SAFETY ANALYSIS OF CENTERLINE RUMBLE STRIPS ALONG RURAL TWO-LANE UNDIVIDED HIGHWAYS IN GEORGIA

A Thesis
Presented to
The Academic Faculty

by

Marisha Sonia Peña

In Partial Fulfillment
of the Requirements for the Degree
Master of Science in the
School of Civil and Environmental Engineering

Georgia Institute of Technology
December 2017

SAFETY ANALYSIS OF CENTERLINE RUMBLE STRIPS ALONG RURAL TWO-LANE UNDIVIDED HIGHWAYS IN GEORGIA

Approved by:

Dr. Michael Rodgers, Advisor
School of Civil and Environmental Engineering
Georgia Institute of Technology

Dr. Michael Hunter
School of Civil and Environmental Engineering
Georgia Institute of Technology

Dr. Angshuman Guin
School of Civil and Environmental Engineering
Georgia Institute of Technology

Date Approved: [December 8, 2017]

Pa' lante, pa' lante, pa' tras ni pa' coger impulso

ACKNOWLEDGEMENTS

First and foremost, I would like to thank my thesis committee advisors Dr. Michael Hunter, Dr. Angshuman Guin, and Dr. Michael Rodgers for allowing me the opportunity to participate in their research projects and for never giving up on me. Their continued sense of humor, patience, and guidance throughout my time as a graduate student were invaluable to my success, and I could not be more grateful to have had the opportunity to work alongside such mentors.

I would also like to express my sincere gratitude to the Georgia Department of Transportation for both funding this research project, as well as awarding me the Transportation Engineer of the Future fellowship. I'm infinitely grateful to GDOT for their support in my academic and professional endeavors.

I would also like to thank my colleagues in the GT Center for Traffic Operations and Safety lab: Abhilasha Saroj, Atiyya Shaw, John Bolen, Nishu Choudhary, Han Gyol Kim, Sung Jung Park and Somdot Roy, as well as my other graduate school colleagues for their technical, stylistic, and personal support. Above all, their friendship served as the utmost encouragement.

I would also like to acknowledge those who proofread my thesis for all of their valuable comments. Special gratitude goes to Claire Petry, my dear friend, cheerleader and confidant, for proofreading many of the texts I wrote throughout my entire academic career at Georgia Tech. And finally, to my family (Team Peña) and especially my niece, Naty, who continually inspire me to push forward in my academic and professional career.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

AADT	Annual Average Daily Traffic
AASHTO	American Association of State Highway and Transportation Officials
CLRS	Centerline Rumble Strips
CMF	Crash Modification Factor
CO	County
CR	County Route
DOT	Department of Transportation
E	East
EB	Empirical Bayes
FARS	Fatality Analysis Reporting System
FB	Full Bayes
FM	From
GIS	Geographical Information System
GDOT	Georgia Department of Transportation
IRB	Institutional Review Board
LOC	Locations
mph	Miles per hour
MP	Milepost
RCLINK	Roadway Characteristic Identification Number
SEV	Several
SPF	Safety Performance Function

SQL	Structured Query Language
SR	State Route
STARS	State Traffic and Report Statistics
TransPi	Transportation Project Information
VIN	Vehicle Identification Number
VMT	Vehicles-miles traveled
Vpd	Vehicles per day

SUMMARY

Vehicle crashes involving crossing over the roadway centerlines are among the most severe types of collisions nationwide. To address this issue, the Georgia Department of Transportation (GDOT) started implementing centerline rumble strips (CLRS) in rural locations across Georgia in 2005 and 2006. CLRS produce both an audible and tactile warning to alert drivers of impending lane departure into the lane of oncoming traffic. As of 2015, approximately 200 miles of CLRS have been installed by GDOT as a countermeasure for crossover crashes along rural two-lane undivided highways. This study evaluates the safety impacts of CLRS deployments in Georgia by analyzing two years of before and two years of after periods to evaluate the safety impacts associated with nine treatment sites and a control group of comparison sites with similar traffic and physical characteristics.

The study dataset consisted of 154 target crashes along 126.46 miles of CLRS treatment sites and 1,391 crashes along control group sites. The empirical Bayes method was used to develop a crash modification factor for CLRS of 0.66, indicating a 34% reduction in crashes involving centerline crossings associated with the installation of centerline rumble strips. The sample size of fatal and injury crashes was too small to obtain separate crash modification factors for fatal crashes and injury crashes. The favorable crash modification factor (0.66) found in this study supports wider use of centerline rumble strips as a safety measure to address crashes involving vehicles that cross the centerline of the roadway. In addition to the safety analysis, this study also provided insights into the crash reporting process by conducting a comprehensive manual review of more than 17,000 crash

reports. Approximately 6% of target crashes were found to be misclassified due to coding errors.

1 INTRODUCTION

In spite of an overall decline in motor vehicle traffic crashes throughout the U.S., crashes—and, most importantly, the injuries and fatalities associated with motor vehicle crashes—continue to be a great concern for federal, state, and local agencies. Furthermore, declines in crashes in rural areas have not been as pronounced as for urban areas. Half of all fatalities occur in rural areas, where only approximately 20% of the population live. Even more striking is that rural fatality rates are 2.4 times higher than in urban areas [1]. The state of Georgia has experienced a similar trend with an overall decline in the number of crashes but double the fatality rate in rural versus urban areas [2]. Consequently, rural road safety is of particular concern in the U.S. and has led transportation agencies to implement countermeasures which include various safety measures.

Along undivided rural highways, crossover crashes involve head-on, sideswipe and run-off road collisions. Rural highways tend to have higher speeds than urban roads of the same functional class. When a vehicle transfers onto the opposing lane of traffic, the severity is compounded by the additive nature of the vehicle speeds at the time of collision [3] [4]. Head-on collisions are among the most severe of crossover crashes due to the increased force sustained by both vehicles. During 2003 in Georgia, head-on and opposite-direction sideswipe crashes comprised only 2% of crashes but head-on collisions are responsible for 12% of all fatalities [3]. Various states have implemented countermeasures which have resulted in an average of a 21% reduction in head-on and opposite-direction sideswipes and a 25% reduction in injuries [3].

1.1 Background

In an effort to address roadway safety concerns throughout the state, the Georgia Department of Transportation (GDOT) developed and executed a comprehensive safety plan in 2005 [3] . The purpose of the GDOT Safety Action Plan was to identify and implement immediate engineering related safety efforts and develop longer range efforts to improve roadway safety in a cost effective manner that is acceptable to the public [3]. One such effort aimed at mitigating crossover crashes on rural two-lane highways was the installation of centerline rumble strips (CLRS), a low-cost countermeasure used by various states for mitigating crossover crashes on two-lane highways. CLRS consists of a series of grooved indentations or devices placed along the centerline that provide both auditory and tactile stimuli alerts for drivers that begin to cross over the centerline [3, 5]. In line with these efforts, this thesis seeks to analyze the safety effectiveness of CLRS in the context of driving and roadway characteristics within the state of Georgia.

GDOT has implemented centerline rumble strips on a variety of roadways in rural areas, with nearly 200 miles statewide installed between 2005 and 2006. Motivated by significant crash reductions in other states, GDOT launched installation of CLRS in stand-alone projects and as part of existing resurfacing projects, with costs averaging between $2,000 to $6,000 per centerline mile [3]. By employing this countermeasure and other engineering-related safety efforts, GDOT set out to achieve a statewide goal of reducing the total number of crashes by 2% annually as well as meet AASHTO's goal of a fatality rate of 1.0 per 100 million vehicle- miles-traveled [3].

1.2 Project Objectives and Justification

The objective of this research is to demonstrate and quantify the safety impacts of CLRS in Georgia by comparative analyses that reveal promising levels of safety effectiveness. Specifically, the project aims to evaluate whether there is any decrease in the number of crashes or any change in the type or severity of crashes after the installation of centerline rumble strips on highway facilities in Georgia.

In this study, the main analysis is performed by utilizing the empirical Bayes (EB) method to account for selection bias and derive a before and an after safety performance function (SPF) and the crash modification factor (CMF) specific to Georgia's rural two-lane undivided highways. Additionally, a naïve before-after analysis of CLRS segments and comparison roads supplements the EB to reveal cursory levels of safety effectiveness. The comparison reference group consists of roadways with similar characteristics to the study segments with centerline rumble strips to account for the effect of the CLRS on these roadways. Using a comparison group, the derivation of a SPF allows the EB method to compare more precise estimates of the after period and account for regression to the mean. The SPF also assists in accounting for time trends and traffic volume changes compared to the comparison group method [6]. The EB method is widely accepted among researchers and is recommended by the Highway Safety Manual (HSM) [7] to control for external factors of before-after evaluations of roadway safety treatments [7-10]. Additionally, the EB method derives a more accurate crash CMF, which is a measure of the safety effectiveness of a particular treatment or design element [7]. For these reasons, the EB method was selected for this thesis.

1.3 Thesis Organization

A literature review briefly discusses applications of CLRS, and presents an in-depth synopsis of the use of the EB approach to evaluate CLRS as a safety treatment. The methodology details the procedure in selecting CLRS study sites and developing the database for treatment and reference sites. Subsequently, this section describes the results from performing a before–after study using an empirical Bayesian analysis to evaluate the impact of a CLRS installation on those crash rates that CLRS are designed to mitigate. The results include the crash statistics derived from the developed CLRS crash database and the derivation of the CMF. As a final point, the conclusions and recommendations include discussion of the impact of the results on the current centerline rumble strips installed, potential application of this countermeasure in Georgia, along with limitations, and potential for future research.

2 LITERATURE REVIEW

As of 2015, approximately 200 miles of CLRS have been installed by GDOT in the state of Georgia as a countermeasure for crossover crashes along rural two-lane undivided highways [11]. Many factors can lead to the aforementioned crash types, the most common being inattentive or drowsy drivers, which account for 86% of fatal head-on crashes on two-lane highways [12]. When coupled with rural roadway conditions, including higher traffic speeds, lower rates of seatbelt use, and longer emergency-response times, safety countermeasures such as CLRS become increasingly attractive. Though CLRS may be constructed in several forms, the majority of installations are of the milled-in type, which is cost effective and can be readily implemented on existing roadways. Alternatively, CLRS can be constructed directly on the centerline, extended into the travel lane, or placed on either side of the centerline pavement markings. CLRS may have the added benefits of improving safety in low-visibility driving conditions, especially in areas with wintry weather or when roadway markings are obscured.

The following literature builds up the effort present in the first phase of the larger research effort conducted by J. Sim et. al. [4]. For a detailed discussion CLRS background, design and use, the reader is directed to [4]. This literature review will focus on the aspect critical to the following to the following analysis: observational before/after studies, which include naïve before and after, empirical Bayes, and full Bayes (FB).

2.1 Observational Before/After Studies

To evaluate the success of any roadway safety improvement program, it is essential to review the change in the number of motor vehicle crashes as well as injuries and fatalities.

These metrics help evaluate the effectiveness of any roadway safety improvement initiative. Safety evaluations offer information and insight for transportation planning and help determine if a particular countermeasure can be used at further sites [13]. At a minimum, a roadway safety project evaluation should include performance measures from both before and after the installation of a roadway treatment or other changes to the roadway [13]. Such a study of the effects of a roadway treatment should consider “what would have been the safety level” in the after period without treatment compared to the safety level with the treatment [10]. The effect of the treatment is represented by the difference in the number of injuries or the crash rate, over time, relative to the after period with and without the treatment [13].

The challenge in this type of comparison lies in estimating “what would have been” with no treatment. In estimating this after period without treatment, one uses the accident data from the before period. This before period may be the main reason that a treatment was implemented and thus, the prediction after period data is subject to a selection bias. Overall, “crash rates can vary significantly from year to year” [13] and, thus any estimates derived from this data are sensitive to selection bias. A roadway with an abnormally high number of crashes may be chosen for implementation of a safety measure. The crashes that occurred in the before period would not represent a suitable estimate to compare to the after period crash count [10]. The change in the crash rate may be due to other factors as well as the countermeasure [6]. The obvious selection bias would be attributed to the fact that the estimate would not reflect a normal or usual crash estimate for the roadway. This consequence presents an obvious example of a RTM bias.

RTM is inherent in crash data. According to the Highway Safety Manual (HSM) [7], RTM is “the tendency for the occurrences of crashes at a particular site to fluctuate up or down, over the long term, and to converge to a long-term average. This tendency introduces regression-to-the-mean bias into crash estimation and analysis, making treatments at sites with extremely high crash frequency appear to be more effective than they truly are" [7]. RTM produces periods that may have a comparatively high or low crash frequency. Attributing a decline in crash frequency to a roadway treatment may be misleading because the overall trend of crash frequency may have already been in decline unrelated to the treatment. A proper comparative analysis effectively accounts for the RTM bias.

Observational before/after studies consist of three methods: naïve before–after, empirical Bayes, and full Bayes. A naïve before–after analysis is based on the assumption that nothing changed in the after period except the treatment in question. Therefore, the before period crashes are used to predict what the after period crashes would have been without treatment [10]. The Bayesian methods (full and empirical) combine before period data with the after period data to develop the expected safety of a treatment. In these methods, the before period is derived from a group of similar sites (the expected crash frequency) and the existing site-specific crash data (the observed crash frequency) [8]. The after period distribution, likewise, involves the same general form as the before period distribution [8, 14]. Empirical Bayes and full Bayes are not different types of studies; they are simply two related approaches to combining prior and current information [8].

2.1.1 *Naïve Before–After*

In transportation safety, a naïve before–after study is one way (albeit not the most accurate way) to estimate the change between a parameter, such as crash frequency, during a before and after period. The before period value is used as the predictor for the after period with no treatment. The naïve before–after study assumes that the passing of time has no effect on the after period and that the expected after crash rate without treatment would be the same as in the before period. This predictor is then compared to the observed metric in the after period with the treatment.

However, the change in safety level can be attributed to several factors in addition to the roadway treatment, such as weather, traffic patterns, driver behavior, driver inclination to report crashes, and RTM. The passing of time is assumed to have no effect on the after period. All these other factors are assumed to be unchanged in a naïve before–after study, and any change in safety is assumed to be caused by the treatment only [10, 13]. This assumption in a naïve before–after study is weak and usually unrealistic. In addition, this method is strongly influenced by the selection bias discussed previously and does not allow researchers to separate out crash rate change attributable to the treatment from the other factors that have also changed over the study period. Any conclusion from this study lessens a researcher's ability to conclusively attribute the measured difference to the treatment of interest. This approach is, therefore, generally not recommended for safety studies [10].

2.1.2 *Full Bayes*

Full Bayesian uses before data to predict future crashes at a treatment site had the treatment not been implemented. However, instead of a single-point estimate of the expected mean and its variance, it predicts a distribution of likely values. The estimate for expected crash frequency in the after period is determined by combining the distribution of likely values with the crash frequency of the specific study sites. The use of a distribution of likely values generally improves the overall estimate of likely crash rates [8]. As the researchers did not use the FB approach in this study, a detailed description is not provided; however, the next sections discuss the reasoning underlying the selection of empirical Bayes over full Bayes.

2.1.3 *Empirical Bayes*

The empirical Bayes method is a simplified version of the full Bayesian method and is well established and accepted by transportation professionals and researchers for roadway safety comparative studies [8, 14]. Through an EB study, the safety effectiveness of a treatment can be based on the model rather than the raw crash data [15]. According to the HSM, the empirical Bayes approach combines "observed crash frequency data for a given site with predicted crash frequency data from many similar sites to estimate its expected crash frequency" [7]. The before period data come from the evaluation sites and a reference group with similar roadway attributes to develop a calibrated safety performance function (SPF) [7, 8, 10]. The consequences of RTM bias and changes in traffic characteristics are explicitly accounted for in the SPF [16]. The safety performance function is an equation that represents the relationship between the expected number of target crashes and the roadway characteristics [8, 16]. The expected crash frequency estimates are combined with "the site-specific crash count to obtain an improved estimate of a site's long-term expected crash frequency" [8]. The EB approach uses the assumption that crashes follow a negative binomial (NB) distribution, and it employs the NB dispersion parameter in the estimation process [8]. Section 3.3 of this report provides a detailed walk-through of the EB method.

2.1.4 *Comparison of EB and FB*

Both EB and FB methods recognize that deriving estimates from just a few years of information from specific sites provides unreliable estimates. To remedy this, central to Bayesian analysis, comparison-group data for the same study period are used to complement the treatment site's data. This addition allows the analysis to formulate more robust estimates and account for RTM bias and traffic volume changes due to various factors, such as general trends, changes in crash reporting, weather, driver behavior, etc. [6, 14].

While both empirical and full Bayesian approaches are suitable and effective methods for conducting a comparative analysis for traffic safety studies, their differences and comparative advantages render them most efficient in different scenarios of study [8, 14]. The FB approach is more complex than the EB approach, and some researchers believe that it more suitably accounts for uncertainty within crash data [8]. The EB approach simplifies the FB approach by assuming the study sites and the comparison sites have similar covariables, such as roadway geometry and signal control. These covariables are accounted for through the SPF derived in the EB method [14]. Furthermore, the FB approach requires substantially fewer data than the EB approach for the untreated control group sites. The FB approach "provides more detailed causal inferences" [8], "more flexibility in selecting crash count distributions" [8], and it does not require the development of safety performance functions to obtain the predicted number of crashes.

However, the FB method is more cumbersome than the EB method because a high level of statistical knowledge is required to carry out the complexity of the full Bayes method. Additionally, it has been more difficult to develop statistical software for an FB application than for an EB application [6]. Finally, research has shown that the EB approach produces similar results to the FB method and reliable analysis when an adequate number of sites exists [6, 8]. Thus, the research team in this effort chose EB for the safety analysis.

2.2 Evaluation of Safety Treatment with an Empirical Bayes Approach

In comparison to the naïve approach, the EB methodology enables a more precise estimation of the number of crashes that would have occurred at an individual treatment site in the after period had the treatment not been implemented [17]. The EB method has been used in various roadway safety analyses [8, 9, 16, 18, 19]. These studies include diverse locations throughout the United States and Canada and varied roadway treatments, such as road diets, conversions of intersections to roundabouts, SRS, and CLRS [5, 8, 9, 16, 18-20].

For example, Persaud et al. [16] conducted a before–after study using the EB procedure for the conversion of intersections to roundabouts. Their study estimated highly significant reductions of 40% for all crash severities and 80% for all injury crashes. Specifically, the crashes with fatalities and incapacitating injuries were reduced by 90%. In a later study, Persaud et al. [8] used the EB and FB methods to examine the safety impacts of a road diet, which involved the conversion of four-lane roadways into three-lane roadways with a two-way left-turn lane in the middle. That study determined the estimated safety effects from both methods to be comparable.

Directly relevant to the current study, the effectiveness of CLRS has also been studied in various locations using EB. One study looked at the effects of 98 treatment sites of CLRS along approximately 210 miles of rural, two-lane roadways in seven states: California, Colorado, Delaware, Maryland, Minnesota, Oregon and Washington [5]. “The results of this research demonstrate that centerline rumble strips are an effective countermeasure on rural two-lane roads” [5]. The study observed positive effects of CLRS on various roadway alignments and geometric configurations, e.g. curved sections, tangents, grades, and level terrain. This analysis revealed that head-on and sideswipe accidents from the opposing direction experienced the most significant reduction, decreasing by 25%. In general, all crash types were reduced by 12%. From this analysis, the authors recommended a wider application of CLRS on rural, two-lane roads [5].

Another study applied the EB method to evaluate the effectiveness of CLRS on 47 segments of undivided, rural, two-lane arterials and divided, rural four-lane freeways in British Columbia [19]. A comparison group with similar roadway characteristics was selected to perform the EB analysis. The authors found that overall, SRS and CLRS can significantly reduce severe collisions and specific collision types. The use of CLRS and SRS demonstrated a reduction of 18.0% of injuries. Individually, SRS reduced off-road right collisions by a statistically significant 22.5%, and CLRS showed a statistically significant reduction of 29.3% in off-road left and head-on collisions. Specifically, installing both CLRS and SRS on undivided, rural, two-lane arterials reduced off-road right, off-road left, and head-on collisions combined by 21.4%. The authors concluded that rumble strips, whether just SRS or CLRS, or the combination of SRS and CLRS, are very effective safety measures to reduce the severity of crashes [19].

In Kansas, a study to assess the safety effectiveness of CLRS was carried out on over 350 miles of highways. The CLRS analyzed in this studied consisted of two patterns of CLRS (rectangular and football). Both naïve and EB before–after comparative analysis were conducted and provided statistically similar results. CLRS reduced total correctable crashes by 29.21%. Crashes with injuries of all severities and fatalities were reduced by 34.05%. Additionally, cross-crashes were reduced by 67.19%. Crashes involving run-off-the-road were reduced by 19.19%. All of the comparisons, except run-off-the-road crashes were found to be statistically significant. Ultimately, the study found that no statistical difference existed between football-shaped and rectangular-shaped CLRS [20].

A recent study in Michigan assessed the safety impacts of a statewide CLRS implementation program, either stand-alone CLRS or in conjunction with SRS, carried out between 2008 and 2010. In this instance the EB method was used to assess the effectiveness of more than 4200 miles of CLRS installed along two-lane highways [9]. On CLRS-only sites, CLRS reduced overall crashes of both target and non-target crashes by 15.8% and by 17.2% at locations with both CLRS and SRS treatments. CLRS were found to reduce target crossover collisions by 27.3% when used alone and by 32.8% when used in conjunction with SRS. The study also found that rumble strips provided the additional benefit of decreasing the number of crashes under adverse pavement conditions and assisted in maintaining drivers in their respective lanes [9].

3 METHODOLOGY

This section details the selection of candidate treatment sites, the creation of treatment and reference crash databases and the development of empirical Bayes safety analysis. The site selection efforts were completed by J. Sin and a detail description may be found in [4]. A summary of those efforts are presented in section 3.1 Site Selection to aid interpretation of findings.

3.1 Site Selection

Sin et. al. [4] state that in the Safety Action Plan, GDOT set out to develop 100 miles of CLRS in the FY 2005. Using 2000-2003 crash data from the Accident Information System (AIS) database, locations with higher frequencies of centerline crossover crashes were identified as potential sites. The Office of Safety & Design, in coordination with the Office of Maintenance, scheduled projects for CLRS installation projects in both stand-alone application and in conjunction with resurfacing projects. Between the fall of 2005 and spring of 2006, GDOT carried out several CLRS installation projects. “By 2006, there were nearly 200 miles of centerline rumble strips installed, primarily on rural two-lane, two-way roadways” [4].

3.1.1 *Georgia Project Database*

Next, Sin et. al. [4] selected the CLRS sites for this study from GDOT’s Transportation Project Information (TransPI) website in 2013. TransPI, now known as GeoPI [11], is a web-based database from which the public can access any related data or documentation for GDOT projects. Project managers and engineers submit project information, including documentation, financial information, and Geographical Information System views, into the TransPI/GeoPI system. That information is accessible to users both within and outside of GDOT [4, 11].

For this study, an initial query by Sin et. al. [4] for projects involving the installation of CLRS resulted in the eight projects listed in Table 1. Although there were only eight projects, several of those involved more than one installation site, such as the project on SR 36 that had segments from SR 74 to SR 7 and also from SR 7 to I-75. After examining the 8 projects, Sin et. al. compiled at least 11 potential CLRS sites, which are listed in Table 2 with their corresponding beginning and ending descriptions [4].

Table 1. Results Obtained from TransPI [4]

Project ID	Project Accounting No.	Project Title	Counties
0006080	—	SR 25 SPUR EAST FM CR 583/SEA ISLAND DR TO E OF SR 25/US 17	Glynn
0006693	CSSTP-0006-00(693)	SR 14|SR 16|SR 154@SEV LOC IN CARROLL &COWETA [CENTERLINE]	Carroll, Coweta
0006945	CSSTP-0006-00(945)	SR 369 FM CHEROKEE CO TO HALL CO – CENTERLINE RUMBLE STRIPS	Forsyth
0006975	CSSTP-0006-00(975)	SR 42@SEV LOC IN HENRY|BUTTS|MONROE – CENTERLINE RUMBLE STRIPS	Butts, Henry, Monroe
0006976	CSSTP-0006-00(976)	SR 204 FM BRYAN COUNTY LINE TO I-95 – CENTERLINE RUMBLE STRIPS	Chatham
0007077	CSSTP-0007-00(077)	SR 36 FM SR 74 TO SR 7 & SR 36 FM SR 7 TO I-75	Butts, Lamar, Upson
0007079	CSSTP-0007-00(079)	SR 136 FROM SR 61/US 411 TO DAWSON COUNTY LINE	Gilmer, Gordon, Murray, Pickens
0007080	CSSTP-0007-00(080)	SR 26 FM E OF BULL RIVER BRIDGE TO TYBEE ISLAND CITY LIMITS	Chatham

Table 2. TransPI Location Description by Installation Site [4]

Project ID	Centerline Rumble Strips Installation Site	Beginning Description	Ending Description
0006080	State Route 25 Spur	Sea Island Drive/ CR 583	State Route 25/US 17
0006693	State Route 14	Herring Road/CR 43	Johnston Circle/CR 7
0006693	State Route 16	Carrolton Bypass	Newnan Bypass
0006693	State Route 154	State Route 54	I-85
0006945	State Route 369	Forsyth County	Forsyth County
0006975	State Route 42	Several Locations in Henry, Butts, and Monroe Counties	Several Locations in Henry, Butts, and Monroe Counties
0006976	State Route 204	Bryan County Line	I-95
0007077	State Route 36	East Main Street	Peach Blossom Trail
0007077	State Route 36	Highway 41	I-75
0007079	State Route 136	State Route 61/US 411	Dawson County Line
0007080	State Route 26	East of Bull River Bridge	Tybee Island City Limits

3.1.2 *Additional Sources for Authentication of Sites*

Sin et. al. [4] noted this query for “centerline rumble strips” in TransPi yielded various entries for a single project; consequently, each entry’s project information required examination. Several project descriptions revealed that some projects consisted of multiple installation sites, and thus, provided conflicting information. Sin et. al. [4] used other sources to authenticate the discrepancies and confirm the details of each study site. To confirm that these roadways did have CLRS, they used Google Maps Street View® to verify its existence at the sites returned by the TransPI query, as shown in Figure 1 for Project ID 0007077. Project Preconstruction Status Reports and Project Plan Sheets were requested from GDOT to define additional project information, “including the total mileage and various dates associated with the project such as the Management Let Date and the Project Completion Date” [4]. An example of a Project Preconstruction Status Report and information taken from the Project Plan Sheets for Project ID 0007077 are showed below in Figure 2 and Table 3 [4].

Figure 1. Google Street View® Roadway Verification [4, 21]

PRECONSTRUCTION STATUS REPORT

PROJ ID	COUNTY	DESCRIPTION
0006080	Glynn	**SR 25 SPUR EAST FM CR 583/SEA ISLAND DR TO E OF SR 25/US 17** This project consists of installing ground-in-place centerline indentation rumble strips centered about the centerline traffic stripe of State Route 25 Spur East from Sea Island Drive to east of State Route 25/US 17 in Glynn County (MP 6.15 to MP 10.53).

PROJ NO.:	CSSTP-0006-00(080)	**SPONSOR :**	GDOT
MPO TIP #:	BATS 06-01	**PROJ MGR:**	Cameron, Derrick
MPO:	Brunswick	**DOT DIST:**	5
PROJ LENGTH (MI)	3.93	**CONG. DIST:**	1
PROGRAM TYPE:	Safety	**House Dist :**	
TYPE WORK:	Rumble Strips	**Senate Dist :**	
LET RESPONSIBILITY:	GDOT Let		
BIKE PROVISIONS INCLUDED?	N		

Phase	*Fiscal Year Approved*	*Approved FY Estimate*	*Fund*	*Phase Status*
Engineering	2006	$ 5,000.00	Q21	AUTHORIZED

**Inflation Included in Estimate*

ACTIVITY	*ACTUAL START*	*ACTUAL FINISH*	*PERCENT COMPLETE*
Environmental Approval	**1/5/2006**	**1/5/2006**	100
PFPR Inspection	1/31/2006	1/31/2006	100
FFPR Inspection			0

Right of Way Acquisition Information

Preliminary Parcel Count: **Total Parcel Count:** **Acquired by:** N/R

Figure 2. Preconstruction Status Report for Project ID 0007077 [4]

Table 3. Project Plan Sheet Information for Project ID 0007077 [4]

Attribute	Description
Project Number	CSSTP-0007-00(077)
Project ID	0007077
Net Length	29.77
Starting Milepost	MP 8.12
Ending Milepost	MP 0.49
Starting County	Upson County
Ending County	Butts County

To determine the exact locations of CLRS along the roadways, the researchers examined maps from the Project Plan Sheets to verify the beginning and ending mileposts of some of the installation sites. However, they discovered that the maps in the Project Plan Sheets did not always match the descriptions found in the projects from the TransPI query. For example, the map in the Project Plan Sheets for Project ID 0007077, as seen in Figure 3, only showed one segment of CLRS installations, although, the project actually had two sections. The two segments of CLRS were detailed in the project documents found in TransPI and verified in Google Maps Street View®. The Project Plan Sheets also listed a Detailed Quantities Estimate, which had values that were used to verify the existence of CLRS in the projects. While information from various sources was not always accurate, it served as reference for determining the correct locations of the CLRS [4].

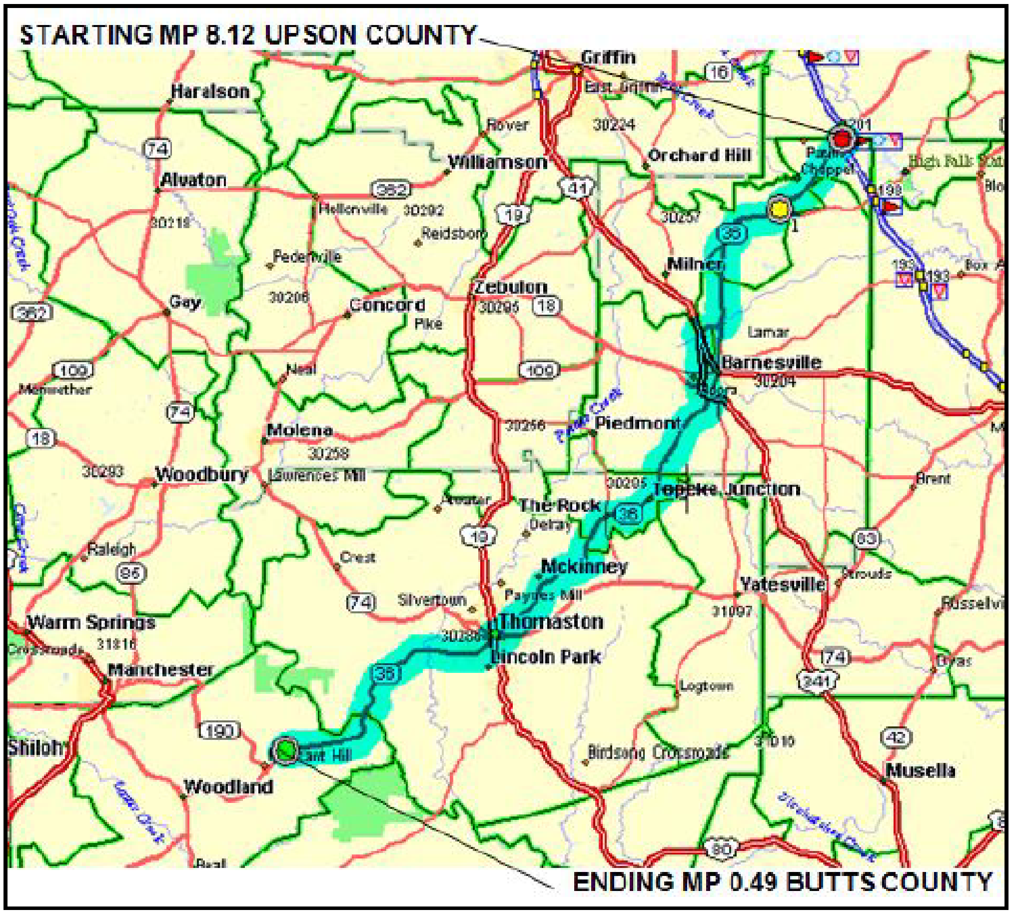


Figure 3. Map of Project ID 0007077 Location from Project Plan Sheet [4]

3.1.3 *Final Study Sites*

After careful examination, Sin et. al. [4] selected 10 CLRS installation sites. From the original query results listed beforehand in Table 1, Project ID 0007080 and Project ID 0006080 were listed as “cancelled” under the Project Completion category and eliminated from the list of potential sites. Project ID 0006975, SR 42, was considered as two separate sections for this analysis. Once the CLRS segments were determined, the last step by Sin et. al. [4] was to verify the exact locations of the start and end mileposts of the CLRS. Initially, the milepost information for each CLRS installation site was extracted from the Project Plan Sheets. However, after careful revision, most of the mileposts did not correspond with the mileposts noted in TransPI. To rectify these inaccuracies, the mileposts from the Project Plan Sheets were plotted in Google Earth® and verified using Google Street View®. Once, the mileposts were confirmed, 126.46 miles along 10 routes was identified as the treatment sites [4].

As part of the current effort, problems encountered during the automated association of crashes to treatment during the data reduction process led a to 12.95-mile decrease in the total miles of treatment sites as given by Sin et. al. [4]. The final treatment database consisted of 113.51 miles in 9 study sites, see Table 4, and a map of their locations is presented in Figure 4. Table 5 shows the segments included in the preliminary sites. The segments that were not included in the final analysis are grayed out.

Table 4. Final List of CLRS Study Sites

Project ID	**Description**
0006693	SR 14
0006693	SR 16
0006945	SR 369
0006975	SR 42 Section A
0006975	SR 42 Section B
0006976	SR 204
0007077	SR 36 Section A
0007077	SR 36 Section B
0007079	SR 136

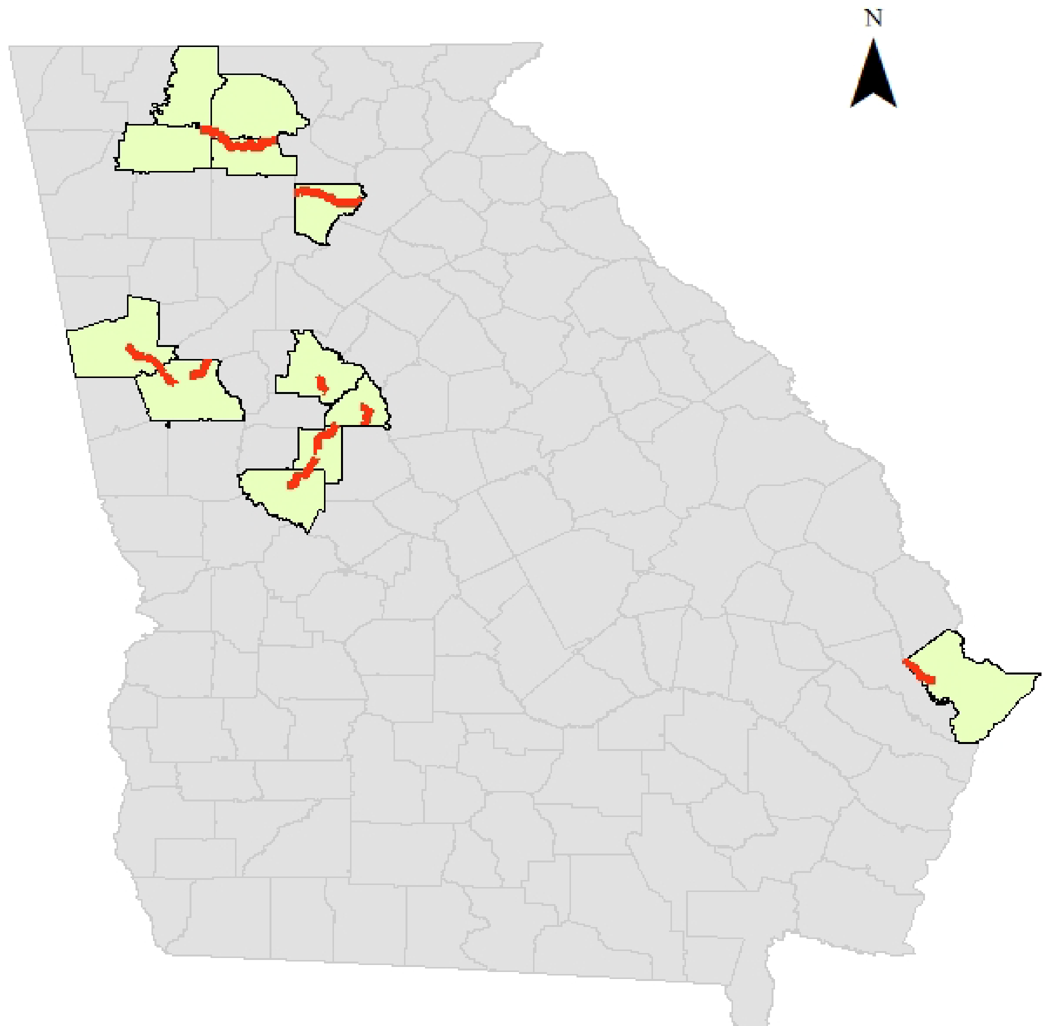


Figure 4. Locations of CLRS Sites [22-24]

Table 5. CLRS Start and End Mileposts for Study Sites

Project ID/Roadway Description	County	Mileposts		Segment Length (mi.)	Total Study Site Length (mi.)
		Begin	End		
0006693/SR 14	Coweta	19.68	19.74	0.06	7.87
		19.74	23.17	3.43	
		23.17	26.72	3.55	
		26.72	27.55	0.83	
0006693/SR 16	Carroll	16.69	17.64	0.95	16.56 *(18.24)*
		17.64	22.65	5.01	
		22.65	26.19	3.54	
		26.19	27.87	1.68	
	Coweta	0.00	3.86	3.86	
		3.86	6.33	2.47	
		6.33	6.98	0.65	
		6.98	7.06	0.08	
0006693/SR 154	Coweta	0.11	0.56	0.45	*0 (7.49)*
		0.56	3.34	2.78	
		3.34	5.31	1.97	
		5.31	7.60	2.29	
0006945/SR 369	Forsyth	0.00	2.71	2.71	19.89
		2.71	5.80	3.09	
		5.80	6.43	0.63	
		6.43	10.07	3.64	
		10.07	11.08	1.01	
		11.08	11.86	0.78	
		11.86	12.82	0.96	
		12.82	19.89	7.07	
0006975/SR 42 A	Butts	0.00	3.18	3.18	7.68 *(7.97)*
		3.18	4.81	1.63	
		4.81	7.44	2.63	
		7.44	7.68	0.24	
		7.68	7.97	0.29	
0006975/SR 42 B	Henry	4.58	8.53	3.95	5.23
		8.53	9.81	1.28	
0006976/SR 204	Chatham	0.00	0.64	0.64	8.14
		0.64	8.14	7.50	

Project ID/Roadway Description	County	Mileposts		Segment Length (mi.)	Total Study Site Length (mi.)
		Begin	**End**		
0007077/SR 36 A	Upson	9.34	11.06	1.72	8.05 *(13.87)*
		11.06	15.72	4.66	
		15.72	19.11	3.39	
	Lamar	0.00	1.93	1.93	
		1.93	4.10	2.17	
0007077/SR 36 B	Lamar	7.21	13.51	6.30	11.84
		13.51	16.83	3.32	
		16.83	18.60	1.77	
		18.60	19.05	0.45	
0007079/SR 136	Gordon	23.56	24.07	0.51	28.25
	Murray	0.00	2.82	2.82	
	Gilmer	0.00	5.21	5.21	
	Pickens	0.00	3.67	3.67	
		3.67	6.32	2.65	
		6.32	7.25	0.93	
		7.25	12.01	4.76	
		12.01	14.14	2.13	
		14.14	17.96	3.82	
		17.96	19.71	1.75	
				Total	**113.51** *(126.46)*

3.1.4 *Analysis Period*

Sin et. al. [4] also noted that the federal and TransPI reports had conflicting start and stop construction dates for each project. To clarify discrepancies, the construction completion dates were confirmed by GDOT engineers to be the Time Charges Stop Date from federal construction reports. The confirmed start and stop dates are listed below in Table 6 [4].

Table 6. Begin and End Dates for CLRS Construction [4]

CLRS Site	Start Date	Stop Date
SR 14	10/11/2005	10/31/2005
SR 16	10/11/2005	10/31/2005
SR 154	10/11/2005	10/31/2005
SR 369	03/06/2006	03/26/2006
SR 42 A	01/17/2006	05/31/2006
SR 42 B	01/17/2006	05/31/2006
SR 204	02/14/2006	02/28/2006
SR 36 A	01/17/2006	05/31/2006
SR 36 B	01/17/2006	05/31/2006
SR 136	01/17/2006	05/31/2006

Initially, the study periods were identified as two whole calendar years before (2003–2004) and two whole calendar years after (2007–2008) the construction of the CLRS sites. This study period would provide time to compensate for changes in driving patterns due to the unfamiliarity of the new roadway features (CLRS) or the presence of construction equipment and changes in roadways, such as closures or detours [4]. However, for this study, police records corresponding to crash data along the CLRS sites and the control sites were available for January 1, 2003, only until May 31, 2008. This limited the use of data from the full 2008 calendar year. To maintain an analysis period, which accounts for seasonal changes consistent with the before and after periods, the final dates for the comparative analysis were determined to be:

- Before period: June 1, 2003, to May 31, 2005
- After period: June 1, 2006, to May 31, 2008

3.2 Crash Database

Investigating officers provide police reports documenting crash data involving motor vehicles, bicycles, and pedestrians. In Georgia, police agencies record motor vehicle crashes with the standardized Georgia Uniform Motor Vehicle Accident Report. GDOT and/or GDOT contractors (Police Report Archive) archive the images of these police reports. Additionally, GDOT and/or its contractors/collaborators extract data from the reports and retain the information in the GDOT Crash Database, a searchable database format to facilitate retrieval of important information for research and other purposes.

While both the Police Report Archive and GDOT Crash Database are public records, they contain sensitive, personally identifiable information that must be protected from inadvertent release. Researchers are granted access to both databases courtesy of GDOT based on approved data protection protocols. The protocols used in this study were originally developed by the Georgia Transportation Institute (GTI). To protect the anonymity of persons involved in crash reports, GTI requires that its research projects use sanitized versions of the databases with all sensitive, personally identifiable information redacted. When the research cannot be conducted using the sanitized databases, it requires approval from the Georgia Tech Institutional Review Board (IRB). To standardize the data management process, GTI has developed protocols to describe how it acquires and subsequently handles, stores, and sanitizes data. These protocols are included in Appendix B [25].

3.2.1 *Data Acquisition, Management, and Storage*

For this study, the researchers downloaded new and updated versions of the Police Report Archive and the GDOT Crash Database via a Secure FTP server, owned and operated by GDOT or their contractor. The researchers were authorized by GDOT to access these databases and download them onto an encrypted hard drive located in a protected location by an operator authorized by GDOT for database access. Downloaded files were immediately copied and verified to two encrypted external hard drives. After being copied, the original files located on the encrypted computer were eliminated using Eraser (http://eraser.heidi.ie/), a software that securely shreds files. The copies located on the encrypted external hard drives were transferred to a safe in a limited access area behind locked doors. Access to the safe was limited to GDOT authorized database users.

3.2.2 *Treatment and Reference Crash Databases*

The crash database were divided into two databases:

- Treatment crashes – crashes occurring along study sites
- Reference crashes – crashes occurring along a control set of roadways with similar characteristics to those of the CLRS sites

The treatment (CLRS-associated) crash data was selected using a database query to filter crashes by the road characteristics and mileposts of the CLRS sites and study period dates. The crash database was associated to the road characteristic data by referencing the Roadway Characteristic Identification Number (RCLinks) ID, a 10-digit identification number that references the county the road is located in, the route, route number, and a two digit suffix, [4]. See Appendix C for query SQL code.

The reference database consisted of 17,381 crashes or 25% of all crashes along rural, undivided two-lane highways in Georgia from 2003 to 2008. Given the verification requirements for the crashes using the police records, a 25% random sampling was used to reduce the reference crash data. After sampling, 17,381 crashes were identified. Physical road characteristics were used to filter out irrelevant cases, as shown below in Table 7. Additionally, the reference crash database excluded the crashes that were selected for the treatment database [4]. The final reference set consisted of 11,706 crashes.

Table 7. Filters Used to Create the Comparison Crashes Database [4]

Variable	**Filter**
Accident Date	Same dates as the crashes in the Treatment Crashes database
Intersecting Road Type	Null
Dividing Highway Barrier Type	0 – No Barrier
Dividing Highway Median Type	0 – Undivided Road
Functional Classification	2 – Rural – Principal Arterial 6 – Rural – Minor Arterial 7 – Rural – Major Collector
Number of Left Lanes	1 – 1 lane on the left side of the roadway
Number of Right Lanes	1 – 1 lane on the right side of the roadway

3.2.3 *Crash Database Sanitation*

The filtered control and treatment databases still contain personally identifiable information and, therefore, require specialized techniques for their removal. The first phase of the sanitation process involves the use of a PERL script to remove initial sensitive information (see Appendix D for full script). Then a GDOT-authorized database user manually reviews the new version of the crash report. These reports are in standard two-page format with supplemental pages provided on some reports (e.g. when multiple vehicles are involved or injuries have occurred). The first page of the report always contains certain personally identifiable information, which is not pertinent to this research effort. Personally identifiable information may, or may not, be present in subsequent pages. Given the non-uniformity of the scanned reports, full automation of the sanitization process is challenging.

The PERL script accomplishes several sanitation tasks [26]. First, all Police Report image files are renamed to replace the crash ID with a unique ID used in the sanitized database. Once this step is completed, the table containing the link between the crash ID and the unique ID is securely destroyed. Hence, this unique ID has no link and cannot be traced by the original crash ID. Next, each Police Report image file is converted to a series of images, each image representing one page. Each image is identified with the unique ID that allows it to be linked back to the sanitized database where other non-personally identifiable information related to the crash is available. Since the information on the first page of each report is not needed, it is then deleted as it contains personally identifiable information. Subsequently, the second page image is verified to ensure proper orientation and inverted if necessary. Portions of the second page, where personally identifiable information is present, are then electronically blanked out. If the original record had only two pages, then the record should be mostly sanitized. It just has to be verified for any extraordinary personally identifiable information. If the record has more than two pages remaining page images are manually checked by a GDOT-authorized database user to identify any personally identifiable information. Any images containing personally identifiable information that are not relevant for research are deleted.

Afterwards, the resulting Police Report still contained some sensitive, personally-identifying information due to the method in which the Police Reports are redacted. Each police officer records the details of the crash. Often times, the officer will include information such as the names of those involved, contact information, Vehicle Identification Numbers

(VINs) and other personally identifiable pertaining to the individuals engage in the crash. A GDOT-authorized database user completed the final sanitation process. For research purposes, any application requiring access to personally identifiable information was conducted under a research protocol approved by the Georgia Tech Institutional Review Board (IRB). Any sensitive information, such as VINs or driver names and contact information, was manually removed with the software XnView, a multi-format graphics viewer with image processing capabilities. Removed data cannot be restored after it has been saved, thus ensuring the privacy of the individuals involved in the Crash Reports. Finally, the sanitized versions are made available to students and other researchers as necessary for analysis in normal research applications. Any application requiring access to personally identifiable information will be covered by a separate protocol approved by the Georgia Tech IRB.

3.2.4 *Crash Database Verification*

The research team verified locations of each crash by comparing the results obtained from the crash database with the corresponding (sanitized) police report. The details of the sanitization process can be found in Appendix B. Because each report was recorded by the investigating officer present at the time of the crash, entries are subject to human error. The researchers assumed that the rate of errors was consistent throughout each study year. Each milepost reported in the police report was considered correct and used to verify if located within the treatment site.

Undergraduate research assistants were charged with verification of the crash database to identify target crashes needed for this research. In this case, target crashes are influenced by the presence of centerline rumble strips. CLRS are intended to prevent specific target crashes: those caused by vehicle centerline crossovers. This research identified target crashes as head-on collisions, opposite-direction sideswipe collisions, or collisions not with motor vehicles [4, 27].

Additionally, the first moments of these collision events likely involve crossing over the centerline due to inattentiveness, distraction, fatigue, or other conditions that are not intentional on the driver's part [4]. Therefore, target crashes exclude several centerline crossovers that occur due to other circumstances, such as vehicles that originally run off the right shoulder of the road and overcorrect and cross over the centerline, or crashes that occurred on locations that did not meet the two-lane, undivided requirement (e.g., at intersections or on three-lane or wider highways). A list of detailed exceptions is provided in Figure 5. Approximately, one-fifth (18.9%) of all target crashes (292 target crashes) experienced some form of hydroplaning. Environmental conditions, such as water, rain, ice, and snow, and spilled fuel on the roadways that caused hydroplaning were considered as target crashes.

- Location outside of study scope
    - Intersections
    - On three-lane or wider highways
    - With separation or barriers between opposite directions of lanes
        - Two-way left-turn lanes
        - Raised medians
        - Turning lanes
- Overcorrection—vehicle first runs off to right
- Passing maneuvers
- Environmental/external factors
    - Animals on roadway, e.g. deer, dogs, etc.
    - Medical reasons
- Vehicular malfunction
    - Faulty brakes, steering failure
    - Steering wheel or tire falling off vehicle
    - Trailer or cargo being pulled falls off vehicle
    - Trailer swaying cause vehicle to lose control

Figure 5. Target Crash Exceptions

Additionally, AADT and segments lengths for each crash analyzed in this study were identified from GDOT public data for 2003 to 2008. This data is publically available on the GDOT website. Annual average daily traffic (AADT) is the counted or estimated total traffic volume in one year divided by 365 days/year [7]. Due to the large number of crashes in the study set, the researchers created a MatLab® code to extract this specific data for all target and non-target crashes by referencing the RCLink and location of crash as specified in the crash report. This code can be found in the Appendix E. Discrepancy found within the resulting tables were manually check by student researchers.

3.3 Empirical Bayes Method/Development of SPF

The Empirical Bayes method was used to evaluate the effectiveness of the CLRS in preventing crossover collisions. The safety performance functions (SPF) for both before and after periods were derived to predict the number of expected crashes in the after period without the installation of CLRS. These estimates were based upon all pre-installation (before–period) crash data for the entire population of study segments, both treatment and reference sites. The basic input for this evaluation includes the number of collisions that occur on the study sites, and their respective AADT values and segment length. According the Highway Safety Manual [7], the SPF for predicted average crash frequency along rural two-lane, two-way roadway segments is calculated by Equation 1.

$$N_{spf\ rs} = e^{\alpha} AADT^{\beta} L \times 365 \times 10^{-6} \quad (1)$$

Where:

$N_{spf\ rs}$ = predicted total crash frequency for roadway segment base conditions;

$AADT$ = average annual daily traffic volume (vehicles per day); and

L = length of roadway segment (miles)

3.3.1 *Before Period SPF Parameters*

The observed parameters for the SPF equation above were determined by a multistep process.

STEP 1: Select the before period SPF

Based on the before period, the predicted number of crashes is found with the SPF, using Equation 2:

$$N_{predicted} = e^{(\alpha_0)} AADT_{before}{}^{\beta_0} L_{before} \times 365 \times 10^{-6} \quad (2)$$

Where:

α_0 = the relationship between crash frequency and roadway characteristic of rural, two-lane, undivided highways in the before period; and

β_0 = the relationship between crash frequency and AADT in the before period

The total number of collisions used in the SPF are derived from the treatment sites and control sites. This total number of collisions is affected by the AADT and other roadway characteristics (including the segment length). The effect of the AADT is evaluated by determining the β coefficient.

STEP 2: Determine the β coefficient for the before period SPF

The specific values for the coefficients α and β are needed to complete the SPF. The SPF in Equation 2 is modified to include vehicle miles traveled (VMT) embedded within the AADT as shown in Equation 3. As per the U.S. Department of Transportation definition, vehicle miles traveled is the measurement of the total miles traveled by vehicles within a specific time-period [28].

VMT for two years was calculated as follows:

$$VMT = AADT \times L \times 730 \times 10^{-6} \quad (3)$$

VMT can be inserted into the before prediction SPF, Equation 2, and simplified as shown below:

$$N_{predicted} = e^{\alpha} AADT^{\beta-1} \times (VMT)$$

$$\frac{N_{predicted}}{VMT} = e^{\alpha} AADT^{\beta-1}$$

$$ln\left(\frac{N_{predicted}}{VMT}\right) = ln(e^{\alpha} AADT^{\beta-1})$$

$$ln\left(\frac{N_{predicted}}{VMT}\right) = (\beta - 1)\ln(AADT) + \alpha \quad (4)$$

VMT is calculated by multiplying the amount of daily traffic on a roadway segment by the length of the segment. The relationship demonstrated in Equation 4, between the observed crashes during the before period and their respective AADT and segment lengths, is fitted to determine the appropriate β coefficient.

STEP 3: Determine the α coefficient of the before period SPF

The α coefficient is determined by accounting for all roadways in the before–after set. Also, the AADT for each section must be corrected to account for all segments in the before period, even those with no crashes. The corrected AADT must be adjusted by a ratio of the standard AADT rate to the treatment AADT, referencing a weighted average.

Using the original SPF equation, the sums of the roadway segments and number of crashes should be included as follows:

$$\frac{\sum N_{observed\ before\ crashes,i}}{Total\ VMT\ for\ all\ roadways} = e^{(\alpha_0)}\overline{corrected\ AADT_{\iota}^{\beta}} \quad (5)$$

$$\overline{corrected\ AADT_i^{\beta}} = ave\left(\left(\frac{AADT_i}{ave(AADT)}\right)^{\beta}\right) \tag{6}$$

$$\alpha_0 = \ln\left(\frac{\frac{\sum N_{observed\ before\ crashes,i}}{Total\ VMT}}{\overline{corrected\ AADT_i^{\beta}}}\right) \tag{7}$$

STEP 4: Using the before period SPF determined, calculate the predicted average crash frequency for the treatment group during the before period

3.3.2 *After Period SPF Parameters*

STEP 5: Select the after period SPF

The after period SPF is calculated as:

$$N_{clrs,observed} = e^{(\alpha_1)} AADT_{clrs,observed}{}^{\beta_1} L_{clrs,observed} \times 730 \times 10^{-6} \tag{8}$$

Where:

α_1 = the relationship between crash frequency and roadway characteristics of rural, two-lane, undivided highways (including the CLRS) in the after period; and

β_1 = the relationship between crash frequency and AADT in the after period

STEP 6: Determine the β coefficients of the after period SPF

The specific values for the coefficients, α_1 and β_1, are needed to complete the SPF. The prediction SPF for the after period is calculated as:

$$N_{predicted,after} = e^{(\alpha_1)} AADT_{after,clrs}{}^{\beta_1} L_{after,clrs} \times 730 \times 10^{-6} \tag{9}$$

$$ln\left(\frac{N_{predicted}}{VMT}\right) = (\beta - 1)\ln(AADT) + \alpha \tag{10}$$

As with the before period, to determine the β coefficient in the after period, the relationship between the observed crashes during the after period and their respective AADT and segment lengths is fitted.

STEP 7: Determine the α coefficient of the after period SPF

The α coefficient for the after period is determined by the same method as for the before period.

$$\overline{corrected\ AADT_i^{\beta}} = ave\left(\left(\frac{AADT_i}{ave(AADT)}\right)^{\beta}\right) \tag{11}$$

$$\alpha_0 = \ln\left(\frac{\frac{\sum N_{observed\ before\ crashes,i}}{Total\ VMT}}{\overline{corrected\ AADT_i^{\beta}}}\right) \tag{12}$$

STEP 8: Using the after period SPF determined, calculate the predicted average crash frequency for the treatment group during the after period

3.3.3 *Determination of Crash Modification Factor*

The crash modification factor (CMF) for CLRS is determined by comparing the observed before and after period data with the predictions from the associated SPFs.

STEP 9: Compare the observed number of crashes at the treatment sites with the predicted crashes in the before period

When comparing the observed collisions to the expected collisions, they are related by the site effects, which include all roadway characteristics found at the study sites, as seen in Equation 13.

$$(Observed\ No.of\ Collisions)_{before} = (Predicted\ No.of\ Collisions)_{before} \times (site\ effect)_{before} \tag{13}$$

Site effects include all the factors that influence the crash rate at a certain site, such as roadway geometry, pavement condition, weather and environment, driver vehicle, etc.

$$site\ effect = effect\ of\ roadway\ geometry \times effect\ of\ weather \times effect\ driver\ behavior \times effect\ of\ pavement\ condition \times effect\ of\ vehicle\ fleet\ and\ age \times etc\ \ldots \tag{14}$$

The study period was chosen to be for the same duration and months so as to reflect the same effects in both the before and after periods. The only difference between the two periods is the addition of CLRS. The site effect in the after period is the same as that of the before period multiplied by the effect of the CLRS, which is quantified in the CMF (see Equation 15).

$$(site\ effect)_{after} = (site\ effect)_{before} \times CMF_{clrs} \tag{15}$$

The relationship between the after period SPF and the before period SPF is affected by the temporal trend in the data.

$$SPF_{after} = SPF_{before} \times temporal\ trend \qquad (16)$$

The after period SPF is used to predict the number of crashes at the treatment sites if the CLRS treatment was not installed.

$$(Predicted\ No.of\ Collisions)_{after} = SPF_{after} \times (site\ effect)_{after} \qquad (17)$$

The observed number of collisions is found in a similar way as for the before period but now includes the presence of CLRS.

$$(Observed\ No.of\ Collisions)_{after} = (Predicted\ No.of\ Collisions)_{after} \times (site\ effect)_{before} \times CMF_{clrs} \qquad (18)$$

The before period site effect is accounted for in the ratio of the average observed and the predicted crash frequencies in the before period.

$$(site\ effect)_{before} = \frac{(Average\ Observed\ No.of\ Collisions)_{before}}{(Predicted\ No.of\ Collisions)_{before}} \qquad (19)$$

Therefore,

$$CMF_{clrs} = \frac{(Observed\ No.of\ Collisions)_{after} \times (Predicted\ No.of\ Collisions)_{before}}{(Predicted\ No.of\ Collisions)_{after} \times (Observed\ No.of\ Collisions)_{before}} \qquad (20)$$

4 RESULTS

4.1 Crash Statistics

This section examines comparative statistics of crashes in the before and after periods. This constitutes a naïve analysis.

4.1.1 *Total Target Crashes*

Overall, 1545 target crashes occurred on all segments during the study period. During the before period (June 1, 2003, to May 31, 2005), 98 and 736 target crashes occurred on CLRS and non-CLRS sites, respectively, for a total of 834 target crashes. In the after period (June 1, 2005, to May 31, 2008), 56 and 655 target crashes occurred on CLRS and non-CLRS sites, respectively, for a total of 711 target crashes. Table 8 summarizes these data by 12-month period. Table 9 shows a site-by-site comparison breakdown.

Table 8. Total Crashes

	Before		**After**		**SUM**	
	6/1/2003–5/31/2004	**6/1/2004–5/31/2005**	**6/1/2006–5/31/2007**	**6/1/2007–5/31/2008**	**Before**	**After**
CLRS	52	46	31	25	**98**	**56**
NON-CLRS	332	404	327	328	**736**	**655**

Table 9. Site-by-Site Comparison, All Crash Types

Study Sites	No. of Crashes in Before Period	No. of Crashes in After Period	Change	
SR 14	11	3	-8	−72.73%
SR 16	27	9	-18	−66.67.%
SR 369	22	16	-6	−27.27%
SR 42 A	6	1	-5	−83.33%
SR 42 B	0	3	3	—
SR 204	3	4	1	+33.3%
SR 36 A	3	1	-2	−66.67%
SR 36 B	9	2	-7	−77.78%
SR 136	17	17	0	+0.00%
Overall	**98**	**56**	**-42**	**−42.86%**

4.1.2 *Analysis of Crash Severities and Types*

A naïve before–after analysis of severity types shows all sites saw a decline of the number of injuries and fatalities. Table 10 shows that CLRS sites experienced declines of −59.1% and −28.6% in injuries and fatalities, respectively. Though not as pronounced, non-CLRS sites had declines of −0.7% and −8.0% for injuries and fatalities.

Figure 6 shows that for CLRS sites, the proportion of crashes with injuries or fatalities declined by −8.17% and −0.34%, respectively. Non-CLRS sites experienced a slight increase (+0.11%) in the proportion of crashes with injuries and a slight decrease (−0.05%) in the proportion for fatalities.

Table 10. Number of Individuals Injured

		Before	After	% Change
Injuries	**CLRS**	88	36	−59.1%
	NON-CLRS	595	591	−0.7%
Fatalities	**CLRS**	7	5	−28.6%
	NON-CLRS	50	46	−8.0%

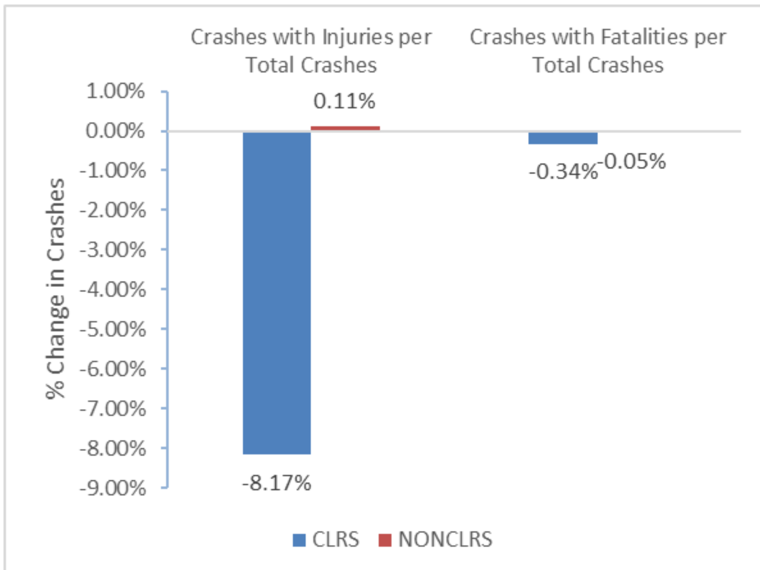


Figure 6. Percent Change of Crashes by Severity in Before versus After Period

4.1.3 *Naïve Analysis of Crashes by Collision Type*

Table 11 summarizes the number of target crashes in the before and after periods. With the exception of opposite-direction sideswipe collisions at the reference (non-CLRS) sites, all crash types showed decreases, albeit with greater decreases at the CLRS sites. Similarly, Figure 7 illustrates the change in these crashes as a portion of all crashes. Comparison of crash types for the before and after periods for all treatment sites for head-on crashes, opposite-direction sideswipe crashes, and not-a-collision-with-a-motor-vehicle crashes is shown in Tables 12–14, respectively.

Table 11. Crash by Collision Type

Type of Collision		Before	After	% Change
Head—On	CLRS	12	7	−41.7%
	NON-CLRS	70	55	−21.4%
Sideswipe—Opposite Direction	CLRS	30	6	−80.0%
	NON-CLRS	95	116	+22.1%
Not a Collision with a Motor Vehicle	CLRS	56	43	−23.2%
	NON-CLRS	570	484	−15.1%

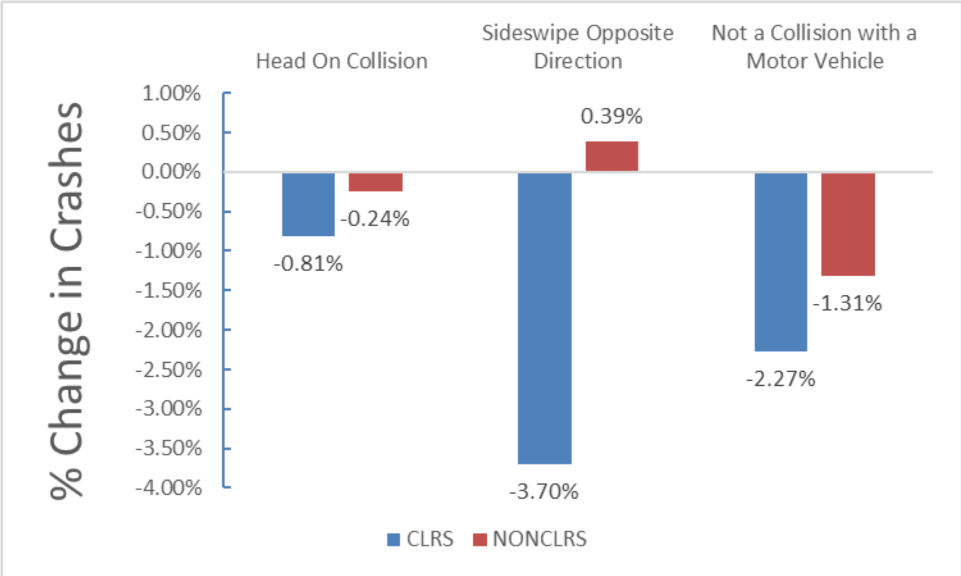


Figure 7. Percent Change in Crashes by Type

Table 12. Comparison of Head-on Crashes

Study Sites	Number of Target Crashes		Change in Crashes	
	Before	After	Number	Percent
SR 14	3	0	−3	−100%
SR 16	3	1	−2	−66.7%
SR 369	3	3	0	0.0%
SR 42 A	0	0	0	—
SR 42 B	0	1	+1	n.d.*
SR 204	0	0	0	—
SR 36 A	0	0	0	—
SR 36 B	2	0	−2	−100.0%
SR 136	1	2	+1	+100.0%
Total	**12**	**7**	**−5**	**−41.7%**
*n.d. = not defined				

Table 13. Comparison of Opposite-Direction Sideswipe Crashes

Study Sites	Number of Target Crashes		Change in Crashes	
	Before	After	Number	Percent
SR 14	2	0	−2	−100.0%
SR 16	10	1	−9	−90.0%
SR 369	12	3	−9	−75.0%
SR 42 A	1	0	−1	−100.0%
SR 42 B	0	0	0	—
SR 204	0	1	+1	n.d.
SR 36 A	1	0	−1	−100.0%
SR 36 B	1	0	−1	−100.0%
SR 136	3	1	−2	−66.7%
Total	**30**	**6**	**−24**	**−80.0%**
*n.d. = not defined				

Table 14. Comparison of Not-a-Collision-with-a-Motor-Vehicle Crashes

Study Sites	Number of Target Crashes		Change in Crashes	
	Before	After	Number	Percent
SR 14	6	3	−3	−50.0%
SR 16	14	7	−7	−50.0%
SR 369	7	10	+3	+42.9%
SR 42 A	5	1	−4	−80.0%
SR 42 B	0	2	+2	n.d.
SR 204	3	3	0	—
SR 36 A	2	1	−1	−50.0%
SR 36 B	6	2	−4	−66.7%
SR 136	13	14	+1	+7.7%
Total	**56**	**43**	**−13**	**−23.2%**
*n.d. = not defined				

4.2 Empirical Bayes Method/Development of SPF

Table 15 lists the critical parameters used in the EB method to determine the respective SPF for each study period.

Table 15. Crash Statistics

	Before	After
Segments with crashes	624	533
All roadway segments w/o CLRS	2,413	2,318
Average AADT (vehicles)	4,302	4,218
Total VMT (millions)	20,384	19,870
Standard crash frequency rate (target crashes/yr/10^6 VMT)	0.0431	0.0377

4.2.1 *Before Period SPF Parameters*

STEP 1: Select the before period SPF

Since the number of crashes in this analysis period is for two years, the general prediction SPF (Equation 2) is modified as shown in Equation 21:

$$N_{predicted,before} = e^{(\alpha_0)} AADT_{before}{}^{\beta_0} L_{before} \times 730 \times 10^{-6} \tag{21}$$

STEP 2: Determine the β coefficient of the before period SPF

Equation 21 is modified to include vehicle miles traveled embedded within the AADT, as shown in Equation 3.

$$VMT = AADT \times \ L \times 730 \times 10^{-6} \tag{22}$$

VMT is incorporated into the before prediction SPF, Equation 21, and simplified as in Equation 4.

The study set included 623 roadway segments without CLRS that experienced crashes. The relationship demonstrated in Equation 4, between the observed crashes during the before period and their respective AADT and segment lengths, is used to determine the β coefficient as illustrated in Figure 8.

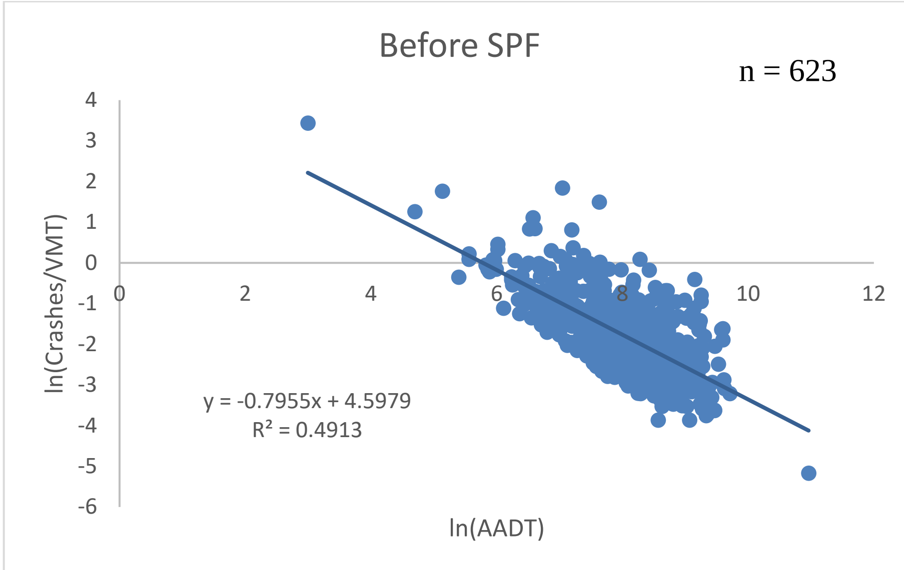


Figure 8. BEFORE: ln(crashes per VMT) versus ln(AADT)

$$-0.7955 = (\beta - 1), \text{ therefore } \beta = 0.2045 \tag{23}$$

STEP 3: Determine the α coefficient of the before period SPF

The α coefficient is determined by accounting for all roadways in the before–after set, even those with no target crashes. The before period study set comprised 2413 roadway segments without CLRS. The average AADT for all these segments was 4302 vehicles. AADT must be adjusted by the ratio of the base rate to the treatment AADT, referencing a weighted average. To do so, the AADT for each section was corrected by dividing each individual AADT by the average AADT. Then these values were replaced in the before period SPF equation and raised to the β coefficient determined in step 2. The standard condition was determined by averaging these values, and was 0.95, as seen in Equation 24.

$$\overline{corrected\ AADT_i^{\beta}} = ave\left(\left(\frac{AADT_i}{ave(AADT)}\right)^{0.2045}\right) = 0.95 \qquad (24)$$

$$ave(AADT) = 4302\ vehicles$$

Alpha was then determined by incorporating the standard condition into the original before period SPF, as shown in Equations 25 and 26.

$$\frac{\sum N_{observed\ before\ crashes,i}}{Total\ VMT\ for\ all\ roadways} = e^{(\alpha_0)}\overline{corrected\ AADT_i^{\beta}} \qquad (25)$$

$$\alpha_0 = \ln\left(\frac{\frac{\sum N_{observed\ before\ crashes,i}}{Total\ VMT}}{\overline{corrected\ AADT_i^{\beta}}}\right) = \ln\left(\frac{\frac{834}{20384}}{0.95}\right) = -3.1440 \qquad (26)$$

The $e^{-3.1435}$constant gives the standard crash frequency rate for the before period, which is 0.0431 target crashes/yr/10^6 VMT.

STEP 4: Using the before period SPF determined, calculate the predicted average crash frequency for the treatment group during the before period

Using the treatment before sites' AADT and segment lengths, the predicted number of crashes is calculated as follows:

$$\mathrm{N_{predicted,before}} = \mathrm{e}^{-3.1440} AADT_{\mathrm{before}_i}^{0.2045} \mathrm{L_{before_i}} \times 730 \times 10^{-6} \qquad (27)$$

However, each segment has different values of AADT per year and only has a certain number of days in that year. Since the before period study is from June 1, 2003, to May 31, 2005, the predicted frequency is more accurately calculated as shown in Equation 28.

$$
\begin{aligned}
N_{\text{predicted,before}} \\
&= (e^{-3.1440} AADT_{2003_i}^{0.2045} L_{\text{before}_i} \times 181 \times 10^{-6}) \\
&+ (e^{-3.1440} AADT_{2004_i}^{0.2045} L_{\text{before}_i} \times 365 \times 10^{-6}) \\
&+ (e^{-3.1440} AADT_{2005_i}^{0.2045} L_{\text{before}_i} \times 184 \times 10^{-6})
\end{aligned}
\tag{28}
$$

Table 16 displays the total predicted crashes on all nine sites.

Table 16. Predicted Crash Frequency in Before Period

Site	Predicted before total crash frequency (vehicles)
SR 14	3.43
SR 16	6.15
SR 369	10.97
SR 42 A	0.83
SR 42 B	1.92
SR 204	1.96
SR 36 A	1.24
SR 36 B	2.37
SR 136	1.61
Total	**30.48**

4.2.2 *After Period SPF Parameters*

STEP 5: Select the after period SPF

As in the before period, the general equation for the after period SPF is:

$$N_{predicted,after} = e^{(\alpha_1)} AADT_{after}{}^{\beta_1} L_{after} \times 730 \times 10^{-6} \quad (29)$$

STEP 6: Determine the β coefficient of the after period SPF

Using the same equation as Equation 4, but with the AADT and segment length values of the after period comparison set, 532 roadway segments with crashes were plotted as shown in Figure 9.

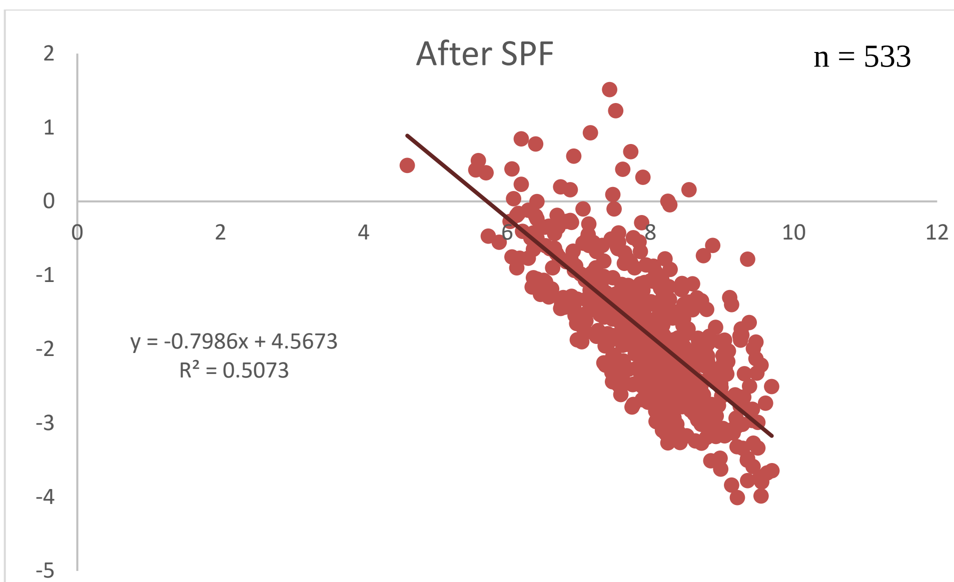


Figure 9. AFTER: ln(crashes per VMT) versus ln(AADT)

$$-0.7986 = (\beta - 1), \text{ therefore } \beta = 0.2014 \quad (30)$$

STEP 7: Determine the α coefficient of the after period SPF

The α coefficient for the after period is determined by the same method as the before period. This set contained a total of 2318 roadway segments without CLRS. The average AADT for all these segments was 4218 vehicles. The standard condition is determined by averaging these values, which was 0.95, as seen in Equation 31.

$$\overline{corrected\ AADT_i^{\beta_1}} = ave\left(\left(\frac{AADT_i}{ave(AADT)}\right)^{\beta_1}\right) = 0.95 \tag{31}$$

$$ave(AADT) = 4{,}218\ vehicles$$

$$Total\ VMT = 19{,}870\ \times 10^6 miles$$

$$\alpha_1 = \ln\left(\frac{\frac{\sum N_{observed\ after\ crashes,i}}{Total\ VMT}}{\overline{corrected\ AADT_i^{\beta_1}}}\right) = \ln\left(\frac{\frac{711}{19870}}{0.95}\right) = -3.2790 \tag{32}$$

The $e^{-3.2790}$ constant gives the standard crash frequency rate for the before period, which is 0.0377 target crashes/yr/10^6 VMT.

STEP 8: Using the after period SPF determined, calculate the predicted average crash frequency for the treatment group during the after period

Using the treatment after sites' AADT and segment lengths, the predicted number of crashes in the after period is generally calculated by Equation 33:

$$N_{clrs,predicted,after} = e^{(-3.2790)} AADT_{clrs,after}^{0.2014} L_{clrs,after} \times 730 \times 10^{-6} \tag{33}$$

The predicted frequency for the after period from June 1, 2006, to May 31, 2008, was more accurately calculated as shown in Equation 34.

$$N_{\text{predicted,after}} = (e^{-3.2790} AADT_{2006_i}^{0.2014} L_{\text{before}_i} \times 181 \times 10^{-6}) + (e^{-3.2790} AADT_{2007_i}^{0.2014} L_{\text{before}_i} \times 365 \times 10^{-6}) + (e^{-3.2790} AADT_{2008_i}^{0.2014} L_{\text{before}_i} \times 184 \times 10^{-6}) \quad (34)$$

Table 17 displays the total predicted crashes on all 9 sites, totaling 26.197 vehicles.

Table 17. Predicted Crash Frequency in After Period

Site	Predicted after total crash frequency (vehicles)
SR 14	2.83
SR 16	5.26
SR 369	9.19
SR 42 A	0.73
SR 42 B	1.86
SR 204	1.73
SR 36 A	1.00
SR 36 B	2.15
SR 136	1.45
Total	**26.20**

4.2.3 *Determination of CMF*

The CMF for CLRS is determined by analyzing the data from the before period and the after period.

STEP 9: Compare the observed number of crashes at the treatment sites with the predicted crashes in the before period

As discussed in Section 3.3.3 (Equations 13–19), the CMF is calculated as:

$$CMF_{clrs} = \frac{(Observed\ No.of\ Collisions)_{after} \times (Predicted\ No.of\ Collisions)_{before}}{(Predicted\ No.of\ Collisions)_{after} \times (Observed\ No.of\ Collisions)_{before}}$$

$$= \frac{56 \times 30.48}{26.20 \times 98} = 0.58329 \approx \mathbf{0.66}$$

4.3 Misclassified Crashes

The research team conducted a comprehensive manual review of 2,203 crashes on CLRS sites and 15,199 crashes on comparison sties without CLRS from 2003 to 2008. Table 18 shows the number of misclassified target crashes and the reason for their misclassification. Misclassifications were found in 6.73% of all target crashes. The category “not a collision with a motor vehicle” had the most misclassifications. This was mainly due to the use of wrong definitions for each classification. A head-on or angle collision involves more than one vehicle. When a motor vehicle collides with anything other than another motor vehicle, it is considered “not a collision with a motor vehicle.” Some police officers misinterpreted a motor vehicle crashing head-on or at angle with an object as “head-on” or “angle.” This error could easily be prevented in the future by more specific training.

Table 18. Misclassified Target Crashes

		Correct Classification: Sideswipe—Opposite direction	Correct Classification: Not a collision with a motor vehicle	Correct Classification: Head—on	Subtotal
Misclassification	**Sideswipe—Opposite direction**		2	3	5
Misclassification	**Not a collision with a motor vehicle**	2		—	2
Misclassification	**Head—on**	2	29		31
Misclassification	**Angle**	23	21	14	58
Misclassification	**Rear**	—	1	—	1
Misclassification	**No classification/ left blank**	—	2	—	2
Misclassification	**Sideswipe—Same direction**	3	1	1	5
	Subtotal	30	56	18	**104**

4.4 Crashes Involving Hydroplaning

For this analysis, crashes that included a vehicle's loss of control due to water, ice, snow, or some other condition that caused the vehicle to hydroplane are included in the analysis. Target crashes that claimed any of the following influences were consider in this hydroplaning subset: hydroplaning, ice, running or standing water, wet pavement, rain, sleet, fuel in road, and sewage in road. Overall, 18.25% of all target crashes experienced some form of hydroplaning. Specifically, 16.23% and 18.48% of the CLRS and non-CLRS target crashes, respectively, were included in this hydroplaning category.

5 CONCLUSIONS AND RECOMMENDATIONS

Vehicle crashes involving crossing over the roadway centerlines (head-on, opposite direction sideswipe, and run-of-road that began crossing the centerline crashes) are among the most severe types of collisions nationwide. Several state transportation agencies have already implemented and conducted safety analyses that concluded that centerline rumble strips have reduced the target crashes. As part of the statewide roadway safety plan, the Georgia Department of Transportation installed approximately 200 miles of centerline rumble strips to prevent crashes stemming from crossing over the centerline. The purpose of this thesis was to analyze the effectiveness of the centerline rumble strips as a safety countermeasure on rural two-lane undivided highways in the state of Georgia. CLRS have contributed to 33% decrease in vehicular crashes throughout the state at a relatively low cost. Ultimately, this study concluded that CLRS are a cost effective countermeasure to improve safety along rural two-lane undivided highways in Georgia.

5.1 Safety Analysis

This study quantified the safety impacts of CLRS along rural two-lane undivided highways in Georgia. A naïve comparison showed an overall decrease of 42.86% in vehicular crashes throughout the state. This reduction estimate is subject to selection bias because the naïve comparison method assumes that the CLRS is the only factor affecting the number of crashes. In reality, other factors at play also lead to any change of the crash rate. To account for these other factors, the empirical Bayes methods was used to determine a decrease of 34% in vehicular crashes that could be associated with CLRS. The bias reflected in the naïve before-after analysis exaggerates the actual effect of CLRS because it does not accurately account for the overall decline in crashes that is attributed to factors other than CLRS.

5.1.1 *Empirical Bayes Method*

The empirical Bayes method accounts for selection bias by combining observed crash frequency with predicted crash frequency from a reference group of roadways with similar roadway characteristics. In a 24-month before period, 98 target vehicular crashes were observed on 113.51 miles of CLRS roadways. These observations were combined with 736 target vehicular crashes on 7,764.91 miles of roadways with similar roadway characteristics occurring during the 24 month after period to develop a calibrated safety performance function for the before period. The SPF explicitly accounts for selection bias and traffic characteristics that may influence the number of crashes. For the after period, 655 target vehicular crashes along 7,906.44 miles of reference roads with similar roadway characteristics was used to develop the calibrated SPF for the before period. Also, in the after period, 56 crashes were observed on the 113.51 miles of CLRS roadways. Subsequently, the crash modification factor of 0.66 was determined by using the predicted and observed number of target crashes. This CMF reveals that a 34% decrease of crashes can be directly associated with the installation of CLRS. With relatively low installations costs, the favorable CMF suggests that a wider application of CLRS should be considered on other rural two-lane undivided highways in Georgia.

5.2 Limitations

Although, the study found favorable results, limitations were not absent from the process. The analysis was limited by the amount of crash data available. Initially, the study was designed to use complete calendar years (from January 1 to December 31) for all of 2003 and 2004, as the before period. An adjustment period in which drivers had time to adapt to the CLRS was designed to be the entire calendar years of 2005 and 2006. Lastly, the after period was designed to use complete calendar years (from January 1 to December 31) for all of 2007 and 2008. However, crash data for all of 2008 was unavailable to the researchers. To rectify this limitation, a complete 24-month period from June 1 to May 31 was used to account for seasonal changes. This time frame still accounted for a period in which the drivers could adjust to the construction and introduction of CLRS on the study routes. Additionally, the sample size for target crashes was too small to obtain crash severity and collision type individual CMFs.

Common to most crash analysis, limitations in the crash reporting process were also revealed in this study. Crash details are compiled from crash reports recorded by police officers at the scene of the crash. Given the high number of police officers that complete reports this leads to significant differences in interpretations, which naturally introduces variability and non-uniformity in the crash data. These errors can be significant to the analysis, especially when the analysis is dependent on the correct description of events that led to the crash and the classification of the incidents and injuries incurred from the crash. For example, crash severity is entered as an injury code, which consist of: 0-not injured, 1-killed, 2-serious, 3-visible, and 4-complaint. The scale of injury is solely up to the interpretation of the officer. A reporting officer may input a 1 or a 0 under injury code to represent 1 or no injured person. However, this would translate to a fatality or no injury, respectively. Additionally, manual verification of crashes found errors in the manner of collision classifications. Manner of collision options include: 1-angle, 2-head on, 3-rear end, 4-sideswipe-same direction, 5-

sideswipe-opposite direction, and 6-not a collision with a motor vehicle. Codes 1 through 5 refer to the manner in which two vehicles collided. Any crash that entails a collision with an object other than a motor vehicle is to be noted as a category 6-not a collision with a motor vehicle. At times reporting officers use classifications 1 through 3 in a collision involving one vehicle. For example, if a vehicle crashed into an object head on, it may have been mistakenly coded as a category 2-head on, instead of the correct category 6-not a collision with a motor vehicle. Ultimately, proper training is essential to safeguard that these details are entered accurately.

The comprehensive manual review of more than 17,000 crash reports proved to be the most resource-intensive part of the analysis effort. This quality assurance process served as a critical step in reducing crash misclassifications and improving the overall derivation of the CMF value. The research team manually checked the base crash data against the crash description recorded by the investigating police officer to verify crash type, as well as obtain a clearer indication as to whether the crash could have been impacted by the presence of CLRS. Misclassification were present in 6.73% of all target crashes. This step was critical to improving the reliability of the CMF value, as it reduced crash misclassifications. A broader methodological recommendation from the lessons learned in the study is to employ sufficient crash verification procedures in any safety study that develops a crash modification factor, especially in cases where the sample size of the crashes is small, or if crash modification factors are desired for specific crash categories.

5.3 Further Research

As shown in this thesis, the EB method provides a valuable analysis of how CLRS augment overall roadway safety along rural, two-lane, undivided roads in Georgia. Other research topics naturally stem from this initial analysis, which would provide GDOT and other transportation agencies a more thorough understanding of the safety impacts of centerline rumble strips. An EB analysis should be completed without vehicles that experienced some sort of hydroplaning, as explained in Section 4.4, to reveal further safety impacts of CLRS. Additional investigations should focus on the safety effects of mixed applications with rumble strips, such as in combination with shoulder rumble strips. Studies in other states, such as that of Kay et. al. (2015) and Sayed et. al. (2010), analyzed the compounded safety effects of the combination of SRS and CLRS [9, 19]. Not only are the safety effects enhanced, but additional crash types can be targeted by the combined treatment, such as road departure collisions that may stem from crossing the right shoulder first. Additionally, rumble strips installed with a treatment that provides a visual warning, such as high-visibility pavement markings, may also prove to modify the effectiveness of CLRS. Different driving, pavement, and weather conditions may also influence the safety effects of CLRS. Some studies have also found that rumble strips promote maintaining drivers in their respective lanes. Perhaps a more in-depth analysis of important subcategories, by crash severity and collision type, would reveal the effects of CLRS on these subcategories. Based on these additional findings, targeted guidelines could be developed that would benefit not only Georgia but other states in the southeast region of the US. This insight gathered from future research can prove useful for transportation agencies when prioritizing the implementation of additional rumble strip projects.

APPENDIX A. SITE CHARACTERISTICS

The following appendix is adapted from [4] and updated for this thesis.

Characteristics pertaining to each of the 9 centerline rumble strips installation site are detailed in this appendix. The order in which the sites are presented is first by project ID number and second by State Route number if the project contained more than one section of roadway with centerline rumble strips. Each site contains information obtained from GDOT, the study beginning and ending mileposts, a map overview, Google Street View® screenshots detailing the extent of the centerline rumble strips' installation, and basic before and after crash statistics. In Georgia, GDOT implemented seven projects which installed centerline rumble strips. The combined total mileage of roadways with centerline rumble strips on these 9 sites is about 126-miles. The installation sites, indicated with a red roadway are spread throughout 12 counties, highlighted in light green, as seen in Figure **10**.

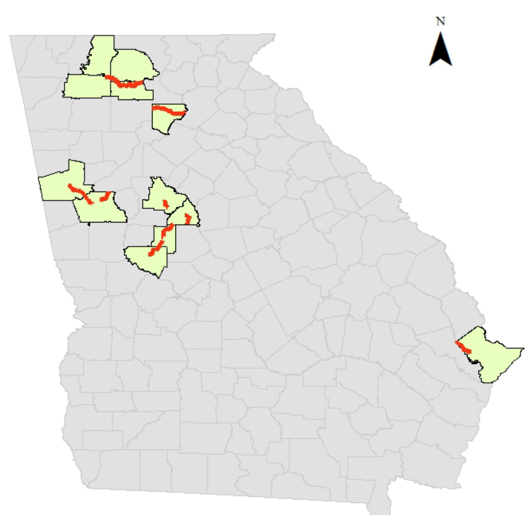


Figure 10. Centerline Rumble Strips Installation Sites in Georgia [22-24]

Project ID: 0006693, SR 14

Figure 11. Project 0006693, SR 14 Details (clockwise from top left): Map [22-24], Location within Georgia, MP 21 Facing South [4, 21], MP 27 Facing North [4, 21]

Table 19. Project 0006693, SR 14, Construction Details [4]

Source	Attribute	Detail
TransPI	Project ID	0006693
	Project Number	CSSTP-0006-00(693)
	Project Title	SR 14|SR 16|SR 154@SEV LOC IN CARROLL&COWETA [CENTERLINE]
	Management Let	6/17/2005
	Project Completion	4/26/2006
	Project Manager	Scott Zehngraff
	Office	Traffic Safety & Design
	Project Type	Safety
	DOT District	3, 6
	Congressional District	3
	Project Description	SR 14 from Herring Road/CR 43 to Johnston Circle/CR 7 (approximately 6.5 miles in Coweta County)
	Construction Contractor	JHC Corporation
	MPO	Atlanta TMA, Not Urban
Federal Report of Completed Projects	Construction Begin Date	10/11/2005
	Time Charges Stop Date	10/31/2005

Table 20. Project 0006693, SR 14, Study Details [4]

Attribute	Detail
Primary Roadway	Georgia State Route 14
Beginning Milepost	19.68
Ending Milepost	27.55
County (Begin)	Coweta
County (End)	Coweta
Length (mi)	7.87
RCLINK	0771001400
Beginning Coordinates	33.436426,-84.750488
Ending Coordinates	33.50653,-84.671309
AADT MP 19.68 – 19.73	2003 – 10,090 2004 – 9,970 2005 – 10,970 2006 – 9,720 2007 – 10,430 2008 – 10,110
AADT MP 19.74 – 23.16	2003 – 14,470 2004 – 14,730 2005 – 14,240 2006 – 13,230 2007 – 14,120 2008 – 12,580
AADT MP 23.17 – 26.71	2003 – 8,450 2004 – 8,600 2005 – 8,660 2006 – 8,170 2007 – 8,700 2008 – 8,210
AADT MP 26.72 – 27.55	2003 – 9,320 2004 – 9,520 2005 – 9,090 2006 – 9,040 2007 – 9,310 2008 – 9,490

Project ID: 0006693, SR 16

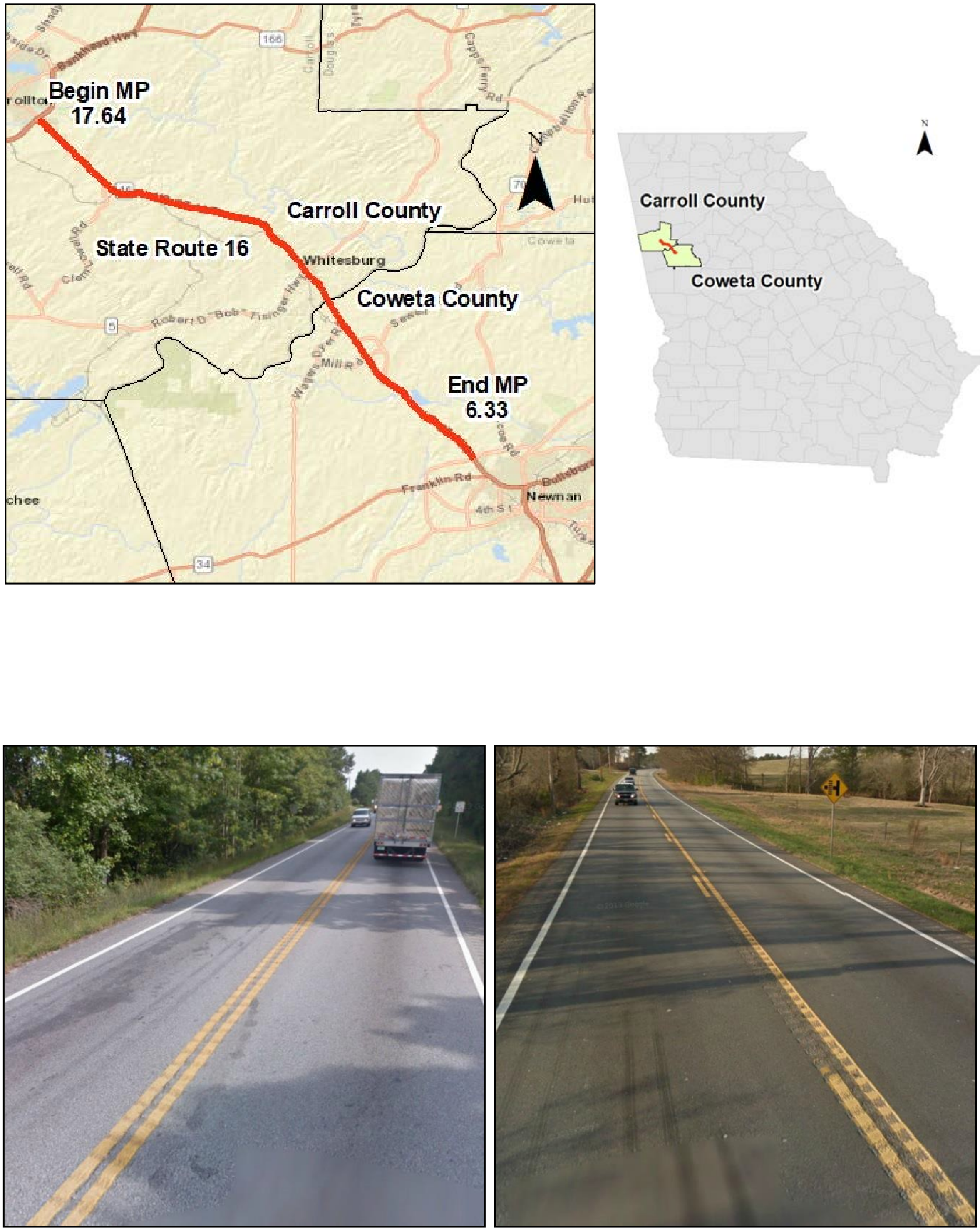


Figure 12. Project 0006693, SR 16 Details (clockwise from top left): Map [22-24], Location within Georgia, MP 6 Facing North [4, 21], MP 17 Facing South [4, 21]

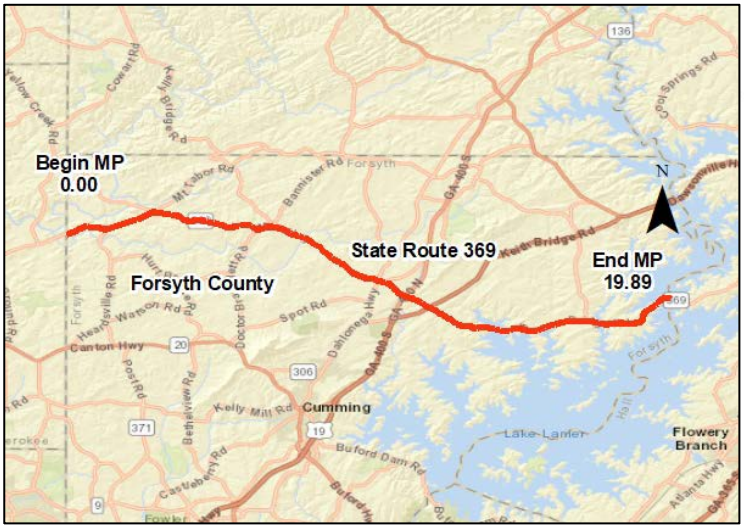
Table 21. Project 0006693, SR 16, Construction Details [4]

Source	Attribute	Detail
TransPI	Project ID	0006693
	Project Number	CSSTP-0006-00(693)
	Project Title	SR 14|SR 16|SR 154@SEV LOC IN CARROLL&COWETA [CENTERLINE]
	Management Let	6/17/2005
	Project Completion	4/26/2006
	Project Manager	Scott Zehngraff
	Office	Traffic Safety & Design
	Project Type	Safety
	DOT District	3, 6
	Congressional District	3
	Project Description	SR 16/US 27 from the Carrollton Bypass to the Newnan Bypass (approximately 17 miles in Carroll and Coweta Counties)
	Construction Contractor	JHC Corporation
	MPO	Atlanta TMA, Not Urban
Federal Report of Completed Projects	Construction Begin Date	10/11/2005
	Time Charges Stop Date	10/31/2005

Table 22. Project 0006693, SR 16, Study Details [4]

Attribute	Detail
Primary Roadway	Georgia State Route 16
Beginning Milepost	17.64
Ending Milepost	6.33
County (Begin)	Carroll
County (End)	Coweta
Length (mi)	16.56
RCLINK	0451001600 0771001600
Beginning Coordinates	33.557189, -85.036286
Ending Coordinates	33.397675,-84.82808
AADT MP 17.64 – 22.65 (Carroll County)	2003 – 10,250 2004 – 10,430 2005 – 10,970 2006 – 12,980 2007 – 11,520 2008 – 10,270
AADT MP 22.65 – 26.18 (Carroll County)	2003 – 9,230 2004 – 10,970 2005 – 11,520 2006 – 10,040 2007 – 10,040 2008 – 9,660
AADT MP 26.19 – 27.87 (Carroll County)	2003 – 9,280 2004 – 9,450 2005 – 8,920 2006 – 9,780 2007 – 8,390 2008 – 9,070
AADT MP 0.00 – 3.85 (Coweta County)	2003 – 8,880 2004 – 8,760 2005 – 8,360 2006 – 8,360 2007 – 8,320 2008 – 7,280
AADT MP 3.86 – 6.33 (Coweta County)	2003 – 10,100 2004 – 10,240 2005 – 9,780 2006 – 9,780 2007 – 9,330 2008 – 8,980

Project ID: 0006945, SR 369



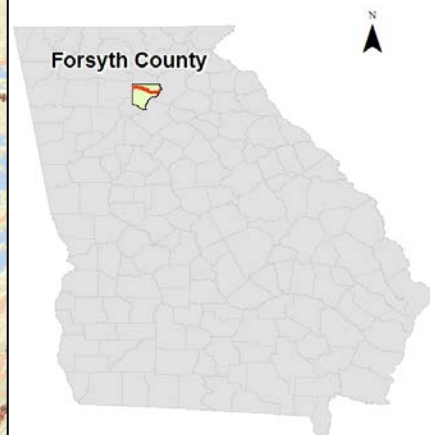


Figure 13. Project 0006945, SR 369 Details (clockwise from top left): Map [22-24], Location within Georgia, MP 10 Facing East [4, 21], MP 1 Facing East [4, 21]

Table 23. Project 0006945, SR 369, Construction Details [4]

Source	Attribute	Detail
TransPI	Project ID	0006945
	Project Number	CSSTP-0006-00(945)
	Project Title	SR 369 FM CHEROKEE CO TO HALL CO - CENTERLINE RUMBLE STRIPS
	Management Let	6/17/2005
	Project Completion	4/26/2006
	Project Manager	Scott Zehngraff
	Office	Traffic Safety & Design
	Project Type	Safety
	DOT District	1
	Congressional District	9
	Project Description	Indentation centerline rumble strips on SR 369 in Forsyth County in District 1
	Construction Contractor	Peek Pavement Marking, LLC
	MPO	Atlanta TMA
Federal Report of Completed Projects	Construction Begin Date	3/6/2006
	Time Charges Stop Date	3/26/2006

Table 24. Project 0006945, SR 369, Study Details [4]

Attribute	Detail
Primary Roadway	Georgia State Route 369
Beginning Milepost	0.00
Ending Milepost	19.89
County (Begin)	Forsyth
County (End)	Forsyth
Length (mi)	19.89
RCLINK	1171036900
Beginning Coordinates	34.295106,-84.258292
Ending Coordinates	34.262606,-83.95333
AADT MP 0.00 – 2.70	2003 – 7,060 2004 – 8,230 2005 – 7,650 2006 – 7,730 2007 – 7,730 2008 – 7,360
AADT MP 2.70 – 5.79	2003 – 8,140 2004 – 9,980 2005 – 10,040 2006 – 9,480 2007 – 9,790 2008 – 9,290
AADT MP 5.80 – 6.42	2003 – 10,180 2004 – 14,760 2005 – 14,310 2006 – 14,100 2007 – 13,990 2008 – 13,310
AADT MP 6.43 – 10.06	2003 – 13,710 2004 – 15,590 2005 – 14,450 2006 – 12,700 2007 – 12,970 2008 – 12,950

Table 24 Continued on Next Page

Table 24 Continued

AADT MP 10.07 – 11.07	2003 – 15,400 2004 – 18,590 2005 – 18,380 2006 – 18,000 2007 – 18,510 2008 – 17,620
AADT MP 11.08 – 11.85	2003 – 11,160 2004 – 12,010 2005 – 12,630 2006 – 12,580 2007 – 12,900 2008 – 12,030
AADT MP 11.86 – 12.81	2003 – 20,040 2004 – 21,030 2005 – 23,640 2006 – 25,320 2007 – 20,420 2008 – 19,430
AADT MP 12.82 – 19.89	2003 – 14,730 2004 – 16,220 2005 – 15,660 2006 – 15,510 2007 – 15,960 2008 – 13,960

Project ID: 0006975, SR 42 Section A

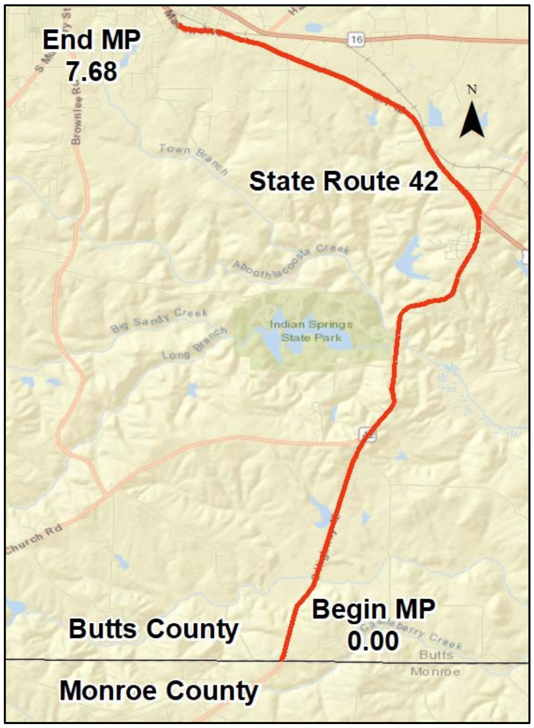


Figure 14. Project 0006975, SR 42, Section A Details (clockwise from top left): Map [22-24], Location within Georgia, MP 7 Facing South [4, 21], MP 3 Facing South [4, 21]

Table 25. Project 0006975, SR 42, Section A, Construction Details [4]

Source	Attribute	Detail
TransPI	Project ID	0006975
	Project Number	CSSTP-0006-00(975)
	Project Title	SR 42@SEV LOC IN HENRY|BUTTS|MONROE-CENTERLINE RUMBLE STRIPS
	Management Let	8/19/2005
	Project Completion	12/3/2009
	Project Manager	Scott Zehngraff
	Office	Traffic Safety & Design
	Project Type	Safety
	DOT District	3
	Congressional District	3, 8
	Project Description	Indentation centerline rumble strips on SR 42 at several locations in Henry, Butts, and Monroe Counties in District 3
	Construction Contractor	-
	MPO	Atlanta TMA, Not Urban
Federal Report of Completed Projects	Construction Begin Date	1/17/2006
	Time Charges Stop Date	5/31/2006

Table 26. Project 0006975, SR 42, Section A, Study Details [4]

Attribute	Detail
Primary Roadway	Georgia State Route 42
Beginning Milepost	0.00
Ending Milepost	7.68
County (Begin)	Butts
County (End)	Butts
Length (mi)	7.68
RCLINK	0351004200
Beginning Coordinates	33.201781,-83.936577
Ending Coordinates	33.290544,-83.950943
AADT MP 0.00 – 3.17	2003 – 1,650 2004 – 1,300 2005 – 1,440 2006 – 1,300 2007 – 1,360 2008 – 1,540
AADT MP 3.18 – 4.80	2003 – 1,500 2004 – 2,690 2005 – 2,590 2006 – 2,280 2007 – 3,080 2008 – 2,930
AADT MP 4.81 – 7.43	2003 – 6,620 2004 – 6,300 2005 – 6,340 2006 – 5,540 2007 – 6,120 2008 – 5,430
AADT MP 7.44 – 7.68	2003 – 2,680 2004 – 2,730 2005 – 8,380 2006 – 9,140 2007 – 8,200 2008 – 7,890

Project ID: 0006975, SR 42 Section B

Figure 15. Project 0006975, SR 42, Section B Details (clockwise from top left): Map [22-24], Location within Georgia, MP 8 Facing North [4, 21], MP 5 Facing North [4, 21]

Table 27. Project 0006975, SR 42, Section B, Construction Details [4]

Source	Attribute	Detail
TransPI	Project ID	0006975
	Project Number	CSSTP-0006-00(975)
	Project Title	SR 42@SEV LOC IN HENRY|BUTTS|MONROE-CENTERLINE RUMBLE STRIPS
	Management Let	8/19/2005
	Project Completion	12/3/2009
	Project Manager	Scott Zehngraff
	Office	Traffic Safety & Design
	Project Type	Safety
	DOT District	3
	Congressional District	3, 8
	Project Description	Indentation centerline rumble strips on SR 42 at several locations in Henry, Butts, and Monroe Counties in District 3
	Construction Contractor	-
	MPO	Atlanta TMA, Not Urban
Federal Report of Completed Projects	Construction Begin Date	1/17/2006
	Time Charges Stop Date	5/31/2006

Table 28. Project 0006975, SR 42, Section B, Study Details [4]

Attribute	Detail
Primary Roadway	Georgia State Route 42
Beginning Milepost	4.58
Ending Milepost	9.81
County (Begin)	Henry
County (End)	Henry
Length (mi)	5.23
RCLINK	1511004200
Beginning Coordinates	33.354986,-84.114869
Ending Coordinates	33.424601,-84.143735
AADT MP 4.48 – 8.52	2003 – 10,020 2004 – 10,920 2005 – 10,160 2006 – 10,950 2007 – 11,700 2008 – 11,260
AADT MP 8.53 – 9.95	2003 – 7,990 2004 – 7,880 2005 – 7,480 2006 – 9,250 2007 – 8,290 2008 – 8,160

Project ID: 0006976, SR 204

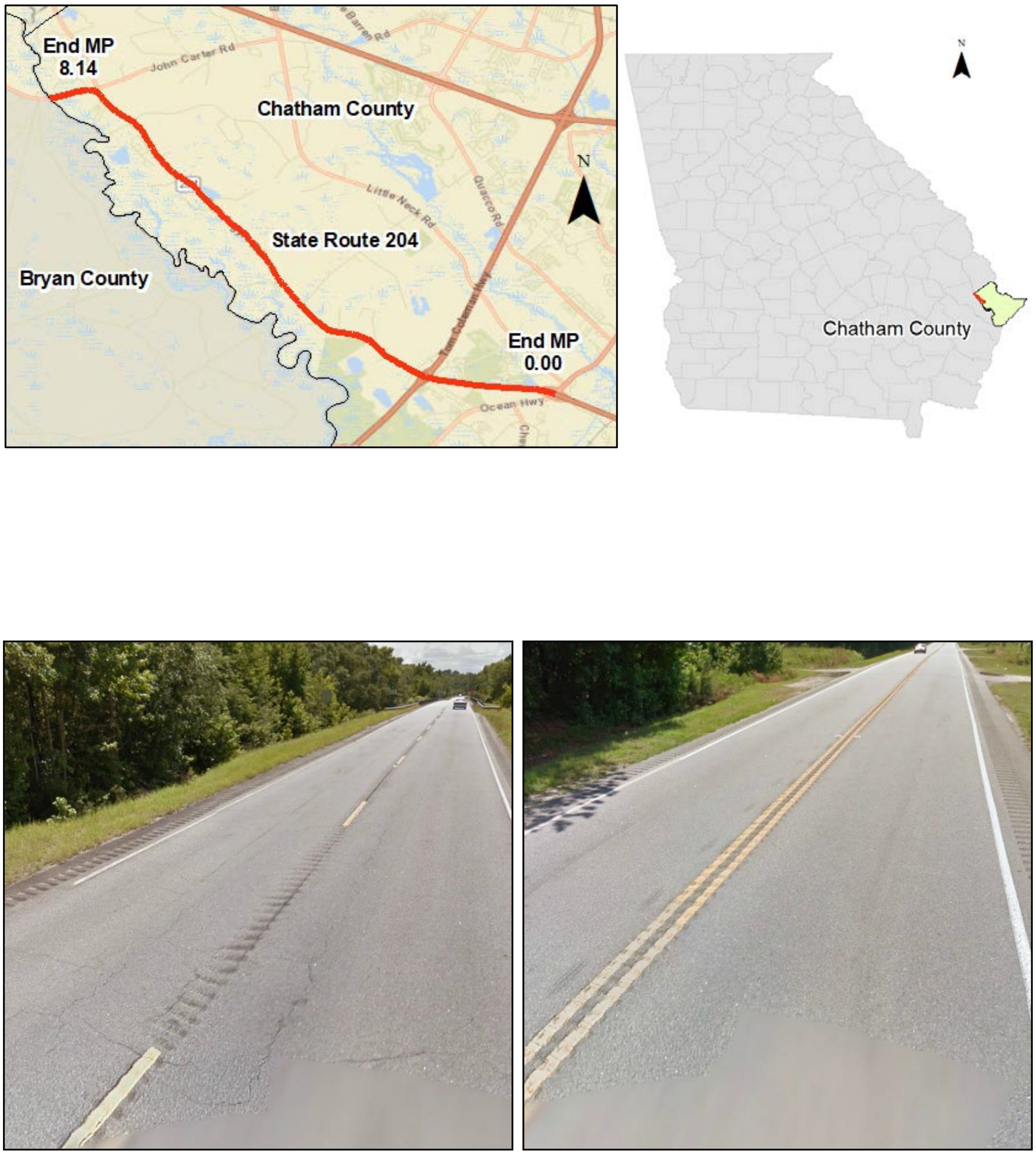


Figure 16. Project 0006976 SR 204 Details (clockwise from top left): Map [22-24], Location within Georgia, MP 8 Facing West [4, 21], MP 1 Facing East [4, 21]

Table 29. Project 0006976, SR 204, Construction Details [4]

Source	Attribute	Detail
TransPI	Project ID	0006976
	Project Number	CSSTP-0006-00(976)
	Project Title	SR 204 FM BRYAN COUNTY LINE TO I-95-CENTERLINE RUMBLE STRIPS
	Management Let	8/19/2005
	Project Completion	12/3/2009
	Project Manager	Scott Zehngraff
	Office	Traffic Safety & Design
	Project Type	Safety
	DOT District	5
	Congressional District	1, 12
	Project Description	This safety improvement project consists of installing centerline ground-in rumble strips on State Route 204 in Chatham County from the Bryan County line to I-95. The intent of this project is to reduce the frequency of head-on and opposite-direction sideswipe crashes.
	Construction Contractor	Peek Pavement Marking, LLC
	MPO	Savannah TMA
Federal Report of Completed Projects	Construction Begin Date	2/14/2006
	Time Charges Stop Date	2/28/2006

Table 30. Project 0006976, SR 204, Study Details [4]

Attribute	Detail
Primary Roadway	Georgia State Route 204
Beginning Milepost	0.00
Ending Milepost	8.14
County (Begin)	Chatham
County (End)	Chatham
Length (mi)	8.14
RCLINK	0511020400
Beginning Coordinates	32.079743,-81.383479
Ending Coordinates	32.006607,-81.28781
AADT MP 0.00 – 0.63	2003 – 3,910 2004 – 3,980 2005 – 3,810 2006 – 3,900 2007 – 4,020 2008 – 3,310
AADT MP 0.69 – 8.14	2003 – 7,180 2004 – 7,000 2004 – 7,530 2004 – 7,470 2007 – 7,420 2008 – 6,710

Project ID: 0007077, SR 36 Section A

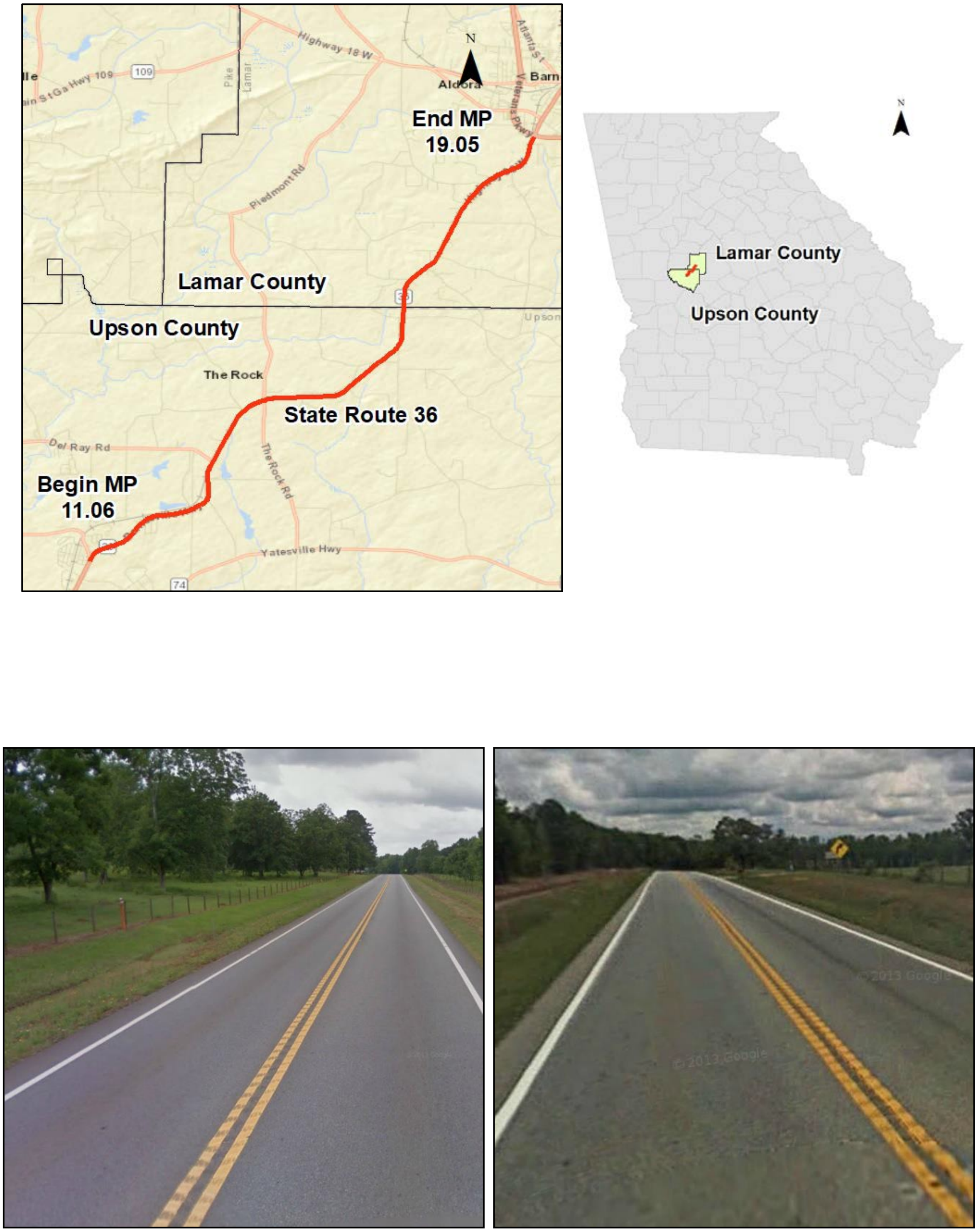


Figure 17. Project 0007077 SR 36, Section A Details (clockwise from top left): Map [22-24], Location within Georgia, MP 0 Facing West [4, 21], MP 10 Facing West [4, 21]

Table 31. Project 0007077, SR 36, Section A, Construction Details [4]

Source	Attribute	Detail
TransPI	Project ID	0007077
	Project Number	CSSTP-0007-00(077)
	Project Title	SR 36 FM SR 74 TO SR 7 & SR 36 FM SR 7 TO I-75
	Management Let	8/19/2005
	Project Completion	12/3/2009
	Project Manager	Scott Zehngraff
	Office	Traffic Safety & Design
	Project Type	Safety
	DOT District	3
	Congressional District	8
	Project Description	Indentation centerline rumble strips on SR 36 from East Main Street to Peach Blossom Trail
	Construction Contractor	Costello Industries, Incorporated
	MPO	Not Urban
Federal Report of Completed Projects	Construction Begin Date	1/17/2006
	Time Charges Stop Date	5/31/2006

Table 32. Project 0007077, SR 36, Section A, Study Details [4]

Attribute	Detail
Primary Roadway	Georgia State Route 36
Beginning Milepost	11.06
Ending Milepost	19.05
County (Begin)	Upson
County (End)	Upson
Length (mi)	8.05
RCLINK	1711003600 2931003600
Beginning Coordinates	32.920363,-84.28954
Ending Coordinates	33.037692,-84.165099
AADT MP 11.06 – 15.71 (Upson County)	2003 – 5,310 2004 – 5,310 2007 – 4,120 2008 – 3,700
AADT MP 15.72 – 19.11 (Upson County)	2003 – 5,030 2004 – 5,030 2007 – 4,910 2008 – 4,720

Project ID: 0007077, SR 36 Section B

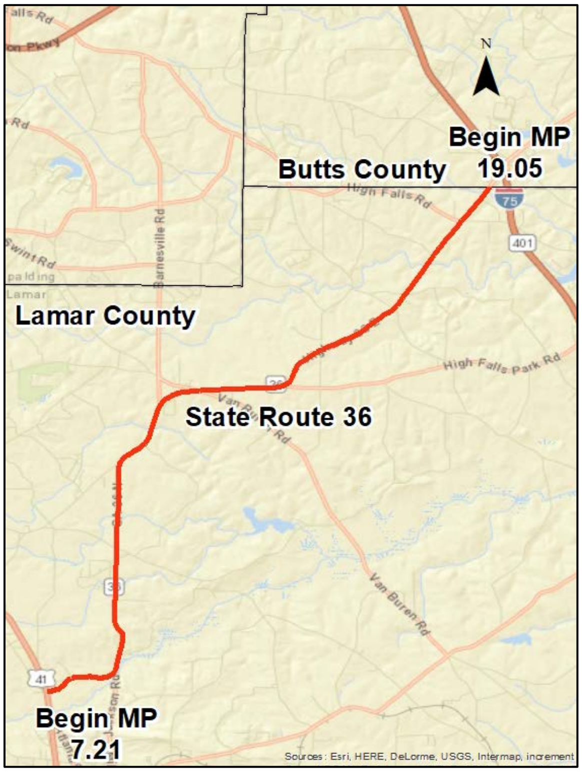


Figure 18. Project 0007077 SR 36, Section B Details (clockwise from top left): Map [22-24], Location within Georgia, MP 15 Facing West [4, 25], MP 8 Facing West [4, 25]

Table 33. Project 0007077, SR 36, Section B, Construction Details [4]

Source	Attribute	Detail
TransPI	Project ID	0007077
	Project Number	CSSTP-0007-00(077)
	Project Title	SR 36 FM SR 74 TO SR 7 & SR 36 FM SR 7 TO I-75
	Management Let	8/19/2005
	Project Completion	12/3/2009
	Project Manager	Scott Zehngraff
	Office	Traffic Safety & Design
	Project Type	Safety
	DOT District	3
	Congressional District	8
	Project Description	SR 36 from Highway 41 to I-75 in District 3
	Construction Contractor	Costello Industries, Incorporated
	MPO	Not Urban
Federal Report of Completed Projects	Construction Begin Date	1/17/2006
	Time Charges Stop Date	5/31/2006

Table 34. Project 0007077, SR 36, Section B, Study Details [4]

Attribute	Detail
Primary Roadway	Georgia State Route 36
Beginning Milepost	7.21
Ending Milepost	19.05
County (Begin)	Lamar
County (End)	Lamar
Length (mi)	11.84
RCLINK	1711003600
Beginning Coordinates	33.080741,-84.170817
Ending Coordinates	33.196695,-84.06902
AADT MP 7.17 – 13.5	2003 – 5,040 2004 – 5,930 2005 – 5,990 2006 – 6,800 2007 – 6,280 2008 – 5,870
AADT MP 13.51 – 16.82	2003 – 5,130 2004 – 6,630 2005 – 6,760 2006 – 4,4500 2007 – 5,460 2008 – 5,200
AADT MP 16.83 – 18.59	2003 – 5,420 2004 – 6,200 2005 – 5,770 2006 – 7,080 2007 – 7,060 2008 – 6,250
AADT MP 18.60 – 19.05	2003 – 6,060 2004 – 6,900 2005 – 7,120 2006 – 8,660 2007 – 7,640 2008 – 7,270

Project ID: 0007079, SR 136

Figure 19. Project 0007079 SR 136 Details (clockwise from top left): Map [22-24], Location within Georgia, MP 10 (Pickens County) Facing East Project [4, 25], MP 2 (Murray County) Facing East [4, 25]

Table 35. Project 0007079, SR 136, Construction Details [4]

Source	Attribute	Detail
TransPI	Project ID	0007079
	Project Number	CSSTP-0007-00(079)
	Project Title	SR 136 FROM SR 61/US 411 TO DAWSON COUNTY LINE
	Management Let	8/19/2005
	Project Completion	1/10/2007
	Project Manager	Scott Zehngraff
	Office	Traffic Safety & Design
	Project Type	Safety
	DOT District	6
	Congressional District	9
	Project Description	Indentation centerline rumble strips on SR 136 from SR 61/US 411 to Pickens County line in District 6
	Construction Contractor	-
	MPO	Not Urban
Federal Report of Completed Projects	Construction Begin Date	1/17/2006
	Time Charges Stop Date	5/31/2006

Table 36. Project 0007079, SR 136, Study Details [4]

Attribute	Detail
Primary Roadway	Georgia State Route 136
Beginning Milepost	23.56
Ending Milepost	19.71
County (Begin)	Gordon
County (End)	Pickens
Length (mi)	28.25
RCLINK	1231013600 1291013600 2131013600 2271013600
Beginning Coordinates	34.589751,-84.704502
Ending Coordinates	34.540836,-84.344769
AADT MP 23.56 – 24.07 (Gordon County)	2003 – 1,920 2004 – 2,030 2005 – 2,070 2006 – 1,980 2007 – 1,980 2008 – 1,860
AADT MP 0 – 2.82 (Murray County)	2003 – 2,400 2004 – 2,870 2005 – 2,520 2006 – 3,050 2007 – 2,990 2008 – 2,850
AADT MP 0 – 5.21 (Gilmer County)	2003 – 2,240 2004 – 2,650 2005 – 3,560 2006 – 2,950 2007 – 3,090 2008 – 2,940
AADT MP 0 – 3.66 (Pickens County)	2003 – 2,010 2004 – 2,340 2005 – 2,260 2006 – 2,050 2007 – 2,240 2008 – 2,130

Table 36 Continued on Next Page

Table 36 Continued

Attribute	Detail
AADT MP 3.67 – 6.31 (Pickens County)	2003 – 4,220 2004 – 3,870 2005 – 3,850 2006 – 3,920 2007 – 4,120 2008 – 3,850
AADT MP 6.32 – 7.24 (Pickens County)	2003 – 2,060 2004 – 1,830 2004 – 1,860 2004 – 1,870 2007 – 1,900 2008 – 1,840
AADT MP 7.25 – 12.00 (Pickens County)	2003 – 1,430 2004 – 1,420 2005 – 900 2006 – 1,110 2007 – 1,260 2008 – 1,100
AADT MP 12.02 – 14.13 (Pickens County)	2003 – 2,230 2004 – 1,690 2005 – 1,970 2006 – 1,760 2007 – 1,820 2008 – 1,730
AADT MP 12.02 – 14.13 (Pickens County)	2003 – 1,200 2004 – 680 2005 – 690 2006 – 650 2007 – 640 2008 – 1,050
AADT MP 17.96 – 19.71 (Pickens County)	2003 – 980 2004 – 510 2005 – 800 2006 – 560 2007 – 630 2008 – 600

APPENDIX B. PROTOCOLS

B.1 Crash Data Download, Handling and Storage Protocol

Background

Research projects related to traffic safety require access to data on crashes involving motor vehicles, bicycles and pedestrians. The primary source for these incident data are police reports prepared by the investigating officer. For the State of Georgia, images of these police reports are archived by the Georgia Department of Transportation (GDOT) and/or GDOT contractors (*Police Report Archive*). In addition to the *Police Report Archive*, GDOT and/or its contractors/collaborators extract data from these police reports and other data sources and place it a searchable database format to facilitate retrieval of important information for research and other purposes (*GDOT Crash Database*).

Although these are public records, both databases (*Police Report Archive* and *GDOT Crash Database*) contain sensitive personally-identifiable information that need to be protected from inadvertent release. Through its contracts with GDOT, the Georgia Transportation Institute (GTI) has access to, and uses, data from both databases. It is the policy of GTI that, whenever possible, all research projects requiring access to GTI crash data use “sanitized” versions of these databases with all sensitive personally-identifiable information redacted. Any application that cannot be conducted using the “sanitized” databases requires Georgia Tech IRB approval.

This protocol describes how GTI obtains updated versions of the *Police Report Archive* and *GDOT Crash Database* and subsequent handling and storage of these data. Procedures for “sanitizing” of these data for routine usage are described in a separate protocol.

Download

New and updated versions of *Police Report Archive* and *GDOT Crash Database* are obtained from GDOT via Secure FTP server owned and operated by GDOT or their contractor.

These downloads will be performed on a computer with an encrypted hard drive located in a protected location by an operator authorized by GDOT for database access.

Handling

Once downloaded, the files will be immediately copied and verified to two encrypted external hard drives. Following the copy process, files on the hard drive of the computer will be securely shredded using a software such as Eraser (http://eraser.heidi.ie/) before the operator leaves the protected download location.

After file shredding on the download computer the operator will immediately transfer the encrypted external hard drives to the secure storage location.

Storage

The encrypted hard drives will be stored in a locked cabinet or safe in a limited access area behind locked doors. Access to keys to the storage cabinet/safe containing the drives will be limited to GDOT authorized database users.

Passwords for the drives will never be stored in the storage location. Passwords will be available to select faculty and graduate students as approved by the GTI director. The passwords will be changed once every year.

B.2. Protocol for "Sanitizing" Crash Databases

Background

Research projects related to traffic safety require access to data on crashes involving motor vehicles, bicycles and pedestrians. The primary source for these incident data are police reports prepared by the investigating officer. For the State of Georgia, images of these police reports are archived by the Georgia Department of Transportation (GDOT) and/or GDOT contractors (*Police Report Archive*). In addition to the *Police Report Archive*, GDOT and/or its contractors/collaborators extract data from these police reports and other data sources and place it in a searchable database format to facilitate retrieval of important information for research and other purposes (*GDOT Crash Database*).

Although these are public records, both databases (*Police Report Archive* and *GDOT Crash Database*) contain sensitive personally-identifiable information that need to be protected from inadvertent release. Through its contracts with GDOT, the Georgia Transportation Institute (GTI) has access to, and uses, data from both databases. It is the policy of GTI that, whenever possible, all research projects requiring access to GTI crash data use "sanitized" versions of these databases with all sensitive personally-identifiable information redacted. Any application that cannot be conducted using the "sanitized" databases requires Georgia Tech IRB approval.

This protocol describes how GTI "sanitizes" the *Police Report Archive* and *GDOT Crash Database* to remove all personally-identifiable information. Procedures for obtaining, handling and storage of the un-sanitized data are described in a separate protocol.

"Sanitization" or Removal of Personally Identifiable Information

GDOT Crash Database

For this electronic database, a GDOT-authorized database user will remove any sensitive (e.g. Vehicle Identification Numbers (VIN)) or personally-identifiable (e.g. driver names or addresses) from the database and create a "sanitized" version for use in normal research applications. The crash ID will be replaced with a locally generated unique ID that has no link or relation to the crash ID. The sanitized copy will be made available to students and other researchers as necessary for analysis. Any research applications requiring access to personally-identifiable information will be conducted under and research protocol approved by the Georgia Tech Institutional Review Board (IRB).

Police Report Archives

The Police Report images always have personally identifiable information present and require specialized techniques for removal. These reports are in standard two-page format with supplemental pages provided on some reports (e.g. when multiple vehicles are involved or injuries have occurred). The first page of the report always contains certain personally-identifiable information. Personally-identifiable information may, or may not, be present in subsequent pages. Given the non-uniformity of the scanned reports, full automation of the sanitization process is challenging. The sanitization process involves the following steps:

1. The Police Report image files are renamed to replace the crash ID with the unique ID used in the sanitized database. Once this step is completed, the table containing the link between the crash ID and the unique ID is securely destroyed.

2. Each Police Report image file is converted to a series of images, each image representing one page. Each image is identified with the unique ID that allows it to be linked back to the sanitized database where other non-personally-identifiable information related to the crash is available.

3. The first page image is deleted as it contains personally identifiable information.

4. The second page image is verified to ensure proper orientation and inverted if necessary. Portions of the second page, where personally identifiable information is present, are then electronically blanked out. If the original record had only two pages the record is now sanitized.

5. If the record has more than two pages remaining page images are manually checked by the operator to identify any personally identifiable information. Any images containing personally identifiable information that are not relevant for research are deleted. If the image contains both research-relevant and personally-identifiable information, the image is manually cropped or blanked out to remove any personally identifiable information.

6. The sanitized images are then collected to form a “sanitized” Police Report database for usage in normal research applications. Any application requiring access to personally-identifiable information will be covered by a separate protocol approved by the Georgia Tech Institutional Review Board (IRB).

APPENDIX C. SQL CODE

This SQL code was used to extract data from the crash database in Microsoft Access and was written by Jerome Sin.

C.1 Treatment Sites: Before Period

```
SELECT *
FROM LOCATION_TBL
WHERE ((((LOCATION_TBL.LOC_RCLINK_IDENTIFIER) Like "0771001400")
AND ((LOCATION_TBL.LOC_ACC_MILELOG)Between
19.68 And 27.55) AND ((LOCATION_TBL.LOC_ACC_JULDT)Between
#1/1/2003# and #12/31/2004#))
OR (((LOCATION_TBL.LOC_RCLINK_IDENTIFIER) Like "0451001600") AND
((LOCATION_TBL.LOC_ACC_MILELOG)Between
16.69 And 27.87) AND ((LOCATION_TBL.LOC_ACC_JULDT)Between
#1/1/2003# and #12/31/2004#))
OR (((LOCATION_TBL.LOC_RCLINK_IDENTIFIER) Like "0771001600") AND
((LOCATION_TBL.LOC_ACC_MILELOG)Between
0.00 And 7.06) AND ((LOCATION_TBL.LOC_ACC_JULDT)Between
#1/1/2003# and #12/31/2004#))
OR (((LOCATION_TBL.LOC_RCLINK_IDENTIFIER) Like "0771015400") AND
((LOCATION_TBL.LOC_ACC_MILELOG)Between
0.11 And 7.60) AND ((LOCATION_TBL.LOC_ACC_JULDT)Between
#1/1/2003# and #12/31/2004#))
OR (((LOCATION_TBL.LOC_RCLINK_IDENTIFIER) Like "1171036900") AND
((LOCATION_TBL.LOC_ACC_MILELOG)Between
0.00 And 19.89) AND ((LOCATION_TBL.LOC_ACC_JULDT)Between
#1/1/2003# and #12/31/2004#))
OR (((LOCATION_TBL.LOC_RCLINK_IDENTIFIER) Like "0351004200") AND
((LOCATION_TBL.LOC_ACC_MILELOG)Between
0.00 And 7.97) AND ((LOCATION_TBL.LOC_ACC_JULDT) Between
#1/1/2003# and #12/31/2004#))

OR (((LOCATION_TBL.LOC_RCLINK_IDENTIFIER) Like "1511004200") AND
((LOCATION_TBL.LOC_ACC_MILELOG)Between
4.58 And 9.81) AND ((LOCATION_TBL.LOC_ACC_JULDT)Between
#1/1/2003# and #12/31/2004#))
OR (((LOCATION_TBL.LOC_RCLINK_IDENTIFIER) Like "0511020400") AND
((LOCATION_TBL.LOC_ACC_MILELOG)Between
0.00 And 8.14) AND ((LOCATION_TBL.LOC_ACC_JULDT)Between
#1/1/2003# and #12/31/2004#))
```

```
OR (((LOCATION_TBL.LOC_RCLINK_IDENTIFIER) Like "2931003600") AND
((LOCATION_TBL.LOC_ACC_MILELOG)Between
9.34 And 19.11) AND ((LOCATION_TBL.LOC_ACC_JULDT)Between
#1/1/2003# and #12/31/2004#))
OR (((LOCATION_TBL.LOC_RCLINK_IDENTIFIER) Like "1711003600") AND
((LOCATION_TBL.LOC_ACC_MILELOG)Between
0.00 And 4.10) AND ((LOCATION_TBL.LOC_ACC_JULDT)Between
#1/1/2003# and #12/31/2004#))
OR (((LOCATION_TBL.LOC_RCLINK_IDENTIFIER) Like "1711003600") AND
((LOCATION_TBL.LOC_ACC_MILELOG)Between
7.21 and 19.05) AND ((LOCATION_TBL.LOC_ACC_JULDT)Between
#1/1/2003# and #12/31/2004#))
OR (((LOCATION_TBL.LOC_RCLINK_IDENTIFIER) Like "1291013600") AND
((LOCATION_TBL.LOC_ACC_MILELOG)Between
23.56 And 24.00) AND ((LOCATION_TBL.LOC_ACC_JULDT)Between
#1/1/2003# and #12/31/2004#))
OR (((LOCATION_TBL.LOC_RCLINK_IDENTIFIER) Like "2131013600") AND
((LOCATION_TBL.LOC_ACC_MILELOG)Between
0.00 And 2.79) AND ((LOCATION_TBL.LOC_ACC_JULDT)Between
#1/1/2003# and #12/31/2004#))
OR (((LOCATION_TBL.LOC_RCLINK_IDENTIFIER) Like "1231013600") AND
((LOCATION_TBL.LOC_ACC_MILELOG)Between
0.00 And 5.15) AND ((LOCATION_TBL.LOC_ACC_JULDT)Between
#1/1/2003# and #12/31/2004#))
OR (((LOCATION_TBL.LOC_RCLINK_IDENTIFIER) Like "2271013600") AND
((LOCATION_TBL.LOC_ACC_MILELOG)Between
0.00 And 19.71) AND ((LOCATION_TBL.LOC_ACC_JULDT)Between
#1/1/2003# and #12/31/2004#)));
```

C.2 Treatment Sites: After Period

```
SELECT *
FROM LOCATION_TBL
WHERE ((((LOCATION_TBL.LOC_RCLINK_IDENTIFIER) Like "0771001400")
AND ((LOCATION_TBL.LOC_ACC_MILELOG)Between
19.68 And 27.55) AND ((LOCATION_TBL.LOC_ACC_JULDT)Between
#1/1/2007# and #12/31/2008#))
OR (((LOCATION_TBL.LOC_RCLINK_IDENTIFIER) Like "0451001600") AND
((LOCATION_TBL.LOC_ACC_MILELOG)Between
16.69 And 27.87) AND ((LOCATION_TBL.LOC_ACC_JULDT)Between
#1/1/2007# and #12/31/2008#))
OR (((LOCATION_TBL.LOC_RCLINK_IDENTIFIER) Like "0771001600") AND
((LOCATION_TBL.LOC_ACC_MILELOG)Between
0.00 And 7.06) AND ((LOCATION_TBL.LOC_ACC_JULDT)Between
```

```
#1/1/2007# and #12/31/2008#))
OR (((LOCATION_TBL.LOC_RCLINK_IDENTIFIER) Like "0771015400") AND
((LOCATION_TBL.LOC_ACC_MILELOG)Between
0.11 And 7.60) AND ((LOCATION_TBL.LOC_ACC_JULDT)Between
#1/1/2007# and #12/31/2008#))
OR (((LOCATION_TBL.LOC_RCLINK_IDENTIFIER) Like "1171036900") AND
((LOCATION_TBL.LOC_ACC_MILELOG)Between
0.00 And 19.89) AND ((LOCATION_TBL.LOC_ACC_JULDT)Between
#1/1/2007# and #12/31/2008#))
OR (((LOCATION_TBL.LOC_RCLINK_IDENTIFIER) Like "0351004200") AND
((LOCATION_TBL.LOC_ACC_MILELOG)Between
0.00 And 7.97) AND ((LOCATION_TBL.LOC_ACC_JULDT)Between
#1/1/2007# and #12/31/2008#))
OR (((LOCATION_TBL.LOC_RCLINK_IDENTIFIER) Like "1511004200") AND
((LOCATION_TBL.LOC_ACC_MILELOG)Between
4.58 And 9.81) AND ((LOCATION_TBL.LOC_ACC_JULDT)Between
#1/1/2007# and #12/31/2008#))
OR (((LOCATION_TBL.LOC_RCLINK_IDENTIFIER) Like "0511020400") AND
((LOCATION_TBL.LOC_ACC_MILELOG)Between
0.00 And 8.14) AND ((LOCATION_TBL.LOC_ACC_JULDT)Between
#1/1/2007# and #12/31/2008#))
OR (((LOCATION_TBL.LOC_RCLINK_IDENTIFIER) Like "2931003600") AND
((LOCATION_TBL.LOC_ACC_MILELOG)Between
9.34 And 19.11) AND ((LOCATION_TBL.LOC_ACC_JULDT)Between
#1/1/2007# and #12/31/2008#))
OR (((LOCATION_TBL.LOC_RCLINK_IDENTIFIER) Like "1711003600") AND
((LOCATION_TBL.LOC_ACC_MILELOG)Between
0.00 And 4.10) AND ((LOCATION_TBL.LOC_ACC_JULDT)Between
#1/1/2007# and #12/31/2008#))
OR (((LOCATION_TBL.LOC_RCLINK_IDENTIFIER) Like "1711003600") AND
((LOCATION_TBL.LOC_ACC_MILELOG)Between
7.21 and 19.05) AND ((LOCATION_TBL.LOC_ACC_JULDT)Between
#1/1/2007# and #12/31/2008#))
OR (((LOCATION_TBL.LOC_RCLINK_IDENTIFIER) Like "1291013600") AND
((LOCATION_TBL.LOC_ACC_MILELOG)Between
23.56 And 24.00) AND ((LOCATION_TBL.LOC_ACC_JULDT)Between
#1/1/2007# and #12/31/2008#))
OR (((LOCATION_TBL.LOC_RCLINK_IDENTIFIER) Like "2131013600") AND
((LOCATION_TBL.LOC_ACC_MILELOG)Between
0.00 And 2.79) AND ((LOCATION_TBL.LOC_ACC_JULDT)Between
#1/1/2007# and #12/31/2008#))
OR (((LOCATION_TBL.LOC_RCLINK_IDENTIFIER) Like "1231013600") AND
((LOCATION_TBL.LOC_ACC_MILELOG)Between
0.00 And 5.15) AND ((LOCATION_TBL.LOC_ACC_JULDT)Between
#1/1/2007# and #12/31/2008#))
OR (((LOCATION_TBL.LOC_RCLINK_IDENTIFIER) Like "2271013600") AND
((LOCATION_TBL.LOC_ACC_MILELOG)Between
```

0.00 And 19.71) AND ((LOCATION_TBL.LOC_ACC_JULDT) Between #1/1/2007# and #12/31/2008#)));

APPENDIX D. PERL SCRIPT

This code was used to join the crash database to the roadway database and was written by Dr. Angshuman Guin.

Joining Treatment Sites to Treatment Roadways

```
use strict;

my $int_filter = 'non_int';
$int_filter  =  'int';
my $road_types = 'all';

#LOC_ACC_ID,LOC_ACC_JULDT,LOC_RCLINK_IDENTIFIER,LOC_ACC_MIL
ELOG,LOC_INTERROUTE_TY
#PE,LOC_AADT_COUNT,LOC_DIVHWYBARRIER_TYPE,LOC_DIVHWYMEDI
AN_ TYPE,LOC_FUNCTIONALCLA
#SS_TYPE,LOC_LANESLEFT_COUNT,LOC_LANESRIGHT_COUNT,Noof
Injured,No of Fatalities,
#Manner of Collision

my $header =
'COUNTY,ROUTE_TYPE,ROUTE_NUM,BEG_MEASURE,END_MEASURE,SECTI
ON_LENGTH,DESCRIPTION,DISTRICT,MAINT_AREA,POPULATION,INVENTO
RY_DATE,DESIGNATED_WAY,TRUCK_ROUTE,TRAVEL_WAY,RURAL_URA
N,SPEED_LIMIT,FAS_NUM,TRUCK_ROUTE_ID,CONGRESS_DIST,STATE_ROU
TE_SEQ,ACCESS_CONTROL,OPERATION,TOTAL_LANES,SPECIAL_CLASS,DI
V_HWY_SHLDR_WIDTH_LFT,DIV_HWY_SHLDR_TYPE_LFT,DIV_HWY_SURF
_WIDTH,DIV_HWY_SURF_TYPE,DIV_HWY_SHLDR_WIDTH_RT,DIV_HWY_S
HLDR_TYPE_RT,DIV_HWY_MEDIAN_WIDTH,DIV_HWY_MEDIAN_TYPE,DIV
_HWY_BARRIER_TYPE,UDIV_HWY_SHLDR_WIDTH_LFT,UDIV_HWY_SHLD
R_TYPE_LFT,UDIV_HWY_SURFACE_WIDTH,UDIV_HWY_SURFACE_TYPE,U
DIV_HWY_SHLDR_WIDTH_RT,UDIV_HWY_SHLDR_TYPE_RT,AUX_LANE_W
IDTH_LFT,AUX_LANE_TYPE_LFT,AUX_LANE_WIDTH_RT,AUX_LANE_TYPE
_RT,MAINT_YEAR,MAINT_TYPE,IMPROVE_YEAR,FUNC_CLASS,TRAFFIC_C
OUNT_TYPE,TRAFFIC_COUNT_YEAR,RIGHT_OF_WAY,RW_TYPE,TC_NUMB
ER,MAINTENANCE_SUR_DES,SIDEWALK_LEFT,SIDEWALK_RIGHT,IMPROV
E_TYPE,TRUCK_PERCENT,TRUCK_PERCENT_TYPE,SIGNAL,AADT_OLD,HP
MS_ID,PACES_RATING,AADT,INTERSECT_ROAD1,INTERSECT_ROAD2,S_FU
NCLASS_ID,DUAL_MAINT_RATING,ROAD_WIDTH,DIVIDED,OPEN_TO_TRAF
FIC,CITY_CODE,T_LANES_LEFT,T_LANES_RIGHT,LAND_DOMAIN,RCLINK,T
otal Crashes,Injury Crashes, Fatal Crashes,Injury with Fatality,Injury without
```

```
Fatality,Fatality without Injury, No Fatality or Injury,Headon,Sideswipe (opposite
dir),Angle,Property Only ';

my $filterf = 'CLRS_Site_Extents.csv';
open FIL, "<$filterf";

my $i = 0;
my @filters;
my $rc_string =
',';
while(<FIL>){
	chomp;
	next unless /^\d/;
	my ($RCLINK,$BEG_MEASURE,$END_MEASURE) = split /,/;
	($filters[$i][0], $filters[$i][1], $filters[$i][2]) =
($RCLINK,$BEG_MEASURE,$END_MEASURE);
	print "($filters[$i][0], $filters[$i][1],
$filters[$i][2])\n";
	$i++;
	$rc_string .= "$RCLINK,";
}
my $filter_size = $i-1;
#

my $baselinef = "2012_BaseLine_Road_Data";
$baselinef = "RC_Data_2007";
my @files = glob ("2*Crash*.txt");

open(BASE, "<$baselinef.txt");

my (%base,%beg_end,%end_beg,@rclinks,%rclink_h);
while (<BASE>) {
	chomp;
	my
($COUNTY,$ROUTE_TYPE,$ROUTE_NUM,$BEG_MEASURE,$END_MEASURE,
$ SECTION_LENGTH,$DESCRIPTION,$DISTRICT,$MAINT_AREA,$POPULATI
ON,
	$INVENTORY_DATE,$DESIGNATED_WAY,$TRUCK_ROUTE,$TRAVEL
_W AY,$RURAL_URAN,$SPEED_LIMIT,$FAS_NUM,$TRUCK_ROUTE_ID,
	$CONGRESS_DIST,$STATE_ROUTE_SEQ,$ACCESS_CONTROL,$OPERA
TION,$TOTAL_LANES,$SPECIAL_CLASS,$DIV_HWY_SHLDR_WIDTH_LFT,
	$DIV_HWY_SHLDR_TYPE_LFT,$DIV_HWY_SURF_WIDTH,$DIV_HWY_
S URF_TYPE,$DIV_HWY_SHLDR_WIDTH_RT,$DIV_HWY_SHLDR_TYPE_RT,
	$DIV_HWY_MEDIAN_WIDTH,$DIV_HWY_MEDIAN_TYPE,$DIV_HWY_
BA
RRIER_TYPE,$UDIV_HWY_SHLDR_WIDTH_LFT,$UDIV_HWY_SHLDR_TYPE_
L FT,
```

```
	$UDIV_HWY_SURFACE_WIDTH,$UDIV_HWY_SURFACE_TYPE,$UDIV
_H
WY_SHLDR_WIDTH_RT,$UDIV_HWY_SHLDR_TYPE_RT,$AUX_LANE_WIDT
H_L FT,
	$AUX_LANE_TYPE_LFT,$AUX_LANE_WIDTH_RT,$AUX_LANE_TYPE_
R T,$MAINT_YEAR,$MAINT_TYPE,$IMPROVE_YEAR,$FUNC_CLASS,
	$TRAFFIC_COUNT_TYPE,$TRAFFIC_COUNT_YEAR,$RIGHT_OF_WAY,
$RW_TYPE,$TC_NUMBER,$MAINTENANCE_SUR_DES,$SIDEWALK_LEFT,
	$SIDEWALK_RIGHT,$IMPROVE_TYPE,$TRUCK_PERCENT,$TRUCK_PE
RCENT_TYPE,$SIGNAL,$AADT_OLD,$HPMS_ID,$PACES_RATING,$AADT,
	$INTERSECT_ROAD1,$INTERSECT_ROAD2,$S_FUNCLASS_ID,$DUAL
_MAINT_RATING,$ROAD_WIDTH,$DIVIDED,$OPEN_TO_TRAFFIC,
	$CITY_CODE,$T_LANES_LEFT,$T_LANES_RIGHT,$LAND_DOMAIN,$
RCLINK) = split /,/;
	#my @fields = split /,/;


	unless ($road_types eq 'all'){
		next unless $rc_string =~ /,$RCLINK,/;
		next unless $RCLINK =~ /^\d\d\d1/;
		next unless $DIV_HWY_BARRIER_TYPE == 0 ;
		next unless $DIV_HWY_MEDIAN_TYPE == 0 ; next
	unless $T_LANES_LEFT == 1 ;
	next unless $T_LANES_RIGHT == 1 ;
	next unless ($FUNC_CLASS == 2 || $FUNC_CLASS == 6
|| $FUNC_CLASS == 7);
	}
	#print "$RCLINK\n"; my $skip = 1;
	foreach my $i (0..$filter_size){
		if ($RCLINK == $filters[$i][0]){
			$skip = 0;
			$skip = 1 unless (( ($filters[$i][1] >=
$BEG_MEASURE) && ($filters[$i][1] <= $END_MEASURE) )||
					( ($filters[$i][1] <=
$BEG_MEASURE) && ($filters[$i][2] >= $END_MEASURE) ) ||
					( ($filters[$i][2] >=
$BEG_MEASURE) && ($filters[$i][2] <= $END_MEASURE) ));#||
#if any of the ends of the RClink is within the CLRS section
#( ($filters[$i][1] >
$BEG_MEASURE) && ($filters[$i][2] < $END_MEASURE) );			#
if the CLRS section is completely within the RCLINK
print "$filters[$i][0]: ($filters[$i][1] <
$BEG_MEASURE) && ($filters[$i][2] > $END_MEASURE)\n"unless
$skip;
		}
	}
	next if $skip != 0;
```

```
	push @{$beg_end{$RCLINK}} , $BEG_MEASURE;
	push @{$end_beg{$RCLINK}} , $END_MEASURE;
	$base{$RCLINK}{$BEG_MEASURE} = $_;
	$rclink_h{$RCLINK} = 1;
}
my @rclinks = sort keys %rclink_h;


my @ext_head = (0..24);

close BASE;
open(ERR,">error.csv");

foreach my $file (@files){
	my
(%total,%fat,%inj,%microfilm,%iwf,%iwof,%fwoi,%nofi,%col2,% col5,%col1,%col6);
	print "processing $file...\n";
	open(IN, "<$file");
	while (<IN>) {
		#my
($LOC_ACC_ID,$LOC_ACC_JULDT,$LOC_RCLINK_IDENTIFIER,$LOC_ACC
_MILELOG,$LOC_INTERROUTE_TYPE,
		#
$LOC_AADT_COUNT,$LOC_DIVHWYBARRIER_TYPE,$LOC_DIVHWYMEDIA
N_T YPE,$LOC_FUNCTIONALCLASS_TYPE,
		#
$LOC_LANESLEFT_COUNT,$LOC_LANESRIGHT_COUNT,$No_of_Injured,$
No_of_Fatalities,$Manner_of_Collision) = split /,/;

	my
($LOC_ACC_ID,$LOC_ACC_JULDT,$LOC_RCLINK_IDENTIFIER,$LOC_CIT
Y_IDENTIFIER,$LOC_COUNTY_IDENTIFIER,

	$LOC_ROUTE_TYPE,$LOC_ROUTE_IDENTIFIER,$LOC_ROUTE_SUFFI
X,$LOC_ACC_MILELOG,$LOC_ACC_MILELOGCUM,

	$LOC_INTERROUTE_TYPE,$LOC_INTERROUTE_IDENTIFIER,$LOC_I
NTERROUTE_SUFFIX,$LOC_ACCESSCONTROL_TYPE,

	$LOC_AADT_COUNT,$LOC_AUXLANELEFT_TYPE,$LOC_AUXLANERI
GH T_TYPE,$LOC_AUXLANELEFT_WIDTH,

	$LOC_AUXLANERIGHT_WIDTH,$LOC_DIVHWYBARRIER_TYPE,$LOC
_D IVHWYMEDIAN_TYPE,$LOC_FEDELIG_TYPE,

	$LOC_FUNCTIONALCLASS_TYPE,$LOC_RURALURBAN_TYPE,$LOC_S
I GNAL_TYPE,$LOC_SPEEDLIMIT_NUMBER,
```

```
	$LOC_LANESLEFT_COUNT,$LOC_LANESRIGHT_COUNT,$LOC_LOCAT
E_DATE,$LOC_LOCATOR_IDENTIFIER,

$LOC_X,$LOC_Y,$Microfilm,$Accident_Number,$NCIC_Number,$Accident_County,
$Accident_Date,

	$Day_of_Week,$Accident_Time,$No_of_Vehicles,$No_of_Inj
ured,$No_of_Fatalities,

	$No_of_Occupants,$Inside_City,$Rd_of_Occurence,$Inters
ect_With,$EMS_Notify,$EMS_Arrival,

	$HSP_Arrival,$Citations,$First_Harmful_Evnt,$Traffic_F
low,$Weather,$Surface_Condition,

	$Light_Condition,$Manner_of_Collision,$Location_of_Imp
act,$Road_Defects,$Other_Damage,

	$Hit_N_Run,$Flag_02,$Flag_03,$Flag_04,$Work_Zone,$Last
_Update,$Supplemental,
		$Supp_Microfilm,$ACC_Num_Suffix) = split /,/;

		if ($int_filter eq 'non_int'){
			if ( $LOC_INTERROUTE_TYPE ) {
				next ;
			}
			if ( $LOC_INTERROUTE_IDENTIFIER ) {
				next ;
			}
			if ( $LOC_INTERROUTE_SUFFIX ) {
				next ;
			}
			if ( $LOC_INTERROUTE_TYPE ) {
				next unless $LOC_INTERROUTE_TYPE =~
/null/i;
			}
			if ( $LOC_INTERROUTE_IDENTIFIER ) {
				next unless $LOC_INTERROUTE_IDENTIFIER
=~ /null/i;
			}
			if ( $LOC_INTERROUTE_SUFFIX ) {
				next unless $LOC_INTERROUTE_SUFFIX =~
/null/i;
			}
		}
```

```
		unless ($beg_end{$LOC_RCLINK_IDENTIFIER}){
			#print ERR $_;
			next;
	}
	my @mps = @{$beg_end{$LOC_RCLINK_IDENTIFIER}} ;
	my @mps2 = @{$end_beg{$LOC_RCLINK_IDENTIFIER}} ;
		@mps = sort {$a<=>$b} @mps;
		@mps2 = sort {$a<=>$b} @mps2;
		$" = "\n";
		#print "$.\t$LOC_RCLINK_IDENTIFIER\t@mps\n";
<STDIN>;
	my $size = scalar(@mps);
		next unless $size;
	my ($beg,$end);
	foreach my $i (0..$size-2){
		if (($LOC_ACC_MILELOG >= $mps[$i]) &&
($LOC_ACC_MILELOG < $mps2[$i])) {
			$beg = $mps[$i];
			$end = $mps2[$i];

			#print
"$i,$LOC_RCLINK_IDENTIFIER,$LOC_ACC_MILELOG,$beg,$end\n";
<STDIN>;
				last;
		}

	}
	next unless $end;
	$total{$LOC_RCLINK_IDENTIFIER}{$beg}{$file} += 1 ;
	$microfilm{$LOC_RCLINK_IDENTIFIER}{$beg}{$file} .=
",$LOC_ACC_ID" ;
	$inj{$LOC_RCLINK_IDENTIFIER}{$beg}{$file} += 1 if
$No_of_Injured > 0;
	$fat{$LOC_RCLINK_IDENTIFIER}{$beg}{$file} += 1 if
$No_of_Fatalities > 0;
	$iwf{$LOC_RCLINK_IDENTIFIER}{$beg}{$file} += 1 if ($No_of_Injured
> 0 && $No_of_Fatalities > 0);
	$iwof{$LOC_RCLINK_IDENTIFIER}{$beg}{$file} += 1 if ($No_of_Injured
> 0 && $No_of_Fatalities == 0);
	$fwoi{$LOC_RCLINK_IDENTIFIER}{$beg}{$file} += 1 if ($No_of_Injured
== 0 && $No_of_Fatalities > 0);
	$nofi{$LOC_RCLINK_IDENTIFIER}{$beg}{$file} += 1 if ($No_of_Injured
== 0 && $No_of_Fatalities == 0);
	$col2{$LOC_RCLINK_IDENTIFIER}{$beg}{$file} += 1 if
$Manner_of_Collision == 2;
	$col5{$LOC_RCLINK_IDENTIFIER}{$beg}{$file} += 1 if
$Manner_of_Collision == 5;
```

```
	$col1{$LOC_RCLINK_IDENTIFIER}{$beg}{$file} += 1 if
$Manner_of_Collision == 1;
	$col6{$LOC_RCLINK_IDENTIFIER}{$beg}{$file} += 1 if
$Manner_of_Collision == 6;
	}

	close IN;
	die unless open(OUT,
">$file.20131203_2007RC.clrs.$int_filter.$road_types.csv");
	$" = ',';
	print OUT "$header,@ext_head\n";
	foreach my $rclink (@rclinks){
		my @beg = keys %{$base{$rclink}};
		foreach my $mp (@beg){
			print OUT
"$base{$rclink}{$mp},$total{$rclink}{$mp}{$file},$inj{$rclink}{$mp}{$file},$fat{$r
clink}{$mp}{$file},$iwf{$rclink}{$mp}{$file},$iwof{$rclink}{$mp}{$file},$fwoi{$r
clink}{$mp}{$fi le},$nofi{$rclink}{$mp}{$file},$col2{$rclink}{$mp}{$file},$
col5{$rclink}{$mp}{$file},$col1{$rclink}{$mp}{$file},$col6{$rclink}{$mp}{$file},$
microfilm{$rclink}{$mp}{$file}\n";
		}
	}
	close OUT;
}
#my @rclinks = keys %base;
#@rclinks = sort @rclinks; close ERR;

_END________

perl join_crash_to_segment_clrs.v2.pl perl
join_crash_to_segment_non_clrs.v2.pl
```

Joining Reference sites to Reference Roadways

```
use strict;


my $int_filter = 'non_int';
$int_filter = 'int';

my $road_types = 'all';

#LOC_ACC_ID,LOC_ACC_JULDT,LOC_RCLINK_IDENTIFIER,LOC_ACC_MIL
ELOG,LOC_INTERROUTE_TY
```

```
#PE,LOC_AADT_COUNT,LOC_DIVHWYBARRIER_TYPE,LOC_DIVHWYMEDI
AN_ TYPE,LOC_FUNCTIONALCLA
#SS_TYPE,LOC_LANESLEFT_COUNT,LOC_LANESRIGHT_COUNT,Noof
Injured,No of Fatalities,
#Manner of Collision

my $header =
'COUNTY,ROUTE_TYPE,ROUTE_NUM,BEG_MEASURE,END_MEASURE,SECTI
ON_LENGTH,DESCRIPTION,DISTRICT,MAINT_AREA,POPULATION,INVENTO
RY_DATE,DESIGNATED_WAY,TRUCK_ROUTE,TRAVEL_WAY,RURAL_URA
N,SPEED_LIMIT,FAS_NUM,TRUCK_ROUTE_ID,CONGRESS_DIST,STATE_ROU
TE_SEQ,ACCESS_CONTROL,OPERATION,TOTAL_LANES,SPECIAL_CLASS,DI
V_HWY_SHLDR_WIDTH_LFT,DIV_HWY_SHLDR_TYPE_LFT,DIV_HWY_SURF
_WIDTH,DIV_HWY_SURF_TYPE,DIV_HWY_SHLDR_WIDTH_RT,DIV_HWY_S
HLDR_TYPE_RT,DIV_HWY_MEDIAN_WIDTH,DIV_HWY_MEDIAN_TYPE,DIV
_HWY_BARRIER_TYPE,UDIV_HWY_SHLDR_WIDTH_LFT,UDIV_HWY_SHLD
R_TYPE_LFT,UDIV_HWY_SURFACE_WIDTH,UDIV_HWY_SURFACE_TYPE,U
DIV_HWY_SHLDR_WIDTH_RT,UDIV_HWY_SHLDR_TYPE_RT,AUX_LANE_W
IDTH_LFT,AUX_LANE_TYPE_LFT,AUX_LANE_WIDTH_RT,AUX_LANE_TYPE
_RT,MAINT_YEAR,MAINT_TYPE,IMPROVE_YEAR,FUNC_CLASS,TRAFFIC_C
OUNT_TYPE,TRAFFIC_COUNT_YEAR,RIGHT_OF_WAY,RW_TYPE,TC_NUMB
ER,MAINTENANCE_SUR_DES,SIDEWALK_LEFT,SIDEWALK_RIGHT,IMPROV
E_TYPE,TRUCK_PERCENT,TRUCK_PERCENT_TYPE,SIGNAL,AADT_OLD,HP
MS_ID,PACES_RATING,AADT,INTERSECT_ROAD1,INTERSECT_ROAD2,S_FU
NCLASS_ID,DUAL_MAINT_RATING,ROAD_WIDTH,DIVIDED,OPEN_TO_TRAF
FIC,CITY_CODE,T_LANES_LEFT,T_LANES_RIGHT,LAND_DOMAIN,RCLINK,T
otalCrashes,Injury Crashes, Fatal Crashes,Injury with Fatality,Injury without
Fatality,Fatality without Injury, No Fatality or Injury,Headon,Sideswipe (opposite
dir),Angle,Property Only ';


my $filterf = 'CLRS_Site_Extents.csv';
open FIL, "<$filterf";

my $i = 0;
my @filters;
my $rc_string = ',';
while(<FIL>){
	chomp;
	next unless /^\d/;
	my ($RCLINK,$BEG_MEASURE,$END_MEASURE) = split /,/;
	($filters[$i][0], $filters[$i][1], $filters[$i][2]) =
($RCLINK,$BEG_MEASURE,$END_MEASURE);
	$i++;
	$rc_string .= "$RCLINK,";
}
my $filter_size = $i-1;
```

```
#

my $baselinef = "2012_BaseLine_Road_Data";
$baselinef = "RC_Data_2007";
my @files = glob ("2*crash*.txt");

open(BASE, "<$baselinef.txt");

my (%base,%beg_end,%end_beg,@rclinks,%rclink_h); while
(<BASE>) {
	chomp;
	my
($COUNTY,$ROUTE_TYPE,$ROUTE_NUM,$BEG_MEASURE,$END_MEASURE,
$SECTION_LENGTH,$DESCRIPTION,$DISTRICT,$MAINT_AREA,$POPULATI
ON,
	$INVENTORY_DATE,$DESIGNATED_WAY,$TRUCK_ROUTE,$TRAVEL
_WAY,$RURAL_URAN,$SPEED_LIMIT,$FAS_NUM,$TRUCK_ROUTE_ID,
	$CONGRESS_DIST,$STATE_ROUTE_SEQ,$ACCESS_CONTROL,$OPERA
TION,$TOTAL_LANES,$SPECIAL_CLASS,$DIV_HWY_SHLDR_WIDTH_LFT,
	$DIV_HWY_SHLDR_TYPE_LFT,$DIV_HWY_SURF_WIDTH,$DIV_HWY_
SURF_TYPE,$DIV_HWY_SHLDR_WIDTH_RT,$DIV_HWY_SHLDR_TYPE_RT,
	$DIV_HWY_MEDIAN_WIDTH,$DIV_HWY_MEDIAN_TYPE,$DIV_HWY_
BARRIER_TYPE,$UDIV_HWY_SHLDR_WIDTH_LFT,$UDIV_HWY_SHLDR_TY
PE_LFT,
	$UDIV_HWY_SURFACE_WIDTH,$UDIV_HWY_SURFACE_TYPE,$UDIV
_H
WY_SHLDR_WIDTH_RT,$UDIV_HWY_SHLDR_TYPE_RT,$AUX_LANE_WIDT
H_LFT,
	$AUX_LANE_TYPE_LFT,$AUX_LANE_WIDTH_RT,$AUX_LANE_TYPE_
RT,$MAINT_YEAR,$MAINT_TYPE,$IMPROVE_YEAR,$FUNC_CLASS,
	$TRAFFIC_COUNT_TYPE,$TRAFFIC_COUNT_YEAR,$RIGHT_OF_WAY,$
RW_TYPE,$TC_NUMBER,$MAINTENANCE_SUR_DES,$SIDEWALK_LEFT,
	$SIDEWALK_RIGHT,$IMPROVE_TYPE,$TRUCK_PERCENT,$TRUCK_PE
RCENT_TYPE,$SIGNAL,$AADT_OLD,$HPMS_ID,$PACES_RATING,$AADT,
	$INTERSECT_ROAD1,$INTERSECT_ROAD2,$S_FUNCLASS_ID,$DUAL_
MAINT_RATING,$ROAD_WIDTH,$DIVIDED,$OPEN_TO_TRAFFIC,
	$CITY_CODE,$T_LANES_LEFT,$T_LANES_RIGHT,$LAND_DOMAIN,$R
CLINK) = split /,/;
	#
	#TWTL control station criteria
	#
	#next unless $ROUTE_TYPE == 1;
	unless ($road_types eq 'all'){
		next unless $RCLINK =~ /^\d\d\d1/;
		next unless $DIV_HWY_BARRIER_TYPE == 0 ;
		next unless $DIV_HWY_MEDIAN_TYPE == 0 ;
		next unless $T_LANES_LEFT == 1 ;
```

```
			next unless $T_LANES_RIGHT == 1 ;
			next unless ($FUNC_CLASS == 2 || $FUNC_CLASS == 6
|| $FUNC_CLASS == 7);
		}
		#my @fields = split /,/;
		if ($rc_string =~ /,$RCLINK,/){
			my $skip ;

			foreach my $i (0..$filter_size){
				if ($RCLINK == $filters[$i][0]){
					#$skip =0;
					$skip = 1 if ( ($filters[$i][1] >=
$BEG_MEASURE) && ($filters[$i][1] <= $END_MEASURE) )||
							( ($filters[$i][1] <=
$BEG_MEASURE) && ($filters[$i][2] >= $END_MEASURE) ) ||
							( ($filters[$i][2] >=
$BEG_MEASURE) && ($filters[$i][2] <= $END_MEASURE) ) ;#||
#if any of the ends of the RClink is within the CLRS section
							#( ($filters[$i][1] >
$BEG_MEASURE) && ($filters[$i][2] < $END_MEASURE) );		# if the CLRS
section is completely within the RCLINK
				}
			}
			next if $skip;
		}
		push @{$beg_end{$RCLINK}} , $BEG_MEASURE;
		push @{$end_beg{$RCLINK}} , $END_MEASURE;
		$base{$RCLINK}{$BEG_MEASURE} = $_;
		$rclink_h{$RCLINK} = 1;
}
my @rclinks = sort keys %rclink_h;

my @ext_head = (0..24);

close BASE;
open(ERR,">error.csv");

foreach my $file (@files){
		my
(%total,%fat,%inj,%microfilm,%iwf,%iwof,%fwoi,%nofi,%col2,% col5,%col1,%col6);
		print "processing $file...\n";
		open(IN, "<$file");
		while (<IN>) {
```

```
			#my
($LOC_ACC_ID,$LOC_ACC_JULDT,$LOC_RCLINK_IDENTIFIER,$LOC_ACC_
MILELOG,$LOC_INTERROUTE_TYPE,
			#
$LOC_AADT_COUNT,$LOC_DIVHWYBARRIER_TYPE,$LOC_DIVHWYMEDI
AN_T YPE,$LOC_FUNCTIONALCLASS_TYPE,
			#
$LOC_LANESLEFT_COUNT,$LOC_LANESRIGHT_COUNT,$No_of_Injured,$No
_of_Fatalities,$Manner_of_Collision) = split /,/;

			My
($LOC_ACC_ID,$LOC_ACC_JULDT,$LOC_RCLINK_IDENTIFIER,$LOC_CITY_
IDENTIFIER,$LOC_COUNTY_IDENTIFIER,

	$LOC_ROUTE_TYPE,$LOC_ROUTE_IDENTIFIER,$LOC_ROUTE_SUFFI
X,$LOC_ACC_MILELOG,$LOC_ACC_MILELOGCUM,

	$LOC_INTERROUTE_TYPE,$LOC_INTERROUTE_IDENTIFIER,$LOC_IN
TERROUTE_SUFFIX,$LOC_ACCESSCONTROL_TYPE,

	$LOC_AADT_COUNT,$LOC_AUXLANELEFT_TYPE,$LOC_AUXLANER
IGH T_TYPE,$LOC_AUXLANELEFT_WIDTH,

	$LOC_AUXLANERIGHT_WIDTH,$LOC_DIVHWYBARRIER_TYPE,$LO
C_DIVHWYMEDIAN_TYPE,$LOC_FEDELIG_TYPE,

	$LOC_FUNCTIONALCLASS_TYPE,$LOC_RURALURBAN_TYPE,$LOC_
SIGNAL_TYPE,$LOC_SPEEDLIMIT_NUMBER,

	$LOC_LANESLEFT_COUNT,$LOC_LANESRIGHT_COUNT,$LOC_LOCA
TE_DATE,$LOC_LOCATOR_IDENTIFIER,

    $LOC_X,$LOC_Y,$Microfilm,$Accident_Number,$NCIC_Number,$Accident_C
ounty,$Accident_Date,

	$Day_of_Week,$Accident_Time,$No_of_Vehicles,$No_of_Injured,$No_of_Fa
talities,

	$No_of_Occupants,$Inside_City,$Rd_of_Occurence,$Inters
ect_With,$EMS_Notify,$EMS_Arrival,

	$HSP_Arrival,$Citations,$First_Harmful_Evnt,$Traffic_Flow,$Weather,$Surfa
ce_Condition,

	$Light_Condition,$Manner_of_Collision,$Location_of_Impact,$Road_Defects,
$Other_Damage,
```

```
	$Hit_N_Run,$Flag_02,$Flag_03,$Flag_04,$Work_Zone,$Last_Update,$Suppl
emental,
		$Supp_Microfilm,$ACC_Num_Suffix) = split /,/;

		if ($int_filter eq 'non_int'){
			if ( $LOC_INTERROUTE_TYPE ) {
				next ;
			}
			if ( $LOC_INTERROUTE_IDENTIFIER ) {
				next ;
			}
			if ( $LOC_INTERROUTE_SUFFIX ) {
				next ;
			}

			if ( $LOC_INTERROUTE_TYPE ) {
				next unless $LOC_INTERROUTE_TYPE =~/null/i;
			}
			if ( $LOC_INTERROUTE_IDENTIFIER ) {
				next unless $LOC_INTERROUTE_IDENTIFIER
=~/null/i;
			}
			if ( $LOC_INTERROUTE_SUFFIX ) {
				next unless $LOC_INTERROUTE_SUFFIX =~ /null/i;
			}
			if ( $LOC_INTERROUTE_SUFFIX ) {
				next unless $LOC_INTERROUTE_SUFFIX =~/null/i;
			}
		}
		unless ($beg_end{$LOC_RCLINK_IDENTIFIER}){
			#print ERR $_;
			next;
		}
		my @mps = @{$beg_end{$LOC_RCLINK_IDENTIFIER}} ;
		my @mps2 = @{$end_beg{$LOC_RCLINK_IDENTIFIER}} ;
			@mps = sort {$a<=>$b} @mps;
			@mps2 = sort {$a<=>$b} @mps2;
			$" = "\n";
			print "$.\t$LOC_RCLINK_IDENTIFIER\t@mps\n" if
$LOC_RCLINK_IDENTIFIER == 1391005300;
		my $size = scalar(@mps);
			next unless $size;
		my ($beg,$end);
		foreach my $i (0..($size-1)){
			next unless $mps[$i+1];
```

```
			die
"LOC_RCLINK_IDENTIFIER,$mps[$i],$mps[$i+1],@mps" if $mps[$i] > $mps[$i+1];
			if (($LOC_ACC_MILELOG >= $mps[$i]) &&
($LOC_ACC_MILELOG < $mps2[$i])) {
				$beg = $mps[$i];
				$end = $mps2[$i];
				last;
			}

		}
			next unless $end;
			print "$beg,$LOC_ACC_MILELOG,$end\n"		if
$LOC_RCLINK_IDENTIFIER == 1391005300;
		$total{$LOC_RCLINK_IDENTIFIER}{$beg}{$file} += 1 ;
		$microfilm{$LOC_RCLINK_IDENTIFIER}{$beg}{$file} .=
",$LOC_ACC_ID" ;
		$inj{$LOC_RCLINK_IDENTIFIER}{$beg}{$file} += 1 if
$No_of_Injured > 0;
		$fat{$LOC_RCLINK_IDENTIFIER}{$beg}{$file} += 1 if
$No_of_Fatalities > 0;
		$iwf{$LOC_RCLINK_IDENTIFIER}{$beg}{$file} += 1 if
($No_of_Injured > 0 && $No_of_Fatalities > 0);
		$iwof{$LOC_RCLINK_IDENTIFIER}{$beg}{$file} += 1 if
($No_of_Injured > 0 && $No_of_Fatalities == 0);
		$fwoi{$LOC_RCLINK_IDENTIFIER}{$beg}{$file} += 1 if
($No_of_Injured == 0 && $No_of_Fatalities > 0);
		$nofi{$LOC_RCLINK_IDENTIFIER}{$beg}{$file} += 1 if
($No_of_Injured == 0 && $No_of_Fatalities == 0);
		$col2{$LOC_RCLINK_IDENTIFIER}{$beg}{$file} += 1 if
$Manner_of_Collision == 2;
		$col5{$LOC_RCLINK_IDENTIFIER}{$beg}{$file} += 1 if
$Manner_of_Collision == 5;
		$col1{$LOC_RCLINK_IDENTIFIER}{$beg}{$file} += 1 if
$Manner_of_Collision == 1;
		$col6{$LOC_RCLINK_IDENTIFIER}{$beg}{$file} += 1 if
$Manner_of_Collision == 6;
	}

	close IN;
	die unless open(OUT,
">$file.20131203_2007RC.non_clrs.$int_filter.$road_types.csv");
	$" = ',';
	print OUT "$header,@ext_head\n";
	foreach my $rclink (@rclinks){
		my @beg = keys %{$base{$rclink}};
		foreach my $mp (@beg){
```

```
                              print OUT
"$base{$rclink}{$mp},$total{$rclink}{$mp}{$file},$inj{$rcli
nk}{$mp}{$file},$fat{$rclink}{$mp}{$file},$iwf{$rclink}{$mp}{$file},$iwof{$rclin
k}{$mp}{$file},$fwoi{$rclink}{$mp}{$file},$nofi{$rclink}{$mp}{$file},$col2{$rcli
nk}{$mp}{$file},$ col5{$rclink}{$mp}{$file},$col1{$rclink}{$mp}{$file},$col6{
$rclink}{$mp}{$file},$microfilm{$rclink}{$mp}{$file}\n";
                    }
          }
          close OUT;
}

#my @rclinks = keys %base;
#@rclinks = sort @rclinks;
close ERR;
```

____END_____

perl join_crash_to_segment_clrs.2007.pl
perljoin_crash_to_segment_non_clrs.2007.pl

APPENDIX E. MATLAB CODE

This code was used to extract AADT, begin MP and end MP for treatment and reference roadway segments and was written by Dr. Chengbo Ai.

```
clear all
close all

load data

% idxFile = {'cov2003','cov2004','cov2005','cov2006','cov2007','cov2008'};
cnt = height(segmentsneeded);
idx_dup = zeros(cnt,1);
record_dup = struct([]);
nRecord = 0;
for i = 1:cnt
    switch segmentsneeded.year(i)
        case 3
            cov = cov2003;
        case 4
            cov = cov2004;
        case 5
            cov = cov2005;
        case 6
            cov = cov2006;
        case 7
            cov = cov2007;
        case 8
            cov = cov2008;
    end
    idx = find(strcmp(segmentsneeded.LOC_RCLINK_IDENTIFIER{i}, cov.RCLINK));
    MP = [cov.BegMP(idx),cov.EndMP(idx)];
    idx_AADT = find(MP(:,1)<= segmentsneeded.LOC_ACC_MILELOG(i) & MP(:,2)
>= segmentsneeded.LOC_ACC_MILELOG(i));
    if ~isempty(idx_AADT)
        % If you want to check your AADT cov files, comment out this line;
        % then change your i value to the next
        idx_dup(i) = length(idx_AADT);
        if idx_dup(i)>1
            for j = 1:idx_dup(i)
                nRecord = nRecord+1;
                record_dup(nRecord).RCLINK =
segmentsneeded.LOC_RCLINK_IDENTIFIER{i};
```

```
            record_dup(nRecord).MP = segmentsneeded.LOC_ACC_MILELOG(i);
            record_dup(nRecord).covfile =
sprintf('cov200%d.csv',segmentsneeded.year(i));
            record_dup(nRecord).ID = idx(idx_AADT(j));
            record_dup(nRecord).MPFrom = cov.BegMP(record_dup(nRecord).ID);
            record_dup(nRecord).MPTo = cov.EndMP(record_dup(nRecord).ID);
         end
      end
      idx_AADT = idx_AADT(1);
      segmentsneeded.BeginMP(i) = cov.BegMP(idx(idx_AADT));
      segmentsneeded.EndMP(i) = cov.EndMP(idx(idx_AADT));
      segmentsneeded.AADT(i) = cov.AADT(idx(idx_AADT));
   else
      segmentsneeded.BeginMP(i) = -1;
      segmentsneeded.EndMP(i) = -1;
      segmentsneeded.AADT(i) = -1;
   end
   i
end
% record_dup includes the information about all the duplicated records
```

REFERENCES

[1] NHTSA, "2014 Traffic Safety Facts "Rural/Urban Comparison"," U. S. D. o. Transportation, Ed., ed. Washington, DC, 2016.

[2] GOHS. (2017). *Georgia Motor Vehicle Fatalities*. Available: http://www.gahighwaysafety.org/research/ga-crashes/injuries/fatalities/

[3] GDOT, "Safety Action Plan," G. D. o. Transportation, Ed., ed, 2005.

[4] J. G. N. Sin, "Safety Impact Study of Centerline Rumble Strips in Georgia," Master of Science, School of Civil and Environmental Engineering, Georgia Institute of Technology, Atlanta, GA, 2014.

[5] B. N. Persaud, R. A. Retting, and C. A. Lyon, "Crash reduction following installation of centerline rumble strips on rural two-lane roads," *Accident Analysis & Prevention,* vol. 36, pp. 1073-1079, 2004.

[6] F. Gross, B. Persaud, and C. Lyon, "A guide to developing quality crash modification factors," 2010.

[7] AASHTO, "Highway Safety Manual," vol. 2, 1st ed. Washington, DC: American Association of State Highway and Transportation Officials, 2010.

[8] B. Persaud, B. Lan, C. Lyon, and R. Bhim, "Comparison of empirical Bayes and full Bayes approaches for before–after road safety evaluations," *Accident Analysis & Prevention,* vol. 42, pp. 38-43, 2010.

[9] J. Kay, P. T. Savolainen, T. J. Gates, T. K. Datta, J. Finkelman, and B. Hamadeh, "Safety Impacts of a Statewide Centerline Rumble Strip Installation Program,"

Transportation Research Record: Journal of the Transportation Research Board, 2015.

[10] E. Hauer, *Observational before-after studies in road safety: estimating the effect of highway and traffic engineering measures on road safety*, 1st ed. Tarrytown, NY: Elsevier Science Inc., 1997.

[11] GDOT. (November 9, 2013). *Transportation Project Information (TransPI)*. Available: http://www.dot.ga.gov/BS/Projects/ProjectSearch

[12] J. Alexander and P. Garder, "Continued research on continuous rumble strips," Final Technical Report 94-4, Maine Department of Transportation1995.

[13] S. Herbel, L. Laing, and C. McGovern, *Highway Safety Improvement Program Manual*: US Department of Transportation, Federal Highway Administration, Office of Safety, 2010.

[14] A. Carriquiry and M. Pawlovich, "From empirical Bayes to full Bayes: methods for analyzing traffic safety data," *White Paper, Iowa State University*, 2004.

[15] W. Outcalt, "Centerline Rumble Strips, Report No. CDOT-DTD-R-2001-8, Interim Report," C. D. o. Transportation, Ed., Interim Report ed, 2001.

[16] B. Persaud, R. Retting, P. Garder, and D. Lord, "Safety effect of roundabout conversions in the united states: Empirical bayes observational before-after study," *Transportation Research Record: Journal of the Transportation Research Board*, pp. 1-8, 2001.

[17] D. W. Harwood, *Use of rumble strips to enhance safety*, 1993.

[18] R. Porter, E. Donnell, and K. Mahoney, "Evaluation of effects of centerline rumble strips on lateral vehicle placement and speed," *Transportation Research Record: Journal of the Transportation Research Board*, pp. 10-16, 2004.

[19] T. Sayed, P. deLeur, and J. Pump, "Impact of rumble strips on collision reduction on highways in British Columbia, Canada: Comprehensive before-and-after safety study," *Transportation Research Record: Journal of the Transportation Research Board*, pp. 9-15, 2010.

[20] D. E. Karkle, M. J. Rys, and E. R. Russell, "Safety effectiveness of centerline rumble strips in Kansas," *Journal of Transportation Safety & Security*, vol. 5, pp. 1-26, 2013.

[21] Google. (2013, Nov. 15, 2013). *Google Street View*. Available: http://maps.google.com

[22] ESRI. World Street Map [Online]. Available: https://services.arcgisonline.com/ArcGIS/rest/services/World_Street_Map/MapServer

[23] R. A. Division. State Highways [Online]. Available: https://arcgis.atlantaregional.com/arcgis/rest/services/OpenData/FeatureServer/36

[24] R. A. Division. Counties Georgia [Online]. Available: https://arcgis.atlantaregional.com/arcgis/rest/services/OpenData/FeatureServer/68

[25] GTI, "Crash Data Download, Handling and Storage Protocol," unpublished.

[26] GTI, "Protocol for Sanitizing Crash Databases," unpublished.

[27] E. R. Russell and M. J. Rys, "NCHRP Synthesis 339: Centerline Rumble Strips," *Transportation Research Record: Journal of the Transportation Research Board*, 2005.

[28] FHWA. *Planning Glossary* Available: www.fhwa.dot.gov/planning/glossary/.