RESEARCH

Keeping Utah Moving

Utah Department of Transportation - Research Division
4501 South 2700 West - P.O. Box 148410 - SLC,UT 84114-8410

Report No. UT-20.04

SEGMENT AND INTERSECTION CRASH ANALYSIS METHODOLOGIES FOR UTAH HIGHWAYS

Prepared For:

Utah Department of Transportation
Traffic & Safety, Research & Innovation Divisions

Submitted By:

Brigham Young University
Dept. of Civil and Environmental Engineering

Authored By:

Grant G. Schultz, PhD, PE, PTOE
Professor

Camille C. Lunt, EIT
Graduate Research Assistant

Timo Pew
Graduate Research Assistant

Richard L. Warr, PhD
Assistant Professor

Matthew J. Heaton, PhD
Assistant Professor

Final Report
August 2020

DISCLAIMER

The authors alone are responsible for the preparation and accuracy of the information, data, analysis, discussions, recommendations, and conclusions presented herein. The contents do not necessarily reflect the views, opinions, endorsements, or policies of the Utah Department of Transportation or the U.S. Department of Transportation. The Utah Department of Transportation makes no representation or warranty of any kind, and assumes no liability therefore.

This report may be protected under 23 USC 409.

ACKNOWLEDGMENTS

The authors acknowledge the Utah Department of Transportation (UDOT) for funding this research, and the following individuals from UDOT on the Technical Advisory Committee for helping to guide the research:

- Jeff Lewis – UDOT Traffic & Safety
- Robert Miles – UDOT Traffic & Safety
- Ivana Vladisavljevic – UDOT Traffic & Safety
- Tim Taylor – WCEC Engineers/UDOT Traffic & Safety
- Travis Jensen – WCEC Engineers/UDOT Traffic & Safety
- Dallas Wall – WCEC Engineers/UDOT Traffic & Safety
- Clancy Black – WCEC Engineers/UDOT Traffic & Safety
- Charles Allen – Parametrix/UDOT Traffic & Safety
- Nathan Taylor – BYU Civil and Environmental Engineering
- Lyvia Zinn – BYU Civil and Environmental Engineering
- Tomas Barriga Aristizabal – BYU Civil and Environmental Engineering
- Rachel Huber – BYU Civil and Environmental Engineering
- Christian Davis – BYU Statistics

TECHNICAL REPORT ABSTRACT

1. Report No. UT- 20.04	2. Government Accession No. N/A	3. Recipient's Catalog No. N/A
4. Title and Subtitle SEGMENT AND INTERSECTION CRASH ANALYSIS METHODOLOGIES FOR UTAH HIGHWAYS		5. Report Date August 2020 6. Performing Organization Code R0402320
7. Author(s) Grant G. Schultz, Ph.D., P.E., PTOE; Camille C. Lunt, EIT; Timo Pew; Richard L. Warr, Ph.D.; Matthew J. Heaton, Ph.D.		8. Performing Organization Report No.
9. Performing Organization Name and Address Brigham Young University Department of Civil and Environmental Engineering 430 Engineering Building Provo, UT 84602		10. Work Unit No. 11. Contract or Grant No. 19-8011
12. Sponsoring Agency Name and Address Utah Department of Transportation 4501 South 2700 West P.O. Box 148410 Salt Lake City, UT 84114-8410		13. Type of Report & Period Covered Final Report July 2018 to August 2020 14. Sponsoring Agency Code PIN No. 14476

15. Supplementary Notes

Prepared in cooperation with the Utah Department of Transportation and the U.S. Department of Transportation, Federal Highway Administration

16. Abstract

This research focuses on the Crash Analysis Methodology for Segments (CAMS) which provides a way for engineers at the Utah Department of Transportation (UDOT) to prioritize safety improvements on state-owned roadways. Unlike the Utah crash analysis methodologies that come before it, the CAMS focuses exclusively on segment-related crashes. The benefits of such an analysis can be found in identifying locations that have safety concerns unbiased from intersections and their related crashes.

The CAMS uses UDOT data to create a spreadsheet of roadway segments and their associated crashes. Each segment is homogeneous with respect to five variables: Annual Average Daily Traffic (AADT), functional class, number of lanes, speed limit, and urban code. In the statistical analyses performed on the data, four years of crash data (2014-2017) are used to predict distributions of crashes for the most recent year of data (2018). Observed crash counts are compared to the predicted distributions and assigned a percentile value within the distributions, and segments are subsequently ranked in order of safety concern according to those percentiles. Two-page technical reports are created for segments that rank high in the state or UDOT Region. These reports consist of concise tables of roadway data and crash trends pertaining to each segment. Research analysts also add observations made in virtual site visits to the reports. In the end, the results and the reports are sent to UDOT where UDOT Region engineers may review and study identified segments in further detail.

This research also includes modifications made to the Intersection Safety Analysis Methodology (ISAM) which focuses exclusively on intersection-related crashes. The modifications made to the ISAM mirror the abilities of the CAMS, thus allowing the pair of methodologies to analyze the entire state route network without overlapping any crash data.

17. Key Words Segment safety analysis, hot spot identification, crash analysis, highway safety research		18. Distribution Statement Not restricted. Available through: UDOT Research Division 4501 South 2700 West P.O. Box 148410 Salt Lake City, UT 84114-8410 www.udot.utah.gov/go/research		23. Registrant's Seal N/A
19. Security Classification (of this report) Unclassified	20. Security Classification (of this page) Unclassified	21. No. of Pages 104	22. Price N/A	

TABLE OF CONTENTS

LIST OF TABLES

<u>**LIST OF FIGURES**</u>

LIST OF ACRONYMS

AADT	Annual Average Daily Traffic
AASHTO	American Association of State Highway and Transportation Officials
BYU	Brigham Young University
CAMS	Crash Analysis Methodology for Segments
CAMS-P	Crash Analysis Methodology for Segments Prediction Model
CAMS-S	Crash Analysis Methodology for Segments Severity Model
CSV	Comma Separated Value
DOT	Department of Transportation
DUI	Driving Under the Influence
FHWA	Federal Highway Administration
GIS	Geographic Information System
GoF	Goodness of Fit
HSM	Highway Safety Manual
ID	Identification Number
ISAM	Intersection Safety Analysis Methodology
ISAR	Intersection Safety Analysis Report
MAD	Median Absolute Deviation
NBL	Negative Binomial-Lindley
NCHRP	National Cooperative Highway Research Program
RGUI	R Graphical User Interface
RPMSE	Root Predicted Mean Square Error
RSAM	Roadway Safety Analysis Methodology
RSAR	Roadway Safety Analysis Report
SPF	Safety Performance Function
SSAR	Segment Safety Analysis Report
UCPM	Utah Crash Prediction Model
UCSM	Utah Crash Severity Model
UDOT	Utah Department of Transportation
UICPM	Utah Intersection Crash Prediction Model
UICSM	Utah Intersection Crash Severity Model

UTM	Universal Transverse Mercator
VBA	Visual Basic for Applications
VMT	Vehicle-Miles Traveled
ZINB	Zero-Inflated Negative Binomial
ZIP	Zero-Inflated Poisson

EXECUTIVE SUMMARY

Every year, efforts are made by the Utah Department of Transportation (UDOT) to reduce the number of fatalities, injuries, and crashes on Utah roads. To help engineers select sites most in need of attention and improvements, UDOT has worked with Brigham Young University (BYU) in a series of safety-focused traffic research. These research efforts aim at finding locations with unusually high numbers of crashes.

Two methodologies are presented in this report. The first is named the Crash Analysis Methodology for Segments (CAMS). Previous to the CAMS were the Road Safety Analysis Methodology (RSAM) which looks at road segments as a whole including intersection crashes along the segments, and the Intersection Safety Analysis Methodology (ISAM) which looks at intersections independently. The purpose of the CAMS is to provide a methodology that looks at road segments without influence from intersection crashes. In other words, this research focuses on segment-related crashes, their hot spot locations, and possible ways to mitigate the safety concerns they cause.

The CAMS begins with a data integration process that combines UDOT roadway characteristics and crash data to create segments of Utah highway. Segments are homogeneous with respect to five variables: Annual Average Daily Traffic (AADT), functional class, number of lanes, speed limit, and urban code. Two statistical models are used to identify hot spots among these segments. One is named the CAMS Prediction (CAMS-P) model and is a Zero-Inflated Poisson (ZIP) model (also referred to as a Poisson Mixture Model) that is used to identify segments with an unusually high number of crashes. The other model, named the CAMS Severity (CAMS-S) model, is used to identify segments with unusually high proportions of injury crashes. Both models compare actual crash counts to predicted distributions of crashes. Two-page technical reports that display information about roadway characteristics and crash history are prepared for identified segments. These reports, named Segment Safety Analysis Reports (SSARs), also provide suggested countermeasures to mitigate the safety concerns present at each segment.

The second methodology is an update of ISAM. In its original publication, the ISAM analyzed intersections that included at least two unique state routes and applied one statistical model to the data. In the updated version, the methodology can analyze state route intersections with at least one of the following characteristics: another state route, a federal aid route, or a traffic signal. Together, the CAMS and ISAM allow for an analysis of the entire Utah state route network.

1.0 INTRODUCTION

1.1 Problem Statement

From 2014 to 2018, an average of 270 people died on Utah roadways annually (UDPS 2020). The Utah Department of Transportation (UDOT) has teamed up with Zero Fatalities: A Goal We Can All Live With® to focus on reducing the number of lives lost on Utah roadways to zero (Zero Fatalities 2020).

Because transportation improvements have a limited budget, it is important for state departments of transportation (DOTs) to put their dollars into projects expected to make a large positive impact. Prioritizing safety improvements can be achieved by evaluating locations that stand out in terms of annual crash frequency compared to similar locations across the state.

In coordination with UDOT, the Brigham Young University (BYU) Civil and Environmental Engineering and Statistics Departments have developed a series of safety-focused research, including methodologies called the Roadway Safety Analysis Methodology (RSAM) and the Intersection Safety Analysis Methodology (ISAM). These methodologies are meant to identify locations around the state of Utah that show a high potential for safety improvement.

The RSAM analyzes segments of roadway in the Utah state route network. This analysis considers all types of crashes that occur along state routes, including a mixture of both intersection-related and segment-related crashes (Schultz et al. 2016). The ISAM, however, analyzes crashes related exclusively to intersections on Utah state routes (Schultz et al. 2018). With the focused nature of the ISAM, more specific countermeasures can be selected to help improve safety at the identified locations. The same focus could be directed toward segment-related safety concerns, yet there was no methodology that did this for Utah roadways previous to the research presented in this report.

1.2 Objectives

The primary objective of this research is to develop a methodology that identifies portions of Utah state routes that indicate a high potential for safety improvement for segment-related crashes. This methodology is named the Crash Analysis Methodology for Segments (CAMS). It includes integrating existing data with the use of Visual Basic for Applications (VBA) programming, analyzing that data with two separate statistical models, and creating two-page technical reports for UDOT engineers that briefly summarize the safety concerns present at high-priority locations. Like the RSAM and ISAM, this research is meant to help UDOT prioritize locations within the broad roadway network the agency oversees.

1.3 Scope

The methodology presented in this research is used to identify crash hot spots within the entire Utah state route network. The scope of this project includes modifying the statistical models used in the RSAM to evaluate segments independently by removing intersections and their associated crashes from the data inputs. In addition, modifications were made to the original ISAM so that the CAMS and ISAM would form a complementary pair; together they analyze the entire state route network but do so in a way that does not double-count any crashes.

1.4 Outline of Report

The body of the report is organized in the following manner.

- Chapter 1 introduces the research topic, objectives, scope, and report outline.
- Chapter 2 provides a literature review exploring topics connected to the research as well as a discussion on previous BYU-UDOT traffic safety research.
- Chapter 3 describes the data used in the CAMS research.
- Chapter 4 explains how the raw data are used to create an input file for the CAMS statistical model.
- Chapter 5 gives a brief description of the statistical model used in the CAMS research.
- Chapter 6 describes the technical reports produced for high-priority segments as well as the process that creates them.
- Chapter 7 provides and discusses the results of the CAMS research.
- Chapter 8 explains the modifications made to the ISAM in conjunction with the

CAMS research.

- Chapter 9 gives some concluding remarks including a review of the CAMS methodology and a brief discussion on future research topics.
- The chapters are followed by a References section.

2.0 LITERATURE REVIEW

2.1 Overview

A literature review was performed to understand existing segment-only crash analyses and the insights they may give into performing such an analysis on Utah roadways. This chapter summarizes the literature review and includes discussion on several key topics. The first discussion is on the determination of segment-related crashes and how they may be distinguished from those that are intersection-related. Next is a discussion on which portions of roadway lengths should be included in a segment safety analysis. Following that is a discussion on segmentation methods present in the literature. Finally, a discussion on previous BYU-UDOT research efforts on segment and intersection safety is provided along with a summary of the literature review.

2.2 The Determination of Segment-Related Crashes

There are several methods for identifying segment-related crashes in the literature. The most common method is to first determine the intersection-related crashes. The *Highway Safety Manual* (HSM), published by the American Association of State Highway and Transportation Officials (AASHTO), recommends that engineers use the intersection-related field of the crash data to determine intersection-related crashes if such a field is given on the crash report. If none is available, the HSM recommends that the engineer evaluate the characteristics of a crash to determine whether the crash was related to the intersection or the segment. The HSM comments that other entities often define intersection crashes as any crash within 250 feet of an intersection. The HSM further explains, "However, not all crashes occurring within 250 feet of an intersection can be considered intersection crashes because some of these may have occurred regardless of the existence of an intersection" (AASHTO 2010). Following this guideline, a radius of 250 feet may be used to search for intersection-related crashes but should not be the only criteria to define them.

If an intersection-related crash report field is not available in the crash data, researchers typically define the segment crashes based on their distance from the intersection. For example, Mountain et al. (1996) and Cafiso et al. (2018) chose to measure a distance of approximately 65 feet (20 meters) and 165 feet (50 meters), respectively, past the edge of the physical area of each intersection and removed all the crashes that occurred either in the intersection or within the measured distance. With only slight variation in methodology but using much larger radii, Borsos et al. (2016) and Jiri et al. (2016) both chose to measure a radius from the center of each intersection and removed all crashes within that radius. Borsos et al. (2016) used a radius of approximately 655 feet (200 meters), and Jiri et al. (2016) used a radius of approximately 330 feet (100 meters).

Some researchers have used combinations of crash type and recorded violation as criteria to define intersection-related crashes. In the segment crash analysis conducted by Pande et al. (2010), crashes with the following characteristics were removed: a left or right turn collision, an angle collision in combination with an improper turn, and an angle collision in combination with a failure to yield right-of-way. The HSM also gives the following examples for determining by the crash type whether it is a segment or intersection crash: rear-end crashes at the end of a queue of vehicles (intersection related), crashes involving a mid-block or driveway turn (segment related), and single-vehicle crashes involving adverse pavement conditions (segment related) (AASHTO 2010).

Previous BYU safety research has not been based on crash type. Although UDOT can determine whether the reporting officer considered a crash to be intersection related, this knowledge was not applied in the original ISAM. The ISAM uses a radius of influence based on the functional area of the intersection to decide which crashes are intersection related. The ISAM uses speed limit to define the functional area of the intersection. The values for the functional area, given in Table 2-1, are measured outward from the stop bar and range from 195 feet for intersections with approach speeds ≤20 mph to 1,320 feet for intersections with approach speeds ≥75 mph. All crashes within this functional area were used in the intersection statistical model (Schultz et al. 2018). These values were derived from the *Access Management Manual*, 2nd Edition, which splits the distance covered by the upstream functional area of an intersection into three parts: d_1, d_2, and d_3—the respective lengths required for perception-reaction time, lane changing and deceleration, and queue length as shown in Figure 2-1 (Williams et al. 2014). The values for d_1 and d_2 were taken from tables in *the Access Management Manual*, 2[nd] Edition, and the average queue length was assumed to be 50 feet for Utah state routes (Schultz et al. 2018).

Table 2-1: Functional Area Values Used in the ISAM

Speed (mph)	d_1	d_2	d_3	*Total*
≤20	75	70	50	195
25	90	105	50	245
30	110	150	50	310
35	130	225	50	405
40	145	290	50	485
45	165	360	50	575
50	185	440	50	675
55	200	525	50	775
60	220	655	50	925
65	240	755	50	1045
70	255	875	50	1180
≥75	275	995	50	1320

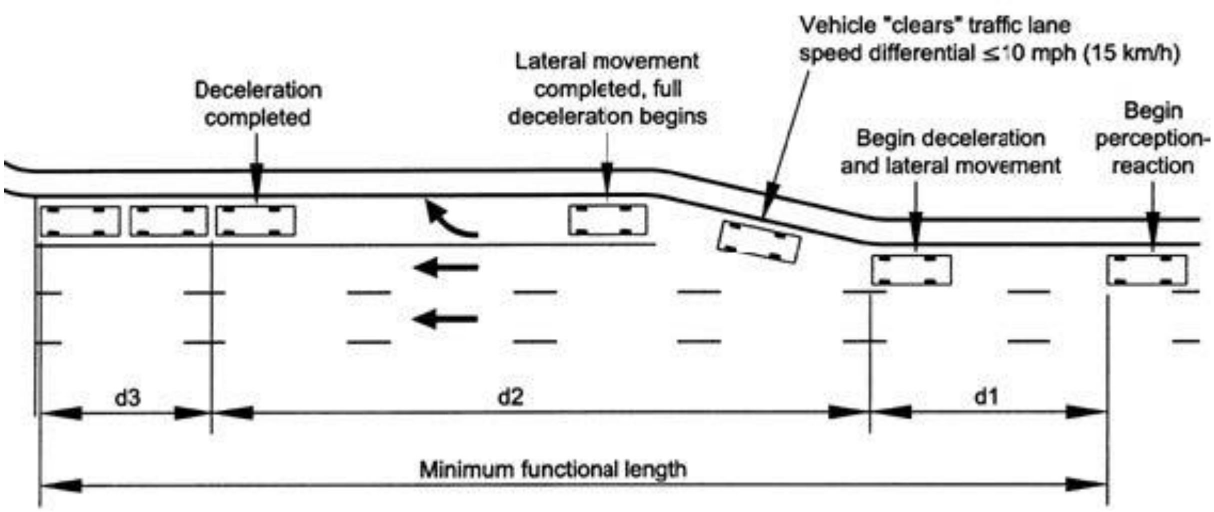


Figure 2-1: Upstream functional distance of an intersection (Rodegerdts et al. 2014).

2.3 Portions of Roadway Lengths to be Included in the Roadway Dataset

Even after deciding what crashes to include in or exclude from a segment-only analysis, decisions about the roadway network remain to be made, specifically which portions of the roadway should or should not be included in the analysis. The HSM discusses highway segments, meaning portions of roadway that may have intersections along them. It explains that all crashes that lie within the physical area of an intersection (area A in Figure 2-2) are to be considered intersection related, but that crashes occurring within the functional area of an intersection (area B in Figure 2-2) may be a mixture of intersection- and segment-related occurrences. The method given in the HSM does not remove the physical area of the intersection in the segmentation process but instead uses the intersection centers as splitting points; in effect, causing some segments to include part of the physical area of an intersection (AASHTO 2010).

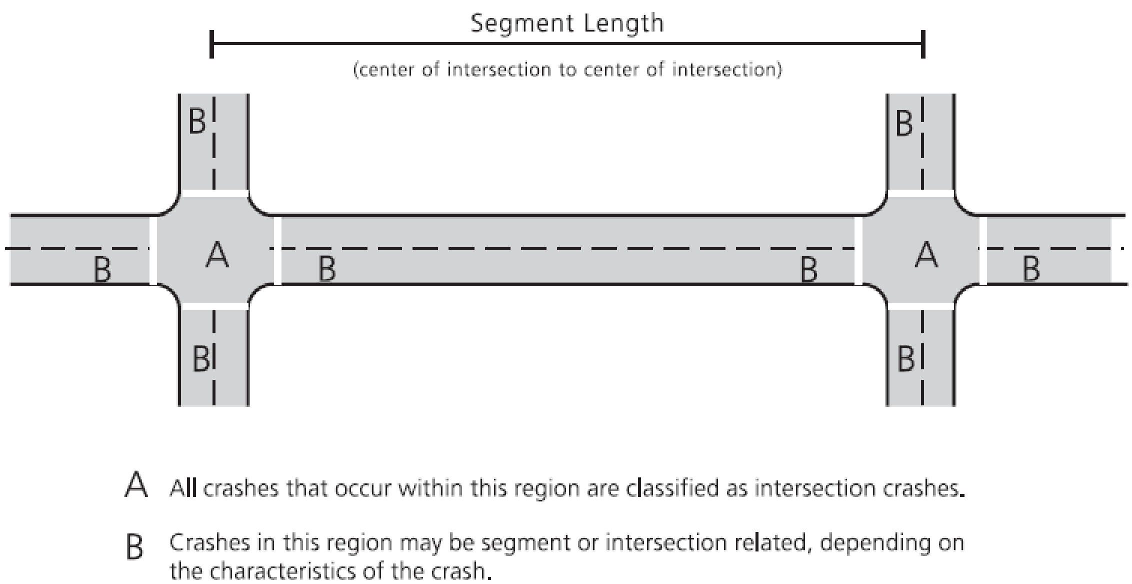


Figure 2-2: Definition of segments and intersections (AASHTO 2010).

Statewide segment analyses would include both highway and freeway segments. Although highway intersections and freeway interchanges differ in their physical characteristics, the assumption made by AASHTO that some crashes within the physical area of an intersection could be related to the segment (and not the intersection) could also be loosely applied to interchange areas of freeway segments: portions of freeway general purpose lanes influenced by nearby interchanges (i.e., portions with significant weaving, merging, and diverging movements) may contain both segment-related crashes and interchange-related crashes. The literature, however, suggests that such an assumption is not usually made for these facilities. In general, researchers who have conducted freeway segment analysis studies tended to remove the distances related to interchange-influenced behavior. In one study, Cafiso et al. (2018) removed the interchanges and their lengths of influence from the roadway database. In a different freeway segment crash study, lengths other than "basic freeway segments" (i.e., not within the merging or diverging areas) as defined in Figure 2-3 were removed from the dataset of the study by the researchers (Zheng et al. 2018).

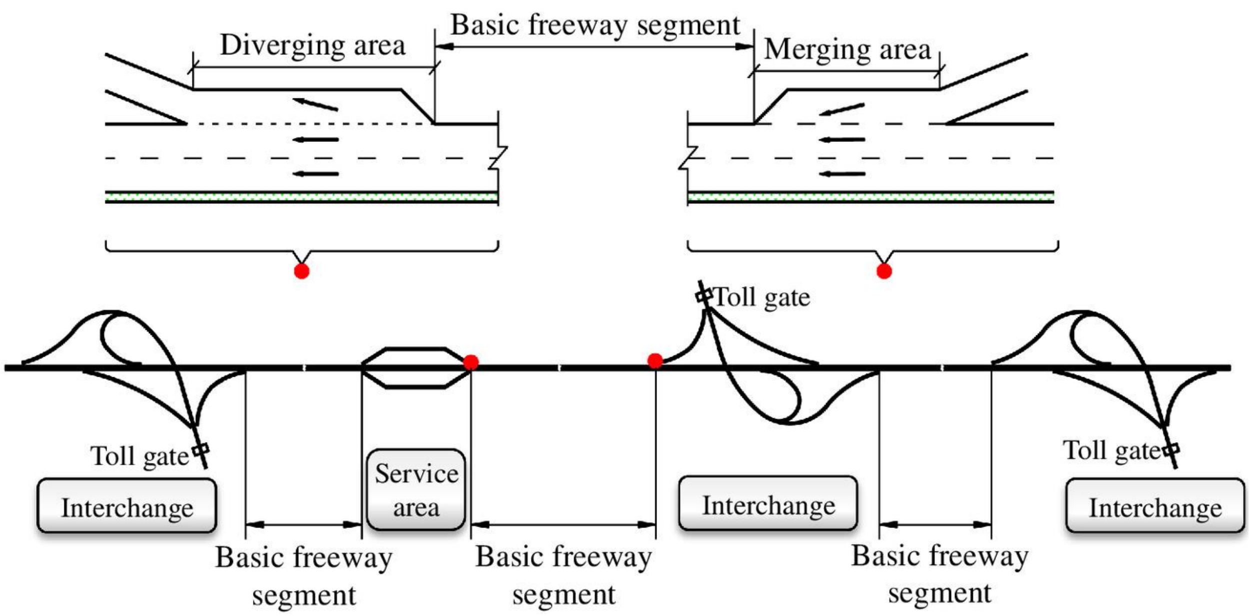


Figure 2-3: Definition for freeway segments (Zheng et al. 2018).

Much of the literature identified for this research study does not contain any reference as to whether the length of the intersection was included in the segmentation process for non-freeway research. The absence of this discussion in most segment-only research endeavors leads the reader to believe that the length of each intersection was not removed in the segmentation process. This would indicate a general agreement in the literature with the method proposed in the HSM: Remove intersected-related crashes but keep all portions of the roadway in a highway segment analysis.

2.4 The Segmentation Process

Beyond deciding what portions of roadway to include in a segment analysis, it is necessary to determine the best segmentation method or way to divide the entire network into manageable portions. Within the literature, the most common segmentation process was that of homogeneous segmentation. Table 2-2 shows a sampling of research teams that implemented a homogeneous segmentation into their crash analyses, including the variables that were used in the process. The starred values in the table represent characteristics that may be included in the HSM definition. The HSM defines a homogeneous segment as “a portion of roadway with similar average daily traffic volumes (veh/day), geometric design, and traffic control features” and typically separates segment analyses by urban/rural and number of lanes (AASHTO 2010).

Table 2-2: Variables Used in Homogeneous Segmentation Methods

	Variables Used (CCR = Curvature Change Rate; RHR = Roadside Hazard Rating)										
Source	*AADT*	*CCR*	*Functional Class*	*Grade*	*Number of Lanes*	*Percent Tunnel*	*RHR*	*Speed Limit*	*Urban Code*	*Width (shoulder)*	*Width (roadway)*
AASHTO (2010)	X	*	*	*	X		*	*	X	*	*
Borsos et al. (2016)	X	X						X		X	X
Cafiso et al. (2010)	X	X					X				X
Cafiso et al. (2018)		X		X		X	X				
Kwon et al. (2013)					X				X		
Schultz et al. (2016)	X		X		X			X	X		

*Represent characteristics that may be included in the HSM (AASHTO 2010) definitions depending on the roadway type and statistical validity.

In addition to the research cited in Table 2-2, Gaweesh et al. (2017) and Ogle et al. (2017) also performed roadway segmentation. The researchers did not use an original set of variables, but instead expressly stated that the AASHTO method was implemented and were thus not included in the table.

The research performed by Schultz et al. (2016) referenced in Table 2-2 was performed on roadway and crash data from UDOT that covered the entire network of state routes. The variables used in the segmentation process have been used in similar BYU research dating back to 2012 where BYU researchers established a framework for crash data analysis that included four roadway characteristics used for homogeneous segmentation: average annual daily traffic (AADT), functional class, number of through lanes, and speed limit (Schultz et al. 2012).

Beginning in 2013, BYU-UDOT crash analysis research has included urban code as a fifth segmentation variable (Schultz et al. 2013).

Despite homogeneous segmentation being the most common method for roadway segmentation, other methods do exist in the literature. For example, Zheng et al. (2018) did not segment their freeway model beyond the breaks made at the interchanges as discussed previously and as shown in Figure 2-3.

As another example of research that differed from the majority, Cafiso et al. (2018) performed four different segmentation methods on their freeway segment analysis to find the method that resulted in the best statistical fit for their data. The first method was to create homogeneous segments with respect to two variables: AADT and curvature. The second method was to create all segments with exactly two curves and two tangents. The third method was to create segments of constant length. The researchers chose this length to be approximately 2,135 feet (650 meters) which is the same length as the largest interchange in their analysis. The fourth method was to create homogeneous segments with respect to four variables: curvature change rate, grade, tunnel presence, and roadside hazard rating. According to their research, the two best-fitting models were the third (fixed length) and second (two curves and two tangents) methods.

2.5 Previous BYU-UDOT Research

Among the UDOT-contracted research performed at BYU are two methodologies related to the present research: the RSAM for segments (Schultz et al. 2016) and the ISAM for intersections (Schultz et al. 2018). This section gives background on these two methodologies as well as their connection to the present research.

For all the UDOT-contracted research discussed in this report, crash severity levels are rated according to the KABCO scale used by the Federal Highway Administration (FHWA) (2017). Severity is coded as an integer between 1 and 5 as outlined in Table 2-3. The term “injury crashes” will be used in this report to mean Severities 3, 4, and 5 (KAB); “total crashes” will mean crashes of Severity 1 through 5 (KABCO).

Table 2-3: Severity Level Codes

Code	*Description*	*FHWA Code (2017)*
5	Fatal	K
4	Suspected Serious Injury	A
3	Suspected Minor Injury	B
2	Possible Injury	C
1	No Apparent Injury	O

2.5.1 Roadway Safety Analysis Methodology

The RSAM was first created by a BYU research team in 2016 and was the first phase in BYU-UDOT research to create a statewide model identifying locations with high potential for safety improvement. Because it analyzes roadway segments, it forms the foundation for the CAMS which looks more specifically into reducing segment-related crashes.

The three parts of the RSAM aim at identifying hot spots along Utah's state route network based on crash data and segments of similar characteristics. First, the data are prepared into one cohesive file of segments, their characteristics, and the crashes pertaining to them; second, the segments undergo statistical analysis; and third, technical reports are created for high-priority segments. The following sections will describe these three parts, all of which can be found in more detail in the UDOT report published by Schultz et al. (2016).

2.5.1.1 Data Preparation

The first part of the RSAM is to prepare the data in such a way that they could be used as an input to statistical models. All the necessary data comes from UDOT, most of which can be found on the UDOT Open Data Portal (UDOT 2017). Due to privacy concerns, sensitive crash data are not available to the general public.

The data preparation is done with the use of VBA programming, and an outline of this part of the RSAM is shown in Figure 2-4. Four crash data files (Crash Data, Crash Rollup, Crash Location, and Vehicles) are combined into one file. The Crash Locations file is used to identify which crashes occurred on a state route and all other crashes are deleted. Information from the three other crash files are then attached to the remaining crashes by matching crash Identification

Numbers (IDs) across the files. In the analysis published in UDOT Report UT-16.13 (Schultz et al. 2016), crash data from the years 2010-2014 were used. In addition to the crash data, five files of roadway characteristic data (AADT, Functional Class, Lanes, Speed Limit, and Urban Code) are also combined into one new file. Route names and mileposts are compared across the five files to combine the data. Integrated into this process is a segmentation method that is used to break down the Utah state route network into small portions. These segments, ranging from one-tenth of a mile to several miles in length, are created in a homogeneous manner, meaning that the AADT, functional class, number of lanes, speed limit, and urban code never change mid-segment and that neighboring segments vary by one or more of these five characteristics.

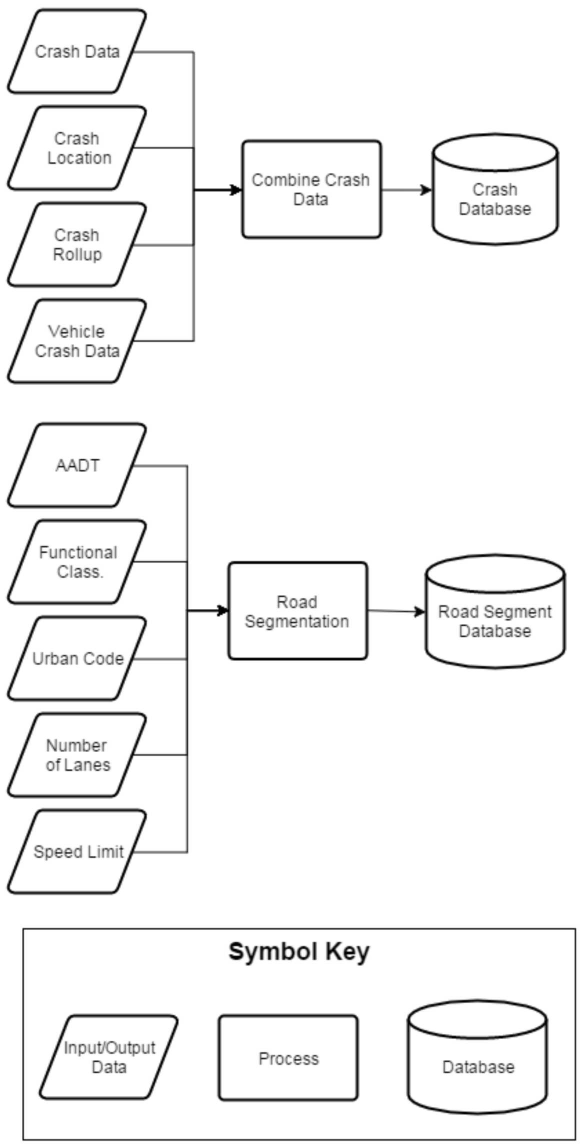
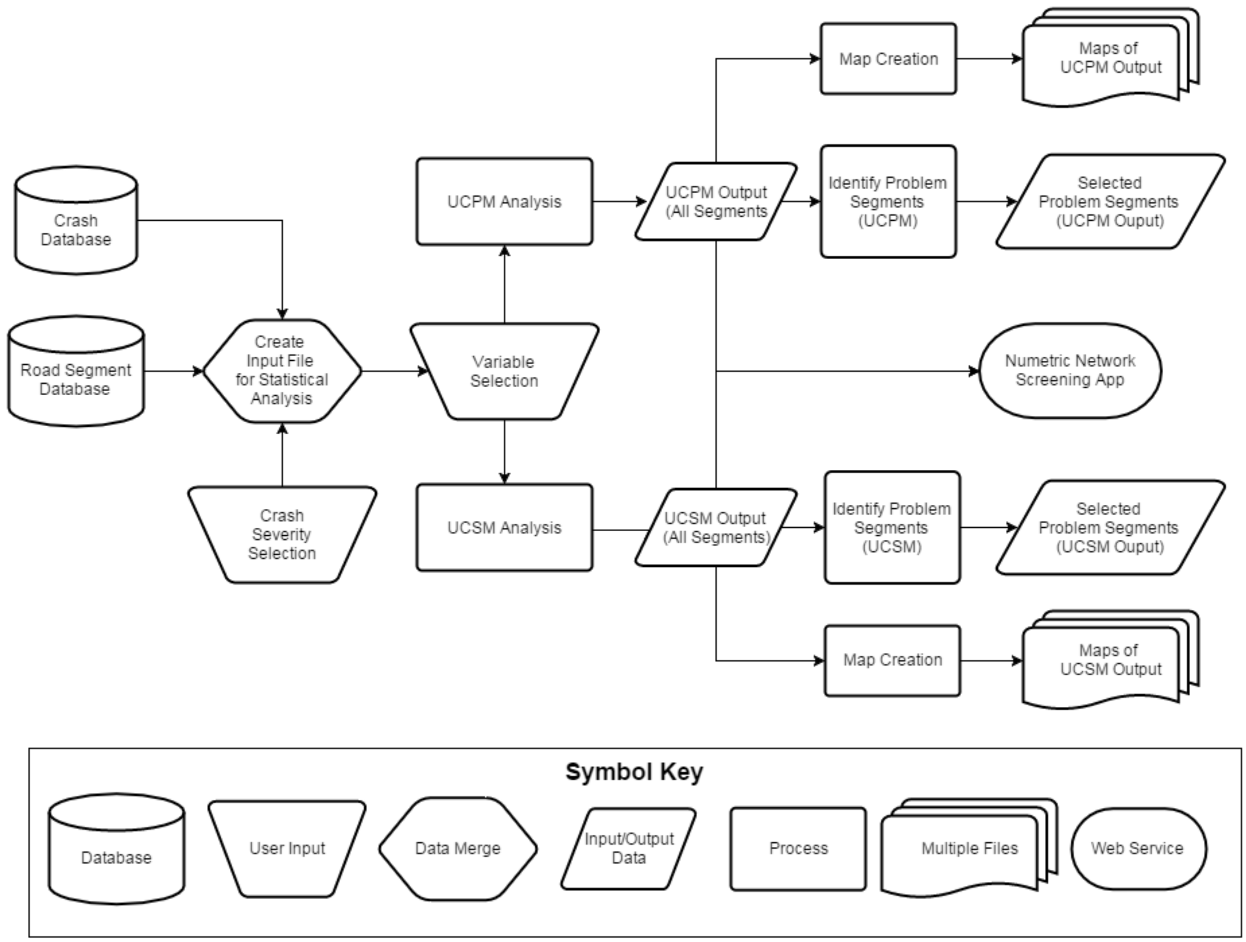


Figure 2-4: Flowchart for RSAM data preparation (Schultz et al. 2016).

The data from the two new files, one containing crash information and the other containing segment information, are then integrated together. Each segment is given a unique ID to distinguish it from the others and to allow for quick reference between files. Crashes are matched to segments based on the route and milepost at which the crash occurred, and crash totals are appended onto each line of segment data. In addition, a column is added to the crash data file that contains the ID of the segment with which the crash is associated. This final data preparation process results in two files: one containing detailed segment information with associated crash totals and the other containing detailed crash information organized by associated segment.

2.5.1.2 Statistical Analysis

The second part of the RSAM is to determine hot spots, or portions of highway that have observed significantly more crashes in a five-year period (2010-2014) than was predicted for that same time span. Two separate statistical analyses were prepared and can be used in the RSAM, and a flowchart showing the application of these two analyses is shown in Figure 2-5. The first is the Utah Crash Prediction Model (UCPM), and the second is the Utah Crash Severity Model (UCSM). The UCPM predicts how many crashes of specified crash severities (e.g. 3, 4, 5) are likely to occur along a segment, whereas the UCSM predicts the number of injury crashes to occur along the segment based on the total number of crashes that occurred. Despite these differences, however, the models have a lot in common. Both of the models take the same input (the detailed segment information created in the data preparation process) and create predicted distributions of crashes for each segment. Furthermore, the observed number of crashes on each segment is compared to the predicted distribution and associated with a percentile value within that distribution. The segments are then ranked according to the percentile values with a higher percentile value representing a greater safety concern. The resultant rankings are then used to determine which segments are considered to be of highest priority for safety improvements.



Note: the Numetric Screening App is no longer in use.

Figure 2-5: Flowchart of the RSAM statistical analyses (Schultz et al. 2016).

2.5.1.3 Technical Reports

The third part of the RSAM is to create two-page technical reports, called Roadway Safety Analysis Reports (RSARs), for high-priority segments. This process begins with a few steps in the ArcMap geospatial software published by Esri (2019) to calculate roadway conditions such as grade, curvature, and number of signs per mile that are displayed in the RSARs. Python scripts compatible with ArcMap were written by the research team specifically for this purpose.

The process also includes using additional VBA code to populate tables found in the RSARs. These tables display information taken from the calculations performed in ArcMap (i.e., roadway characteristics) and from the files created in the data integration process (i.e., segment identification and crash data). Once the automated steps have been completed, research analysts then take individual RSARs and perform virtual site visits using online tools to gather more information on the background and current conditions of each segment. In years past, RSARs for the ten highest-priority segments in each UDOT Region were presented to UDOT for further evaluation. A flow chart of the report creation process is shown in Figure 2-6.

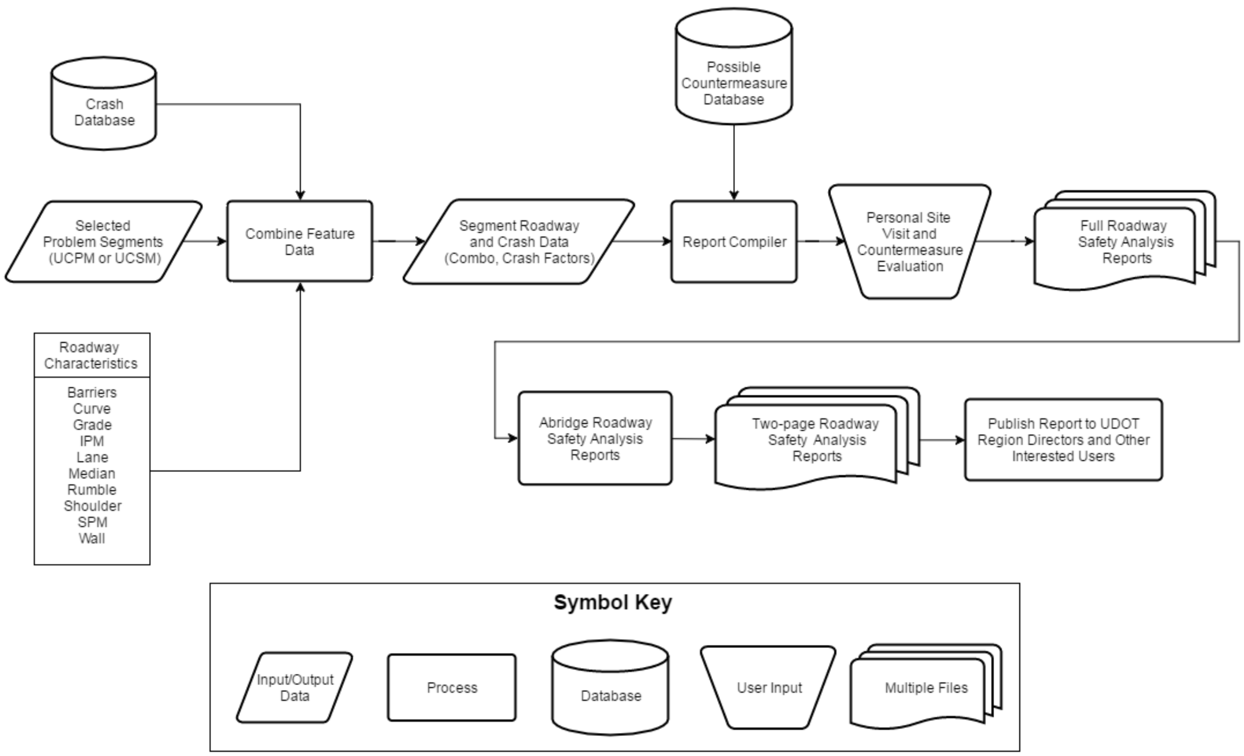


Figure 2-6: Flowchart of the RSAM technical report creation process (Schultz et al. 2016).

2.5.2 Intersection Safety Analysis Methodology

First completed in 2018, the ISAM was developed to analyze and rank state route to state route intersections. The process is shown in Figure 2-7. In a similar fashion to the RSAM process, roadway data and crash data are first combined separately before being merged together. A statistical model performs a predictive analysis and then compares the predicted results with the actual crash counts, ranking the intersections in order of potential for safety improvement. Finally, two-page technical reports are created for top-ranking intersections. This section will describe each of these three steps as implemented in the ISAM when it was first completed in 2018.

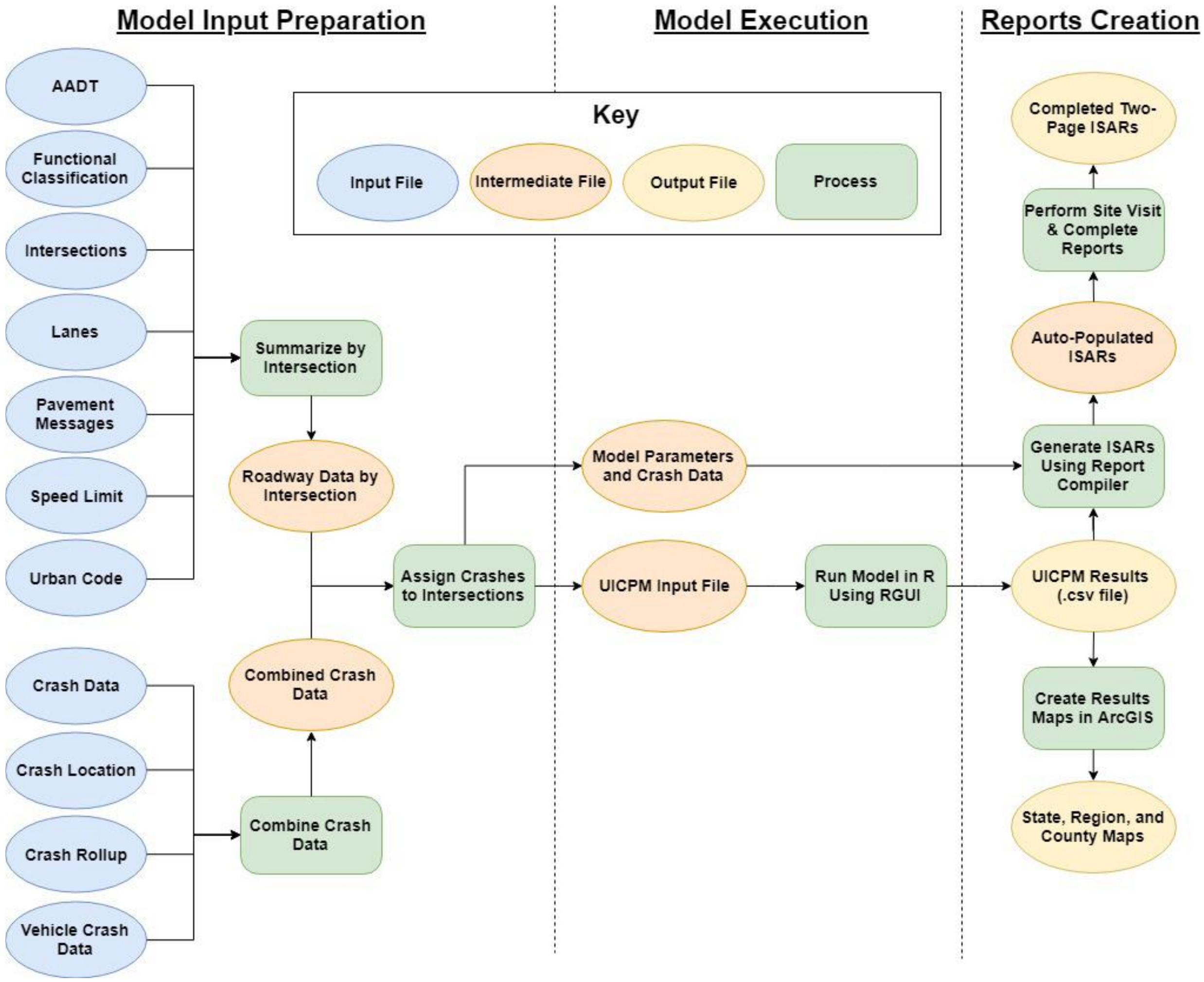


Figure 2-7: The ISAM process (Schultz et al. 2018).

2.5.2.1 Data Preparation

The first part of the ISAM is to combine and prepare the data so that the data can be used to perform a statistical analysis. Like the RSAM, all the necessary data come from UDOT, most of which can be found on the UDOT Open Data Portal (UDOT 2017). Again, due to privacy concerns, sensitive crash data are not available to the general public.

The data preparation is done with the use of VBA programming. In identical manner to the RSAM, four crash data files (Crash Data, Crash Rollup, Crash Location, and Vehicles) are combined into one file. The Crash Locations file is used to identify which crashes occurred on a state route and then information from the other three crash files are attached to those crashes by matching crash IDs. In the analysis published in UDOT Report UT-18.06 (Schultz et al. 2018), crash data from the years 2010-2016 were used. In addition to the crash data, seven roadway files (AADT, Functional Class, Intersection, Lanes, Pavement Messages, Speed Limit, and Urban Code) are also combined to create one new file. Intersections with at least two distinct state routes are identified in the Intersections file, and data from the other files pertaining to each intersection are found and matched by comparing route and milepost information. The input form used to begin both the roadway data preparation and the crash data preparation processes is shown in Figure 2-8.

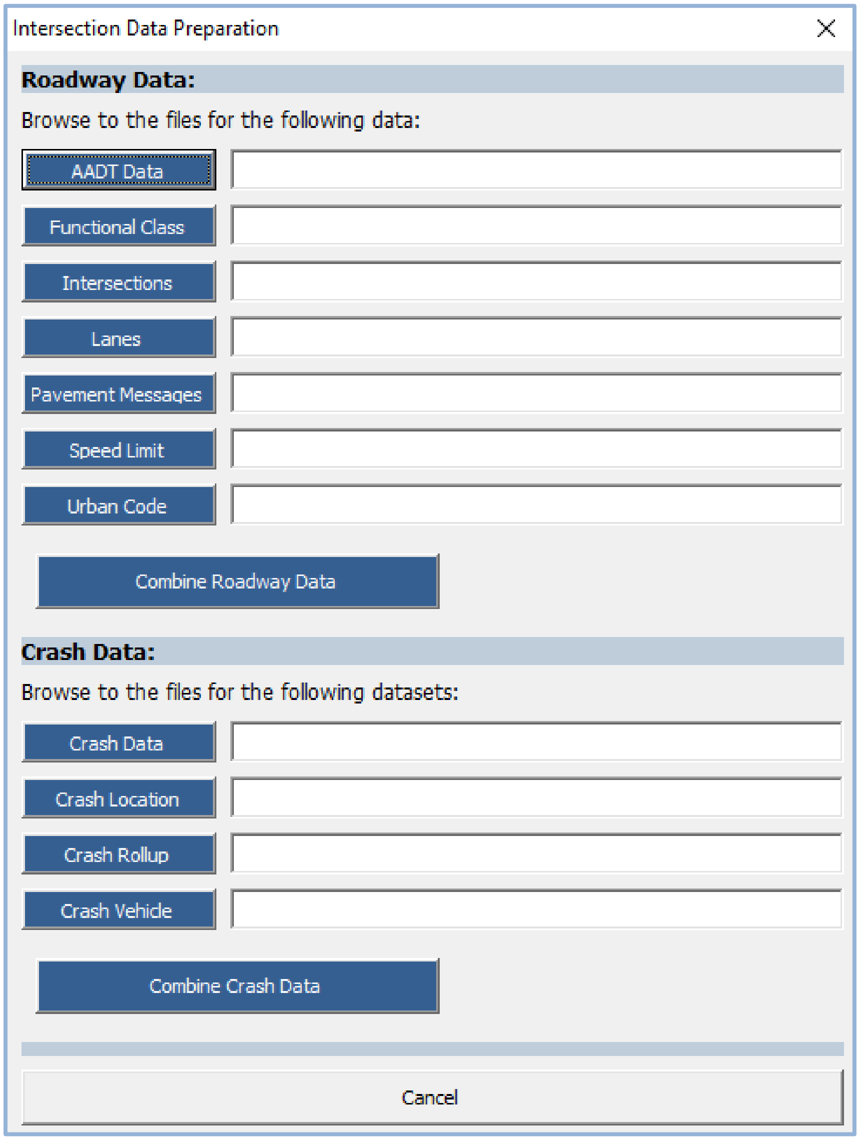


Figure 2-8: Input form for ISAM data preparation.

Once the roadway and crash data have been individually combined into new files, the final data preparation process can begin. This process gives a unique ID to each intersection and then assigns crashes to individual intersections based on their locations. Crashes are associated with an intersection if they occurred within the physical area of the intersection (as bounded by the stop bars) or within the functional area of the intersection. The functional area is represented by a distance beginning at each stop bar and extending away from the intersection along each leg of the intersection. This distance can be defined in several ways, including according to the approach speed limit (recommended), the urban code, and the functional class of the intersection leg. The ISAM has historically only been run with the recommended (approach speed limit) method. The values for the functional area distance based on the approach speed limit were previously given in Table 2-1, and their origin is discussed in Section 2.2.

The process of combining the crash data with the intersection data ends in the creation of two files. One file contains a list of the intersection IDs; roadway characteristics and summarized crash data are given for each intersection in this file. The other file contains a list of the crashes included in the analysis; crash characteristics and the ID of the associated intersection are given for each crash in this file.

2.5.2.2 Statistical Analysis

The second part of the ISAM is to perform a statistical analysis on the data to find intersections that have more injury crashes than predicted. The statistical model is called the Utah Intersection Crash Prediction Model (UICPM) and is a Zero-Inflated Poisson (ZIP) model. In a manner similar to the UCPM for segments, the UICPM uses seven years (2010-2016) of crash data to build the model and creates distributions of predicted crashes for each intersection. The median value of each predicted distribution is compared to the actual (observed) annual crash rate as averaged over the seven years of crash data. Intersections are then ranked according to the percentile value of their observed crash rate within the predicted crash distribution.

2.5.2.3 Technical Reports

The third and final part of the ISAM is the creation of two-page technical reports referred to as Intersection Safety Analysis Reports (ISARs) created for high-ranking intersections. The first page includes tables displaying information about the location and layout of the intersection as well as crash data summed up by severity and by crash factor. The information in this page is automatically populated using VBA macros. The second page is filled out manually by a research analyst and contains summaries of the historical and current conditions of the intersection, an aerial photo of the site, and a list of potential countermeasures.

2.6 Summary

The literature provides several different methodologies for conducting a crash hot spot analysis for roadway segments. The most common way to distinguish between a segment-related crash and an intersection-related crash is to use the distance of the crash from a known intersection, but the distance used to distinguish segment-related crashes from intersection-related crashes varies from one research study to another. However, the literature appears to agree that all portions of the roadway network should be included in a segment safety analysis, even if some roadway portions observe significantly more intersection-related crashes than they do segment-related crashes. The most common method for segmenting the roadway network was homogeneous segmentation, yet the variables used varied between studies.

BYU-UDOT research efforts in years past have included both a roadway crash hot spot analysis and an intersection-specific crash hot spot analysis. The roadway analysis includes all crashes, whereas the intersection analysis includes only crashes that are determined to be intersection related. This review of the literature serves as the basis for obtaining ideas about how a segment-related crash hot spot analysis might be performed for Utah state routes.

3.0 DATA COLLECTION

3.1 Overview

This chapter describes the different data files used in the CAMS research, how they may be obtained, what purpose(s) they serve, and the ultimate products that they help to create. A total of ten raw input files are used in the model; six pertaining to roadway data and four pertaining to crash data. This chapter will explain each file in these two groups (roadway data and crash data).

3.2 Roadway Data

The six roadway data files used in the CAMS are: AADT, Functional Class (for state routes only), Intersections, Lanes, Speed Limit, and Urban Code. Three of these files (AADT, Lanes, and Speed Limit) are accessible to the public via the UDOT Open Data Portal (UDOT 2017).The Urban Code file has previously been accessible on the UDOT Open Data Portal, but it was unavailable on the website at the time of this report. The Functional Class and Intersections files used in this research were recently updated by UDOT to provide data columns needed for this research. These files are not available on the UDOT Open Data Portal but may be obtained upon request. Each of the six roadway data files will be discussed in more detail in the following subsections.

3.2.1 AADT Data

UDOT collects AADT data for state routes and federal aid routes in Utah once every three years according to recommendations in the Traffic Monitoring Guide published by FHWA (2016). The AADT file includes information about where the data were collected, lists the route number as well as the starting and ending mileposts for each line of data, and provides data ranging from the year 1981 until the most recent year of available data (2018 at the time of the report). In addition, the file quantifies single-unit truck traffic and combination truck traffic as percentages of AADT for the most recent year of available data.

According to the surveyed literature, AADT is the most common variable found in the various methods of homogeneous segmentation (See Table 2-2 in Section 2.4). It is one of the five variables that are used to create homogeneous segments in the BYU model. The AADT for each high-priority segment is provided in two-page technical reports (these reports are discussed in more detail in Section 6.3). The truck traffic information is also useful for the analysis. Many interstates and highways within Utah are popular freight routes and, as a result, have relatively high percentages of truck traffic compared to other routes in the state. The percent of truck traffic is used as a variable in the statistical models as explained in more detail in Chapter 5.

3.2.2 Functional Classification Data

The Functional Classification file provides the route number, county, and beginning and ending mileposts for all state routes. The file used in the CAMS was last updated in March 2019. The functional classification is given in both numerical code and text description formats. Codes and their corresponding descriptions are summarized in Table 3-1.

Table 3-1: Functional Classification Codes and Descriptions

Code	*Description*
1	Interstate
2	Other Freeways and Expressways
3	Other Principal Arterial
4	Minor Arterial
5	Major Collector
6	Minor Collector
7	Local

The functional classification data is one of the five variables used to create homogeneous segments in the BYU model. The functional classification for each high-priority segment is provided in the two-page technical reports.

3.2.3 Intersections Data

Updated by UDOT in May 2019, the Intersections file contains a record for every intersection on every Utah state route. Previous to this new file, information involving intersections with only one state route was limited, and data for the non-state routes at these intersections were not available. With the addition of information pertaining to state route intersections with federal aid and local routes in the Intersections file, it is possible for the user to select the intersections of choice and exclude the crashes at or near these intersections from the CAMS analysis.

The Intersections file provides the main route number and milepost of the intersection as well as a brief description of the intersection type and traffic control used, all of which are provided in the two-page technical reports. The file also has columns that tell whether skew, railroad tracks, and/or another state route are present at the intersection. Additional columns include intersection latitude and longitude, and the UDOT Region and maintenance station in which the intersection lies.

3.2.4 Lanes Data

The Lanes file is a compilation of homogeneous stretches of state routes according to their number of lanes and lane width. Each segment has a route number, direction, beginning milepost, and ending milepost. Additional information provided in the Lanes files for each segment includes the UDOT Region, counts of different lane types on that segment (e.g., through lanes, auxiliary lanes, left turn lanes, etc.), beginning and ending coordinates, and beginning and ending elevation. This file was downloaded from the UDOT Open Data Portal in June 2019.

Although the Lanes file provides information on all types of lanes, only the information about through lanes is used in the BYU model. The number of through lanes is one of the five variables used to create segments with homogeneous characteristics.

3.2.5 Speed Limit Data

The Speed Limit file on the UDOT Open Data Portal provides the speed limit and beginning and ending mileposts for segments of the same legal speed limit on all state routes in Utah. This file was most recently updated in 2017.

The posted speed limit for a segment of highway is used to create segments of homogeneous characteristics. It is also the recommended way to calculate intersection functional area.

3.2.6 Urban Code Data

The Urban Code file provides information about defined urban areas including the beginning and ending mileposts for each route that exists in each urban area. The file used in the CAMS was obtained from UDOT in May 2016. Urban areas in Utah consist of the following: Logan, Ogden-Layton, Provo-Orem, Salt Lake City, and St. George. In addition, the urban code file may also identify road segments as small urban, rural, and unknown. Each of these eight urban types (including the five urban areas) has a unique five-digit code.

Because the large scope of a statewide model includes both urban and rural areas, it is important to distinguish between the two, especially since crash patterns on rural roadways often behave differently than those in urban areas. Utah's urban code makes up one of the five variables used to create homogeneous segments in the BYU model.

3.3 Crash Data

Because it may contain personal information, detailed crash data are not available to the general public. The crash data collected by UDOT are saved into four files: Crash Data, Location, Rollups, and Vehicle files. Each of the files contains a different type of information, but the records can be linked together by a unique crash ID for each incident.

3.3.1 The Crash Data File

The Crash Data file provides information about the conditions in which the crash occurred. These conditions include roadway, weather, lighting, pavement, junction, work zone, horizontal and vertical curves, manner of collision, and first harmful event, all of which are recorded with various numerical codes.

3.3.2 The Crash Location File

The Crash Location file provides information including geographical coordinates (given in Universal Transverse Mercator (UTM) X and Y), city, route, and milepost. State routes are numbered by integers less than 1000 (0 to 999), and federal aid routes are numbered by integers between 1000 and 9999. Numbers containing five digits or more represent a city or county code for crashes on local roads or crashes that could not be located.

This file is useful for selecting only crashes that occurred on a state route and for assigning crashes to segments based on route number and milepost.

3.3.3 The Crash Rollup File

The Crash Rollup file provides information about the circumstances of each crash. The number of vehicles and pedestrians involved in each crash are given along with the number of each severity type in each incident. Severity is coded as an integer between 1 and 5 as outlined previously in Table 2-3. In addition, there are more than two dozen Yes/No fields for different crash scenarios (such as pedestrian involved, adverse roadway surface condition, night or dark conditions, and speed related). A complete list of these fields is provided later in this report in Table 7-3.

The Crash Rollup file is useful to the project because the intersection-related Yes/No field is used to filter out crashes that pertain to intersections at select locations. Tallies from the rollup Yes/No fields are also displayed in the final two-page technical reports and are used to help identify appropriate countermeasures.

3.3.4 The Crash Vehicle File

The Crash Vehicle file provides more specific information about each vehicle involved in a crash. It provides the sequence of events for each vehicle as well as the posted speed limit, estimated travel speed, and travel direction.

The sequence of events contained in this file are used in the report-making process as described in Chapter 6. They are inserted into tables to help research analysts identify the potential safety problem on high-priority segments.

3.4 Summary

The data essential for the CAMS process are all collected by UDOT and include files for AADT, functional class, intersections, lanes, speed limit, urban code, crash data, crash location, crash rollup, and crash vehicle data. Each data file provides unique and important information for the CAMS process. Roadway characteristics are important for portioning the roads into homogeneous segments, whereas crash data are used to assign crash totals to those segments. These data files are used in the data integration process as described in the following chapter.

4.0 DATA INTEGRATION PROCESS

4.1 Overview

The purpose of this chapter is to describe how the raw data files are combined and analyzed to produce the input to the statistical model. For these segments to be determined and their associated crashes to be assigned, raw data files of both crash and roadway characteristics information must be merged using VBA macros. These macros are hosted by a Microsoft Excel Macro-Enabled worksheet referred to as the R Graphical User Interface (RGUI) workbook. This chapter will describe the methods for combining the roadway data to create a dataset of segments, combining the crash data to create a list of all relevant crashes, and assigning crashes to segments to create the input to the statistical model.

4.2 Combining the Roadway Data

The purpose of combining the roadway data files is to create a dataset of segments with homogeneous characteristics. Such a dataset allows for statistical comparison and prediction between similar segments and is a crucial part in the crash analysis methodology. Combining several files of roadway data into one useful file is made possible using VBA macros. An outline of this process is shown in Figure 4-1. This section will describe the required user input, the VBA process, and the final segment file.

4.2.1 Required User Input

Upon beginning this process, the user must fill out an input form in the RGUI workbook as shown in Figure 4-2. Five files are required: AADT, Functional Class, Lanes, Speed Limit, and Urban Code. The user can choose between homogeneous segmentation (where the user may also enter a desired minimum segment length) and fixed length segmentation. Homogeneous segmentation is recommended. Once all the required information has been entered and the user has clicked the Combine Roadway Data button, the VBA macros begin the process of combining the data and making any necessary calculations as described in the next section.

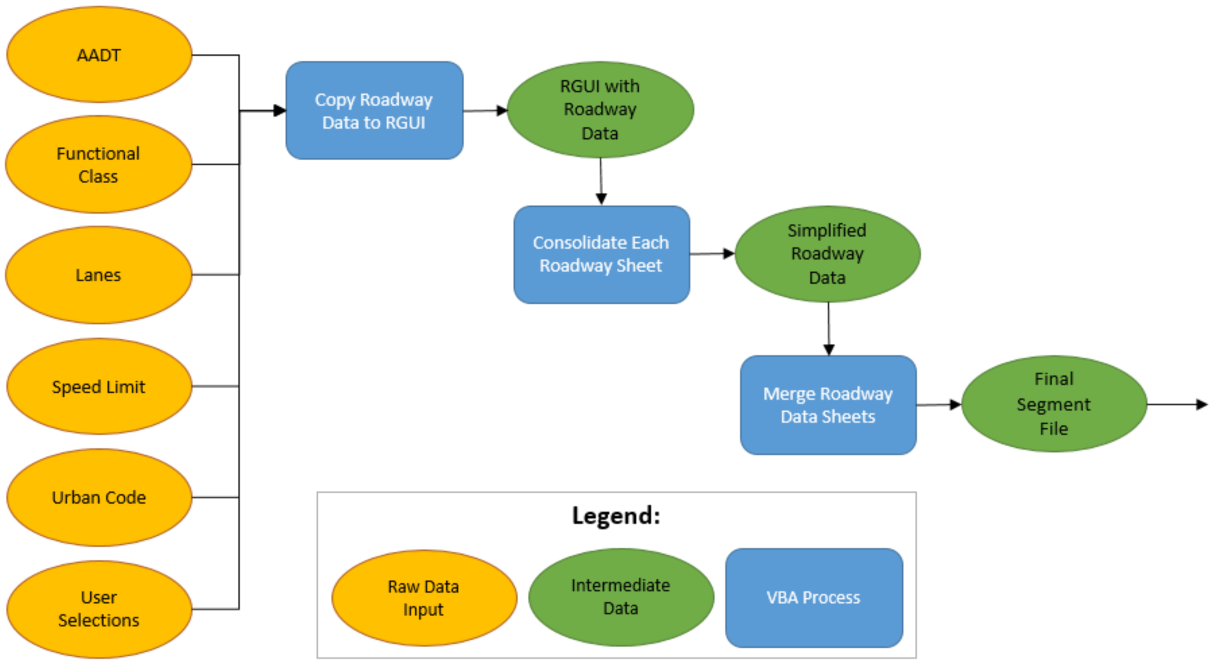


Figure 4-1: Outline of CAMS process to combine roadway files.

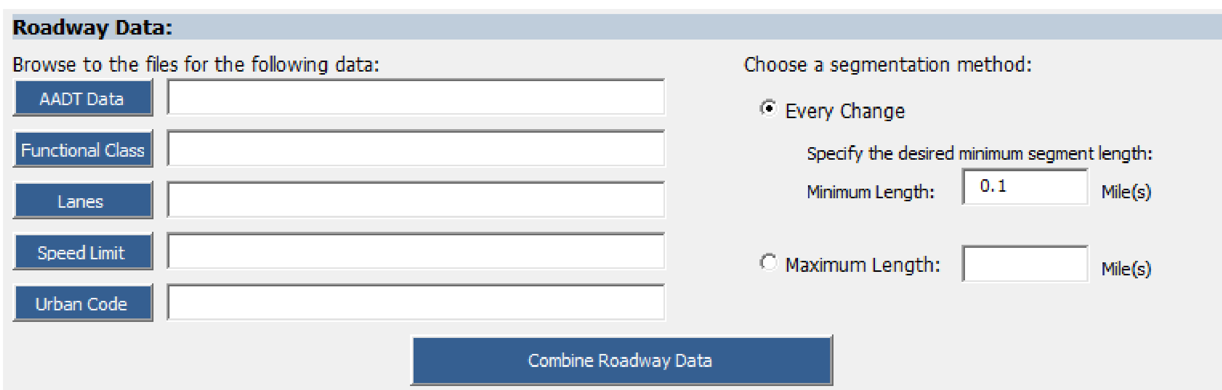


Figure 4-2: CAMS roadway data input form.

4.2.2 VBA Process Description

The initial step in the VBA process, "Copy Roadway Data to RGUI" shown in Figure 4-1, is to copy the data in each of the roadway files and paste the rows and columns into separate sheets in the RGUI workbook. For the Functional Class, Lanes, Speed Limit, and Urban Code files, this step is done with only a few minor formatting and clarification edits. The AADT file goes through a similar editing process and also creates columns for UDOT Region and total percent of trucks.

The second step in the VBA process, “Consolidate Each Roadway Sheet” shown in Figure 4-1, is to consolidate the datasets. For example, there may be two or more consecutive lines in the Speed Limit file that represent adjacent sections of highway with the same posted speed limit. Whenever that is the case, these lines are combined into one—the beginning milepost of the single line is taken from the first section of adjacent sections, and the ending milepost is taken from the final section.

The third and final step in the VBA process, “Merge Roadway Data Sheets” shown in Figure 4-1, is to merge all of the roadway data into one sheet, each line of data representing a unique segment of highway. In this step, the bounding mileposts for each segment are determined and the segment length is calculated. Segments that are smaller than the minimum segment length inputted by the user are combined with the next consecutive segment. Segment lengths, mileposts, and other data are also updated accordingly. To finish, the sheet with the segment data is exported to a new excel workbook and saved for future use.

One thing to note about this process is that the interstate routes (I-15, I-70, I-80, I-84, and I-215) and Mountain View Corridor (UT-85) have been split into two directions each. This means that the positive direction (i.e., northbound or eastbound) is considered a separate segment from the negative direction (i.e., southbound or westbound) of the same route. For these split segments, each direction is assumed to have exactly one-half of the bi-directional AADT. Although other divided highways exist in Utah, the CAMS only splits the six routes mentioned and analyzes all other highways with combined positive and negative directions.

4.2.3 Final Segment File

The file created by the combine roadway files process has several key columns describing each segment, including route, beginning milepost, ending milepost, AADT for 2010 and each year following, functional class, number of through lanes, speed limit, and urban code. Each line of the file lists a unique segment. A segment includes both the positive and negative directions of travel unless it is a segment of an interstate (I-15, I-70, I-80, I-84, and I-215) or Mountain View Corridor (UT-85). Adjacent segments vary from each other in at least one of the considered roadway characteristics (AADT, functional class, number of through lanes, posted speed limit, and urban code). A sample of this file is given in Figure 4-3.

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	AA	AB	AC
1	SEG_ID	LABEL	BEG_MILEPOINT	END_MILEPOINT	Seg_Length	Route_Name	ROUTE_ID	DIRECTION	FC_Code	FC_Type	COUNTY	REGION	AADT_2018	AADT_2017	AADT_2016	AADT_2015	AADT_2014	AADT_2013	AADT_2012	AADT_2011	AADT_2010	Single_Percent	Combo_Percent	Single_Count	Combo_Count	Total_Percent	Total_Count	SPEED_LIMIT	Num_Lanes
2	1	0006P	0	10.88	10.88	6	6	P	3	Other Principal Arterial	MILLARD	4	410	410	400	380	350	330	330	330	340	0.25	0.232	102	95	0.482	197	65	2
3	2	0006P	10.88	20.146	9.266	6	6	P	3	Other Principal Arterial	MILLARD	4	410	410	400	380	350	330	330	330	340	0.25	0.232	102	95	0.482	197	40	2
4	3	0006P	20.146	24.499	4.353	6	6	P	3	Other Principal Arterial	MILLARD	4	410	410	400	380	350	330	330	330	340	0.25	0.232	102	95	0.482	197	65	2
5	4	0006P	24.499	25.248	0.749	6	6	P	3	Other Principal Arterial	MILLARD	4	410	410	400	380	350	330	330	330	340	0.25	0.232	102	95	0.482	197	50	2
6	5	0006P	25.248	27.1	1.852	6	6	P	3	Other Principal Arterial	MILLARD	4	410	410	400	380	350	330	330	330	340	0.25	0.232	102	95	0.482	197	35	2
7	6	0006P	27.1	27.81	0.71	6	6	P	3	Other Principal Arterial	MILLARD	4	410	410	400	380	350	330	330	330	340	0.25	0.232	102	95	0.482	197	50	2
8	7	0006P	27.81	46.017	18.207	6	6	P	3	Other Principal Arterial	MILLARD	4	410	410	400	380	350	330	330	330	340	0.25	0.232	102	95	0.482	197	65	2
9	8	0006P	46.017	77.545	31.528	6	6	P	3	Other Principal Arterial	MILLARD	4	370	370	410	390	360	340	340	340	360	0.175	0.334	65	124	0.509	189	65	2
10	9	0006P	77.545	82.075	4.53	6	6	P	3	Other Principal Arterial	MILLARD	4	530	530	510	480	450	420	420	530	550	0.162	0.265	86	140	0.427	226	65	2
11	10	0006P	82.075	82.357	0.282	6	6	P	3	Other Principal Arterial	MILLARD	4	530	530	510	480	450	420	420	530	550	0.162	0.265	86	140	0.427	226	55	2
12	11	0006P	82.357	82.885	0.528	6	6	P	3	Other Principal Arterial	MILLARD	4	530	530	510	480	450	420	420	530	550	0.162	0.265	86	140	0.427	226	40	2
13	12	0006P	82.885	83.467	0.582	6	6	P	3	Other Principal Arterial	MILLARD	4	2000	2000	1900	1800	1700	1600	1600	1600	1600	0.15	0.196	300	392	0.346	692	40	2
14	13	0006P	83.467	83.897	0.43	6	6	P	3	Other Principal Arterial	MILLARD	4	2000	2000	1900	1800	1700	1600	1600	1600	1600	0.15	0.196	300	392	0.346	692	55	2
15	14	0006P	83.897	87.432	3.535	6	6	P	3	Other Principal Arterial	MILLARD	4	3500	3500	3800	3700	3500	3300	3300	3200	2300	0.137	0.127	480	444	0.264	924	55	2
16	15	0006P	87.432	87.687	0.255	6	6	P	3	Other Principal Arterial	MILLARD	4	3500	3500	3800	3700	3500	3300	3300	3200	2300	0.137	0.127	480	444	0.264	924	40	2
17	16	0006P	87.687	88.352	0.665	6	6	P	3	Other Principal Arterial	MILLARD	4	6100	6100	5900	5600	5400	5100	5000	5100	5100	0.134	0.138	817	842	0.272	1659	30	2
18	17	0006P	88.352	89.402	1.05	6	6	P	3	Other Principal Arterial	MILLARD	4	6100	6100	5900	5600	5400	5100	5000	5100	5100	0.134	0.138	817	842	0.272	1659	30	4
19	18	0006P	89.402	89.504	0.102	6	6	P	4	Minor Arterial	MILLARD	4	3200	3200	3100	3000	2800	2800	2700	2800	2800	0.13	0.15	416	480	0.28	896	30	3
20	19	0006P	89.504	89.676	0.172	6	6	P	4	Minor Arterial	MILLARD	4	3200	3200	3100	3000	2800	2800	2700	2800	2800	0.13	0.15	416	480	0.28	896	30	2
21	20	0006P	89.676	90.612	0.936	6	6	P	4	Minor Arterial	MILLARD	4	3200	3200	3100	3000	2800	2800	2700	2800	2800	0.13	0.15	416	480	0.28	896	45	2
22	21	0006P	90.612	93.846	3.234	6	6	P	4	Minor Arterial	MILLARD	4	3200	3200	3100	3000	2800	2800	2700	2800	2800	0.13	0.15	416	480	0.28	896	65	2
23	22	0006P	93.846	99.72	5.874	6	6	P	4	Minor Arterial	MILLARD	4	2600	2600	2800	2700	2500	2500	2400	2500	2500	0.123	0.174	320	452	0.297	772	65	2
24	23	0006P	99.72	104.408	4.688	6	6	P	4	Minor Arterial	MILLARD	4	1200	1200	1200	1200	1100	1100	1000	2100	1800	0.119	0.186	143	223	0.305	366	65	2
25	24	0006P	104.408	105.279	0.871	6	6	P	4	Minor Arterial	MILLARD	4	2200	2200	2200	2100	2000	1900	1900	1900	1900	0.112	0.21	246	462	0.322	708	55	2

Figure 4-3: Sample of CAMS Final Segment file.

4.3 Combining the Crash Data

The purpose of combining the crash files is to create a list of crashes that will eventually be assigned to segments of state highway for the CAMS-specific analysis. Four related files of crash data are combined by matching unique crash IDs using VBA macros. An outline of this process is shown in Figure 4-4. This section will describe the required user input, the VBA process, and the final crash file.

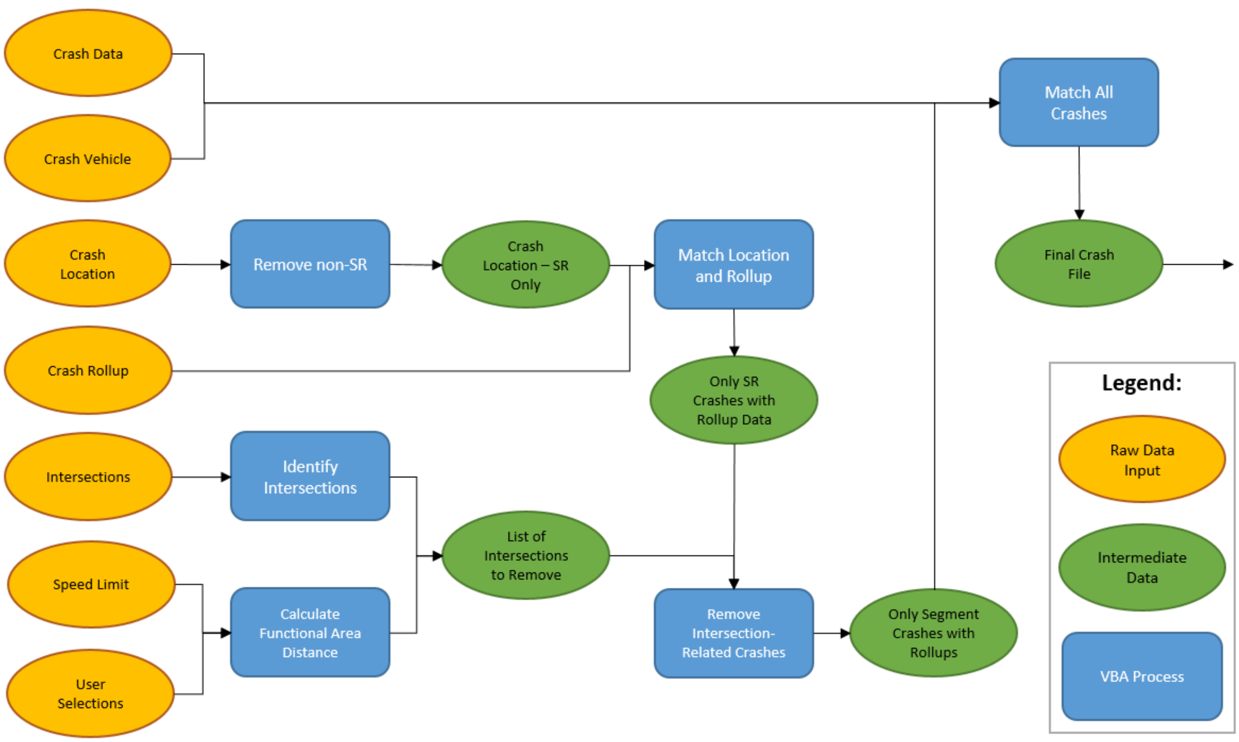


Figure 4-4: Outline of CAMS crash combination process.

<u>4.3.1 Required User Input</u>

To begin, the user must fill out an input form in the RGUI workbook as shown in Figure 4-5. As prompted by the input form, the user must browse for the files that are required by the analysis. File paths for the Crash Data file, the Crash Locations file, the Crash Rollup file, the Crash Vehicle file, and the Intersections file are all required. The user must also select the types of intersection-related crashes to be removed from the segment analysis. The user may select one, two, or all three of the following options: State Route-to-State Route-Intersection Crashes (SR to SR), State Route to Federal Aid Route-Intersection Crashes (SR to Fed Aid), and

Signalized State-Route Intersection Crashes (Signalized SR). Selecting all three options is recommended.

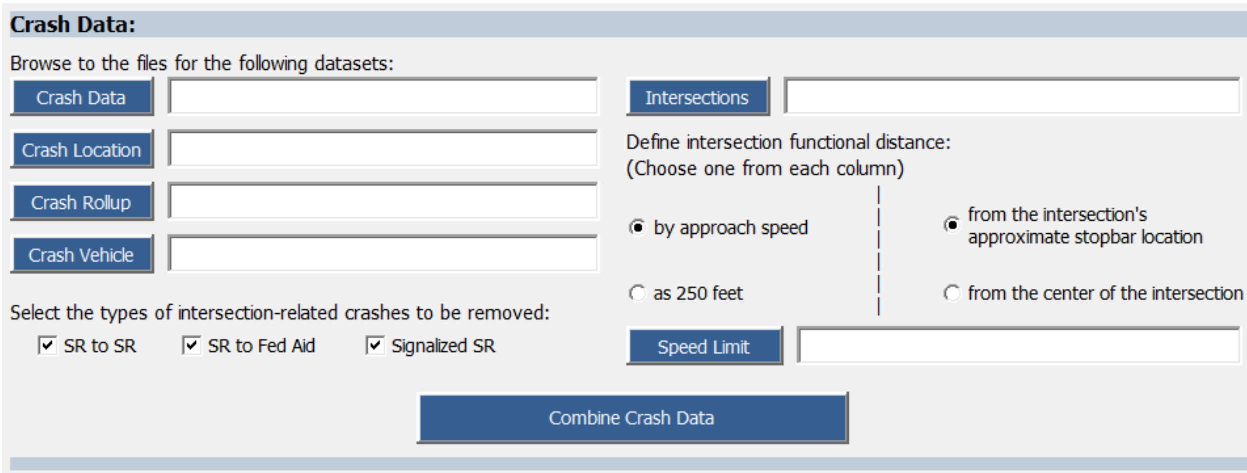


Figure 4-5: CAMS input form for combining crash data.

The user must also choose how to define the effective distance of each intersection. The total effective distance for each intersection is the sum of two parts: the functional area distance and the physical area distance.

4.3.1.1 Functional Area Distance

There are two ways to define the functional area distance in the CAMS, just as there are in the 2019 ISAM. The first option is to define the distance based on the approach speed limit on the main (state) route. These values vary from 195 feet for an approach speed ≤ 20 mph to 1320 feet for an approach speed ≥75 mph as shown in Table 4-1. These values are also used in the ISAM; for a discussion on the origins of Table 4-1, refer to Section 2.2. If the user chooses to define the functional area in this manner, the user will be prompted to browse for the Speed Limit file before beginning the analysis. The other option is to define the functional area distance as 250 feet, a practice referenced in the HSM (refer to Section 2.2 for further discussion on using 250 feet as an estimate of the functional area of an intersection).

4.3.1.2 Physical Area Distance

In addition to the functional area distance, the user is also given the option to add in the physical area of the intersection. After measuring a sample of the intersections in the CAMS dataset, it was determined that the average distance between the center of an intersection and its stop bars was 60 feet. The user can select whether they want to add this 60-foot physical area distance to the functional area distance. Both options (choosing to add or not add the 60 feet) are depicted in Figure 4-6. It is recommended that the user select the options “by approach speed” and “from the intersection’s approximate stop bar location” on the input form. Not only does this increase the likelihood of removing all intersection-related crashes from the segments, it also mirrors the analysis performed in the ISAM and removes the possibility of having overlap in the crashes analyzed by the two models.

Table 4-1: Functional Area Distance According to Approach Speed Limit

Approach Speed (mph)	*Functional Area Distance (ft)*
≤ 20	195
25	245
30	310
35	405
40	485
45	575
50	675
55	775
60	925
65	1045
70	1180
≥75	1320

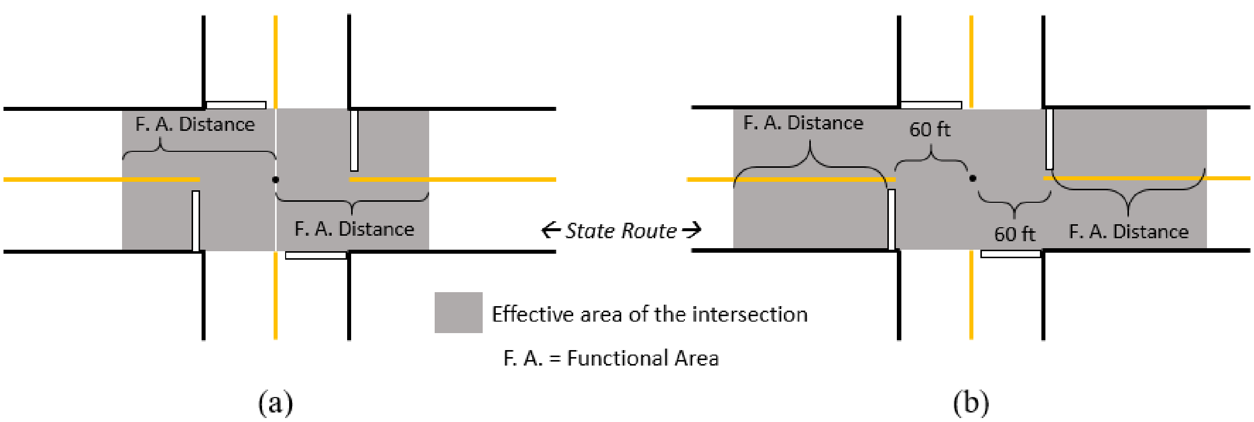


Figure 4-6: Effective area distance with and without considering the physical area distance.

Once all the required information has been entered on the input form and the user has clicked the “Combine Crash Data” button, the VBA macros open each data file, copy the contents, and paste it on a new sheet in the RGUI workbook. This begins the process of combining the data and making any necessary calculations.

4.3.2 VBA Process Description

The first step in the VBA process, “Identify Intersections” shown in Figure 4-4, is to identify and create a list of intersections. To eventually flag and remove all crashes pertaining to the selected intersection types, the VBA code searches for information about the location of each intersection and records that information in a list on a new Excel sheet. The contents of this list, represented by main route numbers and mileposts, depend on the types of intersections selected in the input form. If state route-to-state route-intersection crashes are to be removed from the analysis, then each intersection in the file marked as such is added to the list. If state route to federal aid route-intersection crashes are to be removed, then each intersection with at least one intersecting route with a number between 1000 and 9999 is added to the list. Finally, if signalized state-route intersection crashes are to be removed, then each intersection marked as signal-controlled is added to the list. In the case that multiple types of intersections are checked on the input form, the VBA code will add an intersection to the list if it meets at least one of the criteria for each checked intersection type. These criteria are summarized in Table 4-2.

Table 4-2: Filter Definitions for Intersection Types

Intersection type	*Definition*
SR-SR only	SR-SR column checked YES
SR-Federal Aid	The number of at least one route is between 1000 and 9999
SR-signal	The traffic control type is listed as “Signal”

*Note: “SR” here is used to abbreviate “state route.”

The second step in the VBA process, “Calculate Functional Area Distance” shown in Figure 4-4, is to determine the total effective distance of each listed intersection. The steps programmed in the VBA code are designed to calculate the functional area distance and the physical area distance using the data received from the input form. These values are recorded on the intersection list. The code is then used to sum up the two values on each line and it records this value which is the total effective distance for each listed intersection.

The following step in the VBA process, “Remove non-SR” shown in Figure 4-4, is to pare down the crash files to include only the crashes to be analyzed. First, the location data for crashes that are reported to have occurred on a state route are copied and pasted onto a new sheet in the workbook. This new sheet will serve as the home for the combined crash files and will ultimately contain the complete list of crashes to be analyzed in the CAMS model. In the next step of the VBA process, “Match Location and Rollup” shown in Figure 4-4, the crash rollups data are added to the new crash sheet by matching crash ID numbers between the rollup data and the location data. The combined crash data are then pared down further in the following step of the VBA process, “Remove Intersection-Related Crashes” shown in Figure 4-4, according to the list of intersections created previously. This is done by comparing information about each crash on the combined crash sheet to the different bounds on the intersections list. If the route and milepost of a crash lies within the bounds of the effective distance of any intersection and if that crash is marked as intersection related, it is deleted from the crash data sheet.

Once the crash list includes only the crashes to be analyzed in the CAMS model, the final VBA process, “Match All Crashes” shown in Figure 4-4, begins. The general crash data and the crash vehicle data are added to the combined crash data by matching the crash ID numbers. In the case that multiple vehicles were involved in the crash, only information relating to Vehicle 1 (typically the vehicle at fault) is kept to simplify the dataset. When all the crash data has been matched and added, the sheet with the combined crash information is exported and saved as a comma-separated values (CSV) file.

4.3.3 Final Crash File

The file created by the combine crash process contains information on every crash that is considered segment related in this analysis. Each line of the file lists a unique crash and its characteristics, including location, weather conditions, roadway conditions, number of people injured, severity of injuries, crash factors, event sequence for vehicle number 1, and manner of collision. A sample of this file is given in Figure 4-7.

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V
1	CRASH_ID	CRASH_DATETIME	ROUTE_ID	DIRECTION	LABEL	RAMP_ID	MILEPOINT	UTM_Y	UTM_X	NUMBER_VEHICLES_INVOLVED	NUMBER_FATALITIES	NUMBER_FOUR_INJURIES	NUMBER_THREE_INJURIES	NUMBER_TWO_INJURIES	NUMBER_ONE_INJURIES	PEDESTRIAN_INVOLVED	BICYCLIST_INVOLVED	MOTORCYCLE_INVOLVED	IMPROPER_RESTRAINT	UNRESTRAINED	DUI	AGGRESSIVE_DRIVING
2	10604097	1/1/2014 12:00	15	N	0015N		45.747775	4156238.273	305005.1853	1	0	0	0	0	5	N	N	N	N	N	N	N
3	10604111	1/1/2014 0:24	40	P	0040P		153.098	4471086	634252	1	0	0	0	0	1	N	N	N	N	N	Y	N
4	10604122	1/1/2014 13:42	180	P	0180P		0.9	4469339.961	433326.2446	2	0	0	0	3	1	N	N	N	N	N	N	N
5	10604159	1/1/2014 6:10	101	P	0101P		3.594	4610239.561	427628.4697	2	0	0	1	0	1	N	N	N	N	N	N	N
6	10604195	1/1/2014 2:04	71	P	0071P		7.08	4488327.983	426180.8065	2	1	0	0	0	0	N	N	N	N	Y	N	N
7	10604198	1/1/2014 16:40	15	N	0015N		329.82	4544974.221	419262.9778	1	0	0	0	0	3	N	N	N	N	N	Y	N
8	10604222	1/1/2014 10:02	15	N	0015N		264.829	4452909.477	442495.9102	1	0	0	0	0	2	N	N	N	N	N	N	N
9	10604267	1/2/2014 6:25	6	P	0006P		182.802	4431885	455002	1	0	0	0	0	2	N	N	N	N	N	N	N
10	10604312	1/1/2014 4:50	191	P	0191P		358.1	4487934	625177	1	0	0	0	0	1	N	N	N	N	N	N	N
11	10604349	1/2/2014 13:48	89	P	0089P		374.799	4506353	424964	2	0	0	0	0	2	N	N	N	N	Y	N	N
12	10604386	1/2/2014 15:15	126	P	0126P		1.596	4547160.207	417869.1723	2	0	0	0	0	8	N	N	N	N	N	N	N
13	10604445	1/2/2014 15:07	235	P	0235P		0.06	4568081	418773	2	0	0	0	0	2	N	N	N	N	N	N	N
14	10604452	1/2/2014 13:44	154	P	0154P		8.29	4488652	416651	2	0	0	0	0	2	N	N	N	N	N	N	N
15	10604453	1/2/2014 17:20	151	P	0151P		0.698	4490609	418424	2	0	0	0	0	3	N	N	N	N	N	N	N
16	10604489	1/1/2014 19:20	224	P	0224P		5.731	4500757	457192	2	0	0	0	0	2	N	N	N	N	N	N	N
17	10604505	1/2/2014 20:05	91	P	0091P		35.1	4633919	430964	1	0	0	0	0	2	N	N	N	N	N	N	N

Figure 4-7: Sample of CAMS Final Crash file.

4.4 Assigning Crashes to Segments

The purpose of combining the roadway file with the crash file is to assign crashes to the segments. Merging the crash information with the segment information is made possible using VBA macros. This process is outlined in Figure 4-8. The purple ovals represent the final two outputs of the data integration process as a whole: the CAMS Parameters file and the CAMS Input file. This section will describe the required user input, the VBA process, and the Input and Parameters files.

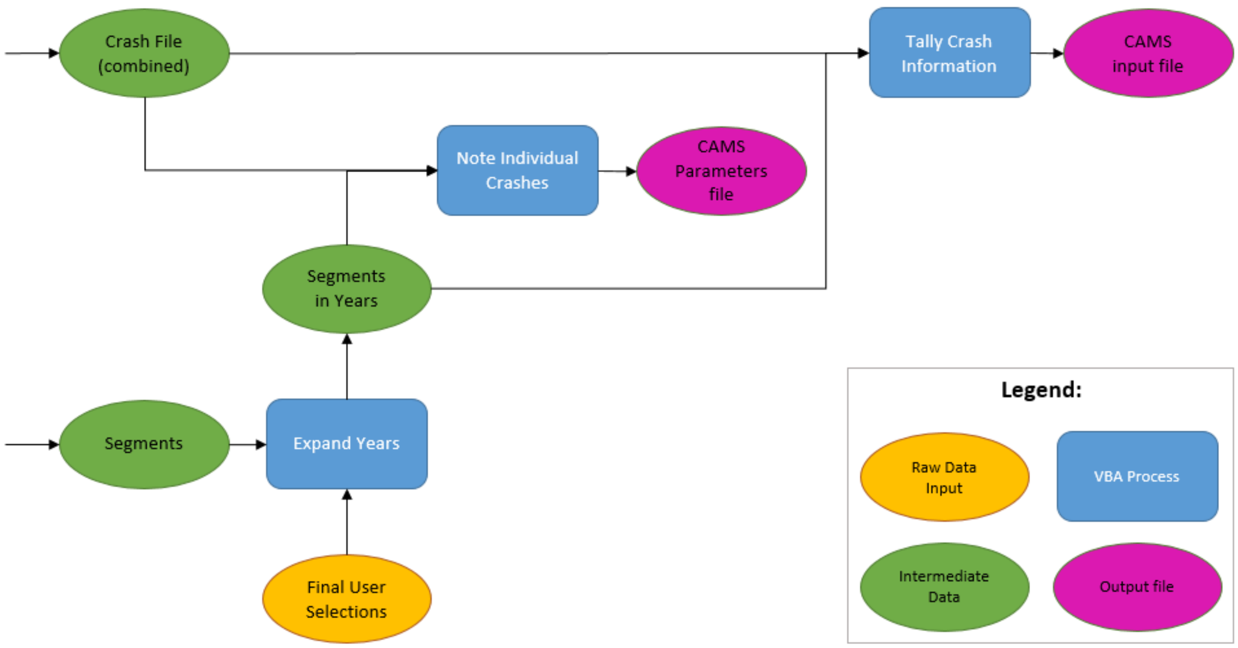


Figure 4-8: Outline of the process to create the CAMS Input file.

4.4.1 Required User Input

Before beginning the combination process, the user must first fill out a brief input form as shown in Figure 4-9. The two files required in this input form are the output files created in the combining crash data and combining roadway data processes (see Sections 4.2.3 and 4.3.3 for description of these files). After selecting the file paths for both the road segment file and the crash file, the user must select at least one of the five crash severities to be analyzed. The requested range of years of crash data to be included in the analysis is also required. This part of the RGUI is usually run with all the severities checked and with the data range including the five most recent years of available crash data (2014-2018 in this study). Once the user has completed the form, they can click the Create Input Data button and initiate the VBA code.

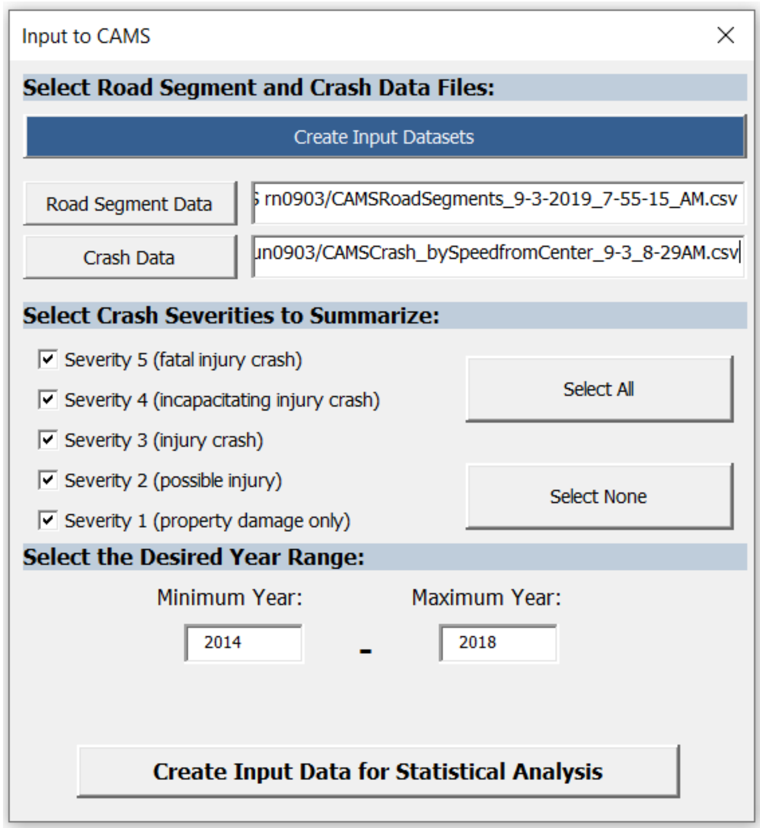


Figure 4-9: Input form for combining the CAMS roadway and crash files together.

4.4.2 VBA Process Description

Once the code has been initiated, the road segment data and the crash data are copied and pasted into the RGUI workbook. In the first step of the VBA process, "Expand Years" shown in Figure 4-8, each segment data line is copied and pasted once for every year the user selected to include in the analysis. For example, if the years 2014 through 2018 were selected, each line of segment data would be copied four times, making a total of five lines, one for 2014, one for 2015, etc., through 2018. All the data pertaining to roadway characteristics (with AADT as the only exception) remain the same for each year; the process assumes that the most recent data provided for the characteristics is accurate for all data years analyzed. Any changes in the physical characteristics of a segment are not accounted for in the data integration process but are instead noted in the produced reports as explained in Section 6.3.2.1.

In the next step of the VBA process, “Note Individual Crashes” shown in Figure 4-8, information about individual crashes is recorded on a separate sheet of the workbook. Each crash is matched to a segment by comparing the route and mileposts, and the matching segment ID number is appended to the crash data. This step is performed for all crashes that occurred on a state route within the data years analyzed (2014-2018 in this study).

In the third and final step of the VBA process, “Tally Crash Information” shown in Figure 4-8, tallies of crash information are appended to the roadway characteristics listed for each segment. This crash information, summed by segment, includes crash totals for each severity type as well as all the crash rollup fields.

4.4.3 Input and Parameters Files

Once all the segments and crashes have been matched, two files are created: the Input file and the Parameters file. The Input file is a CSV file that includes a record for each segment, its roadway characteristics, and its summed crash information. Each line of the file is for an individual year; if the time period requested included five years, then five lines of data would exist for each segment. At the conclusion of the data integration process, the Input file is used in the statistical model. The Parameters file, however, is not used in the statistical model but is used later in the CAMS process to create two-page technical reports for top-ranking segments. The Parameters file is saved as an Excel workbook and contains a list of all the crashes included in the analysis, with the matching segment ID number appended to each crash. Differences in the function of these two files are given in Table 4-3. Samples of the Input file and the Parameters file are given in Figure 4-10 and Figure 4-11, respectively.

Table 4-3: The Input File Compared to the Parameters File

	Input File	*Parameters File*
Contains segment characteristics	X	
Number of crashes per segment is determinable	X	X
Provides crash details		X
Used in the statistical model	X	

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	AA	AB	AC	AD	AE	AF	AG	AH
1	SEG_ID	LABEL	BEG_MILEPOINT	END_MILEPOINT	Seg_Length	YEAR	Route_Name	ROUTE_ID	DIRECTION	FC_Code	FC_Type	COUNTY	REGION	AADT	Single_Percent	Combo_Percent	Single_Count	Combo_Count	Total_Percent	Total_Count	SPEED_LIMIT	Num_Lanes	Urban_Rural	Urban_Ru_1	Total_Crashes	Severe_Crashes	Sum_Total_Crashes	Sev_5_Crashes	Sev_4_Crashes	Sev_3_Crashes	Sev_2_Crashes	Sev_1_Crashes	ADVERSE_ROADWAY_S	ADVERSE_WEATHER
2	1	0006P	0	10.88	10.88	2018	6	6	P	3	Other Principal Arterial	MILLARD	4	410	0.25	0.232	102	95	0.482	197	65	2	99999	Rurall	1	1	1	0	0	1	0	0	0	0
3	1	0006P	0	10.88	10.88	2017	6	6	P	3	Other Principal Arterial	MILLARD	4	410	0.25	0.232	102	95	0.482	197	65	2	99999	Rurall	0	0	0	0	0	0	0	0	0	0
4	1	0006P	0	10.88	10.88	2016	6	6	P	3	Other Principal Arterial	MILLARD	4	400	0.25	0.232	102	95	0.482	197	65	2	99999	Rurall	3	1	3	0	0	1	0	2	0	1
5	1	0006P	0	10.88	10.88	2015	6	6	P	3	Other Principal Arterial	MILLARD	4	380	0.25	0.232	102	95	0.482	197	65	2	99999	Rurall	1	0	1	0	0	0	0	1	0	0
6	1	0006P	0	10.88	10.88	2014	6	6	P	3	Other Principal Arterial	MILLARD	4	350	0.25	0.232	102	95	0.482	197	65	2	99999	Rurall	1	0	1	0	0	0	0	1	0	0
7	2	0006P	10.88	20.146	9.266	2018	6	6	P	3	Other Principal Arterial	MILLARD	4	410	0.25	0.232	102	95	0.482	197	40	2	99999	Rurall	3	1	3	0	1	0	0	2	0	0
8	2	0006P	10.88	20.146	9.266	2017	6	6	P	3	Other Principal Arterial	MILLARD	4	410	0.25	0.232	102	95	0.482	197	40	2	99999	Rurall	0	0	0	0	0	0	0	0	0	0
9	2	0006P	10.88	20.146	9.266	2016	6	6	P	3	Other Principal Arterial	MILLARD	4	400	0.25	0.232	102	95	0.482	197	40	2	99999	Rurall	1	1	1	0	0	1	0	0	1	1
10	2	0006P	10.88	20.146	9.266	2015	6	6	P	3	Other Principal Arterial	MILLARD	4	380	0.25	0.232	102	95	0.482	197	40	2	99999	Rurall	0	0	0	0	0	0	0	0	0	0
11	2	0006P	10.88	20.146	9.266	2014	6	6	P	3	Other Principal Arterial	MILLARD	4	350	0.25	0.232	102	95	0.482	197	40	2	99999	Rurall	1	0	1	0	0	0	0	1	0	0
12	3	0006P	20.146	24.499	4.353	2018	6	6	P	3	Other Principal Arterial	MILLARD	4	410	0.25	0.232	102	95	0.482	197	65	2	99999	Rurall	0	0	0	0	0	0	0	0	0	0
13	3	0006P	20.146	24.499	4.353	2017	6	6	P	3	Other Principal Arterial	MILLARD	4	410	0.25	0.232	102	95	0.482	197	65	2	99999	Rurall	0	0	0	0	0	0	0	0	0	0
14	3	0006P	20.146	24.499	4.353	2016	6	6	P	3	Other Principal Arterial	MILLARD	4	400	0.25	0.232	102	95	0.482	197	65	2	99999	Rurall	1	0	1	0	0	0	0	1	0	0
15	3	0006P	20.146	24.499	4.353	2015	6	6	P	3	Other Principal Arterial	MILLARD	4	380	0.25	0.232	102	95	0.482	197	65	2	99999	Rurall	0	0	0	0	0	0	0	0	0	0
16	3	0006P	20.146	24.499	4.353	2014	6	6	P	3	Other Principal Arterial	MILLARD	4	350	0.25	0.232	102	95	0.482	197	65	2	99999	Rurall	1	0	1	0	0	0	0	1	1	1
17	4	0006P	24.499	25.248	0.749	2018	6	6	P	3	Other Principal Arterial	MILLARD	4	410	0.25	0.232	102	95	0.482	197	50	2	99999	Rurall	0	0	0	0	0	0	0	0	0	0
18	4	0006P	24.499	25.248	0.749	2017	6	6	P	3	Other Principal Arterial	MILLARD	4	410	0.25	0.232	102	95	0.482	197	50	2	99999	Rurall	0	0	0	0	0	0	0	0	0	0
19	4	0006P	24.499	25.248	0.749	2016	6	6	P	3	Other Principal Arterial	MILLARD	4	400	0.25	0.232	102	95	0.482	197	50	2	99999	Rurall	0	0	0	0	0	0	0	0	0	0
20	4	0006P	24.499	25.248	0.749	2015	6	6	P	3	Other Principal Arterial	MILLARD	4	380	0.25	0.232	102	95	0.482	197	50	2	99999	Rurall	0	0	0	0	0	0	0	0	0	0
21	4	0006P	24.499	25.248	0.749	2014	6	6	P	3	Other Principal Arterial	MILLARD	4	350	0.25	0.232	102	95	0.482	197	50	2	99999	Rurall	0	0	0	0	0	0	0	0	0	0
22	5	0006P	25.248	27.1	1.852	2018	6	6	P	3	Other Principal Arterial	MILLARD	4	410	0.25	0.232	102	95	0.482	197	35	2	99999	Rurall	0	0	0	0	0	0	0	0	0	0
23	5	0006P	25.248	27.1	1.852	2017	6	6	P	3	Other Principal Arterial	MILLARD	4	410	0.25	0.232	102	95	0.482	197	35	2	99999	Rurall	0	0	0	0	0	0	0	0	0	0
24	5	0006P	25.248	27.1	1.852	2016	6	6	P	3	Other Principal Arterial	MILLARD	4	400	0.25	0.232	102	95	0.482	197	35	2	99999	Rurall	0	0	0	0	0	0	0	0	0	0
25	5	0006P	25.248	27.1	1.852	2015	6	6	P	3	Other Principal Arterial	MILLARD	4	380	0.25	0.232	102	95	0.482	197	35	2	99999	Rurall	0	0	0	0	0	0	0	0	0	0

Figure 4-10: Sample of CAMS Input file.

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	AA	AB	AC
1	Severities	12345																											
2	Functional Area Definition	bySpeedfromStopbar																											
3	Selected Years	2014-2018																											
4	SEG_ID	CRASH_ID	CRASH_DATETIME	ROUTE_ID	DIRECTION	LABEL	RAMP_ID	MILEPOINT	LATITUDE	LONGITUDE	NUMBER_VEHICLES_INVOL	NUMBER_FATALITIES	NUMBER_FOUR_INJURIES	NUMBER_THREE_INJURIES	NUMBER_TWO_INJURIES	NUMBER_ONE_INJURIES	PEDESTRIAN_INVOLVED	BICYCLIST_INVOLVED	MOTORCYCLE_INVOLVED	IMPROPER_RESTRAINT	UNRESTRAINED	DUI	AGGRESSIVE_DRIVING	DISTRACTED_DRIVING	DROWSY_DRIVING	SPEED_RELATED	INTERSECTION_RELATED	ADVERSE_WEATHER	ADVERSE_ROADWAY_SURF
5	1	10857129	5/20/16 14:13	6	P	0006P		0.1	39.05689726	-114.0469862	1	0	0	0	0	1	N	N	N	N	N	N	N	N	N	N	N	Y	N
6	1	10839235	4/1/16 8:45	6	P	0006P		1.6	39.05476617	-114.0192572	1	0	0	1	1	0	N	N	N	N	N	N	N	N	N	N	N	N	N
7	1	10794902	11/12/15 14:15	6	P	0006P		4.403	39.05082608	-113.9673296	1	0	0	0	0	1	N	N	N	N	N	N	N	N	N	N	N	N	N
8	1	10622203	3/19/14 2:58	6	P	0006P		5.001	39.04997308	-113.956272	1	0	0	0	0	1	N	N	N	N	N	N	N	N	N	N	N	N	N
9	1	11081181	7/17/18 15:04	6	P	0006P		5.9	39.04868936	-113.9396489	1	0	0	2	0	0	N	N	N	N	N	N	N	N	N	N	N	N	N
10	1	10822516	2/3/16 21:40	6	P	0006P		10.05	39.04275507	-113.8628704	2	0	0	0	0	1	N	N	N	N	N	N	N	N	N	N	N	N	N
11	2	11084040	8/4/18 12:05	6	P	0006P		16.4	39.06211924	-113.7511449	1	0	1	0	0	0	N	N	Y	N	N	N	N	N	N	N	N	N	N
12	2	11118702	11/22/18 14:08	6	P	0006P		18	39.07168035	-113.7241064	1	0	0	0	0	1	N	N	N	N	N	N	N	N	N	N	N	N	N
13	2	10823825	2/18/16 15:40	6	P	0006P		18.2	39.07289248	-113.7207293	1	0	0	1	0	0	N	N	N	N	Y	N	N	N	N	Y	N	Y	Y
14	2	11097438	9/20/18 7:09	6	P	0006P		19.395	39.07919324	-113.7002628	1	0	0	0	0	1	N	N	N	N	N	N	N	N	N	N	N	N	N
15	2	10647969	7/13/14 5:00	6	P	0006P		19.799	39.07799104	-113.692919	1	0	0	0	0	2	N	N	N	N	N	N	N	N	N	Y	N	N	N
16	3	10867439	6/28/16 11:00	6	P	0006P		21.2	39.07333997	-113.6676386	1	0	0	0	0	1	N	N	N	N	N	N	N	N	Y	N	N	N	N
17	3	10694632	12/30/14 4:00	6	P	0006P		22.7	39.07364302	-113.6405698	1	0	0	0	0	1	N	N	N	N	N	N	N	N	Y	Y	N	Y	Y

Figure 4-11: Sample of CAMS Parameters file.

4.5 Summary

A total of ten files of UDOT data, ranging from crash location information to the speed limit on any given mile of Utah state highway, are combined in the data integration process. The entire process requires three stages of user input: one for combining the roadway data, another for combining the crash data, and a third for assigning crashes to segments. These three stages constitute the bulk of the VBA code required for the CAMS analysis to be performed in its entirety. Two files are created at the end of the third stage: a file containing prepared segments for the statistical analysis and a file containing details about the crashes included in the analysis for use in the final analysis. Descriptions of the statistical model and final analysis are given in the following two chapters.

5.0 STATISTICAL ANALYSIS

5.1 Overview

The purpose of the statistical analysis is to identify high-priority segments. High-priority segments are those which observe many more crashes than are predicted. Identifying these segments is done by using the data to draw relationships between roadway characteristics and the number of injury crashes, creating distributions of predicted injury crash totals based on roadway characteristics, and flagging segments whose number of observed injury crashes is high compared to the predicted distribution.

Two statistical models were made for the CAMS. One is a model that identifies locations with more injury crashes than predicted; this is called the CAMS Prediction (CAMS-P) model. The other model identifies locations that have higher proportions of injury crashes than predicted based on the total number of crashes that occurred on the segment; this is called the CAMS Severity (CAMS-S) model. The CAMS-S model uses the word "Severity" in its name because it incorporates data from crashes of all severities, whereas the CAMS-P model uses data from injury crashes only. The user need only choose one model to run the CAMS, but each model brings its own strengths to produce meaningful results. This chapter briefly explains the two models and the output that these models produce.

5.2 The CAMS Prediction Model

The CAMS-P model is intended to be an updated version of the UCPM used in the RSAM (see Section 2.5.1.2 of this report as well as Schultz et al. 2016). The UCPM is a ZIP model (also known as a Poisson Mixture Model) and performed well for the RSAM. A similar ZIP model was also used for the original ISAM and the new ISAM (see Section 8.2.2). This research intended to use a ZIP model like the RSAM and ISAM, but two other models were also considered to provide a comparison in performance. The other models considered were a Zero-Inflated Negative Binomial (ZINB) model and a Negative Binomial Lindley (NBL) model. These models were chosen for their ability to analyze data with a high number of zeros, and each is discussed in detail in the BYU Statistics technical report titled "Justification for Considering

Zero-Inflated Models in Intersection Safety Analysis" (Pew 2020). The BYU Statistics report specifically talks about the models considered for the new ISAM, but similar conclusions can be drawn for the segment models as will be discussed in this section.

Performance measures can be used to compare the predictive accuracy and goodness of fit (GoF) for the three models. Root-Predicted Mean-Squared Error (RPMSE) and Median Absolute Deviation (MAD) are representations of predictive accuracy, where a lower number represents a more accurate model for both metrics. GoF is a representation of model fit, where a better-fitting model has a GoF close to 0.05. Of these two performance measure types, predictive accuracy is more important to this research because the purpose of the CAMS is to use predictions to identify potential hot spots. Table 5-1 shows the values for these three metrics for all three models. The ZIP model falls in the middle for predictive accuracy and third for GoF. It is important to note, however, that none of the values in the table are concerning; all three of the models have adequate accuracy and fit, and it would be appropriate to use any of the three models. The metrics indicate that although the performance of the ZIP model may not be distinctly superior to the other two models, it still adequately models the CAMS data. This reasoning justified the use of the ZIP model in the CAMS.

Table 5-1: Performance Measures for the Three Models

Model	*RPMSE*	*MAD*	*GoF*
ZIP	1.265	0.74	0.135
ZINB	1.259	0.74	0.057
NBL	1.271	0.74	0.077

The ZIP model as it has been implemented in the CAMS is a hierarchical Bayesian model. It uses four variables: speed limit, number of lanes, truck percentage of AADT, and natural log of Vehicle-Miles Traveled (VMT). It also uses urban code as an additional parameter to create a hierarchy; this allows for the effects of the variables to vary for different urban codes. The exploratory plots in Figure 5-1 show the relationships the variables in the model have with each other, especially with the number of injury crashes. From Figure 5-1 it can be determined that there is a weak negative correlation between total percent of trucks and injury crashes, weak positive correlations between speed limit and injury crashes as well as between number of lanes and injury crashes, and a moderate positive correlation between the log of VMT and injury crashes.

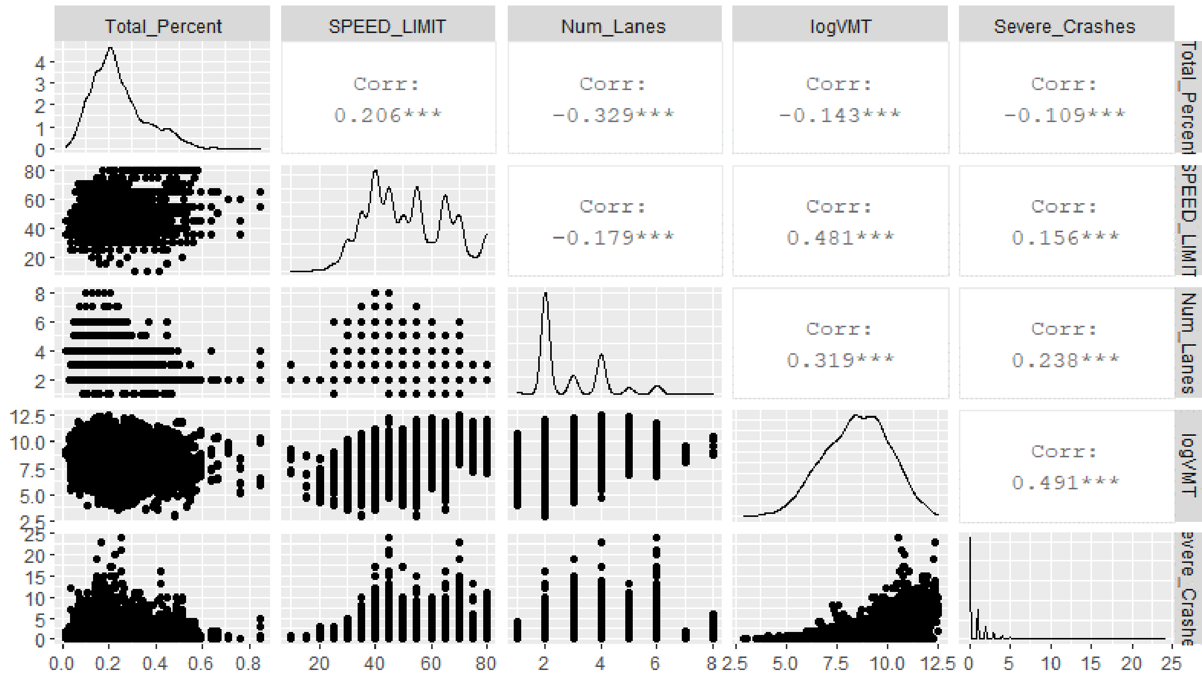


Figure 5-1: Exploratory plots for the CAMS-P (ZIP) model showing variable relationships (Pew 2020).

The probability mass function of the ZIP model with its regression equation is given in Equation 5-1. In this equation, the variable *i* represents each segment, and the variable *j* represents each year. In addition, *Y* represents the number of injury crashes, and π represents the additional probability of observing zero beyond what a Poisson model would typically assume. Further explanation of these two equations can be found in the BYU Statistics report titled "Justification for Considering Zero-Inflated Models in Intersection Safety Analysis" (Pew 2020). Although the BYU Statistics report is written specifically about intersection models, the ZIP model as applied by the BYU Statistics team functions the same way for segments as it does for intersections.

$$P\left(Y_{ij} = y_{ij} \middle| \pi, \lambda_{ij}\right) = \begin{cases} \pi + (1-\pi)e^{\lambda_{ij}} & y_{ij}=0 \\ (1-\pi)\frac{e^{\lambda_{ij}}\lambda_{ij}^{y_{ij}}}{y_{ij}!} & y_{ij}=1,2,\ldots \end{cases}$$
$$\ln(\lambda_{ij}) = \beta_0 + x'_{ij}\beta_k + \eta_{ij} \tag{5-1}$$

The primary input to the model is the Input file created in the data integration process and described in Section 4.4.3. The model is built with four years of recent crash data (2014-2017 in this study), and these years of data are used to create predicted distributions of injury crashes for each segment. Values of actual (observed) crash data from a fifth and most recent year of data (2018 in this study) are compared to the distributions created by the statistical model and assigned a percentile value. Once all segments have been associated with a percentile value, the segments are ranked in order from highest percentile value to lowest.

It is important to note that although the predicted distributions are built with four years of crash data, only one year of observed crash data is used to rank each segment. In the case that a segment had an unusually high number of crashes in the one year being used to assign a percentile value, that segment would end up being ranked significantly higher in the state than it might have been had any other year of crash data been used. The opposite is true for a segment that experienced an unusually low number of crashes in that year. A brief discussion on possible future research to avoid these situations by altering the statistical analysis is given in Section 9.3.1.

A visual representation of the predicted crash distributions created by the statistical model is given in Figure 5-2. In this example, the observed number of injury crashes on a segment of I-15 in Utah County (the vertical dashed line) is plotted against the predicted distribution of crashes for that particular segment. The percentile for this segment is 0.9925, indicating that the observed number of injury crashes (which was 4 as is indicated by the dotted line) is significantly higher than predicted. The high percentile value for this segment resulted in it being ranked 21st out of 3,882 total segments.

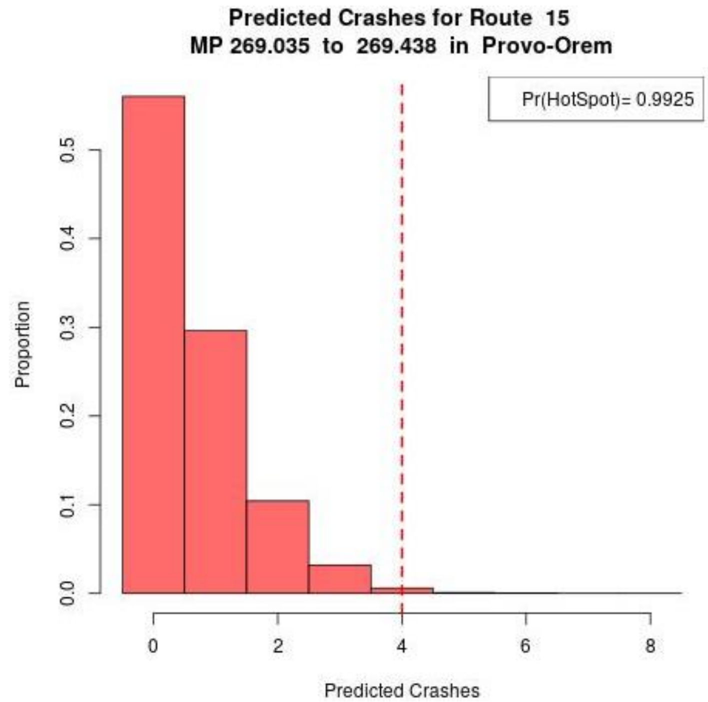


Figure 5-2: Predicted distribution of crashes for a segment in the CAMS-P model.

5.3 The CAMS Severity Model

The CAMS-S model is intended to be an updated version of the UCSM used in the RSAM (see Section 2.5.1.2 of this report as well as Schultz et al. 2016). The CAMS-S model was created to identify segments that may not necessarily have an unusually large number of crashes, but that have an unusually high proportion of injury crashes. In other words, the model answers the question, "If a crash was to occur on any segment, which segments are most likely to experience an injury crash?"

The CAMS-S model shares several similarities with the CAMS-P model, including the variables that it uses. The variables in the CAMS-S model are speed limit, number of lanes, percent trucks, and VMT (no log transformation used). The relationship of these variables is shown in the exploratory plots given in Figure 5-3. The plots indicate that, similar to the CAMS-P model, there is a weak negative correlation between total percent of trucks and injury crashes, weak positive correlations between speed limit and injury crashes as well as between number of lanes and injury crashes, and a moderate positive correlation between VMT and injury crashes.

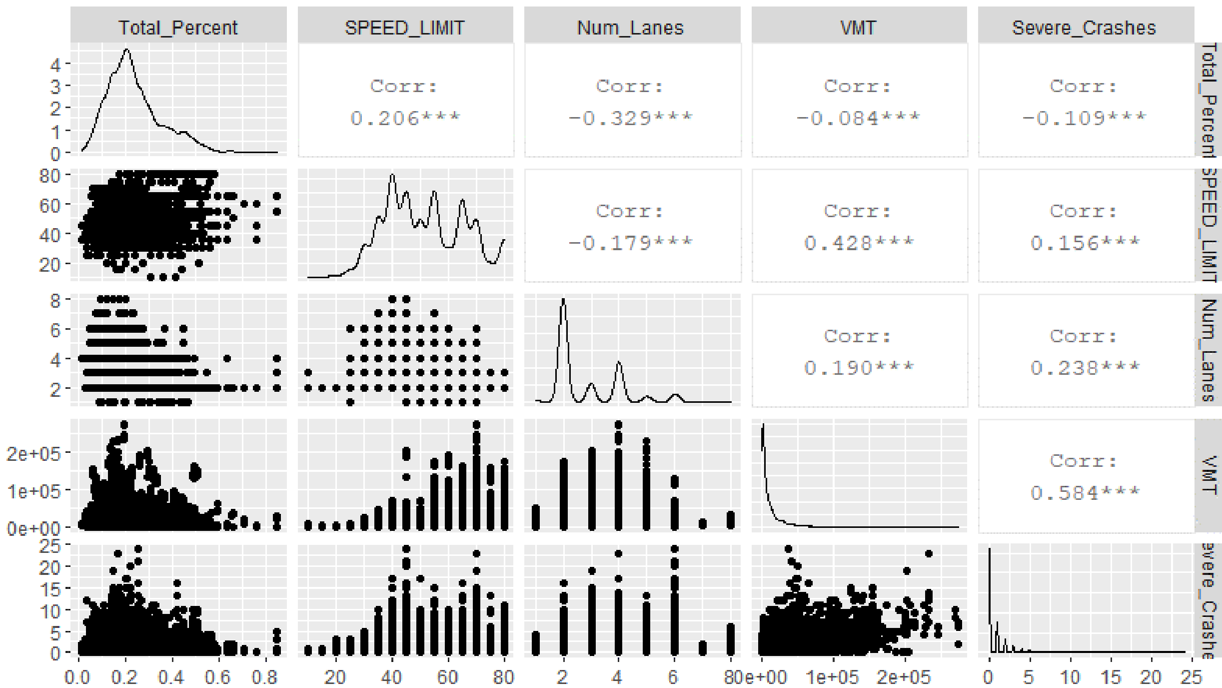


Figure 5-3: Exploratory plots for the CAMS-S model showing variable relationships (Pew 2020).

The probability mass function for the CAMS-S model is given in Equation 5-2. In this equation, all variables are as previously defined for Equation 5-1, with the exception that in this model π represents the probability of a crash being injury-causing. This model also includes n which represents the total number of crashes. The performance measures (RPMSE and MAD) for this model are given in Table 5-2 and indicate that the model provides adequate prediction accuracy. The GoF test does not work well with this model, so it is not included in the table.

$$P\left(Y_{ij} = y_{ij} \middle| \pi_{ij}\right) = \binom{n_{ij}}{y_{ij}} \pi_{ij}^{y_{ij}} (1-\pi_{ij})^{n_{ij}-y_{ij}} \qquad y_{ij}=0,1,\ldots,n_{ij}$$
$$\log\left(\frac{\pi_{ij}}{1-\pi_{ij}}\right) = \beta_{0k} + x'_{ij}\beta \qquad (5\text{-}2)$$

Table 5-2: Performance Measures for the CAMS-S Model

RPMSE	*MAD*
1.359	0.888

The segment ranking system for the CAMS-S model is similar to that of the CAMS-P model. The CAMS-S model analyzes the relationships between roadway characteristics and the ratio of injury crashes to total crashes on each segment using four years (2014-2017 in this study) of data as provided in the Input file. The model predicts how many injury crashes would occur on each segment in a fifth year (2018 in this study) based on the number of total crashes observed on the segment during that same year. These predictions are given as distributions, and each segment is associated with a percentile value based on where the number of observed injury crashes falls in the predicted distribution.

5.4 Output from Statistical Analysis

The output of whichever model is used (the CAMS-P model or the CAMS-S model) is termed the Results file. It is exported as a CSV file and looks exactly like the Input file with the addition of a few columns. These new columns include the following information about each segment: the mean number of predicted crashes for the most recent year of data (2018 in this study), percentile value, rank in state, rank in region, and rank in county. A sample of the Results file from the CAMS-P model is given in Figure 5-4. In addition, the 20 highest-ranking segments from the CAMS-P model and CAMS-S model are given later in this report, in Table 7-1 and Table 7-2, respectively.

5.5 Summary

Two statistical models were applied to the CAMS data: the CAMS-P model and the CAMS-S model. The CAMS-P model is a ZIP model that identifies segments with unusually high numbers of injury crashes in a selected year (2018 in this study) whereas the CAMS-S model identifies segments with unusually high proportions of injury crashes versus total observed crashes in a selected year (2018 in this study). Both models produce a CSV file of their results which is used to create brief, automated technical reports as will be discussed in the following chapter. Results of the two models and their implications are discussed further in Chapter 7.

		SEG_ID	LABEL	BEG_MILEPOINT	END_MILEPOINT	Seg_Length	YEAR	Route_Name	ROUTE_ID	DIRECTION	FC_Code	FC_Type	COUNTY	REGION	AADT	Single_Percent	Combo_Percent	Single_Count
2	5601	1121	0039P	42.67	44.13	1.46	2018	39	39	P	5	Major Collector	WEBER	1	510	0.17	0.11	85
3	6866	1374	0060P	0	0.419	0.419	2018	60	60	P	5	Major Collector	WEBER	1	15000	0.1	0.03	1425
4	14791	2959	0150P	33.29	46.48	13.19	2018	150	150	P	5	Major Collector	SUMMIT	2	200	0.13	0.12	26
5	15711	3143	0171P	13.58	14.09	0.504	2018	171	171	P	3	Other Principal /	SALT LAKE	2	20000	0.06	0.01	1140
6	10141	2029	0089P	103.9	114.3	10.45	2018	89	89	P	3	Other Principal /	KANE	4	1800	0.11	0.32	205
7	7501	1501	0068P	64.37	65.75	1.381	2018	68	68	P	3	Other Principal /	DAVIS	1	10000	0.11	0.11	1100
8	7041	1409	0065P	8.441	14.18	5.734	2018	65	65	P	5	Major Collector	MORGAN	1	420	0.15	0.06	61
9	16306	3262	0189P	25.28	26.54	1.263	2018	189	189	P	3	Other Principal /	WASATCH	3	8300	0.12	0.1	1013
10	2136	428	0015N	296	297.1	1.086	2018	15	15	N	1	Interstate	SALT LAKE	2	130500	0.12	0.06	16052
11	2761	553	0015P	134.9	138.1	3.198	2018	15	15	P	1	Interstate	MILLARD	4	7500	0.15	0.24	1118
12	15851	3171	0173P	4.887	5.031	0.144	2018	173	173	P	3	Other Principal /	SALT LAKE	2	40000	0.11	0.09	4280
13	1671	335	0015N	31.86	37.48	5.614	2018	15	15	N	1	Interstate	WASHINGT	4	13500	0.14	0.19	1877
14	2001	401	0015N	269.4	269.9	0.434	2018	15	15	N	1	Interstate	UTAH	3	95000	0.08	0.07	7695
15	616	124	0008P	0	1.18	1.18	2018	8	8	P	3	Other Principal /	WASHINGT	4	27000	0.14	0.02	3834
16	14961	2993	0154P	6.759	8.274	1.515	2018	154	154	P	3	Other Principal /	SALT LAKE	2	44000	0.11	0.07	4928
17	11676	2336	0091P	26.24	26.89	0.644	2018	91	91	P	3	Other Principal /	CACHE	1	34000	0.05	0.08	1802
18	3376	676	0015P	335.8	336.3	0.495	2018	15	15	P	1	Interstate	DAVIS	1	62500	0.14	0.09	8875
19	17776	3556	0210P	9.855	9.956	0.101	2018	210	210	P	4	Minor Arterial	SALT LAKE	2	6600	0.03	0.01	185
20	1551	311	0014P	25.71	41	15.28	2018	14	14	P	4	Minor Arterial	KANE	4	1000	0.15	0.05	153
21	5116	1024	0035P	29.73	31.51	1.773	2018	35	35	P	5	Major Collector	DUCHESNE	3	390	0.12	0.07	47
22	2981	597	0015P	269	269.4	0.403	2018	15	15	P	1	Interstate	UTAH	3	95000	0.08	0.07	7695
23	8571	1715	0079P	3.298	3.427	0.129	2018	79	79	P	3	Other Principal /	WEBER	1	9900	0.06	0.17	604
24	13556	2712	0126P	0.624	1.7	1.076	2018	126	126	P	3	Other Principal /	DAVIS	1	22000	0.07	0.04	1562
25	8486	1698	0077P	7.456	8.539	1.083	2018	77	77	P	3	Other Principal /	UTAH	3	34000	0.19	0.06	6494
26	14361	2873	0138P	20.3	20.44	0.147	2018	138	138	P	4	Minor Arterial	TOOELE	2	10000	0.09	0.13	850
27	326	66	0006P	204.2	211.2	6.974	2018	6	6	P	3	Other Principal /	UTAH	3	9200	0.13	0.29	1168
28	7596	1520	0070N	25.76	31.69	5.928	2018	70	70	N	1	Interstate	SEVIER	4	2450	0.11	0.26	274

	Combo_Count	Total_Percent	Total_Count	SPEED_LIMIT	Num_Lanes	Urban_Rural	Urban_Ru_1	Total_Crashes	Severe_Crashes	Sum_Total_Crashes	Sev_5_Crashes	Sev_4_Crashes	Sev_3_Crashes	Sev_2_Crashes	Sev_1_Crashes	ADVERSE_ROADWAY_SURF	WORK_ZONE_WORKER_PRI	VMT	logVMT	hier	Crashes18	Predicted_Total	Percentile	State_Rank	Region_Rank	County_Rank
2	56	0.28	141	45	2	99999	Rurall	6	4	6	0	3	1	0	2	1	1	744.6	6.613	4	4	0.0716	1	1	1	1
3	510	0.13	1935	40	3	64945	Ogden -	35	11	35	0	1	10	3	21	4	1	6285	8.746	2	7	0.9062	0.9998	2	2	2
4	24	0.25	50	55	2	99999	Rurall	12	9	12	0	5	4	0	3	1	0	2638	7.878	4	5	0.42	0.9996	3	1	1
5	200	0.07	1340	35	4	78499	Salt Lak	55	10	55	0	1	9	7	38	10	2	10080	9.218	5	7	0.8651	0.9994	4	2	1
6	571	0.43	776	65	2	99999	Rurall	164	15	164	0	7	8	7	142	18	1	18814	9.842	4	8	1.4368	0.9992	5	1	1
7	1130	0.22	2230	50	4	64945	Ogden -	81	6	81	0	0	6	26	49	18	0	13810	9.533	2	5	0.5819	0.999	6	3	1
8	24	0.2	85	35	2	99999	Rurall	27	16	27	2	5	9	5	6	1	1	2408	7.787	4	6	0.9862	0.9986	7	4	1
9	863	0.23	1876	60	4	99999	Rurall	48	5	48	0	2	3	2	41	3	0	10483	9.258	4	4	0.3817	0.9985	8	1	1
10	7438	0.18	23490	70	5	78499	Salt Lak	357	27	357	1	1	25	63	267	31	4	1E+05	11.86	5	13	3.865	0.9984	9	3	2
11	1815	0.39	2933	80	2	99999	Rurall	77	17	77	0	1	16	9	51	38	0	23985	10.09	4	8	1.7887	0.9983	10	2	1
12	3560	0.2	7840	45	6	78499	Salt Lak	105	8	105	0	2	6	22	75	14	14	5760	8.659	5	5	0.6942	0.9976	11	4	3
13	2565	0.33	4442	80	3	99999	Rurall	138	26	138	2	2	22	15	97	58	0	75789	11.24	4	11	3.6434	0.9972	12	3	1
14	6650	0.15	14345	70	6	72559	Provo-C	55	9	55	0	3	6	7	39	35	0	41230	10.63	3	6	0.9402	0.9972	13	2	1
15	567	0.16	4401	40	4	77446	St. Geor	134	18	134	1	2	15	41	75	3	2	31860	10.37	7	9	2.3743	0.9971	14	4	2
16	3080	0.18	8008	60	6	78499	Salt Lak	135	21	135	1	2	18	19	95	27	19	66660	11.11	5	10	3.0274	0.9971	15	5	4
17	2720	0.13	4522	35	4	50959	Logan	193	7	193	0	0	7	34	152	26	0	21896	9.994	1	5	0.7914	0.9959	16	5	1
18	5688	0.23	14563	70	3	64945	Ogden -	95	15	95	1	4	10	11	69	31	5	30938	10.34	2	7	1.7049	0.9953	17	6	2
19	59	0.04	244	40	2	99999	Rurall	3	2	3	0	1	1	0	1	1	0	666.6	6.502	4	2	0.0932	0.9946	18	6	5
20	52	0.21	205	50	2	99999	Rurall	92	15	92	2	5	8	8	69	24	0	15284	9.635	4	7	1.8129	0.9943	19	5	2
21	28	0.19	75	45	2	99999	Rurall	6	3	6	0	1	2	2	1	0	0	691.5	6.539	4	2	0.1049	0.9943	20	3	1
22	6650	0.15	14345	70	4	72559	Provo-C	34	5	34	0	0	5	6	23	15	0	38285	10.55	3	4	0.6039	0.9923	21	4	2
23	1683	0.23	2287	55	5	64945	Ogden -	11	2	11	0	0	2	0	9	2	0	1277	7.152	2	2	0.1109	0.9921	22	7	3
24	880	0.11	2442	40	4	64945	Ogden -	169	22	169	2	2	18	42	105	19	0	23672	10.07	2	9	2.9932	0.9914	23	8	3
25	2040	0.25	8534	45	4	72559	Provo-C	172	7	172	0	3	4	40	125	51	1	36822	10.51	3	5	0.9469	0.9912	24	5	3
26	1250	0.21	2100	40	3	99998	Small U	14	2	14	0	1	1	2	10	4	0	1470	7.293	6	2	0.137	0.9904	25	7	1
27	2705	0.42	3873	65	2	99999	Rurall	247	48	247	2	15	31	40	159	150	1	64161	11.07	4	16	7.4196	0.9896	26	6	4
28	632	0.37	906	80	2	99999	Rurall	15	7	15	0	3	4	0	8	6	0	14524	9.584	4	4	0.7015	0.9896	27	6	1

Figure 5-4: Sample of CAMS Results file.

6.0 REPORT CREATION AND FINAL ANALYSIS

6.1 Overview

The purpose of the report creation portion of CAMS is to help bridge the gap between knowledge and action. Additional VBA macros combine two previously created files—the Parameters file described in Section 4.4.3 and the Results file described in Section 5.4—to create reports that guide UDOT engineers in finding solutions to high-priority segments. This chapter will first describe the Report Compiler which contains the VBA macro process and then discuss the produced two-page technical reports.

6.2 The Report Compiler

The Report Compiler file is an Excel Macro-Enabled workbook containing VBA code to present the results of the statistical model and to create reports for the high-priority segments. The home tab of the workbook shown in Figure 6-1 tells the reader the functional purpose of the file and extends a few warnings about the sensitivity of the code.

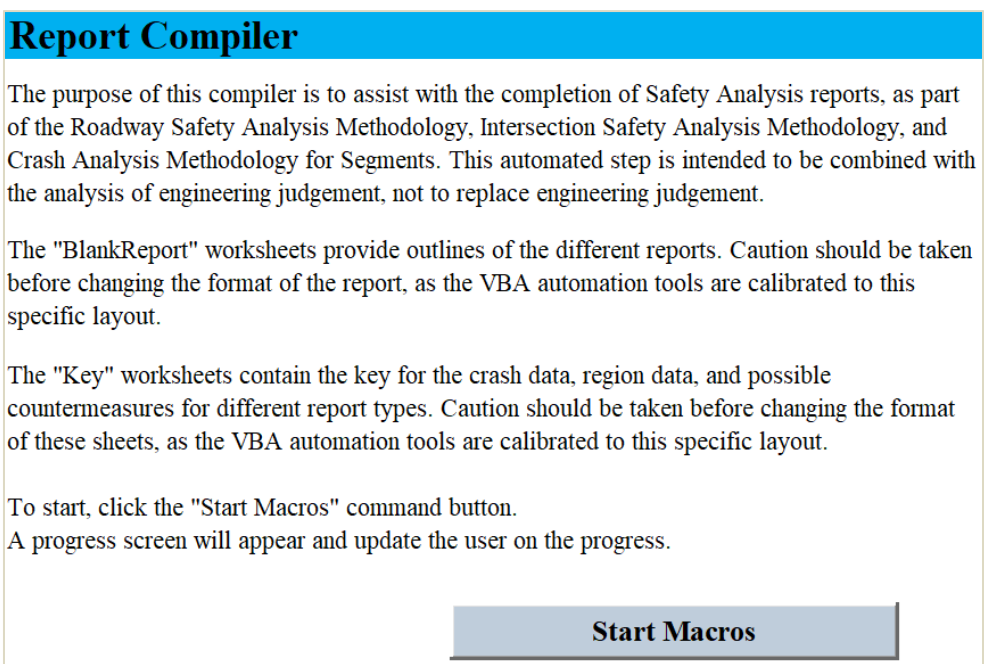


Figure 6-1: Home tab of the Report Compiler.

Upon clicking the “Start Macros” button in the Report Compiler as shown in Figure 6-1 and then selecting “Segment Reports (CAMS; 2020)” in the resulting input form shown in Figure 6-2, the user will be prompted to enter the name of the statistical model used (for referencing purposes only) as shown in Figure 6-3. Once the user has entered and confirmed the model name, the Report Compiler will instruct the user to browse for the Results file and then for the Parameters file. The VBA code is then programmed to load the two files and then subsequently prompt the user to select how many and what scope of reports to create as shown in Figure 6-4. Options include creating reports for a specified number of segments in the state, each UDOT Region, or each county.

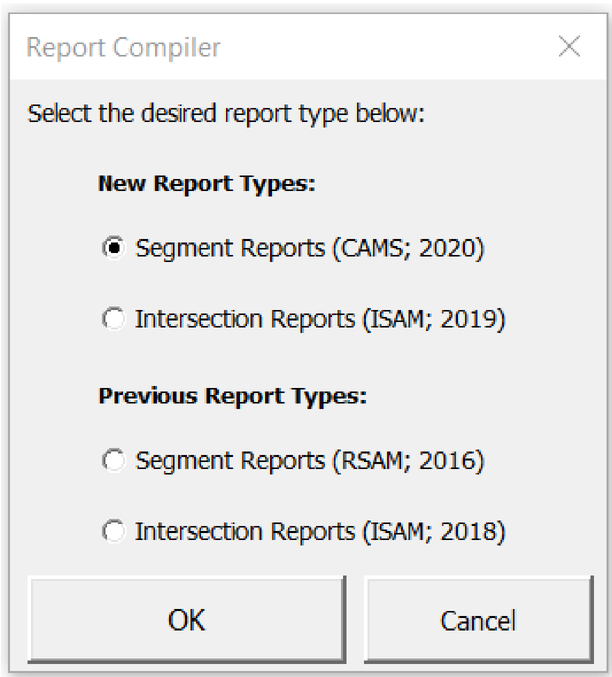


Figure 6-2: Report type input form.

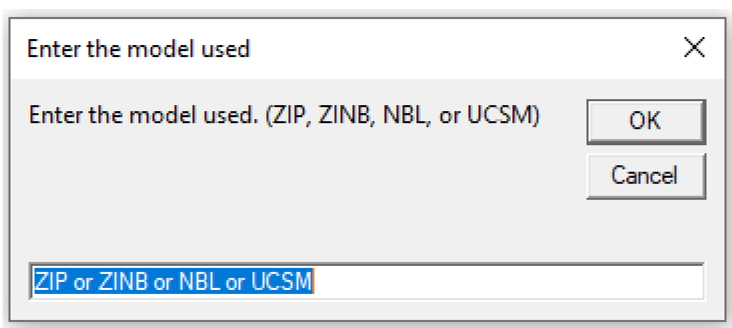


Figure 6-3: Prompt for CAMS model name.

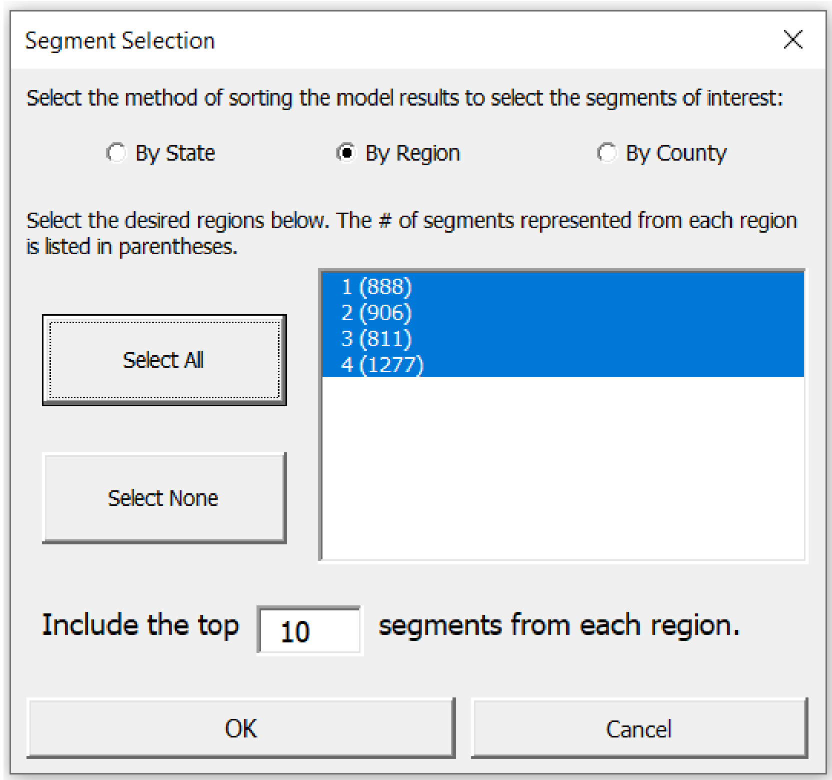


Figure 6-4: CAMS input form for selection of number of reports.

Upon confirming their selection, the user need only wait a few minutes for the reports to be created. The VBA code is programmed to sort through the Results and Parameters files to find information pertaining to the selected segments. Tables pertaining to the roadway characteristics as well as the crash data associated with each segment are subsequently generated.

6.3 The Two-Page Technical Reports

The purpose of creating the two-page technical reports is to provide UDOT engineers with initial summaries of the high-priority segments. The reports are titled Segment Safety Analysis Reports (SSARs). Research analysts review the tabular information in the reports and provide additional notes on the identified segments. Such analysis is purely preliminary, and it is understood that UDOT engineers will dig deeper into the potential issues of the segments themselves using the SSARs as a starting point. This section will describe page one and page two of the SSARs. Images of the blank report are provided in this section in Figure 6-5 and Figure

6-6, and a sample completed SSAR is provided at the end of this chapter in Figure 6-7 and Figure 6-8.

6.3.1 Page One of the Technical Reports

Figure 6-5 shows the first page of the blank report as saved in the Report Compiler workbook. It is meant to display the crash data and roadway characteristics that were compiled in the data integration process and used in the statistical analysis. The purposes of each of the seven tables on this page of the report are explained in the following subsections.

6.3.1.1 Roadway Data Tables

The first three tables of the SSAR display information about the roadway data. Table 1 of the SSAR is the Segment Metadata table. It includes basic information used to identify the segment and its level of priority. The state, region, and county ranks are given in addition to the route number, direction, and milepost limits. Table 2 of the SSAR is the Segment Characteristics table. It displays four of the five variables used in the homogeneous segmentation process for this segment, namely AADT, functional class, number of through lanes, and posted speed limit. Table 3 of the SSAR is the Roadway Characteristics table. It contains information about the roadway design. Portions with a median, shoulders, curves, differing lane widths, barriers, or rumble strips are noted. This table is originally blank, and the research analysts fill out the information manually while they complete virtual site visits as described in Section 6.3.2.1.

6.3.1.2 Crash Data Tables

Tables 4 through 7 of the SSAR contain information about the crashes that occurred on the segment. Table 4 of the SSAR is the Crash Count and Severity table. It lists out how many crashes of each type occurred during the analysis period. For reference, it names the crash severities considered in the statistical model (e.g., “3, 4, 5” for injury crashes) and the method by which the functional area was defined (e.g., “by speed from stop bar”). The table also displays the number of crashes that occurred in the most recent year analyzed (2018 in this study) as well as the number of crashes predicted to occur on that segment by the statistical model. It also details the number of crashes per severity type that occurred on the segment during the entire analysis period (2014-2018 in this study).

Segment Safety Analysis Report

Introduction

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

Segment Identification and Roadway Characteristics

Date:

Street Name:

Table 1: Segment Metadata

Route Number:		UC Model Used:	
Road Direction:		State Rank:	
Beginning, Ending MP:		Region, Rank:	
Length (miles):		County, Rank:	
Data Source Years:			

Table 2: Segment Characteristics

Functional Class:		AADT:	
Number of Thru Lanes:		Speed Limit (MPH):	

Table 3: Roadway Characteristics

MPs	Median	Shoulder	Grade	Curve	Lanes	Wall/ Barrier	Rumble Strips

Micro-Analysis of Crash Data

Crash Data Summary

Table 4: Crash Count and Severity

Crash Severities	Functional Area Method	Crashes During		Total Crashes Between				
		Predicted	Actual	Sev. 5	Sev. 4	Sev. 3	Sev. 2	Sev. 1

Table 5: Top 7 Crash Factors

Crash ID	Latitude	Longitude	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5	Factor 6	Factor 7
Injury Total									
Segment Total									

Table 6: Manner of Collision Data

	Manner 1	Manner 2	Manner 3	Manner 4	Manner 5	Manner 6	Manner 7	Manner 8	Manner 9
Name									
Injury Total									
Segment Total									

Table 7: Data from Crash and Vehicle Datasets

Crash ID	# of Vehicles	First Harmful Event	Manner of Collision	Event Sequence (1)	Event Sequence (2)	Event Sequence (3)	Event Sequence (4)	Most Harmful Event	Vehicle Maneuver

Figure 6-5: Page one of the blank SSAR.

Table 5 of the SSAR is the Top 7 Crash Factors table. It lists in descending order the seven most common crash factors for injury crashes on this segment as reported by the investigating police officer at the scene of the crash. Totals are given as fractions of the injury crashes and of the total crashes. Both sums are given because they can provide different insights into crash patterns on the identified segment. The factors that pertain to just the injury crashes are insightful because it is the injury crashes that are used in the statistical model to identify which segments are high priority. However, the most common factors for any severity can also reveal crash patterns. A list of all the possible crash factors found in the Crash Rollup file is given later in this report in Table 7-3.

Table 6 of the SSAR is the Manner of Collision Data table. It lists in descending order the nine most common manners of collision for injury crashes on the segment. Totals are given as fractions of the injury crashes and of the total crashes. A list of all the possible manners of collision is given later in this report in Table 7-4.

Table 7 of the SSAR is the Data from Crash and Vehicle Datasets table. It contains information about the event sequence of each vehicle involved in an injury crash on the segment. It is there to help analysts recognize patterns in the crashes and encourage them to think critically about potential countermeasures. Once the research analyst has completed page two of the SSAR (discussed in the next section), they remove Table 7 before presenting the final reports to UDOT. Doing so keeps the SSARs brief since the content in Table 7 can be very lengthy.

6.3.2 Page Two of the Technical Reports

The second page of the SSAR hosts less quantitative data and more qualitative information. Figure 6-6 shows the page still blank with the exception of an unfiltered list of potential countermeasures that will appear at the bottom of the page after running the Report Compiler. None of the information on this page comes from the data integration process or statistical model. The following subsections describe the relevance and nature of the sections on this page.

Historical Perspective, Current Conditions, Site Visit Notes *Date:*

[Historical Perspective, document changes to the roadway in the recent years]

[Site Visit Notes, document observations at the site, date, time, etc.]

[Insert 2 photos here:
Street view of something mentioned above
& GIS Location map]

Figure 1:

Figure 2: GIS map showing the location of the segment (ESRI).

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

Engineering Countermeasures

Policy Countermeasures

[begin list here]

Figure 6-6: Page two of the blank SSAR.

6.3.2.1 Historical Perspective, Current Conditions, Site Visit Notes

When the research analysts receive the automated reports, they review the given information presented in the tables to gather an idea of the potential safety problem present at the site. Analysts then use tools such as Mandli's Roadview Explorer (Mandli Communications 2020), Google Maps (Google 2020), and Google Earth Pro (Google 2019) to help identify historical and current conditions of the site. They report their findings from this virtual site visit in the first section of page two of the SSAR.

In addition to the written summary of the virtual site visit, the analysts insert two figures into each report. The first figure is a street view (taken from one of the aforementioned services) depicting something of interest mentioned in the written summary. The second figure is a Geographic Information System (GIS) map showing the location of the segment. The GIS maps are created using the ArcMap software published by Esri as well as an Esri shapefile that can be created from the Results file. An example of these two figures is provided in Figure 6-8.

6.3.2.2 Possible Countermeasures

The VBA code is programmed to paste a list of countermeasures at the bottom of the second page of the SSAR. This list is a compilation of countermeasures found in the National Cooperative Highway Research Program (NCHRP) 500 Report (Neuman et al. 2003) and the Countermeasures That Work report (Goodwin et al. 2015). Research analysts search through this auto-populated list for actions that may help to reduce the types of crashes present at the site. The analysts choose 4 to 12 actions and list the selected countermeasures in one of two columns, engineering countermeasures or policy countermeasures, depending on the nature of the potential solution. The unselected countermeasures are then removed from the SSAR.

6.4 Summary

The output of the statistical model identifies the segments that could most benefit from safety improvements, and the SSARs provide suggestions for improvement for these top-ranking segments. The creation of these reports is performed automatically; the VBA code is programmed to fill in information about the physical characteristics and crash history of the segments and then print each report to a separate Excel workbook. The individual reports are then reviewed by research analysts. Supplemental information about the segments is written and hand-typed into the SSARs to augment the data tables provided.

Segment Safety Analysis Report

Introduction

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

Segment Identification and Roadway Characteristics

Date: 7/17/2020

Street Name: *Ogden River Scenic Byway*

Table 1: Segment Metadata

Route Number:	39	**UC Model Used:**	CAMS-P
Road Direction:	P	**State Rank:**	1
Beginning, Ending MP:	42.67-44.13	**Region, Rank:**	1, 1
Length (miles):	1.46	**County, Rank:**	WEBER, 1
Data Source Years:	2014-2018		

Table 2: Segment Characteristics

Functional Class:	Major Collector	**AADT:**	510
Number of Thru Lanes:	2	**Speed Limit (MPH):**	45

Table 3: Roadway Characteristics

MPs	Median	Shoulder	Grade	Curve	Lanes	Wall/ Barrier	Rumble Strips
42.67 - 44.13	None	Paved - 3ft	Steep	4 sharp curves	2	None	No

Micro-Analysis of Crash Data

Crash Data Summary

Table 4: Crash Count and Severity

Crash Severities	Functional Area Method	Crashes During 2018		Total Crashes Between 2014-2018				
		Predicted	Actual	Sev. 5	Sev. 4	Sev. 3	Sev. 2	Sev. 1
345	by speed from stop bar	0.0716	4	0	3	1	0	2

Table 5: Top 7 Crash Factors

Crash ID	Latitude	Longitude	SINGLE VEHICLE	ROADWAY GEOMETRY RELATED	ROADWAY DEPARTURE	COLLISION WITH FIXED OBJECT	OVERTURN ROLLOVER	NIGHT DARK CONDITION	IMPROPER RESTRAINT
Injury Total			**4/4**	**4/4**	**4/4**	**3/4**	**2/4**	**1/4**	**1/4**
Segment Total			6/6	5/6	5/6	4/6	2/6	3/6	1/6

Table 6: Manner of Collision Data

	Manner 1	Manner 2	Manner 3	Manner 4	Manner 5	Manner 6	Manner 7	Manner 8	Manner 9
Name	Single Vehicle	Angle	Front to Rear	Head On	Sideswipe Same Direction	Sideswipe Opposite Direction	Parked Vehicle	Rear to Side	Rear to Rear
Injury Total	**4/4**	**0/4**	**0/4**	**0/4**	**0/4**	**0/4**	**0/4**	**0/4**	**0/4**
Segment Total	6/6	0/6	0/6	0/6	0/6	0/6	0/6	0/6	0/6

Figure 6-7: Sample SSAR – page one.

Historical Perspective, Current Conditions, Site Visit Notes *Date:* 3/25/2020

This segment has not experienced significant changes since the beginning of the analysis period (2014).

This segment is a portion of SR 39 in Weber County. It is a two lane undivided highway. There is one northbound lane and one southbound lane. There are no rumble strips. The paved shoulder on both sides of the road is 3 ft. There are horizontal and vertical curves on this segment, one of which can be seen in Figure 1. There are no barriers on the curves.

Figure 1: A curve on this segment of the Ogden River Scenic Byway (SR 39) without delineation or barriers (Google).

Figure 2: GIS map showing the location of the segment (ESRI).

Possible Countermeasures

The following is a list of possible countermeasures related to the top 8 crash factors listed in Table 5. The countermeasures listed were compiled from the NCHRP 500 Report volumes and Countermeasures 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*).)

Engineering Countermeasures

Change or mitigate the effects of identified elements in the environment (E)
Delineate roadside objects (E)
Provide adequate sight distance (T)
Provide advance warning of unexpected changes in horizontal alignments (T)
Remove/relocate objects in hazardous locations (P)
Widen the roadway (P)
Design safer slopes and ditches to prevent rollovers (P)

Policy Countermeasures

Encourage trucking companies and other fleet operators to implement fatigue management programs (T)

Figure 6-8: Sample SSAR – page two.

7.0 ANALYSIS OF RESULTS

7.1 Overview

The results of the CAMS are meant to help UDOT prioritize segments of Utah state routes for safety improvements. As discussed in Chapter 5, segments are ranked using statistical models (the CAMS-P and CAMS-S models). This chapter provides discussion on the top 20 high-priority segments as identified by both statistical models. It also discusses the top crash factors and manners of collision found in the CAMS-P model results.

7.2 CAMS Prediction Model Results Discussion

The results of the CAMS-P model analysis are given in Table 7-1 and mapped in Figure 7-1. The route label gives the 4-digit route ID plus the direction code (P stands for positive, meaning the northbound or eastbound direction; N stands for negative, meaning the southbound or westbound direction). The percentile values expressed in the table come from the statistical analysis, and a higher percentile value directly correlates to a higher state rank. The value for "2018 Predicted Injury Crashes" is the mean value of the crash distributions to which the "2018 Injury Crashes" values are compared. As can be seen in Table 7-1, the top 20 segments are spread almost equally between the four UDOT Regions. The top-ranking segments ranged from having 3 crashes to 357 total crashes in the five-year period from 2014 to 2018. The only route to have multiple segments show up in the top 20 was I-15.

It is interesting to note that five of the top 20 segments are from UDOT Region 4. Of all three years BYU has run the UCPM with the RSAM methodology, Region 4 segments never placed in the top 20 in the state. The sudden jump for multiple Region 4 segments is theorized to be explained by understanding the differences between the RSAM and the CAMS. Because the RSAM does not remove intersection-related crashes from the segments, Regions 1, 2, and 3—which tend to have larger and busier intersections than Region 4—are more likely to be flagged as outliers than any segment in Region 4. But in the CAMS, a methodology that removes intersection-related bias, segments would only rank high if they are experiencing an unusually high number of segment-related crashes. Thus, by removing intersection-related bias, it can be seen that there are some segments in Region 4 that may have a significant safety concern when it comes to segment-related crashes.

Although some of the actual crash numbers may not appear very high or concerning, it is important to emphasize the difference between the actual and predicted number of crashes. State rank 1, for example, experienced 4 injury crashes in 2018. Yet according to the model, segments with roadway characteristics like that of this segment are predicted to average less than 0.1 injury crashes per year. The difference, therefore, is significant.

7.3 CAMS Severity Model Results Discussion

The results of the CAMS-S model analysis are given in Table 7-2 and are mapped in Figure 7-2. The columns in the table function the same way as they do in Table 7-1 and the discussion in Section 7.2. This severity model, however, shows which segments have a higher proportion of injury crashes, meaning that if there is a crash on a segment that ranked high in the CAMS-S model, this crash is more likely to be injury-causing than it might be on another segment. As can be seen in Table 7-2, 15 of the top 20 segments are in UDOT Regions 1 and 2, and eight are in Salt Lake County. The top-ranking segments range from having 6 crashes to 412 total crashes in the five-year period from 2014 through 2018. Four routes had multiple segments in the top 20: I-15, I-80, UT-209 (9000 S / 9400 S in Salt Lake County), and UT-171 (3500 S / 3300 S in Salt Lake County).

Historically, RSAM segments in UDOT Region 4 rank higher in the UCSM than they do in the UCPM. However, the CAMS results show fewer Region 4 segments in the CAMS-S top 20 than in the CAMS-P top 20. Only one segment from Region 4 ranked in the CAMS-S top 20: a segment of southbound I-15 in Washington County which also ranked in the top 20 of the CAMS-P results. Seven other segments also showed up in the top 20 for both models. When a segment shows up in both models it indicates that it is a location that has a high number of crashes compared to similar segments and that the proportion of injury crashes is greater than other segments as well. For the segments that rank high in the CAMS-S model but not in the CAMS-P model, that means that although the number of crashes along that segment is not extremely atypical, the proportion of injury crashes is higher than expected.

Table 7-1: CAMS-P Model Top 20 Segments

Route Label	*Beginning Milepoint*	*Ending Milepoint*	*Length (miles)*	*County*	*UDOT Region*	*Total Crashes in 5 yrs*	*Injury Crashes in 5 yrs*	*2018 Injury Crashes*	*Predicted 2018 Injury Crashes*	*Percentile*	*State Rank*	*Region Rank*	*County Rank*
0039P	42.67	44.13	1.46	WEBER	1	6	4	4	0.07	1	1	1	1
0060P	0	0.419	0.419	WEBER	1	35	11	7	0.91	0.9998	2	2	2
0150P	33.292	46.48	13.188	SUMMIT	2	12	9	5	0.42	0.9996	3	1	1
0171P	13.583	14.087	0.504	SALT LAKE	2	55	10	7	0.87	0.9994	4	2	1
0089P	103.866	114.318	10.452	KANE	4	164	15	8	1.44	0.9992	5	1	1
0068P	64.365	65.746	1.381	DAVIS	1	81	6	5	0.58	0.999	6	3	1
0065P	8.441	14.175	5.734	MORGAN	1	27	16	6	0.99	0.9986	7	4	1
0189P	25.278	26.541	1.263	WASATCH	3	48	5	4	0.38	0.9985	8	1	1
0015N	296	297.086	1.086	SALT LAKE	2	357	27	13	3.87	0.9984	9	3	2
0015P	134.875	138.073	3.198	MILLARD	4	77	17	8	1.79	0.9983	10	2	1
0173P	4.887	5.031	0.144	SALT LAKE	2	105	8	5	0.69	0.9976	11	4	3
0015N	31.861	37.475	5.614	WASHINGTON	4	138	26	11	3.64	0.9972	12	3	1
0015N	269.438	269.872	0.434	UTAH	3	55	9	6	0.94	0.9972	13	2	1
0008P	0	1.18	1.18	WASHINGTON	4	134	18	9	2.37	0.9971	14	4	2
0154P	6.759	8.274	1.515	SALT LAKE	2	135	21	10	3.03	0.9971	15	5	4
0091P	26.242	26.886	0.644	CACHE	1	193	7	5	0.79	0.9959	16	5	1
0015P	335.8	336.295	0.495	DAVIS	1	95	15	7	1.70	0.9953	17	6	2
0210P	9.855	9.956	0.101	SALT LAKE	2	3	2	2	0.09	0.9946	18	6	5
0014P	25.711	40.995	15.284	KANE	4	92	15	7	1.81	0.9943	19	5	2
0035P	29.732	31.505	1.773	DUCHESNE	3	6	3	2	0.10	0.9943	20	3	1

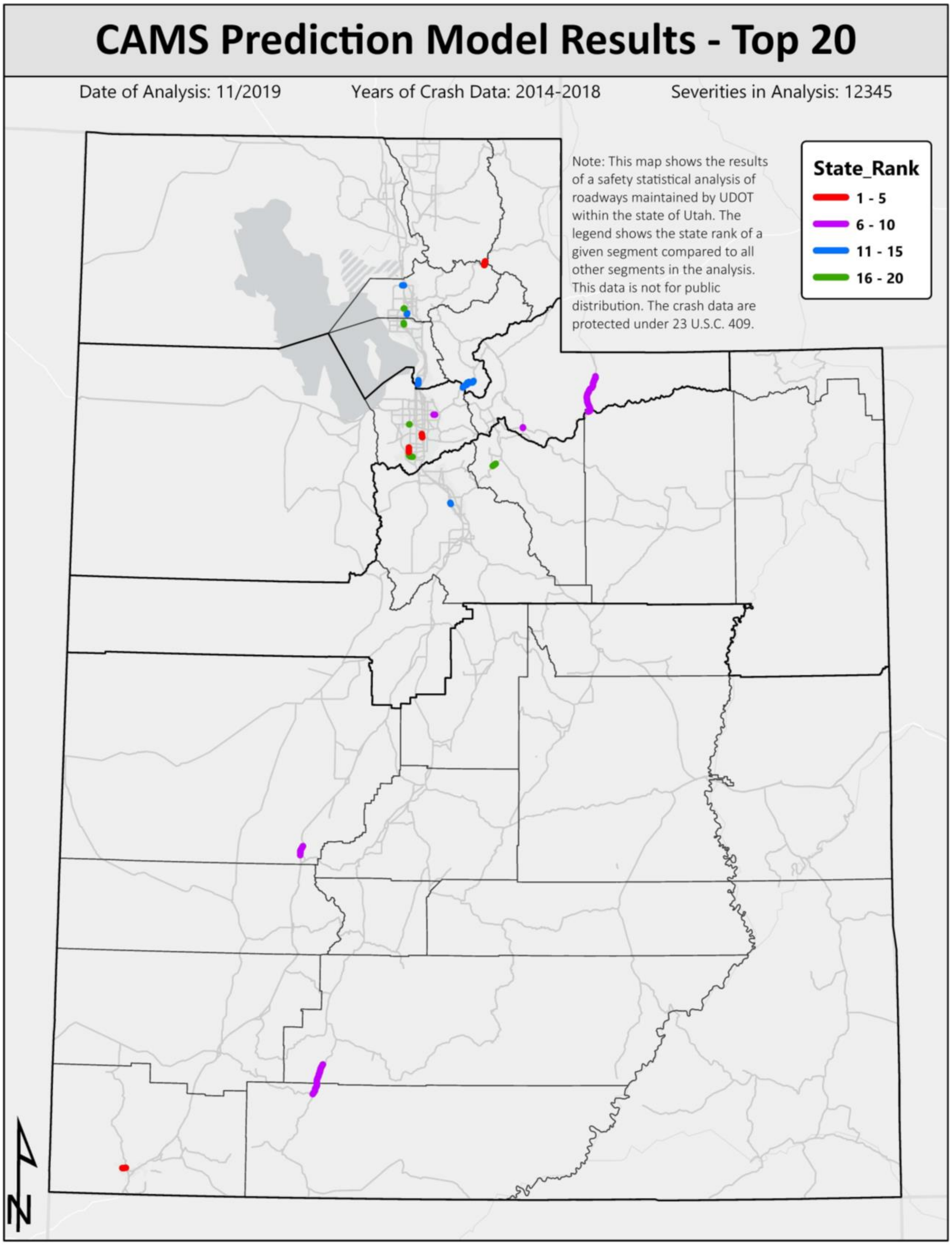


Figure 7-1: Map of CAMS-P model top 20 segments.

Table 7-2: CAMS-S Model Top 20 Segments

Route Label	*Beginning Milepoint*	*Ending Milepoint*	*Length (miles)*	*County*	*UDOT Region*	*Total Crashes in 5 yrs*	*Injury Crashes in 5 yrs*	*2018 Injury Crashes*	*Predicted 2018 Injury Crashes*	*Percentile*	*State Rank*	*Region Rank*	*County Rank*
0060P	0	0.419	0.419	WEBER	1	35	11	7	1.59	0.9999	1	1	1
0080P	56.195	61.837	5.642	TOOELE	2	29	10	4	0.54	0.9997	2	1	1
0150P	33.292	46.48	13.188	SUMMIT	2	12	9	5	0.75	0.9997	3	2	1
0089P	381.727	383.116	1.389	SALT LAKE	2	25	7	4	0.51	0.9995	4	3	1
0039P	42.67	44.13	1.46	WEBER	1	6	4	4	0.67	0.9991	5	2	2
0015P	282.241	283.92	1.679	UTAH	3	329	47	15	5.77	0.999	6	1	1
0065P	8.441	14.175	5.734	MORGAN	1	27	16	6	1.50	0.9989	7	3	1
0171P	13.583	14.087	0.504	SALT LAKE	2	55	10	7	1.95	0.9975	8	4	2
0080P	69.521	76.402	6.881	TOOELE	2	28	7	4	0.77	0.9972	9	5	2
0209P	5.213	6.203	0.99	SALT LAKE	2	22	7	4	0.80	0.9971	10	6	3
0015N	31.861	37.475	5.614	WASHINGTON	4	138	26	11	4.60	0.9969	11	1	1
0126P	10.735	11.247	0.512	WEBER	1	46	17	5	1.14	0.9969	12	4	3
0209P	16.152	17.068	0.916	SALT LAKE	2	37	10	3	0.41	0.9968	13	7	4
0154P	6.759	8.274	1.515	SALT LAKE	2	135	21	10	3.80	0.9963	14	8	5
0085P	3.528	6.565	3.037	SALT LAKE	2	24	6	3	0.42	0.9953	15	9	6
0015N	269.438	269.872	0.434	UTAH	3	55	9	6	1.77	0.9948	16	2	2
0092P	13.824	16.043	2.219	UTAH	3	13	5	3	0.53	0.9946	17	3	3
0171P	11.93	12.533	0.603	SALT LAKE	2	67	12	5	1.30	0.9935	18	10	7
0209P	6.203	7.671	1.468	SALT LAKE	2	176	27	9	3.60	0.9932	19	11	8
0015N	265.842	269.035	3.193	UTAH	3	412	57	14	6.84	0.9922	20	4	4

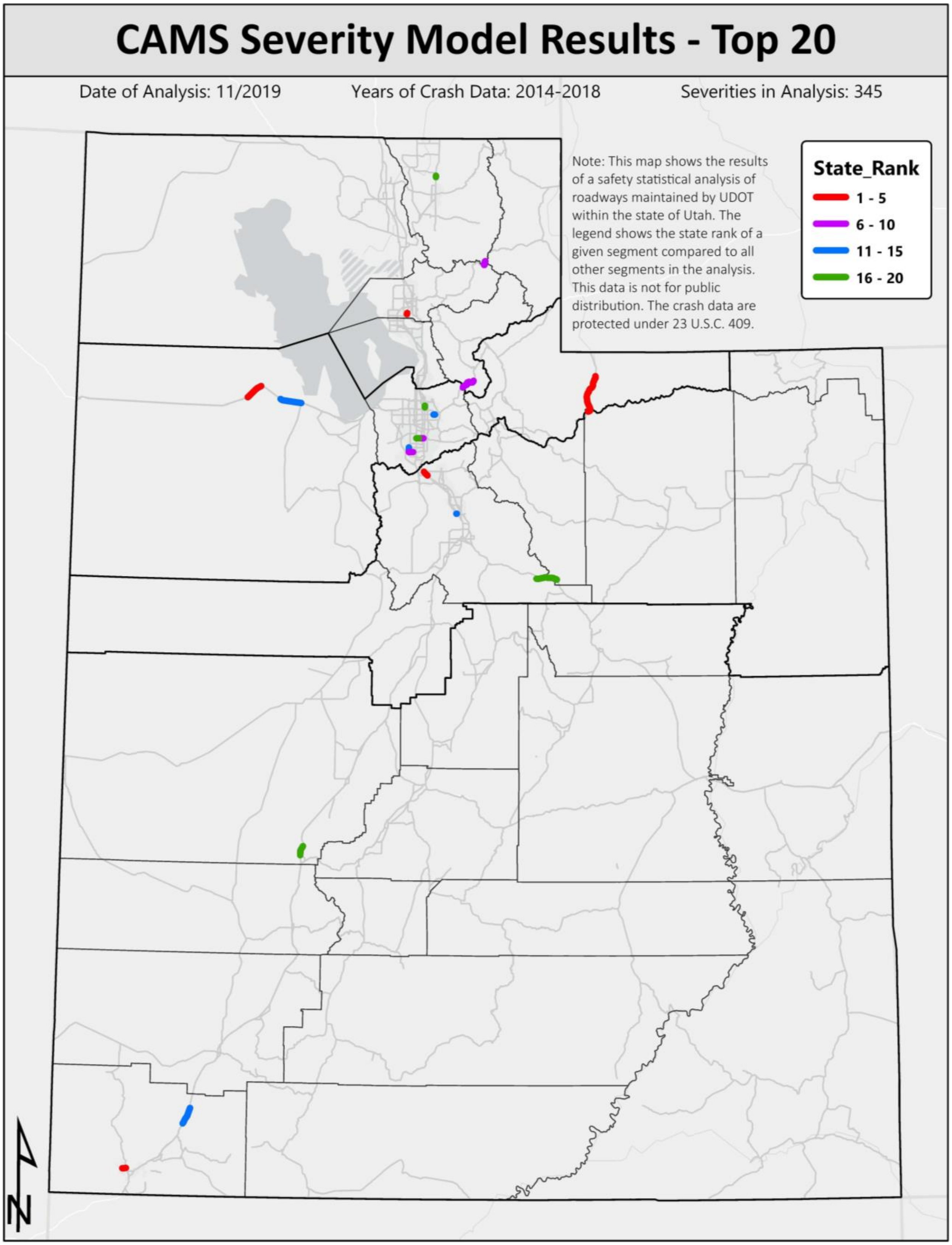


Figure 7-2: Map of CAMS-S model top 20 segments.

7.4 Crash Factors Discussion

Crash factors can be a window into the types of safety concerns present at a site and can aid an engineer in selecting useful countermeasures. Table 7-3 provides a list of all the crash factors found in the Crash Rollup file and gives the count and percentage of crashes in the CAMS-P model top 20 segments associated with each factor. It should be noted that each crash may have more than one crash factor associated with it and that all associated crash factors are accounted for in the table.

Table 7-3: Crash Factors for the Top 20 Segments

Factor	*Injury Crashes*	*%*	*All Crashes*	*%*
Single Vehicle	112	12%	643	13%
Roadway Geometry Related	110	12%	519	11%
Roadway Departure	87	9%	328	7%
Overturn/Rollover	76	8%	130	3%
Speed Related	59	6%	391	8%
Collision with Fixed Object	54	6%	292	6%
Night/Dark Condition	51	5%	433	9%
Adverse Roadway Surface Condition	47	5%	344	7%
Motorcycle Involved	45	5%	65	1%
Teenage Driver Involved	43	5%	287	6%
Adverse Weather	36	4%	267	5%
Intersection Related	32	3%	207	4%
Older Driver Involved	31	3%	235	5%
Distracted Driving	23	2%	147	3%
Driving Under the Influence (DUI)	19	2%	50	1%
Work Zone Related	19	2%	87	2%
Commercial Motor Vehicle Involved	15	2%	113	2%
Unrestrained	15	2%	36	1%
Drowsy Driving	14	2%	39	1%
Improper Restraint	14	2%	35	1%
Wild Animal Related	11	1%	236	5%
Bicyclist Involved	6	1%	8	0%
Aggressive Driving	5	1%	14	0%
Pedestrian Involved	4	0%	6	0%
Domestic Animal Related	2	0%	7	0%
Transit Vehicle Involved	2	0%	10	0%
Railroad Crossing	0	0%	0	0%
Train Involved	0	0%	0	0%

It is interesting to note that the proportion of some crash factors is higher when looking at just injury crashes compared to all crashes, specifically that of Overturn/Rollover and Motorcycle crashes. This indicates that crashes involving these are more likely to be injury-causing. The opposite is true for Wild Animal Related crashes which are more represented in non-injury crashes than in injury crashes. This information can help UDOT understand that protecting vehicles from overturning as well as helping drivers and motorcyclists more safely share the road with each other has potential to help lower injury crash rates, whereas crashes with wild animals may not be as concerning in terms of driver (or passenger) injury.

It is also interesting to note that some of the crashes in the results were marked as Intersection Related. There are a few possible explanations for this. The first possibility is that there are intersections that did not fall under the three categories excluded from the analysis: state route to state route, state route to federal aid route, and signalized intersections on a state route. Intersection-related crashes from these three categories are removed in the data integration process as discussed in Section 4.3.2. If a non-state route or non-federal aid route intersected the state route without a signal, the crashes that occurred on the segment would still be included in the CAMS. Another possibility is that the functional area was not large enough to cover all the intersection-related crashes at one of the three intersection types. A third possibility is that some police officers who filed crash reports considered mid-block crashes related to accesses and driveways to be intersection related and marked the crashes as such on the crash report. Thus, it is very possible that there were intersection-related crashes in the CAMS data.

The fact that a percentage of crashes were intersection related should not be ignored. Knowing the source of these crashes can be insightful. If the crashes come from classifying accesses and driveways as intersections, then there is no need for concern at the high percentage of intersection-related crashes; crashes at mid-block locations would most likely be addressed with a segment-type countermeasure and thus fit in well with the purpose of the CAMS analysis. Yet if the crashes come from an inadequately sized functional area, further research should be performed on this topic, perhaps through the use of a spatial analysis as will be discussed in Section 9.3.2. This is not a likely source of the intersection-related crashes present in the analysis since the values for the functional area are quite conservative and account for perception-reaction time, deceleration, and queuing. Finally, if the crashes come from intersections that do not fit one of the three intersection types removed by the CAMS processes, perhaps all intersection types and their associated crashes should be removed for a segment-only analysis. This action was decided against early in the development of the CAMS due to the intention that the CAMS would be a complement to the ISAM, meaning that the CAMS would analyze only the crashes not analyzed in the ISAM. Due to the dense network of intersections common on Utah roadways, the ISAM would have double-counted a large number of crashes between two intersections if intersections were not selectively chosen to be part of the analysis. Thus, the decision was made to avoid such an action.

Other factors near the top of the list in Table 7-3 include Single Vehicle, Roadway Geometry Related, and Roadway Departure, to name a few. The point of learning what factors were involved in the crashes is to give engineers a greater ability to choose countermeasures in which they can place confidence.

7.5 Manners of Collision Discussion

Manners of collision give insight into how vehicles are crashing and, like crash factors, are useful in aiding an engineer to select countermeasures that will make a difference at the site. Table 7-4 lists all possible manners of collision as well as the count and percentage of crashes in the CAMS-P model top 20 segments associated with each manner.

Table 7-4: Manners of Collision for the Top 20 Segments

Manner of Collision	*Injury Crashes*	*%*	*All Crashes*	*%*
N/A [Single Vehicle]	116	48%	680	37%
Front to Rear	61	25%	673	37%
Angle	37	15%	211	12%
Sideswipe Same Direction	20	8%	191	11%
Head On	5	2%	25	1%
Parked Vehicle	3	1%	15	1%
Sideswipe Opposite Direction	2	1%	23	1%
Rear to Side	0	0%	0	0%
Rear to Rear	0	0%	0	0%
Unknown	0	0%	0	0%

The most common manner of collision is N/A which almost always means that the crash only involved one motorized vehicle. This is expected; single vehicle crashes are likely to be more prominent on a segment than at an intersection. The next most common manner of collision is Front to Rear, commonly called rear-end crashes. In contrast with the first manner of collision, rear-end crashes are more often associated with intersections and queues than with segments. Although surprising, this may be explained via the same reasons as found in the discussion on intersection-related crashes in Section 7.4.

7.6 Summary

The 20 highest ranking segments in the CAMS-P model and CAMS-S model have eight common locations between them. The results from both models also show common trends including the high proportion of segments from UDOT Regions 1 and 2 and segments located in Salt Lake County in particular. Even so, five segments from UDOT Region 4 ranked in the top 20, an outcome previously unobserved in RSAM results. Crash factors and manners of collision data for the CAMS-P model top 20 segments allow engineers to receive insights into what specific safety problems may be occurring at each site. Rollover and motorcycle crashes are shown to be more frequently injury-causing than other crash factors analyzed. The data also indicate that there are still some intersection-related crashes being analyzed by the CAMS, but this is likely not a concern as it allows the model to identify mid-block locations where crashes may be occurring at driveways. This allows an analysis of access management to be performed.

8.0 INTERSECTION ANALYSIS MODIFICATION

8.1 Overview

Part of the research efforts tied to the CAMS was the modification of the ISAM. In 2019, a new and more comprehensive Intersections file (one of the key inputs into the ISAM) was produced and provided by UDOT. Functional class data also became available to BYU for federal aid routes in addition to state routes. These new files allowed an improvement of the ISAM to be possible. This chapter will describe the changes made to the ISAM and will discuss the model results obtained in 2019.

8.2 Changes in the ISAM

A BYU research team updated the ISAM in 2019 to analyze not only state route-to-state route intersections, but also state route to federal aid-route intersections and signalized state route-to-local road intersections. This increased the number of analyzed intersections approximately seven-fold. This section will discuss the main differences between the 2018 ISAM and the 2019 ISAM, including changes made to the data preparation process, the statistical analysis, and the technical reports for high-ranking segments.

8.2.1 Changes in Data Preparation

The first major change between the 2018 and 2019 versions of the ISAM is found in the types of intersections that can be analyzed. In 2018, only intersections with at least two distinct state routes could be analyzed. Information in the UDOT Intersections file for intersections with only one state route was limited, and data for the non-state routes at these intersections were not available. However, more information about the intersecting routes was added to the Intersections file, allowing more detailed identification of all intersections along a state route. In the 2019 version of the ISAM, up to three types of intersections may be included. These three intersection types are: State Route-to-State Route Intersections (SR to SR), State Route-to-Federal Aid-Route Intersections (SR to Fed Aid), and Signalized State-Route Intersections (Signalized SR). These options are presented in the input form used to combine the roadway data together as outlined in red in Figure 8-1. It is recommended to select all three intersection types—doing so maximizes the number of intersections which as a result improves the statistical analysis.

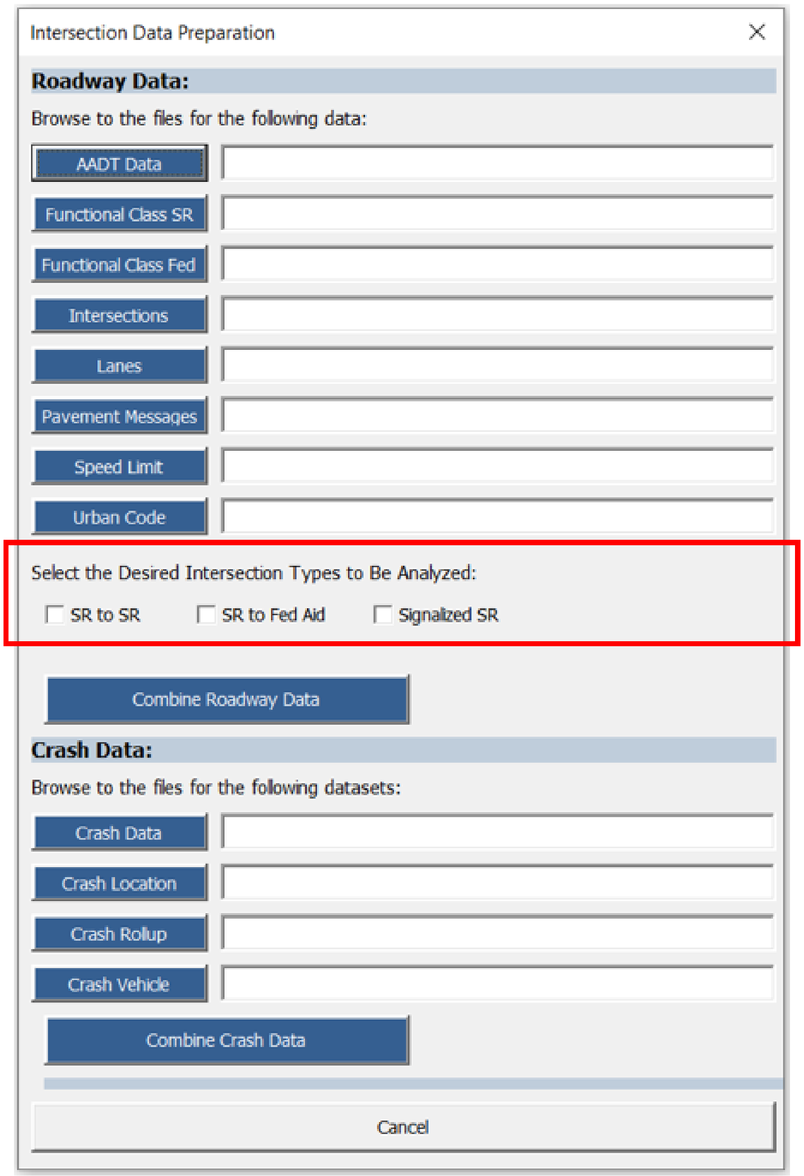


Figure 8-1: 2019 ISAM input form for the data preparation process.

The second major change to the ISAM is found in the types of crashes included in the analysis. Unlike the 2018 version, the 2019 ISAM only includes crashes coded as "Intersection-Related" in the UDOT Crash Rollup file. Furthermore, crashes from all the approaches at the intersection are included, not just the ones that occurred on a state route. These crashes are still associated with an intersection based on the determined functional area of the intersection, and the 2019 ISAM still recommends that the distance from the intersection stop bar to the end of the functional area be calculated using the approach speed limit. The option to define the functional area by an alternative method still exists, but the options are different than they were in the 2018 version. The options to define the functional area based on functional class or urban code were removed (these options had historically never been used by BYU or UDOT to run the ISAM) and the option to define the functional area by a fixed length was made available. The default value for this fixed length is 250 feet, but it may be changed by the user.

8.2.2 Changes in Statistical Analysis

Changes to the statistical analysis were also made to the ISAM. To avoid using the same data to build the model and rank the intersections, the UICPM was altered while continuing to use a ZIP model. Like the CAMS-P model used for segments, the new UICPM uses four years (2014-2017 in this study) of injury crash data to build the model and one year (2018 in this study) to rank the intersections. The model uses number of entering vehicles, percent of trucks, number of lanes, roadway width, and approach speed limit as variables. In addition, there is a hierarchical structure that allows the effect of the variables to be different for intersections in different urban codes. Detailed explanation of the new UICPM (as well as alternative models considered) can be found in the BYU Statistics technical report titled "Justification for Considering Zero-Inflated Models in Intersection Safety Analysis" (Pew 2020).

An additional statistical model was also applied to the data. This model is named the Utah Intersection Crash Severity Model (UICSM) and, unlike the UICPM, it uses all the crashes (Severities 1 through 5) in its analysis. Like the CAMS-S model used for segments, the purpose of the UICSM is to identify intersections with higher proportions of injury crashes (compared with total crashes) than predicted; the higher an intersection is ranked, the more likely it is that a crash occurring on that intersection would be injury-causing. Using the proportions observed in a four-year period (2014-2017 in this study), it creates predicted distributions of injury crashes for a fifth year (2018 in this study). In like fashion to the UICPM, the intersections are then ranked according to the percentile values of the observed crashes within those distributions.

8.2.3 Changes in the Technical Reports

Many of the changes to the ISARs came as a result of a change made to the way the intersection legs are numbered. The 2019 version of the ISAM identifies up to five unique routes, as opposed to three unique routes found in the 2018 ISAM. In the 2019 version, intersection legs are numbered Route 0 through Route 4 as follows: Route 0 increases in the positive milepost direction as it approaches the intersection. The leg to the right is Route 1, the leg straight ahead is Route 2, the leg to the left is Route 3, and for five-legged intersections the fifth leg (not directly 90 or 180 degrees from Route 0) is Route 4. Figure 8-2 provides a visual representation of the route numbering system. For further clarification, a diagram was added adjacent to the Intersection Metadata table as can be seen in Figure 8-3. The Intersection Metadata and Intersection Characteristics tables in the ISARs were also updated to reflect the new numbering system and available data.

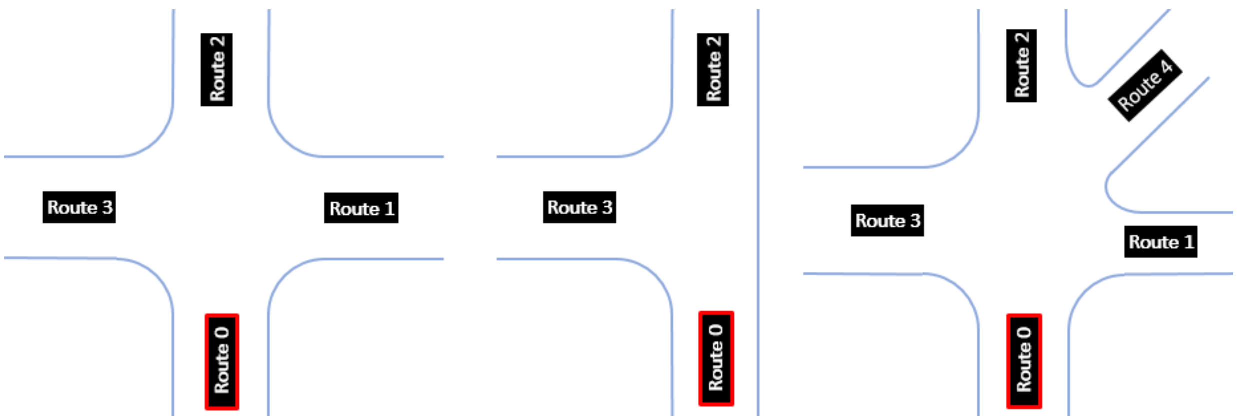


Figure 8-2: Examples of route numbering (Adapted from UDOT Traffic & Safety 2017).

Several additional changes were made to the ISARs in 2019. The 2018 version included a section for a paragraph summary of the crash history of the intersection. This was replaced in the 2019 version with a table containing information about the most common manners of collision in the crash history at that intersection. This new table and the table for the most common crash factors were summed up in two ways instead of only one (the sum per total crashes was provided in addition to the sum per injury crashes previously provided). The implementation of these adjustments can be seen in Figure 8-4. The 2018 ISARs also included a list of suggested countermeasures for the safety problems present at the intersection taken from countermeasures in the NCHRP Report 500 (Neuman et al. 2003) and the Countermeasures That Work report

(Goodwin et al. 2015). In completed 2019 ISARs, this list is separated into two categories—engineering countermeasures and policy or enforcement countermeasures—to highlight the distinct solutions that UDOT engineers, Utah Highway Patrol, and Zero Fatalities (among others) can bring to the table.

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:*

Street Names:

Table 1: Intersection Metadata

Model Used:		**Leg 0 Route & MP:**		
State Rank:		**Leg 1 Route & MP:**		
Region & Rank:		**Leg 2 Route & MP:**		
County & Rank:		**Leg 3 Route & MP:**		
Years of Data:		**Leg 4 Route & MP:**		
City/Area:		**Latitude & Longitude:**		

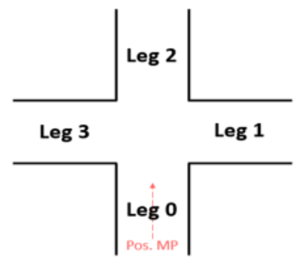


Table 2: Intersection Characteristics

Intersection Control:		**Entering Vehicles in**	
Max. Functional Class:		**# of Lanes on Route 0:**	
Min. Functional Class:		**Max & Min Speed Limit (mph):**	

Figure 8-3: Intersection identification and characteristics section of the 2019 ISARs.

Micro-Analysis of Crash Data

Crash Data Summary

Table 3: Crash Count and Severity

Crash Severities Used	Functional Area Method Used	Crashes during 2018		Total crashes between 2014-2018				
		Predicted	Actual	Sev. 5	Sev. 4	Sev. 3	Sev. 2	Sev. 1

Table 4: Crash Factors

Crash ID	Latitude	Longitude	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5	Factor 6	Factor 7
Injury Total									
Intersect'n Total									

Table 5: Manner of Collision Data

	Manner 1	Manner 2	Manner 3	Manner 4	Manner 5	Manner 6	Manner 7	Manner 8	Manner 9
Name									
Injury Total									
Intrsectn. Total									

Figure 8-4: Crash data summary section of the 2019 ISARs.

8.3 Results and Discussion

The results of the UICPM are given in Table 8-1 and mapped in Figure 8-5. The route label gives the 4-digit route ID plus the direction code (P stands for positive, meaning the northbound or eastbound direction; N stands for negative, meaning the southbound or westbound direction). The percentile values expressed in the table come from the statistical analysis, and a higher percentile value directly correlates to a higher state rank. The value for “2018 Predicted Injury Crashes” is the mean value of the crash distributions to which the “2018 Injury Crashes” values are compared. As can be seen in Table 8-1, 12 of the top 20 intersections are in UDOT Region 2, all of which were in Salt Lake County. The top-ranking intersections ranged from having 2 crashes to 249 total crashes in the five-year period from 2014 to 2018, and most of these intersections are signalized. Only four of the top 20 intersections involve two or more state routes, indicating that the expanded intersection analysis can discover hot spots that the 2018 version cannot.

The results of the UICSM are given in Table 8-2 and are mapped in Figure 8-6. The columns in the table function the same way as they do in Table 8-1. This severity model, however, shows which intersections have a higher proportion of injury crashes than predicted. This means that a crash at a high-ranking intersection is more likely to be injury-causing than it might be at another intersection.

Similar to the results of the UICPM, 12 of the UICSM top 20 intersections are in UDOT Region 2, all of which are in Salt Lake County. The top-ranking intersections range from having 2 crashes to 84 total crashes in the five-year period from 2014 through 2018, and nearly all of these intersections are signalized. Five of the top 20 intersections are on route 71 (700 E / 900 E in Salt Lake County). One final thing to note is that seven of the top 20 intersections were also ranked in the top 20 by the UICPM. Further discussion on the interplay between prediction model results and severity model results was given with the CAMS results in Section 7.3.

Table 8-1: UICPM Top 20 Intersections

Routes	*Traffic Control Device*	*Urban Area*	*County*	*UDOT Region*	*Total Crashes in 5 yrs*	*Injury Crashes in 5 yrs*	*2018 Injury Crashes*	*Predicted 2018 Injury Crashes*	*Percentile*	*State Rank*	*Region Rank*	*County Rank*
171 & 2154	Stop Sign	Salt Lake City	SALT LAKE	2	59	4	4	0.22	0.9997	1	1	1
108 & Local	Signal	Ogden - Layton	DAVIS	1	58	12	8	1.34	0.9993	2	1	1
154 & 175	Signal	Salt Lake City	SALT LAKE	2	180	14	8	1.72	0.9986	3	2	2
111 & 2232	Signal	Salt Lake City	SALT LAKE	2	41	7	5	0.55	0.9985	4	3	3
209 & 2108	Stop Sign	Salt Lake City	SALT LAKE	2	84	18	9	2.08	0.9984	5	4	4
50& 125 & 136	Stop Sign	Rural	MILLARD	4	4	2	2	0.07	0.9971	6	1	1
8 & 3234	Signal	St. George	WASHINGTON	4	30	8	5	0.74	0.9967	7	1	2
172 & 2250	Stop Sign	Salt Lake City	SALT LAKE	2	12	3	3	0.20	0.9965	8	5	5
154 & 173	Signal	Salt Lake City	SALT LAKE	2	249	18	9	2.58	0.9953	9	6	6
71 & 2086	Signal	Salt Lake City	SALT LAKE	2	41	11	6	1.26	0.9944	10	7	7
218 & 1254	Stop Sign	Rural	CACHE	1	9	2	2	0.09	0.9938	11	1	2
92 & 2920	Signal	Provo-Orem	UTAH	3	66	8	5	0.87	0.9935	12	1	1
71 & Local	Signal	Salt Lake City	SALT LAKE	2	21	3	3	0.27	0.9924	13	8	8
173 & Local	Signal	Salt Lake City	SALT LAKE	2	26	7	4	0.73	0.9869	14	9	9
85 & 2036	Signal	Salt Lake City	SALT LAKE	2	110	39	13	5.88	0.9861	15	10	10
89 & 2890	Signal	Provo-Orem	UTAH	3	40	10	5	1.21	0.9851	16	2	2
89 & 2124	Signal	Salt Lake City	SALT LAKE	2	146	9	5	1.22	0.9845	17	11	11
55 & 1338	Stop Sign	Small Urban	CARBON	4	2	2	2	0.16	0.9815	18	1	3
85 & 2036	Signal	Salt Lake City	SALT LAKE	2	111	40	13	6.27	0.9788	19	12	12
178 & 198	Signal	Provo-Orem	UTAH	3	42	5	3	0.49	0.9785	20	3	3

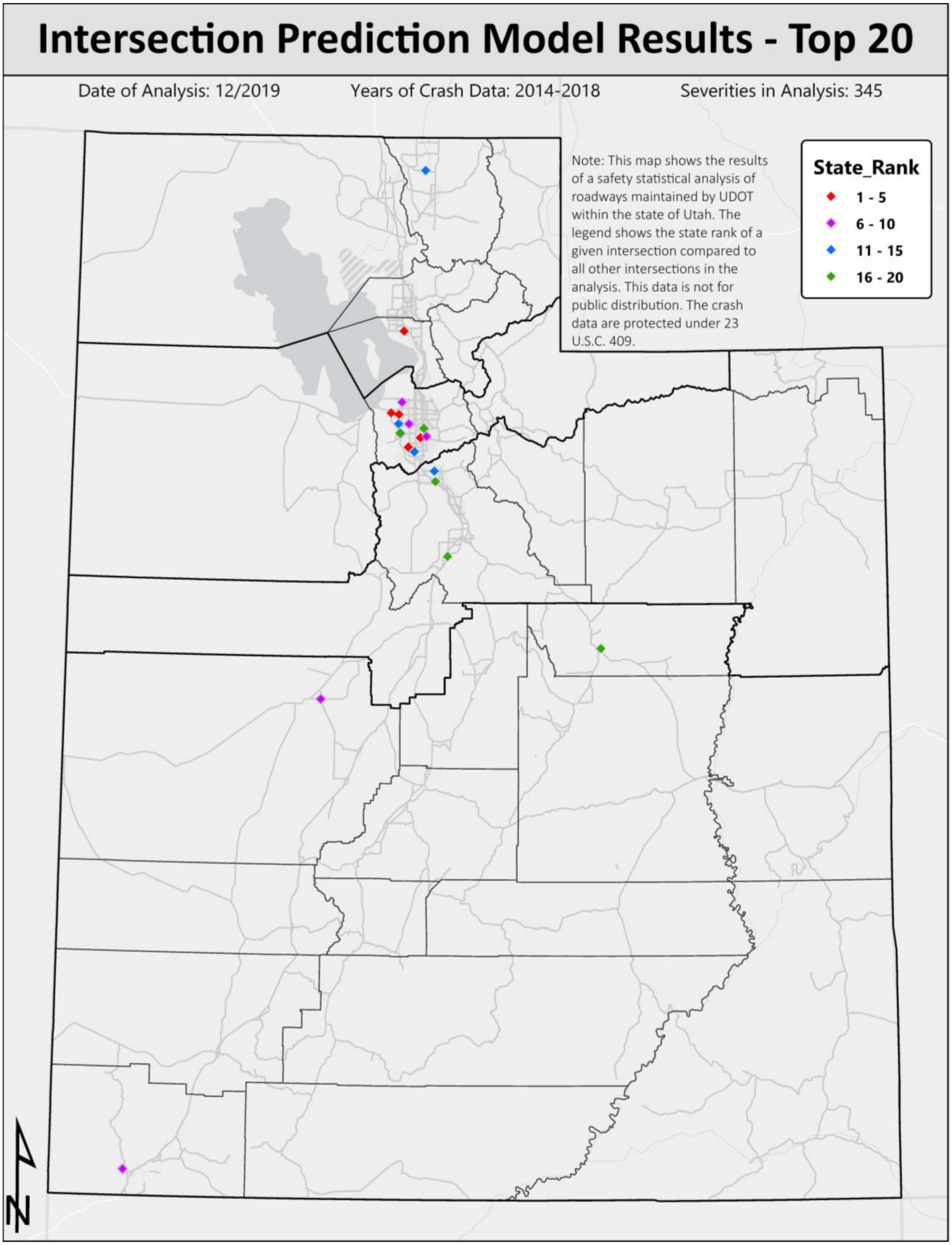


Figure 8-5: Map of UICPM top 20 intersections.

Table 8-2: UICSM Top 20 Intersections

Routes	*Traffic Control Device*	*Urban Area*	*County*	*UDOT Region*	*Total Crashes in 5 yrs*	*Injury Crashes in 5 yrs*	*2018 Injury Crashes*	*Predicted 2018 Injury Crashes*	*Percentile*	*State Rank*	*Region Rank*	*County Rank*
85 & 2034	Signal	Salt Lake City	SALT LAKE	2	43	23	6	1.35	0.9998	1	1	1
209 & 2040	Signal	Salt Lake City	SALT LAKE	2	48	20	7	1.56	0.9996	2	2	2
8 & 3234	Signal	St. George	WASHINGTON	4	30	8	5	0.77	0.9995	3	1	1
189 & 3033	Signal	Provo-Orem	UTAH	3	57	17	6	1.39	0.9991	4	1	1
209 & 2108	Stop Sign	Salt Lake City	SALT LAKE	2	84	18	9	2.91	0.9991	5	3	3
111 & 2232	Signal	Salt Lake City	SALT LAKE	2	41	7	5	1.18	0.9982	6	4	4
71 & 2086	Signal	Salt Lake City	SALT LAKE	2	41	11	6	1.80	0.9963	7	5	5
79 & 126	Signal	Ogden - Layton	WEBER	1	84	28	9	3.67	0.9943	8	1	1
71 & 2254	Signal	Salt Lake City	SALT LAKE	2	44	18	5	1.48	0.9922	9	6	6
108 & Local	Signal	Ogden - Layton	DAVIS	1	58	12	8	3.27	0.9922	10	2	1
89 & 269	Signal	Salt Lake City	SALT LAKE	2	28	9	3	0.53	0.9916	11	7	7
201 & 2258	Signal	Salt Lake City	SALT LAKE	2	47	12	5	1.49	0.9909	12	8	8
71 & 2039	Signal	Salt Lake City	SALT LAKE	2	63	14	6	2.07	0.9904	13	9	9
71 & 2078	Signal	Salt Lake City	SALT LAKE	2	85	28	8	3.40	0.9872	14	10	10
39 & 3404	Signal	Ogden - Layton	WEBER	1	56	17	6	2.24	0.9855	15	3	2
89 & 2890	Signal	Provo-Orem	UTAH	3	40	10	5	1.67	0.9826	16	2	2
52 & 2986	Signal	Provo-Orem	UTAH	3	37	6	3	0.67	0.9824	17	3	3
89 & Local	Signal	Salt Lake City	SALT LAKE	2	42	14	5	1.65	0.9821	18	11	11
55 & 1338	Stop Sign	Small Urban	CARBON	4	2	2	2	0.28	0.9801	19	2	1
71 & 2262	Signal	Salt Lake City	SALT LAKE	2	39	10	2	0.28	0.9797	20	12	12

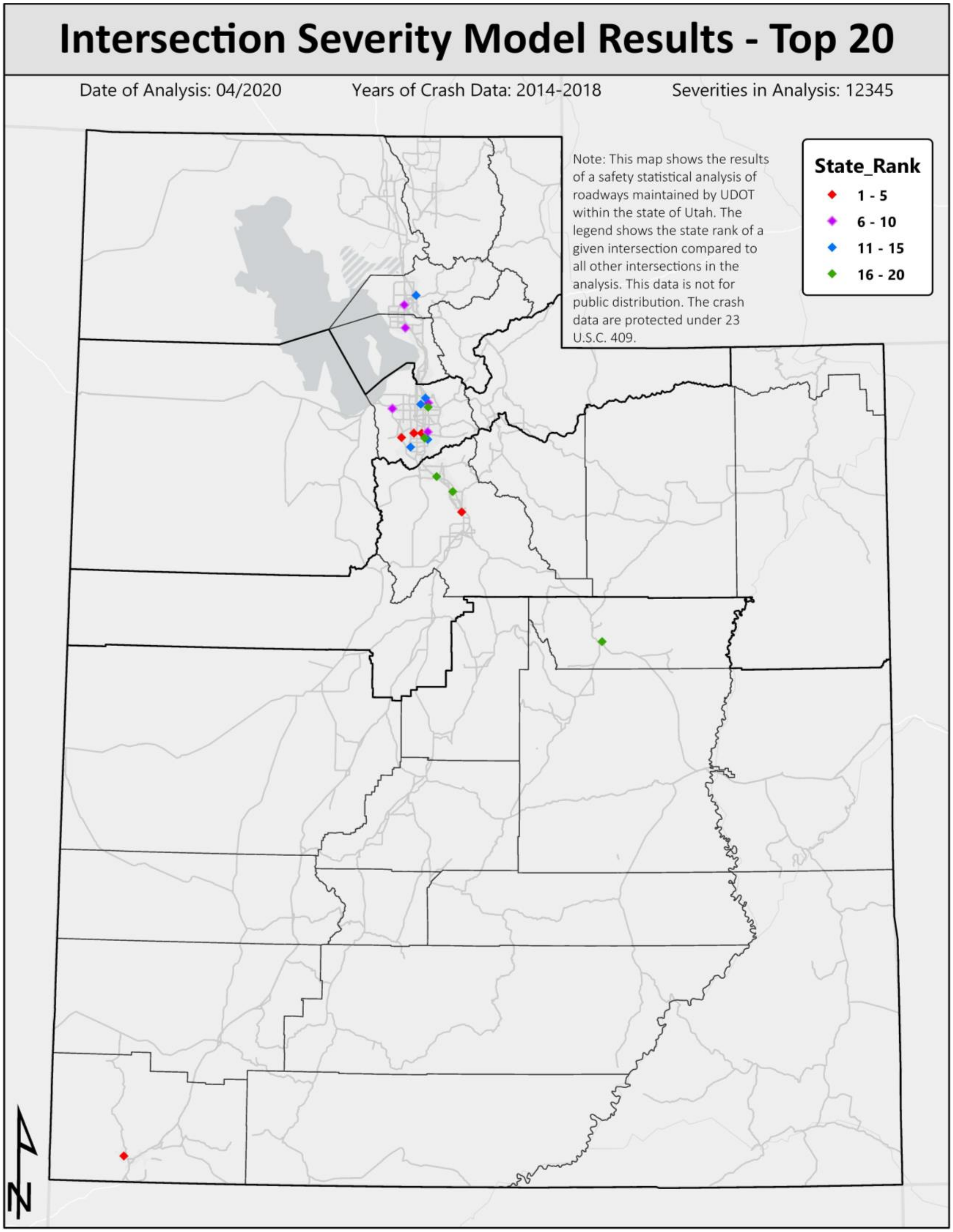


Figure 8-6: Map of UICSM top 20 intersections.

8.4 Summary

The changes discussed in this chapter allow for a more robust hot spot analysis of UDOT’s intersection network. While the building blocks of the ISAM remain intact between the 2018 and 2019 versions, the adjustments made to the methodology allow for a deeper analysis of intersection safety in Utah.

9.0 CONCLUSIONS

9.1 Overview

Every year, efforts are made to reduce the number of fatalities, injuries, and crashes on Utah roads. To aid engineers in selecting the sites most in need of attention and improvements, UDOT has teamed up with BYU in a series of safety-focused research projects. Previous to the CAMS were the RSAM, a methodology that looks at road segments as a whole, and the ISAM, a methodology that looks at intersections. After some modifications were made to the ISAM as discussed in Chapter 8, the intersection analysis became more robust in 2019. Because of these changes, the CAMS and ISAM form a complementary pair of analyses that allow UDOT to focus on both intersection safety and segment safety without interference between the two distinct crash groups. The purpose of the CAMS is to provide a methodology that looks at road segments without influence from intersection crashes. This chapter reviews the CAMS methodology, discusses future research topics, and ends with concluding remarks.

9.2 CAMS Methodology

The CAMS uses Microsoft Excel VBA coding as well as R coding to accomplish its purpose. The overall process has three steps. First, existing UDOT data are taken and used to create spreadsheets of roadway segments and corresponding crash data. Next, these segments are analyzed with a hierarchical Bayesian statistical model that ranks the segments in order of highest risk based on crash history and segment conditions. Finally, two-page technical reports are created for high-priority segments as determined by the statistical model. This section reviews these three steps.

9.2.1 Data Preparation

There are six files of roadway data and four files of crash data used in the data preparation process. The roadway data files include information on AADT, functional classification, location of intersections, number of lanes, speed limit, and urban code pertaining to the entire state route network in Utah. The crash files contain information on crash circumstances and location, crash factors, and manners of collision.

A flow chart for the overall data integration for the CAMS is shown in Figure 9-1. The top left portion outlined in orange shows the process of combining crash data. VBA macros are programmed to first sort through the data, removing any crashes that didn't occur on a state route. They are programmed to then identify crashes that occurred at significant intersections (state route intersections with another state route, a federal aid route, or a signal) and remove associated crashes from the data. Finally, the entirety of the crash data is merged together by a macro programmed to match unique crash IDs.

The bottom left portion of Figure 9-1 outlined in black shows the process of combining the roadway data. The five data files are combined using VBA macros that match roadway characteristics to common segments using beginning and ending milepoints provided in each of the files. These segments are homogeneous, meaning that for the length of each compiled segment, all five characteristics (AADT, functional class, number of lanes, speed limit, and urban code) remain constant. Neighboring segments vary in at least one of the five characteristics.

The top right portion of Figure 9-1 outlined in red shows the process of combining the crash data with the roadway data. During the process of assigning crashes to segments with VBA macros, two files are created. One file is termed the Input file and is used in the statistical model. It is a list of each segment with a tally of crash counts and characteristics. A sample of this data file was given previously in Figure 4-10. The second file is termed the Parameters file. It is a list of each crash with a reference to the specific segment on which it occurred. This file contains crash factors and the manner of collision for each individual crash, and it is used to fill out the two-page technical reports. A sample of the Parameters file was given previously in Figure 4-11.

9.2.2 Statistical Model

Two models were chosen to be used with the CAMS data: a crash prediction ZIP model (the CAMS-P model) and a crash severity model (the CAMS-S model). The CAMS-P model identifies segments that have significantly higher injury crash counts for a user-specified injury range than predicted, and the CAMS-S model identifies segments that have significantly higher proportions of injury crashes than expected based on all crashes on the segment.

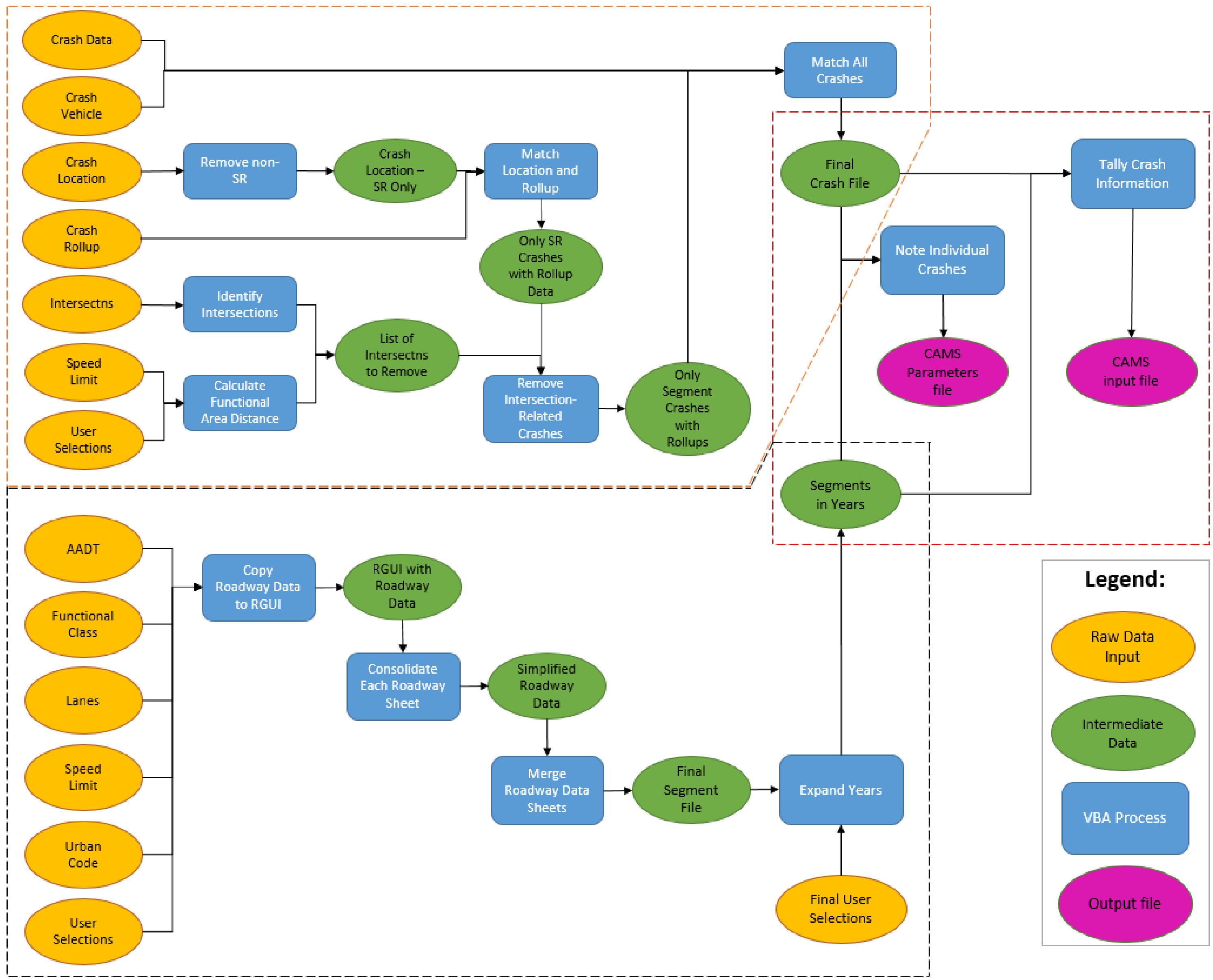


Figure 9-1: CAMS process flowchart.

In the development of the CAMS-P ZIP model, two other options for models were considered. Diagnostic reports indicated that all three of the models would be appropriate to use on the CAMS dataset, so it was decided to continue to use the ZIP model. The variables used in this model are speed limit, number of lanes, percent of trucks, and the natural log of VMT. This model is also hierarchical, meaning that it allows for the effects of the variables to vary across another parameter (in this case, urban code). The results of this model were provided previously in Table 7-1 and indicate that the segments on which to focus safety improvement efforts are spread throughout the state.

Similar to the CAMS-P model, the CAMS-S model uses the following variables: speed limit, number of lanes, percent of trucks, and VMT (no natural log transformation used). The results of this model were provided previously in Table 7-2 and indicate different patterns than found in the CAMS-P model results. The CAMS-S model results are more concentrated in Salt Lake County and only five of the top 20 segments are located in UDOT Region 3 or 4.

9.2.3 Two-Page Technical Reports

The final portion of the CAMS is the creation of two-page technical reports called SSARs. These reports are created for the ten highest-priority segments in each UDOT Region for both the CAMS-P and the CAMS-S model. The purpose of the SSARs is to give UDOT engineers a quick summary about the possible safety concerns on the selected segments.

VBA macros were used to automate the majority of the SSAR creation process. The macros are programmed to compile data from the Parameters file and the output of the statistical model into tables on the reports. These tables contain information ranging from a summary of the segment characteristics to lists of the most common crash types. Each table was discussed previously in Sections 6.3.1.1 and 6.3.1.2. The SSARs also contain sections for the research analyst to describe the current and historical conditions of the segment observed through virtual site visits. They also include a figure depicting the street-view of each segment and a map showing the location and surroundings of the segment. Finally, the SSARs contain two lists of possible countermeasures for the safety problems observed on the segment; one lists engineering countermeasures, and the other lists policy and enforcement countermeasures.

9.3 Future Research Topics

New opportunities for further research into topics on traffic safety in Utah have surfaced during the development of the research presented in this report. This section lists and briefly discusses these possibilities for future research, including alternate statistical analyses, applying Bayesian statistics in a spatial environment, implementing the R Shiny app, developing safety performance functions, weighting crashes by severity according to economic impact, identifying typical crash factor and manner of collision counts, and a summary of other possible topics.

9.3.1 Alternative Statistical Analyses

The statistical analyses used in this research are not the only methods available to analyze the data. There are changes that could be made, even while continuing to use a ZIP model. For example, the analysis could be run five times; for each time, four out of five years of crash data could be used to build the model and the excluded year could be compared against the predicted distributions. The average of the percentiles from each of the five years could be used to create the final rankings. Another example would be to increase the number of years in the analysis; eight years could be used, with four years to build the model and the average of the other four years to compare against the predicted distributions.

The advantage to these alternatives is their ability to account for regression to the mean. Using more than one year of crash data to compare against the predictions avoids flagging an intersection that had one bad year when the rest of its crash history is typically better than average.

9.3.2 Bayesian Statistics in a Spatial Environment

Spatial crash patterns may exist and provide unique insights. BYU Statistics performed research for the FHWA on freeway design and safety that included a portion connected to spatial process modeling (Christensen 2017). Performing spatial modeling on the CAMS data was preliminarily explored in 2019. A description of the findings can be found in the BYU Statistics research paper titled “Hot Spot Identification Analysis for Utah Roadways Using Spatial Poisson Linear Mixed Model” (Davis 2019).

9.3.3 Implementation of R Shiny App

R Shiny is a web-based app for running statistical programs. Having the statistical models for the RSAM, ISAM, and CAMS online rather than on individual computers would allow for BYU and UDOT to run the models without having to download any software or without using a specific computer. The feasibility of this was tested in 2019, and there is currently a beta version available at BYU for researchers to use. In addition to running the models, the R Shiny app can be made to produce charts, figures, tables, and maps.

9.3.4 Development of Safety Performance Functions

The HSM suggests that Safety Performance Functions (SPFs) be made for hot spot analyses. This idea is valuable because SPFs can be shared with other states or agencies in collaboration for better highway safety across the nation. Current research in Utah analyzes locations using predicted crash distributions instead of SPFs. It would be interesting to convert the methodologies in the RSAM, ISAM, and CAMS into SPFs. One big challenge of this would be to account for variability in the data.

9.3.5 Weighting Crashes by Severity According to Economic Impact

In the state of Utah, fatality crashes are assumed to have a significantly greater monetary impact on society than all other crash severities (Saito et al. 2018). Currently, the only differentiation between crash severities in the prediction models of the RSAM, ISAM, and CAMS is that injury crashes are used to make up the dataset and non-injury crashes are left out. It would be interesting to create a hot-spot identification model that weighted crashes differently based on their severity.

9.3.6 Typical Crash Factor and Manner of Collision Counts

As discussed in Section 6.3.1.2, the two-page technical reports provide tables summarizing the top eight crash factors and top nine manners of collision at each identified segment. It would be interesting to compare these values to average values of crash factors or manners of collision across the state. It would also be interesting to perform hot spot analyses on specific crash factors or manners of collision.

9.3.7 Other Related Topics

There are many other topics related to this research that could be explored. The CAMS could be modified to evaluate the two directions of divided highways separately (currently, the CAMS only divides the interstates and Mountain View Corridor, yet there are additional divided highways throughout the state). Another idea could be to create an analysis that focused on rural routes; this analysis could include additional variables such as roadway curvature to help analyze the segments. A third idea could be to analyze the potential impact of roadway changes on the crash severity and manner of collision distributions; this type of analysis could take into consideration variables such as congestion and traffic control devices.

9.4 Concluding Remarks

The CAMS provides a new way for the state of Utah to identify and prioritize segment safety improvement projects. By removing intersection-related crashes at significant intersections (state route intersections with another state route, a federal aid route, or a signal), it is possible to determine locations that are seeing more and higher-severity crashes than similar sites around the state. Identifying and ranking these segments makes it easier for UDOT to focus their efforts and budget on projects that are of highest concern across the state and that show a potential for significant improvement.

REFERENCES

American Association of State Highway and Transportation Officials (AASHTO). (2010). *Highway Safety Manual* (1st ed.), Washington, DC.

Borsos, A., Ivan, J. N., and Orosz, G. (2016). Development of Safety Performance Functions for Two-Lane Rural First-Class Main Roads in Hungary. *Traffic Safety,* 87-100.

Cafiso, S., Dagostino, C., and Persaud, B. (2018). Investigating the Influence of Segmentation in Estimating Safety Performance Functions for Roadway Sections. *Journal of Traffic and Transportation Engineering (English Edition),5*(2), 129-136.

Cafiso, S., Graziano, A. D., Silvestro, G. D., Cava, G. L., and Persaud, B. (2010). Development of Comprehensive Accident Models for Two-lane Rural Highways Using Exposure, Geometry, Consistency and Context Variables. *Accident Analysis & Prevention, 42*(4), 1072-1079.

Christensen, A. (2017). "Want safe travels? Find freeways with these features" Brigham Young University Communications. https://news.byu.edu/ news/want-safe-travels-find-freeways-these-features (Feb. 15, 2019).

Davis, C. (2019) "Hot Spot Identification Analysis for Utah Roadways Using Spatial Poisson Linear Mixed Model." Masters project report, Brigham Young University Department of Statistics. https://github.com/MJHeaton/Spatial-UCPM (Feb. 17, 2020).

Esri. (2019). ArcMap 10.7. Accessed from Brigham Young University.

Federal Highway Administration (FHWA). (2017). "Model minimum uniform crash criteria (MMUCC), 4th edition definitions." https://safety.fhwa.dot.gov/hsip/spm/conversion_tbl/pdfs/kabco_definitions.pdf (Apr. 13, 2020).

Federal Highway Administration (FHWA). (2016). "Traffic Monitoring Guide." https://www.fhwa.dot.gov/policyinformation/tmguide/tmg_fhwa_pl_17_003.pdf (Apr. 11, 2020).

Gaweesh, S. M, Ahmed, I. U., Ahmed, M. M., and Piccorelli, A. V. (2017). Parametric and Nonparametric Approaches in Developing Crash Prediction Models for Rural Mountainous Freeways: A Case Study in Wyoming. *Transportation Research Board 2018 Annual Meeting*, 18-03528.

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

Google. (2019). Google Earth Pro for Desktop.

Google. (2020). "Google Maps." https://www.maps.google.com (Feb. 17, 2020).

Jiri, A., Jiri, S., and Krivankova, Z. (2016). How to Simplify Road Network Screening: Two Case Studies. *Transportation Research Board 2017 Annual Meeting*, 17-01987.

Kwon, O. H., Park, M. J., Yeo, H., and Chung, K. (2013). Evaluating the Performance of Network Screening Methods for Detecting High Collision Concentration Locations on Highways. *Accident Analysis & Prevention, 51*, 141-149.

Mandli Communications. (2020). "Roadview Explorer 5." http://168.178.125.102/utah/? (Feb. 17, 2020).

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

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.

Ogle, J. H., Alluri, P., Zhang, C., Rajabi, M., Sarasua, W., and Bendigeri, V. (2017). Methods to Define Homogenous Segments and Assign Crashes for Highway Safety Manual Applications. *Transportation Research Board 2018 Annual Meeting*, 18-03826.

Pande, A., Abdel-Aty, M., and Das, A. (2010). A Classification Tree Based Modeling Approach for Segment Related Crashes on Multilane Highways. *Journal of Safety Research*, 41(5), 391-397.

Pew, T. (2020). "Justification for Considering Zero-Inflated Models in Intersections Safety Analysis." Masters project report, Brigham Young University Department of Statistics.

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

Saito, M., Seat, C. J., Schultz, G. G., and Clegg, B. W. (2018). "Crash Severity Distributions for Life-Cycle Benefit-Cost Analysis of Safety-Related Improvements on Utah Roadways." Report UT-18.07, Utah Department of Transportation Traffic and Safety, Research Division, Salt Lake City, UT.

Schultz, G. G., Johnson, E. S., Black, C. W., Francom, D., and Saito, M. (2012). "Traffic and Safety Statewide Model and GIS Modeling." Report UT-12.06, Utah Department of Transportation Traffic and Safety, Research Division, Salt Lake City, UT.

Schultz, G. G., Farnsworth, J. S., Roundy, R., Saito, M., Reese, C. S., and Briggs, T. (2013). "Hot Spot Identification and 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., Seigel, 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.

Schultz, G. G., Gibbons, J. D., Saito, M., and Clegg, B. W. (2018). "Intersection Safety Analysis Methodology for Utah Roadways." Report UT-18.06, Utah Department of Transportation Traffic and Safety, Research Division, Salt Lake City, UT.

Utah Department of Public Safety (UDPS). (2020) "Crash Data and Statistics." https://highwaysafety.utah.gov/crash-data/ (Feb. 8, 2020).

Utah Department of Transportation (UDOT). (2017). UDOT Open Data Portal. http://data-uplan.opendata.arcgis.com/ (Feb. 17, 2020).

Utah Department of Transportation (UDOT) Traffic & Safety. (2017). "UDOT Intersection Coding Guide Version 2."

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

Zheng, L., Sun, J., and Meng, X. (2018). Crash Prediction Model for Basic Freeway Segments Incorporating Influence of Road Geometrics and Traffic Signs. *Journal of Transportation Engineering, Part A: Systems, 144*(7), 04018030.

Zero Fatalities. (2020). "Zero Fatalities: A Goal We Can All Live With." https://zerofatalities.com/ (Feb. 17, 2020).