2018-05-01

Intersection Safety Analysis Methodology for Utah Roadways

Joshua Daniel Gibbons
Brigham Young University

Intersection Safety Analysis Methodology for Utah Roadways

Joshua Daniel Gibbons

A thesis submitted to the faculty of
Brigham Young University
in partial fulfillment of the requirements for the degree of

Master of Science

Grant G. Schultz, Chair
Mitsuru Saito
Daniel P. Ames

Department of Civil and Environmental Engineering

Brigham Young University

ABSTRACT

Intersection Safety Analysis Methodology for Utah Roadways

Joshua Daniel Gibbons
Department of Civil and Environmental Engineering, BYU
Master of Science

Roadway safety continues to be a priority for the Utah Department of Transportation (UDOT) Traffic and Safety Division. UDOT has participated in and managed several research projects in recent years to determine the roadway segments of highest safety concern in the state. This research has provided UDOT with more tools to assist in safety project prioritization. Researchers in Department of Civil and Environmental Engineering at Brigham Young University (BYU) have worked with UDOT and the Statistics Department at BYU to create two network screening statistical tools called the Utah Crash Prediction Model (UCPM) and the Utah Crash Severity Model (UCSM) to analyze roadway segment safety. The Roadway Safety Analysis Methodology (RSAM) was developed as a process to run these segment models. Because a significant portion of crashes occur at intersections, there is a need to analyze roadway safety specifically at intersections.

This research focuses on the development of the Utah Intersection Crash Prediction Model (UICPM) and the Intersection Safety Analysis Methodology (ISAM). The UICPM is a Bayesian generalized linear model that determines crash distributions for each intersection based on roadway characteristics and historical crash data. The observed number of crashes at each intersection is compared with the crash distribution, and a percentile value is calculated as the probability that the number of crashes occurring at an intersection in a particular year is less than or equal to the average annual number of crashes. A high percentile value indicates that more crashes were observed than expected and the intersection is a hot spot and should be considered for safety improvements. All intersections are ranked at the state, UDOT Region, and county levels based on the percentile value, the higher ranks having higher percentile values.

The ISAM is the three-step process that was developed to execute the UICPM. The first step is to prepare the model input by formatting and combining the roadway characteristics and crash data files. Crashes are assigned to intersections if they fall with the functional area of an intersection. Due to data limitations, the ISAM is currently being used only for intersections of at least two state routes. It is anticipated that, as more data are made available, the ISAM will function properly for intersections of non-state routes as well. The second step is to execute the UICPM using the R GUI tool and R software. The third step is to create a two-page Intersection Safety Analysis Report (ISAR) for intersections of interest and maps of the state, UDOT Regions, and counties with the model results. Parts of the ISARs are auto-generated and the rest is entered manually by an analyst. The two-page ISARs will be used by UDOT Regions to prioritize intersection safety projects in their respective areas.

Keywords: intersection safety analysis, intersection functional area, highway safety research, UICPM, Numetric, crash analysis

ACKNOWLEDGEMENTS

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

1   INTRODUCTION

1.1   Background

Roadway safety continues to be a priority for the Utah Department of Transportation (UDOT) Traffic and Safety Division. This is evident in the continued campaign of "Zero Fatalities: A Goal We Can All Live With®" (UDOT 2017b). This initiative has helped raise awareness of fatal driver behaviors to achieve a goal of zero roadway fatalities in the state of Utah. UDOT has participated in and managed several research projects in recent years to determine the roadway segments of highest safety concern in the state. This research has provided UDOT with more tools to assist in safety project prioritization.

Researchers at Brigham Young University (BYU) have recently worked with UDOT to create two network screening statistical tools called the Utah Crash Prediction Model (UCPM) and the Utah Crash Severity Model (UCSM). The purpose of the UCPM is to create a crash distribution of the expected number of crashes on state roadway segments, compare the actual number of crashes to the distribution, and rank the roadway segments based on safety concern. The purpose of the UCSM is to create a distribution of the expected severity crash rate on state roadway segments, compare the actual severity crash rate to the distribution, and rank the roadway segments based on safety concern (Schultz et al. 2013c, Schultz et al. 2015). A process called the Roadway Safety Analysis Methodology (RSAM) was created to execute the roadway segment models and provide useful results to UDOT. The RSAM uses automated tools built in Microsoft (MS) Excel with Visual Basic for Applications (VBA) macros to prepare roadway and

crash data for the models, execute the models in R statistical analysis software, and generate reports for segments of interest (Schultz et al. 2016).

Crashes at intersections make up a large part of all the crashes that occur on roadways. Approximately 38 percent of all reported crashes in the state of Utah from 2010 to 2017 were reported as intersection-related. In addition, approximately 23 percent of fatal crashes in the same period were reported as intersection-related (Numetric 2018). The national crash data show similar numbers, with approximately 40 percent of the 5,811,000 crashes in the United States in 2008 being intersection-related (Choi 2010).

Safety at roadway intersections is critical to the safety of the network. For this reason, UDOT proposed to investigate intersection safety to complement the roadway segment safety research that was done previously (Schultz et al. 2013c, Schultz et al. 2015, Schultz et al. 2016). The purpose of this chapter is to provide the objectives of this research and to describe the organization of the research report.

1.2   Objectives

The primary objective of this research is to develop a Utah Intersection Crash Prediction Model (UICPM), similar to the UCPM developed in previous research. It is anticipated that the UICPM will assist UDOT in identifying intersection hotspots and in prioritizing intersection safety projects. The other objective of this research is to create a process similar to the RSAM that will process data for and execute the UICPM using MS Excel workbooks automated using VBA macros. This process works in conjunction with the RSAM and is called the Intersection Safety Analysis Methodology (ISAM).

1.3 Organization

The body of the report is organized into the following chapters:

- Chapter 1 includes an introduction to the research, project objectives, and the organization of the report.

- Chapter 2 includes a literature review of network screening safety statistical models, the RSAM process, the definition of an intersection area, available UDOT data sources, and other intersection safety programs used throughout the United States.

- Chapter 3 includes a general discussion on the processes in the ISAM, which is used to run the UICPM.

- Chapter 4 includes a discussion on the model input data preparation that is needed before running the UICPM.

- Chapter 5 includes a discussion on the development and execution of the UICPM.

- Chapter 6 includes a discussion on the creation of maps and reports that show the UICPM results in a useful way for UDOT employees to use.

- Chapter 7 includes a description of how a user might run the processes developed for the ISAM, including model input preparation, model execution, and report creation.

- Chapter 8 includes an analysis of the results of the UICPM including safety and crash trends.

- Chapter 9 includes the conclusions for this research project, recommendations for future research, and concluding remarks.

2 LITERATURE REVIEW

2.1 Overview

A literature review was performed to better understand roadway safety analysis at intersections and how such can be applied to roadways in Utah. This chapter contains a summary of that literature review with several discussions on key topics. The first topic is a discussion on crash severity and how that is defined by national and Utah authorities. Second, a discussion is provided on the network screening safety statistical models that have been developed in previous research efforts. Next is a discussion on the RSAM that was developed in previous research. The fourth topic is a discussion on the functional area of an intersection as defined in Transportation Research Board (TRB) literature. Fifth, a discussion is provided on the data sources available through UDOT and how those data sources can be accessed. The sixth topic is a discussion on intersection safety programs that have been done in other areas of the United States.

2.2 Crash Severity Definition

Roadway crashes are assigned a severity based on the most severe injury to any person involved in the crash. Crashes in Utah are assigned a number of 1 to 5 for severity (Numetric 2018), and the national severity rating is based on the "KABCO" letter naming system (NHTSA 2012). The numerical system for severity ratings is used in this research; however, each numerical severity rating has a corresponding letter rating in the "KABCO" system. The significance of the severities in each system are described in Table 2-1.

Table 2-1: Crash Severity Rating Descriptions (Numetric 2018 and NHTSA 2012)

Utah Numeric Scale	National Letter Scale	Severity Description
5	K	Fatal injury: injury that results in death within 30 days of crash
4	A	Suspected Serious Injury: serious injury not resulting in fatality; incapacitating injury results from the crash
3	B	Suspected Minor Injury: minor injury evident at the scene of the crash, not serious injury or fatality
2	C	Possible Injury: injuries reported but not evident at the scene of the crash
1	O	No Apparent Injury: the person received no bodily harm; property damage only (PDO)

2.3 Utah Network Screening Safety Statistical Models

Previous UDOT research has resulted in the creation of statistical models that identify the roadway segments in the state of highest safety concern. The two models that have been developed are the UCPM and the UCSM. The UCPM and UCSM are useful in determining roadway segments that are experiencing more crashes than expected and those that are experiencing more severe crashes than expected, respectively. The results of these models provide UDOT with a list of statistically-proven hot spot segments and reports for segments of interest. The purpose of this section is to briefly discuss the development and use of the UCPM and UCSM.

2.3.1 Utah Crash Prediction Model

The UCPM was developed as part of the Hot Spot Identification and Analysis methodology (Schultz et al. 2013c). The purpose of the UCPM is to develop a distribution of the

expected number of crashes on individual roadway segments and compare the median value of the distribution to the actual number of crashes. The UCPM process begins by segmenting roadway data and assigning crashes to each segment based on location. The combined roadway and crash data file is the input to the UCPM. The UCPM uses a regression analysis to develop a distribution of the expected number of crashes by segment based on parameters identified using the Bayesian horseshoe selection method. The actual number of crashes is plotted against the distribution of the expected number of crashes to make this comparison.

For each segment, a percentile value between 0 and 1 is produced that represents the probability that the number of crashes occurring on the segment is less than or equal to the actual observed number of crashes. A high percentile value indicates that a segment experiences more crashes than expected, and a low value shows that a segment experiences less crashes than expected. A percentile value of approximately 0.5 means that the median value of the expected number of crashes is equal to the observed number of crashes on the roadway segment. An example of the UCPM crash distribution and percentile value is shown in Figure 2-1 with the median value shown by the blue line and the actual number of crashes by the red dashed line. Segments are ranked at the state, UDOT Region, and county level based on the percentile value. A full discussion on the development and use of the UCPM is described in the literature (Schultz et al. 2013b, Schultz et al. 2013c, Schultz et al. 2015).
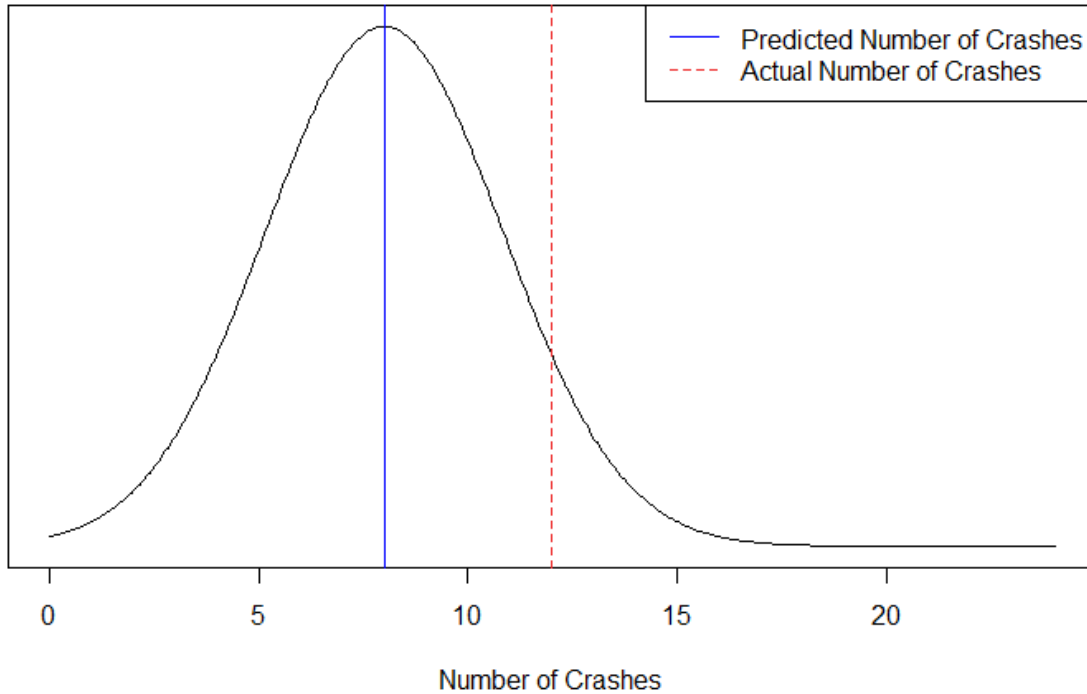
Figure 2-1: Example crash distribution produced by the UCPM (Schultz et al. 2016).

An analysis using the UCPM was completed in 2013 using roadway and crash data from 2008 to 2012. Only the non-incapacitating injury, incapacitating injury, and fatal crashes (i.e., crash severities 3, 4, and 5) severities were included in the analysis to emphasize severe crashes. The Bayesian horseshoe selection method was used to identify variables in the roadway characteristics that provided the best fit for the UCPM when used. Four variables were identified in this process: number of lanes, speed limit, total percent trucks, and vehicle-miles traveled (VMT). These variables are used are multiplied by coefficients in the regression equation to produce crash distribution values for each roadway segment. The top 20 roadway segments from a UCPM analysis completed in 2013 are listed in Table 2-2. As shown, the segments are ranked by the percentile value, which is very high for all segments shown. The analysis of these segments is discussed in detail in the literature (Schultz et al. 2013c, Schultz et al. 2015).

Table 2-2: Top 20 Segments in the 2013 UCPM Analysis (Schultz et al. 2013c)

Segment	Route Label	Beginning MP	End MP	UDOT Region	Percentile	Actual Crashes (Severities 3 to 5)	Model Predicted Crashes	Difference
1	0089P	388.438	389.123	1	1.000000	37	14	23
2	0015P	250.923	253.557	3	0.999989	28	11	17
3	0089P	415.425	415.994	1	0.999911	35	16	19
4	0015P	292.596	293.634	2	0.999733	25	11	14
5	0089P	369.036	369.532	2	0.999311	31	16	15
6	0089P	267.346	276.210	4	0.999144	17	6	11
7	0089P	386.955	388.438	1	0.998678	44	26	18
8	0089P	345.017	346.455	3	0.998622	34	18	16
9	0089P	431.317	433.164	1	0.998589	16	6	10
10	0068P	48.314	49.312	2	0.998567	39	22	17
11	0015P	296.093	297.314	2	0.998389	41	24	17
12	0015P	303.414	304.427	2	0.997989	30	16	14
13	0089P	335.590	336.030	3	0.997944	28	15	13
14	0015X	357.554	361.920	1	0.997600	23	11	12
15	0089P	347.360	347.664	3	0.996500	21	11	10
16	0015X	275.279	276.064	3	0.996278	26	14	12
17	0089P	349.471	350.056	3	0.996256	32	18	14
18	0015P	248.845	250.923	3	0.995800	13	5	8
19	0089P	386.346	386.801	1	0.995600	21	11	10
20	0089P	413.927	414.220	1	0.995211	17	8	9

2.3.2   Utah Crash Severity Model

The UCSM was developed in a later phase of the Hot Spot Identification and Analysis methodology research (Schultz et al. 2015). The purpose of the UCSM is to develop a distribution of severe crash rates for each segment and compare the distribution mean to the actual severe crash rate. The severe crash rate is calculated by dividing the number of severe crashes by the total number of crashes on a roadway segment. Severe crashes in this model are

crashes of severities 4 and 5. Roadway segments with a higher severe crash rate than expected are considered less safe and more prone to severe crashes. Roadway segments with a lower severe crash rate than expected are considered safer and less prone to severe crashes.

Similar to the UCPM process, the UCSM process begins by segmenting roadway data and assigning crashes to each segment based on location. The combined roadway and crash data file is the input to the UCSM. The UCSM creates a distribution of the severe crash rates for each roadway type based on parameters identified using the Bayesian horseshoe selection method. The output of the analysis is a database of the mean severe crash rates of the distribution for each segment, which is then compared with the actual severe crash rate observed on the segment. A percentile value is determined based on the deviation of the actual severe crash rate compared to the predicted severe crash rate. A similar ranking is performed for the UCSM results as is done for the UCPM results. However, the percentile value is multiplied by the difference in the mean of the severe crash rate distribution and the actual number of severe crashes. A full discussion on the development and use of the UCSM is described in the literature (Schultz et al. 2015).

The roadway and crash data from 2008 to 2012 were used to complete a UCSM analysis in 2015. The Bayesian horseshoe selection method identified the following parameters as variables that provided the best model fit: annual average daily traffic (AADT), number of lanes, speed limit, total percent trucks, and VMT. The top 20 roadway segments from the UCSM analysis are shown in Table 2-3. This UCSM analysis is discussed in detail in the literature (Schultz et al. 2015).

2.4   Roadway Safety Analysis Methodology

The RSAM was developed by a BYU research team for UDOT research in 2016 (Schultz et al. 2016). The purpose of developing the RSAM was to implement a series of user-friendly

graphical user interfaces (GUIs) and tools to accomplish the UCPM and UCSM analyses. The

process includes several automation tools that increase the speed and ease of use of the UCPM

and UCSM. The RSAM consists of three parts: (1) crash and roadway data preparation,

(2) statistical network screening, and (3) report compilation for segments of interest. The purpose

of this section is to discuss the RSAM processes and tools for each of the three parts.

Table 2-3: Top 20 Segments in the 2015 UCSM Analysis (Schultz et al. 2015)

Segment	Route Label	Beginning MP	End MP	UDOT Region	Probability (Severe Crash Occurrence)	Total Crashes (Severities 1 to 5)	Severe Crashes (Severities 4 and 5)	Expected Severe Crash Count	Difference
1	0080P	3.993	41.278	2	0.000	83	16	5.242	10.758
2	0068P	11.638	23.934	3	0.000	62	11	3.165	7.835
3	0006P	290.894	300.359	4	0.001	16	5	0.791	4.209
4	0015P	82.253	94.453	4	0.002	84	12	4.747	7.253
5	0173P	8.516	8.775	2	0.002	46	6	1.309	4.691
6	0080P	41.278	48.940	2	0.002	15	5	0.947	4.053
7	0134P	13.451	14.067	1	0.001	6	3	0.239	2.761
8	0048P	7.000	7.400	2	0.003	71	6	1.424	4.576
9	0071P	8.843	9.212	2	0.003	49	6	1.453	4.547
10	0039P	38.173	42.336	1	0.002	15	5	1.040	3.960
11	0089P	303.160	305.530	3	0.002	26	5	0.996	4.004
12	0006P	25.250	27.100	4	0.002	8	3	0.297	2.703
13	0191P	128.890	129.260	4	0.002	2	2	0.087	1.913
14	0089P	328.550	328.847	3	0.006	52	6	1.726	4.274
15	0089P	376.770	377.324	2	0.008	94	8	3.038	4.962
16	0089P	24.910	28.620	4	0.005	13	4	0.774	3.226
17	0080X	3.993	41.278	2	0.009	83	11	5.242	5.758
18	0092P	13.230	22.600	3	0.006	43	4	0.754	3.246
19	0111P	2.811	4.900	2	0.010	75	7	2.528	4.472
20	0089P	351.984	352.710	3	0.007	20	4	0.824	3.176

2.4.1 Crash and Roadway Data Preparation

The purpose of the crash and roadway data preparation process is to compile separate crash datasets into a single file and to segment roadway data into homogeneous roadway segments. This is done in preparation for the statistical network screening. The datasets that are required for this combination and segmentation process are shown in Table 2-4. The roadway data are found on the UDOT Open Data Portal website, where public UDOT data files are stored (UDOT 2017a). The crash data are obtained directly from UDOT, due to the confidential nature of the crash data. The data are protected under 23 USC 409 (USGPO 2012). A flowchart of the crash and roadway data preparation process is shown in Figure 2-2 (Schultz et al. 2016).

Table 2-4: UDOT Datasets required for the Statistical Network Screening

Roadway Data	Crash Data
AADT	(General) Crash Data
Functional Classification	Crash Location
Urban Code	Crash Rollup
Number of Lanes	Vehicle Crash Data
Speed Limit	

The process of combining and segmenting the data was automated into a single MS Excel spreadsheet. The spreadsheet uses macros and functions written in VBA code to perform tasks of summarizing the data. The main sheet of the spreadsheet contains a GUI to make the process user-friendly. The GUI contains buttons and inputs that call macros and functions used in the workbook. The ultimate output of the spreadsheet is a segmented roadway data file and a combined crash data file. These files are used as the inputs to the statistical network screening process. An image of the GUI of the crash and roadway data preparation spreadsheet is shown in

11

Figure 2-3. A thorough explanation of the data preparation process and the functionalities of the

data preparation spreadsheet is found in the literature (Schultz et al. 2016, Gibbons et al. 2016).
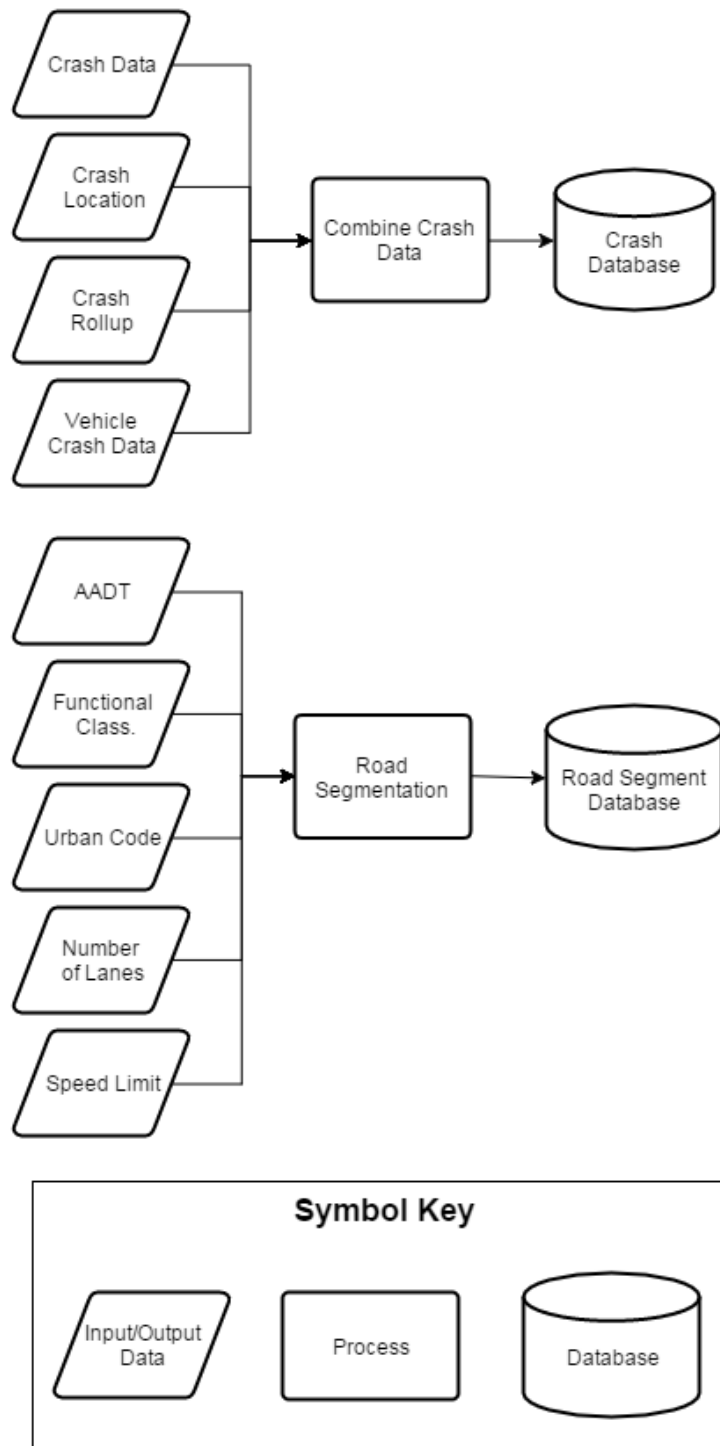
Figure 2-2: Flowchart of crash and roadway data preparation process (Schultz et al. 2016).
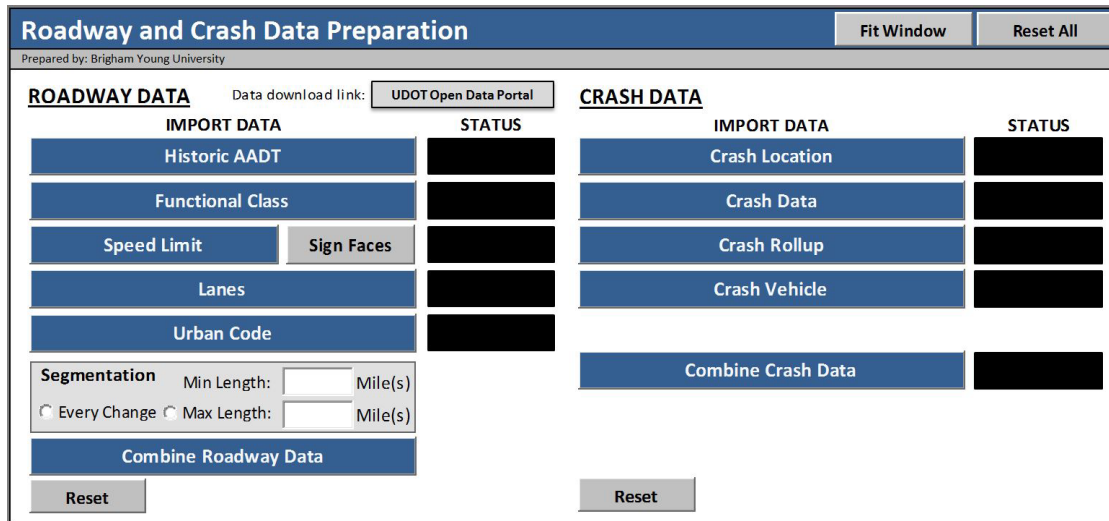
Figure 2-3: GUI of the data preparation spreadsheet (Gibbons et al. 2016).

2.4.2 Statistical Network Screening

The process of the statistical network screening is to use the roadway and crash data file as an input to the UCPM and UCSM and to create output tables and figures that are useful to the user. A flowchart of the statistical network screening process steps is shown in Figure 2-4. The UCPM and UCSM are coded in a statistical analysis software R (RPSC 2016). A GUI was created in MS Excel spreadsheet called the R Graphical User Interface (R GUI) to run the UCPM and UCSM. The R GUI contains macros and functions that automate the processes of inputting the roadway and crash data files and setting the parameters of the models. The user selects the input files and the variables to be used in the analysis, and the model is executed from the Excel-based R GUI. An image of the R GUI is shown in Figure 2-5 (Schultz et al. 2016).

Once the model run has been executed, output files are stored in a location on the local computer. The output data contains rankings for each segment at the state, UDOT Region, and county levels. To create a geographical representation of the data, geographical information systems (GIS) software is used to plot the data. A thorough explanation of the statistical network screening software is found in the literature (Schultz et al. 2016, Siegel et al. 2016).
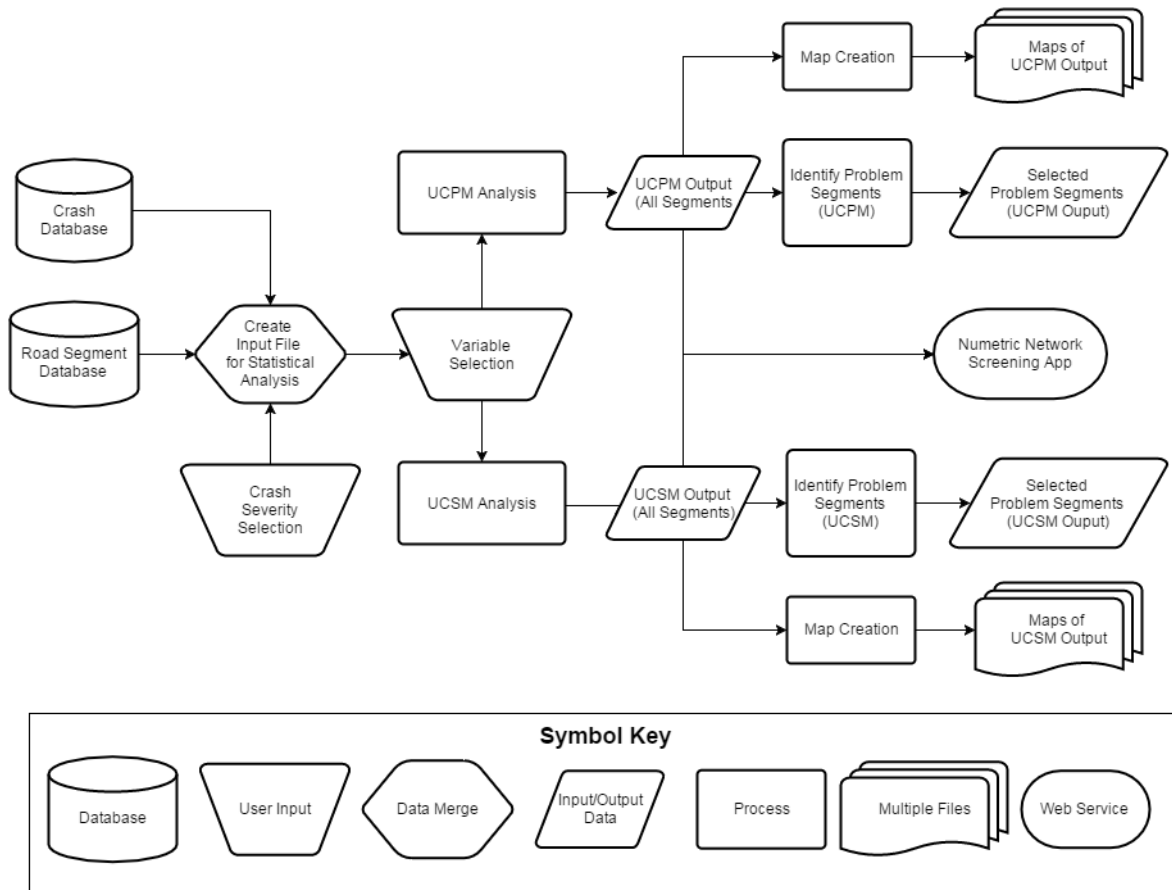
Figure 2-4: Flowchart of statistical network screening process (Schultz et al. 2016).
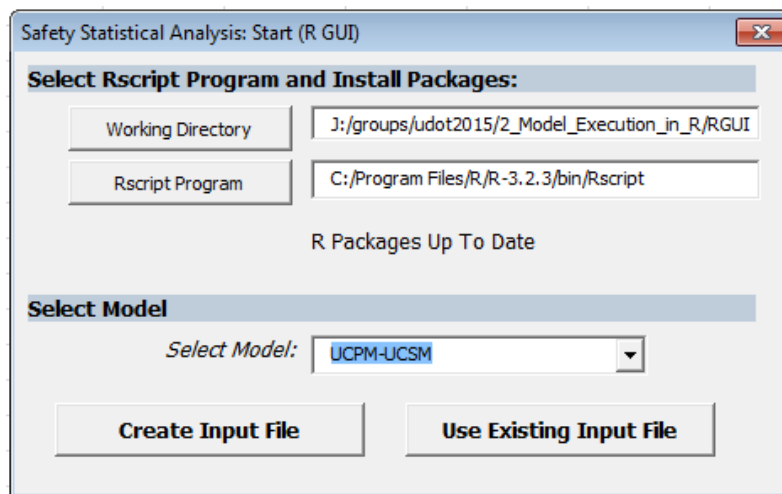
Figure 2-5: R GUI MS Excel workbook interface (Schultz et al. 2016).

2.4.3   Report Compilation for Segments of Interest

The purpose of the report compilation process is to create a useful Roadway Safety Analysis Report (RSAR) for segments of interest that will assist UDOT decision makers in planning projects. This process is completed after the statistical network screening is complete. A flowchart of the report compilation process is shown in Figure 2-6. The process begins by compiling additional roadway characteristics to provide a thorough roadway description in the report. The additional roadway datasets and their sources are shown in Table 2-5. These datasets are combined to the roadway segments file based on location. This is done in an MS Excel spreadsheet with VBA automation tools (Schultz et al. 2016).

The output of combining the roadway feature data is then used to create RSARs. This is done in another MS Excel spreadsheet called the "Report Compiler." The Report Compiler spreadsheet uses VBA macros to auto-populate a report format with the given roadway data. The first section of the report contains metadata of the roadway segment and its network screening analysis results. The second section contains basic segment functional characteristics such as functional classification, number of lanes, AADT, and speed limit. The third section contains additional roadway characteristics. The report also contains additional information regarding the crashes that occurred on the segment, including key crash factors along the corridor. The report is also auto-populated with common countermeasures for the given crash factors. An engineer or analyst must then make a site visit to analyze the location further and make notes of useful information regarding the site. The engineer or analyst may then narrow down the list of countermeasures to the ones that would improve the safety of the segment the most, based on engineering judgement (Schultz et al. 2016).
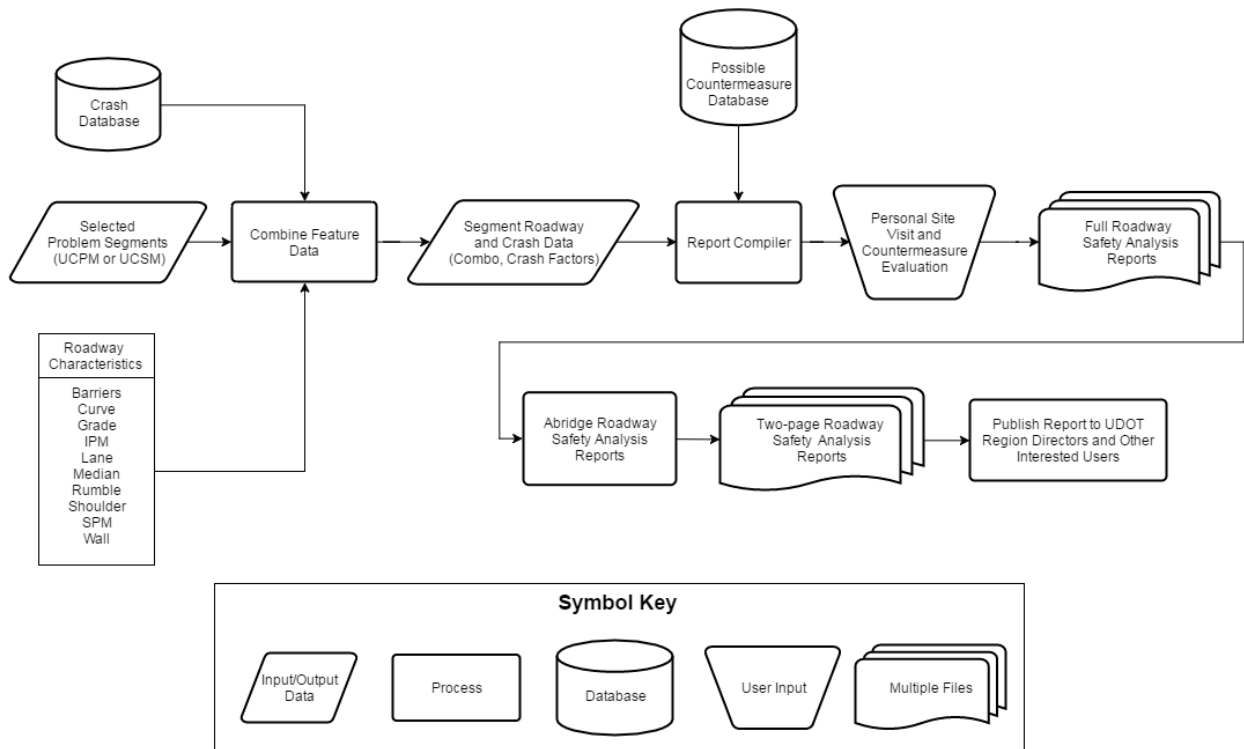
Figure 2-6: Flowchart of report compilation process (Schultz et al. 2016).

Table 2-5: Additional Roadway Characteristics Used for RSARs (Schultz et al. 2016)

Characteristic	Data Source
Median	UDOT Open Data
IPM	Derived from Intersection data, from UDOT Open Data
SPM	Derived from Sign Face data, from UDOT Open Data
Shoulder	UDOT Open Data
Grade	UDOT Open Data
Curve	Derived using HAF Algorithm (Saito et al. 2018)
Lanes	UDOT Open Data
Wall	UDOT Open Data
Barrier	UDOT Open Data
Rumble strips	UDOT Open Data

Once the RSARs have been completed for the segments of interest, the reports are exported to a separate file and condensed to two-page reports. The purpose of the two-page report is to provide a decision-maker with a concise summary of the conditions of the roadway segment in question. The document is short enough that it can be printed on a single sheet of paper, front-to-back. This is the final product of the RSAM process, providing decision-makers with the necessary tools to make key decisions on projects and improvements (Schultz et al. 2016).

2.5 Functional Intersection Area Definition

In order to analyze the roadway safety performance at an intersection, it is necessary to define the boundary of the intersection study area. The intersection study area can be defined by its physical area or functional area. The purpose of this section is to define the area types and to review the method to calculate the functional area distances.

The physical area of an intersection includes the area where the intersecting roadways overlap (Wolshon et al. 2004). The functional area of an intersection is "any area upstream or downstream of an intersection where intersection operations and conflicts significantly influence driver behavior, vehicle operations, or traffic conditions" (Williams et al. 2014). The functional area of an intersection is always larger than the physical area of the intersection (Williams et al. 2014, Wolshon et al. 2004). For the purposes of this research, it was assumed that the upstream and downstream functional area distance of an intersection are equal. This was due to the lack of data for determining the direction of vehicles in a crash. A visual representation of the intersection physical and functional areas is shown in Figure 2-7.
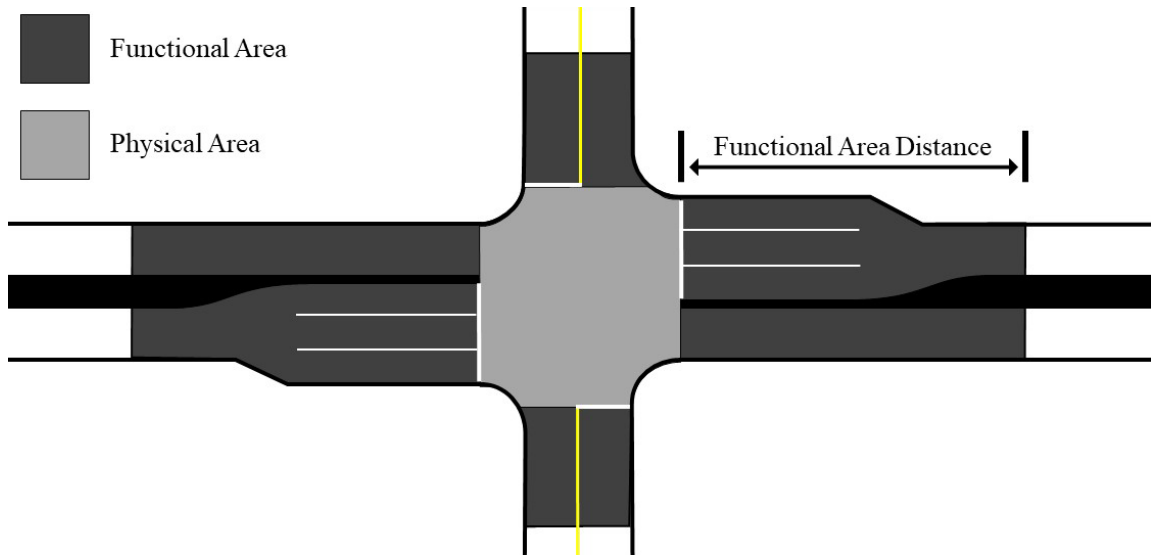
Figure 2-7: Physical and functional areas of an intersection.

The total upstream functional area distance is defined by three sections: (1) perception-reaction distance, (2) deceleration distance, and (3) queue storage (Williams et al. 2014, Stover and Koepke 2002). The perception-reaction distance is the distance traveled while the driver processes the need to stop. The deceleration distance is the distance traveled while slowing down before joining the approach queue. The queue storage is the length that is needed to store the vehicles on the approach. A representation of the functional area distance and its individual sections is shown in Figure 2-8 and summarized in the following subsections.
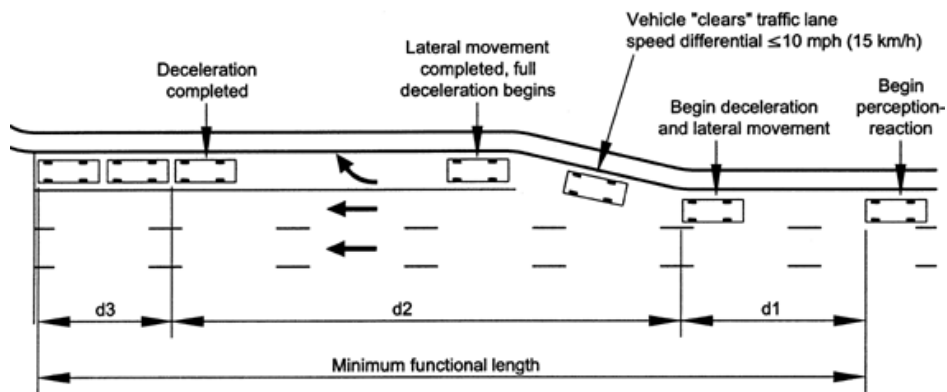
Figure 2-8: Functional area distance of an intersection approach (Rodegerdts et al. 2004).

2.5.1 Perception-Reaction Distance

The perception-reaction distance is dependent on the perception-reaction time (PRT) of the driver and the speed of travel. The perception-reaction distance increases with increasing speed and PRT. The PRT required for a driver to perceive roadway attributes or warnings and react varies from driver to driver. The TRB Access Management Manual suggests using a PRT of 1.5 seconds for urban areas and 2.5 seconds for rural areas (Williams et al. 2014). The perception-reaction distance is calculated by multiplying the PRT by the travel speed of the vehicle, as shown in Equation 2-1.

$$d_{1(PRT)} = 1.47 * PRT * S \qquad (2\text{-}1)$$

where: $d_{1(PRT)}$ = PRT distance (feet)

PRT = perception reaction time (seconds)

S = speed (mph)

For example, if a vehicle were traveling at speed of 35 mph in an urban area (i.e., PRT = 1.5 seconds), the perception-reaction distance would be approximately 77 feet. This value would then be used as the perception-reaction component of the total functional area distance. Several perception-reaction distances for various speeds were calculated and are shown in Table 2-6.

2.5.2 Deceleration Distance

The deceleration distance is the distance traveled while slowing down before joining the approach queue. The deceleration distance is dependent on the initial travel speed, the assumed deceleration value, and the presence of a separate turn lane. There are methods to calculate the standard deceleration distance for through movements and different methods to calculate the deceleration distance for vehicles in a separate turn lane. The distance may also be calculated

using the impact distance at the intersection. The purpose of this section is to describe how to determine each of these distances.

Table 2-6: Perception-Reaction Distances for Various Travel Speeds

Urban[1]		Rural[2]	
Speed, mph	Perception-Reaction Distance, ft	Speed, mph	Perception-Reaction Distance, ft
25	55	25	92
30	66	30	110
35	77	35	129
40	88	40	147
45	99	45	165
50	110	50	184
55	121	55	202
60	132	60	221
65	143	65	239
70	154	70	257

[1]Using 1.5 seconds for the PRT.
[2]Using 2.5 seconds for the PRT.

2.5.2.1 Standard Deceleration Distance

For approaches without separate turning lanes, the deceleration distance is the distance required to stop before the queue storage begins. Therefore, the deceleration distance primarily depends on the travel speed of the vehicle and the deceleration rate. According to the TRB Access Management Manual, researchers have found that drivers travelling at a slower speed use a lower average deceleration rate than drivers travelling at higher speeds. Researchers have found that 85 percent of drivers use a deceleration rate of approximately 7.2 feet per second-squared or more when traveling at 40 mph or less. It was also found that the 50th percentile of drivers use a deceleration rate of 9.9 feet per second-squared or more at the same speed

(Williams et al. 2014). Because deceleration rates are generally higher for roadways with speeds greater than 40 mph, it is appropriate to assume low-speed deceleration rates for high-speed roadways as well. Using the given deceleration rates, the stopping distances can be calculated using Equation 2-2.

$$d = \frac{(1.47*S_0)^2}{2*A}$$
(2-2)

where: d = stopping distance (feet),

          S_0 = initial speed (mph),

          A = acceleration (ft/s^2)

The deceleration distances for various traveling speeds were calculated using the given deceleration rates and Equation 2-2. These values were rounded and are shown in Table 2-7.

Table 2-7: Deceleration Distance Based on Average Deceleration Rates (adapted from Williams et al. 2014)

Speed (mph)	Deceleration Distance (ft)	
	85%[1]	50%[2]
20	60	45
25	95	70
30	135	100
35	185	135
40	240	175
45	305	220
50	375	275
55	455	330
60	540	395
65	635	460
70	735	535
75	840	610

[1]Using average deceleration rate of 7.2 ft/s^2.
[2]Using average deceleration rate of 9.9 ft/s^2.

2.5.2.2 Deceleration Distance for Vehicles in Separate Turn Lane

For vehicles on a separate turning lane movement, the deceleration distance consists of the distance traveled while changing lanes to the turn lane and the full deceleration distance. According to the TRB Access Management Manual, a vehicle will reduce speed by approximately 10 mph while changing lanes to a turn lane before beginning the full deceleration. Average deceleration rates for the lane change maneuver were determined for various speeds and are shown in Table 2-8. Once the vehicle begins the full deceleration in the turn lane, the distance traveled is the same as a standard deceleration distance with a lower initial speed. The distance traveled for the lane change maneuver and full deceleration are shown in Table 2-9 (Williams et al. 2014).

Table 2-8: Average Deceleration Rates for Lane Change Maneuver (adapted from Williams et al. 2014)

Speed, mph	Deceleration Rate, ft/s^2	Time in Lateral Movement, s
< 30	5.9	2.5
30 - 55	4.9	3
> 60	4.2	3.5

2.5.2.3 Impact Distance

The deceleration distance may also be calculated using the impact distance instead of the deceleration method. The TRB Access Management Manual defines the impact distance as the upstream distance from the intersection at which a vehicle brakes in response to a slowing right-turning vehicle in a shared lane. Therefore, the impact distance only applies when the right-most lane is a shared through-right lane as opposed to an approach with an exclusive right-turn lane. The concept of impact distance comes from National Cooperative Highway Research Program

(NCHRP) Report 420 (Gluck et al. 1999). The Access Management Manual provides suggested

percentages of through vehicles that will be impacted by right-turn vehicles based on roadway

functional classification. These values are shown in Table 2-10 (Williams et al. 2014).

Table 2-9: Distance Traveled During Lane Change and Deceleration to a Stop (adapted from Williams et al. 2014)

Speed (mph)	Distance Traveled, ft		
	Lange Change Movement	Full Deceleration Distance	Total Distance
20	55	15	70
25	70	35	105
30	90	60	150
35	130	95	225
40	155	135	290
45	175	185	360
50	200	240	440
55	220	305	525
60	380	375	755
65	310	455	765
70	335	540	875
75	360	635	995

Table 2-10: Suggested Percentage of Impacted Vehicles (adapted from Williams et al. 2014)

Roadway Functional Classification	Through Vehicles Sustaining Impact, %
Principal arterial	2 - 4
Minor arterial	4 - 10
Major collector	5 - 20
Minor collector	10 - 30
Local	N/A

The NCRHP Report 420 researchers produced graphs of the impact lengths based on the percentage of through vehicles that are impacted by turning vehicles. The curve on the graph varies slightly based on travel speed (Gluck et al. 1999). An example of this graph for a 30-mph roadway is shown in Figure 2-9. To calculate the impact distance for a particular intersection approach, the user selects a reasonable percentage of impacted vehicles based on functional classification and the values shown in Table 2-10. The percentage is then used in a graph such as Figure 2-9 to find the impact length.

As an example, a major collector roadway with a speed of 30 mph might have 10 percent of through vehicles impacted. Using the graph in Figure 2-9, the impact length for this scenario would be approximately 220 feet.
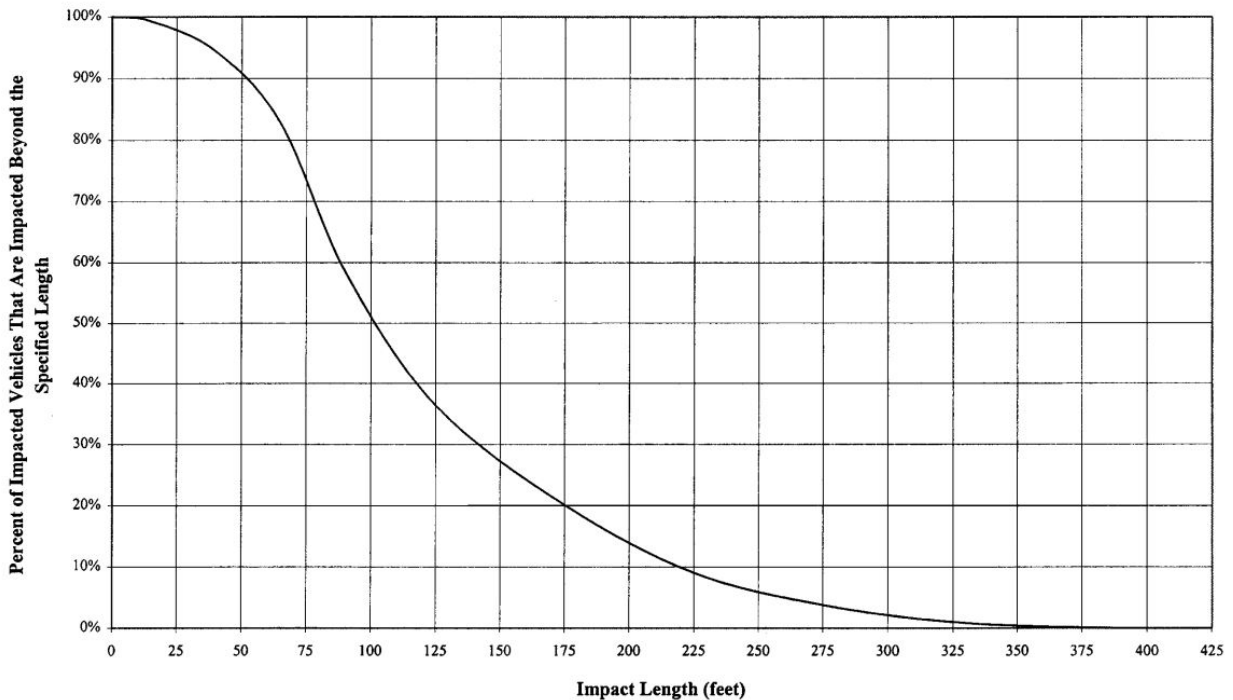
Figure 2-9: Cumulative frequency distribution of impact lengths for a 30-mph roadway (Gluck et al. 1999).

2.5.3   Queue Storage

The queue storage is the length of roadway needed to store vehicles that queue at the intersection. This length varies by intersection and intersection approach. The TRB Access Management Manual discusses using assumed values for the queue storage in the example calculations. This is done by multiplying the number of expected queued vehicles by 25 feet, assuming that each vehicle, including gap space, covers 25 feet of lane length (Williams et al. 2014).

2.5.4   Downstream Functional Distance

The downstream functional distance depends on three factors: (1) geometric factors, (2) operational effects, and (3) human factors. When a vehicle leaves the intersection from a stop to accelerate to travel speed, it requires adequate downstream acceleration distance. If an acceleration lane is included in the downstream movement that the vehicle follows, an additional taper distance must be taken into account. If no taper is needed, then only the acceleration distance is used. The TRB Access Management Manual includes calculated downstream acceleration distances for vehicles leaving from a stop based on American Association of State Highway and Transportation Officials (AASHTO) standards (Williams et al. 2014). These are shown in Table 2-11.

For general intersection safety analyses, it is common to assume that the downstream functional distance is equal to the upstream functional distance. This is generally done when initial crash data does not indicate the direction of travel, but only a location along a roadway, as was the case in this research (Schultz et al. 2008, Schultz et al. 2013a).

Table 2-11: Ideal Downstream Functional Distance (adapted from Williams et al. 2014)

Speed, mph	Acceleration distance, ft	Typical Taper Distance, ft	Downstream Functional Distance, ft
20	100	60	160
25	150	80	230
30	220	100	320
35	320	120	440
40	440	140	580
45	580	160	740
50	770	180	950
55	1,000	200	1,200
60	1,300	220	1,520
65	1,750	240	1,990
70	2,320	260	2,580

2.6   UDOT Data Sources

The access and use of UDOT roadway and crash data has been a critical element of previous research as evident in the RSAM discussion. The following subsections include discussions regarding UDOT data that are pertinent to intersection safety including the UDOT intersection data, UDOT SafeMap and UDOT crash data, and the proposed University of Utah data management system.

2.6.1   UDOT Intersection Data

The UDOT Open Data Portal provides a dataset called "Intersections." This dataset is the result of inventories performed by Mandli Communications. Mandli has most recently completed inventories in 2012, 2014, and 2016. The data are collected using LiDAR and Photolog imagery technology. Collection vehicles installed with this technology make a single pass on Utah state routes and create a three-dimensional point cloud model of the immediate surrounding area. The

point cloud is used to determine roadway characteristics and dimensions (Mandli 2017). The resulting UDOT intersections data file includes useful information regarding each intersection including route numbers, jurisdiction, intersection control type, and the Global Positioning System (GPS) location of the intersection. A list of the data available in the most recent UDOT intersections data file is shown in Table 2-12.

Table 2-12: Data Available in the UDOT 2014 Intersections Data File (UDOT 2017a)

Heading	Description
INTERSECTION_ID	8-digit intersection identification
ROUTE_NAME	Route number with direction letter
START_ACCUM	Milepost location of intersection along given route
STATION	4-digit station number and city location name
REGION	UDOT region number
SIGNALIZED	Yes/no indication if signalized intersection
CONTROL	Intersection control type, if any (stop, yield, other)
SR_SR_INTE	Yes/no indication if intersecting route is a state route
ROUTE_1_IN	Route number of intersecting route, if state route
ROUTE_2_IN	Route number of second intersecting route, if state route
BEGIN_LATITUDE	Latitude of intersection
BEGIN_LONGITUDE	Longitude of intersection
BEGIN_ALTITUDE	Altitude of intersection
COLLECTED_DATE	Date of data collection

2.6.2   UDOT SafeMap and UDOT Crash Data

The UDOT SafeMap tool hosted by Numetric is a data analysis tool developed to provide useful roadway safety data and tools for transportation professionals in the state of Utah. Users need an authorized account to access the website because the Utah crash data are confidential and are protected under 23 USC 409 (USGPO 2012). As of April 2018, there were two sections on UDOT SafeMap: "Numetric" and "Roads." With the roads section were six applications:

"Crash Query," "Citation Query," "Network Screening," "Safety Analysis," "Project Design,"

and "Asset Query" (Numetric 2018). The following sections contain descriptions about these

sections and applications.

2.6.2.1 Numetric Datasets

The purpose of the Numetric section on UDOT SafeMap is to store and view the data that

is used on the website. The user has the option to view existing datasets such as citations,

crashes, and UDOT asset data. These data can be downloaded to a comma-separated value

(CSV) file for additional analysis on a local computer. New datasets can be added to the

Numetric section by uploading an Excel or CSV file. This section is meant primarily for

administrators to upload and update data on UDOT SafeMap, therefore it is recommended by

UDOT that the user avoid making changes in this section.

2.6.2.2 Roads Applications

The Roads section on UDOT SafeMap contains six applications that can be used for

roadway safety and enforcement analyses. The Crash Query application allows the user to

analyze crash data by its various factors. The crashes can be shown geospatially, in tabular

format, or in graphical format. Filters can be searched for and selected to show certain crashes.

For example, a filter for "fatal crashes" can be selected to show only the fatal crashes. Graphical

representations of data are provided for all recorded crash characteristics and factors, providing

useful visuals for the user to understand the crash data and crash factors.

The Citation Query application provides data regarding the police citations that have been

given on Utah roadways. Similar to the Crash Query application, dots are shown on the map for

the locations of the citations that have been issued. Various characteristics are shown for each

point to describe the citation reason, the citation severity, and whether a crash was involved. Filters can be applied to only show certain subsets of the data.

The Network Screening application provides safety data based on roadway segments. The segment data can be shown geospatially or in tabular format. The data includes the segment route, beginning and ending milepoints, and critical safety information such as number of crashes, UDOT Safety Index, and the BYU UCPM and UCSM model results. The segments can be filtered by roadway characteristics such as functional classification, speed limit, and number of lanes.

The Safety Analysis application is a useful tool to analyze certain roadway segments by entering the segment parameters. A new safety analysis screen shows input boxes for a route number and beginning and ending milepoints. The user enters these data and adds the segment to the analysis. A screen appears with crash data for the given segment. The user may provide multiple segments for the analysis. A treatment, or improvement, may be added to calculate a benefit-cost ratio for the treatment.

The Project Design application allows the user to determine the asset status along certain roadway corridors. The user inputs the route number and milepoints of a roadway segment location of a project. The user is then provided with a comprehensive report of all UDOT assets that are currently installed along the corridor.

The Asset Query application provides visual and tabular formats of the UDOT asset data. The user can choose to see data for one of eight provided asset types: barriers, culverts, pavement, pavement messages, rumble strips, signals, signs, and utilities. Once an asset type is selected, the data can be seen on the map or in tables and charts.

2.6.3   University of Utah Data Management System

UDOT has entered an agreement with the University of Utah to develop a data management system where all UDOT crash data will be stored. The purpose of doing this is to have a main data storage and processing system that can be accessed remotely by UDOT and other entities. It is anticipated that the system will be called the "Utah Transportation and Public Safety – Crash Data Initiative" (UTAPS-CDI). It is anticipated that the system will be flexible by allowing data users to obtain data in any desired format, rather than preset data files with a preset format. The system will be similar to other systems that have been created at other universities such as the University of Alabama's Center for Advanced Public Safety (CAPS) and Louisiana State University's Highway Safety Research Group (HSRG) (UofU 2016).

The UTAPS-CDI system will use structured query language (SQL) databases, object-relational mapping, GIS services, user interfaces, and reporting and analysis services available through Shiny by R Studio. Authorized users will receive data from the motor vehicle crash record after it has been processed and gone through quality checks. All quality control tasks will be performed by the University of Utah staff on site. The UTAPS-CDI project had begun as of Fall 2015 (UofU 2016). As of April 2018, the system is still being finalized. It is anticipated that it will be operational by the Fall of 2018.

2.7   Intersection Safety Programs in the United States

A review of other intersection safety programs in the United States was done to learn how other states and agencies have analyzed and improved intersection safety. This section discusses the Georgia Intersection Safety Improvement Program (ISIP) and the Colorado Intersection Safety Performance Functions (SPFs).

2.7.1 Georgia Intersection Safety Improvement Program

The Georgia ISIP was created as a thesis in response to the publication of the Georgia Strategic Highway Safety Plan (SHSP). The Georgia SHSP was created to improve roadway safety and reduce fatalities on Georgia roadways. The Georgia ISIP was created as a five-part program "to better analyze, identify, and implement countermeasures at intersections" to improve overall intersection safety (Thomas 2008).

The first part of the program is a standardized hazardous intersection identification method. Prior to the completion of the Georgia ISIP, counties in the state had their own methods of analyzing safety at intersections. This made it difficult to compare the intersection needs of one county to another. The Georgia ISIP proposes to have a statewide identification method so that safety funds are appropriated to the most urgent projects (Thomas 2008).

The second part of the program is a statewide public involvement tracker. This tool would allow public users of the roadway to submit reports of hazardous intersection locations. Taskforces at the county level would be organized to manage the reports that come from the public. This would allow for a fast response time to fix any issues that may arise (Thomas 2008).

The third part of the program is an automated police crash reporting system through improved technologies. The automated reporting system includes each patrol vehicle being equipped with a mobile data terminal and GPS. This would allow officers to upload crash reports from the crash site to a uniform database. Traffic safety engineers would be able to analyze the data as it is uploaded in order to identify real-time hazardous locations (Thomas 2008).

The fourth part of the program is applying intersection safety strategies from the SHSP of Georgia and other states. This task entails applying the strategies for reducing crashes that have been outlined in the Georgia SHSP. There are also several states with similar roadway networks

as Georgia. It is valuable to analyze the strategies of those states as well, in order to come to the best decision to improve intersection safety (Thomas 2008).

The fifth and final part of the program is using statewide minimal intersection safety equipment. There are several improvements that can be made to intersections that are low cost that improve the safety at the intersections. One example is that of light emitting diode (LED) technology. The LED technology provides better visibility and reliability, which improves the safety of the intersection. This technology can be applied to signal heads and pedestrian walk signs (Thomas 2008).

2.7.2 Colorado Intersection Safety Performance Functions

A SPF is an equation developed to predict the number of crashes per year that will occur at a certain location. According to the Federal Highway Administration (FHWA), SPFs are functions of exposure and roadway characteristics. Exposure is determined using AADT and roadway segment length. The roadway characteristics used may include number of lanes and median type (CDOT 2017).

The Colorado Department of Transportation (DOT) has developed SPFs to predict the number of crashes that will occur at any given intersection in the state. These tools are available on the Colorado DOT website. The SPFs included on the website represent nine typical urban intersection configurations based on number of lanes, number of intersection legs, and intersection control type. SPFs for rural intersections are not included on the website (CDOT 2017). The nine typical intersection SPFs provided by the Colorado DOT are as follows:

- Urban 2-Lane Divided Un-signalized 3-Leg Intersections

- Urban 2-Lane Un-Divided Un-signalized 3-Leg Intersections

- Urban 2-Lane Un-Divided Un-signalized 4-Leg Intersections

- Urban 4-Lane Divided Signalized 3-Leg Intersections

- Urban 4-Lane Divided Signalized 4-Leg Intersections

- Urban 4-Lane Divided Un-signalized 3-Leg Intersections

- Urban 4-Lane Divided Un-signalized 4-Leg Intersections

- Urban 4-Lane Un-Divided Un-signalized 4-Leg Intersections

- Urban 6-Lane Divided Signalized 4-Leg Intersections

Each of the Colorado DOT typical intersections has three files that can be downloaded from the website. The first is an interactive MS Excel file that includes a "UserForm" sheet and an "Expected Means" sheet. The UserForm sheet has a user interface where the user enters data for the intersection street names, roadway AADTs, number of crashes at the intersection, and the date range of the crashes entered. The UserForm sheet also has two embedded graphs that auto-populate as data are entered in the user interface. The Expected Means sheet of the interactive SPF spreadsheet has graphs for total and severe crashes that show the expected CPY values based on major and minor street AADT values. An example of these graphs is shown in Figure 2-10. The graphs on the UserForm sheet are for plotting total crashes and severe crashes, respectively. Once the roadway names and AADT data are entered, the graphs show plots of the expected crashes per year (CPY) based on the roadway AADT. Once the number of crashes and crash date range are entered, the graph shows a plot of the actual CPY relative to the expected CPY. The SPF user interface with a given example for an urban 2-lane divided un-signalized 3-leg intersection is shown in Figure 2-11.

The other two files available for each typical intersection are portable document format (PDF) files with the expected mean graphs that are available in the interactive SPF spreadsheet.

One file is the expected mean graph for total crashes, and the other is the expected mean graph for severe crashes. The benefit of using this SPF information provided by Colorado DOT is that the user can determine how the safety performance of a particular intersection compares to the expected performance. This can be done with the interactive spreadsheet or by using the provided expected mean graphs. In either case, these data are useful for determining the safety of a single intersection; however, it does not compare the safety performance of one intersection to another.
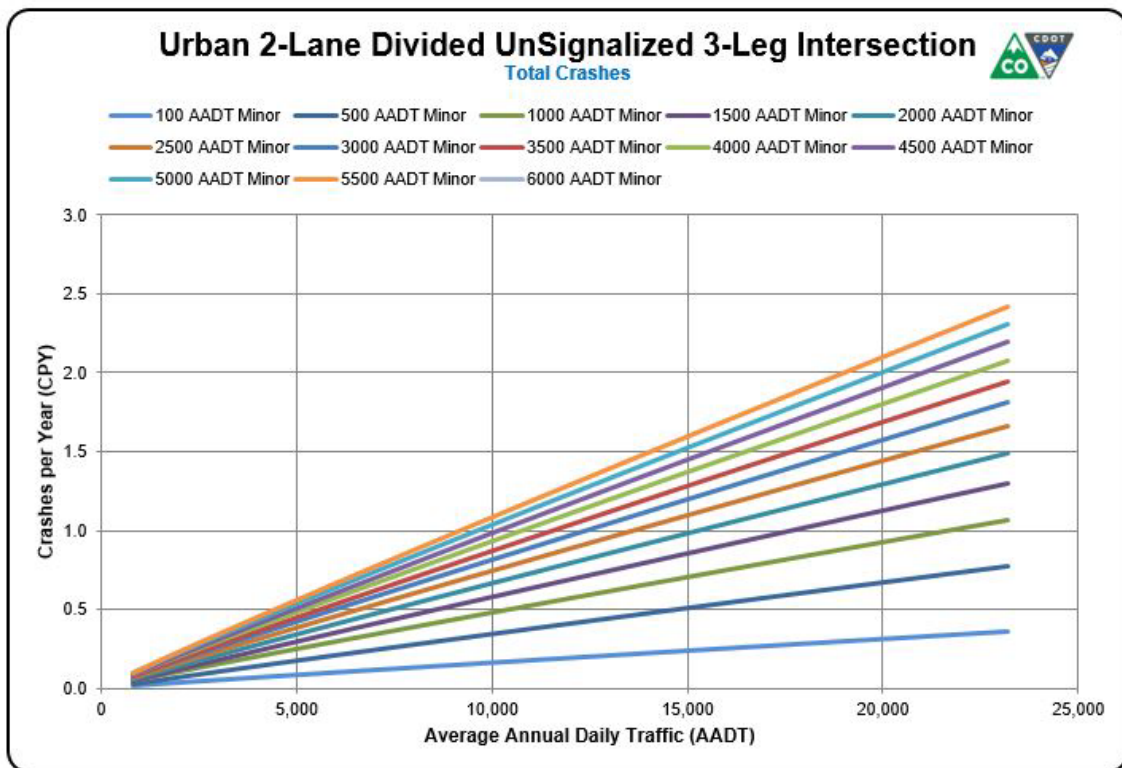
Figure 2-10: Colorado DOT SPF Expected Mean Graphs (CDOT 2017).

Figure 2-11: Colorado DOT SPF analysis interface (CDOT 2017).

Safety Performance Function Analysis - User Interface

Street Name Primary	Street Name Secondary	AADT[A] Road A	AADT[A] Road B	Number of Crashes			Date		CPY[C]	
				PDO[B]	Injury	Fatal	Begin	End	Total	Injury & Fatal
A	B	15,000	2,500	1	1	0	1/1/2016	1/1/2017	1.42	0.30

A. AADT - Average Annual Daily Traffic
   NOTE: If only one AADT is provided, it is assumed as the major AADT and the minor AADT will be calculated at 10%.
B. PDO - Property Damage Only
C. CPY - Crashes per Year (Empirical Bayes Corrected)
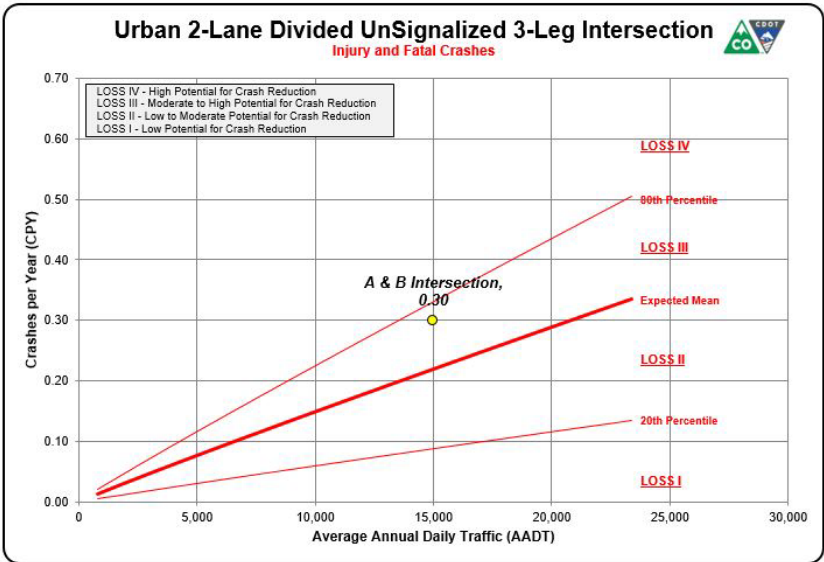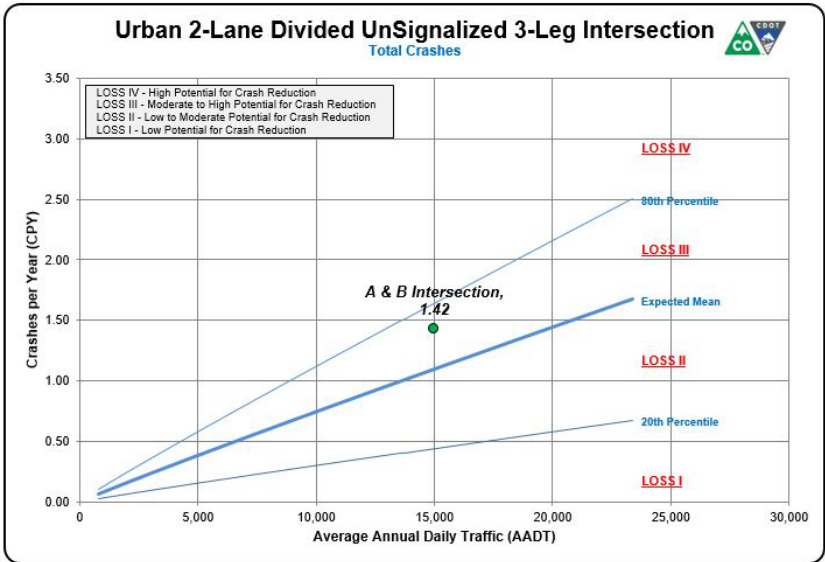
2.8   Summary

The purpose of this chapter is to provide background information regarding intersection safety research and procedures that have already been developed. In recent years, researchers at BYU have developed two models to perform a safety analysis for and rank roadway segments in the state of Utah: the UCPM and UCSM. The RSAM was developed as a process to run these models and provide useful results to UDOT. The functional area of an intersection is the area upstream and downstream of an intersection that is impacted by the operation of the intersection itself. This is critical to understand when analyzing the safety performance of an intersection. There are several datasets and tools provided by UDOT that are useful for intersection safety analysis including the intersections dataset, the UDOT SafeMap online tool, and a future central database hosted by the University of Utah. Intersection safety research has been performed in other states as well including Georgia and Colorado. The information that is discussed in this chapter is useful in developing an intersection safety analysis that will benefit the state of Utah.

3   INTERSECTION SAFETY ANALYSIS METHODOLOGY

3.1   Overview

The ISAM is the analysis process that was developed to run the UICPM. The ISAM includes three major steps: (1) model input preparation, (2) intersection safety model execution, and (3) reports creation. A visual representation of the ISAM process is shown in Figure 3-1. These steps are similar to the steps in the RSAM. The purpose of this chapter is to introduce each step of ISAM, which are discussed in the following sub-sections. Chapters 4 through 7 focus more on the specific processes and tools of each step of the ISAM. First, a discussion is given on the model input preparation. Next, a discussion is given on the execution of the intersection safety model. Last, a discussion is given on the creation of reports based on the model results.

3.2   Model Input Preparation

The intersection safety model relies on accurate roadway and crash data to analyze the safety performance of each intersection. A process, similar to one in the RSAM, has been developed to import and modify datasets to prepare a data input file for the intersection safety model. The purpose of this section is to discuss the sources of the roadway and crash data and how the data are combined to create the UICPM input file. In general, these data are obtained from UDOT in CSV format. This format can easily be imported into a MS Excel spreadsheet format and be modified to fit the needs of the UICPM.
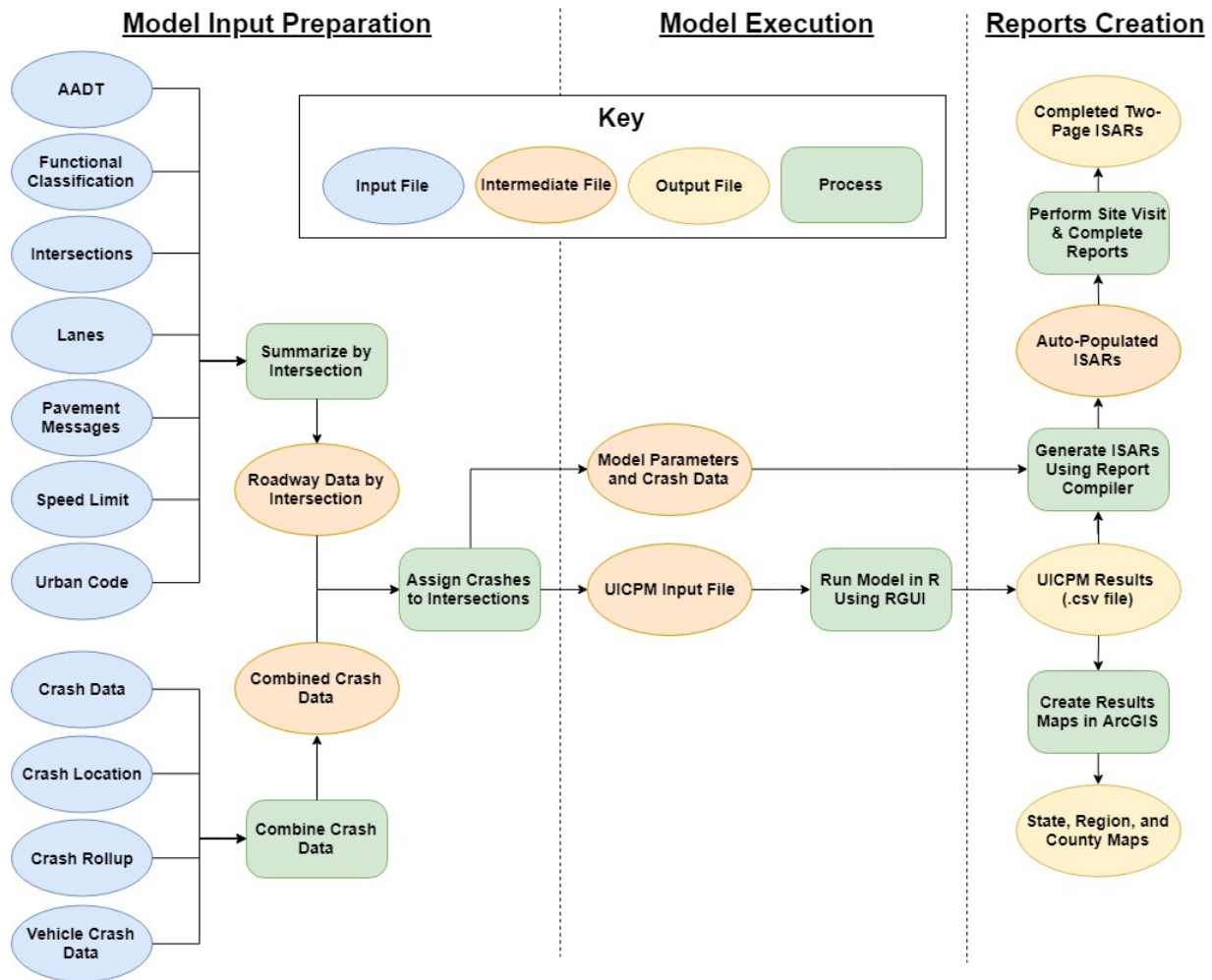
Figure 3-1: Flowchart of the ISAM process.

3.2.1 Roadway Data Input

The UDOT roadway data used in the intersection safety model are obtained from the UDOT Open Data Portal (UDOT 2017a). The data on the UDOT Open Data Portal generally consist of roadway inventory data. The data are open to the public to view and download in GIS or CSV format. It was determined that the data would be used in CSV and spreadsheet format based on the processes built in MS Excel for previous iterations of this safety research. A description of the roadway data used in the intersection safety model is shown in Table 3-1.

Table 3-1: Roadway Data Descriptions

Data	Description
Historical AADT Data	Contains historical dataset of AADT data by roadway segment from each year over several years. Also contains number and percent of trucks on the roadway segment.
Functional Classification	Contains functional classification of Utah roadway segments (i.e., arterial, collector, local road)
Intersections	Contains data related to each intersection including intersecting routes, route milepoints, geographic intersection location (by latitude, longitude and elevation), intersection traffic control device type (i.e., signal and stop), and whether state routes intersect at the intersection
Lanes	Contains data related to the number of lanes along a given roadway segment. The intersection safety model specifically uses the number of thru lanes and width of thru lanes from this dataset.
Pavement Messages	Contains data of all pavement striping messages on Utah state routes, including the location of intersection stop bars. These data are used to calculate the size of intersections.
Speed Limit	Contains speed limit data for each roadway segment
Urban Code	Contains urban code data for each roadway segment, including urban code numbers and corresponding descriptions

The data used for the intersection safety model is the same that was used in the RSAM, except for the intersections and pavement message datasets. The intersections dataset provides necessary information for the intersection safety model such as intersection location and intersecting routes. While in the RSAM the roadway data were segmented into individual roadway segments, the intersection safety model input is separated into characteristics by intersection. The pavement message data allow the user to calculate the size of the physical area of intersections based on the locations of intersection stop bars.

3.2.2 Crash Data Input

The crash data used in the intersection safety model were obtained directly from the UDOT Traffic and Safety Division. The crash data come from crash reports completed by Utah law enforcement officers. Due to the confidentiality of the crash data, the data are not available to the public. Crash reports may be disclosed only in accordance with Utah Code 41-6a-404. A description of the crash data used in the intersection safety model is shown in Table 3-2.

Table 3-2: Crash Data Descriptions

Data	Description
(General) Crash Data	Contains general crash data dealing with the conditions of each crash including weather, lighting, first harmful event, and manner of collision
Crash Location	Contains data regarding the location of each crash including latitude, longitude, elevation, route number, and milepoint
Crash Rollup	Contains data regarding the crash factors of each crash such as "intersection-related," "young driver," and "work zone related"
Crash Vehicle	Contains data regarding the vehicles involved in each crash

3.2.3 Critical Data Columns

Within the roadway and crash data, only certain columns of data are critical and necessary to be included in the final intersection safety model input file. The other columns of data are not needed for the intersection safety model, though they may be useful in other analyses. The necessary columns of data have been identified as "Critical Data Columns." If these data are not found in the input datasets, the process does not work. This is due to the specific data that are required to complete the processes written in the VBA macros and R code. When certain data are not found during the process, the macros or code return an error to the

user. The critical data columns of the roadway and crash data are shown in Appendix A. An example of the critical data columns of the AADT data is shown in Table 3-3.

Table 3-3: AADT Critical Data Columns

Expected Header	Description
ROUTE	Route ID: numeric route number for a given road segment
BEGMP	Beginning Milepoint: beginning milepoint of the road segment
ENDMP	End Milepoint: end milepoint of the road segment
STATION	Station Number: seven-digit number, identifying the traffic counter station number
AADT[YEAR]	AADT [YEAR]: historical dataset of AADT data from each year; at least 7 years of this data are needed (i.e., AADT2012)
SUTRK	Single Truck Percent: percent of single trailer trucks per segment
CUTRK	Combo Truck Percent: percent of combination trailer trucks per segment
SUTRKCOUNT	Single-Unit Truck Count: number of single trailer trucks per segment
CUTRKCOUNT	Combo-Unit Truck Count: number of combination trailer trucks per segment

The headings of the critical data columns in the raw UDOT roadway and crash data files may change occasionally. This makes it difficult to write VBA and R code that looks for specific column headings to important data. A tool called "Check Headers" was created for the RSAM to account for changes in column headings. This same tool was applied for the intersection safety model process as well.

3.2.4   Non-State Route Data Limitations

As of the writing of this report, there are certain limitations in the roadway data that restricts the safety analysis to intersections with at least two intersecting state routes. This is due to lack of data in most of the roadway datasets for non-state routes. A summary of the roadway

data that are available for state routes and non-state routes is given in Table 3-4. As shown, non-state route data are not provided in the functional classification, percent trucks, number of lanes, pavement messages, and speed limit datasets.

Table 3-4: State Route Data Availability

Data	Data Available for	
	State Routes	Non-State Routes
AADT	X	X
Functional Classification	X	
Number of Lanes	X	
Pavement Message	X	
Percent Trucks	X	
Speed Limit	X	
Urban Code	X	X

There is also data lacking in the intersections dataset. As discussed previously, this dataset is used to determine the locations of intersections in the state by coordinates and by route intersecting points. The dataset lists the routes that meet at each intersection only if the routes are state routes. Therefore, non-state routes are not listed as intersecting routes at the intersections. This limits the ability of the automated Excel process to determine if non-state routes intersect at a particular location. Some of the intersections list a single state route at the intersection but not a second one. These are of little value at the present time because no data are available for the cross street of the state routes. This means that the only significant analysis that can be done is analysis on intersections that have at least two state routes listed at the intersection. Because of these limitations, the current ISAM process is being used to only analyze intersections with at least two state routes intersecting. Based on discussions with UDOT officials, it is anticipated that these roadway datasets will be updated in the near future to include data for non-state routes,

including federal aid routes. Once this occurs, the proposed safety analysis process can be applied to more intersections that include non-state routes.

3.3   Intersection Safety Model Execution

Once the model input is created, the UICPM can be executed. The UICPM was developed using R, a statistical analysis software (RPSC 2016). The UICPM is a Bayesian generalized linear model that develops crash distributions for all study intersections. The observed number of crashes at each intersection is then compared with the crash distribution to determine if more crashes are occurring than expected at each intersection. Intersections are ranked by relative safety performance, and the results are provided in a spreadsheet format.

The actual R model file is a text file with R code written to complete the safety analysis. This file can be run manually in R Studio software. However, in order to make the process more user-friendly, the R GUI Excel tool was developed. The R GUI is an Excel-based tool that was developed for the RSAM to run R models from Excel. The R GUI provides an easy way for the user to prepare data and run the model. The R GUI allows the user to set several parameters before running the model such as the input file, the R model file, the number of iterations, and the variables to use in the model itself.

3.4   Reports Creation

After the model runs and results are generated, it is important that the results are provided to UDOT employees in a simple format that helps them make decisions regarding intersection safety improvements. As done with the segment analysis, reports and GIS maps are provided to UDOT for the UICPM results. The Intersection Safety Analysis Report (ISAR) is a simple two-page report, created for each intersection of interest. The report contains information regarding

the intersection physical characteristics, crash history, model results, and potential

countermeasures for the intersection. There is also a section in the report where safety analysis

and site visit notes can be entered by the user. This report is meant for the UDOT Region

employees to bring to the field for further analysis. An Excel tool called the Report Compiler is

used to create these reports. The tool allows the user to choose intersections to create reports for

based on state, region, county ranking, or intersection ID. GIS software is used to create state,

region, and county maps of the model results. The intersections that were analyzed are plotted on

the map, and PDF files of each map are saved. The intersections, represented as dots on the map,

are color-coded to show their ranking compared to other intersections.

3.5   Summary

The ISAM is a complete process designed to accurately and quickly execute the UICPM.

The process is automated with several Excel and GIS tools, making it a simple process for a user

to replicate. The process includes the preparation of data for the intersection model input, the

execution of the statistical model in R, and the creation of reports and maps to display the results.

While this chapter contained a brief overview of the UICPM, the following three chapters have

additional discussions regarding the three steps of the UICPM.

4 MODEL INPUT PREPARATION

4.1 Overview

The preparation of a valid model input file is key to the operation of the UICPM. The purpose of this chapter is to discuss four topics related to the input data and data tools that are used for the intersection safety model. The first topic is a discussion on the preliminary data preparation tools that are available to modify the input roadway and crash data. The next topic is a discussion on the creation of the intersection data file. Third, a discussion is provided on the combination of crash data into a single file. Fourth, a discussion is provided on the assignment of crashes to each intersection to create the intersection model input file.

4.2 Preliminary Data Preparation

The purpose of this section is to discuss the preliminary data preparation tool that modifies the raw UDOT data to create the intersection safety model input file. As with the RSAM, the data preparation processes have been automated using MS Excel VBA macros. The RSAM used the "Roadway and Crash Data Preparation Workbook" to complete the tasks of segmenting roadway data and combining crash data. The tools used in that workbook have now been transferred to be used in the R GUI, which is used to run the R model. In other words, the automated processes of preparing the preliminary data and running the model are now in the same MS Excel workbook.

Before creating the model input file, the roadway data must be combined into a single

intersection roadway file and the crash data must be combined into a single crash dataset. An

additional GUI has been added to the R GUI that allows the user to create the preliminary

roadway and crash data files. As shown in Figure 4-1, the "Create Input Datasets" button has

been added to the model input screen. Once selected, the "Intersection Data Preparation" GUI

appears, as shown in Figure 4-2. To create a new intersection roadway file, the user selects the

blue buttons one at a time in the "Roadway Data" section. Once each button is selected, the user

is asked to browse to and select the corresponding roadway data file. The file path of the data file

is inserted into the text box adjacent to the command button once the user selects the data file.

After all of the roadway data files have been selected, the user selects the "Combine Roadway

Data" button to create the intersection roadway data file.

To create a combined crash data file, the user follows the same steps, but for the crash

data buttons in the "Crash Data" section. The following sections describe the processes of

combining the roadway intersection and crash data, respectively. As mentioned, these processes

have been automated using VBA macros in Excel.
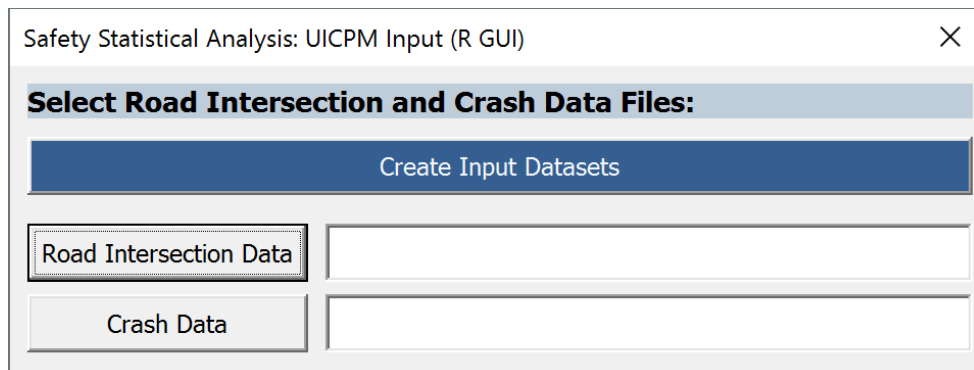
Figure 4-1: "UICPM Input" screen with "Create Input Datasets" button.

Figure 4-2: Intersection Data Preparation window.

4.3   Intersection Data File Creation

The intersection safety model requires the following UDOT roadway data files for the analysis: AADT, functional classification, lanes, pavements messages, speed limit, urban code, and intersections data. Once the user has selected file locations for each of these files and selected the "Combine Roadway Data" button in the Intersection Data Preparation window, a process automated with VBA macros begins to combine the roadway data by intersection. Each

input file is imported one at a time as a separate worksheet into the R GUI workbook. Each file is modified individually to be able to meet the needs of the UICPM. The following sub-sections contain descriptions of the modifications made to each imported file and how these files are compiled into a single roadway intersection data file.

4.3.1   AADT Data

The AADT data are imported to determine the number of entering vehicles at the intersection, the number of intersection legs represented in the data, and percent trucks. Throughout the ISAM research process, it has been found that entering vehicles is an important variable in predicting the number of crashes that occur at a particular intersection. Therefore, it is critical to obtain accurate entering vehicle estimates. The AADT data are listed by roadway segment, providing a segment route number and the beginning and ending milepoints of the segment. Several years of AADT data are provided in the file, and the most recent year of percent truck data is provided. In order to convert this AADT data to entering vehicle data, the AADT value for each data year are divided by two for each leg. The sum of the halved AADT values for each leg is calculated to determine the total number of entering vehicles at the intersection.

These calculations are performed using VBA macros for each represented intersection leg and summed for each intersection. As these calculations are made, the macros also record the number of intersection legs that are represented in the data. The macros also take a weighted average of all percent truck values from the represented intersection legs to determine an overall intersection truck percentage.

4.3.2 Functional Classification

The functional classification dataset is used to determine the functional classification of routes that intersect at each location. Based on the UDOT data, the functional classification ranges from a value of 1 to 7, as shown in Table 4-1 (UDOT 2017a). Functional classifications are assigned to route segments between two milepoints. The functional classification values are recorded for each intersecting route. As will be discussed in Section 4.5.3, functional classification is one alternative to calculate intersection functional area. Recording the functional classification of each intersecting route allows the functional area calculation to be made on each route. Maximum and minimum functional classification values between all intersection routes are recorded to provide a range of functional classification at the intersection. The maximum functional classification is the one with the lowest code number, since lower code numbers represent roadways of higher functional classification (e.g., principal arterial and freeway). The minimum classification is the one with the highest code number (e.g., local road).

Table 4-1: Functional Classification Designation

Code	Description
1	Interstate
2	Other Freeway & Expressway
3	Other Principal Arterial
4	Minor Arterial
5	Major Collector
6	Minor Collector
7	Local

4.3.3 Lanes

The lanes dataset is used to determine the number of through lanes at each leg of an intersection, as well as the width of those lanes. The lanes dataset has information regarding all

49

lanes along state route segments in Utah, including the number of lanes for each lane type from one milepoint to another. The lanes data is an important parameter for each intersection leg because it is a surrogate of the volume and capacity. The number of lanes is multiplied by the width of the lanes to get an estimate for the width of each roadway, which can be used to estimate the size of the intersection. This method was found to be inaccurate due to the difficulty of determining the total number (including shoulders) and width of lanes at an intersection. The pavement messages dataset was incorporated to improve the accuracy of this estimation. This will be discussed further in Sections 4.3.4 and 4.5.2. The maximum and minimum number of lanes values are recorded at an intersection. The maximum and minimum roadway width estimates are also provided.

4.3.4   Pavement Messages

The pavement messages dataset is used to estimate the physical size of intersections. The dataset includes the locations of striped messages and symbols on Utah's state routes. The location is given by a route number and milepoint. The dataset describes the locations of several types of messages, including stop bars and turn arrows. The locations of stop bars or turn arrows are used to find the distance between two stop bars or two arrows on either side of an intersection. This is measured to estimate the internal intersection size. It has been found that not all intersections have stop bar or arrow location data available. However, where available, the distance between either two stop bars or two arrows is helpful in providing a more accurate calculation of intersection size than other alternatives. Because the distance between stop bars is desired, 30 feet is subtracted from the distance between arrows to estimate the distance from stop bar to stop bar. This distance is measured for each route to be used for the functional area calculation and will be discussed in Section 4.5.2.

4.3.5   Speed Limit

The speed limit dataset is used to determine the speed limit of the various legs of an intersection. The speed limit is especially important when calculating the functional area, as will be discussed in Section 4.5.3. The speed limit of each route at the intersection is found using VBA macros. If speed limit is used as the variable that defines the intersection functional area, then these speed limit values are used to determine the functional area distance for each represented intersection leg. After the intersection functional areas have been defined, only the maximum and minimum speed limit values are kept for the intersection model input file to represent the range of speed limits at each intersection.

4.3.6   Urban Code

The urban code dataset is used to determine whether an intersection is in an urban, suburban, or rural area. The dataset provides the urban code as a number value and description for each roadway segment. Rural areas are represented by the number 99999 and with "Rural" as the description. Suburban areas are represented by the number 99998 and with "Small Urban" as the description. Urban areas are described using the metropolitan area name and a corresponding number.

4.3.7   Intersection

The intersection dataset is used to gather various data of the intersections that are analyzed in the UICPM, including the geographical location of each intersection, latitude and longitude coordinates, route numbers of intersecting state routes, and the milepoints at which those routes intersect. The dataset does not list route numbers for non-state routes. There is an indicator in the data showing whether the intersection is a state route to state route intersection. If

it is such, the intersecting routes are listed. The dataset also provides data regarding the intersection control type and whether or not the intersection is signalized.

Several modifications are made to the intersections dataset using automated VBA macros. First, the intersection control type data and signalized indicator are combined into a single data column. This was done because the intersection control type data only shows when there is a stop control. VBA macros are used to refer to the signalized indicator and enter "signal" for the intersection control type where applicable.

The next modification made to the intersection data is the removal of duplicate intersections. In the raw dataset, a single intersection may be listed up to three times with each intersecting route being listed as a main intersection route. The intersections are combined by comparing the difference in latitude and longitude between intersections. If two intersections are found to be closer than 0.001 degrees of latitude and 0.0013 degrees of longitude (approximately 365 feet for both) to each other, the intersections are combined, and the coordinate and elevation values are averaged. The distance of 365 feet was chosen after a sensitivity analysis was conducted to ensure that the correct intersection points were combined into one. When intersection points are combined, the data are also combined so that the final intersection point has data for all intersecting routes and milepoints.

The final modification that is made to the intersection dataset is the conversion of the station data to a metropolitan area. The station data comes with a number value as well as a metropolitan area name. The station number is removed from the data using VBA macros, while the metropolitan area name remains.

Once all roadway data files, including the intersection data file, are imported and modified individually, the roadway data are then combined into a single intersection input file.

The intersection input file contains information regarding the location of each intersection along with the roadway characteristics pertaining to each. A list of the data found in the intersection input file is shown in Table 4-2.

Table 4-2: Intersection Input File Data

Location	Roadway Characteristics
Intersection ID Number	Functional Classification
Latitude	Internal Intersection Distance
Longitude	Percent Trucks
Elevation	Entering Vehicles
Signalized Indicator	Number of Intersection Legs
Intersection Control	Number of Lanes
State Route Intersection	Estimated Roadway Width
Route Numbers	Urban Code
Route Milepoints	Speed Limit
UDOT Region	
County	
City	

4.4 Crash Data Combination

As with previous safety research, UDOT provides crash data each year in four separate data files. Each data file has different information about the crashes that occurred on Utah roadways. A process was developed for the RSAM to combine the four data files into a single dataset (Schultz et al. 2016), and the same process is used for the ISAM. This makes the process of assigning crashes to intersections easier. Crash data are combined based on the crash identification (ID) number that is provided for each crash. The R GUI tool is used to combine this data, as is done with the roadway data. In the Intersection Data Preparation window, the user selects the button for each crash data file to select the location of each file. Once the user selects the "Combine Crash Data" button, each file is imported, and all files are combined into a single

dataset. Minor modifications are made including making the route number format uniform throughout, removing non-state route crashes from the dataset, and removing crashes that occurred on freeway ramps.

The final dataset includes crash data related to the location and time of the crash, the factors that contributed to the crash, crash severity, and the crash sequence of events. A summary of the data that are provided in the combined crash data file is shown in Table 4-3.

Table 4-3: Crash Input File Data

Location / Time	Factors	Other
Crash ID	Light Condition	Crash Severity
Route ID	Weather Condition	Event Sequence
Milepoint	Manner of Collision	Number of Vehicles
Coordinates	Pavement Condition	Number of Injuries by Severity
Crash Date / Time	Work Zone Related	Roadway Departure
Interstate Highway	Roadway Alignment	Overturn / Rollover
Urban / Rural	Pedestrian Involved	Collision with Fixed Object
	Bicyclist Involved	
	Motorcycle Involved	
	Improper Restraint	
	Driving Under the Influence (DUI)	
	Aggressive Driving	
	Distracted Driving	
	Drowsy Driving	
	Speed Related	
	Intersection Related	
	Roadway Geometry Related	
	Animal Related	
	Comm. Motor Vehicle Involved	
	Teenage Driver Involved	
	Older Driver Involved	
	Night / Dark Condition	
	Train / Rail Involved	
	Transit Vehicle Involved	

4.5   Intersection Model Input Creation

Once the intersection and crash data files have been created, the crashes can then be assigned to each intersection in preparation for the intersection safety model. Crashes are assigned to an intersection if they fall within the functional area of the intersection. The method used to calculate the functional area of an intersection is found in Transportation and Land Development (Stover and Koepke 2002) and the TRB Access Management Manual (Williams et al. 2014) as discussed in Section 2.5. The recommended upstream functional area distances provided by Williams et al. (2014) are used as default functional area values. However, the user also has the option of changing those values and defining the functional area by functional classification or urban code. If the user wants to set a single functional area distance for all roadways, the same distance may be entered for all possible characteristic values. For example, a user may want to utilize a value of 250 feet, consistent with the functional area recommended in the Highway Safety Manual (HSM) (AASHTO 2010). Since the functional area is measured from the stop bar of the intersection, additional calculations are made to estimate the internal area of each intersection. The following sub-sections discuss the tool used to create the UICPM input, the internal intersection area, and the intersection functional area, respectively.

4.5.1   UICPM Input Screen

The tool that is used to select the settings for the UICPM input file is the UICPM Input screen in the R GUI, as shown in Figure 4-3. The user selects the intersection and crash data files, the severities to be included in the analysis, and the functional area method. If "Recommended Functional Area" is selected, the functional area distances from the TRB Access Management Manual (Williams et al. 2014) are used. If "User-Defined" is selected, the user defines the functional area distances by speed limit, functional classification, or urban code.

Default functional area distances already exist for these three categories. The user can use the

default distances or enter distances manually. If the user desires to use a single distance for all

roadways, the same distance can be entered for all options in a certain category. Once the desired

options are selected, the user selects the "Create Input Data for Statistical Analysis" button. VBA

macros are used to then assign crashes to intersections to create the UICPM input file.

Figure 4-3: UICPM Input screen.

4.5.2   Internal Intersection Area

To calculate the internal intersection area, there is a need to estimate the distance from

the stop bar to the center of the intersection. As discussed in Section 4.3.4, the pavement

message data are being used to determine the internal intersection distance. However, the data are not fully complete for all intersections in the database. For the intersections that do not have accurate pavement message data available, the internal intersection distance is calculated based on the number of lanes dataset, which provides the number of lanes for each intersecting route, as well as the lane width. By multiplying the number of lanes by the lane width, a roadway width value can be found, which can then be used to estimate the internal intersection distance. However, it has proven difficult to calculate accurate roadway widths from this dataset alone. This is because the summed width of the thru lanes alone is usually much less than the total width of a roadway. Even when turn lanes, shoulders, and other lanes are taken into account when calculating a roadway width, the roadway width could not accurately be estimated.

To address the problem of accurately estimating roadway width, the BYU research team manually measured over 200 of the intersections being studied to compare the actual roadway width (measured stop bar to stop bar) to the calculated roadway width of the intersection legs. When using the number of through lanes to calculate roadway width, the calculated width was approximately 80 feet less than the actual width on average. When using the number of all lanes to calculate the roadway width, the calculated width was approximately 50 feet less than the actual on average. It was determined that this average difference could be added to the calculated values to get a better estimate of the roadway width. It was found that it is slightly more accurate to calculate the roadway width using the width of the number of thru lanes plus the 80-foot adjustment than using the total number of lanes plus the adjustment. Therefore, the roadway width is calculated by multiplying the number of thru lanes by the lane width and then adding 80 feet. This results in most widths being within 50 feet of the actual width.

The internal intersection distance, or distance from the intersection center to the stop bar is then calculated using the maximum and minimum roadway widths at the intersection. The distance is calculated by taking half of the average of the maximum and minimum roadway widths. In other words, the distance is equal to the sum of the maximum and minimum widths, divided by four. Though this distance may not be accurate for all intersection legs, it is the best estimate of the average internal distance that can be determined from the provided data. A visual representation of this calculation is shown Figure 4-4.

$$Internal\ Distance = \frac{Width\ 1 + Width\ 2}{4}$$
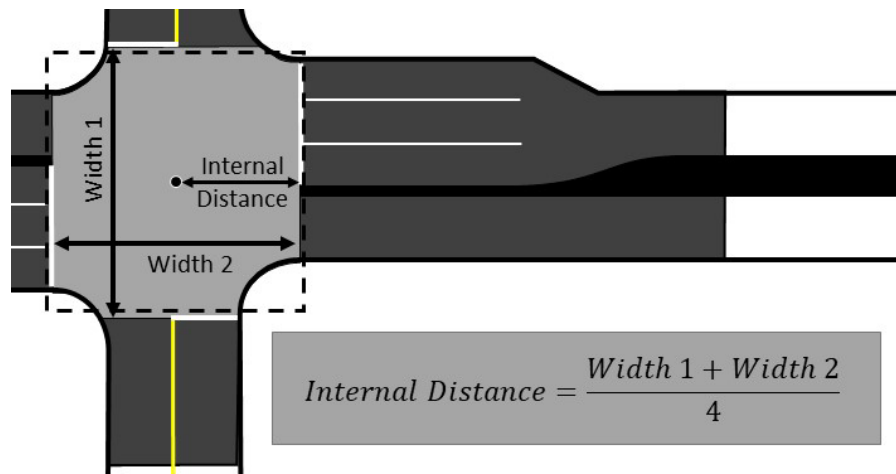
Figure 4-4: Internal intersection distance concept diagram.

4.5.3   Intersection Functional Area

Intersection crashes are assigned to intersections based on the functional area of the intersection, which is generally based on speed limit. In the initial stages of the UICPM, the functional area was used as a radius to assign the crashes to an intersection. The latitude and longitude coordinates of each intersection acted as the origin of that radius, and all crashes within the radius were assigned to the intersection. The functional area distance was based on the maximum speed limit value. This method would use a uniform functional area for all directions and would sometimes assign crashes from side streets to the intersection.

Because of the inaccuracy of the radius methodology, the functional area is now used as a distance that is measured along the routes from the intersection milepoint. The functional area is based on the speed limit of each route; therefore, each route has its own functional area distance. This methodology provides more accurate functional area distances for each route and assignment of crashes at each intersection, especially for the side streets that have different functional areas than the major streets. For now, crashes from non-state routes cannot be assigned using this methodology due to a lack of data; but this will change as more non-state route data becomes available. As discussed in Section 2.5, the downstream functional area distance is generally less than the upstream functional area distance (Williams et al. 2014). However, because of the difficulty in accurately determining the route direction for each crash from the data, it is not possible to determine whether a crash is on the upstream or downstream of an intersection. Therefore, it was assumed that the downstream functional area distance was equal to the upstream functional area distance.

4.6   Summary

The purpose of this chapter was to describe the data that are needed for the intersection safety model and how the data are formatted. The raw UDOT data files first undergo a preliminary data preparation stage to create a combined crash file and a roadway characteristic file defined by intersection. These files are then combined by assigning crashes to intersections if they are located within the functional area of the intersection. A detailed procedure has been developed to ensure that the functional area is an accurate measurement from the stop bar of the intersection, providing an accurate input file for the intersection safety model.

5   INTERSECTION SAFETY MODEL

5.1   Overview

The purpose of the UICPM is to run a regression analysis on existing intersection and crash data to create a distribution of the number of annual crashes that are expected to occur at each intersection in the state. The median of this distribution is assigned to the intersection as the predicted number of crashes, which is then compared to the actual number of crashes to determine the intersections that have more crashes than expected. These intersections are then identified as hot spots, or locations where UDOT can focus their resources to improve safety. The purpose of this chapter is to discuss the development and execution of the UICPM. Three topics are provided in this chapter for this discussion. The first topic is a discussion of the development of the model, including the reasoning behind certain model elements. Second, a discussion is provided of how the model is executed and the tools that are used for model execution. The last topic is a discussion of the results that are generated by the UICPM.

5.2   Model Development

The UICPM was developed using R statistical computing software. R is a free software environment that uses R code to run computations and create useful statistical graphics (RPSC 2016). Also, MS Excel is able to interface with R using VBA macros. For these reasons, R was used to develop the UICPM. The UICPM is a Bayesian generalized linear model. The methodologies used in the development of the UICPM are similar to those used for the UCPM

and UCSM. The following sections contain discussions regarding the variable selection, model methodology, and the intersection ranking involved with the UICPM.

5.2.1 Variable Selection

In developing a Bayesian regression model, it is critical to find variables to use that are good predictors of the response of the model and that create a model that fits the data well. In the case of the UICPM, the response of the model is the number of crashes at an intersection. Common variable selection methods for Bayesian analyses were used to select variables for the UICPM. The Bayesian horseshoe selection method was used to determine which variables would be significant in the intersection safety analysis. This was done by regressing all possible variables onto the response of the actual number of crashes. In order to place more emphasis on the severe crashes, only crashes of severities 3, 4, and 5 were included in the actual crash count.

Instead of using a straightforward regression where the correlation between the response and the variables is analyzed, the Bayesian horseshoe regression algorithm provides probabilities that the variables contribute to the understanding of variation in the response. The resulting probabilities were used to decide which variables to include in the UICPM. Based on the results, only the entering vehicles variable has a probability greater than 90 percent of being contributing to the understanding of variation in the response. A forward variable selection method was used to determine other variables to use. The deviance information criterion (DIC) was used to determine how well each possible model fit the provided data. DIC provides information on model fit by combining the likelihood of the model given the chosen variables and a penalty for adding too many variables that decrease model fit. In this way, DIC helps keep the model simple yet explanatory. In general, a low DIC value means the model has a better fit than models with higher DIC values (Spiegelhalter et al. 2002).

Variables were chosen by comparing DIC values iteratively. All numerical variables were considered in the forward variable selection algorithm. Variables were added to the model iteratively and only retained if they improved the DIC of the model. The DIC of each iteration of the forward variable selection process is shown in Table 5-1. Five variables resulted and were used in the model as explanatory variables, including: entering vehicles, number of intersection legs, minimum number of lanes, maximum roadway width, and maximum speed limit.

Table 5-1: UICPM Forward Variable Selection Results

#	Included Variables	DIC
1	Entering Vehicles	2589
2	Entering Vehicles, Percent Trucks	2592
3	Entering Vehicles, Year	2592
4	Entering Vehicles, Number of Intersection Legs	2583
5	Entering Vehicles, Number of Intersection Legs, Maximum Number of Lanes	2588
6	Entering Vehicles, Number of Intersection Legs, Minimum Number of Lanes	2544
7	Entering Vehicles, Number of Intersection Legs, Minimum Number of Lanes, Maximum Roadway Width	2520
8	Entering Vehicles, Number of Intersection Legs, Minimum Number of Lanes, Maximum Roadway Width, Minimum Roadway Width	2521
9	Entering Vehicles, Number of Intersection Legs, Minimum Number of Lanes, Maximum Roadway Width, Maximum Speed Limit	2519

5.2.2   Model Methodology

As the UICPM was developed, it was assumed that each intersection was independent of every other intersection, such that the number of crashes, y_{ijk}, for intersection reference class i, at intersection j, and year k is approximately equal to a Poisson distribution of an estimated number of crashes (λ_{ijk}), as shown in Equation 5-1. The Bayesian Poisson regression model is then used to estimate the value of λ_{ijk} such that the log of λ_{ijk} is equal to the product sum of select variables and coefficients, as shown in Equation 5-2.

$$y_{ijk} \sim Poisson(\lambda_{ijk}) \qquad\qquad (5\text{-}1)$$

$$\log(\lambda_{ijk}) = \boldsymbol{X\beta} \qquad\qquad (5\text{-}2)$$

where: y_{ijk} = actual number of crashes,

λ_{ijk} = estimated number of crashes,

X = matrix of covariates,

$\boldsymbol{\beta}$ = vector of coefficients,

i = reference class indicator,

j = intersection ID indicator (number assigned to each intersection),

k = data year indicator

It was expected that the effect of increased covariates would be different for certain categories of intersections due to the differences between certain intersection types. Specifically, the effects of using UDOT Region, urban code, and functional classification as reference classes were analyzed. Due to intersections possibly having varying functional classifications between major and minor streets, functional classification was represented with a dual-designation, indicating both the minimum and maximum functional classification at the intersection.

In performing model selection, it was found that the effect of covariates did change for UDOT Region, urban code, and functional classification. Independent and interaction effects were considered for UDOT Region and functional classification, and urban code and functional classification, respectively. It was found that UDOT Region and functional classification with an interaction effect outperformed all other similar models with other reference class specifications. The effect of using UDOT Region and functional classification type as reference class was added to the model.

Predictive accuracy was very important in the development of the UICPM; therefore, the Root Predicted Mean Squared Error (RPMSE) was measured to evaluate predictive ability (Hyndman and Koehler 2006). That is, for every count of crashes at an intersection, y_{ijk}, and the expected number of crashes, λ_{ijk}, the RPMSE was calculated. The equation for RPMSE is shown in Equation 5-3.

$$RPMSE = \sqrt{\frac{1}{N_s} \sum_{i=1}^{N_s} (y_{ijk} - \lambda_{ijk})^2} \tag{5-3}$$

where: N_s = number of intersections

After an extensive model selection process, the best model according to predictive ability and DIC was selected and is shown as Equation 5-4.

$$log(\lambda_{ijk}) = \beta_{0i} + \beta_{1i} * (Entering\ Vehicles)_{ijk} + \beta_{2i} * (Num.\ of\ Intersection\ Legs)_{ijk} \tag{5-4}$$

$$+ \beta_{3i} * (Min.\ Num.\ of\ Lanes)_{ijk} + \beta_{4i} * (Max.\ Roadway\ Width)_{ijk} + \beta_{5i} *$$

$$(Max.\ Speed\ Limit)_{ijk}$$

where: β_{0i} = intercept by reference class, i,

β_{1i} = coefficient for entering vehicles by reference class, i,

β_{2i} = coefficient for number of intersection legs by reference class, i,

β_{3i} = coefficient for minimum number of lanes by reference class, i,

β_{4i} = coefficient for maximum roadway width by reference class, i,

β_{5i} = coefficient for maximum speed limit by reference class, i

5.2.3 Intersection Ranking

The ranking of the intersections based on the model results is an important step to provide UDOT with a list of intersection safety hot spots at the state, region, and county levels.

The ranking for the UICPM is determined based on the actual number of crashes at an intersection compared to the posterior predictive distribution for each intersection, determined by the UICPM. Posterior probability densities are obtained for each parameter, β, showing the most likely values for each coefficient. These probability distributions are used to obtain a posterior predictive distribution for the number of crashes at each intersection, j. The end goal of this process is to estimate the probability that the number of crashes occurring at an intersection in a particular year, y_{ijk}, is less than or equal to the average annual number of crashes. This probability is represented by the percentile value in the model results. The posterior distributions for all coefficients, β, are analyzed to understand the effects the variables have on the total number of crashes. The coefficients are then used to simulate data from each intersection to determine the variability in crashes expected at intersection j, thereby drawing from the posterior predictive distribution.

An extreme percentile value indicates a hot spot or a large amount of variability that cannot be explained adequately with the UICPM. It is suggested that a high value is indicative of a lurking effect, which may indicate that something at the intersection, not accounted for in the model, may be causing a higher number of crashes than expected. Therefore, intersections that have extreme percentile values require the attention of UDOT Region directors to identify potential safety improvements. The percentile values are expressed as proportions. For example, if an observed number of crashes, y_{ijk}, is in the 0.95-percentile, then 95 percent of the draws from the expected distribution for y_{ijk} were below the actual value for y_{ijk}. The actual number of crashes is therefore higher than the estimate of the model for the number of crashes.

5.3 Model Execution

The R GUI Excel tool was developed to provide a simple GUI for the user to prepare the UICPM input file and execute the model in R. Once the user has created the UICPM input file using the R GUI, the UICPM Variable Selection screen appears, as shown in Figure 5-1. Here, the user enters the UICPM input file path, the file path of the R model file, and the number of desired iterations for the model. The user then chooses the variables that will be used in the model run. After selecting the "Start Statistical Analysis" button, a command line is sent using VBA macros to open the R model file and run the model based on the parameters given on the UICPM Variable Selection screen. The model is then run in R and results are provided in the given working directory in the format of Excel files and PDF printouts.
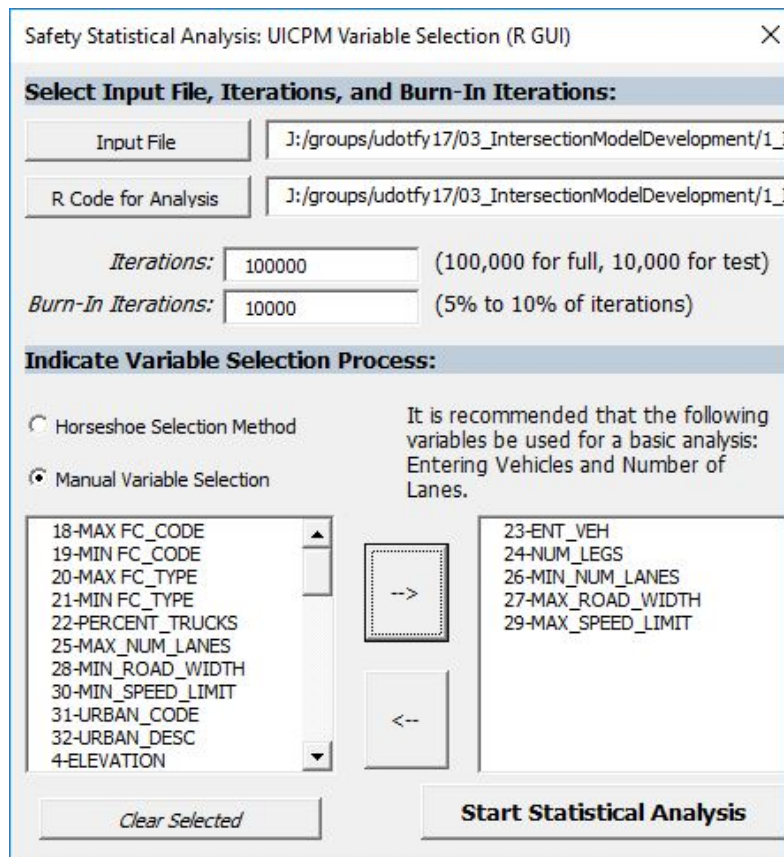
Figure 5-1: UICPM Variable Selection screen.

5.4   Model Results

The R model outputs four files as a result of each model run that are saved in a new folder entitled "Crash Analysis," followed by the date and time of the model run. The first file is a text file that shows the basic format of the regression equation that is produced internally when the model runs. Two of the files are CSV files with the results of the model. One of the results files varies slightly from the other because it is formatted to be uploaded to the UDOT SafeMap online interface. The other results file is meant to be used to create the two-page reports using the Report Compiler tool. The fourth file is a PDF file containing plots of the UICPM results. One plot on the file compares the expected crash count and the actual crash count of all intersections represented in the dataset, as shown in Figure 5-2. The expected annual crash counts are shown along the x-axis, and the actual annual crash counts are shown along the y-axis. A blue dashed line shows where the points would be located if the expected number of crashes were equal to the actual number of crashes. The top 20 intersections on the state level are shown in red on the plot.
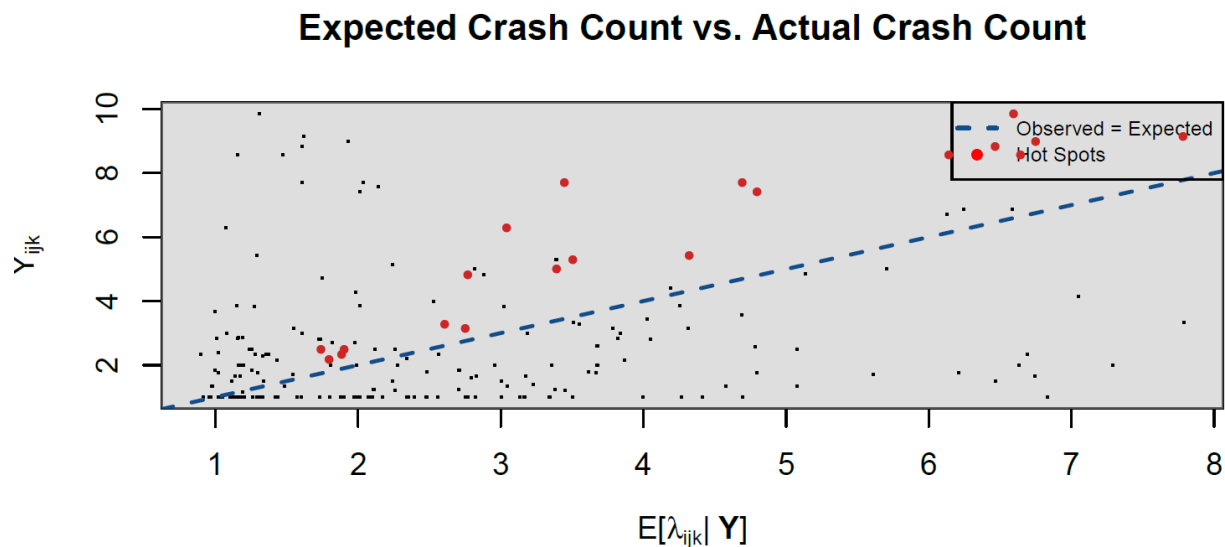
Figure 5-2: UICPM crash count plot.

Also included on the PDF output file are distribution plots for the top 20 intersections on the state level. An example of one of these is shown in Figure 5-3. The plot is labeled with the assigned intersection ID number and the approximate latitude and longitude of the intersection location. The plot shows the distribution of expected annual crashes for an intersection with the same characteristics of that particular intersection. This distribution is based on the regression analysis of the given roadway and crash data. Also shown is a red dashed line showing how many annual crashes actually occurred at the intersection. If this line appears on the far end of a distribution, it means that more crashes happened at the intersection than expected. The plot legend gives the percentile value, which describes where the actual number of annual crashes falls along the distribution for that intersection. As shown in the figure, the example intersection 145 has a percentile of 0.961, meaning that it has significantly more crashes per year than the expected average for an intersection with its same characteristics.
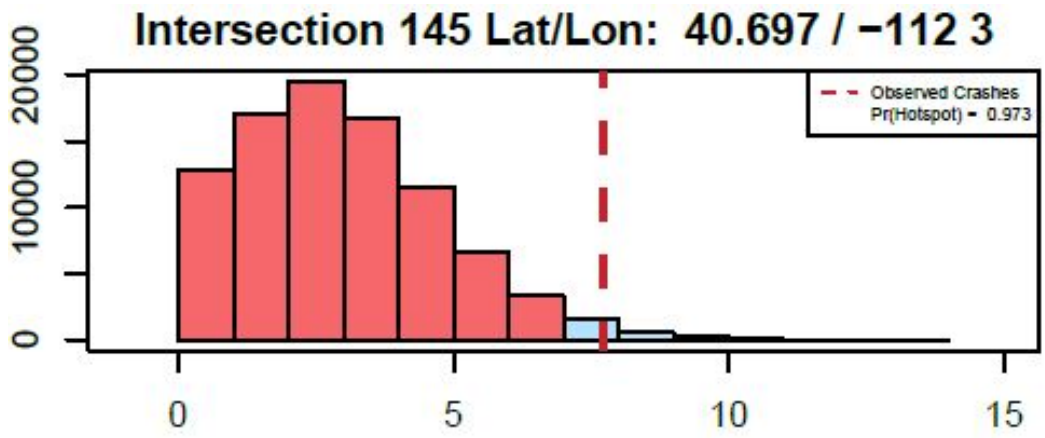
Figure 5-3: UICPM intersection distribution plot.

5.5   Summary

The purpose of the statistical intersection safety model chapter is to discuss the UICPM development, how it is executed, and the model results. The UICPM was developed as a

Bayesian generalized linear regression model using R statistical analysis software. An extensive variable selection process was completed to determine the explanatory and categorical variables that would be included in the model. The Bayesian horseshoe and backward selection methods were used to determine five significant variables to use in the UICPM, including: entering vehicles, number of intersection legs, minimum number of lanes, maximum roadway width, and maximum speed limit. It was noticed that the effects of these variables changed when applying categorical variables for functional classification and UDOT Region. It was found that these two variables have an interaction effect; therefore, they were used in the UICPM as an interaction term to categorize the data. The intersections in the UICPM are ranked based on how extreme the actual number of crashes at each intersection is compared to the number of crashes at the rest of the intersections in the population. The model is executed using the R GUI tool, as was done in the RSAM. The model results are provided in an Excel document, and crash distribution plots of the top-ranked intersections are given in a PDF document. These model results are used to create two-page reports and GIS maps that make the results readily available for UDOT employees.

6   REPORTS CREATION

6.1   Overview

The implementation of the results of any statistical analysis is critical to the success of

the analysis. For this reason, the BYU research team developed processes to create reports and

maps to describe the results of the model and possible actions to be taken based on the results.

These tools have been developed in Excel and ArcGIS software. The purpose of this chapter is to

discuss how the results of the intersection safety model are analyzed and prepared in report

format and using atlas mapping tools. Three topics are provided in this chapter to discuss reports

creation. First, a discussion is provided of the creation of geospatial maps of the model results.

The second topic is a discussion of the safety analysis reports that are created to describe the

model results. The last topic is a discussion of the Report Compiler tool that is used to create the

safety analysis reports.

6.2   Model Results Maps

The purpose of creating mapped intersection safety model results is to provide a

graphical representation of the results for the use of state and region UDOT officials. The maps

created for the UICPM results are very similar to those created for the RSAM process, except

that the UICPM results are points on the map of the intersections that were studied, rather than

roadway segments. Maps are created showing the UICPM results in the state, region, and county

boundaries individually. The intersections are color-coded based on their state rank.

ArcMap GIS software is used to create the atlas maps. An ArcMap file called "StatModel_AtlasCreator" has been created as a template for the maps. The file has shapefiles loaded for the state, region, and county boundaries. A map layout template has also been created in the map file in preparation for the map atlas creation. The user imports the UDOT Linear Referencing System (LRS) Routes shapefile and the model results to the map file. The data are then processed using the following ArcMap tools: "Adjust Route Name," "Plot Statistical Model Results," and "Map Creator." The user makes some manual adjustments before the maps are produced. The "Adjust Route Name" tool changes the format of the UDOT LRS Routes route name so that it matches the route number given in the UICPM model results. The "Plot Statistical Model Results" tool plots points on the map representing intersections. This is done based on a route and milepoint given in the intersection data. The user then adjusts the symbology of the plotted points to show hot spots in red and intersections of low risk safety performance in green. The "Map Creator" tool is then run to create the maps at the state, region, and county levels, respectively. An example state map produced by ArcMap is shown in Figure 6-1.

6.3   Intersection Safety Analysis Reports

The communication of the results of the intersection safety model to UDOT personnel is critical to its effectiveness. As UDOT Region directors learn of the results of the model, plans can be made to mitigate safety concerns at the intersections in their respective region. For this purpose, the ISAR format was created. The ISARs give all the necessary safety information to UDOT Region directors in a simple format. The reports are auto-populated using a tool called the "Report Compiler." A full description of this tool is found in Section 6.4.

State of Utah
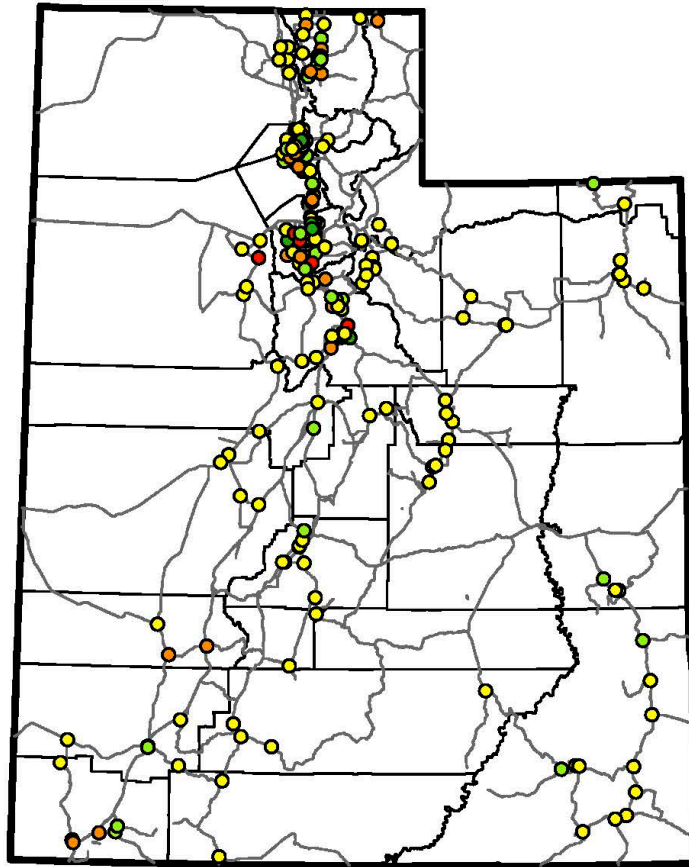
Statistical Model: UICPM
Date of Analysis: 1/11/2018
Years of Crash Data: 2010-2015
Severities in Analysis: 3, 4, 5

Note:
This map shows the results of a
safety statistical analysis of
roadways maintained by UDOT
within the state of Utah.
The legend shows "Rank Percentile"
of a given roadway compared
to all othe roadways in the network.
This data is not for public distribution.
The crash data are protected
under 23 U.S.C. 409.

0 15 30   60   90   120
Miles

Legend

UICPM

State_Rank

● 0-5%

● 5-20%

● 20-80%

● 80-95%

● 95-100%
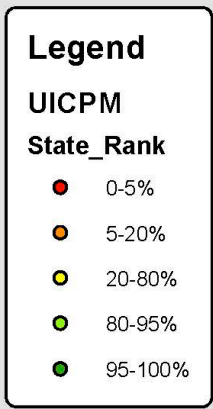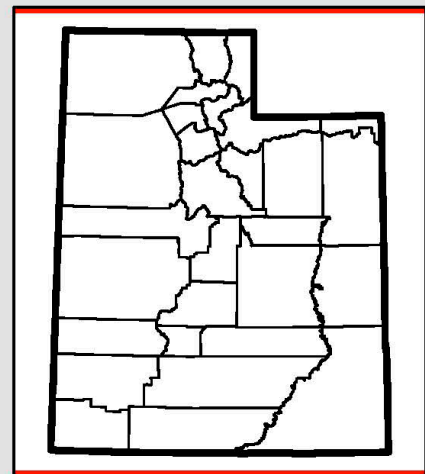
Figure 6-1: Example state map showing UICPM results.

72

The ISAR is divided up into five sections to provide useful intersection and crash data via tables and text: "Introduction," "Intersection Identification and Roadway Characteristics," "Micro-Analysis of Crash Data," "Historical/Current Conditions and Site Visit Notes," and "Countermeasures." A description of each section in the report is given in the following sections.

6.3.1 Introduction

The introduction of the ISAR is simply a description of the purpose and contents of the ISAR. This section reminds the report viewer of the contents of the report and how that information may be useful in their work.

6.3.2 Intersection Identification and Roadway Characteristics

The purpose of the second section of the ISAR is to provide roadway data of the intersection of interest and to describe its ranking amongst other segments in the state, region, and county. This section does not include details of the crashes at the intersection; however, it does provide the years of crash data that were used to assign crashes to the intersection. All of the data in this section is auto-populated by the Report Compiler tool and comes directly from the model results spreadsheet file.

The first table in this section, Table 1, is called "Intersection Metadata." An example of this table is shown in Figure 6-2. This table contains basic information regarding the location of the intersection and how it fared in the intersection safety statistical model compared to other intersections. All intersecting state routes at the intersection of interest are listed in Table 1, along with the respective milepoints where the intersection occurs. All intersections contain route and milepoint data for at least two routes. A third route may intersect at the same location, though this is not common. The latitude, longitude, and general metropolitan area of the

intersection are also included to provide additional information regarding the location of the intersection. Table 1, shown in Figure 6-2, also provides information about the statistical model that was used and its results. The abbreviated name of the model (i.e., "UICPM") is provided, along with the state, region, and county ranks. The region number and county name are also given as are the years of crash data are also provided in the table.

Table 1: Intersection Metadata					
Road 1 Name & MP:	171	3.5	Model Used:	UICPM	
Road 2 Name & MP:	172	3.99	State Rank:	1	
Road 3 Name & MP:	N/A		Region & Rank:	2	1
Latitude / Longitude:	40.69655	-112.0249	County & Rank:	SALT LAKE	1
Years of Data Source:	2010-2015		Metropolitan Area:	Salt Lake West	

Figure 6-2: Intersection Metadata table in the ISAR.

The second table in this section, Table 2, is called "Intersection Characteristics." This table provides key information regarding the geometry, regulations, and function of the intersection of interest. An example of this table is shown in Figure 6-3. The table indicates the control of the intersection, such as signal or stop. The average daily entering vehicles is also provided in the table, along with three intersection characteristics that are given with a maximum or minimum value. This is done to represent the two or three roadways that intersect at that location. The maximum and minimum functional classifications are provided, where the maximum functional classification is the more arterial classification of the two given classifications. The maximum and minimum values for number of through lanes and speed limit are also provided.

Table 2: Intersection Characteristics			
Intersection Control:	Signal	Ave. Entering Vehicles:	29,327
Max. Functional Class:	Other Prinicpal Arterial	Min. Functional Class:	Other Prinicpal Arterial
Max. # of Thru Lanes:	4	Min. # of Thru Lanes:	4
Max. Speed Limit, mph:	45	Min. Speed Limit, mph:	45

Figure 6-3: Intersection Characteristics table in the ISAR.

6.3.3 Micro-Analysis of Crash Data

The purpose of the third section of the ISAR is to provide details of the crashes that occurred at the intersection. There are three tables in this section that report the number of crashes by severity at the intersection, functional area method used, the most prominent crash factors at the intersection, and details about each crash event. The data in this section comes from the model results file as well as a file containing model parameters and crash data details.

The first table in this section, Table 3, is called "Crash Count and Severity." An example of this table is shown in Figure 6-4. This table includes information on the severities that were used in the model and what method was used to calculate functional area. The number of crashes predicted by the model, based on the explanatory variables, is shown alongside the actual number of crashes. This actual number of crashes corresponds only to the number of crashes for the chosen severities. The number of crashes at the intersection are also summarized by severity in this table.

Table 3: Crash Count and Severity								
Crash Severities Used	Functional Area Method Used	Number of Crashes						
		Predicted	Actual	Sev. 5	Sev. 4	Sev. 3	Sev. 2	Sev. 1
345	Speed Limit	22.0	47	1	4	42	87	263

Figure 6-4: Crash Count and Severity table in the ISAR.

The second table in this section, Table 4, is called "Crash Factors." An example of a portion of this table is shown in Figure 6-5. This table provides a summary of the seven most prominent crash factors of the crashes at the intersection. The columns of crash factors are listed in order of prominence. A list of the analyzed crashes at the intersection is provided, with crashes listed by crash ID and geographical coordinates. A "Y" (yes) or "N" (no) is listed under each crash factor for each crash indicating whether the crash factor applied to that particular crash or not. A summary of the number of crashes applicable to each prominent crash factor is listed at the bottom of the table in the "Intersection Total" row.

Table 4: Crash Factors

Crash ID	Latitude	Longitude	ROADWAY GEOMETRY RELATED	NIGHT DARK CONDITION	TEENAGE DRIVER INVOLVED	HEADON COLLISION	DUI	SPEED RELATED	UNRESTRAINED
10541523	40.6518	-112.025	N	Y	N	N	N	Y	N
10681707	40.6518	-112.025	Y	N	N	N	N	N	N
(continued...)									
Intersection Total			22/43	15/43	14/43	7/43	7/43	5/43	4/43

Figure 6-5: Crash Factors table in the ISAR.

The last table in this section, Table 5, is called "Vehicle and Crash Data." An example of a portion of this table is shown in Figure 6-6. This table provides detailed information regarding each of the intersection crashes that were included in the analysis. The crashes are listed by crash ID. The following is a summary of the information that is provided in each data column in the table:

- Number of Vehicles: The number of vehicles that were involved in the crash.

- First Harmful Event: The first harmful event that led to the crash.

- Manner of Collision: The manner in which the vehicle(s) collided (e.g., angle, head-on, sideswipe).

- Events: A list of the events that occurred during the crash for any vehicle ("Not Applicable" means no more events than the first ones listed occurred).

- Most Harmful Event: The most harmful of the events that occurred.

- Vehicle Maneuver: The maneuvering movement being made by the first vehicle listed when the crash occurred.

Table 5: Vehicle and Crash Data

Crash ID	# of Vehicles	First Harmful Event	Manner of Collision	Event (1)	Event (2)	Event (3)	Event (4)	Most Harmful Event	Vehicle Maneuver
10541523	2	Motor Vehicle in Transit	Angle	Crossed Median/Centerline	Operating Motor Vehicle	Not Applicable	Not Applicable	Motor Vehicle in Transit	Straight Ahead
10681707	2	Motor Vehicle in Transit	Angle	Operating Motor Vehicle	Not Applicable	Not Applicable	Not Applicable	Motor Vehicle in Transit	Straight Ahead

Figure 6-6: Vehicle and Crash Data table in the ISAR.

Following Tables 3, 4, and 5 in the ISAR, there is a section called "Safety Problem Summary." This section is not automated and must be completed by the user. The purpose of this section is to define the safety problems that exist at the study intersection based on the crash and vehicle data listed in Tables 3 through 5. A valuable source for determining the safety problems at the intersection is the prominent crash factors. These crash factors tend to describe the prominent safety issues at a certain location. This safety problem can help other viewers understand quickly the safety problems in order to better identify and understand how to respond to them.

6.3.4 Historical/Current Conditions, Site Visit Notes

The purpose of the fourth section of the ISAR is to provide the report user with additional information to help in determining appropriate countermeasures and safety improvements for the study intersection. This section is to be filled out by the user and is not automatically generated. The intent is to have the user complete a historical analysis of the location using online tools including UDOT's Roadview Explorer (Mandli 2018) and Google Earth (Google 2018). The user may then make a site visit to the location to verify given information and further assess the safety conditions.

The historical perspective of the study intersection is completed to fully understand how the intersection has changed over time. This analysis reveals recent improvements that have been made to the intersection including lane geometry and intersection control. This is primarily accomplished using Google Earth historical imagery (Google 2018) and UDOT's Roadview Explorer tool (Mandli 2018).

The purpose of the current conditions and site visit notes is to analyze the intersection in its current state and pinpoint crash factors at the intersection. The current conditions can be determined primarily using current Google Earth imagery (Google 2018) and UDOT's Roadview Explorer tool (Mandli 2018). The intersection geometry can be confirmed with a physical site visit. The purpose of the site visit is to review the information in the other sections of the ISAR and to conduct a more detailed analysis of the intersection. It is expected that all of the ISAR can be filled out except for the site visit section so that the person who goes on the site visit can take a single-page report, front and back, that provides a summary of the background information on the intersection. There is also a spot on the ISAR where the user can put an image taken during the site visit or from the mentioned online tools.

6.3.5 Countermeasures

The purpose of the fifth section of the ISAR is to provide a list of possible countermeasures that are generated based on the prominent crash factors at the intersection. These countermeasures are drawn from the NCHRP 500 Series reports and the Countermeasures that Work (CTW) database. The countermeasures from the NCHRP 500 Series reports are reported with a description of "Tried" (T), "Proven" (P), or "Experimental" (E), which are used to describe the effectiveness of each countermeasure (Neuman et al. 2003). A thorough summary of the countermeasures found in the NCHRP 500 series is found in the literature (Schultz et al. 2013b). The countermeasures from the CTW database use a star rating system of one to five stars for each countermeasure for a similar purpose (Goodwin et al. 2015). These rating systems are listed with each countermeasure that appears in the ISARs to help the user select countermeasures that are most applicable and effective.

6.4   Report Compiler

The process of creating ISARs has been automated using an Excel-based tool, called the Report Compiler. This tool was used for the RSAM process but has now been updated for the intersection model. The Report Compiler allows the user to input model results and relevant crash data, select the intersections of interest, and produce the ISARs.

The Report Compiler requires two input files to create the ISARs. The first file is the intersection model results. This file should be in a comma-delimited spreadsheet format. It contains a data row for each intersection, its roadway and crash information, and the model results. The second file is a parameters file, which contains information regarding the severities used in the model, the functional area method used, and data about each individual crash that was assigned to the study intersections. The parameters file provides critical information about

assumptions made in the model and individual crashes that the model output file does not provide.

The Report Compiler uses VBA macros and GUIs to load the input data and create the ISARs as specified by the user. One significant feature that is now available in the Report Compiler is the ability to choose the intersections of interest. The model results are summarized using VBA macros to determine intersections in the state, each region, and each county. The user is then presented with a GUI that lets them choose the intersections of interest, as shown in Figure 6-7. If intersections are chosen by state, the user enters the number of top-ranked state intersections to include in the reports. If intersections are chosen by region or county, the user first selects the regions or counties to include. The user then chooses the number of top-ranked region or county intersections to include in the reports. The GUI informs the user of how many intersections are represented in each region or county in the model results. If intersections are chosen individually, the user chooses intersections from a list of the represented intersections from the model results. The GUI lists intersections by intersection ID number and intersecting routes.

Once the list of intersections has been generated, an individual ISAR is then created using VBA macros for each of the chosen intersections. The reports are saved as individual Excel files. The user then can review and complete each report. As described previously, this entails the user reviewing the crash data and factors to identify trends that may exist at an intersection. The user then details the historic and current conditions of the intersection. A site visit is completed to confirm any trends that were speculated. The user then selects countermeasures from the auto-populated list that may improve the safety at the study intersection. Once complete, the reports are summarized in a two-page format and sent to the

UDOT Regions for their review. The two-page format allows region representatives to take the

report to the field for further review and decision-making.
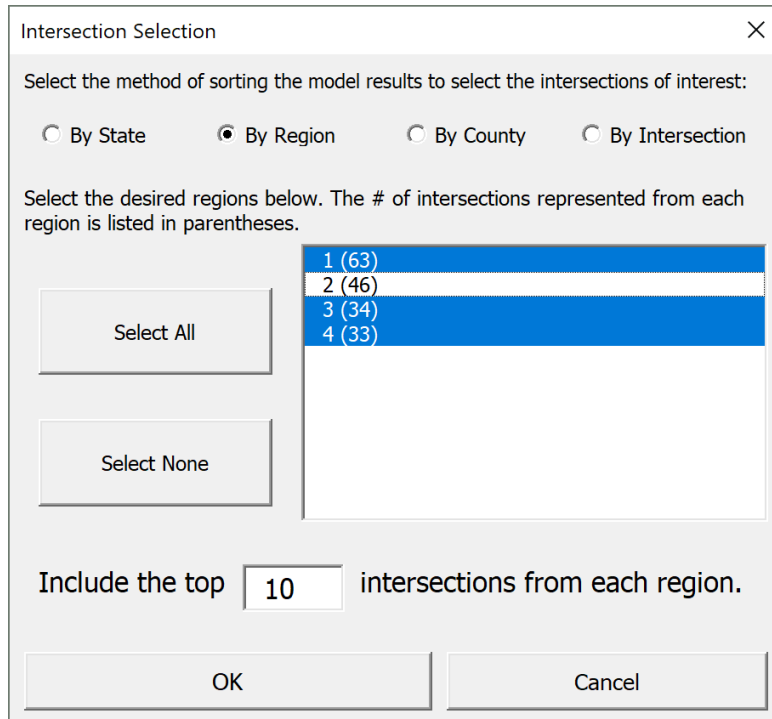
Figure 6-7: Report Compiler Intersection Selection GUI

6.5   Summary

The purpose of this chapter was to discuss how the results of the intersection safety

model are analyzed and prepared in a report format. The model results are mapped geospatially

by state, region, and county to provide a visual representation of the intersection hot spots. The

model results are then published in an ISAR, which contains data regarding the intersection

geometry, roadway classification, crash history, and possible countermeasures. The user

completes a site visit to complete each ISAR. The ISAR reports are created using the Report

Compiler tool.

7   EXAMPLE APPLICATION OF THE INTERSECTION SAFETY ANALYSIS METHODOLOGY

7.1   Overview

The purpose of this chapter is to provide an example of how a user prepares a UICPM input file, runs the model, and prepares output for the results. The discussion in this section is meant only to describe the process and is not intended to be a complete user's manual with step-by-step instructions. First, a discussion is provided on the operation of the R GUI to prepare data for the UICPM. The second topic is a discussion on the operation of the R GUI and R to execute the UICPM. Lastly, a discussion is provided on the creation of GIS maps and reports that show the results of the UICPM.

7.2   Data Preparation

As discussed in Chapter 4, the roadway and crash data preparation tools have been transferred from the RSAM "Roadway and Crash Data Preparation" MS Excel spreadsheet to be integrated in the R GUI. This was done to simplify the data preparation process overall. Now the processes to import data, process the data, and run the intersection safety model are included in one MS Excel spreadsheet tool. Additional screens and GUIs have been added to the R GUI to integrate these processes. This section includes a general overview of the features of the new R GUI.

When the user opens the R GUI, there are two sheets in the MS Excel workbook; these are the "Home" and "Progress" screens. The "Home" screen is where the user starts the GUI, and the "Progress" screen updates the user on the progress of the macros as they run. After the user pushes the "Start GUI" button on the "Home" screen, a start window appears. The user then selects a working directory, or a folder, where all new data files will be stored. The user selects an "Rscript" file from the latest version of R, which allows an R model to be initiated from MS Excel. The GUI then checks that all R packages are up to date with the selected Rscript program. Once the working directory and Rscript program is selected and the R packages are up to date, additional options appear in the window as shown in Figure 7-1. The user then chooses the model to run, so that the correct data can be prepared for the model.
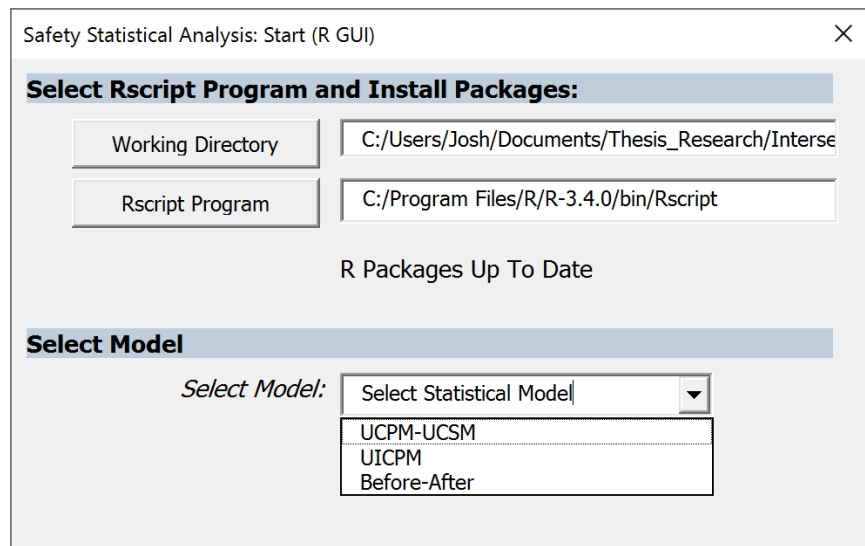

Figure 7-1: R GUI start screen with model selection.

After the model has been selected, the user is then asked to create a new input file or use an existing input file, as shown in Figure 7-2. If the user already has a model input file saved from a previous R GUI iteration, then the "Use Existing Input File" button can be selected to

skip ahead to the variable selection screen. However, if the input has not been created or the user

wants to create a new one, then the "Create Input File" button should be selected.
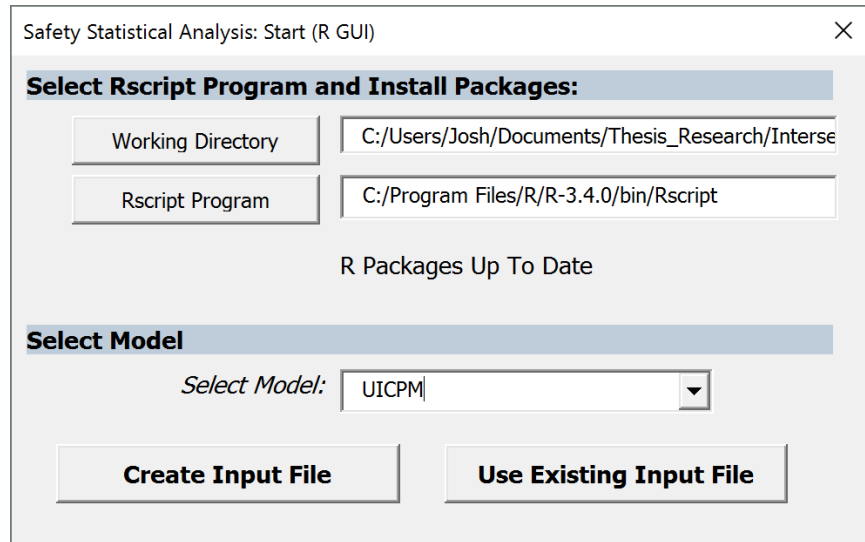
Figure 7-2: R GUI start screen with input file buttons.

Once the "Create Input File" button is selected, the UICPM input window appears, as

shown in Figure 7-3. This window was used in the previous version of the R GUI only to assign

crash data to their respective roadway segments based on previously created roadway and crash

datasets. However, now the user has the option in the R GUI to compile the roadway and crash

data, respectively. This is done by selecting the "Create Input Datasets" button in the UICPM

input window. If the user has already created separate roadway and crash data files in a previous

R GUI iteration, then it is not necessary to create the input datasets again unless the user wants to

create a new file.

After the user selects the "Create Input Datasets" button on the UICPM input window,

the "Intersection Data Preparation" window appears, as shown in Figure 7-4. The purpose of this

window is to help the user input raw roadway and crash data and export combined roadway and

crash data files. This window has several buttons, one for each roadway and crash data input file.

The user selects each button under "Roadway Data" or "Crash Data" to select the data file path

for each file. Once all the file paths have been selected, they appear in the text boxes to the right

of the buttons. The user can then select the "Combine" button to combine the roadway or crash

data into a single file. The raw input data files are formatted using the VBA macros and

combined into a single dataset. Once one dataset has been created, the user is brought back to the

UICPM input window, where the file path of the newly-created data file is entered into its

corresponding text box. The user can then select the "Create Input Datasets" button again to

create the other dataset if needed.

Figure 7-3: UICPM Input screen.

Figure 7-4: Intersection data preparation window.

Once there are file paths listed for the road intersection and crash data, the user can then select which crash severities to include in the analysis and the desired intersection functional area, as shown in Figure 7-5. The crash severities can be selected by each check box or by selecting the "Select All" or "Select None" buttons. The intersection functional area is a distance from the intersection stop bars where the intersection functionality may impact driver and

vehicle behavior. The R GUI assigns crashes to an intersection that fall within the given

functional area radius of the intersection. The UICPM input window allows the user to define the

functional area based off speed limit, functional classification, or urban code. The user can select

either the "Recommended Functional Area" or a "User-Defined" functional area. The

recommended functional area is based on speed limit and comes from the TRB Access

Management Manual (Williams et al. 2014). The user-defined functional area can be based on

speed limit, functional classification, or urban code. The user can select one of these options

from the drop-down menu on the UICPM Input window. The user can use the default values for

the chosen category, or they may change the values, as shown in Figure 7-6.

Figure 7-5: Complete UICPM input window.
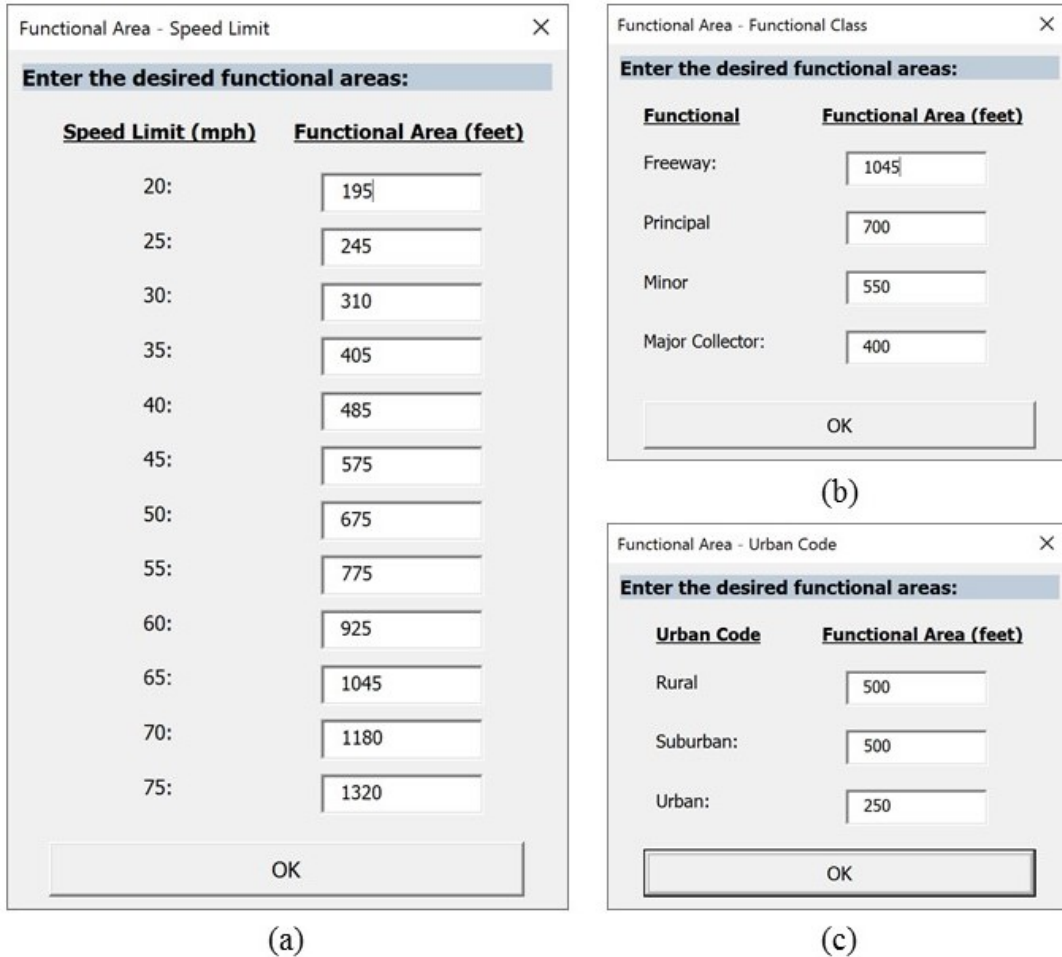
Figure 7-6: Functional area definition windows.

7.3 Model Execution

Once the intersection model input file has been created, the user is ready to run the model in R. The R GUI tool used for the data preparation steps is also used to run the model, as was done with RSAM. The user first navigates to the UICPM Variable Selection screen, as shown in Figure 7-7. This screen automatically appears after the user creates the model input file. Two buttons are provided on this screen for the user to browse to the model input and R model files. The user then chooses the number of iterations to perform in the model. The more iterations that are performed, the more accurate the model is. A full model run should have 100,000 iterations and 10,000 burn-in iterations.
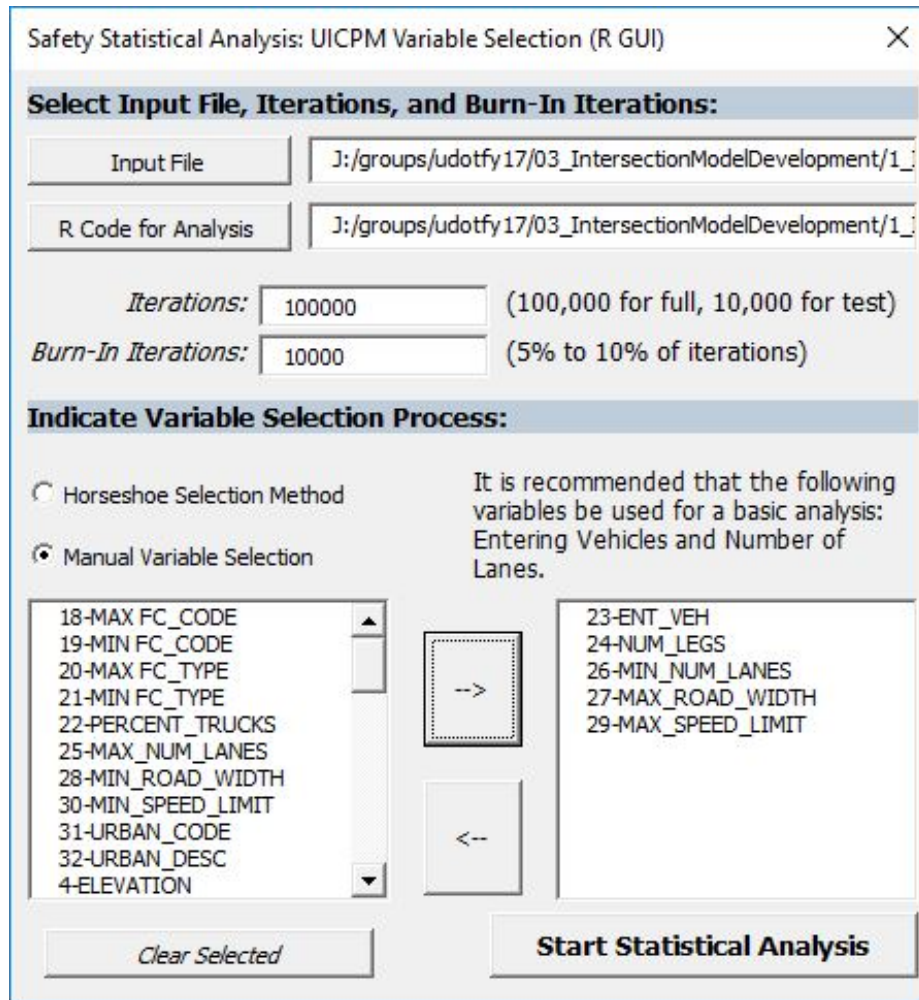
Figure 7-7: UICPM Variable Selection screen.

Once the input file, R model file, and number of iterations has been specified, the variable selection options appear. The user can choose to use the Bayesian horseshoe selection method or manually select variables. As of the writing of this document, the automation of the Bayesian horseshoe selection method is still in development; therefore, the user is only able to manually select the variables at this point. The user can choose from the variables listed in the list box on the left side of the screen. Variables in the list box on the right side of the screen are those that have been selected. By default, the VBA macros are used to select speed limit, minimum number of lanes, number of entering vehicles, and percent trucks for the analysis. As

discussed in Section 5.2.1, it was found that these variables are the best predictors of the number of crashes at an intersection. Once the variables are selected, the user selects the "Start Statistical Analysis" button to run the model.

The chosen parameters are submitted in a command line using VBA macros to be run in R. A Windows terminal prompt appears with populating text on the screen to show the progress of the model. The model runs for at least a few minutes before informing the user that the process is complete. The model results are saved in the working directory under a new folder called "CrashAnalysis" with the date and time of the model run. The working directory was initially chosen by the user upon starting the R GUI. Four files are saved in the new folder. These include a text file with a generic form of the regression equation, a PDF with plots of the crash count and 20 intersection crash distributions, and two Excel spreadsheets with the model results. The plot shows expected annual crash counts on the x-axis and actual annual crash counts on the y-axis. The top 20 ranked intersections at the state level are highlighted in red on the plot.

7.4   Report and Map Creation

After running the UICPM and obtaining results, the user is ready to produce reports and maps that are used to present the results of the model. The ISARs are created using the Report Compiler, and the maps are created in ArcGIS software. The purpose of the Report Compiler is to take the model results and create ISARs for select intersections. Upon starting the Report Compiler, the user chooses whether to create reports for segments or intersections, as shown in Figure 7-8. This is done because the automated processes for each are different. A user would choose to create segment reports if the UCPM or UCSM were run.
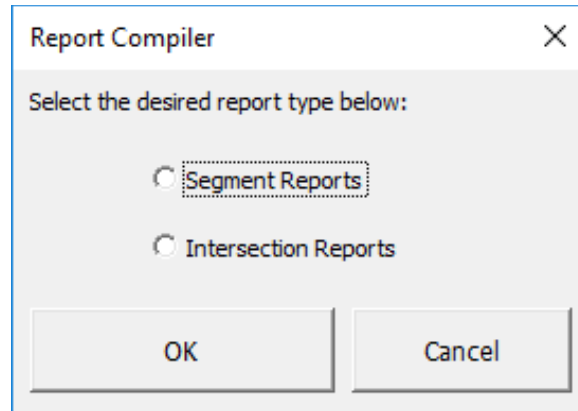
Figure 7-8: Report Compiler initial window.

The user is asked to open the UICPM model results workbook, located in the appropriate "CrashAnalysis" folder produced by the R model. The user is also asked to open the workbook with the crash data and model parameters. This file was saved when the model input file was created using the R GUI. The files are opened using VBA macros and a database is created of the intersections that are available in the model results to create reports for.

The "Intersection Selection" window appears, as shown in Figure 7-9. This GUI is used to select the intersections that will be reported on in the ISARs. The user has the option of selecting intersections by state, region, county, or individual intersection. If intersections are chosen by state, the user enters the number of top-ranked state intersections to include in the reports. If intersections are chosen by region or county, the user first selects the regions or counties to include. The user then chooses the number of top-ranked region or county intersections to include in the reports. The GUI informs the user of how many intersections are represented in each region or county in the model results. If intersections are chosen individually, the user chooses intersections from a list of the represented intersections from the model results. The GUI lists intersections by intersection ID number and intersecting routes.
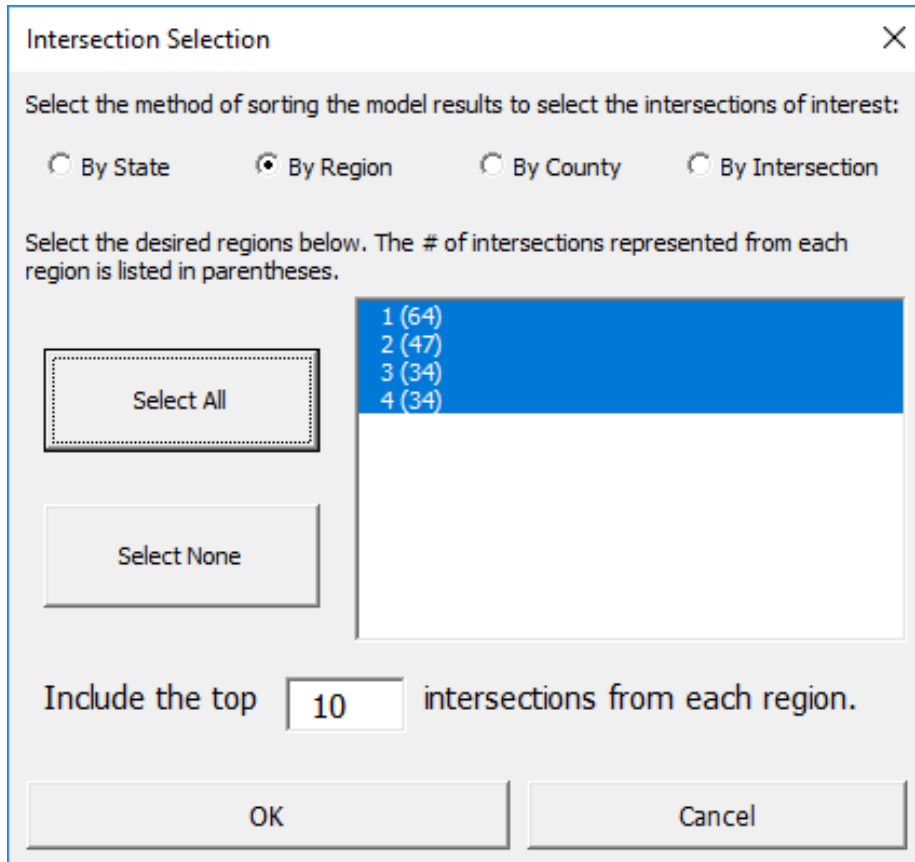
Figure 7-9: Intersection Selection screen.

The user is asked to select the folder to save the ISARs. An additional folder is created in the selected folder to store the ISARs, and the reports are created. The user is notified when they are complete, and the user is directed to the folder with the reports. The Report Compiler workbook automatically closes. The next step after creating the ISARs is to complete site visits at each intersection or virtually if the analysts are unable to go to the site. Once the site visit is complete, the user writes up the remaining sections of the ISAR and compiles the report into a two-page format. Two pages of an example ISAR are shown in Figure 7-10 and Figure 7-11.

The purpose of using ArcGIS with the ISAM is to produce state, region, and county maps with the model results. This process is best done on a local computer and not a network

connection. Before beginning, all relevant files should be copied to a local computer drive location so that the process runs quickly and smoothly.

Intersection Safety Analysis Report

Introduction

The purpose of this report is to summarize and present preliminary results from a safety-specific micro analysis on the identified intersections of interest. This report includes identification of the intersection, micro-analysis of the crash data, site visit notes, and a list of possible countermeasures.

Intersection Identification and Roadway Characteristics   *Date: 1/25/2018*

Table 1: Intersection Metadata

Road 1 Name & MP:	171	3.5	Model Used:	UICPM
Road 2 Name & MP:	172	3.99	State Rank:	1
Road 3 Name & MP:	N/A		Region & Rank:	2	1
Latitude / Longitude:	40.69655	-112.0249	County & Rank:	SALT LAKE	1
Years of Data Source:	2010-2015		Metropolitan Area:	Salt Lake West

Table 2: Intersection Characteristics

Intersection Control:	Signal	Ave. Entering Vehicles:	29,327
Max. Functional Class:	Other Prinicpal Arterial	Min. Functional Class:	Other Prinicpal Arterial
Max. # of Thru Lanes:	4	Min. # of Thru Lanes:	4
Max. Speed Limit, mph:	45	Min. Speed Limit, mph:	45

Micro-Analysis of Crash Data

Crash Data Summary

Table 3: Crash Count and Severity

Crash Severities Used	Functional Area Method Used	Number of Crashes						
		Predicted	Actual	Sev. 5	Sev. 4	Sev. 3	Sev. 2	Sev. 1
345	Speed Limit	**22.0**	**47**	1	4	42	87	263

Table 4: Crash Factors

Crash ID	Latitude	Longitude	NIGHT DARK CONDITION	TEENAGE DRIVER INVOLVED	SINGLE VEHICLE	HEADON COLLISION	ROADWAY GEOMETRY RELATED	OLDER DRIVER INVOLVED	DISTRACTED DRIVING
Intersection Total			13/47	10/47	10/47	8/47	7/47	6/47	5/47

Safety Problem Summary

The most prominent crash factor at this intersection was night/dark conditions. This is most likely due to lack of lighting at the intersection. Other prominent crash factors included teenage drivers and roadway geometry. It appears that Hunter High School is located to the south of the intersection, which may contribute to the high teenage driver crash involvement. Many of the crashes were headon crashes or single vehicle crashes. Others were angle or front to rear crashes. One fatal crash occurred at this intersection between 2010 and 2015.

Figure 7-10: ISAR example, page 1.

It was observed that the study intersection is a four-leg intersection of 3500 South and 5600 West in West Valley City. The intersection has two through lanes and one left-turn lane entering from all approaches. The southbound and northbound approaches each have a right-turn pocket. All approaches have two-way left-turn lanes upstream of the intersection. The eastbound approach to the intersection has only a three-lane cross-section upstream of the intersection. The second eastbound through lane begins approximately 300 feet before the intersection. Based on 2016 street imagery, the westbound and eastbound approaches have permitted-protected left-turn phasing. The northbound and southbound approaches have protected-only left-turn phasing.

According to Google historical satellite imagery, there have not been any significant changes to the intersection layout and the surrounding area since at least 1997. Only roadway striping has changed slightly over time. The accesses to local developments have also remained constant.

A virtual site visit was performed on Saturday, January 27, 2018 at approximately 5:00 p.m. using Google Earth satellite imagery and street view. It appears that there are a significant number of trees on the north and south sides of the east leg of the intersection. In some cases, this may limit the sight distance of westbound, northbound, and southbound vehicles. It was observed that there is close access spacing near the intersection, especially on the west leg of the intersection. The street lighting is fairly frequent along the south side of 3500 South, where a power line is located. There are two light poles at the intersection itself, but providing at least one more on the east side of the intersection may improve lighting at the location. An image of the intersection from the east is shown in Figure 1.
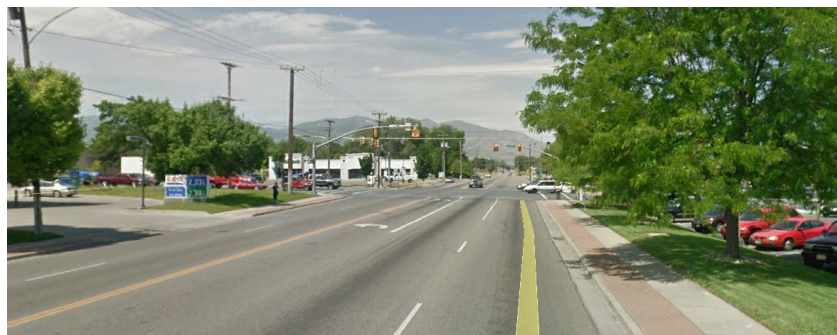
Figure 1: 3500 South / 5600 West intersection, looking to the west (Google).

Possible Countermeasures

The following is a list of possible countermeasure related to the top 8 crash factors listed in Table 5. The countermeasures listed were compiled from the NCHRP 500 Report volumes and Coutermeasures That Work (CTW). (P) = Proven (T) = Tried (E) = Experimental (NA) = Data not available (X*) = Star rating, as designated by CTW. (If countermeasures were listed in both the NCHRP 500 Report and CTW, it is listed with both ratings. For instance, Proven and 4-star rating = (P,4*).)

Communications and outreach on distracted driving (1*)
Decrease the number of poles along the corridor (P)
Eliminate early high school start times [i.e., before 8:30 a.m.] (T)
Enact a graduate licensing system (P,5*)
High visibility cell phone and text messaging enforcement (4*)
Improve design of roadside hardware (T)
Improve lighting at intersections, horizontal curves, and railroad grade crossings (T)
Improve roadway delineation (T)
Increase seatbelt use by older drivers and passengers (P)
Provide adequate sight distance for expected speeds (P)
Provide more protected left-turn signal phases at high-volume intersections (T)
Provide offset left-turn lanes at intersections (T)
Widen and/or pave shoulder (shoulder treatment) (P)

Figure 7-11: ISAR example, page 2.

94

The user opens the "StatModel_AtlasCreator" ArcMap file, developed for the purpose of creating the atlas maps. This ArcMap file contains several feature classes, including state, region, and county boundaries. The user will see an ArcMap toolbox with the "Adjust Route Name" and "Plot Statistical Model Results" tools. The user begins by creating a copy of the UDOT LRS Routes file and adding it to the map. The "Adjust Route Name" tool is then run to change the route name "089A" to "0011" to avoid confusion, and all routes are converted to integer values since that is the format of the intersection routes.

The user then adds the UICPM results CSV file to the map. Since this file doesn't have geospatial ties to the map yet, the "Plot Statistical Model Results" tool is run to plot the intersections on the map. The tool uses the route and milepoint referencing on the LRS Routes and UICPM results files to plot the intersections in the correct locations on the map. An example of how the settings of the "Plot Statistical Model Results" tool should look is shown in Figure 7-12.

When the tool is finished running, the output shapefile is saved to the default geodatabase of the map, called "AtlasMaker.gdb." The user then loads the shapefile into the map by selecting and dragging the shapefile from the Catalog pane to the map. The user then changes the symbology of the point shapefile to show colors based on the state rank of the intersection. The categorical symbology set for the intersections is shown in Table 7-1. The point sizes are changed to 6 points to make them more visible. The symbology of the LRS Routes line shapefile is changed to show as dark grey with a width size of 1. Now that the map contents are prepared, the user is ready to create the maps. The user selects the "Map Creator" Python script from the catalog and sets the parameters of the tool as shown in Figure 7-13. This tool can be run to create the state, region, and county maps. The map files are created as PDFs and saved in same folder

as the "StatModel_AtlasCreator" ArcMap file. A master file with all of the maps is created, as well as individual files for each map. An example of the state map is shown in Figure 7-14.
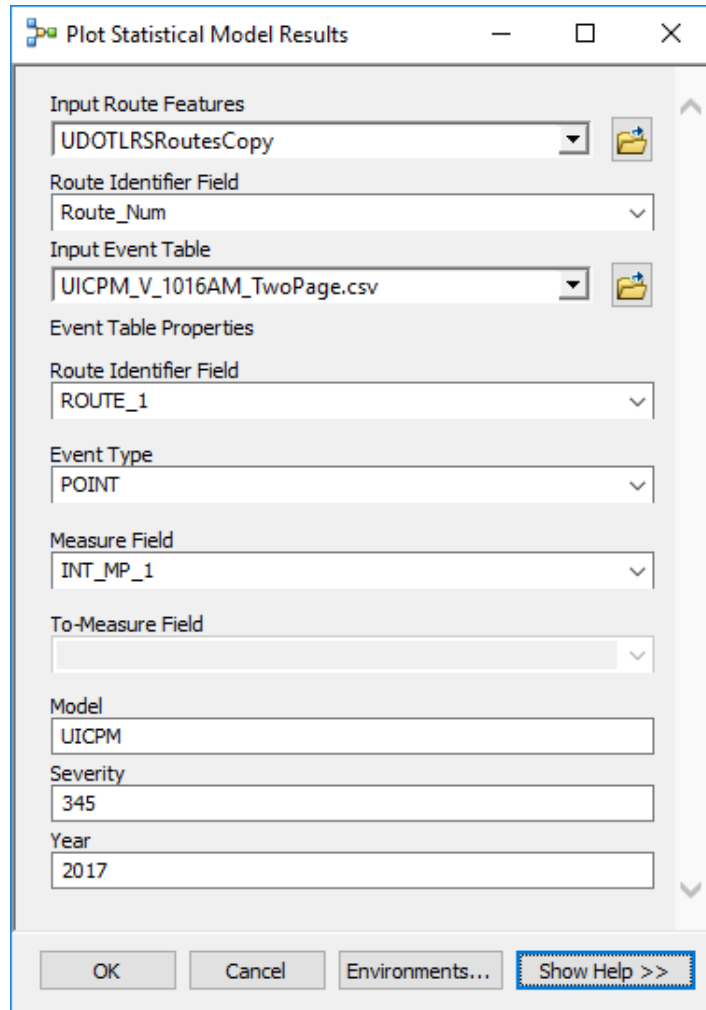
Figure 7-12: Plot Statistical Model Results tool settings.

Table 7-1: UICPM Categorical Symbology Settings

State Rank Percentage	Color
Top 5%	Red
5 – 20%	Orange
20 – 80%	Yellow
80 – 95%	Light Green
Bottom 5%	Dark Green
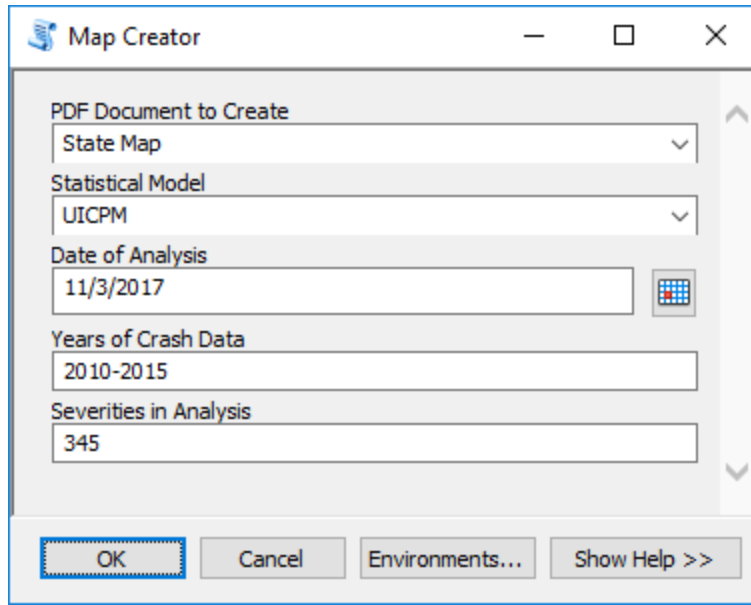
Figure 7-13: Map Creator Python tool settings.

7.5   Summary

The purpose of this chapter was to show an example of how a user completes the ISAM process to produce UICPM results and reports. The first step in this process is to prepare the input data and create an input file for the statistical model. The second step is to run the UICPM with the given input data. The next step is to create ISARs and GIS maps that show the results of the model. The final products of this process are model results in spreadsheet format, a document with plots of the statistical analysis, ISARs for select intersections, and maps of model results by state, region, and county. The results of the UICPM provide UDOT with the information needed to prioritize intersection safety projects in the state of Utah. The results also provide insights to the general safety trends that are seen in the results and the top-ranked intersections.

State of Utah

Statistical Model: UICPM
Date of Analysis: 1/11/2018
Years of Crash Data: 2010-2015
Severities in Analysis: 3, 4, 5

Note:
This map shows the results of a
safety statistical analysis of
roadways maintained by UDOT
within the state of Utah.
The legend shows "Rank Percentile"
of a given roadway compared
to all othe roadways in the network.
This data is not for public distribution.
The crash data are protected
under 23 U.S.C. 409.

0 15 30   60   90   120
Miles

Legend

UICPM

State_Rank

- 0-5%
- 5-20%
- 20-80%
- 80-95%
- 95-100%

Figure 7-14: Example state map of UICPM results.

8   ANALYSIS OF RESULTS

8.1   Overview

Once the UICPM has been run and results have been provided, it is necessary to analyze the results as a whole to find trends that will be helpful for UDOT going forward. The purpose of this chapter is to discuss the most recent results of the UICPM and trends of safety crash factors noticed at the top-ranked intersections in the state. The first topic is a discussion of the most recent UICPM results. The second topic is a discussion of the trends seen in the crash factors of the top-ranked intersections in the state and in each UDOT Region.

8.2   Model Results Discussion

The purpose of the model results discussion is to identify the top-ranked intersections in the state of Utah based on the UICPM and to discuss trends in the results. The latest iteration of the UICPM was completed in March 2018. The model was run with crashes of severities 3, 4 and 5 to focus on severe crashes in the analysis. Crash data and daily roadway volume data were drawn from the years 2010 through 2016. A total of 183 intersections were included in the dataset that was used to run the model. As discussed in Section 3.2, the intersections in the dataset are only intersections with at least two intersecting state routes due to the lack of non-state route data. The dataset contained 64 Region 1 intersections, 46 Region 2 intersections, 34 Region 3 intersections, and 39 Region 4 intersections. The top 25-ranked intersections in Utah based on the UICPM is shown in Table 8-1.

Table 8-1: UICPM Results from March 2018

State Rank	Latitude	Longitude	Region	Metropolitan Area	Actual Crashes/Year	Expected Crashes/Year	Percentile	Description
1	40.6966	-112.0249	2	Salt Lake West	7.7	3.5	0.9728	3500 S / 5600 W
2	40.6530	-112.0246	2	West Jordan	6.3	3.0	0.9636	5400 S / 5600 W
3	41.2443	-111.9700	1	Ogden	7.7	4.6	0.9041	1200 S / Washington Blvd
4	41.0727	-111.9789	1	Clinton	7.4	4.9	0.8805	Hill Field Rd / Main St
5	40.6998	-111.8883	2	Murray	9.9	6.6	0.8684	3300 S / State St
6	41.2312	-111.9787	1	Ogden	4.8	2.7	0.8579	20th St / Wall Ave
7	40.6966	-111.9389	2	Salt Lake Central	8.6	6.1	0.8315	3500 S / Redwood Rd
8	41.7777	-111.8340	1	Logan	5.3	3.7	0.8255	2500 N / Main St
9	40.5877	-111.9387	2	West Jordan	8.8	6.5	0.7931	9000 S / Redwood Rd
10	40.6744	-111.8883	2	Murray	8.6	6.6	0.7734	4500 S / State St
11	40.1150	-111.6548	3	Spanish Fork	2.5	1.7	0.7533	400 N / Main St
12	40.4319	-111.7852	3	Lehi	5.0	3.4	0.7413	Highland Hwy / 5300 W
13	40.6999	-111.8713	2	Murray	5.4	4.3	0.7289	3300 S / 700 E
14	41.3067	-112.0276	1	Clinton	2.2	1.8	0.7288	2700 N / 2000 W
15	41.1981	-112.0258	1	Ogden	3.1	2.6	0.7238	Hinckley Dr / 1900 W
16	40.2751	-111.7272	3	Orem/Provo	2.3	1.9	0.6920	University Pkwy / Geneva Rd
17	41.1674	-112.0260	1	Clinton	4.0	2.8	0.6916	5300 S / 1900 W
18	37.1683	-113.4079	4	Purgatory	2.5	2.0	0.6892	State St / 5300 W
19	40.1152	-111.6351	3	Spanish Fork	1.7	1.1	0.6871	400 N / State St
20	40.5510	-112.2981	2	Tooele	3.4	2.8	0.6869	1000 N / Main St
21	41.2004	-111.9710	1	Ogden	2.3	2.0	0.6837	Washington Blvd / Riverdale Rd
22	41.2419	-111.9454	1	Ogden	5.0	3.7	0.6810	1200 S / Harrison Blvd
23	40.8842	-111.8922	1	Centerville	3.6	2.9	0.6749	500 S / State St
24	37.2867	-109.5454	4	Bluff	1.5	1.2	0.6736	SR-191 / SR-162
25	39.9520	-111.9562	3	Eureka	1.3	1.2	0.6707	15200 S / 12800 W

As shown in the model results in Table 8-1, most of the top-ranked intersections are located in the northern part of the state. It was observed that 10 of the top 25 intersections in the state are located in UDOT Region 1, and eight of the top 25 intersections are located in UDOT Region 2. This is expected due to the high traffic volumes and populations in these urbanized areas. All of the top-ranked intersections are located in urban areas except for the intersections ranked 18, 24, and 25. These intersections are located in Purgatory, Bluff, and Eureka, respectively. Some routes had more than one top-ranked intersection including SR-171, SR-68, US-89, SR-126, and SR-172.

All of the top 25-ranked intersections have percentile values higher than 0.5. This indicates that the actual number of crashes per year of each intersection is higher than the median value of the projected crash distribution. The highest ranked intersections have percentile values just less than 1, indicating that the actual number of crashes for those intersections is included in the calculated crash distribution and is on the extreme high end of the distribution.

8.3   Crash Factor Trends Analysis

The purpose of the crash factor trends analysis is to identify common factors that contribute to crashes at the top-ranked intersections in the state and in each UDOT Region. The identification of these trends could be useful for each UDOT Region to understand the roadway safety needs in each area. This can lead to improved safety project prioritization across the state.

As discussed in Section 6.3.3, the ISARs summarize the most prominent crash factors at each intersection. These data come from the crash rollup data provided by UDOT. The crash factors include various roadway, environmental, vehicle, and crash characteristics that contribute to the event or severity of a crash. These data were summarized for the top 10 intersections in each UDOT Region to determine the crash factor trends of each UDOT Region. The crash factor

data for the four UDOT Regions were then combined to identify the trends statewide. Only the top 10 intersections of each UDOT Region were included in this analysis to emphasize the trends of the worst intersections in the state of Utah. Therefore, the statewide analysis included 40 study intersections total. As discussed in Section 8.2, the latest run of the UICPM only included crash severities 3, 4 and 5, and the crash data used in the analysis were from 2010 to 2016. Therefore, the crash factor trends discussed in this section are only based on crashes of severity 3, 4 and 5 from the years 2010 through 2016.

A summary of the most common crash factors in all UDOT Regions is shown in Table 8-2. As shown, approximately 25 percent of the severe crashes at the study intersections identified night/dark conditions as a crash factor, making it the most prominent crash factor for the 40 study intersections. Other prominent crash factors include teenage drivers, older drivers, and single vehicle crashes. These four factors are consistent through most UDOT Regions, and more variation is seen in the other common crash factors.

Table 8-2: Most Common Crash Factors for All UDOT Regions

Rank	Crash Factor	# of Crashes	% of Crashes
1	Night Dark Condition	238	25%
2	Teenage Driver Involved	175	19%
3	Older Driver Involved	169	18%
4	Single Vehicle	158	17%
5	Adverse Roadway Surface Condition	110	12%
6	Motorcycle Involved	84	9%
7	Roadway Geometry Related	81	9%
8	Distracted Driving	73	8%
9	Adverse Weather	72	8%
10	Pedestrian Involved	71	8%
11	Head-on Collision	69	7%
12	Unrestrained	59	6%

A summary of the most common crash factors for each UDOT Region are shown in Tables 8-3 through 8-6. Because the most common crash factors of each UDOT Region are generally night/dark condition, older drivers, and teenage drivers, it is important to note the other crash factors that are ranked high in each UDOT Region.

The most common crash factors in UDOT Region 1 are shown in Table 8-3. In UDOT Region 1 it was observed that adverse roadway surface conditions, adverse weather, and motorcycles were significant crash factors. The significant adverse weather crash factor may suggest that UDOT Region 1 experiences more extreme weather conditions than other parts of the state. This may contribute to adverse roadway surface conditions, which is another significant crash factor.

Table 8-3: Most Common Crash Factors for UDOT Region 1

Rank	Crash Factor	# of Crashes	% of Crashes
1	Night Dark Condition	77	25%
2	Teenage Driver Involved	66	21%
2	Older Driver Involved	66	21%
4	Adverse Roadway Surface Condition	47	15%
5	Single Vehicle	38	12%
6	Adverse Weather	31	10%
7	Motorcycle Involved	28	9%
8	Head-on Collision	25	8%
9	Unrestrained	24	8%
10	Pedestrian Involved	17	5%
10	Workzone Related	17	5%
10	Overturn Rollover	17	5%

The most common crash factors in UDOT Region 2 are shown in Table 8-4. In UDOT Region 2 it was observed that adverse roadway surface conditions, roadway geometry, and

pedestrians were significant crash factors. The significant adverse roadway surface conditions crash factor may suggest that the maintenance of roadways needs to be increased and improved in UDOT Region 2 in order to improve intersection safety. The significant roadway geometry crash factor suggests that the roadway geometry at certain intersections in UDOT Region 2 needs to be improved. The high-density urban land uses in parts of UDOT Region 2 may have resulted in several crashes involving pedestrians.

Table 8-4: Most Common Crash Factors for UDOT Region 2

Rank	Crash Factor	# of Crashes	% of Crashes
1	Night Dark Condition	121	28%
2	Single Vehicle	83	19%
3	Teenage Driver Involved	67	15%
4	Older Driver Involved	62	14%
5	Adverse Roadway Surface Condition	47	11%
6	Pedestrian Involved	43	10%
6	Roadway Geometry Related	43	10%
8	Motorcycle Involved	38	9%
9	Distracted Driving	37	8%
10	DUI	35	8%
11	Head-on Collision	34	8%
12	Adverse Weather	30	7%

The most common crash factors in UDOT Region 3 are shown in Table 8-5. In UDOT Region 3 it was observed that distracted driving, motorcycles, unrestrained passengers, and roadway geometry were significant crash factors. Also, it was observed that the most prominent crash factor was teenage drivers. This could suggest that there are more young drivers in UDOT Region 3 than other locations in the state. The significant distracted driving crash factor could be

related to the prominence of the teenage driver crash factor, with a high number of young drivers potentially using technology while driving.

Table 8-5: Most Common Crash Factors for UDOT Region 3

Rank	Crash Factor	# of Crashes	% of Crashes
1	Teenage Driver Involved	32	24%
2	Older Driver Involved	30	22%
3	Night Dark Condition	29	21%
4	Single Vehicle	25	19%
5	Distracted Driving	15	11%
6	Motorcycle Involved	13	10%
7	Roadway Geometry Related	12	9%
7	Unrestrained	12	9%
9	Adverse Roadway Surface Condition	10	7%
10	Pedestrian Involved	9	7%
10	Head-on Collision	9	7%
10	Bicyclist Involved	9	7%

The most common crash factors in UDOT Region 4 are shown in Table 8-6. In UDOT Region 4, it was observed that the most common crash factor was roadway geometry. Other significant crash factors include collision with a fixed object and distracted driving. This information suggests that the roadway geometry at certain intersections could be improved to improve safety. Also, the crash factor for collisions with a fixed object suggests that cars sometimes exit the roadway in crashes and hit a fixed object. These crash factors along with distracted driving could all be related to the rural roadways that make up much of UDOT Region 4. There are long stretches of roadway with little change in the roadway or scenery. Once vehicles arrive at an intersection, it is possible that the changes in roadway geometry confuse

drivers, and more so if they are distracted while driving. The high speeds on these roadways in conjunction with this confusion can lead them to exit the roadway in a crash.

Table 8-6: Most Common Crash Factors for UDOT Region 4

Rank	Crash Factor	# of Crashes	% of Crashes
1	Roadway Geometry Related	15	27%
2	Single Vehicle	12	21%
3	Older Driver Involved	11	20%
3	Night Dark Condition	11	20%
5	Teenage Driver Involved	10	18%
6	Collision with Fixed Object	9	16%
7	Distracted Driving	8	14%
8	Overturn Rollover	7	13%
8	Improper Restraint	7	13%
10	Unrestrained	6	11%
10	Adverse Roadway Surface Condition	6	11%
10	DUI	6	11%

8.4   Summary

The purpose of this chapter was to discuss the most recent UICPM results and to identify crash factor trends in the model results. Many of the top-ranked intersections in the state are located in the northern part of the state, particularly in UDOT Regions 1 and 2. Most of the top-ranked intersections are also located in urban areas, with a few exceptions. The percentile values of the top-ranked intersections indicate valid results and the ranking of the intersections makes sense based on local knowledge. The most common crash factors of intersections in all UDOT Regions include night/dark conditions, teenage drivers, older drivers, and single vehicle crashes. Each UDOT Region has unique common crash factors other than those mentioned.

9   CONCLUSIONS

9.1   Overview

There is a need to improve roadway safety at intersections in Utah. The purpose of this research with UDOT was to identify intersection hot spots and provide tools to UDOT to help prioritize intersection safety projects. This purpose has been accomplished with the development of the UICPM. The purpose of the ISAM process is to prepare data for and execute the UICPM, and to create useful reports to assist UDOT in safety project prioritization. The success of this research project has introduced additional research possibilities that can be accomplished in the future. The purpose of this chapter is to summarize the three steps of the ISAM and to provide insights for future research opportunities.

9.2   Intersection Safety Analysis Methodology

The purpose of the ISAM is to prepare roadway and crash data for the UICPM and to generate useful reports of the model results. The ISAM includes the following three steps: (1) model input preparation, (2) model execution, and (3) reports creation. Several automated tools have been developed in MS Excel and R statistical software to make the completion of the ISAM a simple process for any user. A summary of the processes involved in the ISAM is shown in Figure 8-1. They following sections contain discussions that review the three major steps of the ISAM.
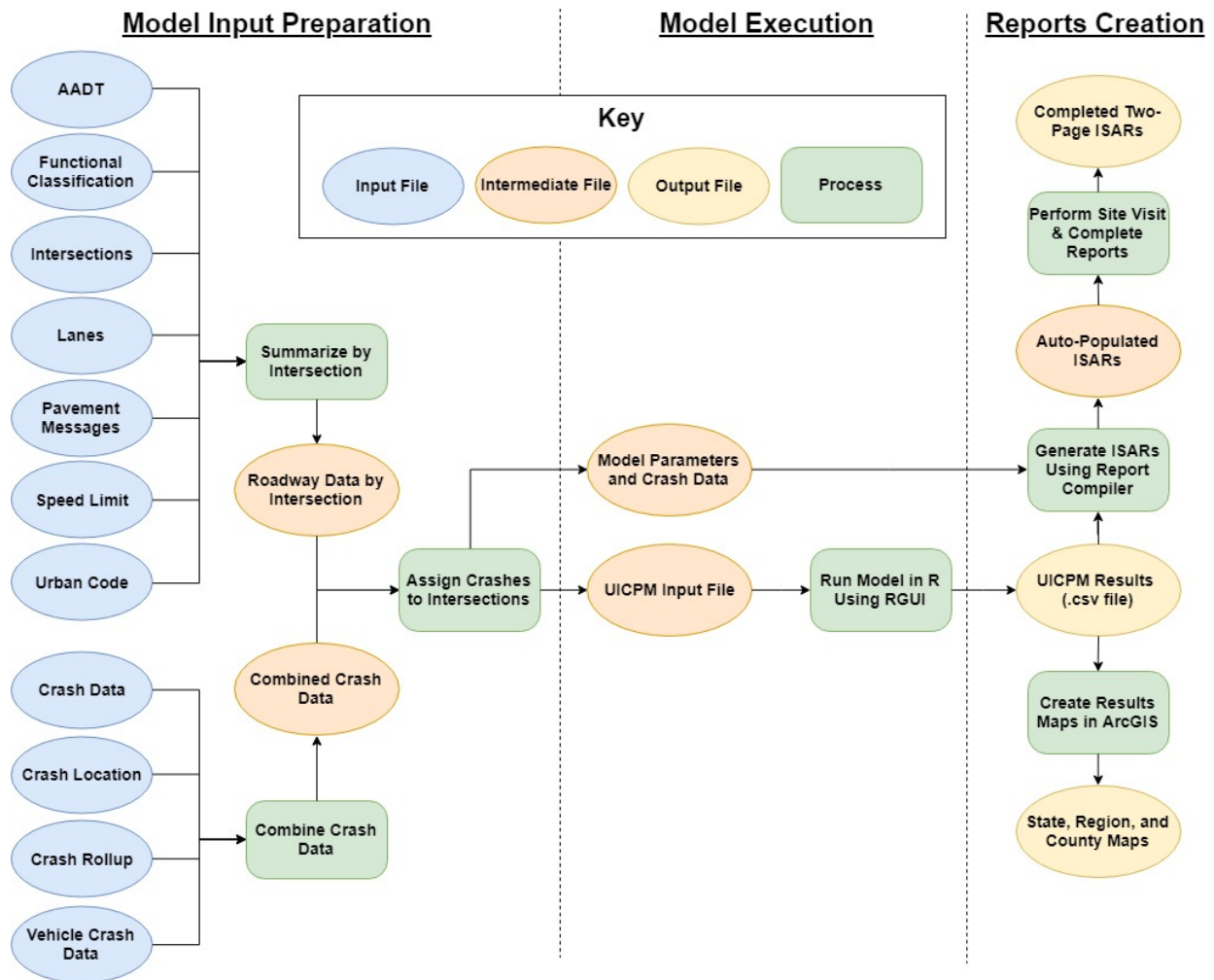
Figure 8-1: Intersection Safety Analysis Methodology summary.

9.2.1 Model Input Preparation

The first step in the ISAM is to prepare the roadway and crash data as inputs to the

UICPM. The roadway data are used in the UICPM as variables in the Bayesian regression

model. The crash data are used as the response of the UICPM. The roadway data are available

publicly on UDOT's Open Data Portal (UDOT 2017a), and the crash data are provided directly

by UDOT. The crash data are not public due to the confidential information and are protected

under 23 USC 409 (USGPO 2012). Automated tools have been developed using VBA macros to

format and combine the raw data files into a final UICPM input file. These macros are found in a tool called the R GUI, as discussed in Chapter 4.

There are six roadway data files that are used in the UICPM: AADT, functional classification, intersections, lanes, pavement messages, speed limit, and urban code. These datasets describe their respective roadway characteristic based on route number and milepoints. The intersection dataset is particularly important for the UICPM, as it provides information for the location of the major intersections in Utah. Based on the intersections dataset, all roadway data files are combined by assigning characteristics to each intersection based on location. The daily volumes on roadway segments at intersections are summed to calculate the number of entering vehicles at each intersection. One limitation with the existing data is that the intersections dataset does not indicate the route numbers for intersecting routes that are non-state routes. This limits the possible intersection dataset to only intersections of at least two state routes.

The UDOT crash data consists of four data files: crash data, crash location, crash rollup, and vehicle crash. These datasets contain various characteristics of the crashes that have occurred in Utah. The crashes are labeled with a crash ID number, which is used to combine all crash datasets into a single crash dataset. The final UICPM input file is created by assigning the crashes to each intersection based on location. If a crash falls within the functional area of an intersection, it is assigned to that intersection. The functional area of an intersection is the area upstream or downstream of an intersection that is impacted by the operation of the intersection itself and can be based on speed limit, functional classification, or urban code. Once the crashes are assigned to intersections, the data are prepared to be input to the UICPM.

9.2.2   Statistical Intersection Safety Model

The second step in the ISAM is to run the UICPM using the data prepared in the previous step. The UICPM is a Bayesian generalized linear model and was developed in a similar process as the UCPM and UCSM. A forward variable selection process was completed to determine the roadway characteristic variables that provided the model with the best fit. The variable selection resulted in five variables to be used for the regression model: entering vehicles, number of intersection legs, minimum number of lanes, maximum roadway width, and maximum speed limit. These variables are multiplied by coefficients for each intersection in the model to estimate a crash distribution for each intersection. It was expected that the effect of increased covariates would be different for certain categories of intersections due to the large differences between certain intersection types. It was determined that adding a categorical variable to the model to group the variables into different groups based on UDOT Region and functional classification improved the model.

The execution of the UICPM was automated using VBA macros in the R GUI tool. The user enters the various parameters of the model into the R GUI, and the R GUI sends a command to run the model in R. The model produces crash distributions for each study intersection, and the actual number of crashes at the intersection is compared to the crash distribution. The median value of the crash distribution is determined to be the expected number of crashes at the intersection. A percentile value is calculated, which represents the probability that the number of crashes occurring at an intersection in a particular year is less than or equal to the average annual number of crashes. A high percentile indicates an extreme number of crashes occurred at the intersection and that the intersection is a hot spot. The intersections are ranked based on the percentile value at the state, UDOT Region, and county levels.

The results of the model are saved in a local folder in MS Excel format. The results file contains all data included in the UICPM input file in addition to the percentile, rankings, and expected number of crashes for each intersection. The UICPM is used to produce plots and graphs showing the crash distributions of the top 20 intersections in the state. Once the UICPM execution is complete, reports are created to discuss and present the results of the model.

9.2.3 Reports Creation

The third step in the ISAM is to create reports of the model results, including ISARs and GIS maps. The ISARs are two-page reports for each intersection that contain intersection information, crash data, safety analyses, and possible countermeasures to fix safety issues. Much of the ISAR is auto-generated using the Report Compiler tool, an MS Excel-based tool operated using VBA macros. The rest of the report is manually written, where the user identifies crash trends, describes historical and current conditions, and proposes possible countermeasures. The ISARs of intersections of interest (i.e., top 10 intersections in each UDOT Region) are provided to UDOT Region offices for consideration when planning safety projects in the region.

Another report that is created is a GIS map of the model results. This is created using ArcMap software and various tools that have been developed in the software. The model results are loaded into the software and plotted on the map. The results are color-coded based on state ranking, with high-ranking intersections in red and low-ranking intersections in green, as shown previously in Figure 6-1. A tool in ArcMap then produces maps of these intersections at the state, UDOT Region, and county level, respectively. These maps are useful in understanding the distribution of intersection hot spots throughout the state of Utah.

9.3   Future Research Topics

The success of this research has resulted in new possibilities for future research in traffic safety in Utah. The following sections contain short descriptions of potential research topics that could be analyzed in the future including segment models without intersections, non-state route intersection analysis, using GIS for UICPM processes, and creating a web application for the statistical models.

9.3.1   Segment Models Without Intersections

The UCPM and UCSM were developed to identify roadway segment hotspots in the state of Utah. These were successful in their methodologies; however, the model included segments that were part of intersections. Seeing that this research project has analyzed only the intersections on Utah roadways, it would be interesting to remove intersection roadway segments from the database to run in the UCPM or UCSM.

9.3.2   Non-State Route Intersection Analysis

As discussed in Section 3.2.4, there are limitations in the roadway datasets for non-state routes. The intersections dataset does not indicate the locations where non-state routes intersect other routes. Also, some datasets only have data for state routes. Based on discussions with UDOT, it is anticipated that data for federal aid routes will soon be available. While the ISAM has been designed to work for all intersections, including non-state route intersections, a future research topic could include ensuring that the ISAM works with non-state route intersections once the data becomes available.

9.3.3   Using GIS for UICPM Processes

The ISAM processes have been automated using MS Excel workbook tools enabled with VBA macros. This was done based on the experience of the research team and the ease of writing VBA macros. However, because the roadway and crash data are spatial, the ISAM processes could be designed in GIS software as well. The implementation of the ISAM in GIS may result in a more user-friendly and clean process. If this is accomplished, it is anticipated that the actual model execution would still need to be run from the R GUI or another non-GIS tool.

9.3.4   Create Web Application for Statistical Models

The existing R GUI is a valuable tool for running the statistical models from MS Excel in R. The research team has discovered a tool called Shiny, developed by R Studio, that can be used to create web applications that run R models (R Studio 2018). The Shiny web applications can be designed to be very user-friendly and can be accessed from any location via a web browser. The implementation of the segment and intersection models with a Shiny web application may provide a better user experience and should be explored further.

9.4   Concluding Remarks

Roadway safety is of high importance for UDOT in the state of Utah. Intersection safety is a significant part of that with approximately 40 percent of crashes on Utah roadways being labeled as "intersection related" (Numetric 2018). The use of the UICPM will assist UDOT in their efforts to achieve "Zero Fatalities" on Utah roadways by identifying intersection hot spots throughout the state. The ISAM has been developed to provide a simple process with automated tools and user interfaces to execute the UICPM. The implementation of this methodology will result in better safety project planning in Utah at the state and UDOT Region levels.

REFERENCES

American Association of State Highway and Transportation Officials (AASHTO). (2010). *Highway Safety Manual* (1st ed.), Volume 3. Washington, D.C.: AASHTO.

Choi, E. (2010). "Crash Factors in Intersection-Related Crashes: An On-Scene Perspective." Report DOT-HS-811-366. U.S. Department of Transportation, Washington, DC.

Colorado Department of Transportation (CDOT). (2017). "Intersection SPF's." <https://www.codot.gov/library/traffic/safety-crash-data/safety-analysis-information/intersection-spfs>. (Jan. 20, 2017).

Gibbons, J., Mineer, S., and Brown, N. (2016). "Roadway Safety Analysis Methodology User Manual: Volume 1." Department of Civil and Environmental Engineering, Brigham Young University, Provo, UT.

Gluck, J., Levinson, H. S., and Stover, V. (1999). "NCHRP Report 420: Impacts of Access Management Techniques." National Cooperative Highway Research Program. Transportation Research Board, Washington, DC.

Goodwin, A., Thomas, L., Kirley, B., Hall, W., O'Brien, N., and Hill, K. (2015). "Countermeasures That Work: A Highway Safety Countermeasure Guide for State Highway Safety Offices." Report DOT-HS-812-202. NHTSA, U.S. Department of Transportation, Washington, DC.

Google. (2018). "Google Earth." <https://www.google.com/earth/>. (Mar. 7, 2018).

Hyndman, R. J. and Koehler, A. B. (2006). "Another Look at Measures of Forecast Accuracy." *International Journal of Forecasting*. <https://www.sciencedirect.com/science/article/pii/S0169207006000239>. (Mar. 7, 2018).

Mandli. (2017). "Mandli Communications." <http://mandli.com/>. (Jan. 21, 2017).

Mandli. (2018). "Roadview Explorer 5." < http://168.178.125.102/UtahRVX3/>. (Mar. 7, 2018).

National Highway Traffic Safety Association (NHTSA). (2012). "MMUCC Guideline: Model Minimum Uniform Crash Criteria, 4th Edition." <https://www.mmucc.us>. (May 5, 2018).

Neuman, T. R., Pfefer, R., Slack, K. L., Raub, R., and Wark, R. (2003). "NCHRP Report 500: Volume 1: A Guide for Addressing Aggressive-Driving Collisions." National Cooperative Highway Research Program. Transportation Research Board, Washington, DC.

Numetric. (2018). "UDOT SafeMap." <https://udot.numetric.com>. (Mar. 6, 2018).

R Project for Statistical Computing (RPSC). (2016). "What is R," <http://www.r-project.org> (Jan. 21, 2017).

Rodegerdts, L. A., Nevers, B., and Robinson, B. (2004). "Signalized Intersections: Information Guide." Report FHWA-HRT-04-091. Federal Highway Administration, U.S. Department of Transportation, Washington, DC.

R Studio. (2018). "Shiny." < https://shiny.rstudio.com>. (Apr. 10, 2018).

Saito, M., Browning, J., and Schultz, G. G. (2018). "Developing a Method to Identify Horizontal Curve Segments with High Crash Occurrences Using the HAF Algorithm." Report UT-18.02, Utah Department of Transportation, Research Division, Salt Lake City, UT.

Schultz, G. G., Allen, C. G., and Eggett, D. L. (2008). "Crashes in the Vicinity of Major Crossroads." Report UT-08.25, Utah Department of Transportation Traffic and Safety, Research Division, Salt Lake City, UT.

Schultz, G. G., Bassett, D., Roundy, R., Saito, M., Reese, C.S. (2015). "Use of Roadway Attributes in Hot Spot Identification and Analysis." Report UT-15.10, Utah Department of Transportation Traffic and Safety, Research Division, Salt Lake City, UT.

Schultz, G. G., Dowell, A. L., Saito, M., and Roundy, R. S. (2013a). "Evaluating the Safety Effects of Signal Improvements." Report UT-13.05, Utah Department of Transportation, Research Division, Salt Lake City, UT.

Schultz, G. G., Farnsworth, J., Bassett, B., Cook, A., Roundy, R., Saito, M. (2013b). "The Utah Crash Prediction Model User's Guide." Department of Civil and Environmental Engineering, Brigham Young University, Provo, UT.

Schultz, G. G., Farnsworth, J. S., Roundy, R., Saito, M., Reese, C.S., and Briggs, T. (2013c). "Hot Spot Identification & Analysis Methodology." Report UT-13.15, Utah Department of Transportation Traffic and Safety, Research Division, Salt Lake City, UT.

Schultz, G. G., Mineer, S. T., Saito, M., Gibbons, J. D., Siegel, S. A., and MacArthur, P. D. (2016). "Roadway Safety Analysis Methodology for Utah." Report UT-16.13, Utah Department of Transportation Traffic and Safety, Research Division, Salt Lake City, UT.

Siegel, S., MacArthur, P., Mineer, S. (2016). "Roadway Safety Analysis Methodology User Manual: Volume 2." Department of Civil and Environmental Engineering, Brigham Young University, Provo, UT.

Spiegelhalter, D. J., Best, N. G., Carlin, B. P., and van der Linde, A. (2002). "Bayesian Measures of Complexity and Fit." *Journal of the Royal Statisical Society*. <https://rss.onlinelibrary.wiley.com/doi/pdf/10.1111/1467-9868.00353>.

Stover, V. G. and Koepke, F. J. (2002). *Transportation and Land Development* (2nd ed.). Institute of Transportation Engineers. Washington, DC.

Thomas, C. G. (2008). "Georgia Intersection Safety Improvement Program." Georgia Department of Transportation, Office of Research, Atlanta, GA.

United States Government Publishing Office (USGPO). (2012). "23 U.S.C. 409 – Discovery and Admission as Evidence of Certain Reports and Surveys." United States Code. <https://www.gpo.gov/fdsys/granule/USCODE-2011-title23/USCODE-2011-title23-chap4-sec409> (Jan. 21, 2017).

University of Utah (UofU). (2016). "Improving Public Safety with Data-Driven Decisions." Civil and Environmental Engineering Newsletter, 2016 Spring edition. <https://www.civil.utah.edu/files/2016/05/2016-Spring-newsletter2016.pdf> (Jan. 23, 2017).

Utah Department of Transportation (UDOT). (2017a). "UDOT Open Data." <udot.uplan.opendata.arcgis.com> (Dec. 1, 2017).

Utah Department of Transportation (UDOT). (2017b). "Zero Fatalities." <http://ut.zerofatalities.com/> (Jan. 21, 2017).

Williams, K. M., Stover, V. G., Dixon, K. K., and Demosthenes, P. (2014). *Access Management Manual* (2nd ed.). Transportation Research Board, Washington, DC.

Wolshon, B., Tindale, S. A., Noyce, D. A., Hagen, L. T., Seyfried, R. K., and Jacquemart, G. G. (2004). "Toolbox on Intersection Safety and Design." Institute of Transportation Engineers, Washington, DC. < http://library.ite.org/pub/e1d08c51-2354-d714-51e9-f3967064dfb9> (Jan. 24, 2017).

LIST OF ACRONYMS

AADT   Annual Average Daily Traffic

AASHTO   American Association of State Highway and Transportation Officials

BYU   Brigham Young University

CAPS   Center for Advanced Public Safety

CPY   Crashes Per Year

CSV   Comma-separated Value

CTW   Countermeasures That Work

DIC   Deviance Information Criterion

DOT   Department of Transportation

DUI   Driving Under the Influence

FHWA   Federal Highway Administration

GIS   Geographic Information System

GPS   Global Positioning System

GUI   Graphical User Interface

HSM   Highway Safety Manual

HSRG   Highway Safety Research Group

ID   Identification

ISAM   Intersection Safety Analysis Methodology

ISAR   Intersection Safety Analysis Report

ISIP   Intersection Safety Improvement Program

LED   Light Emitting Diode

LRS   Linear Referencing System

MS   Microsoft

NCHRP   National Cooperative Highway Research Program

PDF   Portable Document Format

PDO   Property Damage Only

PRT   Perception Reaction Time

R GUI   R Graphical User Interface

RPMSE   Root Predicted Mean Squared Error

RSAM   Roadway Safety Analysis Methodology

RSAR   Roadway Safety Analysis Report

SHSP   Strategic Highway Safety Plan

SPF   Safety Performance Function

SQL   Structured Query Language

TRB   Transportation Research Board

UCPM   Utah Crash Prediction Model

UCSM   Utah Crash Severity Model

UDOT   Utah Department of Transportation

UICPM   Utah Intersection Crash Prediction Model

UTAPS-CDI   Utah Transportation and Public Safety – Crash Data Initiative

VBA   Visual Basic for Applications

VMT   Vehicle-Miles Traveled

APPENDIX A: CRITICAL DATA COLUMNS

Table A-1: AADT Critical Data Columns

Expected Header	Description
ROUTE	Route ID: numeric route number for a given road segment
BEGMP	Beginning Milepoint: beginning milepoint of the road segment
ENDMP	End Milepoint: end milepoint of the road segment
STATION	Station Number: seven-digit number, identifying the traffic counter station number
AADT[YEAR]	AADT [YEAR]: historical dataset of AADT data from each year; at least 7 years of this data are needed (i.e., AADT2012)
SUTRK	Single Truck Percent: percent of single trailer trucks per segment
CUTRK	Combo Truck Percent: percent of combination trailer trucks per segment
SUTRKCOUNT	Single-Unit Truck Count: number of single trailer trucks per segment
CUTRKCOUNT	Combo-Unit Truck Count: number of combination trailer trucks per segment

Table A-2: Functional Classification Critical Data Columns

Expected Header	Description
ROUTE_NAME	Route ID: numeric route number for a given road segment
BEGIN_MP	Beginning Milepoint: beginning milepoint of the road segment
END_MP	End Milepoint: end milepoint of the road segment
FC_CODE	FC_CODE: number representing the functional classification type of the road

Table A-3: Speed Limit Critical Data Columns

Expected Header	Description
Route	Route ID: Route ID number with direction letter (i.e., 0089N)
Direction	Direction: Route direction (P, N)
Beg_MP	Beginning Milepoint: The milepoint where the sign appears
End_MP	End Milepoint: The end milepoint of the road segment
Speed_Limit	Speed Limit: number signifying the speed limit (in mph) of a particular segment.

Table A-4: Lanes Critical Data Columns

Expected Header	Description
ROUTE	Route ID: numeric route number for a given road segment
START_ACCUM	Beginning Milepoint: beginning milepoint of the road segment
END_ACCUM	End Milepoint: end milepoint of the road segment
THRU_CNT	Thru Lanes: number of thru lanes
THRU_WDTH	Thru Lane Width: width of the thru lanes in feet

Table A-5: Urban Code Critical Data Columns

Expected Header	Description
ROUTE_NAME	Route ID: numeric route number for a given road segment
START_ACCUM	Beginning Milepoint: beginning milepoint of the road segment
END_ACCUM	End Milepoint: end milepoint of the road segment
URBAN_CODE	Urban Code: number that represents a description of the surrounding area
URBAN_DESC	Urban Description: description of the surrounding area (i.e., Small-Urban, St. George, rural, etc.)

Table A-6: Pavement Messages Critical Data Columns

Expected Header	Description
ROUTE_NAME	Route ID: numeric route number for a given road segment
START_ACCUM	Milepoint: the location of the pavement message determined by milepoint
MESSAGE_TYPE	Message Type: the type of pavement message (i.e., stop bar)
OMS_SIDE	OMS Side: the side of the street where the pavement marking is (i.e., positive or negative)

Table A-7: Intersections Critical Data Columns

Expected Header	Description
BEGIN_LATITUDE	Intersection Latitude: the latitude coordinate of the intersection location
BEGIN_LONGITUDE	Intersection Longitude: the longitude coordinate of the intersection location
BEGIN_ALTITUDE	Intersection Elevation: the elevation of the intersection location
SIGNALIZED	Signalized Intersection: Y/N to determine whether an intersection is signalized or not
CONTROL	Intersection Control: an indicator of the intersection control (i.e., stop); signalized intersections show no value in this data column
SR_SR_INTE	State Route to State Route Intersection: Y/N to determine whether an intersection is an intersection between two or more state routes or not
ROUTE_NAME	Route Name: the route ID number of the first state route intersecting at the location
ROUTE_1_IN	Route ID of 1st Intersecting Route: the route ID number of the second state route intersecting at the location
ROUTE_2_IN	Route ID of 2nd Intersecting Route: the route ID number of the third state route intersecting at the location
START_ACCUM	Intersection Milepoint: the milepoint location of the intersection referenced on the first route listed in the "ROUTE_NAME" column
STATION	Station/City Information: a combined numerical and text indicator of the station location of the intersection; tends to represent the city or metropolitan area where the intersection is located

Table A-8: Crash Location Critical Data Columns

Expected Header	Description
CRASH_ID	Crash ID: unique crash ID number for each crash
ROUTE	Route ID: numeric route number for a given road segment
ROUTE_DIRECTION	Direction: route direction (i.e., P, N, or X)
RAMP_ID	Ramp ID: ID indicating a ramp and the type (i.e., 1-4, CD)
MILEPOINT	Milepoint: milepoint location of the crash
UTM_Y	UTM Longitude: the longitude coordinate of the crash, referenced in the Universal Transverse Mercator coordinate system
UTM_X	UTM Latitude: the latitude coordinate of the crash, referenced in the Universal Transverse Mercator coordinate system

Table A-9: Crash Data Critical Data Columns

Expected Header	Description
CRASH_ID	Crash ID: unique crash ID number for each crash
CRASH_DATETIME	Crash Date/Time: date and time of crash
CRASH_SEVERITY_ID	Crash Severity ID: numerical severity level of crash (i.e., 1-5)
LIGHT_CONDITION_ID	Light Condition: ID for light condition at time of crash (i.e., 1-6, 88-99)
WEATHER_CONDITION_ID	Weather Condition: ID for weather condition at time of crash (i.e., 1-9, 88-99)
MANNER_COLLISION_ID	Manner Collision: ID for manner of collision in crash (i.e., 1-8, 88-99)
PAVEMENT_ID	Pavement: ID for pavement type (i.e., 1-4, 88-99)
ROADWAY_SURF_CONDITION_ID	Roadway Surface Condition: ID for roadway surface conditions (i.e., 1-9, 88-99)
ROADWAY_JUNCT_FEATURE_ID	Roadway Junction Feature: ID for roadway junction feature (i.e.,1-10, 20-26, 88-99)
WORK_ZONE_RELATED_YNU	Work Zone Related: Y/N to determine whether crash occurred in work zone
WORK_ZONE_WORKER_PRESENT_YNU	Work Zone Worker Present: Y/N to determine whether worker present in work zone
HORIZONTAL_ALIGNMENT_ID	Horizontal Alignment: ID for horizontal curvature of roadway (i.e., 1-2, 88-99)
VERTICAL_ALIGNMENT_ID	Vertical Alignment: ID for vertical curvature of roadway (i.e., 1-4. 88-99)
ROADWAY_CONTRIB_CIRCUM_ID	Roadway Contributing Circumstance: ID for vehicle contributing circumstance related to the crash (i.e., 0-18, 88-99)
FIRST_HARMFUL_EVENT_ID	First Harmful Event: ID for first harmful event resulting from the crash (i.e., 0-62, 88-99)

Table A-10: Crash Rollup Critical Data Columns

Expected Header	Description
CRASH_ID	Crash ID: unique crash ID number for each crash
NUMBER_VEHICLES_INVOLVED	Number Vehicles Involved: number of vehicles involved in the given accident
NUMBER_FATALITIES	Number of Fatalities: number of person-fatalities resulting from a given crash
NUMBER_FOUR_INJURIES	Number of incapacitating injuries: number of person-incapacitating injuries resulting from a given crash
NUMBER_THREE_INJURIES	Number of injuries: number of person-injuries resulting from a given crash
NUMBER_TWO_INJURIES	Number of possible injuries: number of person-possible injuries resulting from a given crash
NUMBER_ONE_INJURIES	Number of property damage only events: number of events for property damage only resulting from a given crash
PEDESTRIAN_INVOLVED	Pedestrian Involved: Y/N to determine whether a pedestrian was involved in the crash
BICYCLIST_INVOLVED	Bicyclist Involved: Y/N to determine whether a bicyclist was involved in the crash
MOTORCYCLE_INVOLVED	Motorcycle Involved: Y/N to determine whether a motorcycle was involved in the crash
IMPROPER_RESTRAINT	Improper Restraint: Y/N to determine whether improper restraint was a factor in the crash
UNRESTRAINED	Unrestrained: Y/N to determine whether a driver/passenger was unrestrained in the crash
DUI	DUI: Y/N to determine whether driving under the influence was a factor in the crash

Table A-10: Crash Rollup Critical Data Columns (continued)

Expected Header	Description
AGGRESSIVE_DRIVING	Aggressive Driving: Y/N to determine whether aggressive driving was a factor in the crash
DISTRACTED_DRIVING	Distracted Driving: Y/N to determine whether distracted driving was a factor in the crash
DROWSY_DRIVING	Drowsy Driving: Y/N to determine whether drowsy driving was a factor in the crash
SPEED_RELATED	Speed Related: Y/N to determine whether speed was a factor in the crash
INTERSECTION_RELATED	Intersection Related: Y/N to determine whether the crash occurred at an intersection
ADVERSE_WEATHER	Adverse Weather: Y/N to determine whether adverse weather was a factor in the crash
ADVERSE_ROADWAY_SURF_CONDITION	Adverse Roadway Surface Conditions: Y/N to determine whether adverse roadway surface conditions were a factor in the crash
ROADWAY_GEOMETRY_RELATED	Roadway Geometry Related: Y/N to determine whether roadway geometry was a factor in the crash
WILD_ANIMAL_RELATED	Wild Animal Related: Y/N to determine whether a wild animal was involved in the crash
DOMESTIC_ANIMAL_RELATED	Domestic Animal Related: Y/N to determine whether a domestic animal was involved in the crash
ROADWAY_DEPARTURE	Roadway Departure: Y/N to determine whether a vehicle departed the roadway as a result of the crash
OVERTURN_ROLLOVER	Overturn/Rollover: Y/N to determine whether a vehicle overturned and/or rolled over as a result of a crash
COMMERCIAL_MOTOR_VEH_INVOLVED	Commercial Motor Vehicle Involved: Y/N to determine whether a commercial motor vehicle was involved in the crash

Table A-10: Crash Rollup Critical Data Columns (continued)

Expected Header	Description
TEENAGE_DRIVER_INVOLVED	Teenage Drive Involved: Y/N to determine whether a teenage driver was involved in the crash
OLDER_DRIVER_INVOLVED	Older Driver Involved: Y/N to determine whether an older driver was involved in the crash
URBAN_COUNTY	Urban County: Y/N to determine whether the crash occurred in an urban area
ROUTE_TYPE	Route Type (L/S/U):
NIGHT_DARK_CONDITION	Night/Dark Condition: Y/N to determine whether night or dark conditions was a factor in the crash
SINGLE_VEHICLE	Single Vehicle: Y/N to determine whether a single vehicle was involved in a crash (i.e. not a collision involving multiple vehicles)
TRAIN_INVOLVED	Train Involved: Y/N to determine whether a train was involved in the crash
RAILROAD_CROSSING	Railroad Crossing: Y/N to determine whether the crash occurred at a railroad crossing
TRANSIT_VEHICLE_INVOLVED	Transit Vehicle Involved: Y/N to determine whether a transit vehicle was involved in the crash
COLLISION_WITH_FIXED_OBJECT	Collision with Fixed Object: Y/N to determine whether the crash involved a fixed object (i.e. not another vehicle, nor a person)

Table A-11: Crash Vehicle Critical Data Columns

Expected Header	Description
CRASH_ID	Crash ID: Specific crash ID number for each crash
INTERSTATE_HIGHWAY	Interstate Highway: Y/N to determine whether the crash occurred on an interstate roadway
VEHICLE_NUM	Vehicle Number: Number assigned to each vehicle involved in a given crash
CRASH_DATETIME	Crash Date/Time: Date and time of crash
TRAVEL_DIRECTION_ID	Travel Direction: Direction value of route at the location of the crash (i.e., 1-5)
EVENT_SEQUENCE_1_ID	Event Sequence #1: ID for first crash sequence for non-collision and collision events (i.e., 0-99)
EVENT_SEQUENCE_2_ID	Event Sequence #2: ID for second crash sequence for non-collision and collision events (i.e., 0-99)
EVENT_SEQUENCE_3_ID	Event Sequence #3: ID for third crash sequence for non-collision and collision events (i.e., 0-99)
EVENT_SEQUENCE_4_ID	Event Sequence #4: ID for fourth crash sequence for non-collision and collision events (i.e., 0-99)
MOST_HARMFUL_EVENT_ID	Most Harmful Event: ID for most harmful event resulting from the crash (i.e., 0-99)
VEHICLE_MANEUVER_ID	Vehicle Maneuver: ID for the controlled maneuver prior to the crash (i.e., 1-14, 88-99)
VEHICLE_DETAIL_ID	Vehicle Detail ID: 8-digit ID number that is specific to a vehicle involved in a crash amongst all other vehicle involved in crashes