2017

Systemic safety evaluation of two lane rural roads using United States road assessment program methodology

Zahra Parvinashtiani
Iowa State University

Systemic safety evaluation of two lane rural roads using United States road assessment program methodology

by

Zahra Parvinashtiani

A thesis submitted to the graduate faculty

in partial fulfillment of the requirements for the degree of
MASTER OF SCIENCE

Major: Civil Engineering (Transportation Engineering)

Program of Study Committee:
Omar Smadi, Major Professor
Jing Dong
Keith K. Knapp

Iowa State University

Ames, Iowa

2017

DEDICATION

I would like to dedicate this thesis to my parents and my husband without whose love and moral support I would not have been able to complete this work. I would also like to dedicate this thesis to my siblings who have always been formative role models throughout my life.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ABSTRACT

The United States Road Assessment Program (usRAP) is a powerful tool for conducting Systemic Safety evaluations. The level of safety of the roads can be assessed through the usRAP Star Rating method, giving one star to least safe and five stars to safest roads. As part of the Star Rating data collection process, a comprehensive list of 40 road attributes are recorded for each 100-meter segment using StreetView imagery. Some of the challenges that are associated with usRAP data collection protocols are human error, inaccurate measurements, and the coder's subjectivity. To examine the effects of these errors on Star Rating results, this study has leveraged the Second Strategic Highway Research Program Roadway (SHRP 2) Information Database (RID) to complement the existing dataset. The RID includes a variety of safety-related roadway attributes collected by a mobile data collection vendor and meets high accuracy requirements by implementing a quality assurance plan. Using benefit-cost analysis, this study aims to compare the objective data collection approach of utilizing a mobile data collection vendor with high quality assurance processes versus the subjective approach of coding data manually. Star Ratings are calculated for a sample of two lane rural roads in North Carolina using the RID and the manually coded dataset.

usRAP uses the risk-based non-crash measure of Road Protection Score (RPS) for assessing the level of safety of the roads by a 1-5 Star Rating scale. The previous validation studies have been mostly limited to the comparison of crash rate and Star Rating averages and have failed to establish a comprehensive statistical relationship. In order to investigate such relationship, this study develops a crash prediction model using a sample of two lane rural roads in North Carolina. The crash frequency was estimated as a function of Road Protection Score and Annual Average Daily Traffic using a negative binomial model. The results of this study showed that the crash frequency consistently increases with Road Protection Score. The safety performance function showed that moving from a 3-star road to a 2-star road would result in 47% more crashes. These findings confirm that Star Rating is a valid risk measure for crash frequency on two lane rural roads.

CHAPTER 1. GENERAL INTRODUCTION

Introduction

In recent decades, several major efforts have been made to improve highway safety at federal, state, and local levels. According to National Highway Traffic Safety Administration (NHTSA) Fatality Analysis Reporting System (FARS), more than 39,000 motor-vehicle related fatalities occurs annually in United States (*1*). In 2014, motor vehicle traffic-related crashes have been the top leading cause of unintentional injury deaths for ages 5 years through 24 years, and the second top leading unintentional cause for ages 25 years through 64 years (*2*). Crashes not only result in loss of hundreds of thousands of lives, but also take their toll on society by costing millions of dollars in medical and work loss expenditures.

Recent federal initiatives have required states to follow a systemic process in order to develop and maintain a highway safety improvement program (*3*). The Federal Highway Administration (FHWA) Systemic Safety Project Selection Tool provides guidance to states on how to develop a comprehensive safety management program (*4*). The first element of the FHWA systemic program includes the problem identification, countermeasure identification, and project prioritization. In the second element, funding sources are allocated to projects for implementation after prioritizing countermeasures for identified problems. Finally, the implemented countermeasures are evaluated in order to determine their effectiveness for making future improvement decisions.

The International Road Assessment Program (iRAP) is one of the most frequently used systemic safety management tools worldwide. The United States Road Assessment Program (usRAP) was initiated in 2004 under the umbrella of iRAP. The usRAP provides Star Rating assessment for roads based on an inventory of roadway elements and then recommends a Safer Roads Investment Plans (SRIPs) for the road segments. The main objective of this study is to gain a more in depth understanding of the usRAP methodology and the challenges associated with it in terms of data collection processes and validation of the relationship between the crashes and Star Ratings.

Research Objectives

While the Road Assessment Programs (RAPs) are widely used in more than 50 countries around the world, more research studies need to be conducted in order to fully investigate the RAP methodology (e.g. Risk Mapping, Star Rating). As part of the usRAP Star Rating process, the roadway elements are collected through an extensive manual data collection protocol. The Star Rating provides an assessment of the level of safety of roads by analyzing the safety-related roadway attributes that have been collected through the review of roadway imagery. Several challenges are associated with the data collection process since errors are likely to occur as a result of human error, inaccurate measurements/estimations, and the coder's subjectivity in data collection. In order to examine the effects of the inaccuracies in usRAP data on the Star Rating results, this study conducts a sensitivity analysis on a sample of two lane rural roads in North Carolina.

Another fundamental premise of the usRAP process is that it is based upon the assumption that the Star Ratings are a valid indicator of crash risk. The usRAP Star Rating methodology does not require crash data and only relies on the information about the built-in features of the road for conducting the safety assessment. A much debated question is that whether or not the Star Rating is a valid measure of level of safety of the road. To answer this question, this study investigates the relationship between crash frequency and Star Rating using a sample of two lane rural roads from North Carolina.

Thesis Organization

This thesis has been organized in the following way. Chapter 1 provides an introduction and research objectives. Chapters 2 and 3 correspond to the research objectives described above. Chapter 2 compares objective and subjective data collection methods using the usRAP systemic safety management tool. Chapter 3 examines the assumption behind the usRAP methodology and validates the relationship between the crashes and Star Rating in two-lane rural roads. The final chapter draws upon the entire thesis, tying up the finding of two research papers and lays out recommendations for future research.

CHAPTER 2. COMPARING OBJECTIVE AND SUBJECTIVE ROADWAY DATA COLLECTION METHODS USING THE UNITED STATES ROAD ASSESSMENT PROGRAM

Modified from a paper to be submitted to *Transportation Research Record: Journal of the Transportation Research Board*

Niloo Parvin[1,*], Omar Smadi [1]

Abstract

The United States Road Assessment Program (usRAP) is a powerful tool for conducting Systemic Safety evaluations. The level of safety of the roads can be assessed through the usRAP Star Rating method, giving one star to least safe and five stars to safest roads. As part of the Star Rating data collection process, a comprehensive list of 40 road attributes are recorded for each 100-meter (328 feet) segment using Google StreetView and/or Aerial imagery. Several challenges are associated with usRAP data collection protocols and extensive quality assurance processes are required to ensure data quality. The sources of error are human error, inaccurate measurements/estimations, and the coder's subjectivity in the data collection. To examine the effects of these errors on Star Rating results, this study has leveraged the Second Strategic Highway Research Program (SHRP 2) Roadway Information Database (RID) to complement the existing dataset. The RID includes a variety of safety-related roadway attributes collected by a mobile data collection vendor and meets high accuracy requirements by implementing a quality assurance plan. Using benefit-cost analysis, this study aims to compare the objective data collection approach of utilizing a mobile data collection vendor with high quality assurance processes versus the subjective approach of coding data manually. Star Ratings are calculated for a sample of two lane rural roads in North Carolina using the RID and the manually coded dataset. The more accurate the input road inventory data are, the more proper safety countermeasure suggestions from the Road Assessment Program tool will be expected.

[1] Department of Civil Engineering, Iowa State University, Ames, IA

[*]Primary researcher and author

Introduction

Systemic approach to traffic safety involves identifying high-risk roadway features associated with specific severe crash types. In this approach, safety improvements are recommended not only based on crash history but also the risk of different roadway elements. Systemic approach to safety is a data-driven proactive process that helps to identify low-density severe crashes that are spread throughout the network in rural areas (*4*). Built-in roadway attributes that have been identified as contributing factors to crash likelihood and severity by previous research studies, can determine the road's level of safety regardless of its historical crash experience.

The International Road Assessment Program (iRAP) is one of the tools that has emerged as a powerful platform for conducting Systemic Safety evaluations. iRAP has shown promising results in Europe, Australia, and has recently started to get more utilized in the United States. Pilot usRAP studies have initiated the investigations of the feasibility of usRAP systemic safety program application in the United States. The level of safety of the roads can be shown through Road Assessment Program (RAP) Star Rating method, giving one star to least safe and five stars to safest roads.

In order to develop Star Ratings, road attribute data are collected through a labor-intensive process. A comprehensive list of 40 road attributes are recorded for each 100-meter (328 feet) segment which will be used later for developing Star Ratings and Safer Roads Investment Plans (SRIPs). These road attributes are recorded from roadway imagery by trained coders. In the RAP data collection method, the accuracy of certain data elements is highly dependent on the judgment of the coder regarding the conditions of the road. Based on the target level of accuracy, iRAP divides the collected data into four categories of absolutely-objective, highly-objective, medium-objective, and low-objective (*5*).

In RAP data collection, alternative datasets can be utilized to complement the manual data collection efforts and minimize the coding errors and data subjectivity issues. It is expected that more accurate input data would result in more valid Star Rating results and more detailed safety countermeasure suggestions from the RAP tool. To examine this argument, Star Ratings are calculated for a sample of two-lane rural roads in North Carolina using both an existing roadway inventory of automatically collected data and a manually coded dataset.

In this study, two datasets are collected and used for the assessment of Star Ratings. The first dataset used for Star Ratings is a subset of the Second Strategic Highway Research Program (SHRP 2) Roadway Information Database (RID) which includes a variety of safety-related roadway attributes collected by a mobile data collection vendor. The second dataset used in this study is a manually coded dataset collected from Google StreetView and aerial images. Although this method of data collection can be achieved with a low budget, the accuracy of data is in question due to subjectivity of the coder. Using benefit-cost analysis, this study aims to compare the objective data collection approach of utilizing a mobile data collection vendor with extensive quality assurance processes, versus the subjective approach of coding data manually used in usRAP.

Literature Review

While there has been much research done on the identification of high frequency and hot-spot crash locations, a systemic approach for prioritizing safety improvement locations has not yet been investigated thoroughly. In this section of the paper, five major systemic safety evaluation tools have been reviewed: the Federal Highway Administration Systemic Safety Project Selection Tool (*4*), SafetyAnalyst (*6*), Interactive Highway Safety Design Model (*7*), Minnesota County Road Safety Plan (CRSP) approach (*8*), and the International Road Assessment Program. Table 2.1 includes a summary of different data needs for the systemic tools discussed in the literature review.

Table 2.1 Data needs for systemic tools discussed in the literature review

Systemic tool	**Data needs**
FHWA Safety Project Selection Tool	Minimum crash, roadway, and traffic data needs: system type, crash type, facility type, crash location type, and location characteristics; roadway geometric and traffic elements, site-specific crash information
SafetyAnalyst	Roadway data (segment length and location, area type, cross-section, roadside, intersection, and ramp data); traffic volume; crash data (crash-level, vehicle-level, and person-level)
Interactive Highway Safety Design Model	Varies by IHSDM evaluation module and highway type. Generally includes roadway data (horizontal alignment, vertical alignment, cross-section, lane, roadside, intersection data); traffic operation data (terrain, functional classification, speed, and volume); crash data (optional)
Minnesota County Road Safety Plan	Rural horizontal curve prioritization (curve radius, traffic volume, presence of intersection in curve, visual trap, and crash experience). Rural stop controlled intersection prioritization (intersection skew angle, presence of curve near intersection, commercial development, distance to previous stop sign, ADT ratio, railroad crossing on minor approach, and crash history). Rural segment prioritization (ADT range, access density, roadway departure crash density, critical radius curve density, and edge risk assessment)
International Road Assessment Program	Vehicle occupant star rating: crash likelihood factors (lane width, curvature, quality of curve, delineation, shoulder rumble strips, road condition, grade, skid resistance, centerline rumble strips, number of lanes, differential speeds, intersection type and quality, street lighting, sight distance, intersection channelization, speed management, property access points, service road, and median type); crash severity factors (roadside object, distance to roadside object, paved shoulder width, median type, intersection type, and property access points), traffic operation factors (operating speed, traffic volume, and median traversability)

Federal Highway Administration has established a Safety Project Selection Tool (Systemic Tool) in order to provide state departments of transportation and local government agencies with general processes and steps to embody a systemic approach into their current road safety management plans (*4*). The FHWA systemic tool incorporates three major elements in a cyclical process. The First element includes instructing agencies on development of a systemic safety plan. The second element provides a framework for setting funding goals and balancing systemic and traditional safety investments. Finally, the third element incorporates instructions for the performance assessment of the implemented systemic safety programs. The FHWA report is revised based on the feedback of pilot studies and contains examples of several counties and government agencies that applied the tool in their jurisdiction.

SafetyAnalyst is another highway safety management tool designed to help state and local agencies to proactively identify sites with highest potential for safety improvement. SafetyAnalyst software consists of six safety management tools: network screening, diagnosis, countermeasure selection, economic appraisal, priority ranking, and countermeasure evaluation. The final product of the software is a systemic safety plan containing site-specific improvement recommendations throughout the network. The countermeasure recommendations are based on the expected effectiveness and economic criterion of net benefits and benefit-cost ratio.

Another safety management tool that is widely used is the Interactive Highway Safety Design Model (IHSDM) which incorporates several highway safety modules. IHSDM supports making data-driven geometric design decisions based on the assessment of safety and operational effects. The latest release of the IHSDM software includes a crash prediction module that covers safety evaluations for rural two-lane and multilane highways, urban/suburban arterials, freeway segments and ramps/interchanges. The crash prediction module is an implementation of the Highway Safety Manual (HSM) Part C Predictive Method. The design consistency, traffic analysis, intersection review, and driver/vehicle modules provide diagnostic tools that complement the HSM Part C. IHSDM is a recommended tool for designers, planners, and reviewers who need to conduct a comprehensive evaluation of expected safety and operational performance of different highway facility types in rural and urban settings.

The County Road Safety Plan is a systemic tool which was developed by Minnesota Department of Transportation for all the counties within its jurisdiction. Emphasis areas of Minnesota's CRSP are identifying candidate roadway segments, horizontal curves, and intersections with risk factors associated with severe crashes. Counties may not always have access to complex and expensive software packages, comprehensive and updated crash and roadway data, and human resources needed to prepare a systemic safety plan. Establishing a methodology that can be used for developing a data driven safety plan helps counties to compete for available safety funds.

The International Road Assessment Program serves as an umbrella organization for EuroRAP, AusRAP, and usRAP with the mission of reducing road fatalities by improving road infrastructure. The Road Assessment Program first started developing its efforts in

1991 by a four-country pilot study in Europe and has grown to a program, which is now active in more than 50 countries throughout Europe, Asia Pacific, North, Central and South America and Africa (*9*). In the United States, the usRAP application was tested during the phase I usRAP study in Iowa and Michigan in 2004. The usRAP studies continued with Star Rating and Risk Mapping analyses in phase II (Florida and New Jersey) and phase III (Illinois, Kentucky, New Mexico, and Utah).

iRAP has established four protocols for assessing the risk of fatal and serious injury crashes along the roadway network: Risk Mapping, Performance Tracking, Star Rating, and Safer Roads Investment Plans. The Risk Mapping protocol includes a series of color-coded maps showing 5 categories of roadway crash density from high to low risk. One of the major philosophies of iRAP is replacing blackspot treatments with the proactive approach of systemic safety management for roads, and ultimately with a network level systemic safety management. The Performance Tracking protocol investigates the changes in fatalities and serious injuries of a road segment over a period of time, and also the effectiveness of implemented countermeasures. The third protocol of the iRAP, Star Rating for roads, is based on Road Protection Scores (RPSs). The objective of Star Rating is to identify road design features that are associated with fatal and serious injury crashes. Finally, the Safer Roads Investment Plans recommend cost-effective countermeasures for improvement of roadway Star Ratings.

Methodology

In order to examine the effect of coder subjectivity in the data collection process, a sensitivity analysis was conducted on four variables. The sensitivity analysis compared the result of Star Rating between usRAP and RID datasets. The usRAP dataset includes all the variables which have been collected through review of Google StreetView imagery. The RID dataset has the same values for all variables except for the following four variables: lane width, curvature, grade, and intersection volume; the SHRP 2 Roadway Inventory Database values are used for these variables.

In the usRAP methodology, separate Star Ratings are assigned for vehicle occupants, motorcyclists, bicyclists, and pedestrians by considering operating speed, traffic volume, and other pertinent roadway contributing factors to each of the road users. The

Star Rating methodology is based on run-off road, head-on, intersection and property access crash types which account for a large proportion of road fatalities and serious injuries. In order to develop Star Ratings, a Road Protection Score is calculated for each of the four road users along each 100-meter road segment using equation 1 (*10*).

$$RPS = \sum \text{Crash Type Scores} \quad [1]$$

Where, RPS shows the risk of death and serious injury for an individual road user; and

$$\text{Crash Type Scores} = \text{Likelihood} \times \text{Severity} \times \text{Operating speed} \times \text{External flow influence} \times \text{Median traversability} \quad [2]$$

Where, likelihood indicates risk factors associated with roadway features that account for crash occurrence; severity indicates risk factors associated with roadway features that account for the severity of crash; operating speed accounts for factors that increase or decrease the crash risk with change of speed; external flow accounts for factors that increase or decrease the crash risk with change of traffic volume; and median traversability indicates potential risk factors associated with deviant vehicles crossing a median (only applies to vehicle occupants and motorcyclists run-off and head-on crashes).

This study focuses on assessing Star Ratings for vehicle occupants. The vehicle occupant Star Rating is calculated by summing up the score for the run-off road, head-on (loss-of-control), head-on overtaking, intersection, property access crash types. Table 2.2 demonstrates the contributing factors to likelihood, severity, operating speed, external flow influence, and median traversability for each of the above crash types.

Star Ratings are then assigned to specific ranges of RPS scores. Each range of Road Protection Scores is equivalent to a Star Rating between 1 and 5 as shown in Table 2.3. The safety assessment for the segments can be presented by both raw and smoothed Star Ratings. The raw Star Ratings are allocated to each 100-m segment based on the RPS scores. The smoothed Star Ratings are the average of raw Star Ratings over longer pre-defined homogenous sections. The smoothed Star Ratings provide more meaningful results, especially for a network level analysis. This study uses smoothed Star Ratings in maps and Safer Roads Investment Plans.

Table 2.2 Contributing factors by crash type (vehicle occupant Star Rating)

Road Protection Score	Contributing factors
Run-off score (driver and passenger sides calculated separately)	Likelihood factors (lane width, curvature, quality of curve, delineation, shoulder rumble strips, road condition, grade, skid resistance/grip); severity factors (roadside object, distance to roadside object, paved shoulder width); operating speed; external flow influence; median traversability
Head-on loss-of-control score	Likelihood factors (lane width, curvature, quality of curve, delineation, centerline rumble strips, road condition, grade, skid resistance/grip); severity factors (median type); operating speed; external flow influence; median traversability
Head-on overtaking score	Likelihood factors (number of lanes, grade, skid resistance/grip, differential speeds), severity factors (median type); operating speed; external flow influence
Intersection score	Likelihood factors (intersection type, intersection quality, grade, street lighting, skid resistance/grip, sight distance, channelization, speed management/traffic calming); severity factors (intersection type); operating speed; external flow influence
Property access score	Likelihood factors (property access points, service road, median type), severity factors (property access points); operating speed; external flow influence

Table 2.3 Star Rating bands and colors

Star Rating	Vehicle occupants and motorcyclists for Road Protection Score
5	0 to < 2.5
4	2.5 to <5
3	5 to < 12.5
2	12.5 to < 22.5
1	22.5 +

Data Description

This study involves a comprehensive data collection effort in order to compile more than 40 different variables for 732 records of undivided two lane rural roads (100-meter segments) in Orange County, North Carolina. These variables incorporate an inventory of roadway attributes including roadway geometric features, roadside object, intersection characteristics, curvature, and grade. Additionally, traffic volume data and operating speeds are also needed to calculate the Star Ratings. usRAP has divided the variables into four categories with respect to their level of subjectivity in the coding process (*5*). Table 2.4 summarizes the usRAP's target level of accuracy for coded road attributes. To conduct a sensitivity analysis, two data sets were compiled as part of the study; usRAP dataset and

RID dataset. The usRAP dataset includes all the different variables in Table 2.4 which have been collected as part of this study through the review of Google StreetView and aerial imagery. The RID dataset has the same values for all variables except for the following four variables: lane width, curvature, grade, and intersection volume. In order to have objective and accurate data, these four variables were extracted from the SHRP 2 Roadway Information Database. RID covers the safety-related roadway attributes of about 12,500 centerline miles in the six Naturalistic Driving Study sites (Florida, Indiana, New York, North Carolina, Pennsylvania, and Washington) and meets high accuracy requirements by implementing a quality assurance plan (*11*). RID includes the mobile data and supplementary sets of data from Highway Performance Monitoring System (HPMS), Highway Safety Information System (HSIS), state transportation agencies, and other resources.

Table 2.4 usRAP target level of accuracy and objectivity for roadway attributes

Target level of accuracy	**Variable**
100% (Absolutely-objective)	Coder name, coding date, road survey date, image reference, road name, length, latitude, longitude, carriageway label
98% (Highly-objective)	Distance, speed limit motorcycle speed limit, truck speed limit, differential speeds, median type, intersection type, number of lanes, street lighting, facilities for motorized two wheelers, bicycle facility, pedestrian crossing - inspected road, sidewalk - driver side, sidewalk - passenger side
95% (Medium-objective)	Upgrade cost, centerline rumble strips, roadside severity - driver side distance, roadside severity - driver side object, roadside severity - passenger side distance, roadside severity - passenger side object, shoulder rumble strips, paved shoulder - driver side, paved shoulder - passenger side, intersection channelization, property access points, lane width, motorcycle observed flow, bicycle observed flow, pedestrian observed flow across the road, pedestrian observed flow along the road driver-side, pedestrian observed flow along the road passenger side, pedestrian crossing facilities - side road, pedestrian fencing, school zone warning, school zone crossing supervisor, curvature, vehicle parking, service road, roadworks
90% (Low-objective)	Landmark, area type, intersecting road volume, land use - driver side, land use - passenger side, pedestrian crossing quality, intersection quality, quality of curve, grade, road condition, skid resistance/grip, delineation, speed management/traffic calming, sight distance
Not applicable	Section, comments, vehicle flow (AADT), operating speed (85th percentile), operating speed (mean), car star rating policy target, annual fatality growth multiplier, motorcycle %, pedestrian peak hour flow across the road, pedestrian peak hour flow along the road driver-side, pedestrian peak hour flow along the road passenger-side, bicycle peak hourly flow, motorcycle star rating policy target, pedestrian star rating policy target, bicycle star rating policy target

As part of the RID quality assurance (QA) process, a comprehensive plan was developed that defined the minimum accuracy requirements for the collected data. The QA plan included random site visits where the ground truth measurements were collected in the field by the research team. In North Carolina, a total of 346 random site visits were conducted on a representative sample of data during the data collection period (2011-2013). Additionally, to assure the proper operation of vendor's data collection equipment, control sites were selected within the six NDS sites.

The ground truth data were compared and checked against the vendor's collected data for these control sites along the different stages of the data collection process. The minimum required accuracy requirements for the 3 data elements are as described below:

- Curvature Radius:
  - 100 ft. for curves less than 1,500 ft. radius
  - 250 ft. for curves between 1,500 ft. and 6,000 ft. radius
  - Within 13% for curves over 6,000 ft. radius
- Grade (+ or -): 1.0%
- Lane width: 1 ft.

Table 2.5 compares the values of RID and usRAP dataset by demonstrating the percentage of segments within each category of lane width, curvature, grade, and intersection volume.

Table 2.5 Comparison of lane width, curvature, grade, and intersection volume between RID and usRAP dataset

Variable		**Percentage of road segments**		**Level of accuracy**
		RID dataset	usRAP dataset	
Lane width	Wide (≥ 10.6 ft.)	13.5%	36.1%	55%
	Medium (≥ 9.0 to <10.6 ft.)	74.3%	63.9%	
	Narrow (< 9.0 ft.)	12.3%	0.0%	
Curvature	Straight or gently curving	68.6%	88.0%	56%
	Moderate	8.5%	11.1%	
	Sharp	16.2%	0.9%	
	Very sharp	6.7%	0.0%	
Grade	≥ 0% to <7.5%	99.6%	88.9%	79%
	≥ 7.5% to <10%	0.4%	10.6%	
	≥ 10%	0.0%	0.4%	
Intersection Volume	1,000 to 5,000 vehicles	6.1%	7.5%	99%
	100 to 1,000 vehicles	5.4%	4.0%	
	1 to 100 vehicles	3.7%	2.5%	
	Not applicable (Not intersection)	84.7%	85.9%	

A general trend of lane width overestimation can be observed from Table 2.5. In the usRAP data collection process, the coder uses online measurement tools over the road imagery to estimate the lane width for the first few segments of each road. The usRAP dataset also underestimates the roadway curvature. The coder would identify the road curvature as one of the following categories (usRAP coding manual):

- Very sharp (The road contains curves which can only be driven at less than 25 mph, approximate radius of curve <656 ft.),
- Sharp (The road contains sharp curves which can only be driven between 25 and 44 mph, approximate radius of curve 656 to 1640 ft.),
- Moderate (The road has fairly tight curves which can be driven at less than 62 mph but more than 44 mph, approximate radius of curve 1640 to 2953 ft.), and
- Straight or gently curving (The road contains only long curves which can be driven at 62 mph or more, approximate radius of curve >2953 ft.).

We can also observe from Table 2.5 that the usRAP coding is underestimating number of the segments with a grade between 0% and 7.5%. The percentage of the segments with a moderate grade of 7.5% to 10% are 10.6% in the usRAP dataset compared to less than 1% in the RID dataset. Table 2.5 shows the difference of the coded intersection volumes between the two datasets. If a segment included an intersection, the intersection volume has been coded based on the following assumptions:

- Undivided unpaved one-lane roads have an estimated volume of 1 to 100 vehicles,
- Undivided paved one-lane roads have an estimated volume of 100 to 1000 vehicles, and
- Undivided paved two-lane roads have an estimated volume of 1,000 to 5,000 vehicles.

The RID dataset uses the North Carolina Department of Transportation AADT data to determine the intersection volume and is based on the actual count of traffic data. The criteria for separating the private property access points from the public intersections has been based on the condition of presence of stop sign or presence of public road name sign. This criteria is different between RID and usRAP and has resulted in different numbers of intersections between the two datasets.

There are several challenges associated with the process of usRAP data collection. Given the large amount of manual data collection needed to evaluate the Star Ratings for roads, it is crucial to undertake extensive quality assurance processes in order to produce consistently high-quality data. While usRAP defines the lane width and curvature as medium-objective variables with 95% accuracy, the coded sample data were found to be 55% and 56% accurate. The grade and intersection volume variables are defined as low-objective data with 90% level accuracy, were found to have 79% and 99% accuracy. Some of the variables that require use of online measurement tools, such as lane width and curvature may be overestimated or underestimated along the course of the data collection. Additionally, there are variables with a subjective nature such as quality of curve, quality of intersection, and sight distance that is highly dependent on the opinion of the coder.

Results and Discussion

Separate analyses were conducted for the RID dataset and usRAP dataset for 45 miles of undivided two lane rural roads in North Carolina. Results of the Star Rating for usRAP and RID dataset are shown in Table 2.6 along with information about the number of segments, total length, and average AADT for each road.

Table 2.6 Star Rating results for usRAP and RID datasets

Road Name	100-m Segments	Length(mi)	Average AADT	RID Dataset - Star Rating	usRAP Dataset - Star Rating
NC 86	304	18.9	6,431	2	2, 3
NC 57	130	8.1	4,419	2, 3	3
St Marys Rd	90	5.6	2,996	2, 3	3
New Hope Church Rd	66	4.1	2,167	2	3
Arthur Minnis Rd	54	3.4	2,433	1, 2	1, 3
Old NC 86	46	2.9	4,794	2, 3	3
Hillsborough Rd	42	2.6	4,954	2	3
Sum or weighted average	732	45.5	2,8194	2	3

The values of Star Rating range from 1 to 3.The Star Rating of 2 is the most common in the RID dataset and the Star Rating of 3 is the most common in the usRAP dataset. Figure 2.1.a and Figure 2.2.b show the resulted Star Rating maps where high risk roads are shown as a black line (1-star) or red line (2-star) on the map. High risk roads are likely to be undivided, with hazardous roadside objects, stop-controlled 3- and 4-leg intersections and relatively high speed limits. The intermediate risk roads are shown as an orange line (3-star) on the map and will likely be undivided, with narrow paved shoulders and some roadside hazards. Turning to the primary factor of interest, the Star Rating results are significantly different between the two datasets. The over-estimation of lane width, moderate grades and under-estimation of sharp curves in the usRAP dataset have resulted in inaccurate estimation of level of safety of the roads. While the Star Rating of most roads are 2 in reality, they have been evaluated as 3-Star Rating falsely in the usRAP dataset.

Among all the 732 segments, 42% of the segments have the same smoothed Star Rating values in both RID and usRAP datasets. Figure 2.3 shows how each variable contributes to the Star Rating differences between the two datasets. It can be observed from the Figure 2.3.a that even though 52% of the segments have at least one different variable, the inaccuracies do not affect the overall assessment of smoothed Star Rating. A possible explanation for this might be that when 100-m segments are smoothed over the longer sections of the road, the raw Star Rating of these segments lose their significance when they are averaged with their adjacent segments. The majority of the coding differences between the two datasets with unchanged Star Rating come from curvature (18%), lane width (14%), and intersection volume (7%).

About 58% of segments were found to have an over-estimation in their Star Rating evaluations. Figure 2.3.b demonstrates a comparison of the investigated variables between the two datasets. The two largest contributing factors to Star Rating over-estimation are inaccurate curvature (25%) and grade (25%) values. Interestingly, 24% of the segments had the exact same values in usRAP and RID dataset. Similar to the segments with unchanged Star Rating, this could be explained by the fact that the error from these segments are neutralized during the smoothing process when they are averaged with the adjacent segments. Overall, it can be concluded that the Star Rating assessments in two lane rural roads are more sensitive to the curvature and grade variables compared to lane width and intersection volume. These findings cannot be extended to other road types (i.e. interstates, freeways, etc.) since these variables may play a different role in assessing their level of the safety.

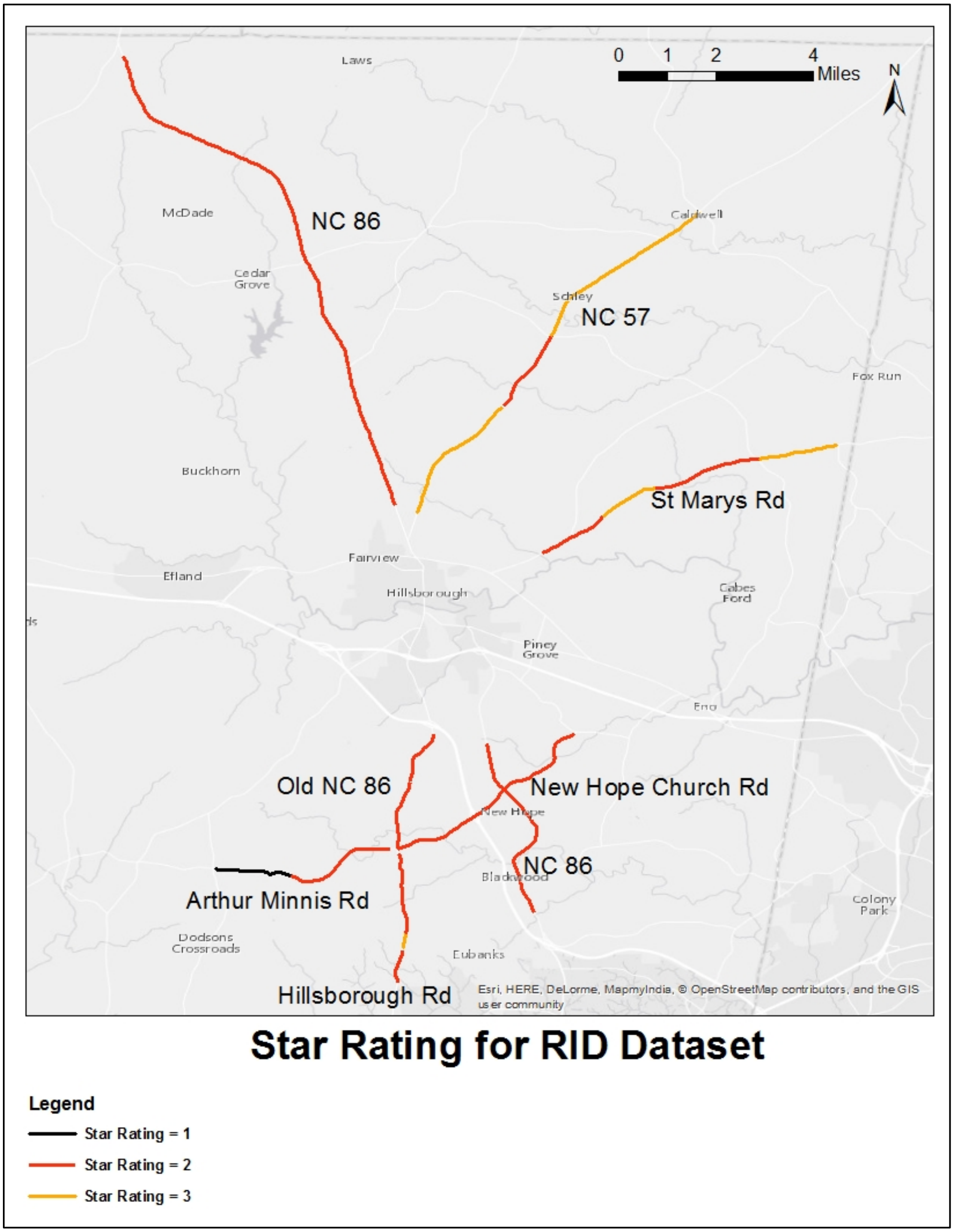

Figure 2.1.a Star Rating map for RID dataset

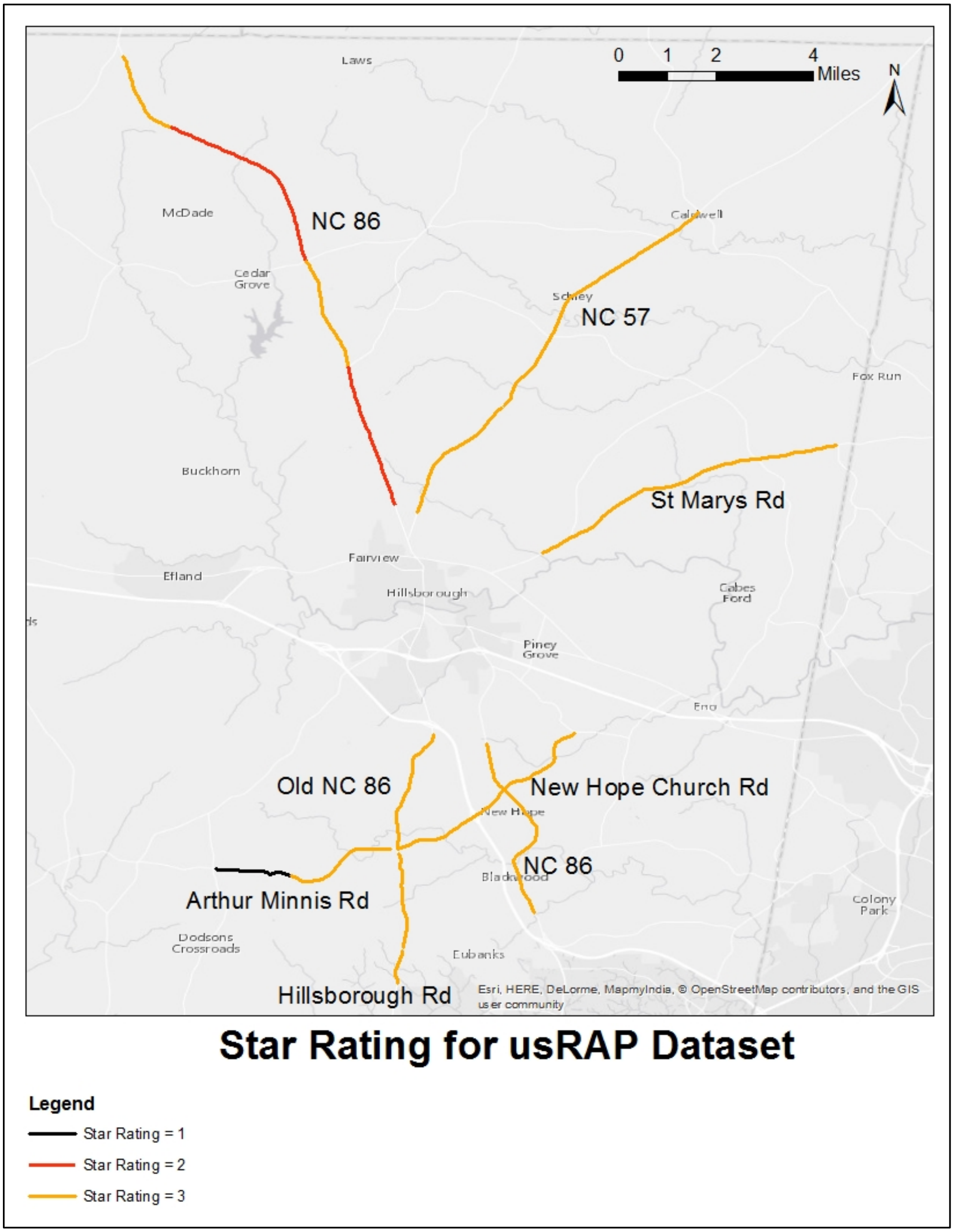

Figure 2.2.b Star Rating map for usRAP dataset

The usRAP methodology also includes the Safer Roads Investment Plans with a list of more than 70 proven countermeasure options. The countermeasures include a variety of options ranging from low-cost improvements such as improving delineation and signing to improvements with higher costs such as roundabouts or additional lanes. In this study, certain economic assumptions are made in order to estimate the benefit and cost of each improvement option. The analysis period is assumed to be 20 years. A Gross Domestic Product per capita of $56115.7 is considered along with discount rate of 3% and minimum attractive rate of return of 0.03. Additionally, the value of life multiplier and value of serious injury multiplier are assumed to be 70 and 0.25 respectively (*12*).

The predicted number of fatal and serious injury crashes are calibrated based on the actual crash experience of the study segments within 8 year period of 2006 to 2013. The SRIP provides the following data assuming all the recommended countermeasures are implemented: an estimate of the total number of fatal and serious injuries (FSIs) that could be prevented over the life of the plan (20 year analysis period), an estimate of the total present value of the economic benefits from crash cost savings, estimated cost of implementation and maintenance of countermeasures, cost per FSI saved, and program's Benefit-cost Ratio (BCR).

In this study, the SRIP for each dataset resulted in different countermeasures. While the recommended countermeasures for the RID dataset ended up in 366 saved fatal and serious injury with benefit-cost ratio of 34, the estimation for usRAP dataset was 258 FSIs and program's BCR was 25. The difference between the SRIP of the two datasets is a result of inaccuracies found in the values of lane width, curvature, grade, and intersection volume in the usRAP dataset.

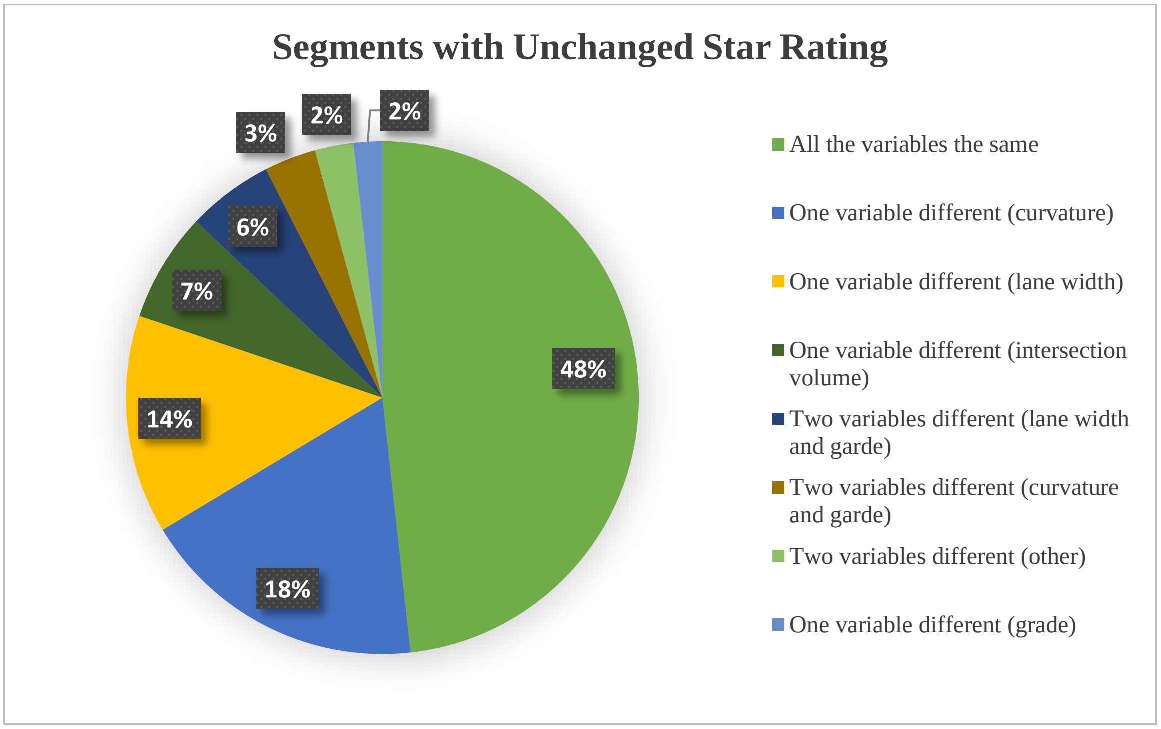

Figure 2.3.a Comparison of segments with unchanged smoothed Star Ratings between RID and usRAP dataset

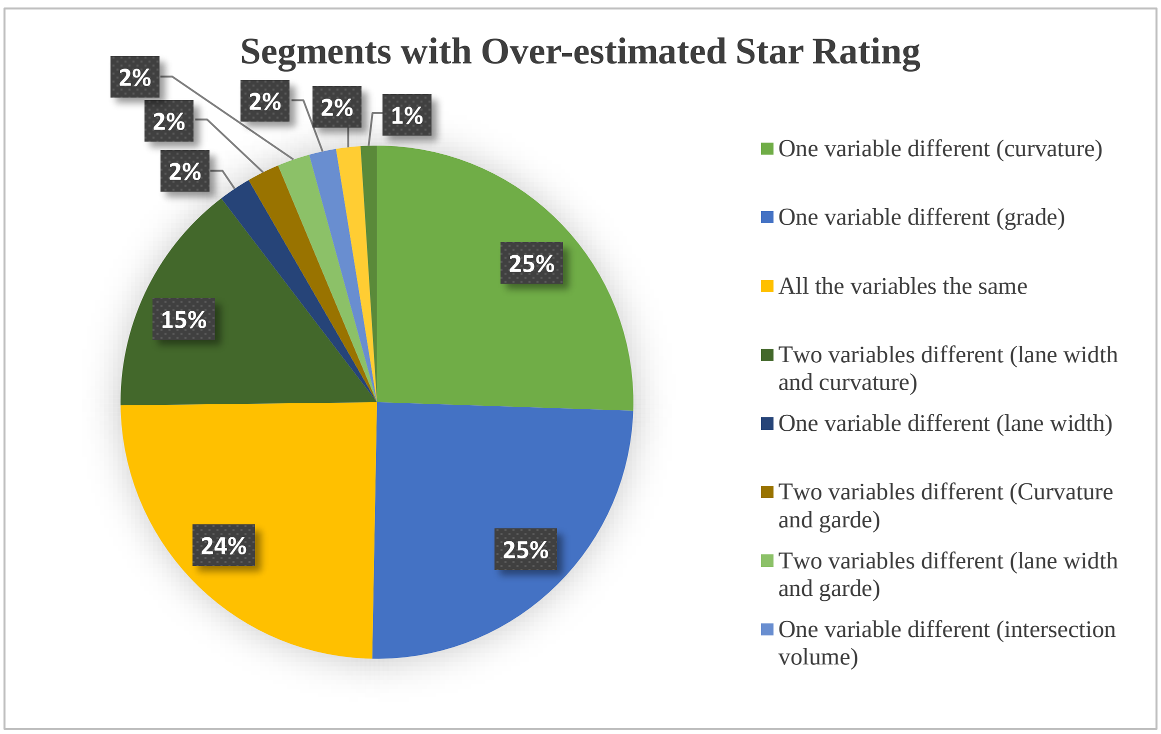

Figure 2.3.b Comparison of segments with over-estimated smoothed Star Ratings between RID and usRAP dataset

Conclusion

This paper analyses the impact of utilizing an objective and automated data collection approach versus the use of the subjective approach of manual data collection for certain roadway attributes. Star Ratings are calculated for 45 miles of high speed two lane rural roads in North Carolina using RID dataset and usRAP dataset. The RID dataset has more accurate data whereas the usRAP dataset has some overestimations or underestimations in the values of variables (i.e. lane width, curvature, grade, and intersection volume). The results of Star Rating analysis for RID dataset showed that most of the study network contains high risk roads with a 2-Star Rating. However, the usRAP dataset assessed most of the network to have medium risk with Star Rating of 3. Another important finding was that the Safer Roads Improvement Plans were significantly different between the two datasets. A comparison of benefit-cost ratio of two datasets revealed that RID dataset not only has a higher BCR ratio but also resulted in a more comprehensive plan to address the safety issues. Overall, these results indicate that there is great benefit in having an objective, comprehensive, and accurate roadway data to assess the level of safety of the roads.

This research extends our knowledge in examining the effect of having high quality data on assessing the level of safety of the roads. Additionally, it has several practical applications. First, it will serve as a base for future projects and studies in the area of systemic safety evaluation of roads for public and private agencies. Second, it demonstrates that investing in acquiring high quality data will end up in more accurate evaluation of road conditions and a more comprehensive list of countermeasures.

Several additional questions can be addressed by further research. First, more investigations are warranted into finding variables with highest level of sensitivity for roads with different functional classes (i.e. freeway, and expressway roads) in the Star Rating methodology. Also, this study assumes the following relationship: the crash rate is expected to decrease, as the Star Rating increases. Future research could explore the relationship between crash frequency or crash rate and Star Rating. Additionally, it would be interesting to compare the more traditional safety evaluation methods such as safety performance functions with the systemic safety evaluation methods.

CHAPTER 3. AN INVESTIGATION OF RELATIONSHIP BETWEEN THE UNITED STATES ROAD ASSESSMENT PROGRAM STAR RATING AND CRASH EXPERIENCE

Modified from a paper to be submitted to *Transportation Research Record: Journal of the Transportation Research Board*

Niloo Parvin[1,*], Omar Smadi[1]

Abstract

Over the recent decades, the International Road Assessment Program (iRAP) has been widely used as a systemic safety management tool. iRAP uses the risk-based non-crash measure of Road Protection Score (RPS) for assessing the level of safety of the roads from a 1 to 5 Star Rating scale. One of the most significant current research needs is the validation of the relationship between the crashes and Star Ratings given that few published studies exist in this area. Moreover, the previous validation studies have been mostly limited only to the comparison of crash rate and Star Rating averages and have failed to establish a comprehensive statistical relationship. In order to investigate such relationship, this study develops a crash prediction model using a sample of two lane rural roads in North Carolina. The crash frequency was estimated as a function of Road Protection Score and Annual Average Daily Traffic using a negative binomial model. The results of this study showed that the crash frequency consistently increases with Road Protection Score. The developed safety performance function showed that moving from a 3-star road to a 2-star road would result in 47% more crashes. These findings confirm that Star Rating is a valid risk measure for crash frequency on two lane rural roads.

[1] Department of Civil Engineering, Iowa State University, Ames, IA

[*]Primary researcher and author

Introduction

In recent years, there has been an increasing interest in utilizing systemic safety management tools in the United States. Transportation agencies have started to balance their efforts between the traditional hot-spot safety identification methods and systemic safety improvement management approaches. The systemic safety management involves a proactive approach towards identifying potential improvement locations with low-density severe crashes (*4*). Systemic safety management tools are specifically helpful where crash data are not available or where the crashes are spread throughout the network (i.e. rural areas). Local agencies with limited or no access to crash data and roadway data will find the systemic safety evaluation methods that have less data requirements more practical to use.

The International Road Assessment Program (iRAP) is one of the most widely used tools of systemic safety management around the globe (*13*). The United States Road Assessment Program (usRAP) is one of the branches of iRAP with the objective of providing a method to benchmark the safety performance of roads in the U.S. using the following process; first, Star Ratings are assigned to roads based on their built-in engineering features. Second, a Safer Roads Investment Plan (SRIP) is recommended for the road segments (*14*). Star Rating assessment is a unique approach in the aspect that it demonstrates crashes likelihood of fatal and serious injury crashes without requiring the historic crash data.

Over the past decade, most research studies on the Road Assessment Programs (RAPs) have emphasized the implementation of the assessment methods and evaluating the Star Ratings for roads. Few studies have investigated the validation of the relationship between Star Rating and crash rate in any systematic way. Critics have questioned if the Star Ratings are demonstrating the road’s crash experience accurately. This paper attempts to investigate this relationship using the data for a sample of 40 miles of two lane rural roads in North Carolina.

Literature Review

Road Protection Score (RPS) is a continuous risk-based measure ranging from 0 to 200 and each range of Road Protection Scores is equivalent to a Star Rating between 1 and 5 (*15*). The RPS assessment methodology has evolved throughout the years with growth and development of iRAP. The first version of RPS, EuroRAP RPS1.0, only considered the safety-related crash protection factors. The second version, EuroRAP RPS2.0, added the crash likelihood factors to the methodology and the RPS values were translated into a Star Rating of 1 to 4 for the vehicle occupants (*16*). The next version, iRAP RPS, provided the scores for four different types of road users, vehicle occupants, motorcyclists, bicyclists, and pedestrians which were converted into a 1 to 5 Star Rating range.

Several studies have focused on the comparison of average crash rates (or costs) and the RAP's Star Rating or Road Protection Score (See Table 3.1). Previous studies provide general findings that partly validate the expected correlations between the crash data and RPS. These findings may be limited by some key challenges such as small sample size, unaccounted human and vehicle-related factors, and crash data quality issues. In many cases, the small samples of Fatal and Serious Injury (FSI) crashes did not allow for a comprehensive analysis. Also, there were situations where high number of FSIs were occurring on safe roads (i.e. 4-star or 5- star roads) that could not be explained by Star Ratings since they were the results of contributing human factors (i.e. speeding, drunk driving). Similarly, the Star Ratings could not explain the crashes that occur as a result of vehicle defects. Finally, some studies had the challenge of using crash data with under-reported FSIs, miscoded locations, and inaccurate traffic flow information.

Two methodologies are common in the literature for making the Crash-Star Rating comparison: qualitative and visual comparisons from maps, and quantitative comparisons of average crash rate versus Star Rating/RPS (*17*). The EuroRAP studies in England and Iceland compared the crash rate Risk Maps and Star Rating maps (*18*)(*19*, *20*). The English motorway network was identified as low risk in both risk maps and Star Rating maps. Also, the Iceland national road network segments that were identified with a high Star Rating, demonstrated lower crash rates on maps.

Table 3.1 Summary of RAP studies on comparison of RPS/Star Rating with crash rate

Study	Sample size (mi)	Road type	Summary
EuroRAP-Sweden (2007)	5,500	Single carriage way, motorway, 2+1 roads, 4-lane roads, and other roads.	A general correlation was found between RPS and FSI crash rate (EuroRAP RPS1.0).
AusRAP-Australia (2008)	3,300	High-speed roads in open rural settings in Queensland.	The average crash costs per kilometer travelled consistently decreases with the increase in Star Rating (AusRAP).
EuroRAP-England (2009)	3,200	Motorways, dual, mixed, and single.	FSI crash rates decreased with increase in Star Rating (EuroRAP RPS1.0).
usRAP-United States (2010)	3,000	Rural and urban roads of various types in Iowa and Washington.	The vehicle occupant Star Rating for 100-m sections decrease with increase in Star Rating values (iRAP).
EuroRAP-Hungary (2016)	900	Rural undivided roads in Hungary.	The study confirmed the expected relationship but the empirical Bayes model outcomes assessed the Star Rating to have a minor impact and an unexpected positive association with crash frequency.

Studies with quantitative methods, involved an investigation into the relationship between predicted risk (RPSs or Star Ratings) and the observed risk (crash rates). The AusRAP study, confirmed the association between the Star Ratings and the crash costs using an extensive sample of Queensland roadway data (*21*). The most interesting finding of the study was that a more meaningful analysis could be conducted if the crash cost per vehicle kilometers travelled is substituted with crash rate. The crash cost values are calculated by multiplying the number of fatal and serious injury crashes by their recovery costs; this means that more weight is put on fatal crashes compared to serious injury crashes. The AusRAP study results show the economic benefits that could be gained by upgrading the Star Ratings of the roads (e.g. the average crash cost per vehicle kilometer travelled would increase $0.051 when moving from a 3-star to a 2-star road). The AusRAP study concluded that the increases in Star Ratings are associated with lower crash costs per vehicle kilometer travelled.

The usRAP phase III pilot study is another comprehensive study that is conducted to establish a relationship between Star Rating and crash rate (*22*)(*23*). The study provided Star Ratings for about 3,000 miles of rural and urban roads in the states of Iowa and Washington. The study concluded that a statistically significant relationship exists between the decrease of crash rate and increase of Star Rating for two-lane undivided roads, four-lane undivided roads, and four-lane divided non-freeways. A possible explanation for finding such consistence correspondences could be that the usRAP study utilized a large sample of high quality data that have been divided to homogenous segments based on their functional class.

A recent study by Ambros et al. discussed an alternative method for validation of iRAP Star Rating using a sample of 900 miles of Hungarian rural road network (*24*). To better understand the relationship between the Star Rating and crashes, the study utilized an empirical Bayes approach to develop a crash prediction model using both observed and expected crash frequency. The study concluded that using the empirical Bayes (EB) approach, results in a better fit for the data compared to a linear approach since the EB method accounts for the effect of regression to the mean in the crash data. Even though the study confirmed the relationship between increasing Star Ratings and decreasing crash frequencies, the model results estimated the Star Rating to have a minor influence with an unanticipated positive correlation with crash frequency.

To summarize, the most robust validation studies took advantage of large sample sizes along with high quality crash data and homogenous road segments in their analysis to show a significant decrease in average crash rate with increase in Star Rating. The findings suggest that a reduction of about a third in fatal and serious injury crash rate (Sweden) and a half in crash cost (Australia) was shown with change from a 2-star to a 3-star road (*18*)(*21*). Moreover, the reduction in crash rate or crash cost is less substantial when a comparison is made between 3-star and 4-star roads or between 4-star and 5-star roads.

Methodology

As discussed in the literature review, the past studies have utilized two approaches for validation of Star Rating/Road Protection Score and crash rate relationship. One of the most common analysis methods is illustrating the general trends of Star Rating by average crash rate through graphical. The alternative approach for ascertaining the relationship between the crashes and Road Protection Score is development a statistical model. In this study, the existence of such relationship is explored by both approaches.

In the first part of the analysis, the general trends of the data are demonstrated as described in the following. The first step of the process is to calculate the average of Star Ratings among the segments for each road. The second step involves smoothing (averaging) crash rates for each road and then summarizing them for each Star Rating. The final outcome of this process is a bar chart presenting crash rate versus the Star Rating. While some studies have used the aggregated crash rate for all the segments, this study has averaged the crash rate data by road. Additionally, crash cost by Star Rating charts have been created. By multiplying the monetary value of life and serious injury by the number of FSI crashes, more weight is assigned to the fatal crashes. The crash costs are divided by Vehicle Miles Traveled (VMT).

This study also compares the use of aggregated and disaggregated Star Rating (SR) approaches in the validation process. The usRAP methodology provides both raw and smoothed Star Ratings. The raw SR is calculated for each 100-m segment, whereas, the smoothed SR is calculated for longer lengths of pre-defined homogenous sections that contain several segments. The smoothed SR provides an opportunity to group similar segments in order to produce more practical results for the Star Rating and the implementation of countermeasure plans. This study compares both smoothed and raw SRs against the crash rate/cost in the first part of the analysis

The second part of the analysis involves finding a statistical relationship between crash rate, traffic volume, and Road Protection Score. Crash data are a form of count data that consist of non-negative integer values, so the standard least squares regression methods cannot be applied (*25*). The Poisson regression model can be used in count data frameworks. An assumption for the Poisson model is that the average of observations should be approximately equal to their variance; however, this is not always true for crash data where the variance of crash counts is generally greater than the mean (over-dispersion in crash data). To address this issue, a negative binomial model has been fitted to the data. Equation 1 illustrates the negative binomial functional form of parameter for each observation i:

$$\lambda_i = \mathrm{EXP}(\beta X_i + \varepsilon_i) \qquad [1]$$

Where X_i is descriptive variable for crashes frequency i and $\mathrm{EXP}(\varepsilon_i)$ is the gamma-distributed error term with the mean of 1 and the variance of α.

Data Description

This study involves an extensive data collection process in order to gather more than 40 safety-related variables for 40 miles of roads in North Carolina. The focus of the study is two-lane rural roads with speed limits of 45 mph and above (See Figure 3.1). The usRAP variables are comprised of several general categories of data (*10*). The first category includes the factors that contribute to crash likelihood; many geometric features of the road are part of this category (e.g. lane width, curvature, grade, intersection characteristics). Another category of collected data belongs to the factors that contribute to crash severity including parameters related to the roadside features (e.g. roadside object, distance to roadside object, paved shoulder width). These two categories of data are collected manually through reviewing satellite and roadway imagery of the 100-meter segments of the road. Supporting data such as operating speed and traffic volume belong to the third category of data.

The usRAP data collection process involves several challenges which warrant discussion. During the data collection process, it is necessary to conduct extensive quality assurance attempts to ensure the avoidance of systematic errors throughout the dataset. The sources of error in the dataset are human error, inaccurate measurements/estimations, and the coder's subjectivity in the data collection. Given the large amount of manual data collection required, the coders are prone to fatigue and false entry of the data and inconsistencies between different coders are likely. To avoid such errors, this study has leveraged the Second Strategic Highway Research Program (SHRP 2) Roadway

Information Database (RID) to complement the existing dataset. The Roadway Information Database has been collected by a mobile data vendor and has gone through an extensive quality assurance plan to meet the accuracy requirements (*11*). Four variables that are expected to have higher levels of inaccuracy and are challenging for the coder to estimate have been substituted by their value from the RID dataset. These variables are lane width, curvature, grade, and intersection volume.

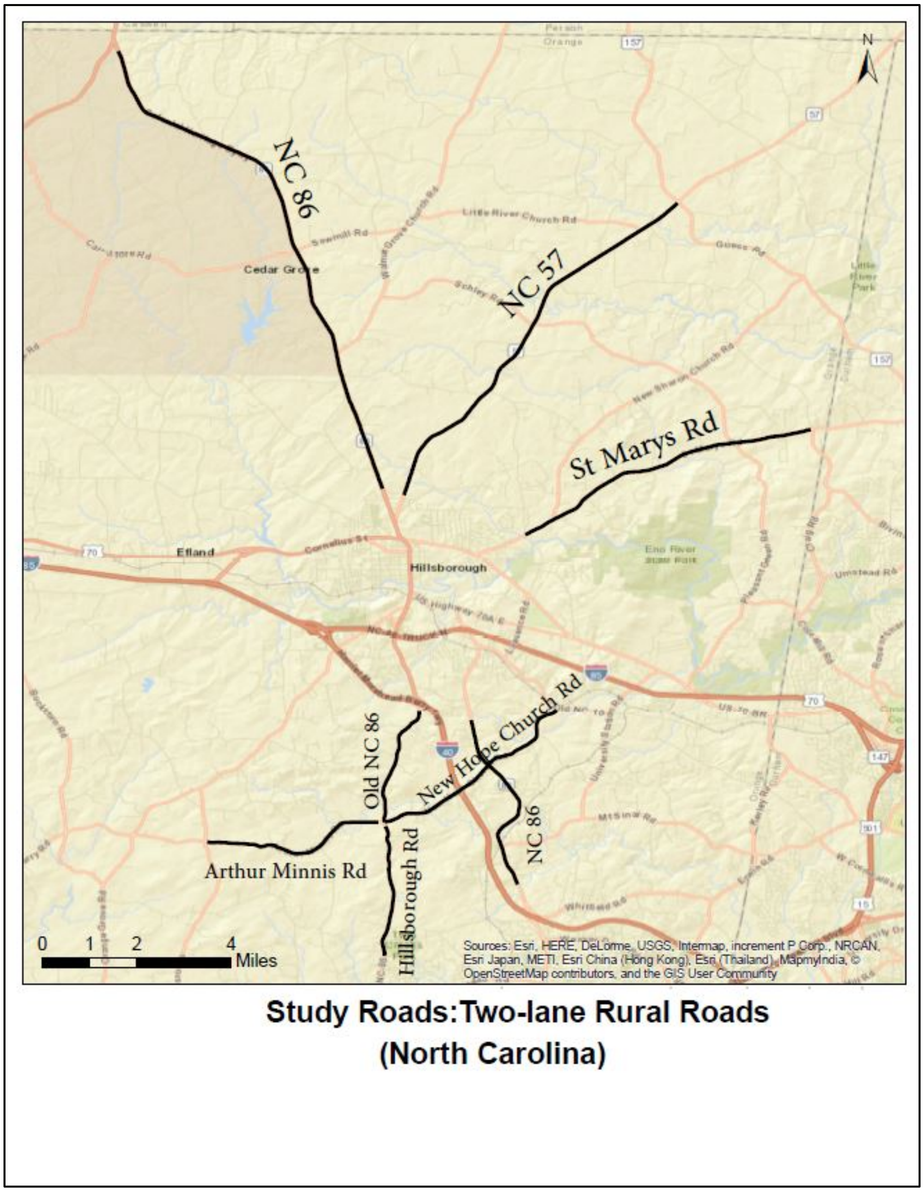

Figure 3.1 Map for the two lane rural roads included in the study

Table 3.2 provides a summary of the key variables in the compiled dataset. Although the dataset contained additional fields, such as indicators for the presence of motorcycles, bicycles, pedestrians, etc., there were too few segments with these features to allow for a meaningful analysis. The sample includes an inventory of roadway characteristics for 647 100-meter segments in the rural areas of Orange County, North Carolina where most of the land use along the segments are undeveloped or farming and agricultural. Additionally, non-frangible signs/posts/poles and trees were frequently found along the road usually located within 3 to 30 feet of the edge lines. Most of the road segments had narrow paved shoulders or no paved shoulders at all. Additionally, intersections were not present in most of the segments, but many roads included one or two residential access points. Many roads in the sample had medium width (9 to 10.6 feet) with no curvature and no grade. The Annual Average Daily Traffic was 4,700 on average for all the segments and its values ranged from 1,867 for the lower functional class roads to 10,133 for the higher functional class roads.

Table 3.2 Frequency and percentage of key variables

Variable	**Variable description**	**Count**	**Percent**
Land use – driver side	Educational	3	0.5%
	Commercial	14	2.2%
	Industrial and manufacturing	0	0.0%
	Residential	51	7.9%
	Farming and agricultural	254	39.3%
	Undeveloped areas	325	50.2%
Land use – passenger side	Educational	7	1.1%
	Commercial	14	2.2%
	Industrial and manufacturing	2	0.3%
	Residential	46	7.1%
	Farming and agricultural	251	38.8%
	Undeveloped areas	327	50.5%
Speed limit	55mph	207	32.0%
	50mph	33	5.1%
	45mph	407	62.9%
Roadside severity – driver side distance	0 to <3 ft.	20	3.1%
	3 to <15 ft.	213	32.9%
	15 to <30 ft.	269	41.6%
	>=30ft	145	22.4%

Table 3.2 continued

Variable	Variable description	Count	Percent
Roadside severity – driver side distance	Tree >=4 in.	413	63.8%
	Non-frangible sign/post/pole >=4 in.	164	25.3%
	Upwards slope - (15° to 75°)	2	0.3%
	Deep drainage ditch	5	0.8%
	Downwards slope (> -15°)	8	1.2%
	Non-frangible structure/bridge or building	4	0.6%
	Frangible structure or building	2	0.3%
	Safety barrier - concrete	1	0.2%
	Safety barrier - metal	11	1.7%
	No object	37	5.7%
Roadside severity – passenger side distance	0 to <3 ft.	12	1.9%
	3 to <15 ft.	198	30.6%
	15 to <30 ft.	310	47.9%
	>=30ft	127	19.6%
Roadside severity – passenger side object	Tree >=4 in.	344	53.2%
	Non-frangible sign/post/pole >=4 in.	241	37.2%
	Upwards slope - (15° to 75°)	1	0.2%
	Deep drainage ditch	7	1.1%
	Downwards slope (> -15°)	6	0.9%
	Frangible structure or building	1	0.2%
	Safety barrier - concrete	1	0.2%
	Safety barrier - metal	11	1.7%
	No object	35	5.4%
Shoulder rumble strips	Not present	471	72.8%
	Present	176	27.2%
Paved shoulder - driver side	None	371	57.3%
	Narrow (≥ 0ft to < 3.0 ft.)	276	42.7%
	Medium (≥ 3.0ft to < 7.9 ft.)	0	0.0%
	Wide (≥ 7.9 ft.)	0	0.0%
Paved shoulder - passenger side	None	371	57.3%
	Narrow (> 0ft to < 3.0 ft.)	276	42.7%
	Medium (≥ 3.0ft to < 7.9 ft.)	0	0.0%
	Wide (≥ 7.9 ft.)	0	0.0%
Intersection type	4-leg unsignalized with no protected turn lane	9	1.4%
	4-leg unsignalized with protected turn lane	1	0.2%
	4-leg signalized with no protected turn lane	3	0.5%
	3-leg unsignalized with no protected turn lane	83	12.8%
	3-leg unsignalized with protected turn lane	5	0.8%
	None	546	84.4%
Intersecting road volume	≥15,000 vehicles	0	0.0%
	10,000 to 15,000 vehicles	0	0.0%
	5,000 to 10,000 vehicles	0	0.0%
	1,000 to 5,000 vehicles	41	6.3%
	100 to 1,000 vehicles	36	5.6%
	1 to 100 vehicles	24	3.7%
	Not applicable	546	84.4%
Intersection quality	Poor	13	2.0%
	Adequate	88	13.6%
	Not applicable	546	84.4%

Table 3.2 continued

Variable	Variable description	Count	Percent
Lane width	Narrow (> 0ft to < 9.0 ft.)	82	12.7%
	Medium (≥ 9.0 to <10.6 ft.)	475	73.4%
	Wide (≥ 10.6 ft.)	90	13.9%
Curvature	Very sharp	40	6.2%
	Sharp	104	16.1%
	Moderate	55	8.5%
	Straight or gently curving	448	69.2%
Quality of curve	Poor	30	4.6%
	Not applicable	521	80.5%
	Adequate	96	14.8%
Grade	≥ 10%	0	0.0%
	≥ 7.5% to <10%	0	0.0%
	≥ 0% to <7.5%	647	100.0%
Sight distance	Poor	164	25.3%
	Adequate	483	74.7%
Vehicle flow (AADT)	Minimum	1,867	
	Maximum	10,133	
	Average	4,698	
Property access points	Commercial Access ≥1	28	4.3%
	Residential Access ≥3	39	6.0%
	Residential Access <3	291	45.0%
	None	289	44.7%

This study has also utilized the crash data from the RID dataset for an 8-year period of time. A total number of 942 crashes, 24 Fatal and Serious Injury (FSI) crashes, 12 fatalities, and 17 serious injuries have occurred within 150 feet of the sample roads from 2006 to 2013. Table 3.3.a provides information on number of segments, length, AADT, and crash experience for the sample roads.

A summary of average Star Ratings, crash rates, and crash costs are shown in Table 3.3.b. Also, the crash costs are calculated for fatal and serious injury crashes, assuming the unit cost of $10,462,000 for each fatality and $590,000 for each serious injury (*26*)

Table 3.3.a Mileage, AADT, and crash experience of study road segments

Road Name	100-m Segments	Length (mi)	Average AADT	Total Crashes	Fatalities	Serious Injuries
NC 86	240	14.9	6431	439	9	7
NC 57	130	8.1	4419	153	1	4
St Marys Rd	90	5.6	2996	101	0	1
New Hope Church Rd	66	4.1	2167	80	0	0
Arthur Minnis Rd	46	2.9	4794	82	2	4
Old NC 86	42	2.6	4954	74	0	1
Hillsborough Rd	33	2.1	2433	13	0	0

Table 3.3.b Summary of the Star Ratings, crash costs, and crash rates

Road Name	Smoothed Star Rating (avg.)	Raw Star Rating (avg.)	FSI crash cost (avg.)	FSI crash cost per Vehicle Miles Travelled (avg.)	Crashes per Million Vehicle Miles Travelled (avg. crash rate)
NC 86	2	3	$409,533	$0.35	1.59
NC 57	3	3	$98,631	$0.11	1.50
St Marys Rd	2	3	$6,556	$0.01	1.87
New Hope Church Rd	2	2	$0	$0	2.86
Arthur Minnis Rd	2	2	$0	$0	2.30
Old NC 86	2	2	$506,174	$0.60	1.62
Hillsborough Rd	2	3	$14,048	$0.01	0.89

Table 3.4 demonstrates the summary statistics of the variables of interest. Each 100-m segment, experiences about 1.5 crashes on average for the 8-year study period. As a result of existing variabilities in the functional class of roads, the variables take wider ranges and have relatively high standard deviations. The variables in Table 3.4 would be used to develop a safety performance function that will predict crash frequency based on AADT and Road Protection Score.

Table 3.4 Summary statistics of crash prediction model variables

Variable	Mean	Std. Dev.	Maximum	Minimum
Crash Frequency (8 years)	1.459	2.198	28	0
Fatal and Serious injury Crash Frequency (8 years)	0.037	0.226	3	0
AADT	4,698	1,989	10,133	1,867
Road Protection Score	14.1	11.9	67.7	1.7

Results and Discussion

Two analyses were conducted for the validation of Star Rating crash rate relationship. In the first analysis, the averages of crash rate and crash cost per Vehicle Miles Travelled for each road are summarized based on the Star Ratings. Figure 3.2.a and Figure 3.2.b show the resulting bar charts for crash rate, where the crash rate decreases with an increase in both smoothed and raw Star Rating for the sample dataset. It can be observed that the average crash rate increases 23% from 3-star to 2-star in the smoothed dataset. The decrease in average crash rate is relatively higher in the raw dataset, where a 54% increase is observed by moving from 3-star to 2-star.

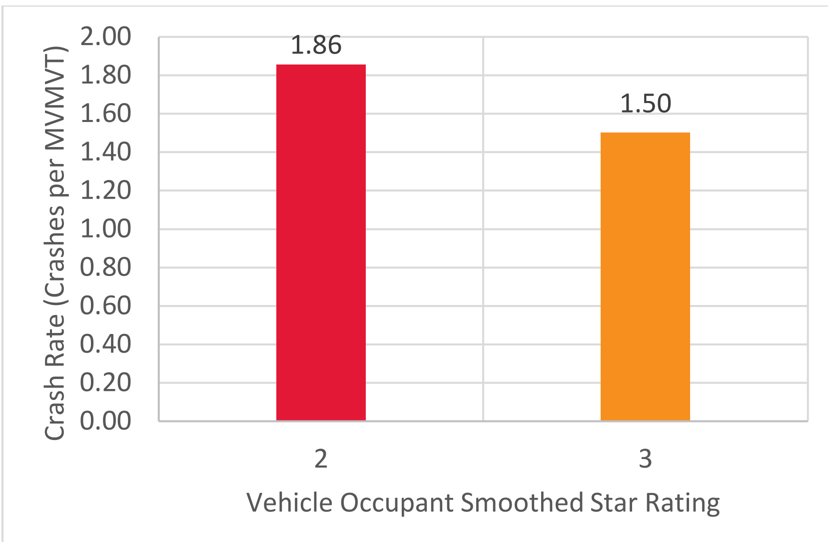

Figure 3.2.a Crash rate by vehicle occupant smoothed Star Rating

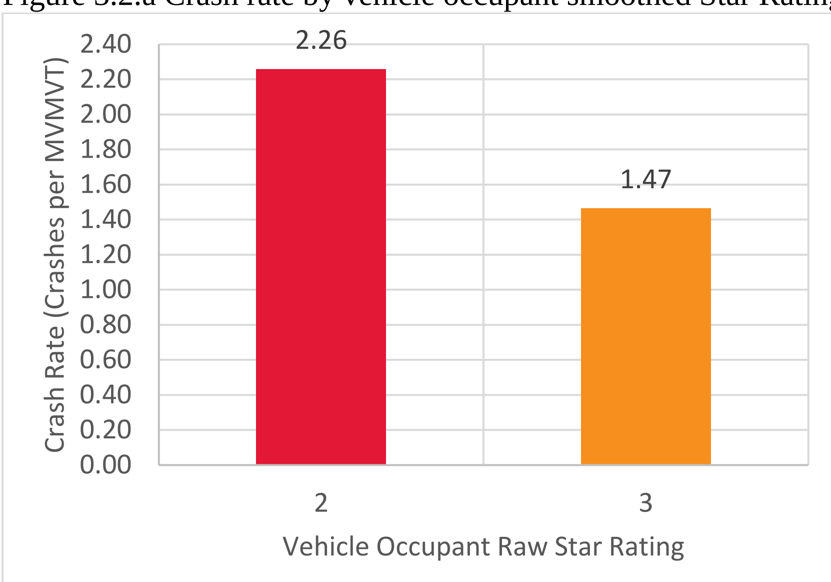

Figure 3.2.b Crash rate by vehicle occupant raw Star Rating

Similar charts were created for crash cost per Vehicle Miles Travelled. The crash cost general trends resembled to the crash rate charts (See Figure 3.3.a and Figure 3.3.b). The crash cost per VMT shows a 44% increase in the smoothed Star Rating chart and 74% increase in the raw Star Rating chart between 3-star and 2-star roads. Given that the crash cost method uses a weighted parameter that takes the severity of crashes into consideration, more significant changes are observed between the 2-star and 3-star roads in this method.

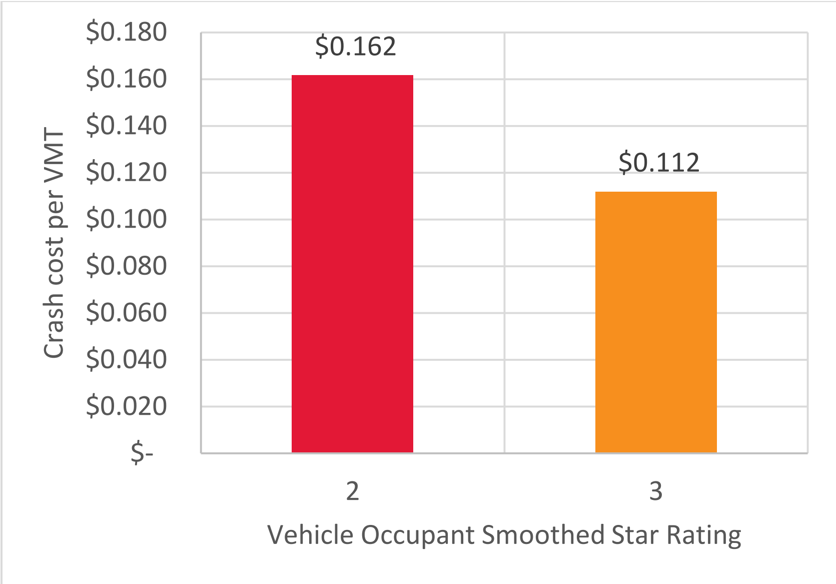

Figure 3.3.a Crash cost per VMT by vehicle occupant smoothed Star Rating

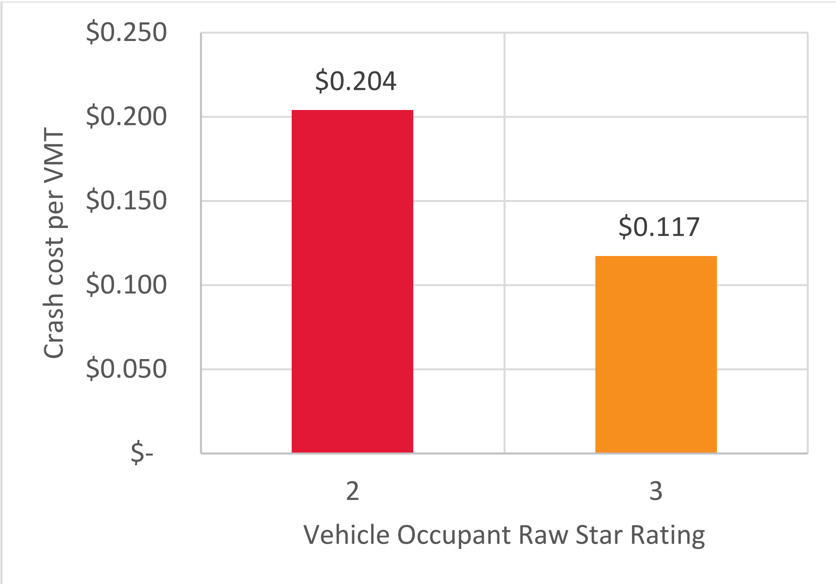

Figure 3.3.b Crash cost per VMT by vehicle occupant raw Star Rating

The second part of the analysis involved developing a statistical model to investigate the relationship between crash frequency, Road Protection Score, and AADT. While the extent of past research has generally been more focused on linear statistical relationships, this research uses a negative binomial model to account for the over-dispersion in crash frequencies. Table 3.5 shows the model results and provides parameter estimates, standard errors, and p-values for each variable, along with the goodness-of-fit statistics.

Table 3.5 Negative binomial model results for crash frequency

Parameter	**Coefficient (Std. Error)**	**p-value**
Intercept	-7.200 (1.040)	<0.001
LN(AADT)	0.823 (0.122)	<0.001
Road Protection Score (Smoothed)	0.044 (0.016)	0.007
Overdispersion parameter	0.885 (0.100)	<0.001
Goodness of Fit		
Sample Size	647	
Log-likelihood	-1040.367	
Akaike's Information Criterion (AIC)	2088.735	

The safety performance function found crashes to increase consistently with traffic volumes. Examining the specific variable of interest, crashes were found to increase consistently with increases in Road Protection Scores (equivalent to decreases in Star Ratings). Specifically, 2-star two lane rural roads would have 47% more crashes on average compared to 3-star roads. These results are generally consistent with the first part of the analysis, showing Star Rating as a valid measure for assessing the risk.

Conclusion

It is essential to answer the question “Is Star Rating a good representative of crash likelihood and crash protection?”. To answer this question, this study incorporates two approaches in order to compare the average crash rate/crash cost and Star Rating/Road Protection Score using a sample of 40 miles of two-lane rural roads in North Carolina using two approaches. Both approaches find the Star Rating/RPS a solid indicator of crash rate and crash cost. In the first approach, the bar charts of Star Rating by crash rate/crash cost revealed a decline in the crash rate/cost when moving from 2-star to 3-star. Specifically, the following conclusions were made:

- Crash rate increases 23% on average moving from a 3-star to 2-star in Smoothed Star Rating chart
- Crash rate increases 54% on average moving from a 3-star to 2-star in Raw Star Rating chart
- Crash cost increases 44% on average moving from a 3-star to 2-star in Smoothed Star Rating chart
- Crash cost increases 74% on average moving from a 3-star to 2-star in Raw Star Rating chart

In general, more increases were observed for the crash cost charts compared to crash rate charts. Moreover, the raw Star Rating graphs projected sharper increases in both crash rate and crash cost charts.

The second approach used a negative binomial model to allow a deeper insight into the statistical relationship between the crash frequency, exposure and Road Protection Score. The analysis results showed a statistically significant correlation between the crash frequency, traffic volume and Road Protection Score with a 99 percent level of confidence. The developed safety performance function showed that moving from a 3-star road to a 2-star road would result in 47% more crashes.

The results of this study are based on total crashes and the aggregated data for the entire sample size where all the roads are high speed two-lane rural roads with Star Rating of 2 or 3. It would be interesting to also examine the existence of such relationship in different types of roads with a wider range of Star Ratings. The small sample size did not allow for assessing the relationship between Star Ratings and fatal and serious injury crashes with targeted crash types (i.e. run-off road, head-on, intersection-related crashes). Moving forward, a larger sample size would confirm the results of this research and extend the findings to wider ranges of Star Ratings and different facility types.

CHAPTER 4. GENERAL CONCLUSION

Conclusions and Recommendations

The Road Assessment Program is a systemic management tool that has been widely recognized and utilized by countries around the world to conduct extensive road safety assessments and set safety target policies. The aim of this research was to investigate data collection protocols and safety assessment methods used in the RAP tool. The research study in Chapter 2 examined the effect of utilizing objective and high quality data on assessing the level of safety of roads using the United States Road Assessment Program. A comparison was made between the manually collected variables versus the automated collected variables and several cases were found with inaccurate roadway attribute values in the manually collected dataset. It was shown that the data with higher quality tend to provide a more accurate evaluation of road's high risk areas along with a more comprehensive countermeasure plan.

Chapter 3 moves on to discuss another aspect of the U.S. Road Assessment Program and examines the assumption behind the Star Rating methodology. The usRAP methodology is based upon the assumption that the Star Rating/Road Protection Score is a valid risk measure for crash experience. A negative binomial model was used to develop a safety performance function for two lane rural roads. The research study findings showed that the increases in Road Protection Score would result in the significant increases in crash frequency. Additionally, comparison of general trends of data associated with lower crash rates with higher Star Ratings. Further investigations are warranted to conduct a more extensive validation analysis using a wider range of road classifications and a more broad range of Star Ratings.

The U.S. Road Assessment Program and other systemic safety assessment management approaches are valuable tools for transportation agencies in order to identify locations with potential for safety improvements. However, in order to provide a more in-depth understanding of challenges associated with these tools, a definite need for more research studies in this area exists.

REFERENCES

1. National Highway Traffic Safety Administration. Fatality Analysis Reporting System. https://www-fars.nhtsa.dot.gov/%0A. Accessed Jan. 1, 2017.
2. Kochanek, K. D., S. L. Murphy, J. Xu, and B. Tejada-Vera. *National Vital Statistics Reports Deaths : Final Data for 2014*. 2016.
3. Federal Highway Administration. *A Systemic Approach to Safety - Using Risk to Drive Action*. 2015.
4. Preston, H., R. Storm, J. Bennett, and B. Wemple. *Systemic Safety Project Selection Tool*. Washington, DC, 2013.
5. International Road Assessment Programme. *iRAP Star Rating and Investment Plan Quality Assurance Guide*. 2014.
6. Federal Highway Administration. *SafetyAnalyst: Software Tools for Safety Management of Specific Highway Sites*. 2010.
7. Federal Highway Administration. *Interactive Highway Safety Design Model*. 2006.
8. Minnesota Department of Transportation. *Otter Tail County Safety Plan: Moving Towards ZERO Deaths.* St. Paul, MN, 2011.
9. Kissinger, J. P. usRAP Moves Beyond Pilots. *ITE Journal*, Vol. 81, No. 11, 2011, p. 40.
10. International Road Assessment Programme. *iRAP Methodology Fact Sheet#6, Star Rating Score Equations*. 2014.
11. Smadi, O., N. Hawkins, Z. Hans, B. Aldemir Bektas, S. Knickerbocker, I. Nlenanya, R. Souleyrette, and S. Hallmark. *SHRP 2 S04A Naturalistic Driving Study: Development of the Roadway Information Database*. 2014.
12. McMahon, K., and S. Dahdah. *The True Cost of Road Crashes*. 2008.
13. World Health Organization. *Global Status Report on Road Safety*. World Health Organization, 2015.
14. International Road Assessment Programme. *iRAP Methodology Fact Sheet #1, Overview*. 2014.
15. International Road Assessment Programme. *iRAP Methodology Fact Sheet#7, Star Rating Bands*. 2013.

16. International Road Assessment Programme. *iRAP Methodology Fact Sheet # 2 Development History*. 2013.

17. International Road Assessment Programme. *Crash Rate – Star Rating Comparisons*. 2011.

18. European Road Assessment Programme. EuroRAP in Sweden. http://www.eurorap.org/partner-countries/sweden/. Accessed Jun. 30, 2017.

19. European Road Assessment Programme. EuroRAP in Iceland. http://www.eurorap.org/partner-countries/iceland/. Accessed Jun. 30, 2017.

20. European Road Assessment Programme. EuroRAP in Great Britain. http://www.eurorap.org/partner-countries/great-britain/. Accessed Jun. 30, 2017.

21. Australian Road Assessment Programme. *Comparing Risk Maps and Star Ratings*. 2008.

22. Harwood, D., K. Bauer, D. Gilmore, R. Souleyrette, and Z. Hans. Validation of US Road Assessment Program Star Rating Protocol: Application to Safety Management of US Roads. *Transportation Research Record: Journal of the Transportation Research Board*, No. 2147, 2010, pp. 33–41.

23. Harwood, D., D. Gilmore, K. Bauer, R. Souleyrette, and Z. N. Hans. US Road Assessment Program (usRAP) Pilot Program—Phase III. 2010.

24. Ambros, J., A. Borsos, and T. Sipos. Exploring an Alternative Approach to iRAP Star Rating Validation. 2016.

25. Washington, S., M. Karlaftis, and F. Mannering. *Statistical and Econometric Methods for Transportation Data Analysis*. . CRC press, Florida, 2010.

26. North Carolina Department of Transportation. *2016 Standardized Crash Cost Estimates for North Carolina*. 2016.

ACKOWLEDGEMENTS

I would like to thank my major advisor Dr. Omar Smadi, and my supervisors Zach Hans and Dr. Basak Bektas for their valuable insights and recommendations. I also would like to extend my thanks to Iowa State University, Institute for Transportation for the support in conducting this research.