RESEARCH

UDOT Keeping Utah Moving

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

Report No. UT-23.15

USING SEVERITY-WEIGHTED RISK SCORES TO PRIORITIZE SAFETY FUNDING IN UTAH

Prepared For:

Utah Department of Transportation
Research and Innovation Division

Final Report
September 2023

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 and Safety
- Robert Miles – UDOT Traffic and Safety
- Ivana Vladisavljevic – UDOT Traffic and Safety
- Travis Jensen – WCG/UDOT Traffic and Safety
- Dallas Wall – WCG/UDOT Traffic and Safety
- Clancy Black – WCG/UDOT Traffic and Safety
- Charles Allen – Parametrix/UDOT Traffic and Safety
- Samuel Runyan – BYU Civil and Construction Engineering

TECHNICAL REPORT ABSTRACT

1. Report No.	2. Government Accession No.	3. Recipient's Catalog No.
UT-23.15	N/A	N/A

4. Title and Subtitle USING SEVERITY-WEIGHTED RISK SCORES TO PRIORITIZE SAFETY FUNDING IN UTAH	5. Report Date September 2023 6. Performing Organization Code n/a
7. Author(s) Grant G. Schultz, Ph.D., P.E., PTOE; Tomas Barriga Aristizabal, EIT; Jace Ritchie; Richard L. Warr, Ph.D.	8. Performing Organization Report No. R0402367
9. Performing Organization Name and Address Brigham Young University Department of Civil and Construction Engineering 430 EB Provo, UT 84602	10. Work Unit No. 55786 15D 11. Contract or Grant No. 21-8977
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 and Period Covered Final Report Jun 2021 to September 2023 14. Sponsoring Agency Code n/a

15. Supplementary Notes

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

16. Abstract

Budgets for transportation improvements are limited so it is important for governments to focus on improving locations most in need of safety funding. The objective of the Two-Output Model for Safety (TOMS) is to provide the Utah Department of Transportation (UDOT) with a reliable method to prioritize safety improvements on state-owned roadways among the different regions. This research will improve the existing Crash Analysis Methodology for Segments (CAMS) and Intersection Safety Analysis Methodology (ISAM) being used to analyze crashes on Utah roadways. The scope of this project is improving on the existing CAMS and ISAM to work together within R, to incorporate segment and intersection severity in safety hot spot analysis, to develop overall severity distributions, and to develop limited recommendations and conclusions related to the research.

TOMS uses UDOT data to create a statistical input. Each segment is homogenous with respect to five variables: average annual daily traffic, functional class, number of through lanes, speed limit, and urban code. Intersections are provided as a separate dataset. In the statistical analyses performed on the data, five years of crash data (2016-2020) are used to determine a weighted risk score for segments and intersections of similar characteristics. Those segments or intersections with excess weighted risk scores are designated as crash hot spots. Two-page technical reports with road characteristics and crash data are created for the top 10 hot spots for segments and intersections in Utah. The reports are sent to UDOT where region engineers may review and determine which locations might be addressed.

17. Key Words	18. Distribution Statement	23. Registrant's Seal
Weighted risk score, hot spot identification, severity distribution, safety improvements	Not restricted. Available through: UDOT Research Division 4501 South 2700 West P.O. Box 148410 Salt Lake City, UT 84114-8410 http://www.udot.utah.gov/go/	N/A

19. Security Classification (of this report)	20. Security Classification (of this page)	21. No. of Pages	22. Price
Unclassified	Unclassified	90	N/A

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ACRONYMS

AADT	Average Annual Daily Traffic
AASHTO	American Association of State Highway and Transportation Officials
ALRS	Linear Referencing System
BYU	Brigham Young University
CAMS	Crash Analysis Methodology for Segments
CAMS-P	Crash Analysis Methodology for Segments – Prediction
CAMS-S	Crash Analysis Methodology for Segments – Severity
CSV	Comma-Separated Values
CRP	Continuous Risk Profile
CUTRK	Combination-Unit Track
EB	Empirical Bayes
EPDO	Equivalent Property Damage Only
EWRS	Excess Weighted Risk Score
GIS	Geographic Information System
HSM	Highway Safety Manual
ID	Identification Number
ISAM	Intersection Safety Analysis Methodology
ISAR	Intersection Safety Analysis Report
JSM	Joint Statistical Model
ML	Maximum Likelihood
PDO	Property Damage Only
PS	Peak Searching
PSI	Potential Safety Improvement
RGUI	R Graphical User Interface
RSAM	Roadway Safety Analysis Methodology
RTM	Regression to the Mean
SAR	Spatial Autoregressive
SMW	Sliding Moving Window
SPF	Safety Performance Function
SSAR	Segment Safety Analysis Report

SSM	Surrogate Safety Measures
SUTRK	Single-Unit Truck
TOMS	Two-Output Model for Safety
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
UTA	Utah Transit Authority
VBA	Visual Basic for Applications

EXECUTIVE SUMMARY

The goal of this research is to assist the Utah Department of Transportation (UDOT) in identifying hot spots along Utah roads and implementing appropriate countermeasures. A new model named the Two-Output Model for Safety (TOMS) was created. While previous iterations of this research had two different models for segments and intersections, TOMS is a singular model that uses roadway and crash data to identify segments and intersections for analysis.

The roadway and crash data files allow for an accurate analysis of crashes on Utah roadways. Roadway data is critical for characterizing segments and intersections, while the crash data is essential to assigning each specific crash to the correct portion of roadway. TOMS uses 16 data files to create the combined crash and roadway data input for analysis. These include 12 roadway files and four crash files. The TOMS compiling process begins with reading in the 16 data files, TOMS then prepares the data files so that the format and contents are consistent, and then two different processes occur. The first process is to segment roadways based on five variables: average annual daily traffic, functional class, lanes, speed limit, and urban code. The second process is to assign the physical characteristics of the roadway to each individual intersection. When the segments and intersections have been created, TOMS outputs a segments file and an intersection file. These two files are used for the statistical analysis and are the "two outputs" referenced by the name of the model.

With the assistance of the Brigham Young University statistics team, the segments and intersections are analyzed by both severity and total number of crashes occurring at the sites. An excess weighted risk score (EWRS) was developed to analyze the severity and number of crashes concurrently. The segments and intersections with the highest EWRS are marked as hot spots. The list of the top 10 hot spots by region for both segments and intersections are screened by UDOT, and subsequent reports are created from the screened list. A report compiler is then executed to create two-page reports which contain roadway and crash information organized in a manner that allows UDOT to see how many crashes are occurring at a site and the manner of collision for the crashes.

1.0 INTRODUCTION

1.1 Problem Statement

By combining roadway and crash data, both roadway segments and intersections can be analyzed across the state of Utah. These segments and intersections can then be evaluated by comparing the expected crash frequency to the observed crash frequency, and then improvements to those sites can be determined. Those areas with the largest difference between expected and observed crash frequency compared to similar sites are known as crash hot spots.

Crash hot spots are the subject of a study of a joint effort between the Utah Department of Transportation (UDOT) and the Brigham Young University (BYU) Civil and Construction Engineering and Statistics Departments. Previous studies on this subject have resulted in the emergence of two tools to assist in the ranking of crash hot spots and proposing potential countermeasures. These are the Crash Analysis Methodology for Segments (CAMS) (Schultz et al., 2020) and the Intersection Safety Analysis Methodology (ISAM) (Schultz et al., 2018; Schultz et al., 2020). These tools were used to perform analysis of all crashes on highways within the state of Utah, the CAMS focusing on segments and the ISAM focusing on intersections. The tools were developed so that crashes do not overlap between the two tools. Technical reports were created with the results from both tools and sent to UDOT's Traffic and Safety Division with roadway and crash information, as well as proposed countermeasures.

The methodologies for CAMS and ISAM each used two different statistical models, one accounting for total crashes named the Prediction model, another accounting for severity of crashes, named the Severity model. The two statistical models have been used to identify hot spots but some of those identified in one model would not be identified in the other. These methodologies did not provide a way to analyze crash count and severity concurrently.

To continue to assist UDOT in prioritizing safety budgets, major improvements to current methodologies were made within this research as well as the development of severity distributions to better identify crash hot spots.

1.2 Objectives

The first goal of this research is to improve the CAMS and ISAM to allow them to be more effective in identifying hot spots and implementing countermeasures. CAMS and ISAM were built within a Macro-Enabled Workbook using Visual Basic for Applications (VBA) code to compile the data necessary to perform the analysis. These two models were rebuilt and combined as the Two-Output Model for Safety (TOMS) in R. R is a programming language built for statistical computing that allows for more efficient and consistent data management resulting in this singular model. The creation of TOMS ensures tool improvements are incorporated including adjustments to hot spot analysis, changes to ranking procedure of sites, and modifications to the "intersection related" criteria using UDOT guidelines.

Another goal of this research is to develop segment and intersection severity distributions. Previously, the CAMS and ISAM models provided rankings based on total crashes and injury crashes separately. The new Joint Statistical Model (JSM) creates a single ranking that compares the number and severity of observed crashes to the expected crashes. The results from the JSM are used in developing segment and intersection severity distributions which are delivered to UDOT to assist in analyzing Utah roads.

UDOT provided the most recent roadway and crash data for this research including the biannual asset inventory data, and the real-time updates of crash data. Hotspot rankings use this data in TOMS and the JSM. Reports generated from the hot spot analysis will help the UDOT regions determine which high-risk areas need to be prioritized. Before the reports were submitted, they were screened by UDOT's Traffic and Safety Division, and only reports for suitable sites are delivered to the regions.

1.3 Scope

The scope of this project was to combine the existing CAMS and ISAM within TOMS, and also, combining the Prediction and Severity Models into the JSM that can analyze crash count and severity concurrently and provide segment and intersection severity distributions requested by UDOT. The data from the TOMS is input into the JSM, and the results are input into the report compiler to generate two-page reports for sites selected by UDOT to deliver to their respective regions.

1.4 Outline of Report

The body of this 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 details how data were acquired and used for the research.
- Chapter 4 outlines the data evaluation process including discussion on the R processes used to create input files for the statistical model.
- Chapter 5 gives a description of the statistical model used in the research.
- Chapter 6 describes the report compiler from which the technical reports for hot spot sites are generated as well as the process that creates them.
- Chapter 7 analyzes and summarizes the results of the research.
- Chapter 8 gives concluding remarks including a review of the TOMS 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 and evaluate the existing methods of network screening used for highway safety analysis. These methods provide insights into how to improve and expand network screening methods for Utah roadways. The first section of the literature review discusses the network screening steps given in the *Highway Safety Manual* (HSM) (AASHTO, 2010). The next section discusses the evaluation of segmentation methods. This is followed by a discussion on ranking sites with potential for safety improvement based on their severity ranking. Finally, the last section summarizes previous BYU-UDOT research efforts on segment and intersection safety analysis.

2.2 Network Screening

The HSM published by the American Association of State Highway and Transportation Officials (AASHTO), defines network screening as “the process to realize a reduction in crash frequency with implementation of countermeasures” (AASHTO, 2010). Network screening arises from the need to use limited funds for safety improvement as effectively as possible. It is also the first step in the roadway safety management process as described in the HSM and as illustrated in Figure 2.1. The roadway safety management process relies first on network screening to determine which sites need evaluation. After diagnosis, proper countermeasures can be selected and appraised to determine the financial cost of such implementation and limited funds can be properly allocated.

The BYU-UDOT model contributes to network screening because it is used to evaluate crash data on the entire network of Utah highways. The model can assist in other steps of the roadway safety management process but does not replace the work of transportation professionals in these areas. For example, the existing model provides a list of useful countermeasures for each of the top-ranked sites, but local traffic and safety personnel must determine the best countermeasures for a specific site. Furthermore, network screening can be performed automatically using modeling techniques whereas other steps in the roadway safety management process require more subjective review.

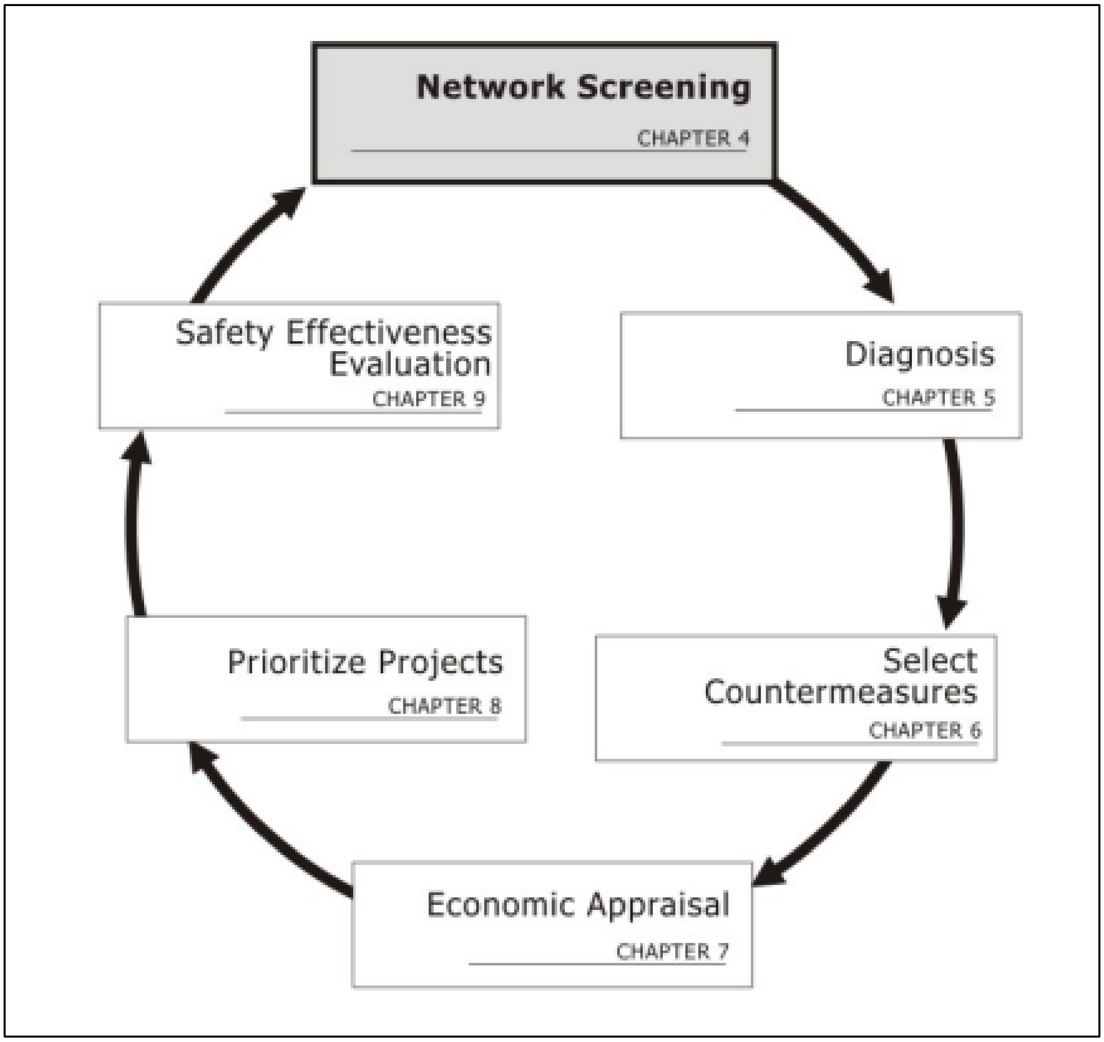

Figure 2.1 Schematic of Roadway Safety Management Process (AASHTO, 2010)

Network screening methodologies are often more generally called “hot spot identification” methodologies (Lyon et al., 2007; Sims and Somenahalli, 2010). In the literature, this term is used somewhat interchangeably with “network screening” to emphasize the purpose of identifying sites that stand out for their potential for safety improvement (PSI). These sites are often called “hot spots” or “black spots” as well as “sites with potential for safety improvement.”

The HSM outlines five steps which constitute network screening: 1) establish focus, 2) identify network and establish reference populations, 3) select performance measures, 4) select screening method, 5) screen and evaluate results (AASHTO, 2010). Each of these steps are described in the following subsections as they pertain to the project.

2.2.1 Establishing Focus for Network Screening

The first step of network screening, establishing the focus, involves defining which sites have PSI. The HSM defines these in two ways; sites that would benefit most from general countermeasures, and sites that would benefit from a specific countermeasure (AASHTO, 2010). The advantage of defining a specific countermeasure is that it allows transportation officials to use funds appropriated for a specific type of safety improvement (e.g., a rumble strip implementation program). The HSM does not give much specific instruction on how to do this, but some researchers have explored the possibility of identifying sites with a high proportion of specific crash types rather than total crashes. This method would be helpful for identifying sites that would benefit from specific countermeasures because of the assumption that specific countermeasures can prevent specific crash types. Unfortunately, due to regression to the mean (RTM), this method tends to return less accurate results since there is a reduced amount of data from counting only specific crash types. However, this may still be a viable option to include in a network screening model for those who want to determine where a specific crash type will be most prevalent.

2.2.2 Identifying Network and Establishing Reference Population for Network Screening

Step two of the network screening process involves separating crashes into networks based on whether they are intersection related or segment related. Previous research by BYU has explained this topic in detail, the results of which are included in this section (Schultz et al., 2020). 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 defines 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 it should not be the only criterion 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 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 Ambros et al. (2017) 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 Ambros et al. (2017) 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 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 marked as intersection related and 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: d1, d2, and d3—the respective lengths required for perception-reaction time, lane changing and deceleration, and queue length (Williams et al., 2014). The values for d1 and d2 were taken from tables in the Access Management Manual, 2nd edition, and the average queue length was assumed to be 50 feet for Utah state routes (Schultz et al., 2018).

2.2.3 Evaluating Performance Measures for Network Screening

The third step in network screening, selecting performance measures, is where the network screening model is chosen. Performance measures are the criteria by which a site is considered to have PSI. The PSI is the difference between the observed crash frequency and the expected crash frequency. Network screening is an essential first step to narrow down the data to a reasonable number of candidate sites for mitigation. Although there is some uncertainty in saying that sites with the highest PSI will have the highest potential for improvement, it is reasonable to assume that the higher-than-expected number of crashes may be related to something an engineer might address. However, there are multiple performance measures which evaluate the expected number of crashes in different ways. The HSM lists many of these performance measures which can be used in network screening (AASHTO, 2010).

Some of these performance measures don't require crash data, meaning they can be used for roads that have no crash data or inaccurate crash data. However, when crash data are available, the most popular performance measures in the literature are "expected average crash frequency with EB [Empirical Bayes] adjustment," and "excess expected average crash frequency with EB adjustment" (Gross et al., 2016). Additionally, some researchers have proposed performance measures not on this list which rely on spatial data. Among these are the "spatial autoregressive (SAR)" model (Gaweesh et al., 2019), and the "Getis-Ord Gi* statistic" which can be combined with a "random forest regressor" model (France-Mensah and O'Brien, 2018; Li and Al-Mahamda, 2020). These methods are explained in more detail in the source material. Furthermore, Sims and Somenahalli (2010), and Lyon et al. (2007) discuss using the "excess proportion of specific crash types" performance measure to determine sites where a specific countermeasure would be most beneficial. Some performance measures also weight crashes by severity which is a topic further discussed in Section 2.4 of this literature review.

One of the biggest problems addressed by many performance functions, and discussed extensively in the literature, is the issue with RTM bias. Because crashes are rare and random events, it is common for crash frequencies to fluctuate over time at a given site. This problem is more pronounced when only short-term crash data are considered instead of long-term crash data, however it is not overcome by long-term crash data because the issues that impact crashes change faster than the time of observation. Additionally, since crashes are rare, there are many instances where a road may have zero crashes. This has the tendency to distort normal distributions when predicting future crash rates. It is instead more useful to use distributions that account for “excess zeros” such as the zero-inflated Poisson distribution. Another issue presented in the HSM is determining an “acceptable” number of crashes for a roadway, to know which roadways need the most improvement. This number is called a “performance threshold” in the HSM because roadways which exceed this threshold are deemed too dangerous. Table 2.1 lists whether various performance measures account for RTM bias and whether they include a performance threshold since these are significant factors of how effective performance measures are.

2.2.4 Selecting Network Screening Methods

The fourth step of the network screening process, selecting the screening method, describes how performance measures are applied to a network of roadways. The literature discusses several network screening methods which can be applied to the model. Of these, the HSM specifically mentions site ranking, sliding moving window (SMW), and peak searching (PS). Kwon et al. (2013) investigated the effectiveness of each of these methods as well as an additional method, continuous risk profile (CRP), and determined that some methods are better for different kinds of roadways than others. SMW, PS, and CRP are only used for screening segments because they use different processes to subdivide segments into smaller parts to find where the crashes are most concentrated. However, little attention is given to SMW, PS, or CRP in the literature and simple ranking is generally the most accepted method used, even for segments. This means that entire segments are ranked based on the total number of crashes on that segment. It is important to determine segments on a roadway network which are relatively homogeneous. The process for defining segments is further explained in Section 2.3 of the literature review.

Table 2.1 Stability of Performance Measures (AASHTO, 2010)

Performance Measure	**Accounts for RTM Bias**	**Method Estimates a Performance Threshold**
Average Crash Frequency	No	No
Crash Rate	No	No
Equivalent Property Damage Only (EPDO) Average Crash Frequency	No	No
Relative Severity Index	No	Yes
Critical Rate	Considers data variance but does not account for RTM bias	Yes
Excess Predicted Average Crash Frequency Using Method of Moments	Considers data variance but does not account for RTM bias	Yes
Level of Service of Safety	Considers data variance but does not account for RTM bias	Expected average crash frequency plus/minus 1.5 standard deviations
Excess Expected Average Crash Frequency Using SPFs	No	Predicted average crash frequency at the site
Probability of Specific Crash Types Exceeding Threshold Proportion	Considers data variance; not effected by RTM Bias	Yes
Excess Proportions of Specific Crash Types	Considers data variance; not effected by RTM Bias	Yes
Expected Average Crash Frequency with EB Adjustments	Yes	Expected average crash frequency at the site
EPDO Average Crash Frequency with EB Adjustment	Yes	Expected average crash frequency at the site
Excess Expected Average Crash Frequency with EB Adjustments	Yes	Expected average crash frequency per year at the site

2.2.5 Screen and Evaluate Results

The final step of the network screening process, screen and evaluate results, refers to the process used to run network screening. Since state roadway networks generally constitute very large datasets, it is necessary to use computer programs to run the network screening model. In the case of the BYU-UDOT model, much of this analysis is done using the coding language, “R: A language and environment for statistical computing” (R), while the data management and report creation is done in Microsoft Excel. Moreover, the BYU-UDOT network screening process is further described in Section 2.5 of the literature review. Alternatively, it is possible to perform network screening using only R since R is capable of statistical analysis as well as data management and report creation. Similar coding languages like Python may also be used for modeling based on the user's preference.

2.3 Evaluation of Segmentation Methods

In addition to choosing an appropriate statistical distribution, it is important to separate crash data by roadway characteristics to create homogenous roadway sites. The process of separating by roadway characteristics is called segmentation. According to an investigation by Cafiso et al. (2018), the most significant factors that contribute to segmentation are the number of curves, length, and average annual daily traffic (AADT). The HSM recommends using AADT, number of lanes, curvature, lane width, shoulder width, median width and clear zone width for segmentation (AASHTO, 2010).

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). In this instance, the HSM is focusing on segmentation of a specific location with similar AADT prior to applying an established safety performance function; however, the concept is still applicable to other scenarios. In addition to the research cited in Table 2.2, Gaweesh et al. (2019) and Ogle et al. (2018) 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. (2020), 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: AADT, functional class, number of through lanes, and speed limit (Schultz et al., 2012). Beginning in 2013, the crash analysis research has included urban code as a fifth segmentation variable (Schultz et al., 2013).

Table 2.2 Variables Used in Homogenous Segmentation Methods

	Variables Used (CCR = Curvature 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	X				
Kwon et al. (2013)					X				X		
Schultz et al. (2020)	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

2.4 Severity-Weighted Hot Spot Analysis

One of the downfalls of most network screening methods is that they don't account for the significance of high severity crashes. Most sources in the literature classify crash severity according to the KABCO severity (FHWA, 2017). Therefore, UDOT has adopted a similar scale using values 1 through 5. The two scales are shown side by side in Table 2.3 along with the UDOT severity descriptions. Crashes which rank between K and B on the KABCO scale and 3 through 5 on the UDOT scale have a greater impact on society than other severities. Therefore, it is useful to weight these crash severities higher when performing hot spot analysis. Fortunately, a few of the performance measures listed in the HSM and shown previously in Table 2.1, as well as alternative methods, allow for some level of severity ranking in hot spot analysis. These include relative severity index, critical rate, and average EPDO crash frequency with statistical adjustment, as well as alternative methods using surrogate safety measures (SSMs) (Stipancic et al., 2019). Additionally, the BYU-UDOT research team developed two models called the Crash Analysis Methodology for Segments – Severity (CAMS-S) and the Utah Intersection Crash Severity Model (UICSM) which attempt to rank sites based on their proportion of severe crashes, (ranked 3-5), to overall crashes (Schultz et al., 2020). This section of the literature review is dedicated to exploring these hot spot analysis methods in greater detail as well as the statistical methods used for modeling crash severity.

Table 2.3 KABCO and UDOT Crash Severity Scales

KABCO	UDOT	Severity
K	5	Fatal Injury
A	4	Suspected Serious Injury
B	3	Suspected Minor Injury
C	2	Possible Injury
O	1	Property Damage Only

2.4.1 Evaluation of Crash Severity Models

In previous years, the BYU-UDOT team developed a crash severity model called the Utah Crash Severity Model (UCSM) for segments and the UICSM for intersections. This model evolved into the CAMS-S for segments and the updated UICSM for intersections but uses the same methodology for evaluating crash severity. 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?" This method works reasonably well for network screening, but it does not actually weigh crashes by their severity as recommended in the HSM. One method of severity weighting given in the HSM is to estimate the monetary cost associated with crashes of various severities. This is usually done by assigning an estimated cost of damages to each crash severity, and then assigning an EPDO score based on the relative cost to the average cost of property damage only (PDO) crashes. The HSM also discusses a few ways to model these weighted values.

One way to model EPDO crashes is to apply the predictive method described in the HSM. This is reliable for modeling EPDO crashes because it uses statistical regression to account for RTM bias. Therefore, the only thing which sets this method apart from the predictive method for total crashes is determining the weights of each crash severity. This is done by determining the average societal cost for each crash severity and dividing this with the average cost of PDO crashes. The actual weights for different crash severities may vary depending on location and the criteria used. Table 2.4 shows the weighted and unweighted crash costs given by UDOT for 2021 (UDOT, 2021a). This EPDO method has the advantage of giving a severity score to each roadway, but it has the disadvantage of ranking sites with a small number of severe crashes higher. This is the case for most severity-weighted ranking procedures, so it is valid to have separate models for total crashes and severity-weighted crashes.

Table 2.4 Updated Crash Costs for Use in Benefit Calculations (2020 Dollars)

Severity	Severity No.	Crash Cost Weighted	Crash Cost Unweighted	Unweighted EPDO
K (Fatal)	5	$ 3,078,500	$14,010,300	828
A (Sus. Serious Injury)	4	$ 3,078,500	$ 805,800	47
B (Sus. Minor Injury)	3	$ 264,000	$ 264,200	16
C (Possible Injury)	2	$ 148,000	$ 148,000	9
O (PDO)	1	$ 17,000	$ 17,000	1
KA (Severe)	5,4	$ 3,078,500	N/A	182
KAB (Injury)	5,4,3	$ 806,700	N/A	48
KABC (Anticipated Injury)	5,4,3,2	$ 415,100	N/A	24
KABCO (Total Crashes)	5,4,3,2,1	$ 134,300	N/A	8

2.4.2 Statistical Modeling of Crash Severity

Part of analyzing hot spots is establishing an expected baseline for similar segments of roadway. This allows for identifying areas that have a lot of potential for improvement while distinguishing between segments and intersections of varying length, volume, and other such uncontrollable factors that could otherwise inflate crash frequencies. However, modeling crash severity presents an interesting challenge (Yasmin and Eluru, 2018). Because crash severity in the KABCO model ranges from PDO to fatal injuries in a series of ordered categories, the modeling is best served by an ordinal model rather than a nominal model (Rudolfer et al., 1995).

There are numerous examples of ordinal models in the literature, and the Bayesian approaches are often used due to their ability to model spatially related data in a hierarchical way. The ordered logistic probit model is prominent among these models and has been used to great effect in modeling crash severity (Hou et al., 2022). Additionally, researchers found that using a generalized logistic probit model with a Leroux conditional autoregressive prior increased model efficiency when predicting crash data due to its increased ability to model spatial correlation (Zeng et al., 2021). Another approach to modeling ordered data is the nested logit, which essentially branches the prediction from general categories to specific ones (e.g., the first step for predicting a crash within the KABCO framework might be to distinguish if it would be either a “KA” crash or a “BCO” crash, then distinguishing if it was a fatal crash or a major injury crash). This has been used successfully in the literature at large (Chen and Shen, 2018), yet another way of predicting ordinal data lies in a model averaging approach that combines the two models to obtain a more robust and generally accurate model. Chen and Shen (2018) contend that model selection itself contains uncertainty, and that by averaging using a model averaging approach smooths that out somewhat and provides better estimates overall, which has been supported empirically. This approach adds a layer of complexity that is difficult to implement.

2.4.3 Using Joint Crash Count and Severity Models

When modeling crash counts and severity, a model that calculates both at the same time while introducing a correlation term between the models can be useful. Recently, such an approach to modeling has been used to great effect to improve upon previous models utilizing negative binomial crash count models and ordered logit fractional split severity models (Afghari et al., 2020; Yasmin and Eluru, 2018). This approach results in more accurate crash counts of both total and severity-specific crashes compared to common models that don’t use the correlation term (Afghari et al., 2020; Yasmin and Eluru, 2018). However, the proportions of severity crashes didn’t show significant differences, and might be worse than if the severity were calculated otherwise (Yasmin and Eluru, 2018). It is possible that other widespread models could be used in order to achieve higher levels of accuracy for severity proportions.

2.4.4 Using Crash Type in Severity Modeling

Sometimes it is useful to perform hot spot analysis with regards to crash type as well as severity. One useful application of modeling crash type comes from the first step of network screening, establishing the focus. The HSM explains this step as deciding whether to perform hot spot analysis for sites that would benefit most from general countermeasures, or sites that would benefit most from a specific countermeasure. The advantage of defining a specific countermeasure is that it allows transportation officials to use funds appropriated for a specific type of safety improvement (e.g., a rumble strip implementation program). Some researchers have explored the possibility of identifying sites with a high proportion of specific crash types to achieve this. This method is helpful for identifying sites that would benefit from specific countermeasures because of the assumption that specific countermeasures can prevent specific crash types. Unfortunately, due to RTM bias, this method tends to return less accurate results since there is a reduced amount of data from only counting specific crash types less (Gladence et al., 2015; Lyon et al., 2007; Srinivasan et al., 2016). However, modeling crash types is still a viable option for network screening and has other applications as well. For example, there is often a correlation between crash type and crash severity, as some crash types are more severe than others. This means that crash type is a useful variable to consider in severity models for network screening.

With regards to the statistical approach to modeling crash types, there are various options. However, within the Bayesian framework the model that is best suited for data with many unordered classes in general is logistic regression (Gladence et al., 2015). This model has been applied to many situations and has proven to be useful (Held and Holmes, 2006). Additionally, it is likely superior to a maximum likelihood (ML) approach due to the relatively small dataset of each intersection, which has proved to be problematic for ML since results are likely to be biased due to the scarcity of data points (Albert and Chib, 1993).

2.5 Previous Utah Safety Research

Among the UDOT-contracted research performed at BYU are two methodologies related to the present research: the updated ISAM and the CAMS (Schultz et al., 2020). The following sections will describe these parts, give background on these two methodologies, and detail their connection to the present research. More detail on this research can be found in the UDOT report published by Schultz et al. (2020).

2.5.1 Crash Analysis Methodology for Segments

CAMS was created in 2019 to model segment-related crashes. It is largely based on the Roadway Safety Analysis Methodology (RSAM) which was first developed by a BYU research team in 2016 and works in much the same way (Schultz, et al., 2016). However, while the RSAM analyzed the entire roadway network, CAMS specifically excludes intersection-related crashes thus only analyzing segment-related crashes.

The three parts of the CAMS 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.

2.5.2 Intersection Safety Analysis Methodology

First completed in 2018, the ISAM was developed to analyze and rank intersections with two or more state routes. It was updated in 2019 to include intersections between state routes and minor roads as well as using updated methods of intersection identification from UDOT. The general process is the same as the CAMS shown in Figure 2.2, except that segment-related crashes are excluded, and intersection-related crashes are analyzed (Schultz et al., 2020).

2.5.3 Data Preparation

The first part of the CAMS and ISAM was to prepare the data for statistical modeling. All the necessary data came from UDOT, most of which can be found on the UDOT Open Data Portal (UDOT, 2021b).

The data preparation was done with the use of VBA programming, a basic language used within Excel to create macros and organize data. Four crash data files (Crash Data, Crash Rollup, Crash Location, and Vehicles) were combined into one file. The Crash Locations file was used to identify which crashes occurred on a state route and all other crashes were deleted. Information from the three other crash files were then attached to the remaining crashes by matching crash Identification Numbers (IDs) across the files. These files include criteria about whether a crash is intersection related or not. These crashes were excluded from the CAMS while all other crashes are excluded from the ISAM.

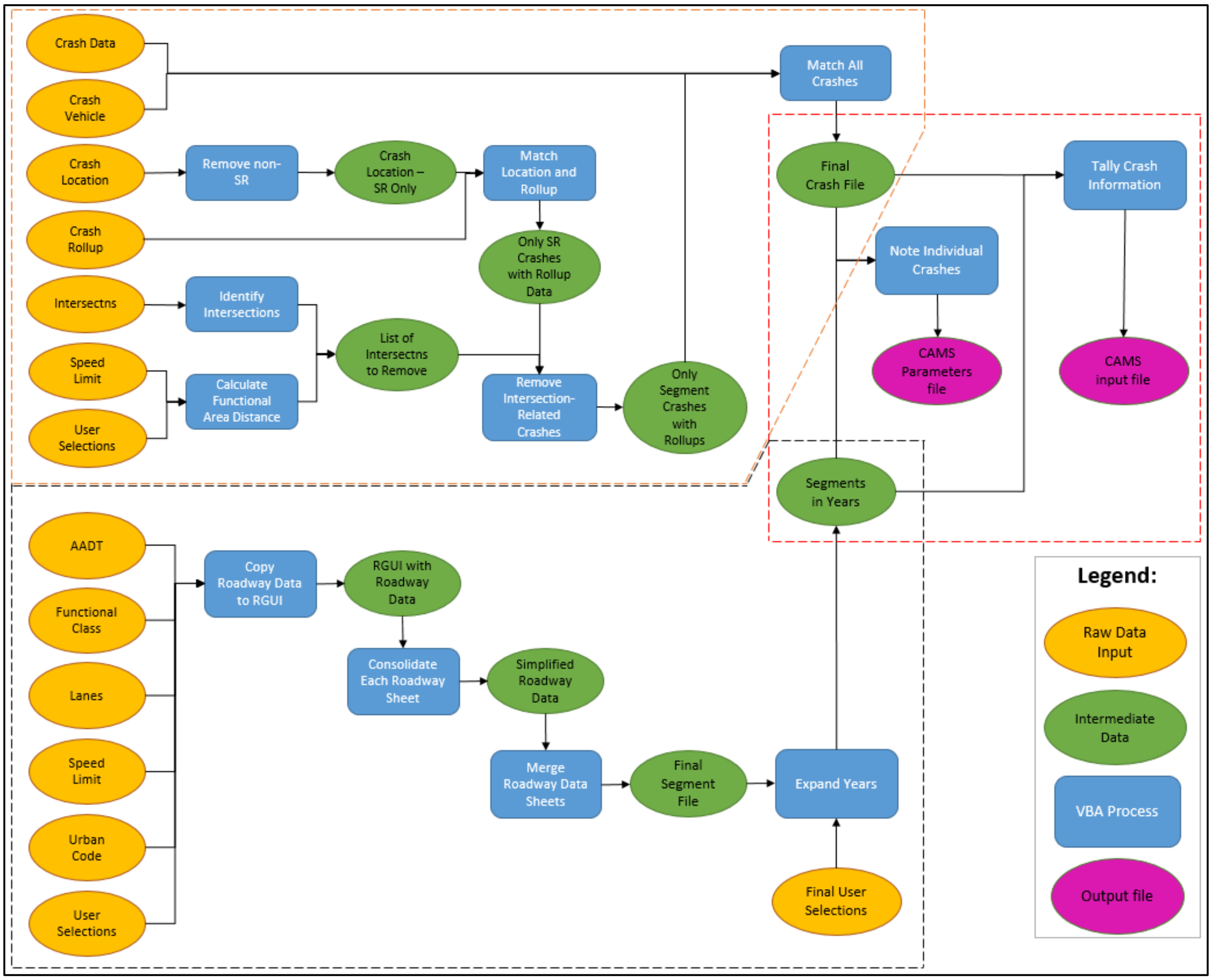

Figure 2.2 The CAMS process (Schultz et al., 2020)

In addition, the CAMS uses six roadway characteristic data files to modify the crash data, (AADT, State Route Functional Class, Intersections, Lanes, Speed Limit, and Urban Code). These files are important for the segmentation method used in CAMS. For the ISAM, eight roadway data files are added to the crash data, (AADT, State Route Functional Class, Intersections, Lanes, Pavement Messages, Speed Limit, Urban Code). The input form used to begin both the roadway data preparation and the crash data preparation processes is shown in Figure 2.3 for the CAMS and Figure 2.4 for the ISAM.

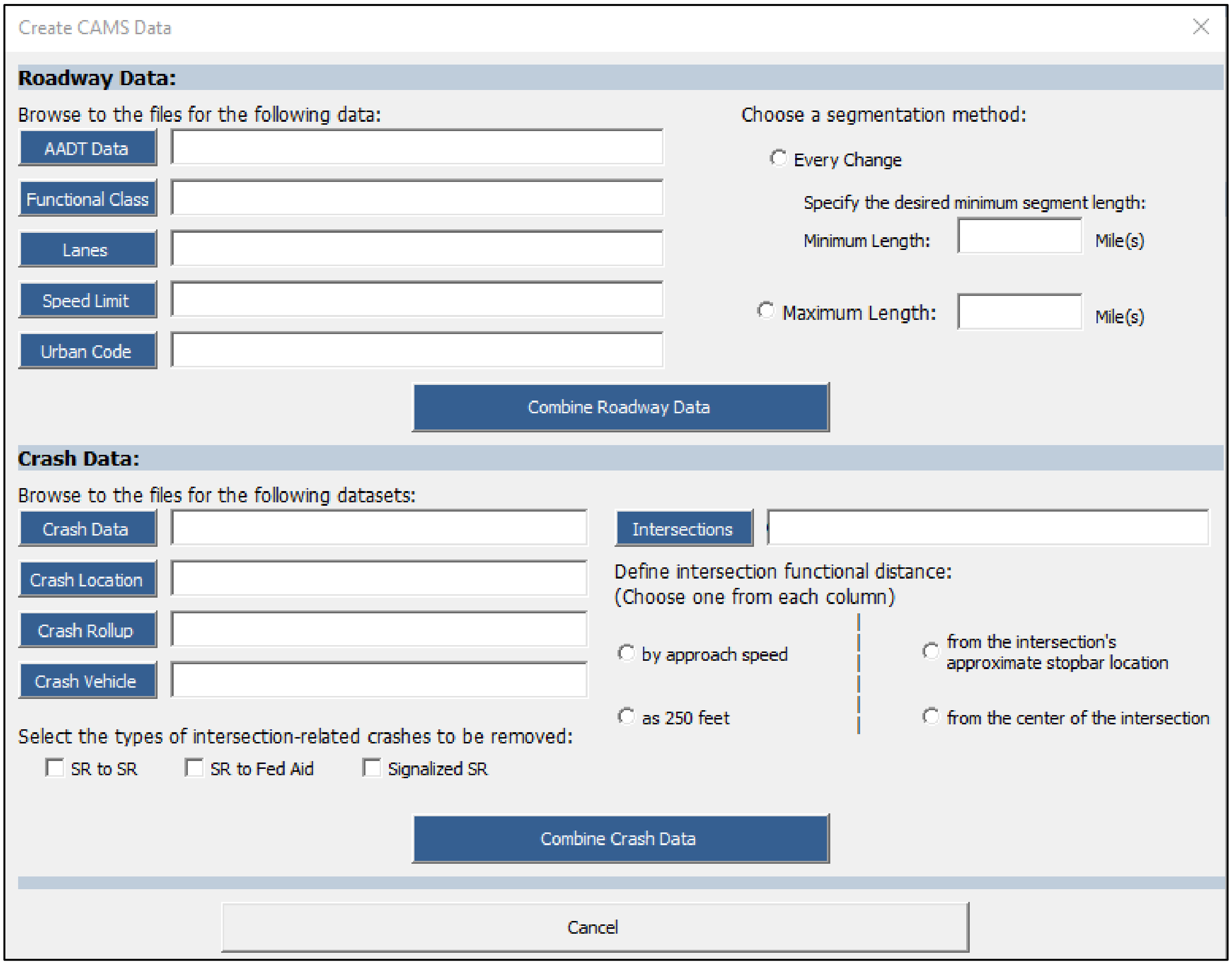

Figure 2.3 Input form for CAMS data preparation

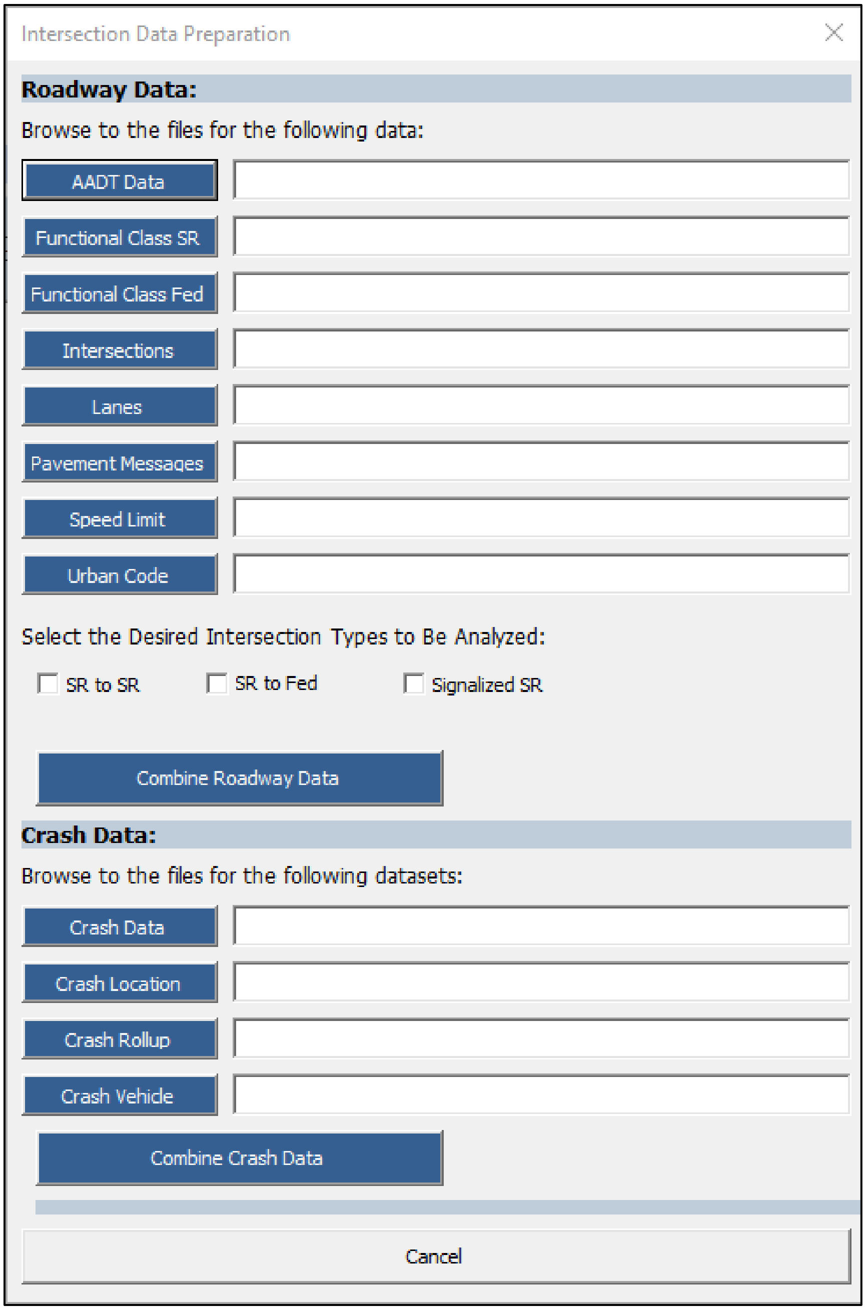

Figure 2.4 Input form for ISAM data preparation

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

2.5.4 Statistical Analysis

The second part of the CAMS and ISAM is to determine hot spots, or portions of the highway network that have observed significantly more crashes in a 5-year period than was predicted for that same time span. Four separate statistical analyses have been developed for this purpose: the Crash Analysis Methodology for Segments – Prediction (CAMS-P), the CAMS-S, the UICPM, and the UICSM. The CAMS-P and the UICPM predict how many crashes of specified crash severities (e.g., 3, 4, 5) are likely to occur at a segment or intersection, respectively, whereas the CAMS-S and the UICSM predict the number of injury crashes to occur at a segment or intersection based on the total number of crashes that occurred. Despite these differences, however, prediction and severity models have a lot in common. Both severity and prediction models take the same input (the detailed roadway information created in the data preparation process) and create predicted distributions of crashes for each segment or roadway. Furthermore, the observed number of crashes at each site is compared to the predicted distribution and associated with a percentile value within that distribution. The segments and intersections are then ranked according to the percentile values with a higher percentile value representing a greater safety concern. The resulting rankings are then used to determine which segments or intersections are of highest priority for safety improvements.

2.5.5 Technical Reports

The third part of the CAMS and ISAM is to create two-page technical reports for high-priority segments and intersections. These are called Segment Safety Analysis Reports (SSARs) for the CAMS and Intersection Safety Analysis Reports (ISARs) for the ISAM. For the CAMS, 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 SSAR. 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 SSAR and ISAR. These tables display information about roadway characteristics, as well as historical and current conditions of the site, and a list of potential countermeasures. Once the automated steps have been completed, research analysts then use individual SSARs and ISARs and perform virtual site visits using online tools to gather more information on the background and current conditions of each segment. The user can choose how many top-priority sites to generate SSARs and ISARs for. They can also be generated for specific UDOT regions, counties, or the entire state. In typical years, SSARs and ISARs from the top 10 intersections and segments in each UDOT region have been presented to UDOT for further evaluation.

2.6 Summary

Network screening for determining sites with potential for safety improvement involves selecting appropriate performance measures and works best when using an effective statistical model to account for RTM in crash data.

Sometimes it is helpful to account for the severity of crashes when performing network screening because higher severity crashes have a much more significant impact on society. The literature discusses many ways to do this. The ISAM and CAMS models only account for the ratio of high severity crashes to total crashes, they do not account for the weight of specific crash severities. The HSM discusses using EPDO to weight crashes by severity and allows for crashes of all severities to be considered. However, professionals focus on fatal and serious injury crashes more than other crashes. Roadways with low crash frequencies may be prioritized higher than roadways with higher crash frequency simply because they have more severe crashes. While this is still a useful measure of a roadway’s PSI, it should not necessarily dictate every safety focus.

The literature also discusses several methods of identifying intersection- and segment-related crashes. The most common method is to use a set distance from the intersection, but this distance varies between studies. Alternatively, intersection-related crashes can be determined by using the approach speed limit to determine the range around the intersection for intersection-related crashes. The literature also recommends using homogeneous segmentation so that roadways can be compared to other roadways with similar characteristics, although the variables used for segmentation varied within the literature.

In the past, BYU-UDOT research has created a model for total crashes (UCPM-UICPM) and for severe crashes (UCSM-UICSM). The UCSM-UICSM gives attention to severe crashes but does not weight them in comparison to less severe crashes. This literature review serves to find ways to improve the model so that it accounts for severity-weighted crashes. This review also identifies best ways to identify intersection-related crashes so that general improvements can be made to the BYU-UDOT model.

3.0 DATA COLLECTION

3.1 Overview

This chapter describes the various data files used, specifically detailing how they were acquired and how they were used within the research. There are two types of data used: roadway data which contains information on the physical characteristics of Utah roadways and crash data which contains information about all the crashes occurring on Utah roadways. Data relevant to the TOMS is used to distinguish specific roadway characteristics and create different segments or intersections, while data essential to the JSM assists in defining specific characteristics that similar segments and intersections share for the hot spot ranking. The following sections list in detail each of the 12 roadway files and four crash files used including the data they contain and how they were obtained.

3.2 Roadway Data

There are 12 roadway files used within the TOMS which are: Routes, AADT, Functional Class, Speed Limit, Urban Code, Lanes, Intersections, Medians, Driveways, Shoulders, Utah Transit Authority (UTA) Stops, and Schools. The UDOT Open Data Portal (UDOT, 2021b) provided seven of these files (Routes, AADT, Functional Class, Lanes, Medians, Driveways, Shoulders). The files are exported directly from the website reported on data from the new adjusted Linear Referencing System (ALRS) that UDOT implemented in the Summer of 2020. The Speed Limit, Urban Code, and Intersections files used in this research were provided directly from UDOT as they included additional data not accessible on the website. The Utah Geographic Information System (GIS) Data Portal provided the UTA Stops and Schools files. The 12 roadway data files will be discussed in more detail in the following sections.

3.2.1 Routes

The Routes data file is a list of all state and federal aid routes from UDOT. The data file includes information on the route name, direction, type, and beginning and ending mile points. The route files serve as a reference point for the other data files also on the ALRS, ensuring that the beginning and ending mile points of the routes match up. A supplemental file provided by UDOT provided an extensive list of divided state routes with the corresponding mile points.

3.2.2 AADT

The AADT data file is a list of the all state and federal routes and their respective daily traffic volumes along with percentages of single-unit trucks (SUTRK) and combination-unit trucks (CUTRK) along each roadway. The data file reports AADT data from 1981 to 2020 and SUTRK and CUTRK data from 2010 to 2020. This data file was updated in the Summer of 2022 and is on the ALRS. AADT is essential to the segmentation of roadways within TOMS as highlighted previously in Table 2.2.

3.2.3 Functional Class

The Functional Classification data file is a list of all the state and federal routes, the county the route is in, and a description of the grouping each route is in depending on their function. This data file was updated in the Summer of 2022 and is on the ALRS. Functional classification is essential to the segmentation of roadways within TOMS.

3.2.4 Speed Limit

The Speed Limit data file is a list of all the state and federal routes and the posted speed limit of each of the segments of the route. It was previously used to calculate functional area of intersections for the intersection models. This data file was provided by UDOT due to the updated data file on the ALRS not being available on the website. There were several missing speed limits along routes which were manually filled either from data provided directly from UDOT or using Google Streetview imagery (Google, 2023b). The speed limit is essential to the segmentation of roadways within TOMS.

3.2.5 Urban Code

The Urban Code data file is a list of all the state and federal routes within urban areas. It was provided directly from UDOT and is not accessible from the website. Utah Urban areas are as follows: Logan, Ogden-Layton, Provo-Orem, Salt Lake City, and St. George. In addition to the five urban areas the roadways may also be classified as small urban, rural, and unknown.

Each urban area has a unique five-digit code that is given to the roadways depending on their location. The urban code is essential to the segmentation of roadways within TOMS.

3.2.6 Lanes

The Lanes data file is a list of all the state routes and their lane configuration and count. It includes information for different lane types such as: auxiliary, through, deceleration, acceleration, turn, and passing. Along with this data it has the width of through lanes. This data file was updated in the Summer of 2022 and is on the ALRS. There were several missing counts of through lanes along routes which were manually filled either from data provided directly from UDOT or using Google Streetview imagery (Google, 2023b). The lanes data are essential to the segmentation of roadways within TOMS.

3.2.7 Intersections

The Intersections data file is a list of all the intersections on state routes. While data on the physical characteristics for non-state routes was limited, this data file did include the mile point of the intersection, traffic control type, UDOT region, and volume counts. UDOT requested that intersections be split up by signalized and unsignalized for the final rankings. Using the traffic control type, intersections were split up by signal type but all other data remained the same. Functional area for each intersection type was provided as well. The functional area and the “intersection involved” crash rollup field was used to determine which crashes were considered intersection related. This data file was provided directly from UDOT since the traffic count data were not accessible from the website. The intersection file is essential to developing the intersection model within TOMS.

3.2.8 Medians

The Medians data file is a list of medians and traffic islands on state routes as well as median/island type and length of median. This is a new data file to be included in the research and statistical model. This data file was updated in the Summer of 2022 and is on the ALRS. While not used for the segmentation of roadways due to the abundance of medians in the data file, these data are essential to the JSM.

3.2.9 Driveways

The Driveways data file is a list of driveways along state routes and their width. This is a new data file to be included in the research and JSM. Driveway type and width are not considered but rather the number of driveways on each segment of roadway. This data file was updated in the Summer of 2022 and is on the ALRS. While not used for the segmentation of roadways due to the abundance of driveways in the data file, these data are essential to the JSM.

3.2.10 Shoulders

The Shoulders data file is a list of shoulders on state routes and their position and width. This is a new data file to be included in the research and statistical model. This data file was updated in the Summer of 2022 and is on the ALRS. While not used for the segmentation of roadways due to the abundance of shoulders in the data file, these data are essential to the JSM.

3.2.11 UTA Stops

The UTA Stops data file is a list of all UTA routes, their stops, and average weekday ridership. This is a new data file to be included in the research and statistical model specifically for the intersection model. The data file has been most recently updated August 2022. The presence of a UTA route near an intersection is used to inform the JSM.

3.2.12 Schools

The Schools data file is a list of all the locations of preschool and K-12 schools in the state of Utah. This is a new data file to be included in the research and statistical model specifically for the intersection model. This data was updated for the 2019-2020 school year. The presence of a school near an intersection is used to inform the JSM.

3.3 Crash Data

There are four crash data files used within TOMS which are: Crash, Location, Rollups, and Vehicle. The Crash file will be referred to as the crash Severity file within this report to avoid confusion with other crash files. UDOT provided these files directly along with a unique crash ID for each file so that they can be joined for the analysis. While the unique crash ID pertains to an actual crash that occurred along Utah roadways this data was only used for crash safety analysis and not made available to the public. Each of the following subsections goes in detail about each of the specific crash files used.

3.3.1 Severity Data

The crash Severity data file is a list of crashes along with the manner of collision and crash conditions. Light, weather, roadway, junction, horizontal and vertical curves, and first harmful event are some of the attributes included in the file. Each of these attributes are assigned a code which corresponds to different conditions. This data file is essential for identifying the severity and manner of collision for each crash.

3.3.2 Location Data

The crash Location data file is a list of crashes along with their location. County, city, route, mile point, latitude longitude, and number of vehicles involved. This data file is essential to assigning each crash to a route or intersection and the corresponding UDOT region.

3.3.3 Rollups Data

The crash Rollups data file is a list of crashes along with specific details of the crash. Number of fatalities, number of injuries, pedestrian/pedacycle/motorcycle involvement, and intersection related are among the attributes included in the file. This data file is essential to determining primarily if crashes occur on segments or intersections as well as determining contributing circumstances to crashes.

3.3.4 Vehicle Data

The crash Vehicle data file is a list of each of the individual vehicles involved in the crash. Estimated speed, event sequence, travel direction, roadway description, and traffic control device description are among the attributes included in this file. This data file is essential for the report compiler.

3.4 Summary

The roadway and crash data files are all essential to TOMS and allow for a robust analysis of Utah roadway segments and intersections and how crashes occur on these roadways. While the roadway data is critical in the segmentation process and describing the characteristics of segments and intersections, the crash data allows each individual crash to be assigned to its corresponding roadway for analysis. The compilation of roadways and crash data is described in the next chapter.

4.0 DATA EVALUATION

4.1 Overview

This chapter describes how roadway and crash data from the previous chapter were screened and compiled to create an input file for the statistical model. In previous research, VBA macros were used within a Microsoft Excel Macro-Enabled worksheet referred to as the R Graphical User Interface (RGUI) workbook to merge all the data files. The RGUI in its entirety has been rebuilt in R, a free software environment for statistical computing. This chapter will describe the methodology within R as well as the required user input and processes used to compile the segment and intersection datasets and assign crashes to the datasets to create the input file for the statistical model.

4.2 Methodology in R

R is a free, open-source software that is a robust upgrade compared to the capabilities of Microsoft Excel. It allows for much more effective data handling and can integrate both comma-separated values (.csv) files and shapefiles into the same data frame. Since it is built around the R language, it can be used to perform complex custom-created functions which greatly reduce the total time of the compilation process and the total amount of code needed. Previously, two different inputs were required to create the segments and intersections input files. Those two files can be created simultaneously within R without the need for extensive user input. The statistical model has historically been built in R, thus the switch to R allows for a computational seamless transition between the two models as they are built in the same programming language.

4.3 TOMS Code

TOMS is set up within R with six different files, one of these files is referred to as an R Markdown file from which code can be executed in a specific order. The other five files are R Scripts, and within them, contain the specific code that is executed. These scripts are meant to run in series and not parallel. The order they were listed in is the order in which they were run.

The following subsections go into more depth on each script and the processes that each script contains. The five scripts are as follows: Functions, Read In, Roadway Prep, Crash Prep, and Compile.

4.3.1 Functions Script

The Functions script contains all functions that were created by the BYU team to conduct the data compilation. Any function that did not already exist within R packages is listed within this script and is essential to the TOMS. There are various types of functions that needed to be created for the TOMS, which include: data prep, segment cleaning, AADT, ALRS correction, segment creation, combine, and simple computational functions.

4.3.2 Read In Script

The Read In script reads in all the necessary data for the TOMS. For the roadway data, various datasets were formatted to make sure only the segments of state routes and intersections along state routes were being analyzed. Unwanted attributes from each dataset were removed and the route names and mile points were adjusted to the same format. Unwanted attributes included: date that the data were collected, metadata describing who collected the data, and attributes that could be found from other data. Duplicated data were removed as well and only unique segments and intersections from each data file were kept. For the crash data, the various datasets were formatted to make sure each crash had its own unique crash ID. Unwanted attributes from each dataset were removed. Much of the removed data were either duplicates or metadata that did not apply for the analysis.

4.3.3 Roadway Prep Script

Twelve data files were used to create the roadway data files used for analysis. The extensive list of those 12 are outlined in the previous chapter along with the specific attributes used from each one. The following subsections outline the R process for combining roadway files, and the subsequent files that are created from the process.

4.3.3.1 R Process

Figure 4.1 outlines the process to combine roadway files. The five data files used to create the homogenous segments and the file from which intersections were listed are individually noted, while supplemental data such as medians, driveways, shoulders, etc. are considered as "other data."

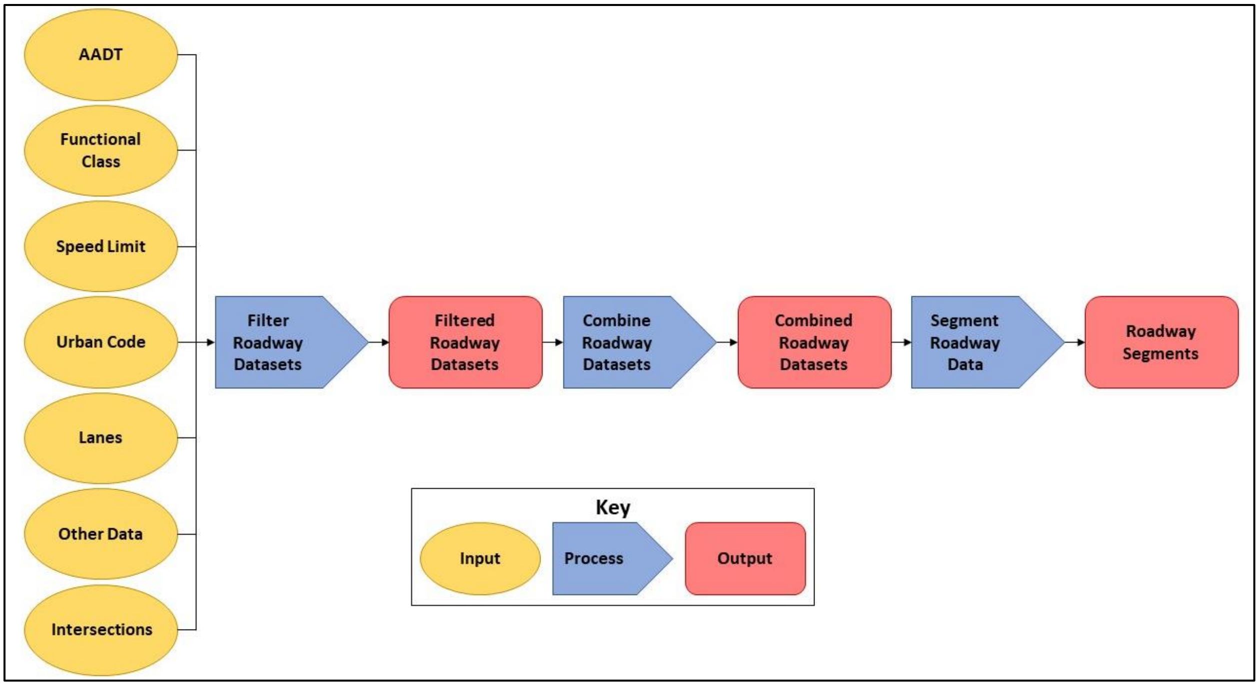

Figure 4.1 Outline of process to combine roadway files

This roadway process produced a data file for both segments and intersections simultaneously. All data were read in and formatted to be combined as detailed in Section 4.3.2. After the "Filter Roadway Datasets" process selected the appropriate data for the analysis the "Combine Roadway Datasets" process began. This process varied for segments and intersections and the "Segment Roadway Data" process only applied to segments.

For segments, the five principal segment datasets were merged to create unique segments for the analysis. Using the standardized formatting of state routes and mile points, the five principal datasets were combined into one. After the "Combine Roadway Datasets" process was finished, the "Segment Roadway Data" process began. If a variable from the five principal datasets changed along a segment then a new segment was created. Supplemental data were added afterwards and were assigned based on the mile points of the segments. Adjacent segments that were homogenous were combined to simplify the analysis. Segments whose total length was less than 0.1 mile were combined with the adjacent segment that was most similar. Any missing data were filled in either manually with corrections provided by UDOT or estimated using the interpolation of the data of adjacent segments. Once the segment roadway file was pivoted for the five-year period of AADT data it was finished and ready to be combined with the crash data.

For intersections, the intersections dataset was used and data from the five principal datasets used to create segments along with any other data were added to each of the intersections. The intersection file was merged with data from all other datasets to create the intersection roadway file. To facilitate assigning crashes to the intersection, an area of influence was determined by intersection type to provide a specific distance around intersections and crashes that fell within the distance would be assigned to the intersection. For example, any crash that occurs within 100 feet of an All Way Stop Control intersection would be assigned to that intersection. Table 4.1 provides a list provided by UDOT of the area of influence assigned to each intersection type.

Table 4.1 Area of Influence of Intersections

Intersection Type	Area of Influence (ft)
Signal Control	300
Minor Leg Stop Control	150
All Way Stop Control	100
Yield Control	100
Uncontrolled	100
Roundabout	300
Offset Left-Turn (CFI)	400
Median Thru-U Turn	400
R-Cut	400
SPUI	500
DDI	400
Active Transportation Only	100
Railroad Crossing	100

The area of influence was notated as Leg Distance within the data. Any missing data were filled in either manually with UDOT corrected data or estimated using data of intersections along the same route. After pivoting the intersections roadway file for the five-year period of AADT data, the file was finished and ready to be combined with the crash data.

4.3.3.2 Final Roadway File

The final roadway file for segments and intersections was formatted with five rows for each unique segment or intersection, with each row representing a different year of data. This allowed for crashes to be assigned to the respective site and year of where and when they occurred. Samples of the roadway files for both segments and intersections are given in Figure 4.2 and Figure 4.3, respectively.

SEG_ID	ROUTE	BEG_MP	END_MP	LENGTH_MILES	LENGTH_FEET	FUNCTIONAL_CLASS	RouteDir	RouteType	COUNTY_CODE	UDOT_Region	SPEED_LIMIT	THRU_CNT	THRU_WDTH	URBAN_CODE	YEAR	AADT	NUM_TRUCKS
1	0006NM	0	0.2	0.2	940.53	Other	N	State	Carbon	4	65	2	12	99998	2020	12744	3524
1	0006NM	0	0.2	0.2	940.53	Other	N	State	Carbon	4	65	2	12	99998	2019	13017	3546
1	0006NM	0	0.2	0.2	940.53	Other	N	State	Carbon	4	65	2	12	99998	2018	13351	2671
1	0006NM	0	0.2	0.2	940.53	Other	N	State	Carbon	4	65	2	12	99998	2017	13064	2667
1	0006NM	0	0.2	0.2	940.53	Other	N	State	Carbon	4	65	2	12	99998	2016	12846	3261
2	0006NM	0.2	3.3	3.2	16679	Other	N	State	Carbon	4	65	2	12	99998	2020	12644	3202
2	0006NM	0.2	3.3	3.2	16679	Other	N	State	Carbon	4	65	2	12	99998	2019	12915	3154
2	0006NM	0.2	3.3	3.2	16679	Other	N	State	Carbon	4	65	2	12	99998	2018	13246	2018
2	0006NM	0.2	3.3	3.2	16679	Other	N	State	Carbon	4	65	2	12	99998	2017	12953	1960
2	0006NM	0.2	3.3	3.2	16679	Other	N	State	Carbon	4	65	2	12	99998	2016	12738	2380

Figure 4.2 Sample of segment roadway file

Int_ID	INT_DESC	INT_TYPE	TRAFFIC_CO	Leg_Distan	SR_SR	STATION	REGION	BEG_LONG	BEG_LAT	BEG_ELEV	INT_RT_0	INT_RT_1	INT_RT_2	INT_RT_3	INT_RT_4	URBAN_CODE	YEAR	ENT_VEH	MEV	ENT_TRUCKS
6	4+ Leg Minor Stop	4-LEG	STOP SIGN	150	NO	4483	4411	-113.758	39.06	5301.237	0006PM	1893PM	0006PM	NA	NA	99999	2020	430	0.16	207
6	4+ Leg Minor Stop	4-LEG	STOP SIGN	150	NO	4483	4411	-113.758	39.06	5301.237	0006PM	1893PM	0006PM	NA	NA	99999	2019	415	0.15	200
6	4+ Leg Minor Stop	4-LEG	STOP SIGN	150	NO	4483	4411	-113.758	39.06	5301.237	0006PM	1893PM	0006PM	NA	NA	99999	2018	412	0.15	199
6	4+ Leg Minor Stop	4-LEG	STOP SIGN	150	NO	4483	4411	-113.758	39.06	5301.237	0006PM	1893PM	0006PM	NA	NA	99999	2017	409	0.15	197
6	4+ Leg Minor Stop	4-LEG	STOP SIGN	150	NO	4483	4411	-113.758	39.06	5301.237	0006PM	1893PM	0006PM	NA	NA	99999	2016	399	0.15	192
23	3-Leg Minor Stop	3-LEG	STOP SIGN	150	NO	4483	4411	-113.03	39.173	4658.806	0006PM	1904PM	0006PM	NA	NA	99999	2020	385	0.14	196
23	3-Leg Minor Stop	3-LEG	STOP SIGN	150	NO	4483	4411	-113.03	39.173	4658.806	0006PM	1904PM	0006PM	NA	NA	99999	2019	372	0.14	189
23	3-Leg Minor Stop	3-LEG	STOP SIGN	150	NO	4483	4411	-113.03	39.173	4658.806	0006PM	1904PM	0006PM	NA	NA	99999	2018	369	0.14	188
23	3-Leg Minor Stop	3-LEG	STOP SIGN	150	NO	4483	4411	-113.03	39.173	4658.806	0006PM	1904PM	0006PM	NA	NA	99999	2017	366	0.13	186
23	3-Leg Minor Stop	3-LEG	STOP SIGN	150	NO	4483	4411	-113.03	39.173	4658.806	0006PM	1904PM	0006PM	NA	NA	99999	2016	412	0.15	210

Figure 4.3 Sample of intersection roadway file

For segments, the beginning and ending mile points, segment length in miles and feet, functional class, route type, county, UDOT region, speed limit, through lane count, through lane width, urban code, AADT, and percent trucks was shown. These columns were taken directly from the five principal datasets shown previously in Figure 4.1. The segments are ordered by route number so that adjacent segments can be seen together as shown in Figure 4.2. The intersections file shows intersection ID, intersection description, intersection type, traffic control, leg distance, state route to state route, station, UDOT Region, longitude, latitude, elevation, primary route, secondary route, tertiary route, quandary route, quinary route, urban code, daily entering vehicles, million entering yearly vehicles, and daily entering trucks. Most of these columns came from the intersection file with the others coming from other datasets. The intersections are ordered by intersection ID. Both samples only show a fraction of the total data included in the file. Several other columns were present which were necessary for the final compiled file. Only the most important columns from the five major datasets are shown.

4.3.4 Crash Prep Script

Four data files are used to create the crash data files used for analysis. The four files are as follows: Severity, Location, Rollups, and Vehicle. They are outlined in the previous chapter along with the specific attributes used from each one. The following subsections explain the R process for combining the four crash files and explain the crash file for both segments and intersections.

4.3.4.1 R Process

Figure 4.4 outlines the process to combine crash files. The four data files used to create the complete crash file are individually noted.

This crash process produced a data file for both segments and intersections simultaneously. All crash data began with the "Filter Crash Datasets" process from the Read In script which is detailed in Section 4.3.2. The "Combine Crash Datasets" process takes the four filtered data files and merges them based on crash ID. Crashes are assigned to the segments file by route and mile point and to the intersections file by latitude and longitude. The two crash files for segments and intersections were not pivoted as the crash datetime columns would be used to assign crashes to the respective year in the analysis period (2016-2020) along whatever segment or intersection they were assigned. In previous iterations of the research, the models to create the segment and intersection files were done separately, as was the process to assign crashes to segments or intersections. Within TOMS, all the crashes can be assigned at once to either segments or intersections. UDOT has outlined intersection-related criteria which is an essential part of the methodology to determine whether a crash is occurring at a segment or intersection. The following section describes these criteria in more detail.

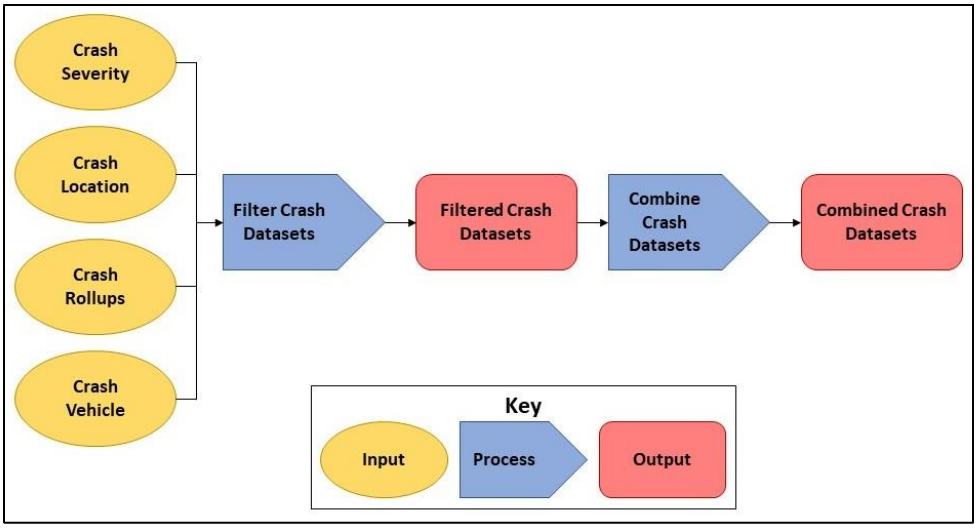

Figure 4.4 Outline of process to combine crash files

4.3.4.2 Intersection-Related Criteria

Crashes within the data have an attribute called “intersection related.” If a crash is marked “intersection related” and falls within the intersection area of influence detailed in Table 4.1, then that crash must be assigned to an intersection. Crashes that are not marked “intersection related” but fall within this area of influence are not assigned to intersections but are assigned to the segment in which they occur. Crashes that are marked “intersection related” but do not fall within the area of influence of an intersection within the data are assigned to segments. Using a spatial join with the area of influence serving as a buffer, crashes that have the “intersection related” attribute are assigned to intersections.

4.3.4.3 Final Crash File

The final crash file for segments and intersections was formatted in a way that each row represents a different crash. The year, location, severity, and number of vehicles involved is listed. These data were used to assign crashes to segments or intersections. Samples of the crash files for both segments and intersections are given in Figure 4.5 and Figure 4.6, respectively.

crash_id	crash_datetime	crash_severity_id	light_condition_id	weather_condition_id	manner_collision_id	roadway_surf_condition_id	roadway_junct_feature_id	horizontal_alignment_id	vertical_alignment_id	roadway_contrib_circum_id	total_number_roadway_lanes	first_harmful_event_id	first_harmful_evt_loc_id	motor_carrier_involved_yn	county_id	city	route	roadway_type	route_direction	ramp_id	milepoint	lat	long	seg_id
11129125	12/21/2018 14:50	1	1	2	2	1	0	1	2	0	6	20	1	N	35	SANDY	0015PM	M	P	0	293	40.54591391	-111.8949944	1001
11130128	12/21/2018 16:16	1	1	1	2	1	0	1	1	0	6	20	1	N	35	DRAPER	0015PM	M	P	0	289	40.48500748	-111.8976574	996
11130295	12/21/2018 16:45	1	6	1	4	1	0	1	1	0	2	20	1	N	35	SALT LAKE CITY	0080NM	M	N	0	117	40.76481353	-111.9654991	2863
11174064	12/21/2018 19:31	1	1	1	2	1	0	1	1	0	6	20	1	N	35	MIDVALE	0015PM	M	P	0	297	40.60456979	-111.9045797	1010
11137268	12/21/2018 22:15	1	3	4	4	3	0	1	2	3	2	20	1	N	43	OUTSIDE CITY LII	0080PM	M	P	0	190	41.1809535	-111.1320243	3013
11130133	12/21/2018 22:30	1	2	1	96	1	0	1	1	0	2	39	1	N	35	MURRAY	0015NM	M	N	0	301	40.66415882	-111.9015232	729

Figure 4.5 Sample of segment crash file

crash_id	crash_datetime	crash_severity_id	light_condition_id	weather_condition_id	manner_collision_id	roadway_surf_condition_id	roadway_junct_feature_id	horizontal_alignment_id	vertical_alignment_id	roadway_contrib_circum_id	total_number_roadway_lanes	first_harmful_event_id	first_harmful_evt_loc_id	motor_carrier_involved_yn	county_id	city	route	roadway_type	route_direction	ramp_id	milepoint	lat	long	int_id
11128323	12/21/2018 17:03	1	6	2	3	1	20	1	1	0	5	20	1	N	35	SOUTH JORDAN	0068PM	M	P	0	46	40.57315552	-111.9386459	3303
11132152	12/21/2018 17:55	1	3	1	2	1	20	1	1	0	2	20	1	N	35	TAYLORSVILLE	0173PM	M	P	0	7.8	40.65301553	-111.9295866	8125
11128687	12/21/2018 20:41	1	2	2	4	1	21	2	1	0	5	20	1	N	43	PARK CITY	0224PM	M	P	0	6.1	40.66039904	-111.5096562	9390
11128682	12/22/2018 9:07	2	1	1	96	1	20	1	1	0	5	22	1	N	11	CLEARFIELD	0193PM	M	P	0	4.2	41.1035277	-112.0069197	8721
11128684	12/22/2018 9:29	2	1	1	2	2	0	1	1	3	4	20	1	N	11	LAYTON	0126PM	M	P	0	1.8	41.07361259	-111.9797975	6609
11133205	12/22/2018 10:27	2	1	1	1	1	20	1	1	0	4	20	1	N	35	WEST JORDAN	0209PM	M	P	0	8.7	40.58775224	-111.9577947	9059

Figure 4.6 Sample of intersection crash file

The segment and intersection data file includes crash ID, crash datetime, crash severity, various contributing factors, county, city, route, roadway type, route direction, ramp ID, mile point, latitude, and longitude. The only major difference between the two files is the inclusion of either a segment or intersection ID. Crashes were only counted as being on a segment or intersection. Both samples only show a fraction of the total data included in the file. Several other columns are present which were necessary for the final compiled file. Only the most important columns from the four crash datasets are shown.

4.3.5 Compile Script

A total of 16 data files were used to create the combined segment and intersection data files used for analysis. The extensive list of those 16 are outlined in the previous chapter along with the specific attributes used from each one. The following subsections outline the R process used to combine the roadway and crash files and describe the final output file which serves as the statistical input.

4.3.5.1 R Process

Figure 4.7 outlines the process to combine the roadway and crash files. While the individual files used are not all noted, the previous figures provide a more detailed description of specific data files used.

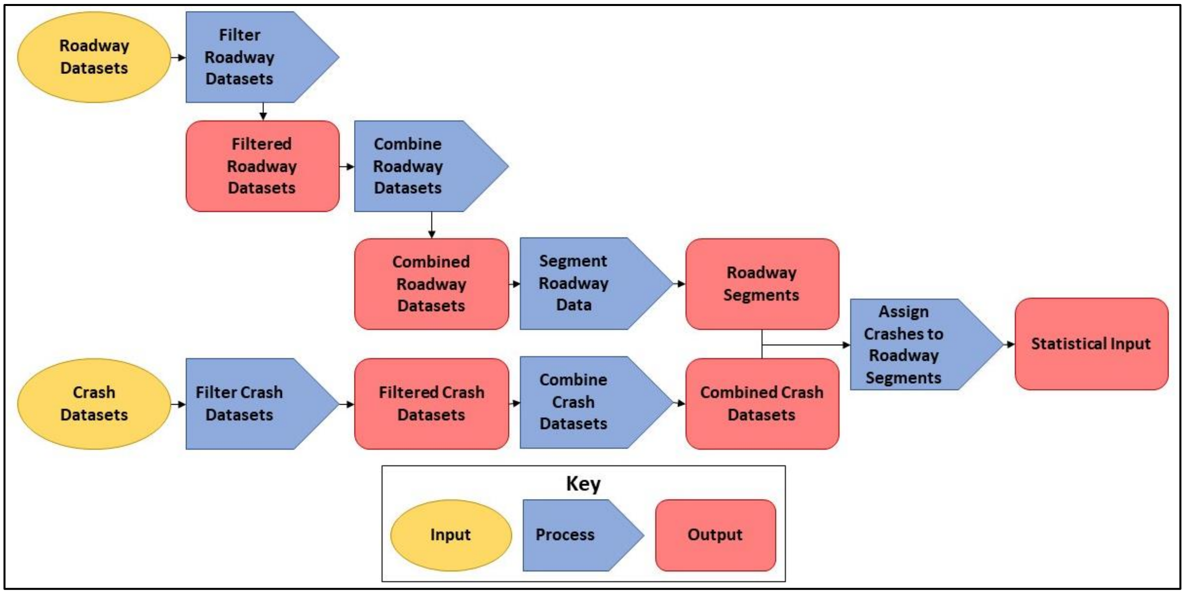

Figure 4.7 Outline of process to combine roadway and crash files

This final combining process produced a data file for both segments and intersections simultaneously. Figure 4.7 shows the entire process which includes much of the information shown previously in Figure 4.1 and Figure 4.4. The "Assign Crashes to Roadway Segments" process took the roadway files and crash files and the location data provided in them and assigned crashes to all segments and then all intersections. This part of the process only assigned location, severity, and vehicle data to each site. Crash attributes were added by looking at the crash IDs associated with each site and provided the number of crashes that contained those attributes. Crash attributes were assigned to the segments and intersections file. The two combined files were then exported into a .csv file that was used for the statistical analysis.

4.3.5.2 Final Combined File

The final combined file for segments and intersections was formatted in a way that there are five rows for each unique segment or intersection, with each row representing a different year of data. Crashes were assigned to the respective site and year of where and when they occurred. The attributes of the site were listed along with the total number of crashes, total crashes per severity, and crash characteristic data. Samples of the combined files for both segments and intersections are shown in Figure 4.8 and Figure 4.9, respectively.

While the roadway data for segments and intersections differs, the addition of crash columns is exactly the same for both. The total number of crashes for each severity along with total crashes for each year of the respective segment or intersection are shown. Both samples only show a fraction of the total data included in the file. Several other columns are present describing other contributing roadway and crash characteristics which were necessary for the statistical analysis. Only the most important columns from the five major datasets and four crash datasets are shown.

SEG_ID	ROUTE	BEG_MP	END_MP	LENGTH_MILES	LENGTH_FEET	FUNCTIONAL_CLASS	RouteDir	RouteType	COUNTY_CODE	UDOT_Region	SPEED_LIMIT	THRU_CNT	THRU_WDTH	URBAN_CODE	YEAR	AADT	NUM_TRUCKS	crash_severity_id_1	crash_severity_id_2	crash_severity_id_3	crash_severity_id_4	crash_severity_id_5	total_crashes
1	0006NM	0	0.2	0.2	940.53	Other	N	State	Carbon	4	65	2	12	99998	2020	12744	3524	0	0	0	0	0	0
1	0006NM	0	0.2	0.2	940.53	Other	N	State	Carbon	4	65	2	12	99998	2019	13017	3546	0	0	0	0	0	0
1	0006NM	0	0.2	0.2	940.53	Other	N	State	Carbon	4	65	2	12	99998	2018	13351	2671	0	0	0	0	0	0
1	0006NM	0	0.2	0.2	940.53	Other	N	State	Carbon	4	65	2	12	99998	2017	13064	2667	0	0	0	0	0	0
1	0006NM	0	0.2	0.2	940.53	Other	N	State	Carbon	4	65	2	12	99998	2016	12846	3261	0	0	0	0	0	0
2	0006NM	0.2	3.3	3.2	16679	Other	N	State	Carbon	4	65	2	12	99998	2020	12644	3202	4	1	1	1	0	7
2	0006NM	0.2	3.3	3.2	16679	Other	N	State	Carbon	4	65	2	12	99998	2019	12915	3154	0	0	0	0	0	0
2	0006NM	0.2	3.3	3.2	16679	Other	N	State	Carbon	4	65	2	12	99998	2018	13246	2018	3	1	1	0	0	5
2	0006NM	0.2	3.3	3.2	16679	Other	N	State	Carbon	4	65	2	12	99998	2017	12953	1960	4	0	0	0	0	4
2	0006NM	0.2	3.3	3.2	16679	Other	N	State	Carbon	4	65	2	12	99998	2016	12738	2380	1	0	0	0	0	1

Figure 4.8 Sample of compiled segment crash and roadway file

Int_ID	INT_DESC	INT_TYPE	TRAFFIC_CO	Leg_Distan	SR_SR	REGION	BEG_LONG	BEG_LAT	INT_RT_0	INT_RT_1	INT_RT_2	INT_RT_3	INT_RT_4	URBAN_CODE	YEAR	ENT_VEH	MEV	ENT_TRUCKS	crash_severity_id_1	crash_severity_id_2	crash_severity_id_3	crash_severity_id_4	crash_severity_id_5	total_crashes
6	4+ Leg	4-LEG	STOP SIGN	150	NO	4411	-113.758	39.06	0006PM	1893PM	0006PM	NA	NA	99999	2020	430	0.2	207	0	0	0	0	0	0
6	4+ Leg	4-LEG	STOP SIGN	150	NO	4411	-113.758	39.06	0006PM	1893PM	0006PM	NA	NA	99999	2019	415	0.2	200	0	0	0	0	0	0
6	4+ Leg	4-LEG	STOP SIGN	150	NO	4411	-113.758	39.06	0006PM	1893PM	0006PM	NA	NA	99999	2018	412	0.2	199	0	0	0	0	0	0
6	4+ Leg	4-LEG	STOP SIGN	150	NO	4411	-113.758	39.06	0006PM	1893PM	0006PM	NA	NA	99999	2017	409	0.1	197	0	0	0	0	0	0
6	4+ Leg	4-LEG	STOP SIGN	150	NO	4411	-113.758	39.06	0006PM	1893PM	0006PM	NA	NA	99999	2016	399	0.1	192	0	0	0	0	0	0
23	3-Leg N	3-LEG	STOP SIGN	150	NO	4411	-113.03	39.173	0006PM	1904PM	0006PM	NA	NA	99999	2020	385	0.1	196	0	0	0	0	0	0
23	3-Leg N	3-LEG	STOP SIGN	150	NO	4411	-113.03	39.173	0006PM	1904PM	0006PM	NA	NA	99999	2019	372	0.1	189	0	0	0	0	0	0
23	3-Leg N	3-LEG	STOP SIGN	150	NO	4411	-113.03	39.173	0006PM	1904PM	0006PM	NA	NA	99999	2018	369	0.1	188	0	0	0	0	0	0
23	3-Leg N	3-LEG	STOP SIGN	150	NO	4411	-113.03	39.173	0006PM	1904PM	0006PM	NA	NA	99999	2017	366	0.1	186	0	0	0	0	0	0
23	3-Leg N	3-LEG	STOP SIGN	150	NO	4411	-113.03	39.173	0006PM	1904PM	0006PM	NA	NA	99999	2016	412	0.2	210	0	0	0	0	0	0

Figure 4.9 Sample of compiled intersection crash and roadway file

4.4 Summary

Sixteen data files are used to create the combined crash and roadway data files used for analysis. Twelve contain roadway data while the other four contain crash data. The entire process is conducted from an R markdown file which requires that the following occur in a specific order: first, the functions are sourced; second, the data is read in; third, the data preparation script is run; finally, the compiling script is executed. Two output files are created from the process: a segments file, and an intersection file. These two files are for the statistical analysis portion which will be described in the next chapter.

5.0 STATISTICAL ANALYSIS

5.1 Overview

This chapter summarizes the statistical analysis to determine roadway hot spots in Utah. In previous research, two separate models, a prediction and severity model were used for the hot-spot ranking procedure. This current iteration uses an estimated monetary cost associated with crashes to account for the significance of high severity crashes on roadways. With the addition of an excess weighted risk score (EWRS) based on the unweighted crash cost, the two previous models in this current iteration are improved and combined. Within this chapter, the phrase "roadway site" will be used to refer to both segments and intersections. The count model looks at the total number of crashes occurring at a roadway site. The severity model looks at how many of the total crashes are severe at a roadway site. The chapter also discusses how the two models are used to rank roadway sites and the final output given to create the two-page reports.

5.2 Count Model

The count model is the current iteration of the previous prediction model. It is an improvement because the crash count and crash severities are jointly modeled. This count model is derived from the techniques described in the literature where a joint model was used to identify road segments with high risk of fatal and serious injury crashes (Afghari et al., 2020). The count model is a negative binomial regression model. The functions used within this model are shown in Equation 5.1.

$$\begin{gathered} X_{ij} \sim NB(\mu_{ij}, r) \\ \boldsymbol{P(X_{ij} = x) = \binom{x+r-1}{x} \left(\frac{\mu_{ij}}{\mu_{ij}+r}\right)^{x} \left(\frac{r}{\mu_{ij}+r}\right)^{r}} \\ \log(\mu_{ij}) = \mathbf{w_{ij}}\boldsymbol{\beta} + \mathbf{u_{ij}}\boldsymbol{\gamma} \end{gathered} \tag{5.1}$$

Where:

- i = Number of years;
- j = Roadway site (segment/intersection);
- X_{ij} = Total number of crashes at a site;

μ_{ij} = Mean number of crashes at a site;

r = Dispersion parameter;

$\mathbf{w_{ij}}$ = Attributes of the site used to estimate crash counts;

$\boldsymbol{\beta}$ = Average effects of the attributes in $\mathbf{w_{ij}}$ on the mean (subset by urban code);

$\mathbf{u_{ij}}$ = Attributes of the site used to estimate both crash counts and crash severity;

$\boldsymbol{\gamma}$ = Average effects of the attributes in $\mathbf{u_{ij}}$ on the mean.

The count model looks at roadway and crash attributes that contribute to the total number of crashes occurring at roadway sites with similar characteristics. The number of crashes for roadway sites with certain characteristics is predicted. The number of crashes is converted to an EWRS, and, using the total number of crashes that occur at the roadway site, the actual EWRS is compared. This comparison is used in the final ranking.

5.3 Severity Model

The severity model is an improvement of the previous severity model used. The severity model is an ordinal multinomial model. The functions used within this model are shown in Equation 5.2.

$$\mathbf{P}(\boldsymbol{Y_{ij}} <= \boldsymbol{y}) = \boldsymbol{\omega}(-\boldsymbol{\rho_{ij}} + \boldsymbol{\theta_{y}}) \tag{5.2}$$

$$\rho_{ij} = \mathbf{z_{ij}}\boldsymbol{\delta} + \mathbf{u_{ij}}\boldsymbol{\gamma}$$

Where:

i = Number of years;

j = Roadway site (segment/intersection);

Y_{ij} = Severity of a crash at a site;

ω = Standard cumulative logistic function;

ρ_{ij} = Latent variable used to estimate severity;

θ_{ij} = Adjustment to the intercept for severity y;

$\mathbf{z_{ij}}$ = Attributes of the site used to estimate crash severity;

$\boldsymbol{\delta}$ = Average effects of the attributes in $\mathbf{z}_{ij}$ on the latent variable.

$\mathbf{u}_{ij}$ = Attributes of the site used to estimate both crash counts and crash severity;

$\boldsymbol{\gamma}$ = Average effects of the attributes in $\mathbf{u}_{ij}$ on the latent variable (subset by urban code).

The severity model analyzes roadway and crash attributes that contribute to the different severities of crashes occurring at roadway sites with similar characteristics. The severity of crashes for roadway sites with certain characteristics is predicted. The crashes are converted to an EWRS which is used in the final ranking.

5.4 Ranking

The joint model is run for both segments and intersections using the statistical input created from combining the roadway and crash data. The sites are compared by their EWRS based on the joint count and severity model. The EWRS from the joint count and severity model are used and those roadway sites with the highest EWRS are ranked the highest.

5.5 Statistical Output

After the statistical analysis is complete, a list of all segments and intersections is created with the ranking of all sites and output as a .csv file. The lists of top 10 segments, signalized intersections, and unsignalized intersections by region are generated for UDOT. These lists can then be screened with reports generated as explained in Chapter 6.

5.6 Summary

Severity is an important factor when creating a hot spot analysis. Instead of using two different models, one accounting for total number of crashes and another for the total number of severe crashes, the two models were combined using an EWRS. All segments and intersections are analyzed using this method. Those with the highest EWRS are marked as hot spots that can be screened by UDOT as outlined in the next chapter.

6.0 REPORT COMPILER

6.1 Overview

This chapter goes over the report compiler that creates two-page reports of the top 10 hot spots for UDOT region engineers. While the report compiler itself has not changed from previous iterations of the research, it will be detailed within this chapter. It is noted that the feasibility of a Dashboard in R was explored as part of this research to replace the two-page reports. Although the Dashboard has potential for future applications, the Technical Advisory Committee chose not to pursue it further at this time. The report compiler uses the statistical input file and a parameters file to create roadway site reports that describe the crashes that occur at the site as well as possible countermeasures. The two-page reports are provided to UDOT Traffic and Safety and then forwarded on to the regions. The following sections describe the parameters file, the report compiler, the two-page reports for both segments and intersections, and each page of the reports for both.

6.2 Parameters File

The parameters file was previously created alongside the combined roadway and crash files within the CAMS and ISAM (Schultz et al., 2020) in VBA but now in R. The parameters file is created after the statistical analysis to assist in the creation of the two-page reports. The primary difference between the statistical output and the parameters file are that while the statistical output includes all the roadway characteristics and crash statistics, the parameters file includes specific details of each crash that occurred at each roadway site. This parameters file is somewhat redundant as within R all the data could be provided in one file. Since the report compiler was unchanged, the parameters file is necessary to create the two-page reports.

6.3 Report Compiler

As previously stated, the Report Compiler was not updated and is accessible as a Macro-Enabled Workbook. Within this workbook, the creation of reports for both the top ranked segment and intersection sites can be completed. Figure 6.1 shows the main sheet from which the report compilation process is conducted.

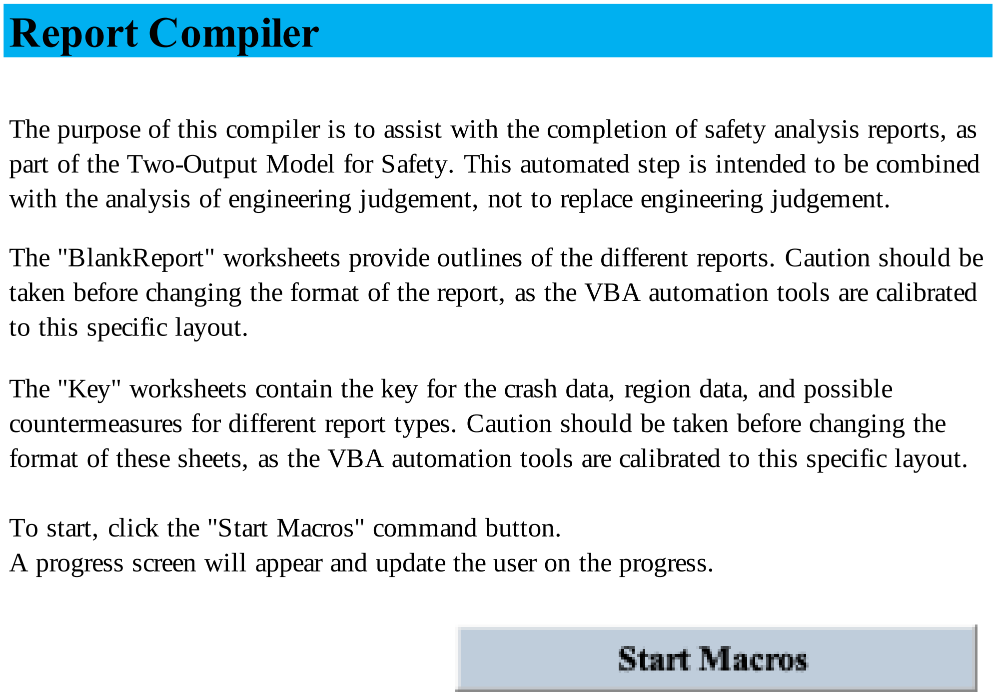

Figure 6.1 Report compiler main sheet

The "Start Macros" button leads to the prompt shown in Figure 6.2 which allows for the selection of report type whether that be for segments or intersections.

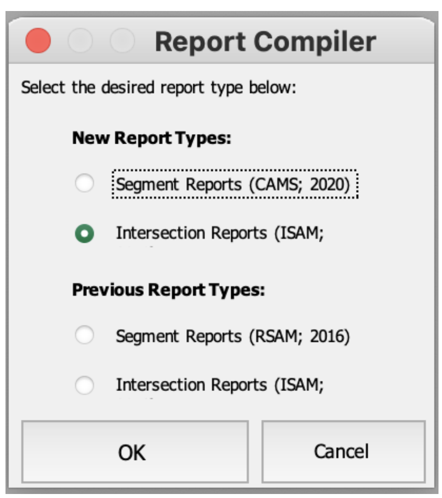

Figure 6.2 Report type selection

Following this prompt, the parameters file along with the statistical output must be selected to begin the report compilation process. Figure 6.3 and Figure 6.4 show the prompts that follow depending on the selection of the segment or intersection report type from Figure 6.2.

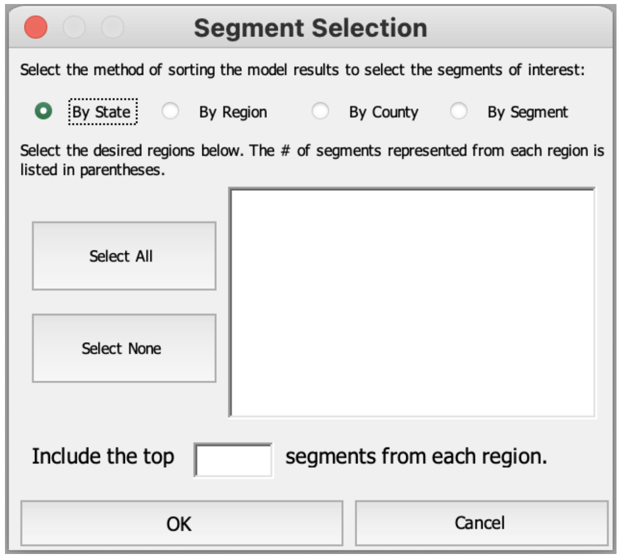

Figure 6.3 Segment selection form

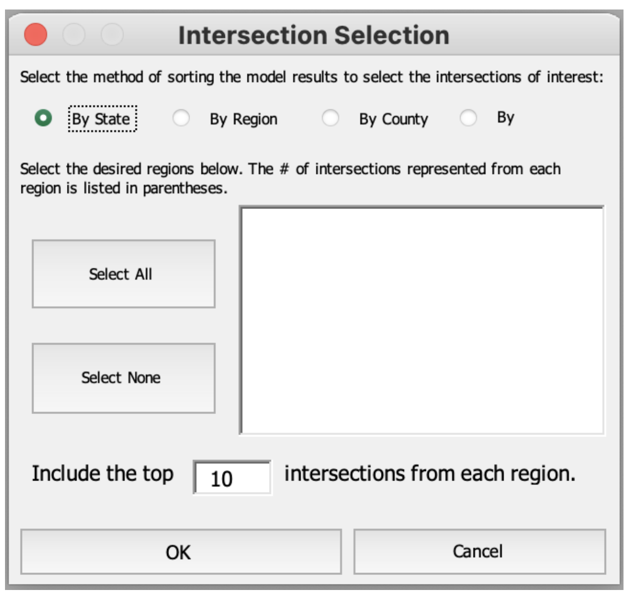

Figure 6.4 Intersection selection form

From these prompts, the number of reports created, and the ranking can be adjusted. For this research, the top 10 reports for each region are created and are ranked by the highest scores across the entire state. After the selection process, the reports are generated. The following section details the reports and the information they contain.

6.4 Two-Page Reports

The two-page reports are the final output of TOMS. While the report compiler has not been transitioned to R like the rest of the current model, it is an integral part of the research that allows the UDOT regions to prioritize funding and establish effective countermeasures to use and potentially save lives. The following subsections detail the different reports for segments and intersections and the information generated on each page organized by the page of the report (page one and page two).

6.4.1 Segments

The purpose of the segment report is to summarize and present preliminary results from a safety-specific microanalysis on an identified segment of interest. The report includes information on segment metadata, characteristics, crash data, historical/current conditions, and possible countermeasures, all of which will be discussed in the following subsections.

6.4.1.1 Page One

Figure 6.5 shows the first page of the two-page reports that are sent to UDOT regions regarding segments that were within the top 10 of the hot spot ranking. There are six tables which are filled through the report compiler that are presented on the first page.

The first three tables contain roadway information. Table 1 describes the location of the segment and general information. This includes route, direction, mile points, length (in miles), and ranking. Table 2 follows and details some data from the five essential segmentation variables described previously in Section 3.2. This includes functional class, AADT, number of lanes, speed limit, and urban code. Table 3 contains the supplemental data from variables other than the essential five. This table was previously manually filled in but within the switch to R is filled through the report compiler.

The next four tables of the two page-reports contain crash information. Table 4 includes the crash and severity counts for crashes that occurred on the segment. Table 5 details the top seven crash factors for crashes along the segment and details how many of those crashes were injury crashes (Severity 3-5). Table 6 includes manner of collision data and like Table 5, details how many crashes per manner of collision were considered injury crashes.

6.4.1.2 Page Two

Figure 6.6 shows the second page of the two-page reports that are sent to UDOT regions regarding segments that were within the top 10 of the hot spot ranking. The second page is manually filled through virtual site visits conducted with Google Maps (Google, 2023b), Google Earth (Google, 2023a), and Roadview Explorer 5 (Mandli Communications, 2022). Photos from these sites are included within the report to show the surrounding geographical area along with a street view. Any major changes made to the roadway within the years of study (2016-2020) are noted in the site visit notes section.

Following this section, possible engineering and policy countermeasures are suggested based on the data located in Table 5. These countermeasures are generated automatically through the report compiler and at least five are selected with a maximum of 12 to be included in the report. Following the site visit and selection of countermeasures the reports are sent to the UDOT regions that they correspond to.

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: 7/27/2023

Street Name: Interstate 80 Westbound

Table 1: Segment Metadata

Route Number:	80	UC Model Used:	SegEWRS
Road Direction:	N	State Rank:	1
Beginning, Ending MP:	170.255-178.9433782	Region, Rank:	2, 1
Length (miles):	8.6883782	County, Rank:	Summit, 1
Data Source Years:	2016-2020		

Table 2: Segment Characteristics

Functional Class:	Interstate	AADT:	16631
Number of Thru Lanes:	2	Speed Limit (MPH):	80

Table 3: Roadway Characteristics

MPs	Median	Shoulder	Grade	Curve	Lanes	Wall/ Barrier	Rumble Strips
170.255-178.9433782	Depressed Median	Paved-5 ft	Slight	Y	2	N	Y

Micro-Analysis of Crash Data

Crash Data Summary

Table 4: Crash Count and Severity

Crash Severities	Functional Area Method	Crashes During 2020		Total Crashes Between 2016-2020				
		Predicted	Actual	Sev. 5	Sev. 4	Sev. 3	Sev. 2	Sev. 1
12345	UDOT	15.12184	18	4	4	9	8	50

Table 5: Top 7 Crash Factors

Crash ID	SINGLE VEHICLE	COLLISION WITH FIXED OBJECT	ROADWAY DEPARTURE	ROADWAY GEOMETRY RELATED	OVERTURN ROLLOVER	COMMERCIAL MOTOR VEH INVOLVED	SPEED RELATED
Injury Total	17/21	15/21	15/21	13/21	12/21	7/21	7/21
Segment Total	67/89	56/89	55/89	50/89	21/89	32/89	24/89

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	Sideswipe Same Direction	Angle	Front to Rear	Head On	Sideswipe Opposite Direction	Parked Vehicle	Rear to Side	Rear to Rear
Injury Total	18/21	2/21	1/21	0/21	0/21	0/21	0/21	0/21	0/21
Segment Total	74/89	10/89	1/89	4/89	0/89	0/89	0/89	0/89	0/89

Figure 6.5 Segment report example - page one

Historical Perspective, Current Conditions, Site Visit Notes	*Date:*	7/27/2023

This segment has not experienced significant changes during the analysis period (2016-2020).

This segment is a portion of WB I-80 in Summit County. It is a two-lane divided highway. There are two lanes. There are rumble strips. The paved shoulder on the right varies from 3 ft to 5.1 ft and on the left varies from 6.5 ft to 11.4 ft. There are slight horizontal curves on this segment as shown in Figure 1.

Figure 1: Streetview of WB I-80 with slight curve (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

Engineering Countermeasures
Increase and strengthen truck maintenance programs and inspection performance (E)
Implement roadway improvements to reduce the likelihood and severity of run-off-road and/or head-on collisions (P)
Install interactive truck rollover signing (P)

Policy Countermeasures
Incorporate information on distracted/fatigued driving into education programs and materials for young drivers (T)
Encourage trucking companies and other fleet operators to implement fatigue management programs (T)

Figure 6.6 Segment report example - page two

6.4.2 Intersections

The purpose of the intersection report is to summarize and present preliminary results from a safety-specific microanalysis on an identified intersection of interest. Both signalized and unsignalized intersections will use the same report template. The report includes information on intersection metadata, characteristics, crash data, historical/current conditions, and possible countermeasures, all of which will be discussed in the following subsections organized by the page of the report (page one and page two).

6.4.2.1 Page One

Figure 6.7 shows the first page of the two-page reports that are sent to UDOT regions regarding intersections that were within the top 10 of the hot spot ranking. There are five tables which are filled through the report compiler that are presented on the first page.

The first two tables contain roadway information. Table 1 describes the location of the intersection and general information. This includes all intersecting routes, mile points of each of the routes, and ranking. Table 2 follows and details the intersection control type at the intersection along with some data from the segmentation variables described previously in Section 3.2. This includes functional class of the largest and smallest roadway, entering vehicles, number of lanes on main route, and maximum/minimum speed limit.

The next four tables of the two-page reports contain crash information. Table 3 includes the crash and severity counts for crashes that occurred on the segment. Table 4 details the top seven crash factors for crashes along the segment and details how many of those crashes were injury crashes (Severity 3-5). Table 5 includes manner of collision data and like Table 4, details how many crashes per manner of collision were considered injury crashes.

6.4.2.2 Page Two

Figure 6.8 shows the second page of the two-page reports that are sent to UDOT regions regarding intersections that were within the top 10 of the hot spot ranking. The second page is manually filled through virtual site visits conducted with Google Maps (Google, 2023b), Google Earth (Google, 2023a), and Roadview Explorer 5 (Mandli Communications, 2022). An aerial photo of the intersection is included to show the surrounding geographical area along with a street view along the main route. Any major changes made to the intersection within the years of study (2016-2020) are noted in the site visit notes section. Changes such as an increase in the number of lanes or any change in intersection control are noted here.

Following this section, possible engineering and policy countermeasures are suggested based on the crash factors data located in Table 4. Similar to the segment report, these countermeasures are generated automatically through the report compiler and at least five are selected with a maximum of 12 to be included on the report. Following the site visit and selection of countermeasures the reports are sent to the UDOT regions that they correspond to.

6.5 Summary

The report compiler creates two-page reports of the top 10 hot spots of segments and intersections for UDOT Region engineers. The report compiler uses the statistical input file and a parameters file to create roadway site reports as well as to suggest possible countermeasures. These reports contain roadway and crash information organized in a manner that allows UDOT to quickly understand how many crashes are occurring at a site and potential causes for the crashes. While countermeasures are suggested, they are not meant to be definitive solutions. These deliverables are sent directly to the regions, and it is up to the region engineers to appropriately address the safety issues at each roadway site.

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: *7/27/2023*

Street Names: Main St & 1700 S

Table 1: Intersection Metadata

Model Used:	IntEWRS	Leg 0 Route & MP:	165	9.7
State Rank:	1	Leg 1 Route & MP:	1193	0.88
Region & Rank:	1, 1	Leg 2 Route & MP:	165	9.7
County & Rank:	Cache, 1	Leg 3 Route & MP:	1193	0.88
Years of Data:	2016-2020	Leg 4 Route & MP:	-	-
City/Area:		Latitude & Longitude:	41.70268	-111.8348

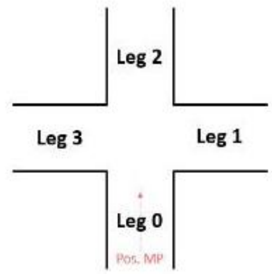

Table 2: Intersection Characteristics

Intersection Control:	SIGNAL	Entering Vehicles in 2020:	13,720	
Princ. Functional Class:	Minor Arterial	# of Lanes on Route 0:	4	
		Max & Min Speed Limit (mph):	55	-

Micro-Analysis of Crash Data

Crash Data Summary

Table 3: Crash Count and Severity

Crash Severities Used	Functional Area Method Used	Crashes during 2020		Total crashes between 2016-2020				
		Predicted	Actual	Sev. 5	Sev. 4	Sev. 3	Sev. 2	Sev. 1
12345	UDOT	0.85864	5	3	0	1	8	19

Table 4: Crash Factors

Crash ID	Latitude	Longitude	OLDER DRIVER INVOLVED	DISTRACTED DRIVING	TEENAGE DRIVER INVOLVED	NIGHT DARK CONDITION	ADVERSE ROADWAY SURF CONDITION	UNRESTRAINED	OVERTURN ROLLOVER
Injury Total			2/4	2/4	1/4	1/4	1/4	1/4	1/4
Intersect'n Total			12/41	6/41	14/41	11/41	10/41	2/41	1/41

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	Angle	Front to Rear	Sideswipe Same Direction	Head On	Single Vehicle	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
Intrsectn. Total	25/41	11/41	3/41	1/41	1/41	0/41	0/41	0/41	0/41

Figure 6.7 Intersection report example - page one

Historical Perspective, Current Conditions, Site Visit Notes *Date:* *7/27/2023*

This intersection has not experienced significant changes during the analysis period (2016-2020).

This intersection is located in Providence, Utah just south of Logan. SR-165 (Main St) is the primary route of this intersection and has four lanes, two thru,one left-turn, and one right-turn lane on the North and South approaches. 1700 S is the secondary route and while the West approach has a single lane, the East Approach has two lanes, a left-turn/thru land and a right-turn lane. An aerial view of the intersection and surrounding area is shown in Figure 1.

Figure 1: Aerial view of the intersection (Google).

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
Provide advanced warning signs (T)
Provide all-red clearance intervals at signalized intersections (T)
Delineate roadside objects (E)
Delineate trees or utility poles with retroreflective tape (E)

Policy Countermeasures
Increase seatbelt use by older drivers and passengers (P)

Figure 6.8 Intersection report example - page two

7.0 ANALYSIS AND SUMMARY OF RESULTS

7.1 Overview

The TOMS output allows UDOT region engineers to review and determine which areas of roadway need improvement. After the data compilation and statistical analysis, the two-page reports are created for the top 10 “hot spots” in each UDOT region. These reports contain tables of roadway and crash data. This section will review the top 10 hot spot segments and intersections on Utah roadways. The following subsections describe the results of the segments, then intersections, the development of severity distributions, and applications of this research.

7.2 Segments

Table 7.1 contains the results of the statistical analysis conducted on the compiled segments file. The rank, route, mile point, location, and functional class information of the segment are given in the first seven columns while the crash information is given in the next seven. The state ranking is determined by the EWRS score for each segment and is how they are subsequently ranked. The crash data is from 2016-2020. The top-ranking segments ranged from having nine crashes to 75 total crashes in the five-year period from 2016 to 2020. All these segments have had three or more fatal crashes within the study period. Eight of the top 10 segments have a functional class of either Interstate or Other Principal Arterial. Region 3 is the only one without a segment in the top 10 while the other three regions have at least three segments.

Figure 7.1 shows a map of the top ranked segments. This map was developed by UDOT using ArcGIS (UDOT, 2023). This map shows the top 10 segments from each region which are ranked in the state from 1-54. The map shows that these top-ranked segments are all over the state in both urban and rural areas.

Table 7.1 Top 10 Segment Hot Spots

State Rank	Region	Main Route	Beginning Mile point	Ending Mile point	County	Functional Class	Observed Crashes (2016-2020)						
							Total	Injury (3-5)	Severity				
									1	2	3	4	5
1	2	0080NM	170.255	178.943	Summit	Interstate	75	17	50	8	9	4	4
2	4	0089PM	25.199	39.070	Kane	Other Principal Arterial	66	8	49	9	3	1	4
3	2	0080NM	76.535	81.185	Tooele	Interstate	26	7	13	6	3	1	3
4	4	0006PM	296.136	299.370	Emery	Other Principal Arterial	10	5	5	0	2	0	3
5	2	0036PM	6.084	13.368	Tooele	Major Collector	19	6	11	2	2	1	3
6	4	0163PM	12.436	17.786	San Juan	Major Collector	9	4	4	1	1	0	3
7	1	0084NM	7.117	11.963	Box Elder	Interstate	36	8	27	1	3	2	3
8	1	0084PM	11.963	15.797	Box Elder	Interstate	32	6	19	7	2	1	3
9	4	0089PM	280.826	283.832	Sanpete	Other Principal Arterial	41	10	27	4	5	2	3
10	1	0126PM	3.355	3.860	Davis	Other Principal Arterial	39	11	24	4	7	1	3

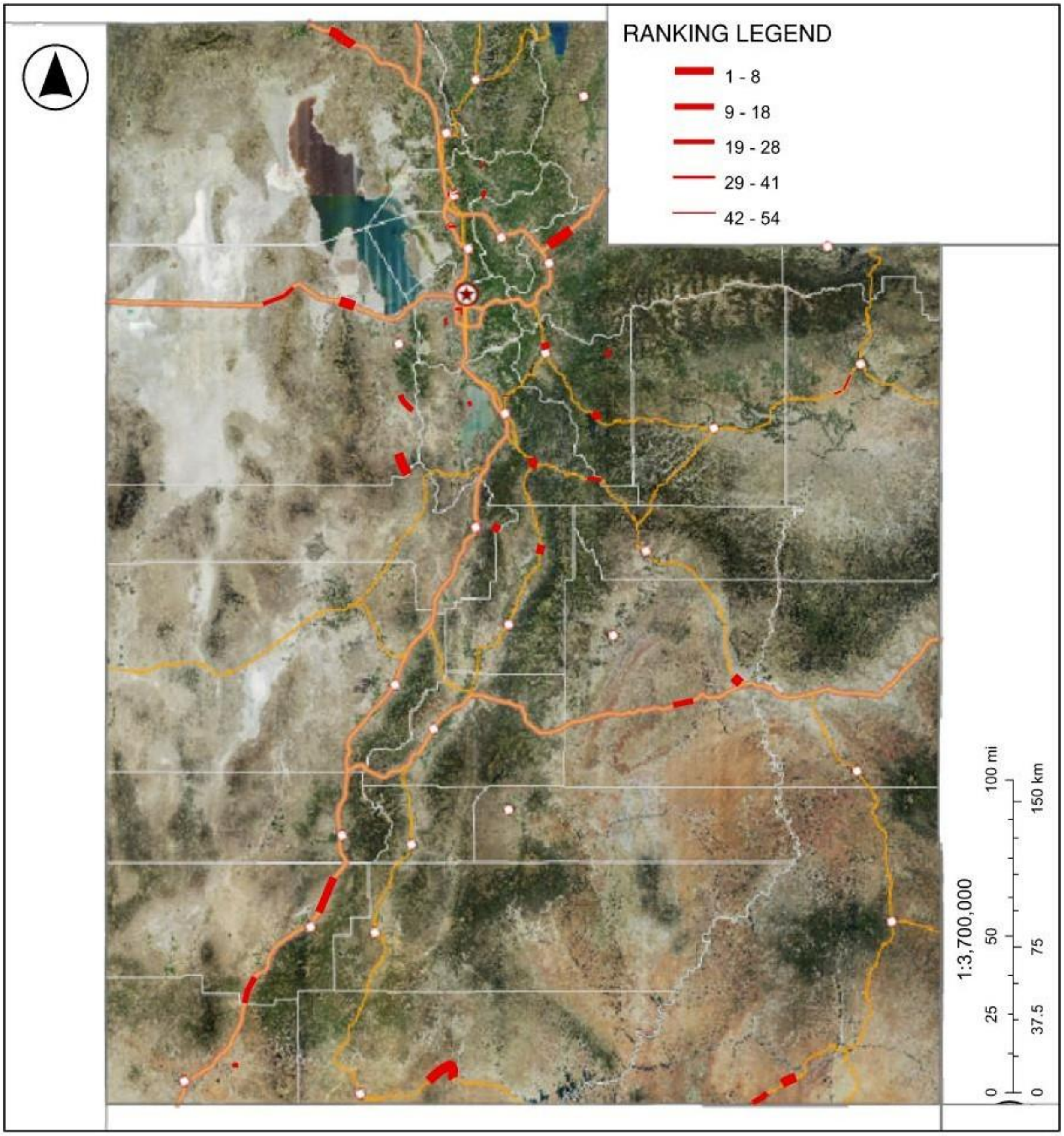

Figure 7.1 Map of top-ranked segments (UDOT, 2023)

7.3 Intersections

Table 7.2 contains the results of the statistical analysis conducted on the compiled intersections file. The rank, region, traffic control device, location, and intersection type information of the segment are given in the first nine columns, while the crash information is shown in the following seven columns. The state ranking is determined by the EWRS score for each intersection and is how they are subsequently ranked. The crash data is from 2016-2020. The top 10 intersections are not spread equally across the regions with most of the hot spots occurring in Region 1. The top-ranking intersections ranged from having six crashes to 195 total crashes in the five-year period from 2016 to 2020. All of the top-ranked intersections have at least one fatal crash within the study period.

Figure 7.2 shows a map of the top-ranked intersections. The cluster of intersections are mainly in Salt Lake and Davis counties. Many of the top-ranked signalized intersections are shown to be in urban areas. More specifically, urban areas in Region 1 and Region 2. Unsignalized intersections are shown to be spread out in both urban and rural areas. They also are spread out amongst all four regions compared to the signalized intersections.

7.4 Development of Severity Distributions

Using TOMS and the JSM, the severity distributions were derived.

Table 7.3 shows the predicted average crash severity by segment type from 2016-2020. Segment types were split into urban and rural (with urban segments appearing first). The first column lists the segment type and below it, in parentheses, are the number of miles of that segment type that were analyzed. The next five columns present the predicted percentage of crash severity, from PDO to fatal crashes. The final column is a 95 percent uncertainty interval on the predicted percentage of severe crashes (Suspected Serious Injury and Fatal) for that segment type.

Table 7.5 shows the average crash severity by intersection type for 2016-2020. Like the segment severity distributions, intersections are split up into urban and rural. Signalized and unsignalized intersections are both included within the table. The first column lists the intersection type and below it, in parentheses, are the number of intersections of that type that were analyzed. The next five columns present the predicted percentage of crash severity, from PDO to fatal crashes. The final column is a 95-percent uncertainty interval on the predicted percentage of severe crashes (Suspected Serious Injury and Fatal) for that intersection type.

One issue that was observed in both the segment and intersection analysis is that the proportion of predicted rural severe crashes is biased low. The hierarchical nature of the statistical model attempts to link the multiple populations and tie them closer to each other (e.g., statistical borrowing of strength). Usually this is a desirable model characteristic. However, in this case it had an undesirable effect. Via a hypothesis test, the data provides strong evidence that the proportion of rural severe crashes is different (and higher) than the proportion of urban severe crashes. The model does not capture that effect. If interested in comparing rural and urban severe crash proportions, it is suggested that the empirical crash severity data found in Table 7.4 and Table 7.6 be used. The 95-percent uncertainty intervals were fit with a binomial model using a Bayesian framework with a uniform prior distribution on the proportion of severe crashes.

7.5 Applications of Research

The model results including the top 10 segments, signalized, and unsignalized intersections from each region were provided to UDOT. UDOT checked to verify that no recent project that may have improved safety conditions and no planned or upcoming projects were already being considered for these sites. All remaining roadway sites were recommended for further consideration.

Table 7.7 shows the selected segments, Table 7.8 shows the selected signalized intersections, and Table 7.9 shows the selected unsignalized intersections from the analysis. The tables show general information and the state ranking of the selected sites. There were 26 segments, 20 signalized intersections, and 30 unsignalized intersections selected for a total of 76 roadway sites.

Table 7.2 Top 10 Intersection Hot Spots

State Rank	Region	Traffic Control	Main Route	Intersecting Route	City	Street Names	County	Intersection Type	Observed Crashes (2016-2020)						
									Total	Injury (3-5)	Severity				
											1	2	3	4	5
1	1	Signalized	0165PM	1193PM	Providence	Main St & 1700 S	Cache	4-LEG	31	4	19	8	1	0	3
2	1	Unsignalized	0235PM	Local	Ogden	Washington Blvd & 900 N	Weber	3-LEG	6	4	1	1	2	0	2
3	1	Unsignalized	0091PM	Local	Lewiston	Hwy 91 & 13400 N	Cache	4-LEG	12	4	8	0	1	1	2
4	2	Signalized	0173PM	2156PM	West Valley City	5415 S & Northwest Ave	Salt Lake	4-LEG	37	8	19	10	3	3	2
5	2	Signalized	0173PM	Local	West Valley City	5400 S & Mountain View Corridor	Salt Lake	4-LEG	22	6	10	6	4	0	2
6	1	Unsignalized	0126PM	Local	Farr West	2000 W & 4000 N	Weber	3-LEG	8	2	3	3	0	0	2
7	1	Signalized	0126PM	0108PM	Clearfield	SR-193 & H St	Davis	4-LEG	195	33	120	42	27	4	2
8	2	Signalized	0171PM	0154PM	West Valley City	3500 S & Bangerter Hwy	Salt Lake	CFI CENTRAL	195	20	119	56	17	1	2
9	2	Signalized	0089PM	Local	Salt Lake City	State St & Kensington Ave	Salt Lake	4-LEG	20	7	5	8	3	3	1
10	1	Signalized	0204PM	0079PM	Ogden	Harrison Blvd & 20th St	Weber	4-LEG	76	30	32	14	27	2	1

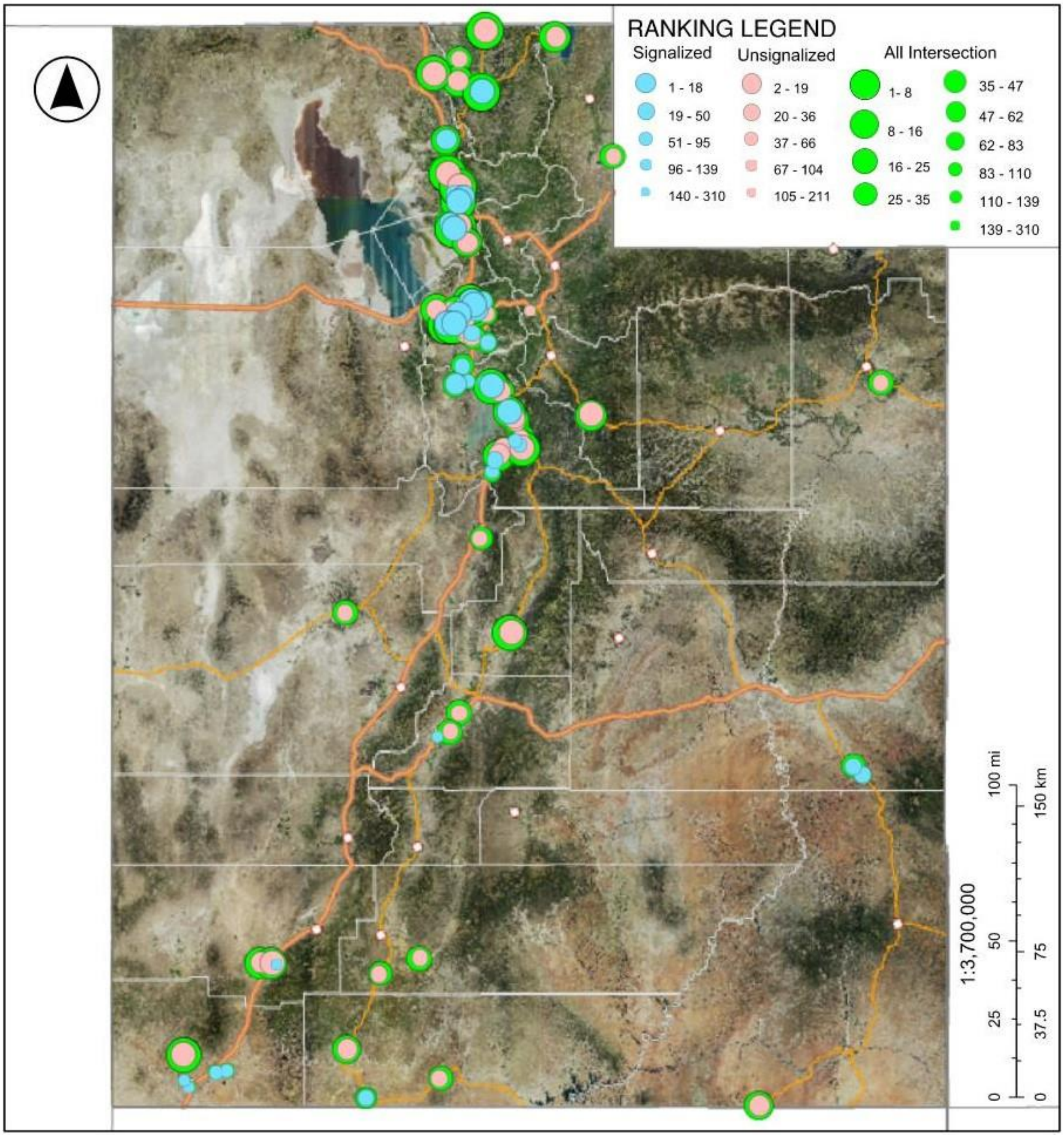

Figure 7.2 Map of top-ranked intersections (UDOT, 2023)

Table 7.3: Average Crash Severity by Segment Type

Segment Type (#)	No Injury (PDO)	Possible Injury	Suspected Minor Injury	Suspected Serious Injury	Fatal	Severe Crashes 95% Uncertainty Interval
Urban						
2-Lane (1091)	72.88%	16.05%	8.51%	2.01%	0.55%	(2.26%, 2.89%)
3-Lane w/TWLTL (53)	71.90%	16.57%	8.84%	2.12%	0.57%	(1.06%, 4.55%)
4-Lane Divided (371)	73.49%	15.76%	8.27%	1.95%	0.53%	(2.26%, 2.71%)
4-Lane Undivided (303)	72.73%	16.11%	8.57%	2.03%	0.56%	(2.13%, 3.09%)
5-Lane w/TWLTL (23)	69.40%	17.75%	9.83%	2.36%	0.66%	(1.28%, 5.05%)
6-Lane Divided (214)	72.69%	16.16%	8.57%	2.02%	0.56%	(2.21%, 2.98%)
6-Lane Undivided (58)	70.72%	17.14%	9.30%	2.22%	0.62%	(2.03%, 3.69%)
Freeway 4-Lane (230)	74.01%	15.49%	8.08%	1.90%	0.52%	(1.96%, 2.90%)
Freeway 6-Lane w/HOV (56)	73.67%	15.66%	8.21%	1.93%	0.53%	(2.06%, 2.87%)
Freeway 6-Lane w/o HOV (174)	73.83%	15.59%	8.15%	1.91%	0.52%	(2.11%, 2.78%)
Freeway 8-Lane w/HOV (111)	73.36%	15.82%	8.32%	1.96%	0.54%	(2.22%, 2.78%)
Freeway 8-Lane w/o HOV (72)	74.31%	15.33%	7.98%	1.87%	0.51%	(1.90%, 2.87%)
Freeway 10-Lane w/HOV (54)	72.45%	16.28%	8.66%	2.05%	0.56%	(2.24%, 2.99%)
Freeway 10-Lane w/o HOV (14)	74.50%	15.24%	7.90%	1.85%	0.51%	(1.22%, 3.69%)
Freeway 12-Lane (8)	73.12%	15.93%	8.42%	1.99%	0.54%	(1.29%, 3.95%)
Rural						
2-Lane (2724)	73.37%	15.82%	8.31%	1.96%	0.54%	(2.28%, 2.71%)
Multilane Divided (57)	76.21%	14.35%	7.29%	1.69%	0.46%	(1.57%, 2.78%)
Multilane Undivided (201)	74.62%	15.18%	7.86%	1.84%	0.50%	(1.53%, 3.24%)
Freeway 4-Lane (454)	74.06%	15.47%	8.06%	1.89%	0.52%	(2.09%, 2.73%)
Freeway 6-Lane (47)	75.55%	14.69%	7.53%	1.75%	0.48%	(1.52%, 3.02%)

Table 7.4: Empirical Crash Severity by Segment Type

Segment Type (#)	No Injury (PDO)	Possible Injury	Suspected Minor Injury	Suspected Serious Injury	Fatal	Severe Crashes 95% Uncertainty Interval
Urban						
2-Lane (1091)	73.21%	15.19%	8.76%	2.32%	0.52%	(2.54%, 3.17%)
3-Lane w/TWLTL (53)	76.99%	12.27%	8.90%	1.53%	0.31%	(0.86%, 3.95%)
4-Lane Divided (371)	73.39%	18.29%	6.86%	1.25%	0.21%	(1.31%, 1.63%)
4-Lane Undivided (303)	73.15%	16.81%	7.99%	1.60%	0.45%	(1.67%, 2.51%)
5-Lane w/TWLTL (23)	60.49%	21.91%	14.51%	2.78%	0.31%	(1.70%, 5.59%)
6-lane Divided (214)	72.14%	17.88%	8.49%	1.20%	0.29%	(1.23%, 1.81%)
6-Lane Undivided (58)	67.99%	20.68%	9.73%	1.28%	0.32%	(1.09%, 2.35%)
Freeway 4-Lane (230)	76.03%	13.11%	8.12%	2.30%	0.44%	(2.29%, 3.26%)
Freeway 6-Lane w/HOV (56)	73.11%	19.36%	6.55%	0.91%	0.07%	(0.76%, 1.26%)
Freeway 6-Lane w/o HOV (174)	74.98%	15.73%	7.49%	1.44%	0.36%	(1.54%, 2.09%)
Freeway 8-Lane w/HOV (111)	72.63%	19.50%	6.59%	1.14%	0.14%	(1.11%, 1.48%)
Freeway 8-Lane w/o HOV (72)	74.37%	16.55%	7.55%	1.22%	0.31%	(1.19%, 1.97%)
Freeway 10-Lane w/HOV (54)	70.20%	20.02%	8.18%	1.43%	0.17%	(1.35%, 1.92%)
Freeway 10-Lane w/o HOV (14)	73.70%	18.04%	7.08%	0.51%	0.67%	(0.58%, 2.41%)
Freeway 12-Lane (8)	75.24%	17.85%	6.53%	0.38%	0.00%	(0.12%, 1.38%)
Rural						
2-Lane (2724)	74.32%	10.93%	9.94%	3.54%	1.27%	(4.56%, 5.08%)
Multilane Divided (57)	76.98%	13.44%	6.91%	2.23%	0.44%	(2.09%, 3.41%)
Multilane Undivided (201)	78.70%	10.41%	7.69%	2.48%	0.72%	(2.36%, 4.33%)
Freeway 4-Lane (454)	74.25%	11.20%	9.87%	3.42%	1.26%	(4.29%, 5.11%)
Freeway 6-Lane (47)	75.92%	14.14%	7.07%	2.52%	0.35%	(2.12%, 3.87%)

Table 7.5: Average Crash Severity by Intersection Type

Intersection Type (#)	No Injury (PDO)	Possible Injury	Suspected Minor Injury	Suspected Serious Injury	Fatal	Severe Crashes 95% Uncertainty Interval
Urban						
Signal controlled, 4-Leg (988)	63.50%	21.61%	12.31%	2.24%	0.34%	(2.34%, 2.80%)
Signal controlled, 3-Leg (134)	62.69%	21.99%	12.66%	2.31%	0.35%	(2.03%, 3.33%)
2-way stop controlled, 4-Leg (916)	63.62%	21.57%	12.25%	2.22%	0.34%	(2.12%, 3.00%)
Minor leg stop controlled, 3-Leg (2459)	63.05%	21.82%	12.50%	2.28%	0.35%	(2.24%, 3.01%)
Rural						
Signal controlled, 4-Leg (18)	65.30%	20.81%	11.51%	2.06%	0.32%	(0.00%, 5.45%)
2-way stop controlled, 4-Leg (1048)	62.92%	21.84%	12.59%	2.30%	0.35%	(1.45%, 4.03%)
Minor leg stop controlled, 3-Leg (2528)	61.28%	22.60%	13.29%	2.45%	0.38%	(1.75%, 4.04%)

Table 7.6: Empirical Crash Severity by Intersection Type

Intersection Type (#)	No Injury (PDO)	Possible Injury	Suspected Minor Injury	Suspected Serious Injury	Fatal	Severe Crashes 95% Uncertainty Interval
Urban						
Signal controlled, 4-Leg (988)	62.91%	22.20%	12.52%	2.07%	0.30%	(2.23%, 2.54%)
Signal controlled, 3-Leg (134)	65.10%	21.10%	11.76%	1.81%	0.23%	(1.57%, 2.65%)
2-way stop controlled, 4-Leg (916)	62.30%	21.98%	12.70%	2.75%	0.27%	(2.60%, 3.50%)
Minor leg stop controlled, 3-Leg (2459)	64.79%	20.21%	12.21%	2.37%	0.42%	(2.44%, 3.20%)
Rural						
Signal controlled, 4-Leg (18)	69.33%	16.00%	12.67%	1.33%	0.67%	(0.73%, 5.70%)
2-way stop controlled, 4-Leg (1048)	62.41%	18.50%	13.22%	4.70%	1.17%	(4.35%, 7.90%)
Minor leg stop controlled, 3-Leg (2528)	64.63%	13.95%	14.32%	5.39%	1.71%	(5.54%, 9.07%)

Table 7.7: UDOT Selected Segments List

State Rank	Region	Main Route	Beginning Mile point	Ending Mile point	County
2	4	0089PM	25.199	39.070	Kane
4	4	0006PM	296.136	299.370	Emery
5	2	0036PM	6.084	13.368	Tooele
6	4	0163PM	12.436	17.786	San Juan
7	1	0084NM	7.117	11.963	Box Elder
8	1	0084PM	11.963	15.797	Box Elder
9	4	0089PM	280.826	283.832	Sanpete
13	3	0132PM	39.144	41.754	Juab
16	3	0006PM	184.202	187.570	Utah
17	2	0173PM	2.910	3.564	Salt Lake
18	4	0015PM	82.459	94.697	Iron
19	2	0068PM	54.973	55.815	Salt Lake
20	3	0035PM	16.155	18.870	Wasatch
21	4	0163PM	0.614	6.387	San Juan
23	2	0073PM	7.122	13.340	Tooele
24	4	0059PM	20.670	21.797	Washington
25	4	0015PM	42.224	51.248	Iron
31	3	0006PM	205.834	210.521	Utah
32	1	0104PM	1.176	1.786	Weber
34	1	0089PM	414.286	414.393	Weber
35	2	0089PM	381.148	381.298	Salt Lake
36	1	0167PM	8.013	11.002	Weber
37	3	0068PM	23.955	25.315	Utah
41	1	0203PM	5.413	5.960	Weber
51	1	0158PM	8.161	11.611	Weber
54	1	0039PM	6.402	6.521	Weber

Table 7.8: UDOT Selected Signalized Intersections List

State Rank	Region	Main Route	Other Route	Street Names	City	County	Intersection Type
4	2	0173PM	2156PM	5415 S & Northwest Ave	West Valley City	Salt Lake	4-LEG
7	1	0126PM	0108PM	State St & Antelope Dr	Layton	Davis	4-LEG
10	1	0204PM	0079PM	Wall Ave & 31st St	Ogden	Weber	4-LEG
18	3	0089PM	Local	State St & 550 W	Provo	Utah	3-LEG
26	1	0193PM	Local	SR-193 & H St	Clearfield	Davis	4-LEG
35	1	0091PM	0091PM	SR-91 I-15 Interchange	Brigham City	Box Elder	DDI
50	1	0203PM	Local	Harrison Blvd & 20th St	Ogden	Weber	4-LEG
55	1	0039PM	Local	1200 S & Depot Dr	Ogden	Weber	3-LEG
56	4	0191PM	Local	Hwy 191 & Arches Nat'l Park Rd	Moab	Grand	3-LEG
70	4	089APM	0089PM	100 E & 300 S	Kanab	Kane	4-LEG
80	3	0178PM	Local	800 S & Turf Farm Rd	Payson	Utah	4-LEG
90	4	0191PM	1699PM	Main St & Center St	Moab	Grand	4-LEG
93	1	0204PM	Local	Wall Ave & 23rd St	Ogden	Weber	4-LEG
119	3	0006PM	3035PM	Hwy 6 & 2550 E	Spanish Fork	Utah	4-LEG
120	3	0089PM	2888PM	State St & 900 W	American Fork	Utah	4-LEG
123	3	0068PM	Local	Redwood Rd & Exchange Dr	Saratoga Springs	Utah	4-LEG
193	4	0008PM	3234PM	Sunset Blvd & Westridge Dr	St. George	Washington	4-LEG
203	4	0289PM	0289PM	Center St & 300 W	Cedar City	Iron	4-LEG
220	4	0008PM	3184PM	Sunset Blvd & Dixie Downs Rd	St. George	Washington	4-LEG
296	4	0120PM	Local	1300 S & College Ave	Richfield	Sevier	4-LEG

Table 7.9: UDOT Selected Unsignalized Intersections List

State Rank	Region	Main Route	Other Route	Street Names	City	County	Intersection Type
2	1	0235PM	Local	Washington Blvd & 900 N	Ogden	Weber	3-LEG
11	4	0018PM	Local	SR-18 & Diamond Valley Dr	St. George	Washington	3-LEG
16	4	0089PM	Local	Hwy 89 & Jensen Rd	Manti	Sanpete	3-LEG
19	4	0056PM	Local	SR-56 & Beacon Dr	Cedar City	Iron	3-LEG
21	4	0056PM	Local	SR-56 & Iron Springs Rd	Cedar City	Iron	3-LEG
24	1	0030PM	0023PM	SR-30 & SR-23	Mendon	Cache	4-LEG
27	3	0089PM	Local	State St & 550 N	Lindon	Utah	3-LEG
28	1	0089PM	3457PM	Hwy 89 & 2000 N	Harrisville	Weber	3-LEG
31	1	0273PM	1446PM	Main St & Nicholls Rd	Kaysville	Davis	3-LEG
32	1	0235PM	Local	Washington Blvd & Canfield Dr	Ogden	Weber	3-LEG
33	3	0089PM	Local	State St & 800 S	Provo	Utah	3-LEG
39	4	0257PM	1934PM	400 W & 500 N	Hinckley	Millard	4-LEG
40	2	0068PM	Local	Redwood Rd & Sequoia Vista Cir	Salt Lake City	Salt Lake	4-LEG
41	4	0119PM	2540PM	SR-119 & N 3380 E	Richfield	Sevier	4-LEG
42	4	0118PM	Local	SR-118 & 1520 N	Sigurd	Sevier	3-LEG
43	1	0218PM	Local	SR-218 & 6600 N	Newton	Cache	3-LEG
44	1	0016PM	Local	SR-16 & Co Rd 101	Rich County	Rich	3-LEG
45	3	0040PM	Local	Hwy 40 & 4625 E	Naples	Uintah	3-LEG
47	3	0089PM	Local	Hwy 89 & California Ave	Provo	Utah	3-LEG
48	4	0089PM	Local	Hwy 89 & Kitchen Corral Wash	Kane County	Kane	3-LEG
49	4	0089PM	Local	Hwy 89 & Fish Hatchery Rd	Hatch	Garfield	4-LEG
57	3	0028PM	Local	SR-28 & Four Mile Rd	Nephi	Juab	3-LEG
60	2	0089PM	Local	State St & 8840 S	Sandy	Salt Lake	3-LEG
66	2	0171PM	Local	3300 S & 3040 E	Millcreek	Salt Lake	3-LEG
76	2	0266PM	Local	Taylorsville Expy & 1175 W	Taylorsville	Salt Lake	4-LEG
81	2	0173PM	Local	5400 S & 5160 W	Kearns	Salt Lake	3-LEG
101	2	0068PM	Local	Redwood Rd & 7310 S	West Jordan	Salt Lake	3-LEG
104	2	0172PM	Local	5600 W & Lampert Ln	West Valley City	Salt Lake	3-LEG
202	2	0224PM	Local	SR-224 & Bear Cub Dr	Park City	Summit	3-LEG
211	2	0068PM	Local	Redwood Rd & 8600 S	West Jordan	Salt Lake	3-LEG

7.6 Summary

The results from the segment and intersection analysis are useful to UDOT region engineers in helping them review and determine which areas of roadway need improvement. A ranking based on economic impact vastly improves the significance of the model results. The list of both segments and intersections was provided to UDOT, and with their feedback, the reports from the screened list were generated and sent out to the regions.

<u>8.0 CONCLUSIONS AND IMPLEMENTATION</u>

8.1 Overview

The goal of zero fatalities is a major priority for UDOT. Efforts are made to reduce the number of fatalities, injuries, and crashes on Utah roads every year. 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. Building off previous safety analysis research, TOMS provides a new way for the state of Utah to identify and prioritize segment and intersection safety improvement projects. Along with improvements to the methodology and statistical analysis, TOMS uses a joint model to rank segments and intersections based on the highest EWRS. UDOT can better focus their efforts and manage their budget on projects that are of highest concern across the regions with the identifying and ranking of these segments and intersections. This chapter will review the methodology developed for the research, discuss implementation of the research, and summarize future research topics that this research can contribute to.

8.2 Methodology

The TOMS is a model built within R that allows for the compilation of roadway and crash data to conduct a statistical analysis that will create hot spot identification reports to help UDOT region engineers determine how best to address safety concerns. The following subsections go over the data compilation process, then the statistical analysis, then the report compiler.

<u>8.2.1 Data Compilation</u>

Sixteen data files are used to create the combined crash and roadway data files used for analysis. Twelve contain roadway data while the other four contain crash data. The entire process is conducted from an R markdown file which requires that the following occur in a specific order: first, the functions are sourced; second, the data are read in; third, the data preparation scripts run; finally, the compiling script creates the two output files. There is a segments file and an intersection file. These two files are for the statistical analysis.

8.2.2 Statistical Analysis

Instead of using two different models, one accounting for total number of crashes and another for the total number of severe crashes, the two models can be combined by using an EWRS score. All segments and intersections are analyzed using this method. Those with the highest EWRS scores are marked as hot spots. The list of the top 10 hot spots for both segments and intersections are screened by UDOT and subsequent reports are created from the screened list.

8.2.3 Report Compiler

The report compiler creates two-page reports of the hot spots of segments and intersections selected by UDOT for the region engineers. The report compiler uses the statistical input file and a parameters file to create roadway site reports as well as to suggest possible countermeasures. These reports contain roadway and crash information organized in a manner that allows UDOT to quickly understand how many crashes are occurring at a site and potential causes for the crashes. While countermeasures are suggested, they are not meant to be definitive solutions. These deliverables are sent directly to the regions, and it is up to the region engineers to appropriately address the safety issues at each roadway site.

8.3 Implementation of Research

UDOT will implement the results of this research by screening the hot spots list, selecting sites that fall under their criteria, distributing the two-page reports to the regions, and using the reports to prioritize safety projects across the state. Along with the reports, the severity distributions will assist UDOT and the regions to analyze segments and intersections and to determine if the proportion of severe crashes occurring at a site is significantly higher than at other similar sites. These distributions provide a more overarching metric that may be useful in identifying hot spots even without a two-page report.

8.4 Future Research Topics

The switch to R has allowed for more topics to be explored in future research. The following sections explain potential topics that could be explored such as a safety dashboard, manners of collision, the development of safety performance functions (SPFs), and running statistical models to compare urban vs. rural safety.

8.4.1 Safety Dashboard

A prototype dashboard that could be used to generate SPFs and identify hot spots given certain roadway characteristics was created early in the research and presented to UDOT. It was determined that this would not be a focus of the research at this time. With more time and with refinement of the current data, a dashboard that would allow the user to input roadway data could be created and used to identify safety parameters.

8.4.2 Manner of Collision Research

While this research focuses on the severity of crashes and how expected crashes differ from observed, future research might include finding segments and intersections that have a disproportionate number of general crash types. For example, looking at angle crashes at intersections with atypical roadway geometry, looking at roadway departure crashes along poorly lit segments, or looking at midblock turning crashes along roadways without raised medians.

8.4.3 Development of Safety Performance Functions

With the creation of severity distributions for both segments and intersections the implementation of this to SPFs for Utah roadways is a possible next step in the research. In coordination with the safety dashboard, the results from the Bayesian hierarchical model could contribute to having accurate and practical SPFs that UDOT can use.

8.4.4 Urban vs. Rural

While the severity distributions were created within this research there are some questions about the application of those results. Specifically, those distributions between urban and rural areas. Due to the difference in urban and rural populations, a hierarchal model is not always the best to detect a difference. A statistical model could be run on the overall urban and rural populations to test if there is a significant difference in the proportions of severe crashes.

8.4.5 Other Future Research Topics

TOMS has the capability of easily adapting to new datasets and formats. Potentially using TOMS with more recent roadway and crash data to bring more immediate attention to sites that may need improvement is a possible future research topic that could greatly benefit UDOT.

8.5 Concluding Remarks

The TOMS has significantly improved previous models to assist the state of Utah in identifying crash hot spots. By weighing severity using weighted risk scores, TOMS can analyze crash count and severity concurrently with a singular statistical model. This research will allow for much more data to be analyzed in the future. The switch to R has allowed for more freedom in how the data can be used to further improve traffic safety in the state of Utah. The capabilities of TOMS with further improvements can continue to aid UDOT in their goal of reaching Zero Fatalities.

REFERENCES

Afghari, A., Haque, M., and Washington, S. (2020). Applying A Joint Model of Crash Count and Crash Severity to Identify Road Segments with High Risk of Fatal and Serious Injury Crashes. *Accident Analysis and Prevention, 144*.

Albert, J., and Chib, S. (1993). Bayesian Analysis of Binary and Polychotomous Response Data. *Journal of the American Statistical Association, 88*(422), 669-679.

Ambros, J., Sedonik, J., and Křivánková, Z. (2017). How to Simplify Road Network Safety Screening: Two Case Studies. *TRB 2017 Annual Meeting*.

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

Borsos, A., Ivan, J., and Orosz, G. (2016). Development of Safety Performance Functions for Two-Lane Rural First-Class Main Roads in Hungary,. (G. Yannis, and S. Cohen, Eds.) *Traffic Safety, 4*, 87-100. doi:10.1002/9781119307853.ch6

Cafiso, S., D'Agostino, C., and Persaud, B. (2018). Investigating the Influence of Segmentation in Estimating Safety Performance Functions for Roadway Sections. *Journal of Traffic and Transportation Engineering, 5*(2), 129-136. doi:10.1016/j.jtte.2017.10.001

Cafiso, S., Di Graziano, A., Di Silvestro, G., La Cava, G., and Persaud, B. (2010). Development of Comprehensive Accident Models for Two-Lane Rural Highways Using Exposure, Geometry, Consistency and Context Variables. *Accident Analysis and Prevention, 42*(4), 1072-1079. doi:10.1016/j.aap.2009.12.015

Chen, C., and Shen, C. (2018). Model Selection Based on Resampling Approaches for Cluster Longitudinal Data with Missingness in Outcomes. *Statistics in Medicine, 37*(20), 2982-2997.

Esri. (2019). ArcMap 10.7.

Federal Highway Administration (FHWA). (2017). *KABCO Injury Classification Scale and Definitions*, <https://safety.fhwa.dot.gov/hsip/spm/conversion_tbl/pdfs/kabco_ctable_by_state.pdf> (Accessed 26 February 2023).

France-Mensah, J., and O'Brien, W. (2018). A Hybrid Logistic Regression and Spatial Analysis Approach for Identification of Candidate Highway Safety Projects. *Construction Research Congress, 94*. doi:10.1061/9780784481288.064

Gaweesh, S., Ahmed, M., and Piccorelli, A. (2019). Developing Crash Prediction Models Using Parametric and Nonparametric Approaches for Rural Mountainous Freeways: A Case Study on Wyoming Interstate 80. *Accident Analysis and Prevention, 123*, 176-189. doi:10.1016/j.aap.2018.10.011

Gladence, L., Karthi, M., and Anu, V. (2015). A Statistical Comparison of Logistic Regression and Different Bayes Classification Methods for Machine Learning. *APRN Journal of Engineering and Applied Sciences*.

Google. (2023a). *Google Earth*., <https://earth.google.com/web/> (Accessed 14 August 2023).

Google. (2023b). *Google Maps*. (Google), <https://www.google.com/maps> (Accessed 14 August 2023).

Gross, F., Harmon, T., Albee, M., Himes, S., Srinivasan, R., Carter, D., and Dugas, M. (2016). *Evaluation of Four Network Screening Performance Measures.* U.S. Department of Transportation. Federal Highway Administration.

Held, L., and Holmes, C. (2006). Bayesian Auxiliary Variable Models for Binary and Multinomial Regression. *Bayesian Analysis, 1*(1), 145-168.

Hou, Q., Mu, S., and Zhang, M. (2022). Marginal Effects for Random Parameters Logit Models: A Case Study of Crash Severity Analysis. *22nd COTA International Conference of Transportation Professionals*.

Kwon, O., Park, M., Yeo, H., and Chung, K. (2013). Evaluating the Performance of Network Screening Methods for Detecting High Collision Concentration Locations on Highways. *Accident Analysis and Prevention, 51*, 141-149. doi:10.1016/j.aap.2012.10.019

Li, D., and Al-Mahamda, M. (2020). Collective Risk Ranking of Highway Segments on the Basis of Severity-Weighted Crash Rates. (K. Wang, Ed.) *Journal of Advanced Transportation, 2020*. doi:10.1155/2020/8837762

Lyon, C., Gotts, B., Wong, W., and Persaud, B. (2007). Comparison of Alternative Methods for Identifying Sites with High Proportion of Specific Accident Types. *Transportation Research Record, 2019*(1). doi:10.3141/2019-25

Mandli Communications. (2022). *Roadview Explorer 5*. <https://roadview.udot.utah.gov/utah/index.php> (Accessed 14 August 2023).

Mountain, L., Fawaz, B., and Jarrett, D. (1996). Accident Prediction Models for Roads with Minor Junctions. *Accident Analysis and Prevention*, *28*(6), 695-707. doi:10.1016/S0001-4575(96)00042-5

Ogle, J., Alluri, P., Zhang, C., Rajabi, M., Sarasua, W., and Bendigeri, V. (2018). Methods to Define Homogenous Segments and Assign Crashes for Highway Safety Manual Applications. *Transportation Research Board 97th Annual Meeting*.

Pande, A., Abdel-Ay, 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.

Rudolfer, S. M., Watson, P. C., and Lesaffre, E. (1995). Are Ordinal Models Useful for Classification? A Revised Analysis. *Journal of Statistical Computation and Simulation, 52*(2), 105-132.

Schultz, G. G., Farnsworth, J. S., Roundy, R., Saito, M., Reese, C. S., and Briggs, T. (2013). *Hot Spot Identification and Analysis Methodology.* Brigham Young University, Departement of Civil And Environmental Engineering. Salt Lake City, Utah: Utah Department of Transportation Traffic and Safety, Research Divisions.

Schultz, G. G., Gibbons, J. D., Saito, M., and Clegg, B. W. (2018). *Intersection Safety Analysis Methodology for Utah Roadways.* Brigham Young University, Department of Civil and Environmental Engineering. Salt Lake City, UT: Utah Department of Transportation Traffic and Safety, Research Divisions.

Schultz, G. G., Johnson, E. S., Black, C. W., Francom, D., and Saito, M. (2012). *Traffic and Safety Statewide Model and GIS Modeling.* Brigham Young University, Department of Civil and Environmental Engineering. Salt Lake City, Utah: Utah Department of Transportation Traffic and Safety, Research Divisions.

Schultz, G. G., Lunt, C. C., Pew, T., Warr, R. L., and Heaton, M. J. (2020). *Segment and Intersection Methodologies for Utah Highways.* Brigham Young University, Department of Civil and Environmental Engineering. Salt Lake City, UT: Utah Department of Transportation Traffic and Safety, Research & Innovation Divisions.

Schultz, G. G., Mineer, S. T., Saito, M., Gibbons, J. D., Siegel, S. A., and MacArthur, P. D. (2016). *Roadway Safety Analysis Methodology for Utah.* Brigham Young University, Civil Engineering. Salt Lake City, UT: Utah Department of Transportation Traffic and Safety, Research Divisions.

Sims, A., and Somenahalli, S. V. (2010). Hot Spot Identification Using Frequency of Distinct Crash Types Rather Than Total Crashes. *Australasian Transport Research Forum*.

Srinivasan, R., Gross, F., Lan, B., and Bahar, G. (2016). *Reliability of Safety Management Methods: Network Screening.* U.S. Department of Transportation. Washington, D.C.: Federal Highway Administration.

Stipancic, J., Miranda-Moreno, L., Saunier, N., and Labbe, A. (2019). Network Screening for Large Urban Road Networks: Using GPS Data and Surrogate Measures to Model Crash Frequency and Severity. *Accident Analysis and Prevention, 125*, 290-301. doi:10.1016/j.aap.2019.02.016

Utah Department of Transportation (UDOT). (2021a). *2021 Update of Crash Costs Based on Severity.* Salt Lake City, Utah.

Utah Department of Transportation (UDOT). (2021b). *UDOT Open Data*. <https://data-uplan.opendata.arcgis.com/> (Accessed 8 April 2023).

Utah Department of Transportation (UDOT). (2023). BYU_Model_Location. <https://uplan.maps.arcgis.com/apps/mapviewer/index.html?webmap=60b5879fee344f4d8af2a05264051936> (Accessed 29 July 2023).

Williams, K., Stover, V., Dixon, K., and Demosthenes, P. (2014). *Access Management Manual* (2nd Edition ed.). Washington, D.C.: Transportation Research Record.

Yasmin, S., and Eluru, N. (2018). A Joint Econometric Framework for Modeling Crash Counts by Severity. *Transportmetrica A: Transport Science, 14*(3), 230-255.

Zeng, Q., Wang, Q., Wang, F., and Sze, N. N. (2021). Revisiting Spatial Correlation in Crash Injury Severity: A Bayesian Generalized Ordered Probit Model with Leroux Conditional Autoregressive Prior. *Transportmetrica A: Transport Science, 18*(3), 1084-1102.