain up to 15% of the given data.
# TABLE OF CONTENTS

ABSTRACT........................................................................................................................iii

LIST OF FIGURES ..............................................................................................................vi

LIST OF ACRONYMS AND ABBREVIATIONS ..................................................................vii

ACKNOWLEDGMENTS ......................................................................................................viii

CHAPTER 1 INTRODUCTION ..............................................................................................1

CHAPTER 2 METHODOLOGY ..............................................................................................7
  Identification of Study Intersections ................................................................................8
  Classification of RTOR Crash Types .............................................................................9
  Data Collection ............................................................................................................10
  Model Construction ....................................................................................................11
    2.1. Response Variables ..........................................................................................11
      2.1.1. RTRC ........................................................................................................11
      2.1.2. RTRC / MEV ............................................................................................12
      2.1.3. RTRC / Total ............................................................................................13
    2.2. Independent Variables .....................................................................................14
      2.2.1. RT Volume and RTOR Volume ...............................................................15
      2.2.2. Red Cycle Time / Cycle Length ................................................................15
      2.2.3. Cross Product ............................................................................................15
      2.2.4. Gaps < 6.5 s and Gaps > 6.5 s ..................................................................16
      2.2.5. Transformations of Predictor Variables ....................................................17
    2.3. Bivariate Regression Analysis .........................................................................17
    2.4. Multiple Regression Analysis ..........................................................................18
      2.4.1. RTRC ........................................................................................................19
      2.4.2. RTRC / MEV ............................................................................................19
      2.4.3. RTRC / Total Intersection Crashes ..........................................................20

CHAPTER 3 RESULTS .........................................................................................................23
  Selection of Study Intersections ....................................................................................23
  Identification of RTOR Crash Cases ...........................................................................26
  Bivariate Regression Analysis .......................................................................................26
  Multiple Regression Analysis .......................................................................................27

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
LIST OF FIGURES

Figure 1. Time Gap (AASHTO, 2004) ........................................................................... 5
Figure 2. Eliminated RTRC Responses .........................................................................11
Figure 3. Similar RTRC Responses...............................................................................12
Figure 4. Sample Skewed RTRC / MEV Response Distribution ..................................13
Figure 5. Eliminated RTRC / Total Responses.............................................................14
Figure 6a. Model Construction Flowchart ......................................................................21
Figure 6b. Model Selection Flowchart ...........................................................................22
Figure 7. Map of Study Intersections............................................................................25
Figure 8. Plot of (In (RTRC)) vs. (In (RTOR volume))................................................28
Figure 9. Qualitative Analysis Plots for (In (RTRC)) vs. (In (RTOR volume)) ..........29
Figure 10. Plot of log RTRC / MEV vs. log RTOR volume ........................................31
Figure 11. Plot of ln Cross Product vs. ln RTRC / Total ...............................................33
Figure 12. Behavior of Cross Product Components .......................................................34
Figure 13. Plot of Crashes vs. Gaps > 6.5 seconds .........................................................36
Figure 14. Plot of Crashes vs. Gaps < 6.5 seconds .........................................................37
Figure 15. Gap Supply Graph .........................................................................................38
<table>
<thead>
<tr>
<th>AASHTO</th>
<th>American Association of State Highway and Transportation Officials</th>
</tr>
</thead>
<tbody>
<tr>
<td>ADT</td>
<td>Annual Daily Traffic</td>
</tr>
<tr>
<td>ANOVA</td>
<td>Analysis of Variables</td>
</tr>
<tr>
<td>C.I.</td>
<td>Confidence Interval</td>
</tr>
<tr>
<td>DF</td>
<td>Degrees of Freedom</td>
</tr>
<tr>
<td>FAST</td>
<td>Freeway and Arterial System of Transportation</td>
</tr>
<tr>
<td>HCL</td>
<td>High Crash Locations</td>
</tr>
<tr>
<td>HCM</td>
<td>Highway Capacity Manual</td>
</tr>
<tr>
<td>ITE</td>
<td>Institute of Transportation Engineers</td>
</tr>
<tr>
<td>MEV</td>
<td>Million Entering Vehicles</td>
</tr>
<tr>
<td>NDOT</td>
<td>Nevada Department of Transportation</td>
</tr>
<tr>
<td>ROW</td>
<td>Right of Way</td>
</tr>
<tr>
<td>RT</td>
<td>Right-Turn</td>
</tr>
<tr>
<td>RTOR</td>
<td>Right-Turn-on-Red</td>
</tr>
<tr>
<td>RTRC</td>
<td>Right-Turn Related Crashes</td>
</tr>
</tbody>
</table>
ACKNOWLEDGMENTS

I am extremely grateful to the following parties:

- John Tobin, PE, PTOE, Scott Plummer, PE, the G.C. Wallace Transportation Group, and the G.C. Wallace Executive Committee.
- Dr. Mohamed Kaseko. My fateful conversation with him during the 2003 Fall Transportation Conference got me back into school after a four-year hiatus.
- Dr. Sandra Catlin for serving on my committee and for her significant contribution to the statistical analysis.
- Dr. Shashi Nambisan and Dr. Hualiang (Harry) Teng for serving on my committee.
- Dr. Harriet Barlow for her patient guidance of the formatting.
- Dr. Samir Moujaes and Dr. Moses Karakouzian.
- Dr. David James for introducing me to research methods during CEG 700.
- Jay Van Sickle of the Nevada Department of Transportation (NDOT).
- Jim Neubert of Freeway and Arterial System of Transportation (FAST).
- Susan Thomason and Diana Lellieu of Silver State Traffic.
- UNLV Lied Library Inter-Library Loan Staff.
- Heidi Benson and Stephanie Benson.
- Rudy Villaluz, Carol Villaluz, Ruby Boyden, and Jason Boyden.
CHAPTER 1

INTRODUCTION

The purpose of this research is to identify a mathematical model that relates right-turn related crashes (RTRC) to volumetric factors at high crash locations (HCL). This mathematical model will help identify the causes of RTRC at HCL in the Las Vegas Valley.

According to data compiled by the Nevada Department of Transportation (NDOT) (NDOT, 2003), unincorporated Clark County and the cities of Henderson, North Las Vegas, and Las Vegas experienced 28,621 vehicular crashes between October 1999 and October 2002 and 15,834 injuries as a result of these crashes.

NDOT prepares a yearly crash mitigation review that identifies the intersections that are classified as HCL. An intersection must have thirty (30) crashes within three (3) years in order to be considered a HCL. The traffic engineers of these public entities meet yearly to discuss mitigation measures at each HCL.

It is suspected that Right-Turn-on-Red (RTOR) is a contributing factor to these types of crashes at HCL since drivers may not be able to perceive the safest gaps to enter opposing traffic.

RTOR is a policy that permits vehicles at signalized intersections to turn right against a solid red indication. Vehicles must come to a complete stop and yield to pedestrians and to vehicles with ROW before performing the turning maneuver.
Since the nationwide adoption of RTOR in 1975, transportation professionals have been studying the effects of this policy upon intersection safety. In 1976, preliminary data indicated that “accidents were occurring because of RTOR. However, compared to all intersection accidents, the frequencies are small” (McGee & Warren, 1976). It was also suggested that RTOR led to degradation of vehicular traffic safety not only for right turns but for all movements at the signalized intersection (Galin, 1981).

Further research concluded that “the traffic laws . . . permitting motorists to turn right on steady red at signalized intersections result in statistically significant and substantial increases in the numbers of right-turn crashes at these intersections” (Zador, Moshman, & Marcus, 1982).

Other findings included:

- A 20.7% increase in right-turn accidents at signalized intersections following introduction of RTOR.
- A 57% increase in pedestrian crashes in urban areas.
- RTOR had the greatest proportionate effect on crashes involving a single vehicle and pedestrian. This was most pronounced in urban areas and in areas with high elderly pedestrian traffic (Zador, Moshman, & Marcus, 1982).

Opposition to the conclusion reached by Zador, Moshman, and Marcus arose when Frith stated that their research “should not be seen as an indictment of the safety of RTOR” because the number of “incapacitating” accidents stayed constant— the only injury accident category reported upon in their paper. Frith also suggested that the question of whether permissive RTOR was “consistent with safety” would remain open until the data presented either by Zador or by other researchers were to include all
accidents in which injury is reported and publish the results (Frith, 1984). This additional
data should help evaluate the contribution of RTOR to overall accident potential.

In response to the criticism, Zador stated that “the available research evidence
indicates that allowing vehicles to turn right on red at signalized intersections increases
all right turning crashes by about 23%” (Zador, 1984).

The American Association of State Highway and Transportation Officials
(AASHTO) published a study on RTOR in 1979 based on data obtained for 732
signalized intersections in 14 large cities. The study included accident information both
before and after permitting RTOR. AASHTO concluded that the number of overall
accidents did not increase at signalized intersections following the adoption of RTOR
despite a 37 percent increase in accidents involving right-turning vehicles (Jaleel, 1984).
This implies that the overall proportion of the accident types has changed as a result of
the RTOR policy.

There are five basic types of RTOR accidents:

a. Sideswipe crashes between an opposing left-turn vehicle and a right
turning vehicle,

b. Right-angle crashes between a vehicle in the major road and a right-
turning vehicle,

c. Pedestrian crashes in either the approach or departure crosswalks,

d. Rear-end crashes in the approach between a car at the front of the
queue and a car that is second in the approach queue that either travels
too fast or anticipates a gap developing in the mainstream that the
driver at the front of the queue is not willing to accept, and
e. Rear-end crashes between a vehicle just entering the major road and a vehicle on the major road that is unable to slow down in time (McGee & Warren, 1976).

The vast majority (85%) of the pedestrian crashes are classified as RTOR-right. In this situation, the victim, who is coming from the driver's right, is not seen because the driver is looking to his left for a gap in traffic (Preusser, Leaf, DeBartolo, Blomberg, et al., 1982).

In order to perceive a safe gap to enter a traffic stream, a driver must be able to see oncoming vehicles clearly. Prior to 2000, AASHTO based its model on the kinematic behavior of the minor road vehicle turning onto the roadway and the deceleration performance of the following major road vehicle. The current AASHTO Green Book bases its sight distance model upon the optimum gap acceptance. The optimum gap is generally understood as a gap that is longer than the critical gap. The critical gap is defined as the minimum time interval in the major-traffic stream that allows intersection entry for one minor-street vehicle (Highway Capacity Manual, 2000).

A study prepared by Harwood, Mason, and Fitzpatrick lay the groundwork for the new AASHTO intersection sight distance model by observing the largest rejected and smallest accepted gaps at stop-controlled intersections nationwide (Stover & Koepke, 2002). These observed gaps were then implemented into the AASHTO model. Figure 1 illustrates the time gaps assumed for a vehicle turning right from a stopped condition on a minor road onto an intersecting major road.
Harwood noted that major-road ADT and cross-road ADT variables accounted for most of the variability in accident data that was explained by the models using negative binomial regression for three (3) of the five (5) specified intersection types (Harwood, 1996). As a result, most variables used in the model are related to volume and to stop control.

Harwood also noted that “the results of the statistical analyses...indicated that geometric design features explain relatively little of the variability in intersection accident data for at-grade intersections”. This was corroborated by an evaluation by three (3) independent reviewers of hard-copy police accident reports for a sample of eight (8) urban, four-leg, signalized intersections that found that only 5% to 14% of accidents had causes that appeared to be related to the geometric design features of the intersections (Harwood, 1996). This confirms the abandonment of the original theory that geometric designs affected variability of the accident data.

It is intuitive that the RTOR policy contributes to RTRC potential; but the extent that it contributes is unknown. Throughout the literature review, it has become apparent that other factors such as volume and gap supply may be contributing factors as well.

<table>
<thead>
<tr>
<th>Design Vehicle</th>
<th>Time gap (s) at design speed of major road ($t_g$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Passenger car</td>
<td>6.5</td>
</tr>
<tr>
<td>Single-unit truck</td>
<td>8.5</td>
</tr>
<tr>
<td>Combination truck</td>
<td>10.5</td>
</tr>
</tbody>
</table>
This study intends to identify the extent of the contributions of each of these factors to RTRC potential. The dependent variables (also known as response variables), the independent variables (also known as predictor variables), and the methodology used for the regression models will be developed in Chapter 2. Results of the models will be presented in Chapter 3 and interpreted in Chapter 4. The final conclusion of this study will be presented in Chapter 5.
CHAPTER 2

METHODOLOGY

A mathematical model was catered to the HCL that were most likely to have high numbers of RTRC. Once these types of HCL were identified, geometric and volumetric data at these locations were collected. Each intersection was divided into four (4) separate approaches because of the distinct differences in the variables at each approach.

Separation of each individual approach was supported in the Harwood paper. Another conclusion of his was that “the consideration of major-road ADT and crossroad ADT as separate independent variables provided better modeling results than consideration of a single variable representing either the sum or the product of the two variables” (Harwood, 1996).

The following data were used to formulate the predictor variables for the model:

- Right-Turn Volume (RT Volume), expressed in vehicles per hour (vph),
- Right-Turn-on-Red Volume (RTOR Volume), expressed in vph,
- Red Time / Cycle Time, expressed as a percentage of red time / cycle time (%),
- Cross Product (per 1000 vehicles) of Right-Turn Volumes and Opposing Through Volumes [Cross Product], expressed in conflicts per 1000 vehicles during the peak hour,
• Frequency of Gaps greater than \( t \) seconds (\( t = \) critical gap = 6.5 seconds)
  
  \[ \text{Gaps} > 6.5 \text{ s} \],

• Frequency of Gaps less than \( t \) seconds (\( t = \) critical gap = 6.5 seconds)
  
  \[ \text{Gaps} < 6.5 \text{ s} \].

Simple bivariate regression analysis was performed to identify the variables with strong relationships with the following response variables.

• RTRC,

• RTRC / Million Entering Vehicles (RTRC / MEV),

• RTRC / Total Intersection Crashes (RTRC / Total).

A regression analysis was performed with the predictor variables left intact and with the transformations of the predictor variables.

The best variables from the bivariate analyses were simply combined in various multiple linear regression models. Results from the bivariate and multiple linear regression analyses were compared to identify the best fitting model for each response. The models for each response were then compared to identify the best response for this particular phenomenon.

Identification of Study Intersections

In 2003, NDOT identified 135 High Crash Locations (HCL) in Clark County, 35 in the City of Henderson, 192 in the City of Las Vegas, and 40 in the City of North Las Vegas. These were based on collision reports collected between October 31, 1999 and October 30, 2002.
A student's t-distribution was used with a 95% confidence interval (C.I.) for the following data at the 402 study intersections:

- Ratio of RTRC to total crashes
- Total number of crashes (all types)
- Total number of RTRC
- Crashes per million entering vehicles (MEV)

Investigating the 402 study intersections with these parameters should provide a balanced representation of the most dangerous intersections with a high concentration of RT-related crashes. Intersections that performed within the top 2.5% of each criterion were identified as study intersections. The results of the statistical analyses are included in Section One of the Appendix.

Classification of RTOR Crash Types

From the selected intersections, detailed classifications of RTOR crashes were made based on McGee's identification of the types of RTOR crashes:

a. Case 1- Sideswipe crashes between an opposing left-turn vehicle and a right turning vehicle
b. Case 2- Right-angle crashes between a vehicle in the major road and a right-turning vehicle
c. Case 3a and 3b- Pedestrian crashes in either the approach or departure crosswalks
d. Case 4a- Rear-end crash in the approach between a car at the front of the queue and a car that is second in the approach queue that either
travels too fast or anticipates a gap developing in the mainstream that the driver at the front of the queue is not willing to accept.

e. Case 4b- Rear-end crash between a vehicle just entering the major road and a vehicle on the major road that is unable to slow down in time.

The crash diagrams used to identify each case are included in Section One of the Appendix.

Data Collection

Field visits at each of the selected study intersections were performed to identify potential factors that may contribute to these crashes, including:

- Presence of RTOR
- Red Cycle Time
- Opposing traffic stream speed limit
- Presence of vertical or horizontal curve or skew

The field visit reports are included in Section Three of the Appendix.

Traffic volumes were obtained from Silver State Traffic. All of these counts were performed from March 3, 2004 to February 24, 2005.

Signal timing information was obtained from the Freeway and Arterial System of Transportation (FAST). All of these timings were in force as of March 25, 2005.

The traffic volume and signal timing reports are included in Section Four of the Appendix.
Model Construction

The simple and multiple linear regression methods are selected in order to analyze the relationship of the predictor variables to the given responses. These methods are applied in order to find a predictive model that relates an increase or decrease in crashes to an increase or decrease in a given factor. Methodology flowcharts are included at the end of this section as Figures 6a and 6b.

2.1. Response Variables

There are three responses (dependent variables) identified for this study. Frequency histograms were prepared to see which version of each variable had the most normal distribution. These histograms are included in Section Five of the Appendix.

2.1.1. RTRC

No transformations of this response exhibit a normal distribution. Figure 2 illustrates the quadratic and the logarithmic transformations of the response that were eliminated from consideration because they were slightly skewed.

Figure 2. Eliminated RTRC Responses
Since the distributions of the untransformed variable and the natural logarithmic transformations are similar (as illustrated in Figure 3), it is assumed that the untransformed response variable and the natural logarithmic transformation of the response variable relate similarly to a given predictor variable.

![Figure 3. Similar RTRC Responses](image)

2.1.2. RTRC / MEV

The second response is the ratio of RT-related crashes / Million Entering Vehicles (RTRC / MEV) at an intersection. The number of crashes / MEV is a traditional safety measure at intersections. No transformations of this response exhibit a normal distribution. Figure 4 illustrates an example of the response removed from consideration because it exhibited a skew. This is true of the untransformed response and all of the transformations of the response.
2.1.3. RTRC / Total

The third response is the ratio of RT-related crashes / total crashes during the PM peak hour. This ratio represents the probability of RT-related accidents during a given peak hour. No transformations of this response exhibit a normal distribution.

The quadratic, natural logarithmic, and the logarithmic transformations of the response are eliminated from consideration because they were slight skewed. Figure 5 illustrates examples of these types of responses.
Since the distributions of the untransformed variable and the logarithmic and the natural logarithmic transformations are similar, it is assumed that the untransformed response variable, the logarithmic transformation of the response variable, and the natural logarithmic transformation of the response variable relate similarly to a given predictor variable.

2.2. Independent Variables

The independent (predictor) variables for the model represent the volumetric and geometric characteristics of each intersection that may contribute to the number of RTRC
at that intersection. The raw variables and the frequency histograms of each of these
variables are included in Section Five of the Appendix.

2.2.1. RT Volume and RTOR Volume

RT Volume (Right-Turn) Volume and RTOR (Right-Turn-on-Red) Volume are
modeled as continuous variables.

RT Volume includes the total amount of traffic performing the right-turning
movement during the PM peak hour (i.e., the traffic turning not only on red, but also on
green).

RTOR Volume is calculated as the right-turning peak hour volume during the PM
peak hour multiplied by the ratio of red cycle time to cycle length for the corresponding
approach.

2.2.2. Red Cycle Time / Cycle Length

The ratio of red cycle time to cycle length is modeled as a continuous variable. This
ratio is an expression of the percentage of red time per cycle length at a given approach.
Absolute red time is not used since there is no guarantee that the current red times are
similar to those in force when the crash data were collected. This variable implicitly
models the length of exposure of a right-turning vehicle to the main stream traffic during
the red cycle.

2.2.3. Cross Product

The Cross Product is modeled as a continuous variable. This variable is the product
of the total right-turning traffic from an approach during the PM peak hour and the
through traffic on the main stream during the same peak hour. Only the through
volumes in the departure lane nearest to the right-turning vehicle are used. These
volumes are calculated by dividing the total through traffic on the main stream approach by the total number of lanes on that approach. This variable implicitly models the conflict potential for a through / right-turn movement pair.

2.2.4. Gaps < 6.5 s and Gaps > 6.5 s

A vehicle that performs a RTOR maneuver must wait to merge into the traffic stream on the main stream. This is also true for vehicles that turn right from a minor street at an unsignalized intersection to the major street.

Drivers turning onto a major street at an unsignalized intersection must accept a gap equal to or greater than the critical gap. The critical gap is the minimum time headway (in seconds) between vehicles in the main stream that a driver can accept before comfortably executing a merging maneuver.

The expected number of gaps $h$ greater than or lesser than $t$ seconds can be calculated with the following equations (Garber & Hoel, 2001):

Freq. $(h \geq t) = (V - 1)e^{-\lambda t}$

Freq. $(h < t) = (V - 1)(1 - e^{-\lambda t})$

Where: $V =$ Volume (vph) on main stream flow,

$\lambda =$ arrival rate $= (V / T)$.

These equations assume Poisson distribution for the main stream flow. For calculation purposes, $T$ will be assumed to be 1 hr (3600 seconds) and the critical gap $t$ will be 6.5 seconds per AASHTO recommendations for Case B2.

The Poisson distribution is reasonable for light-to-medium traffic flows but may not be acceptable for conditions of heavy traffic. Under heavy traffic, the gaps are very small.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
and less random. Theoretically, gaps are more likely to be small in heavy traffic conditions.

2.2.5. Transformations of Predictor Variables

Transformations of the predictor variables were performed in order to discover the best distribution. These transformations are listed in Table 1.

Table 1. Transformations and Powers

<table>
<thead>
<tr>
<th>Transformation Type</th>
<th>Power</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reciprocal</td>
<td>-1</td>
</tr>
<tr>
<td>Cube Root</td>
<td>1/3</td>
</tr>
<tr>
<td>Square Root</td>
<td>1/2</td>
</tr>
<tr>
<td>Square</td>
<td>2</td>
</tr>
<tr>
<td>Cube</td>
<td>3</td>
</tr>
<tr>
<td>Ln</td>
<td>n/a</td>
</tr>
<tr>
<td>Log</td>
<td>n/a</td>
</tr>
</tbody>
</table>

It was discovered through qualitative analyses of each of the frequency histograms that the transformations shifted the original distribution along the x-axis. None of these transformations creates a more normal distribution for any of the predictor variables. Frequency histograms of the transformations each predictor variable are available in Section Five of the Appendix.

2.3. Bivariate Regression Analysis

Data for each intersection approach are distinct because of the different volumetric and conditions on each approach. Each approach represents one data point for the Bivariate Regression Analysis.
A regression analysis was performed for each response with respect to each of the six independent variables. Scatter plots and ANOVA tables were generated for simple linear regression of the untransformed variables. The quadratic, logarithmic, and natural logarithmic transformations of the variables were also used in order to provide a variety of models to evaluate. One (1) unit was added to each value of the crash data in the data sets in order to eliminate zero (0) values due to the infinite calculations of the logarithmic and natural logarithmic functions. Histograms, normal probability distribution plots, and plots of residuals to the fitted variable were created in MINITAB version 12.21. These histograms and plots are included in Section Five of the Appendix.

The fit of each predictor variable to each of the responses was evaluated quantitatively and qualitatively.

The quantitative analyses identify variables that related to a given response with a p-value less than 0.05 and a high F-value. High F-values identify the most significant models.

The qualitative analyses identify variables that have a histogram of residuals that resemble a bell curve, a normal probability distribution plot of the residuals that is as close to a straight line as possible, and a randomly scattered plot of residuals to the fitted variable. The variables that satisfied the criteria of both the quantitative and qualitative analyses were identified for further investigation.

2.4. Multiple Regression Analysis

The variables that satisfied the bivariate linear regression criteria (i.e. p-value less than 0.05 and high F-value) were identified for further investigation after the bivariate
analyses. They were placed into groups of multiple regression models for each response. The ultimate goal was to find the best fitting model for each response.

MINITAB analysis with the “Best Subsets Regression” command was used to analyze the models in each group. With this command, the two regression models with one predictor that have the highest $R^2$ are selected. This process is repeated for regression models with two predictors. The process ends when all predictors are used in the model.

This method selects the smallest subsets that had the highest adjusted $R^2$ and the lowest $C_p$ statistic. $R^2$, or the Coefficient of Multiple Determination, is the percent of the variance in the dependent variable that can be explained by all of the independent variables taken together. A $C_p$ statistic that is close to the number of $p$ parameters in the model infers that the subset model is not biased when compared to the overall model. Models that met these criteria were identified as the best multiple regression models for the response.

2.4.1. RTRC

The two predictor variables that satisfied the bivariate regression analysis criteria for this response were RTOR volume and RT volume. Since RTOR volume is merely the product of RT volume and the corresponding red cycle length, these variables are too correlative to be combined into a multiple regression model.

2.4.2. RTRC/MEV

The two predictor variables that satisfied the bivariate regression analyses criteria for this response were RTOR volume and RT volume. Since RTOR volume is merely the
product of RT volume and the corresponding red cycle length, these variables are too
correlative to be combined into a multiple regression model.

2.4.3. RTRC / Total Intersection Crashes

The four predictor variables that satisfied the bivariate regression analyses criteria
were Cross Product, RTOR volume, RT Volume, and Red Time / Cycle Time.

Combinations of the Cross Product with the RTOR volume, RT Volume, and Red
Time / Cycle Time were eliminated because of their correlation.

Combinations of Red Time / Cycle Time and RTOR Volume were eliminated for the
same reason.

The regression based on natural logarithmic transformation is used because it
provides better quantitative results than the regression based on the untransformed
variables.

The following correlation matrix identifies the variables that can be used together in a
prospective multiple regression model. Variables that have a correlation less than 0.05
are used in combinative models. The resulting multiple regression models for the
response use RT volume and Red Cycle Time.

Table 2. Model Correlation Matrix

<table>
<thead>
<tr>
<th>Correlation (p-value)</th>
<th>Cross Product</th>
<th>RTOR Volume</th>
<th>RT Volume</th>
<th>Red Cycle Time %</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cross Product</td>
<td>1.000</td>
<td>0.700 (0.000)</td>
<td>0.672 (0.000)</td>
<td>0.257 (0.048)</td>
</tr>
<tr>
<td>RTOR Volume</td>
<td>0.700 (0.000)</td>
<td>1.000</td>
<td>0.993 (0.000)</td>
<td>0.115 (0.383)</td>
</tr>
<tr>
<td>RT Volume</td>
<td>0.672 (0.000)</td>
<td>0.993 (0.000)</td>
<td>1.000</td>
<td>0.016 (0.905)</td>
</tr>
<tr>
<td>Red Cycle Time %</td>
<td>0.257 (0.048)</td>
<td>0.115 (0.383)</td>
<td>0.016 (0.905)</td>
<td>1.000</td>
</tr>
</tbody>
</table>

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
Figure 6a. Model Construction Flowchart

Model Construction

Enter Data
- Identify responses
- Identify predictor variables

Bivariate Regression Analysis
- Which predictor variables are significant predictors of dependent variables (responses)?
- Analyses used:
  - Simple Regression
  - Linear and Quadratic Relationships
  - Logarithmic + 1 Transformation
  - Linear and Quadratic Relationships
  - Natural Logarithmic + 1 Transformation
  - Linear and Quadratic Relationships

Discard Variable

NO

Does predictor variable have a p-value less than or equal to 0.05 with a relatively high $R^2$ and high F-statistic?

YES

Quantitative Analysis
- Scatter plots
  - Histograms of error terms of fitted model = bell curve
  - Normal plot of residuals = straight line
  - Residuals versus fits = no pattern

Discard Variable

NO

Does predictor variable meet standards of Quantitative Analysis?

YES

Multiple Regression Analysis
- Develop two and three variable models.
- Check collinearity of predictor variables.
Model Selection

Goal:

What is best predictor variable?
What is best response?

Response: RTRC
- Compare following models:
  - Best bivariate models

Response: RTRC / MEV
- Compare following models:
  - Best bivariate models

Response: RTRC / Total Crashes
- Compare following models:
  - Best bivariate models
  - Best multiple regression models
  - Groups I, and II

Evaluate best models
- Compare following models:
  - Best RTRC
  - Best RTRC / MEV
  - Best RTRC / Total Crashes
  - Identify trends

Compare best models for each response
- Identify best models per $p$, $R^2$, and $F$
- Perform qualitative analysis
CHAPTER 3

RESULTS

Selection of Study Intersections

A student’s t-distribution was used with a 95% confidence interval (C.I.) for the absolute number of crashes at the four hundred and two (402) study intersections. The calculation sheets are located in Section Two of the Appendix.

Table 3. Intersection Selection Criteria

<table>
<thead>
<tr>
<th>Criterion</th>
<th>Mean</th>
<th>Range within 95% C.I.</th>
<th>Number of intersections in top 2.5% of criterion</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Lower Bound</td>
<td>Upper Bound</td>
</tr>
<tr>
<td>RTRC /Total Crashes</td>
<td>0.0502</td>
<td>0.0468</td>
<td>0.0536</td>
</tr>
<tr>
<td>Total Crashes (all types)</td>
<td>71</td>
<td>66</td>
<td>76</td>
</tr>
<tr>
<td>RTRC</td>
<td>4</td>
<td>3.674</td>
<td>4.326</td>
</tr>
<tr>
<td>Crashes per MEV</td>
<td>1.44</td>
<td>1.19</td>
<td>1.69</td>
</tr>
</tbody>
</table>

Fifteen (15) intersections fit the criteria of being in the top 2.5% for each of these factors. These locations consistently intersect at 90-degrees and have similar signal control.

1. Alta Drive @ Rampart Boulevard
2. Bonanza Road @ Eastern Avenue
3. Bonanza Road @ Martin Luther King Boulevard
4. Charleston Boulevard @ Buffalo Drive
5. Charleston Boulevard @ Martin Luther King Boulevard
6. Craig Road @ Martin Luther King Boulevard
7. Flamingo Road @ Arville Street
8. Las Vegas Boulevard @ Charleston Boulevard
9. Las Vegas Boulevard @ Sahara Avenue
10. Paradise Road @ Flamingo Road
11. Sahara Avenue @ Buffalo Drive
12. Sahara Avenue @ Valley View Boulevard
13. Spring Mountain Road @ Valley View Boulevard
14. Sunset Road @ Stephanie Street
15. Tropicana Avenue @ Eastern Avenue

Figure 7 illustrates the locations of these intersections in the Las Vegas Valley.
Figure 7. Map of Study Intersections
Identification of RTOR Crash Cases

There were one hundred and twenty-nine (129) total RTRC at the fifteen (15) study intersections. Table 4 illustrates the cases, types and frequencies of RTRC.

Table 4. Crash Cases, Types, and Frequencies

<table>
<thead>
<tr>
<th>Crash Case</th>
<th>Crash Type</th>
<th>Number</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Sideswipe with opposing left-turn vehicle</td>
<td>27</td>
</tr>
<tr>
<td>2</td>
<td>Right-Angle with vehicle in major road</td>
<td>49</td>
</tr>
<tr>
<td>3a</td>
<td>Pedestrian at Approach</td>
<td>24</td>
</tr>
<tr>
<td>3b</td>
<td>Pedestrian at Departure</td>
<td>22</td>
</tr>
<tr>
<td>4a</td>
<td>Rear-end at Approach</td>
<td>7</td>
</tr>
<tr>
<td>4b</td>
<td>Rear-end at Departure</td>
<td>0</td>
</tr>
</tbody>
</table>

The response variables for the regression models were built with the sums of the crash types. Specific regression models were not constructed for the pedestrian crash cases because pedestrian volumes were not available.

Bivariate Regression Analysis

The equations in Table 5 relate to the given response with a $p < 0.05$. High F-statistics were also considered. More information is included in Section Six of the Appendix.
Table 5. Eligible Bivariate Linear Regression Models

<table>
<thead>
<tr>
<th>Model</th>
<th>Equation</th>
</tr>
</thead>
<tbody>
<tr>
<td>RTRC</td>
<td>$(\ln (\text{RTRC})) = -0.7693 + 0.3401 (\ln (\text{RT Volume}))$</td>
</tr>
<tr>
<td></td>
<td>$(\ln (\text{RTRC})) = -0.7009 + 0.3451 (\ln (\text{RTOR Volume}))$</td>
</tr>
<tr>
<td>RTRC / MEV</td>
<td>$(\log (\text{RTRC} / \text{MEV})) = -0.0099 + 0.0099350 (\log (\text{RTOR Volume}))$</td>
</tr>
<tr>
<td></td>
<td>$(\log (\text{RTRC} / \text{MEV})) = -0.0104 + 0.0096260 (\log (\text{RT Volume}))$</td>
</tr>
<tr>
<td>RTRC / Total</td>
<td>$(\ln (\text{RTRC} / \text{Total})) = -0.0066 + 0.017164 (\ln (\text{X prod per 1000}))$</td>
</tr>
<tr>
<td></td>
<td>$(\ln (\text{RTRC} / \text{Total})) = -0.0355 + 0.019527 (\ln (\text{Right Turn Traffic}))$</td>
</tr>
<tr>
<td></td>
<td>$(\ln (\text{RTRC} / \text{Total})) = -0.0377 + 0.021053 (\ln (\text{RTOR Volume}))$</td>
</tr>
<tr>
<td></td>
<td>$(\log (\text{RTRC} / \text{Total})) = 0.0508 + 0.18385 (\log (\text{Red Cycle Time}))$</td>
</tr>
</tbody>
</table>

Multiple Regression Analysis

The model below satisfied R-adj and c-p criteria per the MINITAB “Best Subsets regression” analysis. More information is included in Section Six of the Appendix.

Table 6. Eligible Multiple Linear Regression Model

<table>
<thead>
<tr>
<th>Model</th>
<th>Equation</th>
</tr>
</thead>
<tbody>
<tr>
<td>RTRC / Total</td>
<td>$\ln (\text{RTRC} / \text{Total}) = 0.01333 + 0.018909 \ln \text{RT volume} + 0.16465 \ln \text{Red Cycle Time}$</td>
</tr>
</tbody>
</table>
CHAPTER 4

DISCUSSION

RTRC

After the conclusion of the Bivariate Linear Regression Analysis, two predictors were identified to have a relatively strong correlation with this particular response: RT Volume and RTOR Volume.

The linear relationship with the natural logarithmic transformations for the RTOR variable had the best fit. Figure 8 is the scatter plot of the data.

$$(\ln \text{(RTRC)}) = -0.7009 + 0.3451 \ln \text{(RTOR volume)}$$

Figure 8. Plot of $(\ln \text{(RTRC)})$ vs. $(\ln \text{(RTOR volume)})$
It was originally assumed that the number of RTRC would be directly related with both RT and RTOR volumes. It was also assumed that RTOR volumes would provide a slightly higher number of crashes. The general ascent of the fitted line infers this to be true as far as the volumes are concerned (i.e. greater volumes = greater number of crashes). Figure 9 shows the qualitative plots used to analyze this response / variable pair.

Figure 9. Qualitative Analysis Plots for (\(\ln(\text{RTRC})\)) vs. (\(\ln(\text{RTOR volume})\))

The histogram of the residuals is not quite a bell curve. The normal probability plot does not show a straight line relationship and the residuals vs. fits plot is not totally random. Therefore, one may infer through inspection of the respective scatter plots and
residual plots that there is not a very strong correlation between the ratio of RTRC and RTOR volume at a particular intersection.

Multiple regression analyses using RT and RTOR volumes were not used since they are closely correlated within this given response. The best subsets run justifies this since the best equation involved solely one variable (RTOR Volume).

The best regression models are illustrated in Table 7. The Bivariate Linear Regression involving the natural logarithmic transformation seems to be the better model.

Table 7. Best Regression Models for RTRC

<table>
<thead>
<tr>
<th>RTRC</th>
<th>Bivariate Linear Regression</th>
</tr>
</thead>
<tbody>
<tr>
<td>(ln (RTRC)) = -0.7009 + 0.3451 (ln (RTOR volume))</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Constant</th>
<th>Variable</th>
<th>Coeff</th>
<th>t</th>
<th>p</th>
<th>Coeff</th>
<th>t</th>
<th>p</th>
<th>R2</th>
<th>R2 (adj.)</th>
<th>Std. Error</th>
<th>F</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>-0.7009</td>
<td>-1.29</td>
<td>0.203</td>
<td>0.3451</td>
<td>3.18</td>
<td>0.002</td>
<td>14.8%</td>
<td>13.4%</td>
<td>0.5102</td>
<td>10.11</td>
<td>0.002</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Best Subsets Regression

<table>
<thead>
<tr>
<th>Constant</th>
<th>Variable</th>
<th>Coeff</th>
<th>t</th>
<th>p</th>
<th>Coeff</th>
<th>t</th>
<th>p</th>
<th>R2</th>
<th>R2 (adj.)</th>
<th>Std. Error</th>
<th>F</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.2533</td>
<td>3.57</td>
<td>0.001</td>
<td>0.0052</td>
<td>2.98</td>
<td>0.004</td>
<td>13.2%</td>
<td>11.8%</td>
<td>1.4030</td>
<td>8.86</td>
<td>0.004</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

RTRC / MEV

After the conclusion of the Bivariate Linear Regression Analysis, the only two predictors that were identified to have a relatively strong correlation with this particular response were RT Volume and RTOR Volume. The linear variation of the logarithmic transformations for RTOR volume had the best fit. Figure 10 is the scatter plot of the data.

(log (RTRC / MEV)) = -0.0099 + 0.0099350 (log (RTOR volume))
The same assumptions made for the RTRC response were also made for the RTRC / MEV response. Inspections of the fitted line plot, histograms of the residuals, normal probability plot and the residuals vs. fits plot lead to a conclusion similar to the one reached for the RTRC response. Therefore, one can assume that there is no strong correlation between RTRC / MEV and RTOR.

Multiple regression analyses using RT and RTOR volumes were not used since they are closely correlated within this given response. The best subsets run justifies this since the best equation involved solely one variable (RTOR Volume).

The best regression models are illustrated in Table 8. The Bivariate Linear Regression involving the logarithmic transformation seems to be the better model.
Table 8. Best Regression Models for RTRC / MEV

RTRC / MEV

Bivariate Linear Regression

\[
\log (\text{RTRC} / \text{MEV}) = -0.0099 + 0.009935 \log (\text{RTOR volume})
\]

<table>
<thead>
<tr>
<th>Coeff</th>
<th>t</th>
<th>p</th>
<th>Coeff</th>
<th>t</th>
<th>p</th>
<th>(R^2)</th>
<th>(R^2) (adj.)</th>
<th>Std. Err.</th>
<th>F</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.0099</td>
<td>1.29</td>
<td>0.202</td>
<td>0.0099</td>
<td>2.81</td>
<td>0.007</td>
<td>12.0%</td>
<td>10.4%</td>
<td>0.0072</td>
<td>7.88</td>
<td>0.007</td>
</tr>
</tbody>
</table>

Best Subsets Regression

RTRC / MEV = 0.0174 + 0.000057 RTOR volume (C-p = 1.4)

<table>
<thead>
<tr>
<th>Coeff</th>
<th>t</th>
<th>p</th>
<th>Coeff</th>
<th>t</th>
<th>p</th>
<th>(R^2)</th>
<th>(R^2) (adj.)</th>
<th>Std. Err.</th>
<th>F</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.0174</td>
<td>4.01</td>
<td>0.000</td>
<td>0.0001</td>
<td>2.64</td>
<td>0.011</td>
<td>10.7%</td>
<td>9.2%</td>
<td>0.0173</td>
<td>6.97</td>
<td>0.011</td>
</tr>
</tbody>
</table>

RTRC / Total Amount of Crashes

After the conclusion of the Bivariate Linear Regression Analysis, four predictors were identified to have a strong correlation with this particular response: Cross Product, RTOR volume, RT volume, and Red Time / Cycle Time. The linear variation of the natural logarithmic transformations for the Cross Product variable had the best fit. Figure 11 is the scatter plot of the data.

\[
\log (\text{RTRC/Total}) = -0.0066 + 0.0172 \log (\text{X Product per 1000 veh})
\]
The response of "ln RTRC / Total" is directly related to the "ln of Cross Product".

Figure 12 illustrates the behavior of the individual components of the Cross Product. As Cross Product increases, right-turn and through volume both increase for a given Cross Product. The sharp increase in the slope of the through volume implies that through volume has more influence on Cross Product than right-turn volume.
Cross Product also affects gap supply and the opposing through volume. Through volume can affect gap supply because right-turn volume and opposing through volume are inversely related within the same given Cross Product. Examples of this behavior are illustrated on Table 9. Total gaps are directly related to through volume for the same given Cross Product. Increasing Cross Product exacerbates relationships.
Table 9. Gap and Cross Product Relationships

<table>
<thead>
<tr>
<th>GAPS</th>
<th>RT Volume</th>
<th>Opposing Through Volume</th>
<th>Cross Product</th>
</tr>
</thead>
<tbody>
<tr>
<td>99</td>
<td>10</td>
<td>100</td>
<td>1000</td>
</tr>
<tr>
<td>9</td>
<td>100</td>
<td>10</td>
<td>1000</td>
</tr>
<tr>
<td>RT Volume vs Opposing Through Volume</td>
<td>INVERSE</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GAPS vs Opposing Through Volume</td>
<td>DIRECT</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GAPS vs RT Volume</td>
<td>INVERSE</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

If Cross Product increases and Opposing Through Volume has more influence on the increase of the Cross Product, then the number of gaps has a similar influence as the opposing through volume.

There may not be a meaningful mathematical relationship between gap supply and the absolute number of crashes because of the poor fit of both the linear and polynomial functions. However, an interesting trend can be described by the polynomial function if it is assumed to be meaningful.
Figure 13 illustrates a loosely fitting polynomial function. At the leftmost end of the curve, there are not a lot of Gaps > 6.5 s available. This implies jam density and a high supply of Gaps < 6.5 s available. As a result, the propensity for a crash between a right-turning vehicle and an opposing through vehicle is very low because the right-turn vehicle would have to wait for signal control before advancing. At the rightmost end of the curve, there are a lot of Gaps > 6.5 s available. This implies free flow. As a result, the propensity for a crash between a right-turning vehicle and an opposing through vehicle is very low because of the high availability of safe gaps. This implies that the possibility for driver error is described by the Gaps > 6.5 s vs. RTRC curve. Driver error is maximized in the middle of the curve, where a driver can misjudge a gap.
Figure 14 illustrates the relationship between RTRC and Gaps < 6.5 s. When RTRC is compared to Gaps < 6.5 s, the results are not as conclusive. The leftmost end of the resulting curve implies safer conditions under free flow because of a low Gaps < 6.5 s and a corresponding high Gaps > 6.5 s. The high Gaps < 6.5 s on the rightmost end of the curve implies unsafe conditions under jam density.

However, the curve illustrates crash numbers that are the reverse of what is assumed. There are high amounts of crashes on the leftmost end of curve where free flow is assumed. This contradicts the conclusions drawn from the previous curve. The open-ended nature of the gap supply could be responsible for the discrepancy.
Figure 15 relates the number of Gaps > 6.5 s to Gaps < 6.5 s. The number of Gaps > 6.5 s approaches an asymptote of 200. Since it approaches an asymptote, Gaps > 6.5 s is a better predictor variable. The maximum number of gaps can be derived from a partial derivative of the gap equation. This is included in Section Six of the Appendix.

Inspections of the fitted line plot, histograms of the residuals, normal probability plot and the residuals vs. fits plot lead to a conclusion similar to the one reached for the other responses. This is true for the plots from both the bivariate and multiple regression models. Therefore, one can assume that there is no strong correlation between RTRC / Total and Cross Product.

The best regression models are illustrated in Table 10. The Bivariate Linear Regression involving the natural logarithmic transformation seems to be the better model.
Table 10. Best Regression Models for RTRC / Total

<table>
<thead>
<tr>
<th>RTRC / Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bivariate Linear Regression</td>
</tr>
<tr>
<td>((\ln (\text{RTRC/Total})) = -0.0066 + 0.0172 (\ln (X \text{ Product per 1000 veh})))</td>
</tr>
<tr>
<td>Variable</td>
</tr>
<tr>
<td>Constant</td>
</tr>
</tbody>
</table>

| Best Subsets Regression |
| \(\ln \text{RTRC/Total} = 0.0133 + 0.0189 \ln \text{RT volume} + 0.165 \ln \text{Red Cycle Time % (C-p = 3.0)}\) |
| Variable | Coeff | \(t\) | \(p\) | Coeff | \(t\) | \(p\) | Coeff | \(t\) | \(p\) | \(R^2\) | \(R^2 (\text{adj.})\) | Std. Err. | \(F\) | \(p\) |
| 0.0133 | 0.26 | 0.798 | 0.0189 | 2.17 | 0.034 | 0.1647 | 2.03 | 0.047 | 13.8% | 10.8% | 0.0406 | 4.57 | 0.014 |

Model Selection

The model with the best values of \(p\), \(R^2\), and \(F\) is the model for RTRC as a function of RTOR volume. This model explains 15% of the data. The qualitative plots for all the responses are too similar to use for model selection.

Success of Statistical Modeling

ITE defines intersections with more than 1.2 crashes per MEV as HCL. The only study intersections that do not satisfy this criterion are Alta Drive and Rampart Boulevard, Sahara Avenue and Buffalo Drive, and Sahara Avenue and Valley View Boulevard. These intersections do not affect any analyses by providing any outlying data that affect the predictive ability of any of the models.

Harwood provided general guidelines for the methodology of this project and concluded that “traditional multiple linear regression is generally not an appropriate statistical approach to modeling of accident relationships because accidents are discrete, non-negative events that often do not follow a normal distribution” (Harwood, 1996).
This is exemplified by the results of multiple linear regression and simple linear regression results.

Harwood also states that “Poisson, negative binomial, lognormal, and logistic distributions appear to be better suited to the modeling of accident relationships than the normal distribution” (Harwood, 1996). However, these were not attempted because of the low correlation that Harwood eventually discovered. He found that “regression models to determine relationships between accidents and intersection geometric design, traffic control, and traffic volume variables based on the negative binomial distribution explained between 16% and 38% of the variability in accident data.” (Harwood, 1996). As a result, models with different distributions were not attempted.

Harwood also cast doubt upon the practicality of the models. He writes “while the models presented in this report are the best that can be developed from available data, they do not appear to be of direct use to practitioners.” He added that “furthermore, goodness of fit of models is not as high as would be desired. Therefore, models presented here are appropriate as a guide to future research; but do not appear to be appropriate for direct application by practitioners” (Harwood, 1996).

An addendum to this paper used single collision data (the previous project used multi-vehicle collisions only) and reached the same conclusions (Harwood, 1999).
CHAPTER 5

CONCLUSIONS

The purpose of this research is to identify a mathematical model that relates right-turn related crashes (RTRC) to volumetric factors at high crash locations (HCL). Three dependent variables (also known as response variables) and six independent variables (also known as predictor variables) at fifteen (15) study intersections were developed for use in bivariate and multiple linear regression models. The best model calculates RTRC as a function of RTOR volume and explains 15% of the data.

It is intuitive that as RTOR increases, crash potential increases; but the Gaps > 6.5 s seems to be a good predictor variable as well. A loose polynomial relationship between Gaps > 6.5 s and RTRC suggests the likelihood of a driver to misjudge a gap size.

No conclusive relationship could be found between any predictor variable and any of the given responses. Regression models of the relationships between these particular responses and predictor variables were found to explain only up to 15% of the given data. Multiple linear regression does not provide a modeling advantage. Although simple linear regression assuming normal distribution was used in lieu of regressions with more sophisticated distributions (such as Poisson, lognormal, negative binomial, and logistic), the poor fit confirms Harwood’s original conclusion that the linear regression models are not appropriate for direct application.
The poor fit in the Harwood models also suggests that the temporal difference between the collection of the traffic counts and the crash data for this particular investigation was insignificant. Although a future analysis could be conducted with such data being collected simultaneously, the results could be expected to be just as poor. Therefore, it is impractical to devise a predictive model for crashes based on volumetric factors.

This study is focused on the causal factors of RTRC at HCL. In order to remove this bias and provide more variability of data, future efforts can integrate more randomly selected intersections. A simultaneous regression approach for all the RTRC cases can also be performed. Data for pedestrian volumes will be needed for the pedestrian crash cases.

Data can also be collected for the number of conflicting vehicles per lane and the number of conflicting lanes on the roadway with the opposing through volumes. These data will identify the “escape potential” on the roadway and the resulting reduction of RTRC.
BIBLIOGRAPHY


*ITE Journal*, 51.1, 24-29.


New York: Thomson Learning.


VITA

Graduate College
University of Nevada, Las Vegas

Paul John Villaluz, PE, PTOE

Home Address:
9307 Avon Park Avenue
Las Vegas, NV 89149

Degrees:
Bachelor of Science. Civil Engineering. 1997
University of Nevada, Las Vegas

Thesis Title: Statistical Models for Right-Turn Related Crashes at High Crash Locations in the Las Vegas Valley

Thesis Examination Committee:
Chairperson, Dr. Mohamed Kaseko, Ph. D.
Committee Member, Dr. Shashi Nambisan, Ph. D.
Committee Member, Dr. Hualiang (Harry) Teng, Ph. D.
Graduate Faculty Representative, Dr. Sandra Catlin, Ph. D.