RED LIGHT RUNNING PREDICTION AND ANALYSIS

By

Mr. Stephen E. Hill and Dr. Jay K Lindly
Department of Civil and Environmental Engineering
The University of Alabama
Tuscaloosa, Alabama

Prepared by

UTCA

University Transportation Center for Alabama
The University of Alabama, The University of Alabama at Birmingham, and
The University of Alabama in Huntsville

UTCA Report Number 02112
July 7, 2003

Technical Report Documentation Page

1. Report No FHWA/CA/OR-	2. Government Accession No.	3. Recipient Catalog No.
4. Title and Subtitle Red Light Running Prediction and Analysis	**5. Report Date** July 7, 2003 **6. Performing Organization Code**	
7. Authors Mr. Stephen E. Hill and Dr. Jay K. Lindly	**8. Performing Organization Report No.** UTCA Final Report 02112	
9. Performing Organization Name and Address University Transportation Center for Alabama P O Box 870205 University of Alabama Tuscaloosa, AL 35487-0205	**10. Work Unit No.** **11. Contract or Grant No.** HPP-1602(524)	
12. Sponsoring Agency Name and Address University Transportation Center for Alabama P O Box 870205 University of Alabama Tuscaloosa, AL 35487-0205	**13. Type of Report and Period Covered** Final Report; January 1, 2002 – December 31, 2002 **14. Sponsoring Agency Code**	
15. Supplementary Notes		

16. Abstract

Transportation professionals and engineers have identified red light running as a major traffic safety hazard. However, there are currently only limited tools available to assist transportation professionals in selecting intersections for remediation or law enforcement based upon red light running violation risk.

The goal of this research was to develop statistical models to predict red light running violation frequency (violations/hour) based upon traffic operational and intersection geometric characteristics for four-approach intersections. Red light running violation data over the time period of 2-8 PM was gathered from 19 intersection approaches in four states (Alabama, Texas, Iowa, and California) and was compiled and analyzed. A total of 1,775 violations were observed over 554 hours (for an observed rate of 3.2 violations/hour). Additionally, fourteen geometric and traffic operational characteristics were recorded for each intersection.

The results of this research present several regression equations (linear, curvilinear, and multiple linear) which can be used to predict the red light running frequency that can be expected at an intersection approach based upon several geometric and traffic operational characteristics. Variables showing predictive power included average daily traffic (ADT), number of approach lanes, speed limit, number of lanes crossed by the approach, and distance to preceding and following intersections.

17. Key Words		18. Distribution Statement	
Red Light Running, Safety			
19. Security Class (of this report) Unclassified	**20. Security Class. (Of this page)** Unclassified	**21. No of Pages**	**22. Price**

Contents

List of Tables

List of Figures

Executive Summary

Transportation professionals and engineers have identified red light running as a major traffic safety hazard. However, only limited tools are available to assist transportation professionals in selecting intersections for remediation or law enforcement based upon red light running violation risk.

The goal of this research was to develop statistical models to predict red light running violation frequency (violations/hour) based upon traffic operational and intersection geometric characteristics for four-approach intersections. Red light running violation data over the time period of 2-8 PM was gathered from 19 intersection approaches in four states (Alabama, Texas, Iowa, and California) and was compiled and analyzed. A total of 1,775 violations were observed over 554 hours, for an observed rate of 3.2 violations/hour. Additionally, fourteen geometric and traffic operational characteristics were recorded for each intersection.

The results of this research present several regression equations (linear, curvilinear, and multiple linear) which can be used to predict the red light running frequency that can be expected at an intersection approach based upon several geometric and traffic operational characteristics. From simple linear regression analysis, the number of approach lanes and the average daily traffic volume (ADT) were shown to be predictors of red light running frequency with moderate degrees of predictive power.

Curvilinear regression analysis suggested that the number of approach lanes, approach speed limit, and the ADT of the observed approach in addition to the distance to the preceding and following intersections possessed predictive power with regard to red light running frequency. Analysis by best subsets multiple regression revealed that a combination of the number of approach lanes, the number of lanes crossed by the approach, approach ADT, and the compass direction of the approach produced a regression equation with excellent predictive power.

Section 1
Introduction

Problem Identification

Transportation professionals and engineers have identified red light running as a major traffic safety hazard. However, there are currently only limited tools available to assist transportation professionals in selecting intersections for remediation or law enforcement based upon red light running violation risk.

Objective

The goal of the current research work was to develop statistical models to predict red light running violations based upon traffic operational and intersection geometric characteristics for four-approach intersections. These models were envisioned as practical tools to allow transportation professionals to determine the likelihood that particular intersection configurations and traffic flow characteristics may lead to red light running violations. Providing remediation or law enforcement to those sites should reduce the likelihood of red light running violations and the likelihood of red light running crashes.

Report Organization

This report is organized as follows:

- Section Two – Literature Review – This chapter discusses the red light running problem and previous efforts to predict red light running incidents.
- Section Three – Methodology – This chapter gives a detailed account of the steps involved in the collection and reduction of relevant data.
- Section Four – Model Construction and Analysis – This chapter details the construction of the red light running violation prediction models and provides an analysis of the model results.
- Section Five - Conclusions and Opportunities for Future Work – This chapter presents the conclusions derived from this research and gives suggestions for enhancements to this research.

Section 2
Literature Review

This chapter provides a review of relevant red light running research that has been previously conducted. Red light running is defined, the red light running problem is discussed, and previous attempts to predict red light running incidents are examined.

Red Light Running Overview

Red light running may be defined as follows (Passetti, 1997):

Deliberately entering an intersection after the signal light has turned red.

Although any vehicle that enters an intersection after the signal has turned red has "run the red light," jurisdictions vary in their enforcement of red light violations. In all jurisdictions in which data was collected for this project (Alabama, California, Iowa, and Texas), a vehicle is considered to have "run the red light" if it enters the intersection after the traffic signal has turned red. A vehicle is considered to have "entered the intersection" when a portion of the vehicle crosses the stopbar.

Specific red light running laws are codified for each state in the follow manner:

- Alabama – Code of Alabama, Section 32-5A-32 (3)
- California – California Vehicle Code, Section 21543
- Texas – Texas Statutes, Transportation Code, Chapter 544.007 (d)
- Iowa – Iowa Code, Title 8, Section 321.257.2a

Estimates indicate that more than 3.8 million crashes occur at intersections each year, accounting for nearly 60 percent of all reported crashes. Of these intersection crashes, 8,500 are fatal (one or more persons killed), and over 1 million result in injuries. Red light running accounted for 92,000 crashes in 1999 resulting in nearly 950 fatalities and 90,000 injuries. The Federal Highway Administration (FHWA) has estimated that the public cost of red light running crashes amounted to over $7 billion in 1998 alone (Hasson, 2000).

As a result of red light running crashes, large numbers of persons are subjected to injury or death. In fact, research has indicated that vehicle occupants are more likely to be injured in red-light running crashes than in most other types of crashes. In one study, occupant injuries occurred in 45 percent of red-light running crashes, compared with 30 percent for other crash types (IIHS, 1993).

Frequency of RLR Violations

The literature indicates that motorists frequently run red lights. Research conducted in Arlington, Virginia at two intersections in 1994 and 1995 revealed 8,121 red light runners over a period of 2,694 hours for a rate of 3.0 red light runners per hour (Retting et. al.,1998). This violation rate climbed to as many 12.0 violations per hour during peak traffic conditions (Hill et. al., 2001). Violation data from 13 intersections in Iowa yielded violation rates ranging from 0.45 to 38.50 violations per 1,000 vehicles and from 0.09 to 9.78 violations per hour (Kamyab et. al., 2002). Additional research conducted at twelve intersections in California suggests a violation rate of 1.3 violations per 1,000 vehicles (Meadow et. al., 1999).

Red Light Running Violation Frequency and Crash Relationship

Bonneson et al. (2002) developed a regression model to quantify the relationship between total crashes and the red light running violation frequency for an intersection approach:

$$C_3 = m_y e^{b_0} ADT_s^{b_1} ADT_c^{b_2} RLR_r^{b_e}$$

where:
C_3 = three-year count of right-angle and left-turn crashes
m_y = number of years associated with the crash data
ADT_s = average daily traffic volume on the subject approach
ADT_c = average daily traffic volume on the cross street
RLR_r = red-light running rate on the subject approach (number of red light violations per 1,000 vehicles)
b_i = regression coefficients

This model was calibrated using data obtained from 12 intersection approaches located throughout Texas. The resultant model is represented by the following equation:

$$C_1 = 0.00278 ADT_c^{0.614} RLR_r^{0.387}$$

where C_1 is the predicted annual crash frequency rate (crashes/year) for an intersection approach. The ADT of the subject approach term (ADT_s) is not included in the final calibrated model because its regression coefficient (b_1) was not significantly different from zero. The R^2 value for this model is 0.57 with a Pearson's χ^2 of 8.86. Because the χ^2 value for the model is less than the associated χ^2 critical value of 16.9 (calculated by Bonneson et. al.) the hypothesis that the model fits the data cannot be rejected.

Examination of this calibrated crash frequency model suggests an exponential relationship between crash frequency and red light running violation frequency (when ADT is fixed).

Characteristics of Red Light Runners

Retting et. al. conducted research to develop a profile of red light runners. To identify characteristics of drivers that run red lights, 1,373 observations (462 violators and 911 red light compliers) were taken in Arlington, Virginia in 1994 and 1995. Violators were found to be younger, less likely to wear seat belts, possess poorer driving records, and drive smaller and older vehicles than those that complied. Red light runners were also found to be more likely to have multiple speeding convictions on their driving records (Retting and Williams, 1996).

Intersection Characteristics and Red Light Running

The Federal Highway Administration compiled an overview list of conditions that may contribute to red light running frequency (FHWA and NHTSA, 2003):

- Grade – May increase the time and distance needed by a motorist to stop a vehicle at an intersection.
- Poor visibility – May prevent a motorist from reacting to the traffic signal in time to stop.
- Temporary roadside obstructions – May block a motorist's field of view and contribute to driver confusion.
- Line of sight – Insufficient line of sight may reduce reaction times.
- Sign reflectivity – May effect sign legibility and contribute to driver confusion.
- Traffic volume – Increased traffic volumes may lead to increased numbers of red light running violations/crashes.
- Signal timing – Inadequate signal timing may increase red light running frequency.
- Weather – May distract drivers or result in increased stopping distances.

Several of these factors have been examined by researchers and are included in this research effort.

Research conducted by Mohamedshah et. al. in 2000 sought to address the following questions concerning red light running (RLR) crashes (Mohamedshah et. al., 2000):

- Does the width of the cross-street have any effect on RLR crash risk?
- What is the relationship of other select intersection characteristics to RLR crashes?
- Using this information, how can one better target urban intersections for traffic law enforcement techniques such as RLR cameras or heightened intersection enforcement coupled with publicity?

To answer these questions, Highway Safety Information System (HSIS) databases for a four-year period (1993-1996) were used. These databases provided information about

RLR crashes at 1,756 four-approach signalized urban intersections and 4,709 two-vehicle RLR crashes in California (Mohamedshah et. al., 2000).

Regression models were employed to examine the effects of the following factors on RLR crashes (Mohamedshah et. al., 2000):

- Number of lanes on both streets (a surrogate for street width)
- Average Daily Traffic (ADT)
- Traffic control type

Analysis of the regression models revealed that when a vehicle entered from a minor street and crossed a mainline street, a seven percent increase in RLR crashes occurred for each additional lane increase on the mainline. This effect was not seen for mainline vehicles crossing a minor street (Mohamedshah et. al., 2000).

The researchers devised two hypotheses with regard to ADT. The first hypothesis stated that as approach street ADT increased, more vehicles would approach the red light and thus more RLR violations/crashes would result. The second hypothesis was that higher ADT on the cross-street (minor street) would reduce the number of safe gaps available and produce a higher chance of an RLR vehicle colliding with another vehicle (Mohamedshah et. al., 2000).

The first hypothesis was shown to be true for Mohamedshah's dataset, as the regression models indicated that RLR crashes on the mainline increased with higher mainline ADT. The second hypothesis was also shown to be true; as the ADT of the cross street increased so did the number of RLR crashes (Mohamedshah et. al., 2000).

The positively correlated relationship between ADT and red light running frequency was also noted by Baguley (1998) at seven rural intersections in England. Baguley also noted a small positive correlation between red light running frequency and approach speed and an inverse correlation with cross-street ADT.

Mohamedshah also examined the effect of differing types of traffic controls on red light running frequency. He analyzed three types of signals: fully actuated, semi-actuated, and pre-timed. Based upon the analysis of RLR at these signal types, approaches equipped with fully actuated signals were found to have more crashes than approaches with semi-actuated and pre-timed signals (approximately 35 to 39 percent more crashes for fully actuated than for pre-timed signals) (Mohamedshah et. al., 2000). This finding suggests that the actuation of a signal may "surprise" the driver and leave the driver unable to successfully stop prior to entering the intersection or enter the intersection prior to the red signal.

Retting et. al. investigated the effect of signal timing on red light compliance. The *Manual on Uniform Traffic Control Devices* (MUTCD) suggests that typical yellow (change) intervals be from three to six seconds in length, with allowances for longer intervals where traffic speeds are higher. The length of all-red (clearance) intervals is largely a function of traffic speed and width of the intersecting street (Retting and Greene, 1997). In 1994, the Institute of Transportation Engineers (ITE) published "Determining Vehicle Signal Change and Clearance Intervals" which addressed these factors and established computational methods for determining appropriate intersection signal timing schemes (ITE, 2001).

Retting et al. examined the degree to which adherence to ITE signal timing recommendations influenced red light running. Twenty approaches at ten intersections were studied over the course of one year beginning in October 1992. The results of this research indicate that as yellow interval lengths approached or surpassed ITE recommended lengths, the incidence of red light violations decreased (Retting and Greene, 1997).

With regard to the timing of all-red intervals, the effect on the incidence of red light running was not as pronounced. The greatest reductions in red light running violations occurred as the all-red interval approached 80 percent of the ITE recommended timing and continued to decrease for timings above and beyond the ITE recommended interval length. However, for all-red intervals that were less than 80 percent of the ITE recommended length, increases in interval length yielded increases in the number of red light violators. This result suggested that all-red interval length was an inconclusive predictor of red light violations (Retting and Greene, 1997).

Bonneson et al. (2002) also examined the relationship between signal cycle length and red light running violation frequency. Their research concluded that the frequency of red light running violations is positively correlated with the frequency of yellow-signal presentation (i.e., how often the yellow signal appears within a period of time). The research further suggests that an increased cycle length, and thus a reduction in the number of times the yellow signal is presented, should be accompanied by a reduction in red light running violation frequency. Furthermore, Bonneson et. al. note (in agreement with Retting and others) that as yellow interval duration is increased, the frequency of red light running violations is decreased.

Section 3
Methodology

Overall Research Approach

The goal of this research work was to investigate the use of statistical modeling techniques to predict red light violations based upon intersection traffic operational and geometric characteristics. Rather than rely on cumbersome, long-term data collection techniques, the models were built using a limited amount of readily-obtainable red light running violation data and intersection field studies. Although previous research (by Mohamedshah) utilized red light running crash data to analyze the relationship of intersection characteristics to red light running, this research opted to use violation data for the following reasons:

- Crashes are relatively rare events. An intersection may experience only a small number of intersection crashes per year but would experience possibly hundreds of red light running violations.
- Crashes may be recorded improperly. Crash data recording can often be subjective. For example, one official may attribute accident causation to "failure to yield," while another may attribute causation to "red light running," leading to difficulty in crash data analysis. Given the relatively small number of crashes for any particular intersection, errors in data recording can lead to an incorrect analysis.
- At intersections with low traffic volumes, RLR violations occur, but the possibility of a crash is assumed to be small due to the low volumes.

This chapter provides a detailed account of the methodologies used in data collection and analysis.

Red Light Violation Data Collection

The following criteria were established to obtain a uniform set of red light violation data and useful intersection/traffic parameters:

- Intersections were chosen to represent a cross-section of geometric and traffic operational parameters (limited to four-legged intersections). Intersection selection was not based upon red light running-attributed crash history.
- Due to the potential modifications of driver behavior, violation data must not have come from intersections where automated enforcement systems had been previously deployed. Data could be taken from intersections that were

included in red light camera demonstration projects (i.e., no automated red light violation citations were being issued).

- Violation data must have come from the hours of 2-8 PM on "normal" (i.e., non-holiday) weekdays. At least one 2-8 PM collection window for each intersection approach was required for the approach to be included in the report. Collecting and using data from entire 24-hour periods was found to be too cumbersome, so the 2-8 PM data collection period was chosen. Figure 3.1, drawn from Kamyab et. al., illustrates the distribution of red light running violations by time of day and suggests that the 2-8 PM data collection period allows collection of data during afternoon high and low violation periods. Figure 3.2, also from Kamyab et. al., suggests that weekdays typically have a higher occurrence of violations than weekends.
- A consistent definition of a red light running violation was necessary to ensure data compatibility. For this research, any vehicle which had not passed over the stopbar prior to the red signal and then proceeded into the intersection was considered to have committed a red light running violation. The violation definition was extended to also include vehicles which failed to stop before turning right with a red signal. Any red light violation recording technique which did not adhere to the above red light violation definition was excluded from consideration for this research. This consideration specifically eliminated those techniques (e.g., some automated enforcement systems) which allow a "grace" period of a fraction of a second after the red signal indication before violations are recorded.

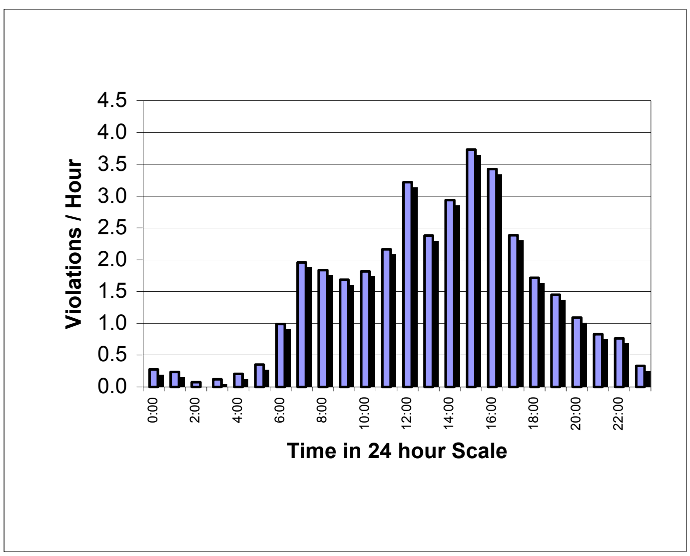

Figure 3-1: Hourly red light running violations/hour distribution (Kamyab, et. al., 2000)

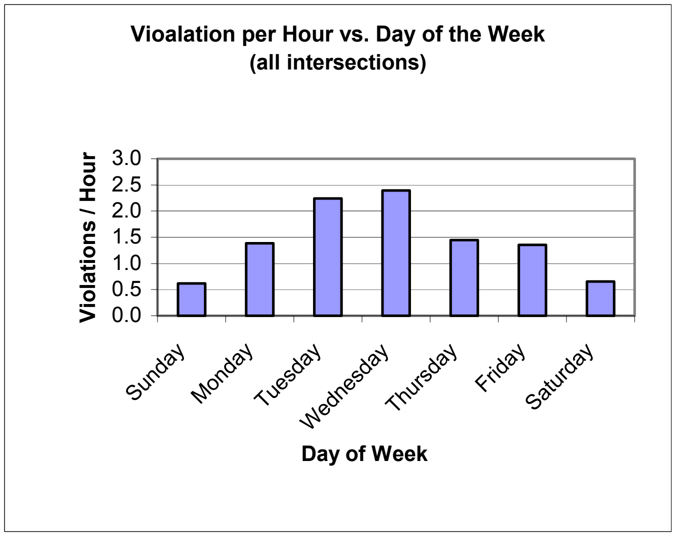

Figure 3-2: Daily red light running violations/hour distribution (Kamyab, et. al., 2000)

Red light running violation data was obtained from the following sources:

- The Center for Transportation Research Excellence (CTRE) at Iowa State University in Ames, IA.
- The City of Tuscaloosa Department of Transportation (TDOT) in Tuscaloosa, AL.
- Precision Traffic Systems, Austin, TX.
- The City of Los Angeles Police Department (LAPD) in Los Angeles, CA.
- University Transportation Center for Alabama (UTCA) Project 00470, The University of Alabama. "A Demonstration of Red Light Running Camera Technology in Alabama."

Center for Transportation Research Excellence

CTRE researchers provided red light violation data obtained during the course of their work on "Red Light Running in Iowa: The Scope, Impact, and Possible Implications" (Kamyab et. al., 2000). Violation data was collected at 12 intersections within Iowa over the course of three months in mid-2000 as part of a red light running camera demonstration project (i.e., no citations issued, camera data used for research purposes only). Not all of the Iowa data was used, for the following reasons:

- Insufficient amounts of 2-8 PM data were available for some locations, which prohibited their use in the model construction.
- Intersection ADT values were not available for several intersections. ADT was assumed to be a key variable in the prediction model, so these intersections were eliminated from consideration.

Seven Iowa intersections were removed during the data reduction process, leaving five intersections for input into the prediction model. A total of 140 hours of 2-8 PM violation data were obtained.

Tuscaloosa Department of Transportation

TDOT operates 26 cameras as part of its traffic management system. These cameras are deployed at a number of major intersections and heavily traveled locations within Tuscaloosa. TDOT granted the researchers permission to use several of these cameras to record traffic at intersections and then manually count red light violations. Six intersections were chosen because of their proximity to TDOT cameras and because a high quality image of the approach could be obtained (i.e., all lanes and traffic signals could be easily seen). From each of these intersections, 36 hours (216 hours in total from the six intersections) of 2-8 PM weekday traffic was recorded, and red light violations were manually counted.

Precision Traffic Systems (PTS)

PTS is a private engineering firm based in Austin, Texas which specializes in red light running camera development and deployment. Red light violation data from a red light camera demonstration project at one intersection in Richardson, Texas and one intersection in San Antonio, Texas was provided by PTS. This data was collected during field demonstration projects conducted by PTS. PTS provided a large amount of violation data which resulted in the inclusion 36 hours of 2-8 PM weekday violation counts for both intersections.

Los Angeles Police Department

In 1996, the City of Los Angeles conducted a red light running field demonstration project. This project involved the installation of red light running cameras at three intersections. Red light violation data recorded by the demonstration cameras was made available to the researchers. From this data, a total of 18 hours of 2-8 PM red light running violation data was obtained.

University Transportation Center for Alabama Project 00470

Concurrent with this research project, UTCA and the Alabama Department of Transportation conducted a field demonstration of red light running camera equipment in Tuscaloosa, Alabama. Red light running cameras were placed at three locations in Tuscaloosa and recorded violation data (no citations were issued) for approximately one month at both sites. Data from two of the three demonstration sites was available in time for inclusion in this research. A total of 72 hours of 2-8 PM violation data was obtained from this demonstration.

Site Visits

Additionally, site visits were conducted at each intersection included in this study to obtain intersection characteristics data. These site visits were conducted during July-October 2002. Data was obtained during the site visits by both visual inspection and measurement of the intersection.

Summary

From the above sources, data from 19 intersection approaches was compiled. 1,775 violations were observed over 554 hours (for an overall 2-8 PM violation rate of 3.2 violations/hour). Table 3.1 provides a summary of the red light violation data and the intersection and traffic parameters associated with the violations.

Model Variables

The variables that were considered for inclusion in the Red Light Running Prediction model are discussed below. They were chosen because they were readily obtainable and were suspected to contribute to the frequency of red light violations. For example, the "Approach Speed Limit" was readily obtainable from speed limit signs at or near the intersection. Many of the other variables were obtained by a brief intersection survey. Note that signal type and signal timing elements were excluded from the variables examined in this research. These elements were not included because they were not available to the researchers for a majority of the intersections studied (i.e., signal timing data was not recorded by the Iowa researchers or the Los Angeles Police Department).

2:00-8:00 PM Violation Count

This variable records the number of violations counted during the 2-8 PM counting periods for each intersection approach site. This count is an aggregate of all 2-8 PM counting periods conducted at the particular site. The violation counts obtained during the course of this research ranged from four (at 14th St. and Central Ave. in Dubuque, IA) to 253 (at Beltline Road and Highway 75 in Richardson, TX).

2:00-8:00 PM Violations Per Hour

This variable is determined by dividing the "2:00-8:00 PM Violation Count" by the total number of hours of red light violation data available from the intersection site. For example, 82 2-8 PM red light running violations were recorded over 26 hours at the intersection of U.S. 75 and 18th Street in Sioux City, Iowa for a red light running frequency of 3.15 2-8 PM violations per hour. The average 2-8 PM violation rate from the nineteen sites was determined to be 3.2 violations/hour and ranged from 0.13 violations per hour (14th Street/Central Avenue, Dubuque, IA) to 7.03 violations per hour (Beltline Road/Highway 75, Richardson, TX).

Table 3-1. Intersection summary

Intersection	Location	Direction	Approach Speed	Crossing Speed	Vi	Hours of Data	Vi/Hr	ADT	ADT (Cross)
I-74 Off Ramp at State St.	Iowa	North	55	32.5	81	36	2.25	14384	13224
14th St. at Central Avenue	Iowa	East	25	25	4	30	0.13	8770	10745
Riverside Dr. at Iowa 1/6	Iowa	South	55	32.5	122	24	5.08	20815	27202
35th St. at University Ave.	Iowa	North	35	35	78	24	3.25	9600	13229
US 75 at 18th St.	Iowa	South	55	32.5	82	26	3.15	10339	4138
Beltline Rd. at Hwy 75	Texas	West	30	45	253	36	7.03	39780	NA
AL 69 at Skyland Blvd.	Alabama	South	55	45	89	36	2.47	34495	26268
University Blvd. at 19th Ave.	Alabama	West	30	25	77	36	2.14	13456	420
L. Wallace N at Stillman Blvd.	Alabama	North	45	35	194	36	5.39	27338	6057
12th Ave. at University Blvd.	Alabama	North	25	30	64	36	1.78	6374	NA
L. Wallace N at 8th St.	Alabama	North	45	25	176	36	4.89	23992	17814
15th St. at Lake Ave.	Alabama	East	45	25	101	36	2.81	34558	NA
McFarland Blvd. at McFarland Mall	Alabama	North	50	25	188	36	5.22	38225	NA
Skyland Blvd. at McFarland Mall	Alabama	East	45	25	104	36	2.89	36642	NA
University Blvd. at 22nd Ave.	Alabama	West	30	25	65	36	1.81	12885	1988
La Brea at Rodeo	California	North	35	35	20	6	3.33	42485	NA
Florence at Figueroa	California	North	35	35	10	6	1.67	31394	NA
Soto St. at Olympic Blvd.	California	North	35	35	12	6	2.00	25112	NA
New Braunfels Ave.at E. Commerce St.	Texas	North	30	30	55	36	1.53	9060	NA

Vi = Number of violations during data collection period
Speed limits given in miles/hour

Table 3-1. Intersection summary (continued)

Intersection	Lanes	Crossed Lanes	DP	DN	Locale #1	Locale #2	Pavement Type	Surface Condition	Grade
I-74 Off Ramp at State St.	3	5	NA	375	Urban	Commercial	Concrete	Heavily Worn	Downhill
14th St. at Central Avenue	2	2	270	285	Urban	Commercial	Asphalt	Worn	Downhill
Riverside Dr. at Iowa 1/6	4	6	NA	NA	Urban	Commercial	Concrete	Worn	Negligible
35th St. at University Ave.	3	5	350	NA	Urban	Commercial	Asphalt	Good	Downhill
US 75 at 18th St.	3	2	NA	NA	Rural	Industrial	Concrete	Worn	Negligible
Beltline Rd. at Hwy 75	3	2	NA	440	Dense Urban	Commercial	Concrete	Worn	Negligible
AL 69 at Skyland Blvd.	2	6	NA	NA	Urban	Commercial	Asphalt	Good	Downhill
University Blvd. at 19th Ave.	2	2	365	430	Urban	Commercial	Asphalt	Good	Negligible
L. Wallace N at Stillman Blvd.	4	5	360	360	Dense Urban	Commercial	Asphalt	Good	Negligible
12th Ave. at University Blvd.	2	3	675	NA	Urban	Commercial	Asphalt	Good	Negligible
L. Wallace N at 8th St.	3	2	360	360	Dense Urban	Commercial	Asphalt	Good	Negligible
15th St. at Lake Ave.	3	2	930	NA	Urban	Commercial	Asphalt	Good	Downhill
McFarland Blvd. at McFarland Mall	3	4	380	600	Urban	Commercial	Asphalt	Worn	Negligible
Skyland Blvd. at McFarland Mall	2	3	275	800	Urban	Commercial	Asphalt	Worn	Uphill
University Blvd. at 22nd Ave.	2	2	430	430	Dense Urban	Commercial	Asphalt	Worn	Uphill
La Brea at Rodeo	3	7	975	NA	Dense Urban	Commercial	Asphalt	Good	Negligible
Florence at Figueroa	3	7	300	800	Dense Urban	Commercial	Asphalt	Worn	Negligible
Soto St. at Olympic Blvd.	3	7	NA	NA	Dense Urban	Industrial	Asphalt	Worn	Negligible
New Braunfels Ave.at E. Commerce St.	1	4	225	NA	Urban	Residential	Asphalt	Good	Negligible

Note: DP = Distance to Preceding Intersection and DN = Distance to Next Intersection (both in feet)

Approach Direction

This variable gives the compass direction of the approaching street. This direction was determined from street maps of the intersections/approaches. Categorizing the nineteen intersection approaches resulted in the following breakdown:

- Northbound – 10 Approaches
- Southbound – 3 Approaches
- Eastbound – 3 Approach
- Westbound – 3 Approaches

It is hypothesized that Approach Direction may be relevant to red light violations because of potential visibility problems (i.e., sun in the driver's eyes obscuring the traffic signal). Because data was collected during the afternoon hours, it was possible that westbound approaches experienced a greater number of red light violations due to restricted visibility.

Approach Speed Limit

This variable records the posted speed limit for the approach street. The average posted speed limit of the nineteen approaches was 40 mph. The distribution of posted speed limits was as follows:

- 25 mph – 2 Approaches
- 30 mph – 4 Approaches
- 35 mph – 4 Approaches
- 45 mph – 4 Approaches
- 50 mph – 1 Approach
- 55 mph – 4 Approaches

It is hypothesized that as approach speed (given in this research by the approach speed limit) increases, the likelihood for red light running violations increases as well, because drivers are less likely to be able to stop at higher speeds.

Crossing Speed Limit

This variable records the posted speed limit for the crossing street. For cases where the crossing street was found to have different posted speed limits for each direction, the speed limits were averaged. For example, at the I-74 Off Ramp/State Street intersection, State Street (the crossing street) has a speed limit of 35 MPH in one direction and 30 MPH in the other. For this crossing street, an average posted speed limit of 32.5 MPH was used in model development. The average crossing speed limit for the nineteen approaches was 31.45 mph. The distribution of posted crossing speed limits at the nineteen approaches follows:

- 25 mph – 7 Approaches
- 30 mph – 2 Approaches
- 32.5 mph – 3 Approaches
- 35 mph – 5 Approaches
- 45 mph – 2 Approaches

The speed limit of the crossing street is included in the variable set because it provides information about the type of roadway that is being crossed (e.g., low speed limit may indicate that the road being crossed is a minor road).

Average Daily Traffic (ADT)

This variable records the number of vehicles passing a point (in either direction on the approach road) for an entire non-holiday weekday. For this research, ADT was obtained from several sources:

- Iowa Intersections – ADT data provided by the Iowa Department of Transportation.
- Texas Intersections – ADT data provided by the Texas Department of Transportation via PTS.
- Alabama Intersections – ADT data provided by the Tuscaloosa Department of Transportation.
- Los Angeles Intersections – ADT data provided by the City of Los Angeles.

The most recent ADT counts were used for each approach. ADT counts from the five Iowa intersections were generated in either 2001 or 2002; the Texas ADTs were recorded in 2002 and 2003; the Tuscaloosa intersection ADT values were generated in 1998 and 2000; the Los Angeles intersections ADT values were generated in 1996. The average ADT for the nineteen sites was 23,142 vehicles per day with a range from 6,374 (12^{th} Ave./University Blvd., Tuscaloosa, AL) to 42,485 vehicles per day (La Brea/Rodeo, Los Angeles, CA).
Previous research (Mohamedshah et. al., 2000 and Baguley, 1998) suggests that as ADT increases, more vehicles have the opportunity to violate the red signal, thus leading to more red light running violations.

Crossing ADT

This variable records the ADT of the crossing street (in both directions if the crossing street is not one-way). ADT counts for crossing streets were available for ten of the nineteen intersection approaches included in this study. The average Crossing ADT for the ten sites was 12,109 vehicles per day with a range from 420 (University Boulevard./19^{th} Avenue., Tuscaloosa, AL) to 27,202 vehicles per day (Riverside Dr./Iowa 1/6, Iowa City, IA). Mohamedshah et. al., 2000 hypothesized that as Crossing ADT increases, the opportunity for “safe gaps” may decrease, resulting in increases red light running crashes.

Number of Approach Lanes

This variable records the number of lanes on the monitored intersection approach. For clarification, this variable records only the number of lanes on the approach, not inclusive of the lanes on the opposite approach. The number of approach lanes ranged from one to four (see "Special Cases" below) with the following distribution:

- 1 Lane – 1 Approach
- 2 Lanes – 6 Approaches
- 3 Lanes – 10 Approaches
- 4 Lanes – 2 Approaches

It was hypothesized that as the number of approach lanes increased that there was greater opportunity for vehicles to violate a red signal (i.e., if the lead vehicle on a one lane roadway stopped at a red signal, all trailing vehicles stopped as well and did not have the opportunity to run the signal).

Special Cases

The number of approach lanes included in the analysis portions of this report for Alabama Highway 69 South at Skyland Boulevard. intersection approach in Tuscaloosa, AL was given as two. However, there were actually a total of five lanes at this approach. Neither left nor right turn violations were recorded at this intersection; thus, two left turn lanes and a right turn only lane were removed from the analysis. The ADT (not crossing ADT) was adjusted, after consultation with Tuscaloosa DOT officials, to more accurately reflect traffic in the two recorded lanes.

A similar situation arose at Beltline Road at Highway 75 in Richardson, TX. In this situation, the center two lanes of a four-lane approach were monitored. The outer lanes were not selected for monitoring. The outer lane on the left hand side of the approach was a left-turn only lane, while the outer lane on the right hand side of the approach was a straight-only lane. The ADT of the approach was modified to reflect only vehicles traveling in the two monitored center lanes, and the approach is listed in the model as having two approach lanes.

Number of Crossing Lanes

This variable records the number of lanes crossed by traffic on the monitored approach as it traverses the studied intersection from the stopbar to the other side of the intersection. Medians which were as large as a lane were counted as one lane. The number of crossing lanes at the nineteen intersections ranged from two to seven with the following distribution:

- 2 Lanes – 7 Approaches
- 3 Lanes – 2 Approaches
- 4 Lanes – 2 Approaches
- 5 Lanes – 3 Approaches
- 6 Lanes – 2 Approaches
- 7 Lanes – 3 Approaches

The number of crossing lanes was viewed as representative of intersection width and was included in the variable set for that reason.

Distance to Preceding Intersection

This variable measures the distance from the stop bar of the intersection approach to the stop bar of the preceding intersection. It was hypothesized that as distance to the preceding intersection increased, vehicles had an opportunity to achieve greater speeds and were less likely to be able to stop at a red indication.

This measurement was available for thirteen of the nineteen intersections. For two cases (Alabama Highway 69/Skyland Boulevard and I-74 Off-Ramp/State Street) the measurement was not applicable because the roadway leading into the intersection was limited-access in nature (i.e., no traffic signals/intersections present). For the other four cases, the distance was so great (over ¼ mile) that it was assumed to have a negligible effect on red light violation behavior at the studied intersection. For the thirteen intersections where measurements were made, the distance to the preceding intersection ranged from 270 to 975 feet with an average distance of 453.5 feet. Because of the limited dataset for this variable, results drawn using this variable should be used with caution.

Distance to Next Intersection

This variable measures the distance from the stop bar of the intersection approach to the stop bar of the next intersection. This measurement was available for ten of the nineteen intersections. The other nine locations were not measured because the distance was so great (over 1/4 mile) that it was assumed to have a negligible effect on red light violation behavior at the studied intersection. For the ten intersections where measurements were made, the distance to the next intersection ranged from 285 to 800 feet with an average distance of 488 feet. Because of the limited dataset for this variable, observations made using this variable should be used with caution.

Locale

The locale variables record the surrounding environment of the intersection. These variables may provide important information about the likelihood of red light violations within certain environments. Previous research has suggested that the setting of an intersection may provide clues as to the likelihood of red light running violations (Porter and England, 2000).

Locale 1 (Rural, Urban, Dense Urban)

The first locale variable records the relative density of surrounding structures in three categories:

- Rural – Few or no structures or buildings in the immediate area of the intersection. This is typically characterized by the presence of undeveloped land.
- Urban – Several buildings likely in the immediate area of the intersection. The intersection is likely located within either a small town or a suburb.
- Dense Urban – The areas immediately adjacent to the intersection are highly developed. The intersection is likely located in the downtown or central business district of a large city.

When performing statistical calculations, Locale (Rural, Urban, Dense Urban) is a multiple categorical variable given by two indicator variables (LocRUDUInd1 and LocRUDUInd2). The following describes the use of the two indicators to categorize intersections:

For Rural	-	LocRUDUInd1 = 1
	-	LocRUDUInd2 = 1
For Urban	-	LocRUDUInd1 = 1
	-	LocRUDUInd2 = 2
For Dense Urban	-	LocRUDUInd1 = 2
	-	LocRUDUInd2 = 2

Of the nineteen intersections studied for this report, eleven were characterized as being in an "Urban" locale, seven were characterized as being in a "Dense Urban" locale, and one was characterized as being in a "Rural" locale.

Locale 2 (Residential, Commercial, Industrial)

The second Locale variable categorizes the intersection surroundings by the type of immediately adjacent development.

- Commercial – Intersections with "Commercial" locales are characterized by being surrounded by small businesses, shopping malls, or other retail/wholesale establishments.
- Industrial – Intersections with "Industrial" locales are characterized by being surrounded by factories, mills, or other similar production facilities.
- Residential – Intersections with "Residential" locales are characterized by being surrounded by homes, apartments, or other housing developments.

When performing statistical calculations, the Locale (Commercial, Industrial) variable is a multiple categorical variable composed of two indicator variables.

For Residential Locale	- Locale Indicator 1 = 2
	- Locale Indicator 2 = 2
For Commercial Locale	- Locale Indicator 1 = 1
	- Locale Indicator 2 = 1
For Industrial Locale	- Locale Indicator 1 = 1
	- Locale Indicator 2 = 2

Of the nineteen intersections studied for this report, one was categorized as "Residential", 16 were categorized as "Commercial," and two were categorized as "Industrial."

Pavement Type

This variable records the type of pavement present at the intersection. The two most common pavement types are "Concrete" and "Asphalt". The Pavement Type variable is a binary indicator variable. A Pavement Type of "Concrete" received a "1," while a Pavement Type of "Asphalt" received a "2." Fifteen of the nineteen intersections studied for this report were asphalt, while four were concrete.

Surface Condition

This qualitative variable records the condition of the pavement on the approach street and within the intersection itself. It was hypothesized that poor surface conditions could lead to driver uneasiness that altered driving behavior. This variable was divided into three categories that are described below.

- Good – Pavement is either brand-new or contains little to no cracking and no loose pavement material. Pavement markings are in excellent condition and retain much or all of their original visibility.
- Worn – Pavement has experienced moderate cracking. Loose pavement material may be present in small quantities, and pavement markings may have experienced some wear and are less visible than when initially placed.
- Heavily Worn – Pavement is in poor condition. Surface is heavily cracked, loose pavement material is likely present, and pavement markings are heavily worn or no longer visible. Pavement is generally in need of rehabilitation.

Surface condition is represented by two indicator variables: Surface Condition Indicator 1 and Surface Condition Indicator 2. These two variables are binary (i.e., only possess two values). The following describes the use of the two indicator variables for surface condition cases:

For Good Pavement	- Surface Condition Indicator 1 = 2
	- Surface Condition Indicator 2 = 1
For Worn Pavement	- Surface Condition Indicator 1 = 1
	- Surface Condition Indicator 2 = 2

For Heavily Worn Pavement	- Surface Condition Indicator 1 = 2
	- Surface Condition Indicator 2 = 2

Of the nineteen intersections studied for this report, nine were found to be "Good," nine were found to be "Worn," and one was found to be "Heavily Worn."

Grade

It was hypothesized that drivers on a downhill grade were less likely (able) to stop for a red signal, and thus intersections with a downhill approach might have a higher frequency of red light violations. The grade (slope) of the incoming approach was measured qualitatively. To allow a rapid field survey, a qualitative measurement of roadway grade was obtained. Grade was included in the model development in three subjective categories:

- Downhill – Significant downhill slope (negative grade)
- Negligible – Little or no roadway slope
- Uphill – Significant downhill slope (positive grade)

Each of these categories was described through the use of two indicator variables:

For Downhill Grade	- Grade Indicator 1 = 2
	- Grade Indicator 2 = 1
For Negligible Grade	- Grade Indicator 1 = 1
	- Grade Indicator 2 = 2
For Uphill Grade	- Grade Indicator 1 = 2
	- Grade Indicator 2 = 2

Of the nineteen intersections studied for this report, five were found to have a "Downhill," grade, thirteen were found to have a "Negligible" grade, and two were found to have an "Uphill" grade.

Section 4
Model Construction and Analysis

The creation of the Red Light Running Prediction model was accomplished through the use of regression analysis. Regression analysis is "a statistical technique for modeling and investigating the relationship between two or more variables" (Montgomery and Runger, 1998). Regression models were constructed and then analyzed for their predictive ability at 95% level of confidence (as indicated by the regression and individual regression variable p-values).

Modeling Techniques

The SPSS Version 11.0.1 software package was used to perform the statistical analysis for this project. SPSS can perform a variety of statistical operations including regression analysis (linear, multiple linear, curvilinear, and curve estimation). The software provides tools to fit 11 regression models: Linear, Logarithmic, Inverse, Quadratic, Cubic, Power, Compound, S, Logistic, Growth, and Exponential.

Initial Variable Analysis

Prior to the regression analysis, the correlation between variables in the analysis was examined via a correlation matrix, given in Table 4-1. Variables with significant or notable correlations are given in bold type. These variables likely have a direct relationship to each other and are likely not independent of each other; thus, such variables should not both be included in a regression analysis.

Linear Regression Analysis

Simple linear regression analysis was conducted to examine the relationship between red light running violation frequency and individual intersection geometric and traffic operational characteristics. Each variable was regressed with the violation rate (Violations/Hour) and a "best-fit" regression equation was developed for each response/predictor pair. Table 4-2 provides the results of the simple linear regression analysis. The analysis suggests that none of the variables examined are strong predictors of red light running violation frequency, but does suggest that several variables may possess some degree of predictive power.

Of the predictor variables, only Approach Lanes and ADT of Approach proved to be significant predictors of red light running violation frequency at a 95% confidence level (p-values of 0.003 and 0.024, respectively). From the regression equation with Approach Lanes, the predicted number of red light running violations increases by 1.426 violations/hour per additional lane on the approach street. From the regression equation with ADT, the predicted number of red light running violations increases by 0.7168 violations/hour per 10,000 additional vehicles/day.

Table 4-1. Correlation matrix

	Approach Speed	Crossing Speed	ADT	Crossing ADT	Approach Lanes	Crossed Lanes	Dist. Preceding	Dist. Next
Crossing Speed	0.123							
	0.616							
ADT	0.285	0.301						
	0.237	0.210						
Crossing ADT	0.553	0.529	0.568					
	0.098	0.116	0.087					
Approach Lanes	**0.486**	0.259	0.339	0.269				
	0.035	0.283	0.156	0.452				
Crossed Lanes	0.214	**0.505**	0.298	**0.678**	0.343			
	0.378	**0.027**	0.215	**0.031**	0.150			
DP	0.053	0.122	0.384	-0.435	0.268	0.126		
	0.864	0.691	0.196	0.389	0.376	0.681		
DN	0.168	0.003	0.626	-0.609	-0.109	0.468	-0.333	
	0.643	0.994	0.053	0.200	0.765	0.173	0.420	
Pavement Type	-0.435	-0.359	0.078	-0.200	-0.401	0.068	0.000	0.231
	0.063	0.131	0.750	0.579	0.089	0.781	1.000	0.520

Note: Upper Cell = Pearson's Correlation, Lower Cell = p-value

DP = distance to preceding intersection, DN = distance to next intersection (both in feet)

Table 4-2. Linear regression results

Variable	Equation	R^2	Regression p
Approach Lanes	Vi/Hr = -.829 + 1.462X	0.411	0.003
ADT of Approach (2-direction)	Vi/Hr = 1.437 + (7.168 X 10^{-5})X	0.267	0.024
Locale (Rural/Urban/Dense Urban)	Vi/Hr = 2.391 + 1.484X_1 - .721X_2	0.186	0.192
Grade	Vi/Hr = 2.017 - 1.254X_1 + 1.419X_2	0.159	0.250
Pavement Type	Vi/Hr = 6.004 - 1.625X	0.159	0.091
Approach Speed Limit	Vi/Hr = .724 + .05930X	0.137	0.118
Crossing Speed Limit	Vi/Hr = .282 + .08963X	0.108	0.171
Approach Direction	Vi/Hr = 3.454 + .425X_1 - 1.202X_2 + .467X_3	0.104	0.635
ADT of Cross Street	Vi/Hr = 2.386 + (5.541 X 10^{-5})X	0.099	0.375
Locale (Area Type)	Vi/Hr = 3.940 + 1.049X_1 - 1.731X_2	0.065	0.584
Surface Condition	Vi/Hr = 1.435 - .155X_1 + .970X_2	0.016	0.876
Distance to Preceding Intersection	Vi/Hr = 2.577 + (5.639 X 10^{-4})X	0.008	0.774
Distance to Next Intersection	Vi/Hr = 3.594 - (5.173 X 10^{-4})X	0.002	0.904
Crossed Lanes	Vi/Hr = 3.188 - 0.02315X	0.001	0.915

Curvilinear Regression Analysis

In addition to standard linear regression, several curvilinear regression models were considered. Variables which contain two or more indicators (e.g., Approach Direction is comprised of Approach Indicators 1 and 2) were excluded from the curvilinear regression analysis, as the analyses used were strictly univariate. A summary of the results of the curvilinear regression analysis is given in Table 4-3. As in the case of linear regression analysis, no variable was found to be a strong predictor of red light running violation frequency, however, several variables demonstrated potential predictive power.

Table 4-3. Curvilinear regression results

Variable	Model	Equation	R^2	Regression p
Approach Lanes	Quadratic	Vi/Hr = 0.8575 + 0.0483X + 0.2725X^2	0.425	0.012
Distance to Preceding Intersection	Cubic	Vi/Hr = -13.723 + 0.1004X - 0.0002X^2 + 1X$10^{-7}$$X^3$	0.335	0.276
Approach Speed Limit	S	Vi/Hr = $e^{(2.6711+(-65.532/X))}$	0.324	0.011
Distance to Next Intersection	Quadratic	Vi/Hr = -10.845 + 0.0570X - 5X$10^{-5}$$X^2$	0.300	0.287
ADT of Approach (2-direction)	Logarithmic	Vi/Hr = -11.520 + 1.4783 ln (X)	0.285	0.019
Crossing Speed Limit	Quadratic	Vi/Hr = 9.5651 - 0.4863X + 0.0086X^2	0.155	0.259
ADT of Cross Street	Quadratic	Vi/Hr = 2.5278 + 2X10^{-5}X + 1.3X$10^{-9}$$X^2$	0.109	0.685
Crossed Lanes	Cubic	Vi/Hr = 10.6815 - 6.7775X + 1.7781X^2 - 0.1401X^3	0.086	0.708
Grade	NA	NA	NA	NA
Pavement Type	NA	NA	NA	NA
Approach Direction	NA	NA	NA	NA
Locale (Area Type)	NA	NA	NA	NA
Surface Condition	NA	NA	NA	NA
Locale (Rural/Urban/Dense Urban)	NA	NA	NA	NA

Of the predictor variables in the curvilinear analysis, Approach Lanes, Approach Speed Limit, and ADT of Approach are significant predictors of red light running frequency at 95% level (p-values of 0.012, 0.011, and 0.019, respectively). The Distance to Preceding Intersection and Distance to Next Intersection variables have considerable R^2 values (> 0.3), but because data for these variables was only available for thirteen and ten intersections, respectively, and because the regression p-values associated with these variables (0.276 and 0.287, respectively) were found to be greater than 0.05, the predictive ability of these variables is suspect.

Plots of the curvilinear regression equations for Approach Lanes, ADT, and Approach Speed Limit versus red light running frequency are given in Figures 4-1, 4-2, and 4-3, respectively. The figures indicate that as each variable increases, the predicted number of red light running violations/hour increases as well.

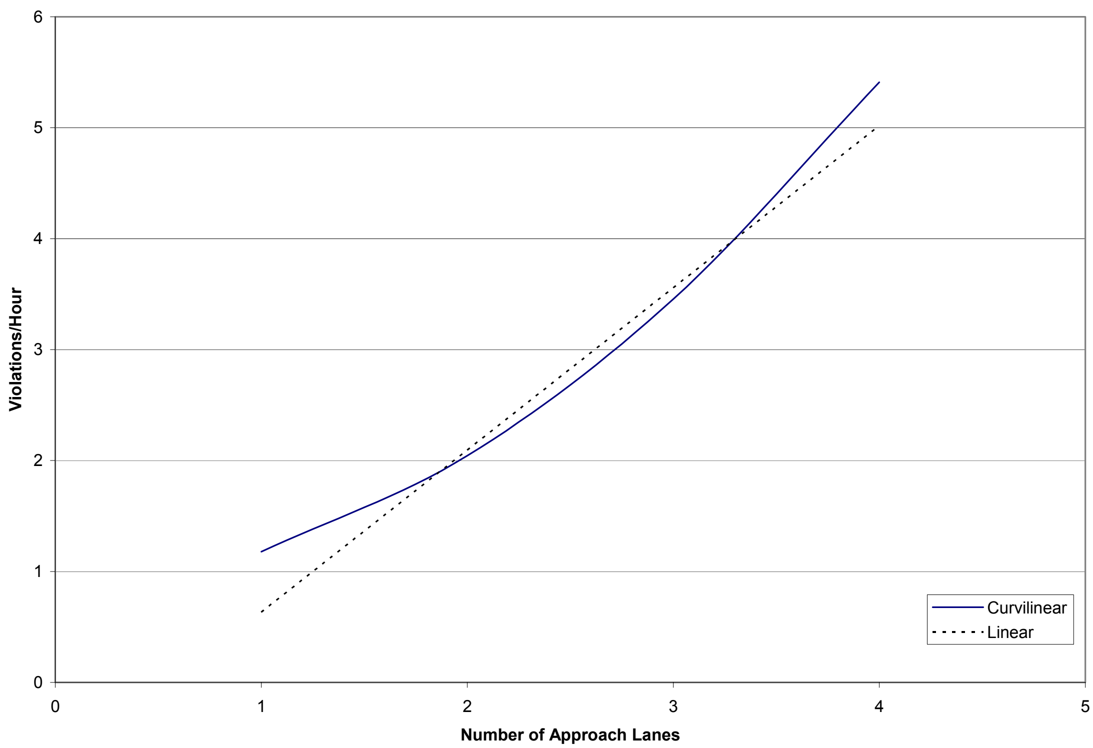

Figure 4-1. Linear and curvilinear models – Number of approach lanes vs. violations/hour

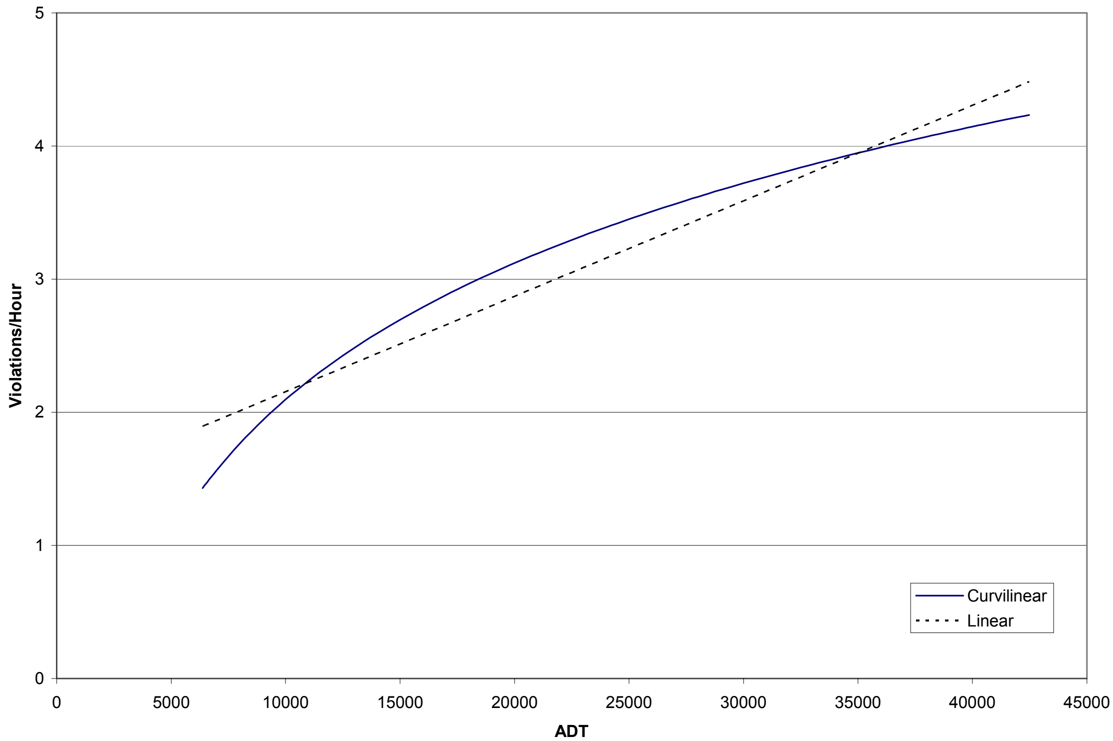

Figure 4-2. Linear and curvilinear models – ADT vs. violations/hour

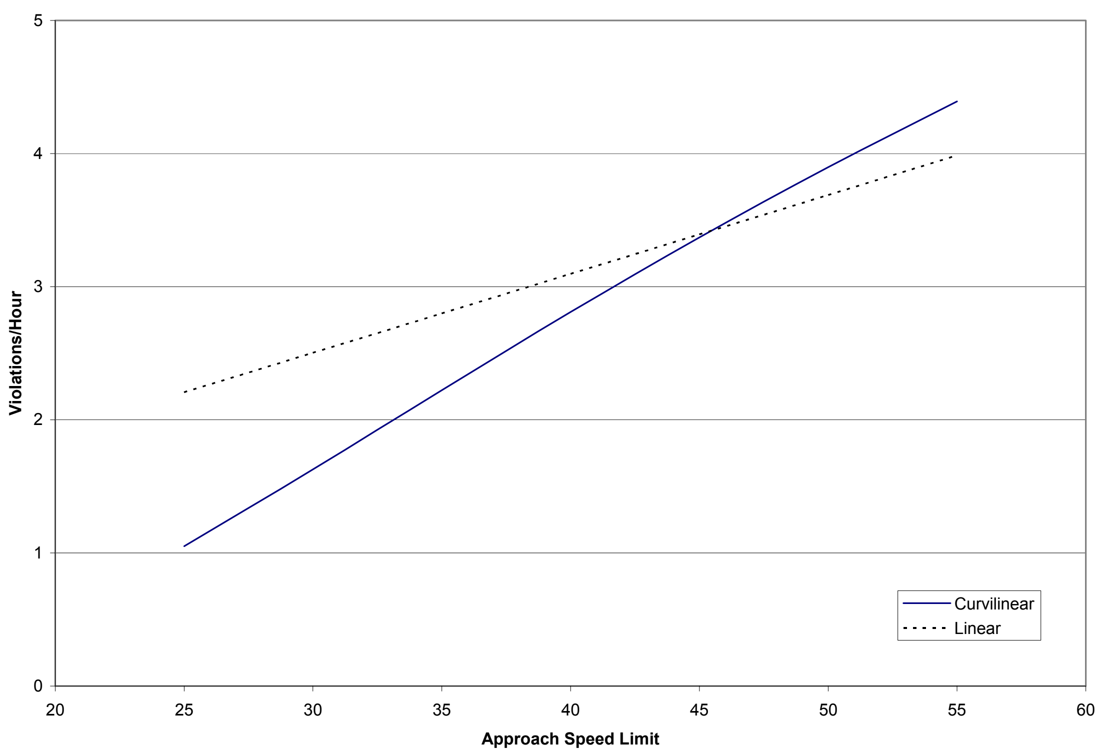

Figure 4-3. Linear and curvilinear models – Approach speed limit vs. violations/hour

Best Linear/Curvilinear Models

For each variable, the best model from the linear and curvilinear regression analyses was selected. Selection was based upon analysis of the R^2 and regression p-value. The results of the best model selection are given in Table 4-4.

Table 4-4. Best linear/curvilinear model selection results

Variable	Best Model	R^2	Regression p
Approach Lanes	Cubic	0.434	0.032
ADT (2-direction)	Cubic	0.314	0.120
Locale (Rural/Urban/Dense Urban)	Linear	0.19	0.192
Grade	Linear	0.159	0.091
Pavement Type	Linear	0.16	0.091
Approach Speed Limit	S	0.324	0.011
Crossing Speed Limit	Quadratic	0.155	0.259
Approach Direction	Linear	0.1	0.635
ADT of Cross Street	Cubic	0.109	0.863
Locale (Area Type)	Linear	0.07	0.584
Surface Condition	Linear	0.02	0.876
Distance to Preceding Intersection	Cubic	0.335	0.276
Distance to Next Intersection	Quadratic	0.300	0.287
Crossed Lanes	Cubic	0.086	0.708

Multiple Linear Regression Analysis

A “Best Subsets” analysis was employed to isolate a satisfactory multivariate linear regression equation. The “best subsets” routine examines all of the models created from all possible combinations of predictor variables and uses the R^2 statistic to select the best model. Variables without full datasets (data not available for all 19 intersections) were excluded, as were variables that were shown (via linear regression analysis) to be extremely poor predictors of red light running violation frequency and multiple categorical variables (variables composed of more than one indicator variable). These multiple categorical variables were excluded from the analysis because the “best subsets” routine was unable to distinguish between multiple categorical variables and non-categorical variables. The variables considered in the “best subsets” routine follow:

- Approach Speed Limit
- Crossing Speed Limit
- Number of Approach Lanes
- Number of Crossing Lanes
- Pavement Type (a binary categorical variable)
- ADT of Approach Street

The results of the “best subsets” analysis are given in Table 4-5.

Table 4-5. Best subsets routine results

Vars	R^2	C-p	S	Approach Speed	Crossing Speed	Approach Lanes	Crossed Lanes	Pavement Type	ADT
1	41.1	8.4	1.3488			X			
1	26.7	14.1	1.505						X
2	51.2	6.4	1.2655			X			X
2	48	7.7	1.3063			X	X		
3	62.4	3.9	1.1466			X	X		X
3	58.5	5.5	1.2054		X	X	X		
4	69.5	3.1	1.0693		X	X	X		X
4	65.6	4.7	1.1358			X	X	X	X
5	69.7	5	1.1055		X	X	X	X	X
5	69.7	5	1.1059	X	X	X	X		X
6	69.8	7	1.1493	X	X	X	X	X	X

Vars = number of variables in model. X indicates variable was included in regression.

The model chosen from the best subsets routine (based upon consideration of the R^2, C-p, and S statistics) contains the Approach Lanes, Crossed Lanes, and ADT variables and is given by the linear equation:

$$Violations / Hour = -0.78 + 1.43AL - 0.321CL + 5.7X10^{-5} ADT$$

where:
AL = number of lanes on the subject approach
CL = number of lanes on the crossing approach
ADT = ADT of subject approach

This model has an R^2 value of 0.624 and a regression p-value of 0.002. Note that although the number of lanes on the subject approach and the ADT of the subject approach likely have a practical correlation, the statistical analysis (shown via Table 4-1) revealed no statistically significant relationship between the variables for this dataset.

Each of the multiple categorical variables found via linear regression to be fair to excellent predictors of red light running violation frequency was then added to the "best subsets" multiple linear regression equation in an attempt to improve the equation's predictive ability. The following multiple categorical variables were added (one-at-a-time) to the "best subsets" equation:

- Locale (Rural, Urban, Dense Urban)
- Grade
- Approach Direction

The results of the additional variable inclusion are given in Table 4-6.

Table 4-6. Variable addition to best subsets model

Variable	R^2 Change	Regression p-value Change	Significant?
Locale (Rural/Urban/Dense Urban)	+ 0.008	+ 0.012	NO
Grade	+ 0.071	+ 0.003	NO
Approach Direction	+ 0.181	- 0.001	YES

Of the multiple categorical variables added to the "best subsets" regression variables, only Approach Direction yielded a significant improvement. The "best subsets" regression with Approach Direction included results in the following equation:

$$Violations / Hour = 3.10 + 1.21AL - 0.520CL + 8.3X10^{-5} ADT + 0.120AI1$$
$$- 2.31AI2 - 0.476AI3$$

where:
AI1 = Approach Direction Indicator 1
AI2 = Approach Direction Indicator 2
AI3 = Approach Direction Indicator 3

This model has an R^2 value of 0.805 and an associated regression p-value of 0.001, which suggest that the model possesses significant predictive power with regard to red light running violation frequency for the limited data used in this study. However, inclusion of approach direction in this regression should be undertaken with caution given the lack of a variety of approach directions available in this research (i.e., ten northbound approaches versus a total of nine approaches from the three other compass directions).

Summary

Several models of the linear, curvilinear, and multiple linear regression form were created to predict red light running frequency based upon one or several of fourteen intersection geometric and traffic operational characteristics. Models created by each form of analysis were shown to have varying levels of red light running frequency predictive ability, with several models demonstrating excellent predictive abilities.

Section 5
Conclusions and Opportunities for Future Work

This section provides a summary of the work conducted during the course of this research and the conclusions drawn from that work. Opportunities for future work are also identified.

Overall

This research identified several models that describe the relationship between red light running violation frequency and intersection and traffic parameters. Data was obtained from 19 intersections in four states. Further intersection characteristic and red light running violation frequency data should be collected and integrated into the database that resulted from this research. Additional data may add further statistical significance to the findings of this research and allow the resultant regression equations to be applied more generally. Until additional intersection and violation frequency data are included in the regression analysis, users should exercise caution in the application of the current regression models.

Data Collection

Data for this project was gathered from five separate sources: the Center for Transportation Research Excellence in Ames, IA; the Tuscaloosa Department of Transportation; Precision Traffic Systems in Austin, TX; the Los Angeles Police Department; and the University Transportation Center for Alabama in Tuscaloosa, AL. After removal of incompatible data, 19 intersection approaches (nine in Alabama, five in Iowa, three in California, and two in Texas) were selected. From each of these intersection approaches, 14 geometric and traffic operational parameters were collected, in addition to a minimum of six hours of red light running violation data during the hours of 2-8 PM on weekdays. A total of 1,775 red light running violations were observed over the course of 554 hours of data collection, for an overall rate of 3.2 red light running violations per hour.

Variable Modeling and Results

Each of the 14 intersection parameters was input into linear, multiple linear, and curvilinear regression models against Vi (average number of violations per hour for each intersection). These models were evaluated based upon their predictive power and goodness-of-fit. The following conclusions were drawn from the results of the model construction:

- The single best predictor of red light running frequency based upon simple linear regression is the number of lanes on the approach (R^2 = 0.411, p = 0.003), followed by the ADT of the approach (R^2 = 0.267, p = 0.024).
- The best predictors of red light running frequency found via curvilinear regression analysis were number of lanes on the approach (R^2 = 0.425, p = 0.012), the speed limit on the approach (R^2 = 0.324, p = 0.011), and the ADT of the approach (R^2 = 0.285, p = 0.019).
- A “best subsets” multiple linear regression routine was applied to six non-multiple categorical variables (either non-categorical variables or only binary categorical variables) which were found through simple linear regression to be fair to good predictors of red light running frequency. The resultant best subsets model, composed of the three variables: Number of Approach Lanes, Number of Crossing Lanes, and ADT of the approach, was found to have significant predictive abilities (R^2 = 0.624, p = 0.002). The three multiple categorical variables (Locale (Rural, Urban, Dense Urban), Grade, and Approach Direction) with fair to excellent predictive power (as shown via simple linear regression) were then added (one variable at a time) to the best subsets model. Of these three variables, only Approach Direction made a significant impact on the regression equation. The resultant multiple linear regression equation (composed of Number of Approach Lanes, Number of Crossing Lanes, Approach ADT, and Approach Direction) had an R^2 value of 0.805 and an associated p-value of 0.001.

Based upon these conclusions, it is apparent that various models exist that possess significant predictive ability relative to red light running violation frequency as a function of intersection traffic operational and geometric characteristics. These models may be applied by transportation/traffic professionals or law enforcement officials to estimate the potential level of red light running violations for a particular intersection. Application of these models to several intersections may allow jurisdictions to identify those with a greater potential for red light running. These intersections may then be targeted for remediation or increased red light running enforcement.

Future Work

As a product of this research, there are several potential opportunities for future work. These opportunities are discussed below:

- Additional intersections should be incorporated into the database compiled during the course of this research. Given the limited dataset available for this work, additional intersection/violation data would potentially add to statistical validity of the models, allow fine-tuning, and possibly reveal additional violation/intersection or traffic parameter relationships. Expansion of the dataset is envisioned as an iterative process where additional intersections would be added and new regression models created.

- The red light running violation prediction models created as part of this research may be aided by the inclusion of a signal type or signal timing element. Research by Bonneson et.al., Mohamedshah, and others has indicated that these elements may significantly affect the red light violation rate.
- Models should be tested and validated over a broad spectrum of intersection configurations and traffic operational conditions. Such testing would provide additional credibility for the models.
- Interaction within violation prediction models between variables that likely have a practical correlation (e.g., number of approach lanes and ADT) should be further examined.
- Effort should be made to incorporate a potential crash severity element into the red light running violation frequency prediction models. Inclusion of crash severity prediction potentially would allow for further identification of intersections subject to a high frequency of severe red light running crashes.

Section 6
References

Baguley, C.J. “Running the Red at Signals on High-Speed Roads.” *Traffic Engineering & Control.* July/August, Vol. 29, Issue 7/8, pp. 415-420. 1988.

Bonneson, J, K. Zimmerman, and M. Brewer. *Engineering Countermeasures to Reduce Red-Light Running*, Federal Highway Administration Report FHWA/TX-03/4027-2, College Station, Texas, 2003.

Federal Highway Administration and National Highway Traffic Safety Administration, “Guidance for Using Red Light Cameras”. URL:http://safety.fhwa.dot.gov/rlcguide/index.htm.

Hasson, P. “Red Lights Mean Stop”, *Public Roads*. September/October, Vol. 66, No. 2, pp. 16-19. 2002.

Hill, S., *An Overview of Automated Traffic Enforcement Programs: Red Light Running and Speed*, University Transportation Center for Alabama Report 00470-1, Tuscaloosa, AL, 2001.

Institute of Transportation Engineers. “Dick Armey Calls Attention to Intersection Safety and Traffic Congestion”. URL:http://www.ite.org/library/rlr.htm (18 Jan 2002)

Insurance Institute for Highway Safety. “Urban Crashes: Identifying Types Is 1st Step”, *Status Report*. Vol. 28, No. 2, pg. 4. February 6, 1993

Kamyab, A., T. McDonald, J. Stribiak, B. Storm, and M. Anderson-Wilk. *Red Light Running in Iowa: The Scope, Impact, and Possible Implications*, Center for Transportation Research and Education. Ames, IA, December 2000.

Kleinbaum D. and L. Kupper. *Applied Regression Analysis and Other Multivariate Methods*. Duxbury Press. North Scituate, MA, 1978.

Meadow, L.J., J.V. Moulden, R.A. Retting, D.M. Smith, and C.E. Walker. “Red-Light Camera Seminar”, Conference Proceedings of the Institute of Transportation Engineers International Conference, March, 1999.

Mohamedshah, Y., L. Chen, and F. Council. “Association of Selected Intersection Factors with Red Light Running Crashes”. Conference Proceedings of the Institute of Transportation Engineers Annual Meeting, August, 2000.

Montgomery, D. and G. Runger. *Applied Statistics and Probability for Engineers*. Wiley and Sons. New York, 1998.

Passetti, K.A. "Use of Automated Enforcement for Red Light Violations", Prepared for CVEN 677, Advanced Surface Transportation Systems, Department of Civil Engineering, Texas A&M University, August 1997.

Porter, B.E. and K.J. England, "Predicting Red-Light Running Behavior: A Traffic Safety Study in Three Urban Settings." *Journal of Safety Research,* Vol. 31, Issue 1, pp. 1-8. 2000.

Retting, R.A.and M.A. Greene. "Influence of Traffic Signal Timing on Red-Light Running and Potential Vehicle Conflicts at Urban Intersections." *Transportation Research Record*, Issue 1595, pp. 1-7. 1997.

Retting, R.A. A.F. Williams, and M. A. Greene. "Red-Light Running and Sensible Countermeasures: Summary of Research Findings". *Transportation Research Record*, Issue 1640, pp. 23-26.1998.

Retting, R.A. and A.F. Williams. "Characteristics of Red Light Violators: Results of a Field Investigation", *Journal of Safety Research*, Vol. 27, Issue 1, pp. 9-15. 1996.

SPSS Inc., *SPSS Regression Models*. Prentice Hall. Upper Saddle River, NJ, 2001.