MID-AMERICA TRANSPORTATION CENTER

Report # MATC-UNL: 004-72

Final Report

WBS: 27-1121-0005-004-72

Developing a Safety Management System including Hazardous Materials for Highway-Rail Grade Crossings in Region VII

Li Zhao, PhD, PE
Postdoctoral Research Associate
Nebraska Transportation Center
Department of Civil and Environmental Engineering
University of Nebraska-Lincoln

Muhammad Umer Farooq
Graduate Research Assistant
Department of Civil and Environmental Engineering
University of Nebraska-Lincoln

Aemal Khattak, PhD
Professor and Director
Mid-America Transportation Center
Department of Civil and Environmental Engineering
University of Nebraska-Lincoln

2022

A Cooperative Research Project sponsored by U.S. Department of Transportation- Office of the Assistant Secretary for Research and Technology

Developing a Safety Management System including Hazardous Materials for Highway-Rail Grade Crossings in Region VII

Li Zhao, Ph.D., P.E.
Postdoctoral Research Associate
Nebraska Transportation Center
Department of Civil and Environmental Engineering
University of Nebraska-Lincoln

Aemal Khattak, Ph.D.
Professor and Director
Mid-America Transportation Center
Department of Civil and Environmental Engineering
University of Nebraska-Lincoln

Muhammad Umer Farooq
Graduate Research Assistant
Department of Civil and Environmental Engineering
University of Nebraska-Lincoln

A Report on Research Sponsored by Mid-America Transportation Center

University of Nebraska-Lincoln

November 15, 2022

Technical Report Documentation Page

<table>
<tr><td>1. Report No.
25-1121-0005-004-72</td><td>2. Government Accession No.</td><td colspan="2">3. Recipient's Catalog No.</td></tr>
<tr><td colspan="2" rowspan="2">4. Title and Subtitle
Developing a Safety Management System including Hazardous Materials for Highway-Rail Grade Crossings in Region VII</td><td colspan="2">5. Report Date
November 15, 2022</td></tr>
<tr><td colspan="2">6. Performing Organization Code</td></tr>
<tr><td colspan="2">7. Author(s)
Li Zhao, Aemal Khattak, Muhammad Umer Farooq</td><td colspan="2">8. Performing Organization Report No.
25-1121-0005-004-72</td></tr>
<tr><td colspan="2" rowspan="2">9. Performing Organization Name and Address
Mid-America Transportation Center
Prem S. Paul Research Center at Whittier School
2200 Vine St. Lincoln, NE 68583-0851</td><td colspan="2">10. Work Unit No. (TRAIS)</td></tr>
<tr><td colspan="2">11. Contract or Grant No.
69A3551747107</td></tr>
<tr><td colspan="2" rowspan="2">12. Sponsoring Agency Name and Address

Office of the Assistant Secretary for Research and Technology
1200 New Jersey Ave., SE
Washington, D.C. 20590</td><td colspan="2">13. Type of Report and Period Covered
Final Report (5/2020 – 11/2022)</td></tr>
<tr><td colspan="2">14. Sponsoring Agency Code
MATC TRB RiP No. 91994-75</td></tr>
<tr><td colspan="4">15. Supplementary Notes</td></tr>
<tr><td colspan="4">16. Abstract
Highway-rail grade crossings (HRGCs) rank among the leading locations for fatal crashes on the railroad network in the United States, and safety at HRGCs is a top priority for the railroads and the Federal Railroad Administration (FRA). This project studied HRGC safety needs by investigating crash data and HRGC inventory characteristics, and then developed a systematic framework for HRGC safety management in three steps. First, the project started with preparing a comprehensive database that included 1) HRGC crashes with geographic coordinates, 2) HRGC inventory data, 3) highway and train traffic operations data, and 4) HRGC crash-related hazardous materials release data. These data were obtained from the FRA and spanned across the four states in Region VII, i.e., Nebraska, Kansas, Iowa, and Missouri. Second, different accident prediction models for HRGC crash prediction were compared. These models included the accident prediction and severity (APS) model recommended by the FRA and other commonly used crash prediction models such as general linear regression models with fixed or random effects, zero-inflated models, and hurdle models. The APS model was found the best fit for the HRGC data in Region VII area. Therefore, the APS model was calibrated and validated including the index of hazardous materials released and its impact on the surrounding areas resulting from HRGC crashes. A risk score model was developed to rank the HRGCs. Finally, a prototype HRGC Safety Management System (SMS) was developed. The prototype underwent testing utilizing crash data from Nebraska and was initially implemented specifically for the state of Nebraska. The prototype SMS structure was designed so that it could be adopted by state Departments of Transportation (DOTs) in Region VII and across the United States. This project benefits the quality of information provided to decision-makers and enhances the statewide safety management of HRGCs. In particular, the development of this SMS can assist HRGC managers in being proactive to safety and risk situations at HRGCs.</td></tr>
<tr><td colspan="2">17. Key Words
Highway-rail grade crossing (HRGC), Safety management system, crash analysis, HRGC inventory, hazardous materials</td><td colspan="2">18. Distribution Statement</td></tr>
<tr><td>19. Security Classif. (of this report)
Unclassified</td><td>20. Security Classif. (of this page)
Unclassified</td><td>21. No. of Pages
80</td><td>22. Price</td></tr>
</table>

Table of Contents

List of Figures

List of Tables

List of Abbreviations

Accident Prediction and Severity (APS)

American Society of Civil Engineers (ASCE)

BNSF Railway Company (BNSF)

Federal Motor Carrier Safety Administration (FMCSA)

Federal Railroad Administration (FRA)

Highway-Rail Grade Crossing (HRGC)

Mid-America Transportation Center (MATC)

National Cooperative Highway Research Program (NCHRP)

National Safety Council (NSC)

Nebraska Transportation Center (NTC)

Safety Management System (SMS)

State Action Plan (SAP)

Transportation Research Board (TRB)

Transportation Research International Documentation (TRID)

Union Pacific Railroad Company (UP)

United States Department of Transportation (US DOT)

Zero-Inflated Negative Binomial (ZINB)

Acknowledgments

The research team would like to thank Dr. Huong Pham and Mr. Martin Herz for the data coding and visualization. The research team would also like to thank Mr. Roger Figard, Executive Director of the City of Lincoln, and Lancaster County Railroad Transportation Safety District, for his support and information sharing.

Disclaimer

The contents of this report reflect the views of the authors, who are responsible for the facts and the accuracy of the information presented herein. This document is disseminated in the interest of information exchange. The report is funded by a grant from the USDOT's University Transportation Centers Program. However, the US Government assumes no liability for the contents or use thereof.

Abstract

Highway-rail grade crossings (HRGCs) are among the top locations for fatal crashes on the railroad network in the United States, and safety at HRGCs is a top priority for the railroads and the Federal Railroad Administration (FRA). This project studied HRGC safety needs by investigating crash data and HRGC characteristics, and then developed a systematic framework for HRGC safety management in three steps, as described below.

First, the project started with preparing a comprehensive database that included 1) HRGC crashes with geographic coordinates, 2) HRGC inventory data, 3) highway and train traffic operations data, and 4) hazardous materials release data in HRGC crashes. These data were obtained from the FRA and covered the four states in Region VII, i.e., Nebraska, Kansas, Iowa, and Missouri.

Second, different accident prediction models for HRGC crashes were compared. These models included the accident prediction and severity (APS) model recommended by the FRA and other commonly used crash prediction models such as general linear regression models with fixed or random effects, zero-inflated models, and hurdle models. The APS model was found the best fit for the HRGC data in Region VII area. Therefore, the APS model was calibrated and validated including the index of hazardous materials released into and impact on the surrounding areas resulting from HRGC crashes. A risk score model was developed to rank the HRGCs.

Finally, a prototype HRGC Safety Management System (SMS) was developed. The prototype was assessed using Nebraska's crash data and initially implemented for Nebraska. The prototype SMS structure was designed so that it could be adopted by state Departments of Transportation (DOTs) in Region VII and across the United States.

This project benefits the quality of information provided to decision-makers and enhances the statewide safety management of HRGCs. In particular, the development of this SMS can assist HRGC managers in being proactive to safety and risk situations at HRGCs.

Chapter 1 Introduction

Crashes between highway vehicles and trains represent a leading cause of injuries and fatalities in the railroad industry. At highway-rail grade crossings (HRGCs), crashes tend to be more severe due to the greater possibilities of fatal and disabling injury crashes compared to crashes reported at non-HRGC locations. HRGCs continue to be a major concern for the Federal Railroad Administration (FRA), railroad companies, and other transportation agencies despite an ever-increasing focus on improved design and engineering practices.

In 2004, the FRA established the goal of “zero tolerance” for rail-related crashes, injuries, and fatalities (United States Department of Transportation (USDOT), 2006). To achieve this, efforts have been done in the past two decades to enhance HRGC safety including implementing new safety regulations, increasing inspections and audits, improving training for railroad employees, and most importantly, investing in technology and infrastructure upgrades. Key enhancements include the installation of median barriers, flashing warning signals, and automatic gates at HRGCs were deemed essential.

In recent years, each state has been required to develop the HRGC State Action Plan (SAP) in accordance with FRA State HRGC Safety Guidance, which outlines key challenges and visions to address those challenges (FRA, 2020). Despite significant efforts that have been made in reducing the number of HRGC crashes, there is still considerable work ahead to consistently meet the "zero tolerance" goal. In the context of this HRGC safety research, this chapter introduces (1) the background and motivation, (2) the specific research needs and objectives, (3) the work plan and tasks to achieve the established needs and objectives, and (4) the results that the project is expected to deliver.

1.1 Background and Motivation

HRGCs are a major safety and economic priority in the United States. According to the National Safety Council (NSC), motor vehicles related HRGC fatalities totaled 149 in 2021, an increase of 35% compared to 2020 (NSC, 2021). Despite various efforts to improve safety, there are still approximately 2,200 HRGC incidents each year (FRA, 2019), which can result in injuries and fatalities for both drivers and pedestrians, as well as damage to trains, vehicles, and surrounding property.

As a result, traffic authorities (e.g., federal, and local Departments of Transportation (DOTs) have made HRGC safety one of the top priorities. Numerous safety enhancement strategies are implemented, including the installation of warning signs, lights, and gates, educational initiatives aimed at educating drivers and pedestrians about the risks and importance of following traffic laws. Additionally, a lot of safety models are developed to better understand HRGC risk, aiding in the management of HRGC safety.

Accurate prediction models and systematic HRGC hazard ranking are important as they provide evidence for the HRGC project selection process, and these are also essential ingredients of HRGC safety management. A literature review by Sperry et al. (2016) found that 39 out of 50 states in the United States utilize some type of hazard ranking formula or other systematic method for project prioritization. Approximately half of the states followed the USDOT Accident Prediction Model as the primary hazard ranking method. Other models include the NCHRP 50 expected crash frequency model, the New Hampshire Hazard Index, and other State-Specific formulas.

With respect to Nebraska, the NCHRP 50 formula (used before 1999) and the USDOT Accident Prediction Model (used after 1999) have been used for hazard ranking of HRGCs. A recent project, "Nebraska Railway Intersection Safety Study", funded by the Nebraska Department of Transportation (NDOT) and investigated by Khattak et al., updated NDOT's Accident Prediction Model for HRGCs using the latest crash and rail crossing inventory data (Khattak, et al., 2020).

The objective of this project is to improve safety management at HRGCs by leveraging historical data to identify potential critical contributors to HRGC hazards. A prototype Safety Management System (SMS) is developed allowing transportation agencies to conduct advanced risk assessment studies of all HRGCs in the Region VII areas and beyond. The motivation of this project is threefold:

(1) HRGC safety management can take a more proactive approach using this SMS, rather than reacting to safety concerns as they appear.

(2) Traffic authorities may need the assistance of this SMS as a reference tool for the development of a set of policies, strategies, and practices associated with safety.

(3) Current SMSs focus on highways, and the Federal Region VII areas (e.g., Nebraska, Kansas, Missouri, Iowa) do not have an SMS platform on HRGC inventory, HRGC crashes, and rail hazardous material monitoring.

The development of the SMS for HRGCs is a critical step towards improving safety management by transportation agencies. It will enable them to conduct advanced risk assessment studies of all HRGCs in the Region VII and beyond as well as identify potential critical contributors to HRGC hazards based on historical data. Overall, the implementation of this SMS prototype is expected to contribute to improving safety management and preventing potential HRGC hazards among Region VII states (Nebraska, Iowa, Kansas, and Missouri).

1.2 Research Needs

The safety of HRGCs has long been one of the main concerns of both traffic management officials as well as scholars in the field. While there is national-level guidance on HRGC safety management provided by the FRA, it is unclear how effective it is at the local level (e.g., the US Midwest region). Regardless, it would be helpful for each region to establish its own dedicated SMS that adapts to its unique traffic conditions, travel behavior, and infrastructure characteristics. Specifically, the system developed in this research must address these HRGC safety needs.

1. *There is a need to establish an HRGC comprehensive database and a geographic information-based data search system.*

   According to NDOT, when potential HRGC safety improvement projects are proposed, a Diagnostic Team Review will determine if these improvements are warranted (Title 415, 2020). This requires comprehensive consideration through analyzing train data, vehicle data, crossing data, accident history, and hazardous materials, etc. Unfortunately, there is no such database that integrates all this HRGC related information. Therefore, the establishment of an integrated HRGC database will enhance the review and evaluation process by determining project implementation priority and decision-making support. This will allow for better allocation of limited safety improvement resources.

2. *There is a need to develop a systematic method for identifying and ranking HRGCs using a hazard index that best fits the Nebraska crash data.*

   Federal Region VII states (i.e., Nebraska, Iowa, Kansas, and Missouri) follow the FRA recommended APS model as a safety performance practice for HRGC crash assessments. However, its effectiveness in HRGC crash prediction may be debatable. For example, Khattak et al. (2020) model for Nebraska was based on logistic regression that outperformed in Nebraska HRGC crash predictions compared to predictions obtained from the APS model. Currently, there is no broad study of the APS model's validity in the Region VII states and there are no consistent indicators to identify high-risk HRGCs in the region using the APS model. As such, it would be helpful to establish a systematic method that identifies and ranks HRGCs for safety or operational improvements in the region. Specifically, a comprehensive hazard index can be used to rank HRGCs, which can prioritize safety improvements to those HRGCs. A benefit to this approach is that the research complements the recent NDOT-sponsored HRGC research by Khattak et al. (2020) and enables the existing NDOT research results to be disseminated to the regional audience.

3. *There is a need to develop an HRGC SMS that integrates crash database, crash prediction algorithms, hazardous materials movement data, and graphical displays.*

Currently, state-wide SMS does not exist in Nebraska. Khattak and Iranitalab (2016) conducted a survey for the need of a statewide highway SMS (including HRGCs) for Nebraska. The work implemented in this research developed a prototype of an SMS for HRGCs that functioned as: 1) a tool to identify risks at HRGCs through crash hotspot analysis, 2) an integrated forecasting model for agencies that accounts for crash risks, 3) a tool for traffic authorities, staff, and the public in exploring HRGC safety information, and 4) a roadmap for other state DOTs who want to adopt a proactive HRGC modeling approach. By integrating various data sources and providing a comprehensive management tool for identifying risks, SMS can assist transportation agencies develop proactive responses to reduce crash risks, enabling them to make informed decisions on project funding prioritization.

1.3 Work Plan and Tasks

In this section, a work plan with six specific tasks is provided that shows the framework for achieving the objectives. These six tasks are literature search, data preparation, database development, hazard ranking model development, SMS prototype development (with crash prediction models), and final report preparation. Note that task implementation is divided according to their contents and are elaborated in the following chapters.

During project implementation, the first step was to integrate HRGC crash data, inventory data, traffic data, hazardous materials data, and other environmental data in a single, comprehensive database. This serves as the main repository of safety data and was created so that key analyses may be conducted. In the second step, established HRGC safety models were compared to determine the best model that incorporates discrete and non-negative crash frequency data, HRGC inventory, and crash likelihood prediction variables. The explored models included:

1) Negative binomial/Poisson regression model with fixed effects
2) Negative binomial/Poisson model with random effects
3) Zero-inflated negative binomial/Poisson model
4) Hurdle negative binomial/Poisson model

These predictive models were compared to the state-of-the-practice and state-of-the-art models that are currently used for crash prediction. Finally, a comprehensive SMS for the HRGC crashes was developed based on the integrated data and accident prediction and risk models.

1.3.1 Task 1. Literature search

A review of published literature was made to note the latest developments in SMS and similar systems research. The purpose of this task is to ensure that no research which might contribute to this study is overlooked or duplicated. Literature sources, such as the transportation specific databases, e.g., Transportation Research Board (TRB), Transportation Research International Documentation (TRID), and multidisciplinary databases, e.g., American Society of Civil Engineers (ASCE) and Web of Science, were utilized for the literature review.

1.3.2 Task 2. Data preparation

Nebraska HRGC related crash data, which are available from the FRA and NDOT, were prepared for integration and analysis. Specifically, data obtained from FRA included the HRGC inventory, geometric layout, traffic control system, etc. Hazardous materials data such as hazardous materials release and hazardous materials-involved crashes were obtained from the FRA safety database for HRGC accidents/incidents (safetydata.fra.dot.gov/OfficeofSafety). Other data such as crash details, vehicle information, driver information, and environmental data were obtained from the crash report database provided by NDOT. In addition, hazardous materials movement data were obtained from the FRA Federal Motor Carrier Safety Administration (FMCSA).

1.3.3 Task 3. Develop an HRGC safety database in a spatial system

The data prepared in Task 2 were compiled in a Geographic Information System (GIS). Consequently, all the data were integrated into a unified data format for modeling. Data integration in this task involved data fusion, data coding, data cleaning and data quality assessment, and missing data imputation. This database serves as the central repository of HRGC safety data and analysis.

1.3.4 Task 4. Develop a systematic method for the HRGC hazard ranking

Most common HRGC safety models (e.g., generalized linear models with fixed or random effects, zero-inflated models, and hurdle models) were used to evaluate and predict the likelihood of the HRGC crashes and crash severity. The models in this task were compared to the current FRA Accident Prediction Model updated in 2020. The goal of the modeling is to develop an equation that identifies key factors (e.g., train volume, speed, pavement, control type, etc.) and estimates the crash risk at HRGCs to obtain an aggregate hazard index in accordance with Nebraska's crash data. Note the developed models should be pertinent to the Region VII states without major restrictions.

1.3.5 Task 5. Develop a prototype of the HRGC safety management system

The SMS incorporates the HRGC crash database module, the crash prediction module, hazardous materials release module, and the hazardous index module. The structure of the prototype is designed so it can be readily expanded to a full version of the HRGC safety management, is flexible enough that it can be coordinated easily with other management systems and can be easily adopted by other states in Region VII and the United States.

1.3.6 Task 6. Final report and presentation

A final report detailing all the research results was developed (i.e., this document). The key section of the report included the process of establishing the HRGC safety management system for the state DOTs in Region VII (i.e., Nebraska, Kansas, Iowa, and Missouri DOTs).

1.4 Expected Results and Products

A geospatial user-friendly tool, i.e., the SMS, was developed in this study to filter crash data at selected HRGCs. Scenarios with different combinations of the factors (e.g., roadway traffic volume, train volume, etc.) can be chosen as a what-if scenario study. The background operation of the SMS requires the support of a comprehensive HRGC crash database and the best accident prediction model, which was developed and compared to the FRA Accident Prediction and Severity (APS) model.

The prototype SMS intends to function as 1) a tool to identify risk HRGC issues such as crash hotspots, 2) an integrated forecasting model for agencies that accounts for both crash likelihood and severity, and 3) a tool to connect traffic authority, staff, and the public in exploring the potential alternatives for safety improvement at HRGCs.

The output of this research was used to develop the SMS as a search tool associated with the HRGC crash database and the accident prediction and risk score models used in this project. The tool provides its interface and analytical capabilities, as well as the crash databases and related information. Note that the SMS framework can be calibrated using data from all four state DOTs in Federal Region VII. Should the DOTs choose to implement it, the research team can work closely with safety engineers from across Federal Region VII to help ensure the prototype meets their needs. In the end, this project ensures that the system is accessible to transportation professionals and ensures that all graphical displays are both useful and easily understandable.

This report is structured into several chapters that cover various aspects of the research project. The second chapter is dedicated to the literature review, which provides a comprehensive overview of the existing research on the topic. In chapter 3, the data preparation process is explained in detail, outlining the steps taken to collect, clean, and organize the data used in the analysis. Chapter 4 is focused on the development of a hazmat hazardous model. Chapter 5 evaluates the FRA's APS model, compares different crash prediction models, and develops a risk score model. Finally, the report concludes with the development of an SMS prototype, which serves as a practical implementation of the research findings.

Chapter 2 Literature Review

2.1 HRGC Crash Modeling Methodology

The USDOT Accident Prediction Formula for HRGCs incorporates comprehensive explanatory factors for HRGC crash prediction and hazard ranking and thus is the most commonly used model by state DOTs (Abioye et al., 2020). Unlike other formulae such as the New Hampshire Index, which are also frequently used in practice, the USDOT formula calculates a collision prediction value using a combination of three independent calculations (Ogden, 2007):

(1) an initial hazard ranking based on a crossing's characteristics.

(2) a collision prediction value using crash history over a period (e.g., 5 years).

(3) a normalizing constant add-on as a periodic adjustment term.

In literature, HRGC crash modeling studies focus on the two primary topics of crash frequency and crash severity. Crash frequency focuses on modeling the probability or frequency of accidents at HRGCs. The aim is to identify the risk factors and factors that contribute to crashes and develop predictive models that can help reduce the frequency of accidents. Crash severity focuses on modeling the consequences of an accident and predicting the severity of injuries, fatalities, and property damage caused by a crash, which can help inform decisions about safety improvements and emergency response planning. These two types of crash studies forecast and determine two sets of factors (not necessarily the same) that influence the probabilities of crash frequency or crash severity levels, as described below.

Due to the nature of crash data (e.g., discrete, non-negative, count), generalized linear models (GLMs) such as the Poisson model and Negative Binomial model (Hu et al., 2012; Chadwick et al., 2014; Oh et al., 2006) are commonly used to predict crash frequency. In recent years, the Zero-inflated Negative Binomial Model (USDOT, 2020, Mathew and Benekohal, 2021) was adopted to address data with zero excess crashes under dispersions.

In addition, the impact of geospatial factors on HRGC crash frequency prediction has also begun to receive attention (Heydari et al., 2018). Machine learning technicques are increasingly being used to improve crash prediction at HRGCs. For example, approaches such as hierarchical tree-based regression, random forests, gradient boosting, and neural networks are gaining popularity for crash frequency prediction at HRGCs (Yan et al., 2010; Lee et al., 2019; Zheng et al., 2019; Zhou et al., 2020; Lu et al., 2020).

Crash severity, on the other hand, has been modeled to investigate its relationship with potential influencing factors at HRGCs using techniques such as the ordered probit model, latent ordered response model, generalize logit model, binary logit model, generalized linear mixed model, mixed logit model, etc. (Hao and Daniel, 2013; Abdel-Aty and Keller, 2005; Eluru et al., 2012; Hu et al., 2010; Zhao et al. 2018; Haleema and Ganb, 2015). It should be noted that these approaches have shed light on the available models to understand expected changes in crash severity, however, they are performed independently from models for crash frequency as mentioned before. Very few studies have been conducted to model both at the same time (Keramati et al., 2020).

Many factors have been utilized in modeling to explain changes in HRGC safety. It has been found that vehicle and train volume, crossing geometry (e.g., sight distance), train speeds, land use of the crossing, surface type, HRGC control type, and driver age and gender, etc. all impact the injuries and fatalities at HRGCs (Abioye et al., 2020; Haleema and Ganb, 2015). In a study of the year-after-year decrease in the number of collisions and fatalities at HRGCs, Mok and Savage (2005) found that, 40% of the reduction was due to reduced drunk driving and improved emergency medical response, with another 20% attributed to the installation of automated gates and/or flashing lights. Addressing these factors and implementing effective safety measures can help improve the overall safety of HRGCs and reduce the risk of accidents and fatalities.

2.2 Safety Management System (SMS) Framework

So far, the SMS on road traffic has received extensive research and attention. For example, since 2012, the International Organization for Standardization (ISO) offered an international scope that specified minimum requirements for road traffic SMS (ISO-39001). The aim of the standard was to reduce death and serious injuries related to road traffic crashes based on a quality management approach, such as the Plan-Do-Check-Act process that allows safety management to ensure safety risks are identified and effectively controlled (Williams, 2020), and best practices for the road traffic safety improvement requirements.

Kravets et al. introduced an automated traffic management system for urban traffic safety and road condition analysis (Kravets et al., 2018). The system was developed based on mathematical models, including neural networks and decision tree algorithms, and site data, including parameters and traffic conditions that resulted in the road accident.

In 2016, Khattak et al. (2016) assessed the needs of the highway SMS in Nebraska and proposed a structural concept of the system, as can be seen in Figure 2.1.

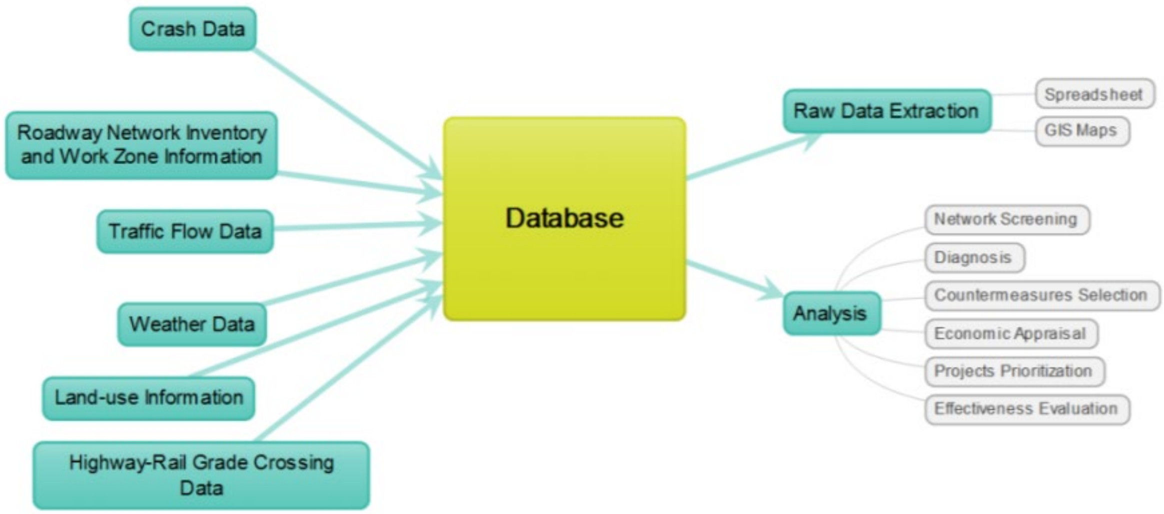

Figure 2.1 Safety management system concept (Khattak and Iranitalab, 2016)

Intelligent transportation systems provide a viable solution to traffic safety management and their role in the improvement of road traffic safety significant. Traffic safety management systems should include a data information system, a data query and filter platform, a safety risk scoring algorithm, and a result display interface. Safety management systems serve as a platform that bridges the communications between traffic managers (authorities), traffic engineers (practitioners), and the public (users).

This project developed a safety management system like the conceptual structure shown in Figure 2.1. The core of the system is driven by hazard-ranking algorithms that prioritize crossings of interest. At present, approximately half of the states used the USDOT Accident Analysis and Prediction model (Sperry et al., 2017), which is also examined in this study. The HRGC SMS is expected to provide transportation decision makers and agencies technical support for project selection and safety improvement priorities for HRGC investments.

Chapter 3 Data Preparation

The database prepared in the development of the HRGC safety management system in this study includes four datasets. They are 1) HRGC inventory data, 2) HRGC crash data, 3) Hazardous Materials (Hazmat) movement approval data, and 4) Hazmat incident data. These datasets were obtained from the FRA website (https://railroads.dot.gov/). Among them, the HRGC crash data was verified with data obtained from NDOT. The inventory data was verified through manual observations by the research team in the Khattak et al. study (2020) of some sample locations.

3.1 HRGC Inventory Data

Table 3.1 provides the railroad companies and counts of HRGCs that are open to railroad traffic in Nebraska. The main railroad companies operating in Nebraska are BNSF and Union Pacific (UP), accounting for more than seventy-five percent of the total HRGCs.

Table 3.1 HRGC Ownership in Nebraska in 2017 (https://railroads.dot.gov/).

No	Railroad Owner	Count
1	BNSF Railway Company (BNSF)	1412
2	Union Pacific Railroad Company (UP)	1117
3	Nebraska Central Railroad Company (NCRC)	352
4	Nebraska, Kansas, Colorado Railnet (NKCR)	295
5	Nebraska Northwestern Railroad, Inc. (NNW)	22
6	Fremont & Elkhorn Valley Railroad (FEVR)	22
7	Manning Rail, Inc. (MAN)	12
8	Sidney & Lowe Railroad, Inc. (SLGG)	8
9	Omaha, Lincoln & Beatrice Railway Company (OLB)	4
10	Chicago, Central & Pacific Railroad Company (CC)	3
11	Nebkota Railway, Incorporated (NRI)	2
Total		**3252**

Figure 3.1 shows the distribution of HRGCs and their affiliation to eleven different railroad owners in Nebraska.

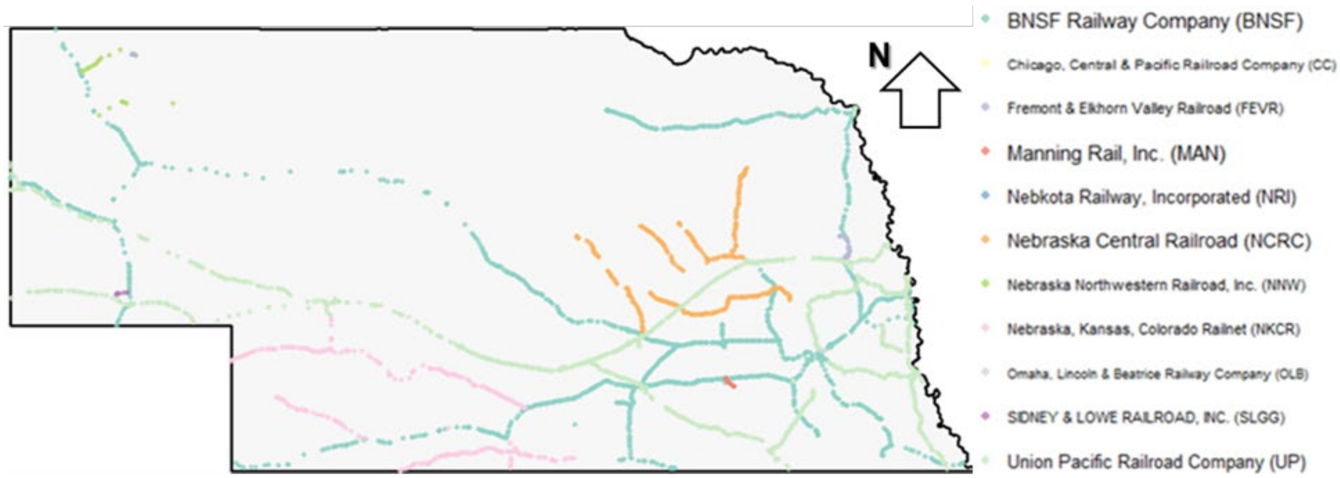

Figure 3.1 HRGC locations of different railroad companies

HRGC traffic control is designed to facilitate safe and efficient operations of both highway and railway traffic in the conflicting area (i.e., the crossing). At all times, motor vehicles on the road should be prepared to take appropriate actions as required by traffic controls when approaching an HRGC. Highway-rail grade crossing traffic control strategies can be divided into (1) no control, (2) passive control, such as yield or stop signs, and (3) active control, such as flashing lights or gates. Table 3.2 lists the counts of HRGCs by different traffic control types in Nebraska.

Table 3.2 HRGC traffic control devices in Nebraska

Control Type	**Number of HRGCs**	**Percentage**
No control	2810	54.1%
Passive- Yield signs	693	13.3%
Passive - Stop signs	815	15.7%
Active - Flashing lights only	156	3.0%
Active - Gates with flashers	722	13.9%
Total	**5196**	**100%**

It should be noted there are other types of passive signs and markings such as crossbuck as well as supplemental active control devices, such as advance warning signs (with Flashers), etc. The Manual on Uniform of Traffic control Devices (MUTCD, 2009) provides the full list of HRGC control devices and their functions.

3.2 HRGC Crash Data

Nebraska, like many other states, tracks and publishes data on HRGC crashes. The NDOT maintains crash data for all reported crashes in the state, including those reported at HRGCs. Figure 3.2 shows trends in HRGC crashes that resulted in injuries and fatalities in Nebraska over the past two decades.

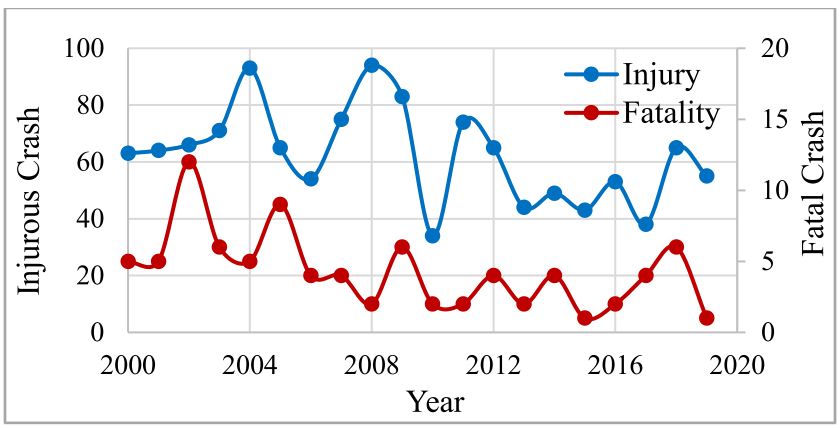

Figure 3.2 HRGC crashes in Nebraska

Figure 3.2 shows a general decreasing trend in the number of HRGC crashes resulting in injuries and fatalities in Nebraska over the past two decades. While the number of crashes fluctuated from year to year, the overall trend shows a gradual decrease from more than 60 crashes in the early 2000s to less than 20 in recent years. This indicates that the safety measures and regulations implemented in Nebraska over the years have been effective in reducing the number of crashes at HRGCs. Measures include upgrading HRGC traffic controls from stop or yield signs to flashing warning signals and automatic gates, installing medians or centerline barriers on the approach lanes, grade separation, and crossing closure. However, there is room for further improvement in HRGC safety.

A closer look at the Nebraska crashes that led to fatalities and injuries, in Figure 3.3, shows the distribution of cases from 2016 to 2020. Taking the circled data for an example, six cases with one fatality (i.e., yellow color in 2018) means there were six cases of crashes that involved one death. Other data in Figure 3.3 are explained in the same way.

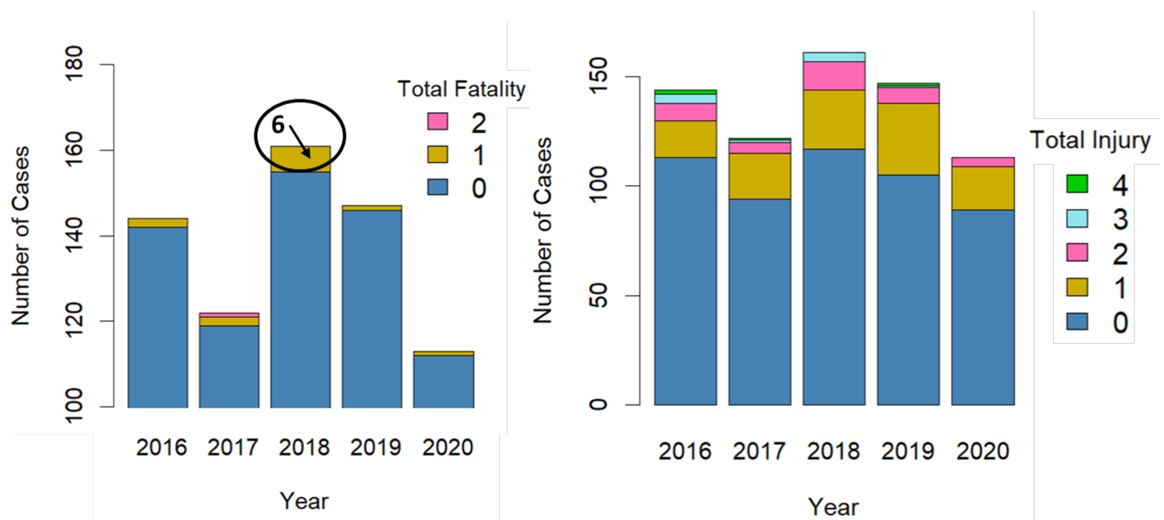

Figure 3.3 Nebraska HRGC fatalities and injuries in the past five years

The dataset shows most crashes involved no injuries or fatalities and likely resulted in property damage only (PDO). The crash severity percentage can also be found in the pie chart in Figure 3.4, where five types of severity were reported: fatal, suspected serious injury, visible injury, possible injury, and PDO.

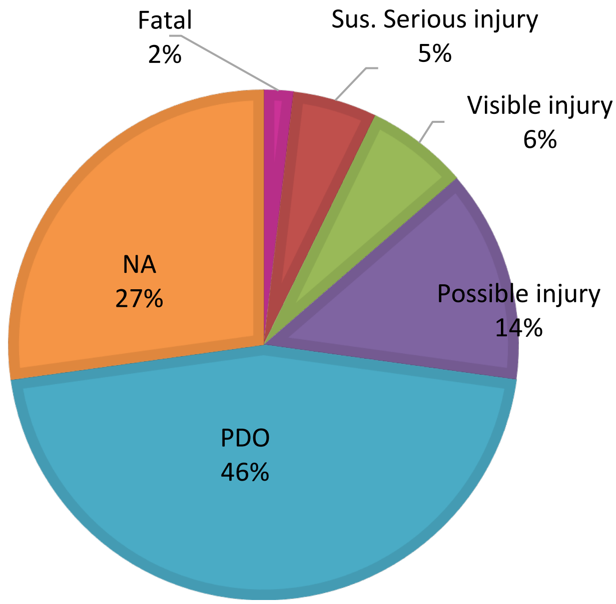

Figure 3.4 crash severity type and the percentage

A lot of factors may have contributed to the crash frequency as well as the crash severity. Table 3.3 summarizes some of the relevant factors that were commonly listed in crash reports filled out by police officers.

Table 3.3 Environmental conditions at the occurrence of an HRGC crash

Factor	No of cases	Percent
Urban	381	55.5%
Rural	306	44.5%
Daylight	495	72.1%
Dark	156	22.7%
Dusk/Dawn	29	4.2%
Unknown	7	1.0%
Clear	512	74.5%
Cloudy	104	15.1%
Rain	37	5.4%
Snow	21	3.1%
Unknown	13	1.9%
Straight	642	93.4%
Curved	42	6.1%
Level	574	83.6%
Slope	110	16.0%
Unknown	3	0.4%

As can be seen in Table 3.3, the HRGC crashes are nearly evenly distributed between rural and urban areas. Other than that, most of the crashes occurred during the daytime (72.1%), in good weather conditions (74.5%), and on straight (93.4%) and level (83.6%) road geometry.

Crashes under other conditions such as night, snow or rain, curved and sloped roads, etc., account for a small number.

To summarize, the HRGC inventory data typically includes information about the location and characteristics of the crossing, such as the number of tracks, type of warning devices, and presence of gates or lights. Crash data, on the other hand, includes information about accidents that occur at these crossings, such as the date, time, location, and severity of the crash, as well as information about the vehicles, drivers, and environment involved. This data can be useful for policymakers, transportation officials, and other stakeholders who are interested in improving railroad safety and implementing targeted interventions to reduce crashes and fatalities at HRGCs. Specifically, these data will be used in accident prediction modeling in Chapter 5 of this report.

Chapter 4 Hazmat Material Transport

In the United States, hazardous materials transported over long distances by rail is considered the safest method (USDOT, 2017). According to USDOT, rail accounts for 16% of ton-miles shipped, which is the second largest transportation mode in tonnage-miles. Table 4.1 shows the ton-miles of hazardous material shipped via the different transportation modes truck, rail, water, air, parcel, and pipeline.

Table 4.1 Hazardous Materials Shipments by Transportation Mode, 2017 (USDOT, 2017)

Transportation mode	**Value $**		**Tons**		**Ton-miles**		**Average miles/shipment**
	billions	%	millions	%	billions	%	
Truck	1091.3	64.9	1814.8	61.1	126.8	33.2	63
Rail	39.0	2.3	90.4	3.0	61.7	16.1	640
Water	137.1	8.2	304.2	10.2	60.9	15.9	72
Air	4.8	0.3	0.3	0.0	0.2	0.1	1,333
Pipeline	339.9	20.2	679.8	22.9	NA*	NA	NA
Parcel, USPS	13.5	0.8	0.3	0.0	0.2	0.1	949
Others	54.6	3.3	78.1	2.6	75.0	24.4	4.0
Total	**1,680**	**100**	**2,968**	**100**	**383**	**100**	**189**

*NA = Not Applicable

In the past two decades from 2000 through 2020, hazardous materials released in railroad incidents showed a decreasing trend nationwide. As shown in Figure 4.1, the number of incidents reported on rails transporting hazardous materials in Nebraska has overall declined over the past 20 years.

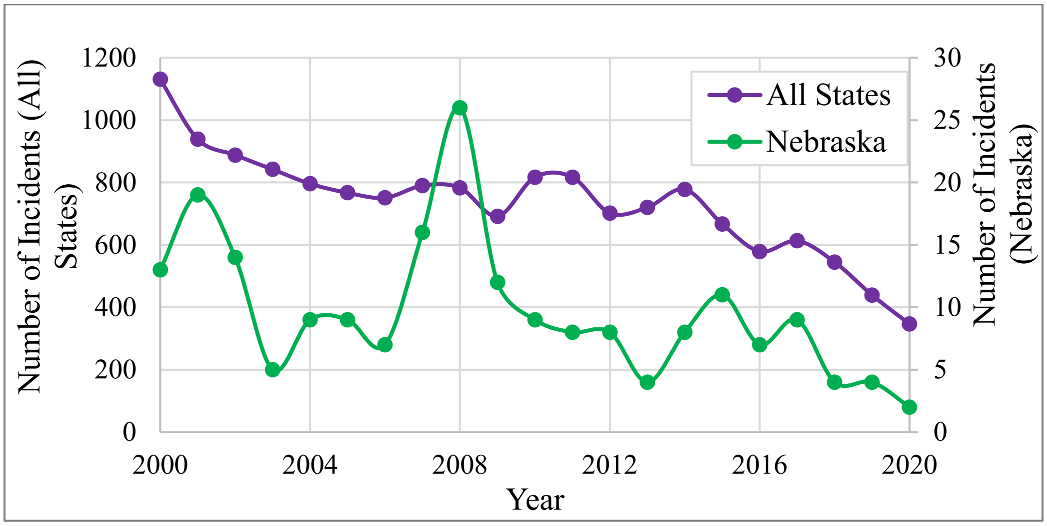

Figure 4.1 Incident Report by Rail Carrying Hazmat

According to the NDOT crash report data, there were 203 crashes involving hazmat release from 2000 to 2020. As shown in Figure 4.2, the hazmat crashes show a decreasing trend over the years, and most occurred during daytime hours (7 am – 7 pm).

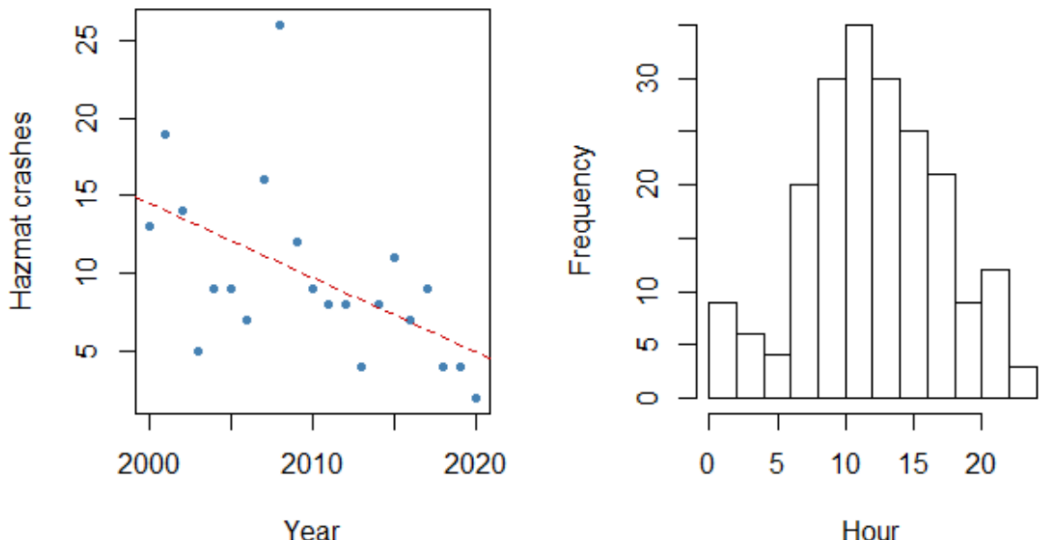

Figure 4.2 Hazmat crashes by year and hour

The Federal Motor Carrier Safety Administration (FMCSA) has grouped hazardous materials into nine classes (USDOT, 2022). They are listed in Table 4.2 with the corresponding number of reported incidents in Nebraska over the past 20 years.

Table 4.2 Classification of hazardous materials

Class	Description (Hazardous materials)	No of Incidents
1	Explosives	0
2	Gases	40
3	Flammable and combustible liquids	103
4	Flammable substances	2
5	Oxidizing substances and organic peroxide	5
6	Poison (Toxic) and poison inhalation hazard	0
7	Radioactive materials	0
8	Corrosives	34
9	Miscellaneous hazardous materials	19
Total (Nebraska, 20 years)		**203**

Based on this classification, the data indicated that more than half of the crashes were related to Class 3: flammable/combustible liquid (e.g., alcohol). The second most hazmat crashes were related to Class 2: gases, including both flammable gases (e.g., propane and spray paints) and non-flammable gases (e.g., liquified helium). Figure 4.3 shows the hazmat classification distribution of the 203 traffic crashes in Nebraska.

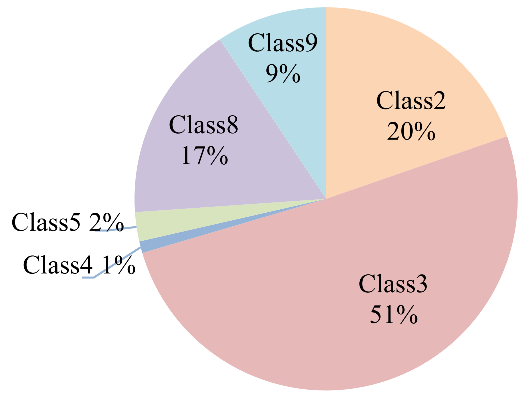

Figure 4.3 Traffic crashes by hazmat classifications

Among the hazmat crashes, over 80% (163 cases) experienced hazmat spillage. There were three types of units that quantify the hazmat release in the crash record data: (1) LGA = liquid gallon, measures liquid hazmat in gallons, (2) SLB = solid pound, measures solid hazmat in pounds, (3) GCF = gas cubic feet, measures gas hazmat in ft^3. All the hazmat release quantities were converted and presented using LGA - liquid gallon, so that they were comparable, i.e., 1 SLB = 0.12 LGA. 1 GCF = 7.48 LGA. The quantities were further converted to the logarithmic scale due to the large value range, i.e., from 0.0078 gallons to 17200 gallons. Figure 4.4 shows the histogram of the hazmat release quantities in all Nebraska crashes after the unit conversion. As an example, the mean of the release quantity in Figure 4.4 is 0.73 log gallons, which when converted back to the actual hazmat quantity release is 2.07 gallons.

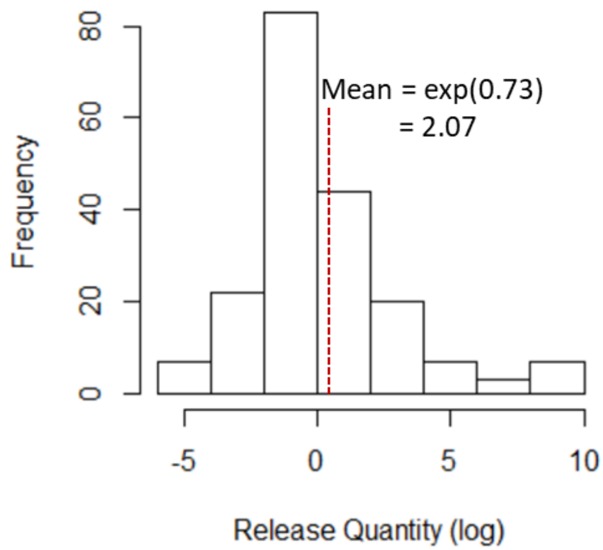

Figure 4.4 Hazmat release quantity in log gallons (exp = exponential function).

The distributions of hazardous materials release quantity (log gallon), the type of incident, and the corresponding damage cost along the railway network in Nebraska are shown in Figure 4.5 and Figure 4.6.

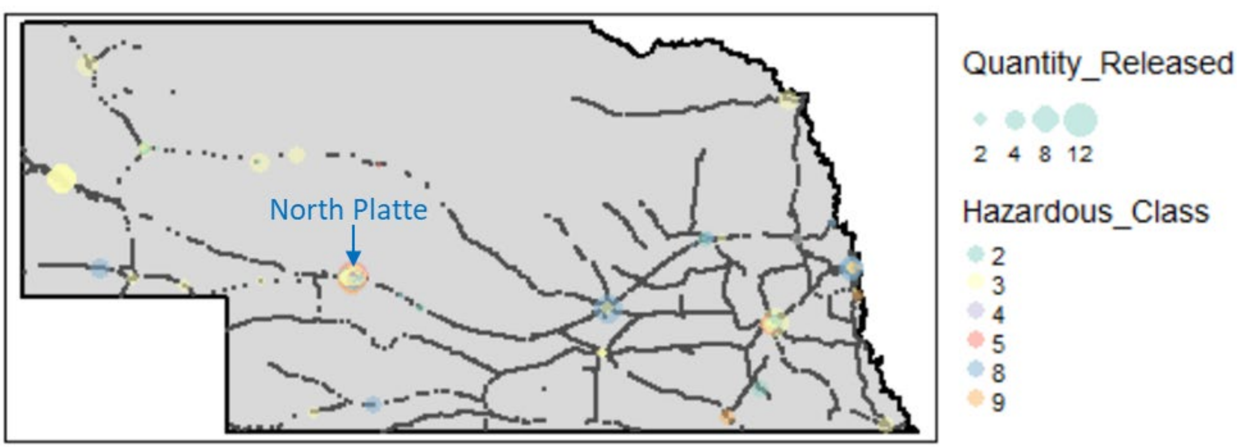

Figure 4.5 Quantity (log gallons) released by hazmat classification

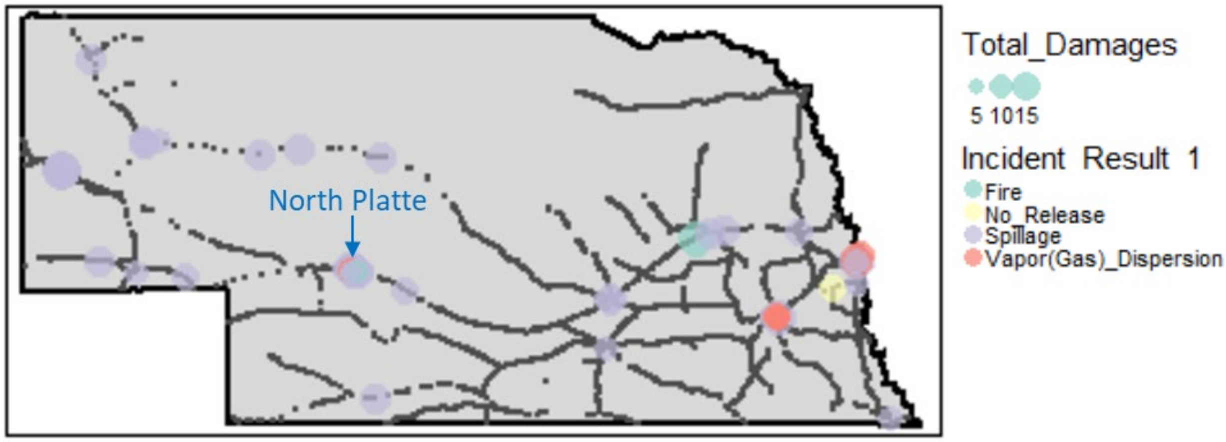

Figure 4.6 Total damages by different outcome

The number of hazmat incidents and the amount of hazmat releases were mostly located in rural areas (west Nebraska) rather than in urban areas (Omaha, Lincoln). As shown in Figure 4.5 and Figure 4.6, North Platte was the site of the largest number of hazmat incidents and quantity releases, accounting for more than half of the total cases. It should be noted that North Platte Nebraska has the largest railroad classification yard in the world—the Union Pacific's Bailey Yard, which is amid key east-west and north-south corridors, making it a critical component of Union Pacific's rail network.

The causes of hazmat incidents are many. Table 4.3 lists the causes of failure and the number of Nebraska incidents they were reported in occurring from 2000 to 2020. The percentage calculated in column 3 and column 4 are slightly different since some cases have multiple causes. Column 3 (percentage – cause) is the percentage out of the total number of causes. Column 4 (percentage – case) is the percentage of cases that cause was reported in.

Table 4.3 Summary of cause of failure lead to hazmat incidents in Nebraska

Cause of failure	Count	Percentage - cause	Percentage - case
Corrosion-Exterior	3	1.2%	1.5%
Deterioration or aging	30	12.0%	14.7%
Human error	6	2.4%	2.9%
Defective component or device	42	16.9%	20.6%
Derailment	10	4.0%	4.9%
Freezing	3	1.2%	1.5%
Inadequate preparation for transportation	24	9.6%	11.8%
Missing component of device	13	5.2%	6.4%
Loose closure, component, or device	76	30.5%	37.3%
Over-pressurized or overfilled	10	4.0%	4.9%
Incorrect sized component or device	1	0.4%	0.5%
Misaligned material, component, or device	4	1.6%	2.0%
Valve open	12	4.8%	5.9%
NA*	15	6.0%	7.4%

*NA = Not Applicable

As can be found in Figure 4.3, the main causes of hazmat incidents were loose closure, component, or device (37.3%), defective component or device (20.6%), deterioration or aging (14.7%), and inadequate preparation for transportation (11.8%). Collectively, these factors account for 84.4% of all cases.

A hazmat release hazard index model was developed to account for all historical hazmat releases in railway networks. The model can be expressed in Equation 1.

$$h_i = \sum_j \frac{q_j}{1 + dist(c_i, m_j)} \tag{1}$$

Where h_i is the hazard index for hazmat release at HRGC number i (i.e., c_i); q_j is the historical hazmat release quantity at location number j (i.e., m_j); dist(c_i, m_j) represents the distance between HRGC location c_i and hazmat release location m_j.

The hazard index model was based on the exposure of HRGCs to historical hazmat release. Specifically, when an HRGC was close to the hazmat release locations and the release quantity was large, the HRGC was significantly affected. As a result, there was a potentially higher risk of the HRGC being impacted in the future. Figure 4.7 illustrates how to calculate this impact through a couple of HRGC locations (in blue) and hazmat release locations (in red).

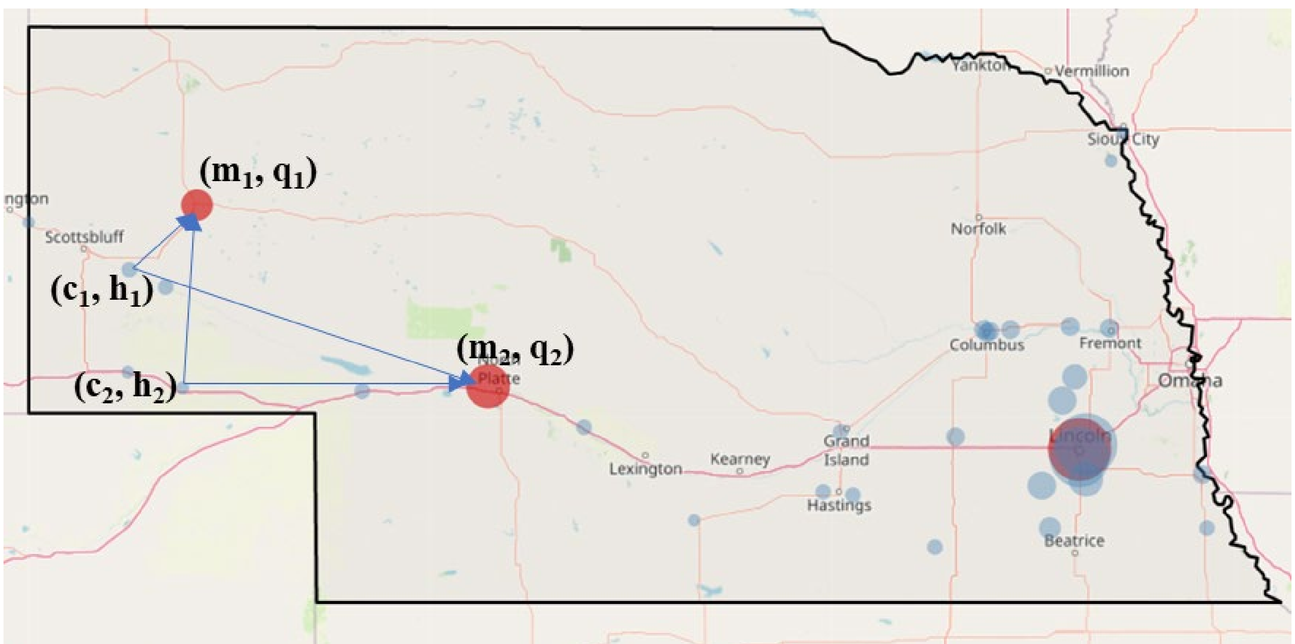

Figure 4.7 Examples for calculating the hazard index of hazmat release at HRGCs

It should be noted that the hazard index relies on historical data and can be updated as more data is collected. This hazard index was normalized, and together with the predicted crashes in the next Chapter, they resulted in a risk score for a given HRGC.

Chapter 5 Accident Prediction and Risk Score Models

The FRA's Office of Safety Analysis maintains a database of railroad accidents, incidents, and casualties across the United States. This database also holds valuable information regarding the HRGC inventory, encompassing crucial physical and dynamic characteristics of HRGCs. Information such as geographical location of the grade crossings, type of gates, type of warning lights, material used in rail tracks, exposure, total daily trains, maximum speed of trains, and functional classification, etc., are recorded for each HRGC in the inventory database (Appendix A). Appendix A provides a complete list of the HRGC inventory data that FRA provided.

The United States Department of Transportation used these variables to develop an Accident Prediction and Severity (APS) model for rail crossings, which has been widely used by federal, state, and local authorities to assess accident risk at highway-rail grade crossings since the late 1980s. However, the FRA commissioned a study, published in 2020, that established a new model using current consensus analysis methods and data trends. The FRA's new APS model addressed several shortcomings of the 1986 APS model. This project aimed to expand upon that research to provide analysts with a more efficient tool.

5.1 FRA's New Model for Accident Prediction and Severity on HRGC

The FRA officially released an update to its accident prediction model (Brod and Gillen, 2020) to help with grade crossing monitoring by allowing for more accurate and reliable ranking of HRGCs, more rational allocation of resources for public safety improvements, and the ability to assess the statistical significance of variances in measured risk.

According to the FRA, the new model helps with grade crossing management by allowing more accurate risk ranking of grade crossings, more rational allocation of resources for public safety improvements at grade crossings, and the ability to assess the statistical significance of variances in measured risk at grade crossings. However, in some instances the inventory data provided by the FRA is outdated. Considering that the proposed model relies on this data, there is a possibility that it may not accurately predict HRGCs crashes.

This research identified the existing data gaps in FRA highway-rail grade crossing inventory data by taking a sample from the FRA database of highway-rail grade crossings from Nebraska and using the new APS model to estimate predicted crashes. Furthermore, this project ranked HRGCs in descending order based on crash predictions, with the most crash vulnerable crossings at the top of the list.

Crash database frequently labels locations as having "no reported crashes," thus zero-inflated models are employed in research by splitting roadway segments into crash-free and crash-prone categories. The framework for this research was designed on zero-inflated negative binomial (ZINB) regression and the Empirical Bayes (EB) method, which takes into account crash history while adjusting for "regression to the mean" bias. The new ZINB regression model can be expressed in two components: count and zero-inflated, as shown in Equations 2 - 4 (Brod and Gillen, 2020).

Count component:

$$N_{CountPredicted} = \exp\left(\beta_0 + \beta_1 * \mathrm{l}Expo + \beta_2 * D_2 + \beta_3 * D_3 + \beta_4 * RurUrb + \beta_5 * XSurfID2 + \beta_6 * \mathrm{l}Aadt + \beta_7 * \mathrm{l}MaxTtSpd\right) \quad (2)$$

Zero-inflated component:

$$\pi = logit(\exp\left(\gamma_0 + \gamma_1 * lTotalTrains\right)) \quad (3)$$

The ZINB model is combined as:

$$N_{Predicted} = N_{CountPredicted} * (1 - \pi) \quad (4)$$

Table 5.1 listed the parameters contained in Equations 2 - 4 above and their corresponding descriptions.

Table 5.1 Variables in the FRA's new ZINB model

Variable code	Variable name	Description
D2	Flashing light indicator	Number of flashing lights > 0, Yes = 1, No = 0
D3	Automated gate indicator	Number of automatic gates > 0, Yes = 1, No = 0
RurUrb	Rural/Urban	Urban = 1; Rural = 0
XSurfID2s	Crossing surface type	Timber = 1; Asphalt = 2; Asphalt & Timber, or Concrete, or Rubber = 3; Concrete & Rubber = 4
lMaxTtSpd	Transformed maximum timetable speed	Note: the four variables, i.e., lMaxTtSpd, lAadt, lTotalTrains, lExpo, were transformed into log format using formula: lx = log(1+αx), where x is the original variable, α is a factor, which was selected so that for the median value of x, ln(1+ αx) = ln(x)
lAadt	Transformed average annual daily traffic	
lTotalTrains	Transformed total number of daily trains	
lExpo	Transformed traffic exposure, Expo = Aadt*TotalTrains	
Crash5y	Counts of traffic crashes at crossings in 5-year	Traffic crashes in 2014 - 2018

The results from the ZINB model output using the FRA’s new approach for accident prediction are shown in Table 5.2.

Table 5.2 ZINB regression model coefficients results (Brod and Gillen, 2020)

Variable	**Estimate**	**Std.Error**	**Z value**	**Pr(>|z|) (p-value)**	**Sig. Code**
ZINB regression count model coefficients (negative binomial with log link)					
(Intercept)	-8.35922	0.32079	-26.059	<2e-16	***
IExpo	0.19023	0.02866	6.638	3.18e-11	***
D2	-0.28478	0.04806	-5.926	3.10e-09	***
D3	-0.85770	0.04089	-20.976	<2e-16	***
RurUrb	0.39346	0.03162	12.444	<2e-16	***
XSurfaceID2s	0.13182	0.01715	7.686	1.52e-14	***
IMaxTtSpd	0.68760	0.68760	22.702	<2e-16	***
IAadt	0.10626	0.10626	3.511	0.000446	***
Log(theta)	-0.25934	0.08867	-2.925	0.003447	**
ZINB regression zero-inflation coefficients (binomial with logit link)					
(Intercept)	1.17084	0.19001	6.162	7.19e-10	***
ITotalTr	-1.01008	0.08452	-11.961	<2e-16	***

Significance code: 0’***” 0.001’**’ 0.01’*’ 0.05’.’ 0.1’’ 1

To assess if the new APS model is a better fit, the first step involves estimating a comparable APS model using identical variables and HRGCs crash data over the previous five years, obtained from the Nebraska Department of Transportation. Specifically, the FRA’s new model for accident prediction (Brod and Gillen, 2020) used Nebraska HRGCs from the FRA database in the same format as the APS model. The model results were reproduced as shown in Table 5.3.

Table 5.3 Results using Nebraska HRGC data from the FRA database

Variable	Estimate	Std.Error	Z value	Pr(>|z|) (p-value)	Sig. Code
Count model coefficients (negbin with log link)					
(Intercept)	-5.960	0.434	-13.74	< 2e-16	***
IExpo	0.001	0.021	0.049	0.9609	
D2_Light	-0.576	0.185	3.109	0.0019	**
D3_Gates	-0.316	0.183	-1.724	0.0447	*
RurUrb	1.508	0.122	12.32	< 2e-16	***
XSurfID2s	1.483	0.300	4.95	7.43E-07	***
IAadt	0.299	0.036	8.198	2.44E-16	***
IMaxTtSpd	0.284	0.076	3.752	0.0002	***
Log(theta)	-1.073	0.128	-8.412	< 2e-16	***
Zero-inflation model coefficients (binomial with logit link)					
(Intercept)	-0.640	0.398	-1.607	0.1081	
ITotalTrains	-0.692	0.195	-3.539	0.0004	***

Theta = 0.3419

Number of iterations in BFGS optimization: 29

Log-likelihood: -2349 on 14 Df

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

A comparison of values reported in Table 5.2 and Table 5.3 shows that most estimated coefficients do not change much in terms of (1) the sign of the coefficient values (e.g., plus/positive and minus/negative), and (2) the significance of the coefficient values (i.e., $p < 0.05$) except for two coefficients (i.e., IExpo and Intercept for Zero-inflation model). The minor alterations in coefficients can be attributed to the utilization of the Nebraska-based HRGCs data sample in the model estimation.

5.2 Data Preparation for Model Comparison and Correction

The FRA makes a significant effort to keep crash inventory data up to date, but there are still data gaps. Local coordinators typically offer updated information to HRGC inventory data, which are then submitted using FRA-approved forms. The forms include entries for specialized field names and value assignments, which are entered by authorized personnel into the inventory.

Since reporting updates for the inventory database does not always necessitate verification from other agencies, data for some fields, such as highway and train traffic volumes, may not be updated on a regular basis. This could result in out-of-date or incorrect data, which could affect crash predictions from models based on the database.

The Federal Railroad Administration (FRA)'s HRGC inventory data were used in both the old APS model (Farr, 1987) and the new model (Brod and Gillen, 2020) for HRGC crash prediction. However, outdated, missing, and incorrect data were found that could skew model results. Take HRGCs in Nebraska as an example. The FRA reported 5200 HRGCs that were open (i.e., in operation) in Nebraska at the time of data collection. Among them, there were three main problems with the data.

(1) Outdated: the annual average daily traffic recorded at 78% of HRGCs were before 2010, and 26% were even 50 years ago.

(2) Missing: 40% of HRGCs did not indicate their locations in urban or rural areas, while 33% of the crossing surface types were missing.

(3) Incorrect: quite a few HRGCs were recorded with traffic control strategy (e.g., signs, signals) that mismatch reality.

The FRA inventory data were validated by the research team from a previous NDOT project (Khattak et al., 2020), who visited and checked 539 HRGCs spread across nine Nebraska counties. While details about the validation efforts can be found in the final report of the project (Khattak et al., 2020), Table 5.4 provides a summary of the work conducted. Among the 539 HRGCs visited, 27 HRGCs were found closed and thus were excluded in the research, resulting in 512 HRGCs used in later analysis. In total, the research team of the previous NDOT project corrected 3973 items in the inventory data in the validation efforts of the visited HRGCs.

Table 5.4 HRGC sites visited in Nebraska counties (Khattak et al., 2020)

County	HRGC sites recorded	HRGC sites closed	HRGC sites visited	Total corrected items
Lancaster	203	7	112	1033
Cass	126	2	55	390
Douglas	184	3	67	394
Gage	75	4	41	462
Jefferson	94	2	46	199
Otoe	133	4	79	331
Saline	87	0	38	156
Sarpy	70	2	25	203
Saunders	141	3	76	805
Total	**1113**	**27**	**539**	**3973**

The correction of the HRGC inventory includes missing, incorrect, and outdated data for various information such as traffic control type (e.g., number of crossbucks, bells, signs, signal), traffic lanes, tracks, surface type, posted highway speed, land development (e.g., urban, or rural), roadway storage distance, crossing angle, etc.

According to the FRA's new 2020 model, accident predictions at HRGCs are affected by exposure, warning device type (such as lights or gates), functional classification, surface material used in rail tracks, maximum timetable speed of trains, average annual daily traffic, and number of total of trains passing through a selected HRGC. During the estimation of predicted crashes using FRA inventory data, it was already determined that a few variables would have no effect on physical inventory, and which would keep changing. These variables are exposure, total trains, and average annual daily traffic. In this study, we used the same values for AADT, total trains, and exposure as the FRA did in estimating the new model for accident prediction.

To that end, for all the corrected information conducted by the research team of the previous NDOT project (Khattak et al., 2020), four variables that are directly related to physical inventory of crossings and have a significant effect on crash prediction were focused and validated. They are: RurUrb, XSurfID2s, D2_Light, and D3_Gates. The detailed summary of the data correction for the four variables are listed in Table 5.5.

Table 5.5 Data correction summary for HRGC inventory

	Identified by the FRA (Crossing Closed = No)	Visited by the previous research team (Khattak et al., 2020)	Data Percentage
HRGC sites	5200	539	10.4% (539/5200)
Land use function (RurUrb)	Rural = 2709 Urban = 424 Sum = 3133 (NA* = 2067)	Rural = 378 Urban = 134 Sum = 512	0.1% (2/3133) 0.4% (2/512)
Crossing surface type (XSurfID2s)	Timber (1) = 1804 Asphalt (2) = 137 Asp.&Tim. (3) = 22 Concrete (3) = 1427 Rubber (3) = 42 Con.&Rub. (4) = 61 Sum = 3493 (NA = 1707)	Timber (1) = 179 Asphalt (2) = 23 Asp.&Tim. (3) = 3 Concrete (3) = 228 Rubber (3) = 2 Con.&Rub. (4) = 68 Sum = 503 (NA = 9)	3.3% (116/3493) 23.1% (116/503)
Gate control (D3_Gates)	No (0) = 4318 Yes (1) = 722 Sum = 5040 (NA = 160)	No (0) = 356 Yes (1) = 156 Sum = 512	0.3% (14/5040) 2.7% (14/512)
Flashing light (D2_Light)	No (0) = 4164 Yes (1) = 872 Sum = 5036 (NA = 164)	No (0) = 307 Yes (1) = 205 Sum = 512	0.2% (8/5036) 1.6% (8/512)

*NA = Not Applicable

As seen in Table 5.5, the variable with the largest correction in the data was the crossing surface type (XSurfID2s), which accounts for 3.3% of all Nebraska HRGCs and 23.1% of visited

HRGCs. On the other hand, the variable with the smallest correction in the data is the land use function (RurUrb), which accounts for 0.1% of all Nebraska HRGCs and 0.4% of visited HRGCs. The correction of gates and flashing light traffic control information is about 0.2% and 0.3%, respectively (2.7% and 1.6% of visited HRGCs, respectively).

Many of the crossing pavements have been renovated and reinforced over the past decades. However, little correction was found in land use functions as the attributes of an HRGC located in rural or urban areas remained unchanged over the past decades. The HRGC traffic control type was found upgraded, for example, from signs to signals or vice versa. It should be noted that, although the Khattak et al. research team (2020) recorded quantity changes (e.g., the number of flashing lights from two to four), these types of changes were not shown and summarized in Table 5.5 since they did not change the binary variable of flashing lights from 0 (No, without flashing lights) to 1 (Yes, with flashing lights).

5.3 Model Check

After examining the data used in model variables (e.g., missing data exclusion), 2901 of the total 5200 HRGCs inventory data from the FRA were used in the model check. As illustrated in Figure 5.1, this dataset is labeled as DAT_VN. For the 512 HRGCs visited in the previous project (Khattak et al., 2020), corrections in the data were made and incorporated with the other 2389 unvisited HRGCs into the dataset DAT_PC.

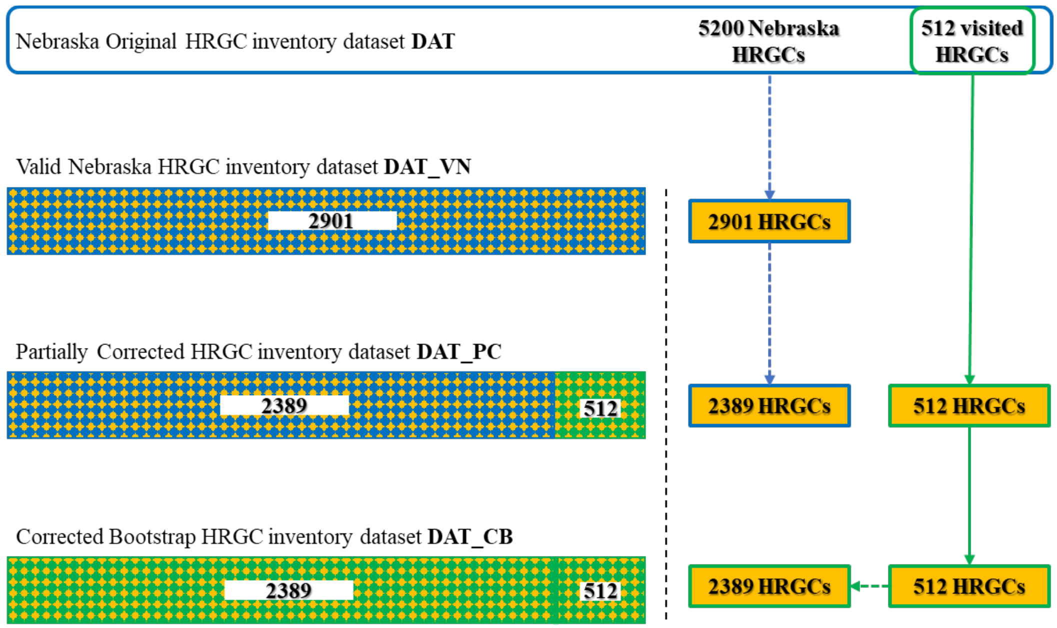

Figure 5.1 Illustration of the dataset composition for DAT_VN, DAT_PC, and DAT_CB

To understand the impact of adding the corrected data to the model results, a bootstrap technique was applied to the 512 visited HRGC data. A sampling of 2380 HRGC data was performed to construct a dataset DAT_CB. Based on the three datasets, the ZINB models were estimated, and the results are shown in Table 5.6.

Table 5.6 ZINB predicted crashes estimated using the three datasets

	DAT_VN (Valid Nebraska dataset)		DAT_PC (Partially Corrected dataset)		DAT_CB (Corrected Bootstrapped dataset)	
	Estimate	**Pr(>|z|)**	**Estimate**	**Pr(>|z|)**	**Estimate**	**Pr(>|z|)**
Count model coefficients (negbin with log link)						
(Intercept)	-7.51	<0.001	-7.64	<0.001	-3.96	<0.001
IExpo	0.09	0.308	0.11	0.2326	-0.14	0.131
D2	-0.18	0.688	-0.37	0.4291	0.20	0.371
D3	-0.76	0.078	-0.65	0.152	-0.02	0.936
RurUrb	0.36	0.247	0.38	0.2317	0.15	0.541
XSurfID2s	0.19	0.041	0.23	0.0149	0.11	0.285
IAadt	0.20	0.081	0.20	0.0878	0.33	0.002
IMaxTtSpd	0.94	0.001	0.95	0.001	0.28	0.230
Log(theta)	0.80	0.673	1.03	0.6901	0.65	0.665
Zero-inflation model coefficients (binomial with logit link)						
(Intercept)	1.21	0.312	1.15	0.363	3.06	<0.001
ITotalTrains	-0.43	0.115	-0.35	0.165	-0.79	<0.001
Model comparisons						
Theta	2.221		2.789		1.912	
Log-likelihood:	-526.2		-522.9		-395.5	
AIC	1074.4		1067.8		813.0	

As can be seen in Table 5.6, DAT_PC dataset performed slightly better than the DAT_NV dataset (AIC: 1074.4 vs 1067.8), indicating that the correction in the inventory data may have positive impact on the model results. Furthermore, it can be found that the DAT_CB dataset yielded the best goodness of fit (AIC = 813.0) on the FRA recommended accident model.

5.4 Model Comparison

As discussed in the Chapter 2 literature review, the APS model is widely accepted by state DOTs, including the four states in Region VII (i.e., NE, IA, KS, MO). In recent years, investigation efforts have been done to reexamine the appropriateness of the APS model developed by the FRA and explore possible optimized HRGC accident prediction models best suited to local conditions. For example, Khattak et al. (2020) suggested a Poisson regression model, instead of the zero-inflated negative binomial model (used in the FRA APS model), was best fit for crash prediction in Nebraska.

To investigate the best possible model fit for HRGC crash prediction among Region VII states, this research developed and compared four types of models assuming both Poisson and Negative Binomial distribution for crash frequency. These models included (1) general linear models, using the NDOT recommended fixed effects as the basic model (Khattak et al., 2020), (2) general linear mixed models, with random effects that allowed for flexibility in the regression relationship for each state, (3) zero-inflated models, which were embedded in the APS model and recommended by the FRA (Brod and Gillen, 2020), and (4) hurdle models, which were also generally used in the setting of excess zeroes. Note the main difference between (3) and (4) was that, typically, (3) were used when there were excess structural and sampling zeroes, while (4) were used when there were only excess sampling zeroes in the data. More details can be found elsewhere (Feng, 2021).

The dataset used in the models included the FRA inventory and crash data of the four states in Region VII, with a total number of 26,757 HRGCs in operation. As outdated and missing data can lead to inaccurate model results, only HRGC inventory data that was updated after 2013 and provided by AADT was used for analysis. In addition, HRGCs with outlier and missing data in all other variables were also manually checked and removed. As a result, data from 3978 HRGCs in Region VII areas were used in the model comparisons in this subsection. Table 5.7 shows the models developed in R (open-sourced programming language), the corresponding function specifications, and the packages used.

Table 5.7 Models specifications in R

Model	Name	Specification	Package
po.fixd	Poisson model/ fixed effects	glm(C5y~gate+light+urban+surface+laadt+train+speed+track,family = "poisson",data = regionvii)	stats
nb.fixd	Negative binomial/ fixed effects	glm.nb(C5y~gate+light+urban+surface+laadt+train+speed+track,data = regionvii)	MASS
po.rand	Poisson model/ random effects	glmer(C5y~gate+light+urban+surface+laadt+train+speed+track+(1|state), family="poisson", data=regionvii)	lme4
nb.rand	Negative binomial/ random effects	glmer.nb(C5y~gate+light+urban+surface+laadt+train+speed+track+(1|state), data = regionvii)	lme4
po.zinf	Zero-inflated Poisson model	zeroinfl(C5y~gate+light+urban+surface+laadt+train+speed+track|train ,dist = "poisson",data = regionvii)	pscl
nb.zinf	Zero-inflated negative binomial	zeroinfl(C5y~gate+light+urban+surface+laadt+train+speed+track|train ,dist = "negbin",data = regionvii)	pscl
po.hurd	Hurdle Poisson model	hurdle(C5y~gate+light+urban+surface+laadt+train+speed+track|train,dist = "poisson", data = regionvii)	pscl
nb.hurd	Hurdle negative binomial model	hurdle(C5y~gate+light+urban+surface+laadt+train+speed+track|train,dist = "negbin", data = regionvii)	pscl

The variables included in the models were the same as in the APS model and were mostly coded the same way as described in subsection 5.2, with a few exceptions: (1) only AADT was rescaled using log format (i.e., laadt), (2) exposure resulting from highway AADT and number of trains (i.e., aadt*train) was not included as an intersection term, (3) the number of main tracks (i.e., track) was added according to the recommendation by the NDOT crash prediction model (Khattak et al., 2020).

Four metrics—sample-size adjusted AIC (second order estimate Akaike Information Criterion), BIC (Bayesian Information Criterion), MSE (Mean Squared Error), and RMSE (Root Mean Squared Error)—were used to evaluate the performance of different models. The AIC and BIC were used to evaluate the fit of the models, while MSE and RMSE were used to evaluate the prediction error of the models. All four metrics were commonly used in model comparison to select the best model that fit the data and provided accurate predictions, while also preferring smaller values. The model performance results are shown in Table 5.8.

Table 5.8 Model performance results in four different metrics

Model	AICc	BIC	MSE	RMSE
po.fixd	1618	1683	67.35	3.05
nb.fixd	1592	1663	56.62	3.02
po.rand	1617	1688	0.41	3.06
nb.rand	1591	1668	0.32	3.03
po.zinf	1591	1668	0.12	0.31
nb.zinf	1582	1664	0.12	0.31
po.hurd	1665	1742	25.23	0.45
nb.hurd	1664	1747	3.65	0.37

Results in Table 5.8 show the zero inflated models were a good fit for the data in Federal Region VII, and the negative binomial regression seemed better given it had the lowest values for all four performance metrices.

5.5 HRGC Risk Ranking

Based on crash prediction using the ZINB model (i.e., model nb.zinf in Table 5.7) and the hazard index of hazmat release impact modeled in Chapter 4, the risk score *r* at a given HRGC i was calculated using the following Equation 5.

$$r_i = w\widetilde{N_i} + (1 - w)\widetilde{h_i} \tag{5}$$

Where $\widetilde{N_i}$ is the normalized predicted number of crashes N_i at HRGC i; $\widetilde{h_i}$ is the normalized hazard index for hazmat release impact h_i at HRGC i (see Chapter 4); and w is a factor that moderates the weights between the crashes and the hazmat release impact. Each state DOT may determine the weight w. A value of 0.5 is recommended if no further information is given.

Using Equation 5, all HRGCs were consistently ranked by crash risk in descending order, with the most crash vulnerable crossings at the top of the list. Table 5.9 shows the top 10 HRGCs in Nebraska identified as the riskiest, which are located on the map shown in Figure 5.2.

Table 5.9 Top 10 riskiest HRGCs in Nebraska

CrossingID	UpdatedDate	Train	Gate	Light	AADT	Urban	Track	Crs.Prd	Haz.Ind	Risk.Score	Risk.Rank
089157E	2/21/2022	14	Y	Y	8100	Y	1	0.344	34.374	0.607	1
817970R	2/21/2022	59	N	N	5600	Y	1	0.223	37.929	0.596	2
816899F	2/21/2022	45	N	N	12700	Y	2	1.144	3.363	0.522	3
089156X	2/21/2022	14	Y	Y	3600	Y	1	0.285	29.391	0.505	4
064364C	2/21/2022	12	N	N	27100	Y	1	0.654	9.954	0.398	5
064128X	11/30/2021	23	Y	Y	10900	Y	2	0.602	8.091	0.349	6
089154J	2/21/2022	14	Y	Y	5800	Y	1	0.133	22.779	0.347	7
817965U	2/21/2022	37	Y	Y	400	N	3	0.220	19.884	0.345	8
064129E	11/30/2021	22	Y	Y	11000	Y	2	0.595	7.798	0.342	9
089153C	2/21/2022	14	Y	Y	10100	Y	1	0.151	20.446	0.323	10

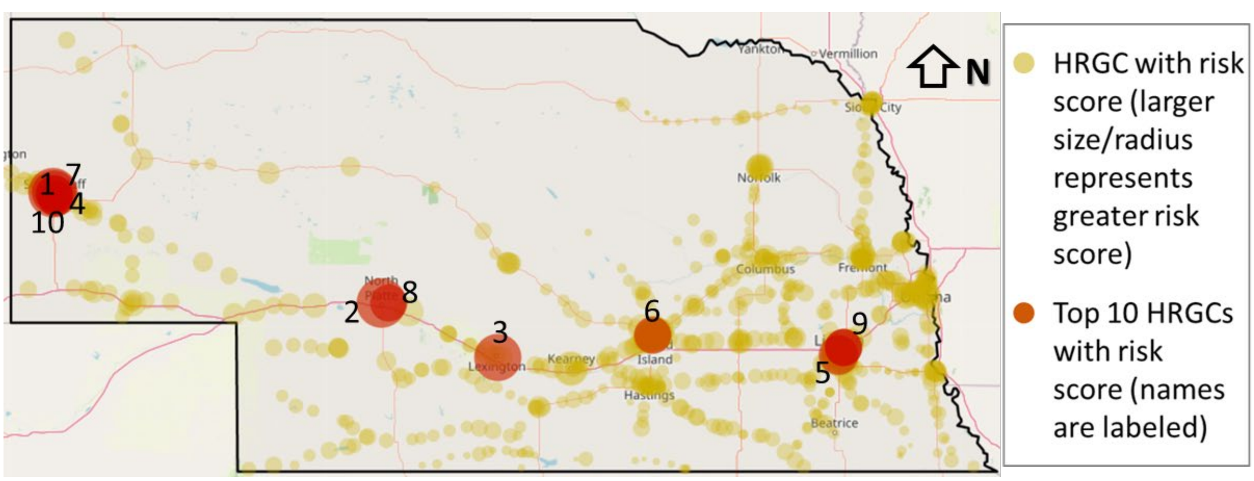

Figure 5.2 Locations of the top 10 riskiest HRGCs in Nebraska

In summary, this subsection discusses the method for ranking HRGCs based on crash risk and the potential impact of hazardous material releases. It used a formula in Equation 5 to calculate risk scores, combining crash predictions and hazard indices, and provides a list of the top 10 riskiest HRGCs in Nebraska, indicating their potential for crashes and hazardous material impacts. All this information will be visualized in the next chapter.

Chapter 6 Development of the Safety Management Systems

As one of the main objectives of this project, a safety management system (SMS) for the HRGC, driven by data and models, was developed. The role of SMS acts as a toolbox for traffic practitioners and stakeholders. This chapter introduces the framework, interface, and functions of the developed HRGC SMS toolbox. The benefits of the SMS are that it can be used by decision makers in various cases such as infrastructure planning, safety investment, forecast, management, control and monitoring, risk exposure and hotspot identification, and hazmat evacuation and responses, in different scales of the transportation network (e.g., crossing level, corridor, or area).

Multiple data-driven platforms exist for road safety management. Examples include the GIS Safety Analysis Tools and Safety Analyst from FHWA, Roadsoft from Michigan Technological University, and Regional Transportation Safety Information Management System (RTSIM) from Arizona Department of Transportation. However, these safety management systems mainly focus on highway safety. Due to the unique circumstances of railway, e.g., HRGC inventory collection and crashes between a vehicle and a train, a separate SMS focusing on HRGCs is necessary to cover the Midwest region.

6.1 SMS Framework

The framework design for the SMS included three levels: (1) database, (2) models and query algorithms, and (3) interface for users, as can be seen in Figure 6.1. The bottom level (e.g., Level 1) was the HRGC database. It consisted of all the available data in this project, which were HRGC inventory information, HRGC crashes, and the shipping and incidents information of hazmat carried through the railroad. The inventory data came from two sources, (1) all HRGC inventories in Region VII downloaded from the FRA website and stored separated by State, and (2) the Nebraska HRGC inventories verified by Khattak et al. (2020) during field visits. The crash data also included two sources, i.e., downloaded from the FRA website and obtained from the NDOT. The datasets in Nebraska were used as a case study to demonstrate data verification, model development, and result display, while datasets from the other three states (i.e., Kansas, Iowa, Missouri) also fit the SMS framework without loss of generality. In other words, the local or field data was used as a supplement to the FRA data when incorrect information was identified or when data was missing.

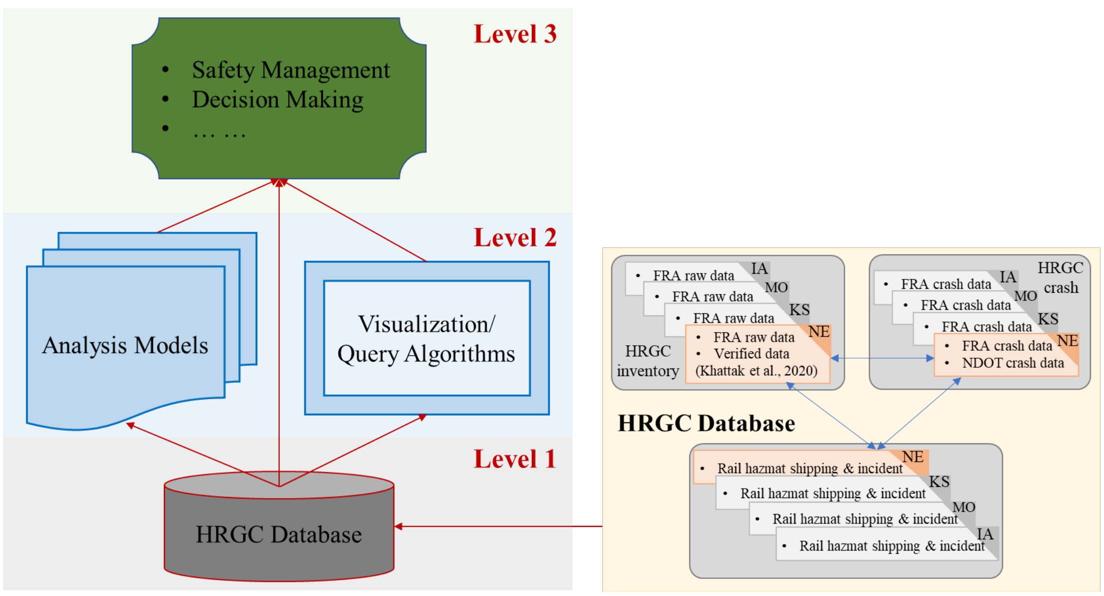

Figure 6.1 An overview of the framework of the SMS

The middle level (e.g., Level 2) was the core part of the framework. It included two parts: the analysis models for accident risk prediction and the data visualization/query algorithms. Note the analysis models referred to accident prediction and risk score models which were introduced in Chapter 5. Several packages from R statistic software were used for data visualization, such as "tmaptools" and "tmap", which were sets of tools for reading and processing spatial data and supplying the workflow to create thematic maps.

An example of the database using Nebraska data can be seen in Figure 6.2, where the HRGC locations (black dots), crash locations (blue dots), and hazmat release locations (red dots) and range (in gallon) are reflected. In the HRGC safety management systems developed in this project, all the information associated with HRGC inventory, crash, and hazmat features are integrated.

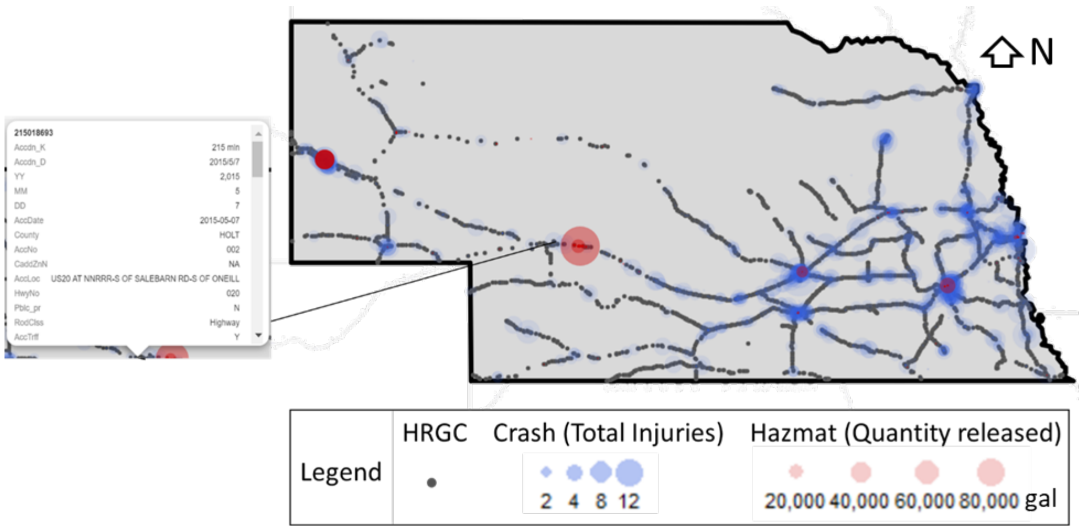

Figure 6.2 HRGC data querying and visualization

The top level (e.g., Level 3) was the encapsulation of models and data for user interface, and the design can be found in the next subsection. The framework developed in this project can be adapted to different traffic safety management systems (e.g., highway-highway intersections, HRGCs) if the corresponding database and crash prediction and severity risk models support it.

For hazmat data management, the quantity of hazmat released in the rail corridor and the probability associated with crashes at the HRGC was integrated. Specifically, when an HRGC was queried in the SMS, both the crash risks (predicted) and the hazmat incident history were associated. In this way, it was given the capability to provide traffic safety departments and first responders with informed decisions about emergency rescues and evacuation plans.

6.2 Shiny App

As part of the technology transfer, this project developed a tool that incorporates the framework of the HRGC SMS provided in the previous subsection. The tool is an application that can facilitate data visualization, filtering, model estimation, result presentation, and report download that are built in R Shiny, a R statistical programming language package that enables building of interactive web applications. It was chosen for the SMS development because of its convenience of hosting a standalone application on a webpage, along with its powerful capabilities for statistical analysis and modeling in R language. Figure 6.3 demonstrates the interface of the application developed by the research team. The codes for this application are attached in Appendix B.

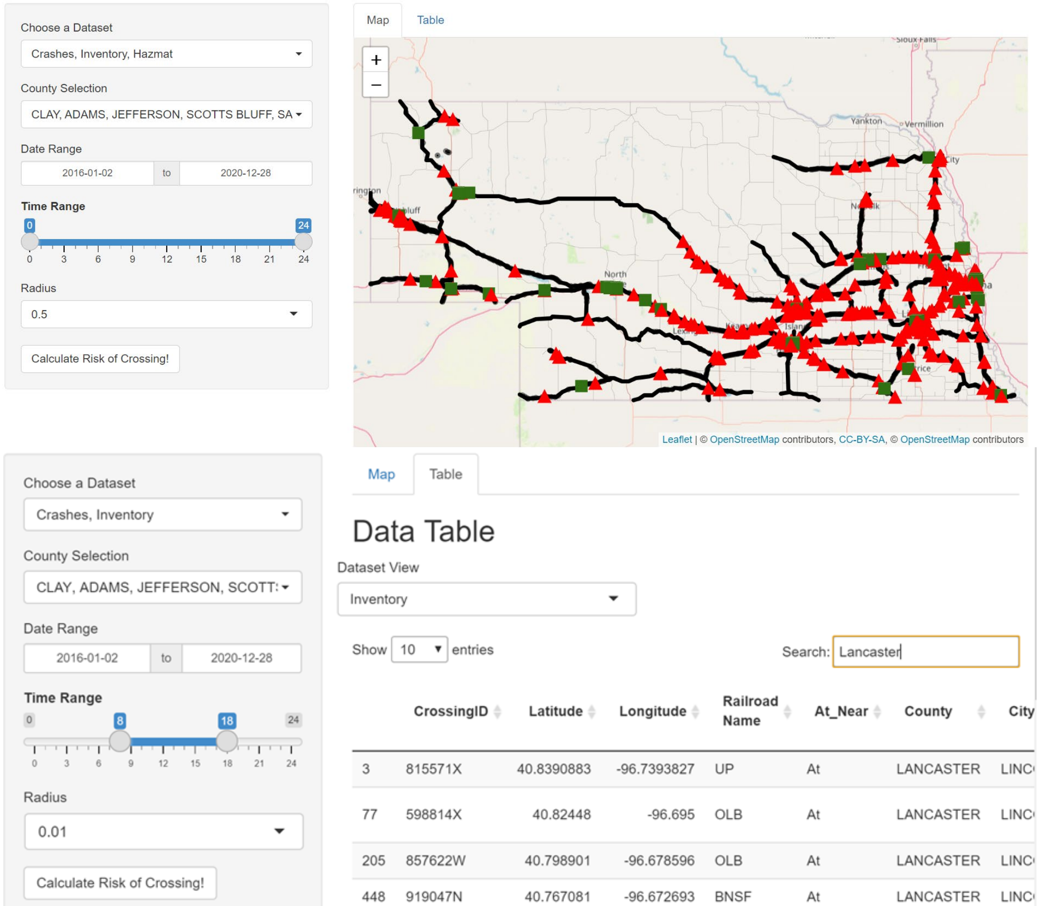

Figure 6.3 Shiny App interface of the HRGC SMS

As a prototype and initial version, the HRGC SMS application can perform the following functions, allowing users to gain a comprehensive view of the historical safety situation and obtain a risk prediction for HRGCs of interest.

1. Select and visualize different datasets (e.g., HRGC crash, inventory, and hazmat) individually as well as in combination. All the data selected was displayed in both the map (the first label) and table (the second label) simultaneously.
2. Filter data of interest in space (i.e., single county or multiple counties in Nebraska) and time (i.e., date range in the past 20 years and hours of day).

3. All detailed information associated with the data was shown on the map when a user’s computer mouse hovered over the data point.
4. The risk scores of the selected HRGCs were calculated according to the accident prediction models in Chapter 5. They were updated on the map and on the table after selecting/pressing the calculation button. All HRGCs were ordered by the risk score (the higher risk is ranked at the top), and users could show the top 10 results, for example the top 10 highest risk HRGCs on the map or in the downloadable table.
5. The radius was set for a range around the crossing to determine the number of crashes that occurred. It counted the number of crashes within a crossing given the radius in miles. For example, when set radius = 0.01 mile (50 ft), all the crashes located within 50-ft of the crossings were counted and associated with the corresponding crossing.
6. All the data, including the risk scores calculated, were updated in the table, and could be downloaded according to the selected criterion.

Chapter 7 Conclusions, Recommendations, and Implementations

This project completed an effort for safety assessment and management at HRGCs in Nebraska. Through investigating the hazmat release data, crash data, and the crossing inventory data, the project evaluated previously established HRGC crash frequency and risk prediction models, which are the FRA's APS model and other four most common models (i.e., regression models with fixed and random effects, zero-inflated model, and hurdle model) under two distributions (i.e., Poisson and Negative Binomial). A risk score model was developed to indicate the hazards at HRGCs that account for both crash and hazmat release impacts. Finally, a systematic framework for HRGC safety management was provided.

To model crash frequency and risk prediction, this project started with an examination of the FRA's prediction model updated by Khattak et al. (2020) to determine whether it was a better fit of the HRGC inventory data and crash data in Region VII areas. Specifically, three different datasets of Nebraska HRGCs from (1) FRA, (2) FRA partially calibrated in the field, and (3) all filed calibration data (involving bootstrapping) were used to fit the FRA's new proposed accident prediction models. In the model fitting, variables remained the same, and five-year crashes on crossings from the Nebraska DOTs crash database were used.

Four different types of commonly used crash prediction models with two types of crash distributions (i.e., Poisson, Negative binomial) were examined. All the Region VII HRGC data were used to check which model was the best fit for the Region VII area. Variables were modified from the recommendation of the FRA APS model and the NDOT updated APS model. Results show that the ZINB model was still the best model format in fitting the HRGC inventory and crash data in Region VII.

Finally, a risk score model was developed to rank the HRGCs according to their potential hazardous impact. The risk score accounted for both the predicted number of crashes and the hazardous index of the hazmat release impact at a given HRGC. This risk score was leveraged in the SMS when a user needed to know information regarding a certain HRGC's safety or the safety of HRGCs in a region, which provided comparative safety statuses among the inquiring HRGCs.

The SMS developed in this research project provides a handy tool for the transportation agencies in HRGC safety data management and risk prediction. It can be used in spotting HRGCs that may need extra attention in safety issues, or in project prioritization when multiple HRGC safety improvement projects are needed but there is a limited budget. Note the prototype SMS structure was designed so that it could be adopted by state DOTs in Region VII and across the United States.

This project improves the quality of the information provided to decision-makers to enhance the statewide safety management of HRGCs. In particular, this SMS can assist HRGC managers in being proactive instead of reactive when it comes to safety and risk situations at HRGCs.

Based on the findings of this research project, the research team concluded that (1) the existing FRA inventory data may be used carefully in the crash modeling, although field data verification may be necessary; (2) Region VII states may consider developing a framework to ensure that key variables in HRGC inventory data are updated; (3) Region VII can continue use of USDOT APS for HRGC crash prediction; and (4) Region VII states may consider developing a grade crossing risk score and ranking model, as introduced in this research project, for its own use with project prioritization in the state.

References

Abdel-Aty, M. and J. Keller. "Exploring the overall and specific crash severity levels at signalized intersections". *Accident Analysis & Prevention*, 37(3): 417–425, 2005

Abioye, O.F., M.A. Dulebenets, J. Pasha, M. Kavoosi, R. Moses, J. Sobanjo, and E. E. Ozguven. "Accident and hazard prediction models for highway-rail grade crossings: a state-of-the-practice review for the USA". *Railway Engineering Science*, 28: 251–274, 2020. https://doi.org/10.1007/s40534-020-00215-w

Brod, D. and D. Gillen. "A new model for highway-rail grade crossing accident prediction and severity". Report No. DOT/FRA/ORD-20/40, US Department of Transportation, October 2020. Accessed date: May 21, 2021. https://railroads.dot.gov/sites/fra.dot.gov/files/2022-09/GX%20APS2-A.pdf

Chadwick, S.G., N. Zhou, and M.R. Saat. "Highway-rail grade crossing safety challenges for shared operations of high-speed passenger and heavy freight rail in the US". *Safety Science*, 68: 128-137, 2014

Cooper, D.L. and D.R. Ragland. Addressing inappropriate driver behavior at rail-highway crossings. Transportation Research Board 87th Annual Meeting Compendium of Papers, Washington, D.C., 2008

Eluru, N., M. Bagheri, L.F. Miranda-Moreno, and F. Fu. "A latent class modeling approach for identifying vehicle driver injury severity factors at highway-railway crossings". *Accident Analysis & Prevention*, 47: 119–127, 2012

Federal Highway Administration (FRA). "Highway-Rail Grade Crossing Incidents", 2019. Accessed date: February 21, 2022. https://railroads.dot.gov/accident-and-incident-reporting/highwayrail-grade-crossing-incidents/highwayrail-grade-crossing

Federal Highway Administration (FRA). "State Highway-Rail Grade Crossing Action Plans". U. S. Department of Transportation, 2020. Accessed date: February 15, 2022. https://railroads.dot.gov/sap

Feng, C.X. A comparison of zero-inflated and hurdle models for modeling zero-inflated count data. *Journal of statistical distributions and applications*, 8(8), 2021

Hao, W. and J. Daniel. "Motor vehicle driver injury severity study under various traffic control at highway-rail grade crossings in the United States". *Journal of Safety Research*, 51: 41-48, 2014

Heydari, S., L. Fu, L. Thakali, and L. Joseph. "Benchmarking regions using a heteroskedastic grouped random parameters model with heterogeneity in mean and variance: Applications to grade crossing safety analysis". *Analytic methods in accident research*, 19: 33-48, 2018

Hu, S.R., C.S. Li, and C.K. Lee. "Investigation of key factors for accident severity at railroad grade crossings by using a logit model". *Safety Science*, 48(2): 186-194, 2010

Hu, S.R., C.S. Li, and C.K. Lee. "Model crash frequency at highway–railroad grade crossings using negative binomial regression". *Journal of the Chinese Institute of Engineers*, *35*(7): 841-852, 2012

International Standard 39001 (ISO-39001). "Road traffic safety (RTS) management systems - Requirements with guidance for use". NS-ISO 39001: 2012 (E). https://pdfcoffee.com/iso390012012e-road-safety-management-systempdf-pdf-free.html. Accessed date: June 28, 2022

Keramati, A., P. Lu, X. Zhou, and D. Tolliver. "A simultaneous safety analysis of crash frequency and severity for highway-rail grade crossings: The competing risks method". *Journal of advanced transportation*, *2020*: 1-13, 2020

Khattak, A. and A. Iranitalab. “Safety management system needs assessment”. Final Report No. SPR-M025. Nebraska Department of Transportation, 2016

Khattak, A., Y. Kang, and H. Liu. “Nebraska rail crossing safety research”. Final Report No. SPR-P1 (19) M091. Nebraska Department of Transportation, 2020

Haleem, K. and A. Gan. “Contributing factors of crash injury severity at public highway-railroad grade crossings in the US”. *Journal of Safety Research*, 53: 23-29, 2015

Kravets, A.G., D.A. Skorobogatchenko, N.A. Salnikova, N.Y. Orudjev, and O.V. Poplavskaya. “The traffic safety management system in urban conditions based on the C4.5 algorithm”. *2018 Moscow Workshop on Electronic and Networking Technologies (MWENT)*: 1-7, IEEE, 2018

Lee, D., J. Warner, and C. Morgan. “Discovering crash severity factors of grade crossing with a machine learning approach”. 2019 Joint Rail Conference. *American Society of Mechanical Engineers Digital Collection*, 2019.

Lu, P., Z. Zheng, Y. Ren, X. Zhou, A. Keramati, D. Tolliver, and Y. Huang. “A gradient boosting crash prediction approach for highway-rail grade crossing crash analysis”. *Journal of advanced transportation*, *2020*.

Mathew, J. and R.F. Benekohal. “Highway-rail grade crossings accident prediction using zero inflated negative binomial and empirical Bayes method”. *Journal of Safety Research*, 79: 211-236, 2021.

Mok, S.C. and I. Savage. “Why has safety improved at rail-highway grade crossings?” *Risk Analysis: An International Journal*, *25*(4): 867-881, 2005.

National Safety Council (NSC). “Every four hours someone is hit by a Train”. Accessed date: January 26, 2022. https://www.nsc.org/road/safety-topics/every-four-hours-someone-is-hit-by-a-train?#:~:text=Roadway%2Drail%20crossing%20fatalities%20involving,fatalities%20totaled%2081%20in%202021.

Ogden, B.D. Railroad-highway grade crossing handbook - Revised Second Edition, Section 3: Assessment of Crossing Safety and Operation. No. FHWA-SA-07-010; NTIS-PB2007106220. Federal Highway Administration, United States Department of Transportation, 2007. Accessed date: November 16, 2021. https://toolkits.ite.org/gradecrossing/sec03.htm#:~:text=The%20U.S.%20DOT%20collision%20prediction,and%20the%20New%20Hampshire%20Index.

Oh, J., S.P. Washington, and N. Doohee. "Accident prediction model for railway-highway interfaces". *Accident Analysis and Prevention*, 38(2): 346–356, 2006.

Sperry, B.R., B. Naik, and J.E. Warner. "Evaluation of grade crossing hazard ranking models". Final Report No. FHWA/OH-2016/10. Ohio Department of Transportation, 2016. Accessed date: February 13, 2022. https://rail.transportation.org/wp-content/uploads/sites/30/2019/09/Formula-Research-Final-Report.pdf

Sperry, B.R., B. Naik, and J.E. Warner. "Current issues in highway–rail grade crossing hazard-ranking and project development". *Transportation Research Record: Journal of the Transportation Research Board*, 2608: 19-26, 2017

Title 415, Nebraska Department of Transportation: Local Assistance Division. "Chapter 6 – Highway-Rail Crossings - Construction, Repair and Maintenance". Accessed date: December 2, 2020. https://www.nebraska.gov/rules-and-regs/regsearch/Rules/Transportation_Dept_of/ Title-415/Chapter-6.pdf

USDOT. "Railroad Safety Statistics 2005 Annual Report". *Federal Railroad Administration, United States Department of Transportation*. December 2006

USDOT. “Hazardous Materials Shipments by Transportation Mode”, 2017. Bureau of Transportation Statistics (BTS). Federal Highway Administration, United States Department of Transportation. Accessed date: March 22, 2022. https://www.bts.gov/content/us-hazardous-materials-shipments-transportation-mode-2007

USDOT. Federal Motor Carrier Safety Administration (FMCSA). United States Department of Transportation. Accessed date: March 24, 2022. https://www.fmcsa.dot.gov/regulations

USDOT. “Highway-Rail Grade Crossings Overview”. Federal Railroad Administration, United States Department of Transportation. Accessed date: February 6, 2022. https://www.fra.dot.gov/ Page/P0156

USDOT. “Traffic Crash Facts 2018 Annual Report”. Prepared By Highway Safety/Accident Records Section Nebraska Department of Transportation, 2018. Accessed date: July 2, 2021. https://dot.nebraska.gov/media/13521/facts2018.pdf

USDOT. “A New Model for Highway-Rail Grade Crossing Accident Prediction and Severity”, 2020. Accessed date: March 11, 2022. https://railroads.dot.gov/sites/fra.dot.gov/files/2020-10/GX%20APS-A.pdf

Williams, A. “PLAN-DO-CHECK-ACT”. *Professional Safety*, 65(2): 18-19, 2020

Yan, X., S. Richards, and X. Su. “Using hierarchical tree-based regression model to predict train–vehicle crashes at passive highway–rail grade crossings”. *Accident Analysis and Prevention*, 42(1): 64–74, 2010

Zhao, S., A. Iranitalab, and A. Khattak. “A clustering approach to injury severity in pedestrian-train crashes at highway-rail grade crossings”. *Journal of Transportation Safety and Security*, 9962: 1–18, 2018

Zheng, Z., P. Lu, and D. Pan. “Predicting highway–rail grade crossing collision risk by neural network systems”. *Journal of Transportation Engineering, Part A: Systems*, *145*(8), 04019033, 2019

Zhou, X., P. Lu, Z. Zheng, D. Tolliver, and A. Keramati. “Accident prediction accuracy assessment for highway-rail grade crossings using random forest algorithm compared with decision tree”. *Reliability Engineering & System Safety*, 200: 106931, 2020

Appendix A HRGC inventory data from (U.S. DOT Form-71)

U. S. DOT CROSSING INVENTORY FORM

DEPARTMENT OF TRANSPORTATION

FEDERAL RAILROAD ADMINISTRATION

OMB No. 2130-0017

Instructions for the initial reporting of the following types of new or previously unreported crossings: For public highway-rail grade crossings, complete the entire inventory Form. For private highway-rail grade crossings, complete the Header, Parts I and II, and the Submission Information section. For public pathway grade crossings (including pedestrian station grade crossings), complete the Header, Parts I and II, and the Submission Information section. For Private pathway grade crossings, complete the Header, Parts I and II, and the Submission Information section. For grade-separated highway-rail or pathway crossings (including pedestrian station crossings), complete the Header, Part I, and the Submission Information section. For changes to existing data, complete the Header, Part I Items 1-3, and the Submission Information section, in addition to the updated data fields. Note: For private crossings only, Part I Item 20 and Part III Item 2.K. are required unless otherwise noted. An asterisk * denotes an optional field.

A. Revision Date *(MM/DD/YYYY)*	**B. Reporting Agency**	**C. Reason for Update** *(Select only one)*	**D. DOT Crossing Inventory Number**
____/____/________	☐ Railroad ☐ Transit ☐ State ☐ Other	☐ Change in Data ☐ New Crossing ☐ Closed ☐ No Train Traffic ☐ Quiet Zone Update ☐ Re-Open ☐ Date Change Only ☐ Change in Primary Operating RR ☐ Admin. Correction	

Part I: Location and Classification Information

1. Primary Operating Railroad	**2. State**	**3. County**
4. City / Municipality ☐ In ☐ Near	**5. Street/Road Name & Block Number** *(Street/Road Name)* | * *(Block Number)*	**6. Highway Type & No.**

7. Do Other Railroads Operate a Separate Track at Crossing? ☐ Yes ☐ No
If Yes, Specify RR

8. Do Other Railroads Operate Over Your Track at Crossing? ☐ Yes ☐ No
If Yes, Specify RR

9. Railroad Division or Region	**10. Railroad Subdivision or District**	**11. Branch or Line Name**	**12. RR Milepost**
☐ None	☐ None	☐ None	*(prefix)* | *(nnnn.nnn)* | *(suffix)*

13. Line Segment *	**14. Nearest RR Timetable Station** *	**15. Parent RR** *(if applicable)*	**16. Crossing Owner** *(if applicable)*
		☐ N/A	☐ N/A

17. Crossing Type	**18. Crossing Purpose**	**19. Crossing Position**	**20. Public Access** *(if Private Crossing)*	**21. Type of Train**	**22. Average Passenger Train Count Per Day**
☐ Public ☐ Private	☐ Highway ☐ Pathway, Ped. ☐ Station, Ped.	☐ At Grade ☐ RR Under ☐ RR Over	☐ Yes ☐ No	☐ Freight ☐ Transit ☐ Intercity Passenger ☐ Shared Use Transit ☐ Commuter ☐ Tourist/Other	☐ Less Than One Per Day ☐ Number Per Day ____

23. Type of Land Use
☐ Open Space ☐ Farm ☐ Residential ☐ Commercial ☐ Industrial ☐ Institutional ☐ Recreational ☐ RR Yard

24. Is there an Adjacent Crossing with a Separate Number?
☐ Yes ☐ No If Yes, Provide Crossing Number ____

25. Quiet Zone *(FRA provided)*
☐ No ☐ 24 Hr ☐ Partial ☐ Chicago Excused Date Established ____

26. HSR Corridor ID	**27. Latitude in decimal degrees**	**28. Longitude in decimal degrees**	**29. Lat/Long Source**
☐ N/A	*(WGS84 std: nn.nnnnnnn)*	*(WGS84 std: -nnn.nnnnnnn)*	☐ Actual ☐ Estimated

30.A. Railroad Use *	**31.A. State Use** *
30.B. Railroad Use *	**31.B. State Use** *
30.C. Railroad Use *	**31.C. State Use** *
30.D. Railroad Use *	**31.D. State Use** *
32.A. Narrative *(Railroad Use)* *	**32.B. Narrative** *(State Use)* *

33. Emergency Notification Telephone No. *(posted)*	**34. Railroad Contact** *(Telephone No.)*	**35. State Contact** *(Telephone No.)*

Part II: Railroad Information

1. Estimated Number of Daily Train Movements

1.A. Total Day Thru Trains *(6 AM to 6 PM)*	1.B. Total Night Thru Trains *(6 PM to 6 AM)*	1.C. Total Switching Trains	1.D. Total Transit Trains	1.E. Check if Less Than One Movement Per Day ☐ How many trains per week? ____

2. Year of Train Count Data *(YYYY)*

3. Speed of Train at Crossing
3.A. Maximum Timetable Speed *(mph)* ____
3.B. Typical Speed Range Over Crossing *(mph)* From ____ to ____

4. Type and Count of Tracks

Main ____ Siding ____ Yard ____ Transit ____ Industry ____

5. Train Detection *(Main Track only)*
☐ Constant Warning Time ☐ Motion Detection ☐AFO ☐ PTC ☐ DC ☐ Other ☐ None

6. Is Track Signaled?	7.A. Event Recorder	7.B. Remote Health Monitoring
☐ Yes ☐ No	☐ Yes ☐ No	☐ Yes ☐ No

FORM FRA F 6180.71 (Rev. 3/15) OMB approval expires 01/31/2026 Page 1 OF 2

U. S. DOT CROSSING INVENTORY FORM

A. Revision Date *(MM/DD/YYYY)*	PAGE 2	D. Crossing Inventory Number *(7 char.)*

Part III: Highway or Pathway Traffic Control Device Information

1. Are there Signs or Signals? ☐ Yes ☐ No

2. Types of Passive Traffic Control Devices associated with the Crossing

2.A. Crossbuck Assemblies *(count)*

2.B. STOP Signs *(R1-1)* *(count)*

2.C. YIELD Signs *(R1-2)* *(count)*

2.D. Advance Warning Signs *(Check all that apply; include count)* ☐ None
☐ W10-1 ________ ☐ W10-3 ________ ☐ W10-11 __________
☐ W10-2 ________ ☐ W10-4 ________ ☐ W10-12 __________

2.E. Low Ground Clearance Sign *(W10-5)* ☐ Yes *(count________)* ☐ No

2.F. Pavement Markings ☐ Stop Lines ☐ Dynamic Envelope ☐ RR Xing Symbols ☐ None

2.G. Channelization Devices/Medians ☐ All Approaches ☐ Median ☐ One Approach ☐ None

2.H. EXEMPT Sign *(R15-3)* ☐ Yes ☐ No

2.I. ENS Sign *(I-13)* Displayed ☐ Yes ☐ No

2.J. Other MUTCD Signs ☐ Yes ☐ No
Specify Type ______________ Count __________
Specify Type ______________ Count __________
Specify Type ______________ Count __________

2.K. Private Crossing Signs *(if private)* ☐ Yes ☐ No

2.L. LED Enhanced Signs *(List types)*

3. Types of Train Activated Warning Devices at the Grade Crossing ***(specify count of each device for all that apply)***

3.A. Gate Arms *(count)* Roadway _____ Pedestrian _____

3.B. Gate Configuration ☐ 2 Quad ☐ 3 Quad ☐ 4 Quad ☐ Full *(Barrier)* Resistance ☐ Median Gates

3.C. Cantilevered *(or Bridged)* Flashing Light Structures *(count)* Over Traffic Lane _____ ☐ Incandescent Not Over Traffic Lane _____ ☐ LED

3.D. Mast Mounted Flashing Lights *(count of masts)* __________ ☐ Incandescent ☐ LED ☐ Back Lights Included ☐ Side Lights Included

3.E. Total Count of Flashing Light Pairs

3.F. Installation Date of Current Active Warning Devices: *(MM/YYYY)* ______/__________ ☐ Not Required

3.G. Wayside Horn ☐ Yes Installed on *(MM/YYYY)* ______/__________ ☐ No

3.H. Highway Traffic Signals Controlling Crossing ☐ Yes ☐ No

3.I. Bells *(count)*

3.J. Non-Train Active Warning ☐ Flagging/Flagman ☐ Manually Operated Signals ☐ Watchman ☐ Floodlighting ☐ None

3.K. Other Flashing Lights or Warning Devices Count ___________ Specify type ____________________

4.A. Does nearby Hwy Intersection have Traffic Signals? ☐ Yes ☐ No

4.B. Hwy Traffic Signal Interconnection ☐ Not Interconnected ☐ For Traffic Signals ☐ For Warning Signs

4.C. Hwy Traffic Signal Preemption ☐ Simultaneous ☐ Advance

5. Highway Traffic Pre-Signals ☐ Yes ☐ No Storage Distance * ____________ Stop Line Distance * ____________

6. Highway Monitoring Devices *(Check all that apply)* ☐ Yes - Photo/Video Recording ☐ Yes – Vehicle Presence Detection ☐ None

Part IV: Physical Characteristics

1. Traffic Lanes Crossing Railroad ☐ One-way Traffic ☐ Two-way Traffic ☐ Divided Traffic Number of Lanes _______

2. Is Roadway/Pathway Paved? ☐ Yes ☐ No

3. Does Track Run Down a Street? ☐ Yes ☐ No

4. Is Crossing Illuminated? *(Street lights within approx. 50 feet from nearest rail)* ☐ Yes ☐ No

5. Crossing Surface *(on Main Track, multiple types allowed)* Installation Date * *(MM/YYYY)* _______/__________ Width * ______________ Length * _______________
☐ 1 Timber ☐ 2 Asphalt ☐ 3 Asphalt and Timber ☐ 4 Concrete ☐ 5 Concrete and Rubber ☐ 6 Rubber ☐ 7 Metal
☐ 8 Unconsolidated ☐ 9 Composite ☐ 10 Other *(specify)* ____

6. Intersecting Roadway within 500 feet? ☐ Yes ☐ No If Yes, Approximate Distance *(feet)* _______________

7. Smallest Crossing Angle ☐ 0° – 29° ☐ 30° – 59° ☐ 60° - 90°

8. Is Commercial Power Available? * ☐ Yes ☐ No

Part V: Public Highway Information

1. Highway System
☐ (01) Interstate Highway System
☐ (02) Other Nat Hwy System (NHS)
☐ (03) Federal AID, Not NHS
☐ (08) Non-Federal Aid

2. Functional Classification of Road at Crossing ☐ (0) Rural ☐ (1) Urban
☐ (1) Interstate ☐ (5) Major Collector
☐ (2) Other Freeways and Expressways
☐ (3) Other Principal Arterial ☐ (6) Minor Collector
☐ (4) Minor Arterial ☐ (7) Local

3. Is Crossing on State Highway System? ☐ Yes ☐ No

4. Highway Speed Limit ____________ MPH ☐ Posted ☐ Statutory

5. Linear Referencing System *(LRS Route ID)* *

6. LRS Milepost *

7. Annual Average Daily Traffic *(AADT)* Year _______ AADT ____________

8. Estimated Percent Trucks __________________ %

9. Regularly Used by School Buses? ☐ Yes ☐ No Average Number per Day ___________

10. Emergency Services Route ☐ Yes ☐ No

Submission Information - *This information is used for administrative purposes and is not available on the public website.*

Submitted by _________________________________ Organization ______________________________________ Phone ______________ Date ____________

Public reporting burden for this information collection is estimated to average 30 minutes per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed and completing and reviewing the collection of information. According to the Paperwork Reduction Act of 1995, a federal agency may not conduct or sponsor, and a person is not required to, nor shall a person be subject to a penalty for failure to comply with, a collection of information unless it displays a currently valid OMB control number. The valid OMB control number for information collection is 2130-0017. Send comments regarding this burden estimate or any other aspect of this collection, including for reducing this burden to: Information Collection Officer, Federal Railroad Administration, 1200 New Jersey Ave. SE, MS-25 Washington, DC 20590.

FORM FRA F 6180.71 (Rev. 3/15) OMB approval expires 01/31/2026 Page 2 OF 2

Appendix B Shiny Application Code

```
library(shiny) #install.packages('shinyWidgets')

library(shinyWidgets) #install.packages('leaflet')

library(leaflet) #install.packages('DT')

library(DT) #install.packages(tigris)

library(tigris)

library(shinyFeedback)

library(tidyverse)

library(spatialrisk)

library(sf)

library(here)


Railroad <- here("data_update/data_update/Railroads-shp/TRANS_Railroad_DOT.shp") %>%

st_read() %>% sf::st_transform("+proj=longlat +datum=WGS84 +units=m")

HighWay <- here("data_update/data_update/Highways-shp/ TRANS_Highways_DOT.shp") %>%

 st_read() %>% sf::st_transform("+proj=longlat +datum=WGS84 +units=m")

aps_data <- read_rds("Data/aps.rds") %>% select(CrossingID, Risk)

crashes <- readRDS("Data/Crashes.rds")

inventory <- readRDS("Data/Inventory.rds") %>% mutate(TotalCrashes = 0, TotalHazmats = 0)

hazmat <- readRDS("Data/Hazmat.rds") %>% filter(!(is.na(Longitude)) || !(is.na(Latitude)))

nebraska <- tigris::counties(state = "NE", class = "sf") %>%

sf::st_transform("+proj=longlat +datum=WGS84 +units=m")

listOfdfs <- list(crashes, inventory, hazmat)
```

```
options(tigris_use_cache = TRUE)

Counties <- unique(c(crashes$County, inventory$County, hazmat$County))


# Define UI for application that sets up map and table

ui <- fluidPage(titlePanel("Nebraska Train Map"), includeCSS("www/styles.css"),

sidebarLayout(#Our inputs that can be used on both tabs simaltenously

sidebarPanel(pickerInput(inputId = "df", "Choose a Dataset",

choices = c("Crashes" = 1, "Inventory" = 2, "Hazmat" = 3), multiple = T, selected = 1),

pickerInput("county", "County Selection", choices=Counties,

options = list(`actions-box` = TRUE),multiple = T,

selected = unique(crashes$County)),

conditionalPanel(condition = "input.df.includes('1')",

dateRangeInput('dateRange', label = "Date Range",

start = "2016-01-02", end = "2020-12-28"),),

conditionalPanel(condition = "input.df.includes('1')",

sliderInput(inputId = "time", label = strong("Time Range"),

min = 0, max = 24, step = 1, value = c(0,24))),

conditionalPanel(condition = "input.df.includes('2')",

selectInput("radius", "Radius", c(0.01, 0.1, 0.25, 0.5 , 0.75, 1, 2), selected = .5)),

conditionalPanel(condition = "input.df.includes('2')",

actionButton("apsCalculator", "Calculate Risk of Crossing!")), textOutput("error")),


#Main showing of our app, which shows the map on one tab and a table on the another
```

```
mainPanel(tabsetPanel(
tabPanel(title = "Map", leafletOutput("map", height = "500px",width = "100%")),
tabPanel("Table", fluid = TRUE, titlePanel("Data Table"),
fluidRow(conditionalPanel(condition = "input.df.length > 1",
selectInput("dataview", "Dataset View", c("Crashes", "Inventory", "Hazard")))),
fluidRow(column(12, dataTableOutput('table'))),
fluidRow(downloadButton('download', "Download")))))))

# Define server logic required to create a map based on inputs of train crashes
server <- function(input, output) {  #ask about this
square_green <- makeIcon(iconUrl =
"https://www.freeiconspng.com/uploads/green-square-1.png",
iconWidth = 18, iconHeight = 18)
crash_triangle <-   #Need to find red
makeIcon(iconUrl =
"https://www.freeiconspng.com/uploads/red-triangle-png-20.png",
iconWidth = 18, iconHeight = 18)
train_crossing <- #change to yellow
makeIcon(iconUrl =
"https://www.freeiconspng.com/uploads/wrong-cross-x-free-icon-24.png",
iconWidth = 8, iconHeight = 8)

#Create a Map Error function. Returns a map when one of the datasets is empty no matter what.
```

```
ErrorMap <- function(){Map <- leafletProxy("map") %>% clearMarkers() %>%
clearControls() %>% clearShapes()
return(Map)}
CrashesInCircle <- function(radius, Crashmat, Crossings){
Rad <- as.numeric(radius) * 1609.34
if( nrow(Crashmat) == 0 || nrow(Crossings) == 0){ return(Crossings) #Output an error
# stop(safeError(paste("There is an empty dataframe input. Cannot generate points in circle for
Crashes")))}
else{for(i in 1:nrow(Crossings)){
pic <- spatialrisk::points_in_circle(Crashmat, Crossings$Longitude[i], Crossings$Latitude[i],
Longitude, Latitude, radius = Rad)
Crossings$TotalCrashes[i] <- nrow(pic)}
return(Crossings)}}
HazmatsInCircles <- function(radius, Crashmat, Crossings){Rad <- as.numeric(radius)
if( nrow(Crashmat) == 0 || nrow(Crossings) == 0){
return(Crossings)
# stop(safeError(paste("There is an empty dataframe input. Cannot generate points in circle for
Hazmat")))
}
for(i in 1:nrow(Crossings)){
pic <- spatialrisk::points_in_circle(Crashmat, Crossings$Longitude[i], Crossings$Latitude[i],
Longitude, Latitude, radius = Rad)
Crossings$TotalHazmats[i] <- nrow(pic)}
```

```
return(Crossings)}

HazmatMap <- function(hazmat){#Three Map Functions
leafletProxy("map", data = hazmat) %>% clearMarkers() %>% clearControls() %>%
clearShapes() %>% addMarkers(data = hazmat, lat = ~Latitude, lng = ~Longitude,
label = ~ReportID, icon = ~square_green, popup = ~paste0(
"Overall Risk Score: ", "<br>", "City: ", IncidentCity, "<br>", "Date: ", IncidentDate, "<br>",
"Time: ", IncidentTime, "<br>", "Quantity Released: ", QuantityReleased, "<br>",
"Commodity: ", `Commodity Long Name`, "<br>", "Result: ", IncidentResult, "<br>")) %>%
addPolylines(data = HighWay, color = 'gray', weight = .5, highlightOptions =
highlightOptions(color = 'white', weight = 3)) %>%
addPolylines(data = Railroad, color = 'black', weight = 8, opacity = 1, highlightOptions =
highlightOptions(color = 'white', weight = 3))}
CrossingsMap <- function(crossings, radius){leafletProxy("map", data = crossings) %>%
clearMarkers() %>% clearControls() %>% clearShapes() %>% addMarkers(data = crossings,
icon = ~train_crossing, lat = ~Latitude, lng = ~Longitude, label = ~CrossingID,
popup = ~paste0("Overall Risk Score: ", "<br>", "Railroad Name: ", `Railroad Name`, "<br>",
"Crossing Type: ", CrossingType, "<br>",
"Total Crashes within radius of ", input$radius, " miles: ", TotalCrashes, "<br>",
"Total Hazmats within radius of ", input$radius, " miles: ", TotalHazmats, "<br>",
"Gate or Sign: ", GateOrSign, "<br>", "Land Use: ", `Land Use`, "<br>",
"Surface Type: ", SurfaceType, "<br>", "AADT: ", AADT, "<br>")) %>%
addCircles(data = crossings, lng = ~Longitude, lat = ~Latitude,
```

```
color = 'yellow', opacity = .01, fillOpacity = .1,
radius = as.numeric(input$radius) * 1609.344) %>%
addPolylines(data = HighWay, color = 'gray', weight = .5, highlightOptions =
highlightOptions(color = 'white', weight = 3)) %>%
addPolylines(data = Railroad, color = 'black', weight = 5, opacity = 1, highlightOptions =
highlightOptions(color = 'white', weight = 3))}
CrashMap <- function(crash){leafletProxy("map", data = crash) %>%
clearMarkers() %>% clearControls() %>% clearShapes() %>% addMarkers(data = crash,
lat = ~Latitude, lng = ~Longitude, label = ~AccidentID, icon = ~crash_triangle,
popup = ~paste0("Crash Day: ", Accident_Date, "<br>", "Time: ", Accident_Time, "<br>",
"Weather at Crash: ", Weather, "<br>", "Type of Crash: ", CrashType, "<br>",
"Accident Severity", AccSeverity, "<br>", "Total Vehicles: ", TotVeh, "<br>",
"Total Injured: ", TotInj, "<br>", "Total Fatalaties", TotFatal, "<br>")) %>%
addPolylines(data = HighWay, color = 'gray', weight = .5, highlightOptions =
highlightOptions(color = 'white', weight = 3)) %>%
addPolylines(data = Railroad, color = 'black', weight = 8, opacity = 1, highlightOptions =
highlightOptions(color = 'white', weight = 3))}
Datasets <- reactiveValues(data = inventory, crashdata = crashes, hazmatdata = hazmat)
output$map <- renderLeaflet({leaflet(options = leafletOptions(preferCanvas = TRUE)) %>%
addTiles() %>% addPolylines(data = HighWay, color = 'gray', weight = .5,
highlightOptions = highlightOptions(color = 'white', weight = 3)) %>%
addPolylines(data = Railroad, color = 'black', weight = 8, opacity = 1, highlightOptions =
highlightOptions(color = 'white', weight = 3)) %>%
```

```
addProviderTiles("OpenStreetMap", group = "OpenStreetMap",
options = providerTileOptions(updateWhenZooming = FALSE,
# map won't update tiles until zoom is done
updateWhenIdle = TRUE # map won't load new tiles when panning
))}) %>%
bindCache(input$df , input$time ,  input$county, input$dateRange, input$radius)
#Need logic when both are selected
#Use observe event
#Observer event for creating a map based on all four inputs. Looks for dataset
observeEvent( c(input$df , input$time ,  input$county, input$dateRange, input$radius), {
Datasets$data <- inventory
if(length(input$df) == 3){Datasets$data <- Datasets$data %>%
filter(County %in% input$county)
Datasets$data  <- CrashesInCircle(input$radius, Datasets$crashdata, Datasets$data)
Datasets$data  <- HazmatsInCircles(input$radius, Datasets$hazmatdata,Datasets$data)
Datasets$crashdata <- Datasets$crashdata %>% filter(County %in% input$county) %>%
filter(Accident_Date > input$dateRange[1] & Accident_Date < input$dateRange[2]) %>%
filter(Accident_Time > as.numeric(input$time[1]) * 100 & Accident_Time <
as.numeric(input$time[2]) * 100)
Datasets$hazmatdata <- Datasets$hazmatdata %>% filter(County %in% input$county)
if( Datasets$data == 0 || Datasets$crashdata  == 0 || Datasets$hazmatdata == 0){ErrorMap()}
else{CrashMap(Datasets$crashdata) %>% addMarkers(data = Datasets$data,
icon = ~train_crossing, lng = ~Longitude, lat = ~Latitude, label = ~CrossingID,
```

```
popup = ~paste0("Overall Risk Score: ", "<br>", "Railroad Name: ", `Railroad Name`, "<br>",
"Crossing Type: ", CrossingType, "<br>",
"Total Crashes within radius of ", input$radius, " miles: ", TotalCrashes, "<br>",
"Total Hazmats within radius of ", input$radius, " miles: ", TotalHazmats, "<br>",
"Gate or Sign: ", GateOrSign, "<br>", "Land Use: ", `Land Use`, "<br>",
"Surface Type: ", SurfaceType, "<br>", "AADT: ", AADT, "<br>")) %>%
addCircles(data = Datasets$data, lng = ~Longitude, lat = ~Latitude, color = 'yellow', opacity = 1,
radius = as.numeric(input$radius) * 1609.344) %>%
addMarkers(data = Datasets$hazmatdata, lat = ~Latitude, lng = ~Longitude, label = ~ReportID,
icon = ~square_green, popup = ~paste0("Overall Risk Score: ", "<br>",
"City: ", IncidentCity, "<br>", "Date: ", IncidentDate, "<br>", "Time: ", IncidentTime, "<br>",
"Quantity Released: ", QuantityReleased,"<br>",
"Commodity: ", `Commodity Long Name`, "<br>", "Result: ", IncidentResult, "<br>")) %>%
addLegend(color = c('green', 'blue', 'red'), title = 'Map Legend',
labels = c('Hazards','Crossings','Crashes'), position = 'bottomright', opacity = 0.9)}}
else if(length(input$df) == 2){if( 1 %in% input$df ){
Datasets$crashdata <- Datasets$crashdata %>% filter(County %in% input$county) %>%
filter(Accident_Date > input$dateRange[1] & Accident_Date < input$dateRange[2]) %>%
filter(Accident_Time > as.numeric(input$time[1]) * 100 & Accident_Time <
as.numeric(input$time[2]) * 100)
if(2%in% input$df){Datasets$data <- Datasets$data %>% filter(County %in% input$county)
Datasets$data <- CrashesInCircle(input$radius, Datasets$crashdata, Datasets$data)
Datasets$data <- HazmatsInCircles(input$radius, Datasets$hazmatdata,Datasets$data)
```

```
if( nrow(Datasets$data) == 0 || nrow(Datasets$crashdata) == 0){ErrorMap()}
else{CrashMap(Datasets$crashdata) %>% addMarkers(data = Datasets$data,
icon = ~train_crossing, lng = ~Longitude, lat = ~Latitude, label = ~CrossingID,
popup = ~paste0("Overall Risk Score: ", "<br>", "Railroad Name: ", `Railroad Name`, "<br>",
"Crossing Type: ", CrossingType, "<br>",
"Total Crashes within radius of ", input$radius, " miles: ", TotalCrashes, "<br>",
"Total Hazmats within radius of ", input$radius, " miles: ", TotalHazmats, "<br>",
"Gate or Sign: ", GateOrSign, "<br>", "Land Use: ", `Land Use`, "<br>",
"Surface Type: ", SurfaceType, "<br>", "AADT: ", AADT, "<br>")) %>%
addCircles(data = Datasets$data, lng = ~Longitude, lat = ~Latitude, color = 'yellow',
opacity = 1, radius = as.numeric(input$radius) * 1609.344) %>%
addLegend(color = c('blue', 'red'), title = 'Map Legend', labels = c('Crossings','Crashes'),
position = 'bottomright', opacity = 0.9)}}
else{Datasets$hazmatdata %>% filter(County %in% input$county)
if(nrow(Datasets$hazmatdata) == 0 || nrow(Datasets$crashdata) == 0){ErrorMap()}
else{HazmatMap(Datasets$hazmatdata) %>% addMarkers(data = Datasets$crashdata,
lat = ~Latitude, lng = ~Longitude, label = ~AccidentID, icon = ~crash_triangle,
popup = ~paste0("Overall Risk Score: ", "<br>", "Crash Day: ", Accident_Date, "<br>",
"Time: ", Accident_Time, "<br>", "Weather at Crash: ", Weather, "<br>",
"Type of Crash: ", CrashType, "<br>", "Accident Severity", AccSeverity, "<br>",
"Total Vehicles: ", TotVeh, "<br>", "Total Injured: ", TotInj, "<br>",
"Total Fatalaties", TotFatal, "<br>")) %>%
addLegend(color = c('green',  'red'), title = 'Map Legend', labels = c('Hazards','Crashes'),
```

```
position = 'bottomright', opacity = 0.9)}}}
else{Datasets$data <- Datasets$data %>% filter(County %in% input$county)
Datasets$data  <- CrashesInCircle(input$radius, Datasets$crashdata, Datasets$data)
Datasets$data  <- HazmatsInCircles(input$radius, Datasets$hazmatdata,Datasets$data)
Datasets$hazmatdata <- Datasets$hazmatdata %>% filter(County %in% input$county)
if(nrow(Datasets$hazmatdata) == 0 || nrow(Datasets$data) == 0){ErrorMap()}
else{HazmatMap(Datasets$hazmatdata) %>% addMarkers(data = Datasets$data,
icon = ~train_crossing, lng = ~Longitude, lat = ~Latitude, label = ~CrossingID,
popup = ~paste0("Overall Risk Score: ", "<br>", "Railroad Name: ", `Railroad Name`, "<br>",
"Crossing Type: ", CrossingType, "<br>",
"Total Crashes within radius of ", input$radius, " miles: ", TotalCrashes, "<br>",
"Total Hazmats within radius of ", input$radius, " miles: ", TotalHazmats, "<br>",
"Gate or Sign: ", GateOrSign, "<br>", "Land Use: ", `Land Use`, "<br>",
"Surface Type: ", SurfaceType, "<br>", "AADT: ", AADT, "<br>")) %>%
addCircles(data = Datasets$data, lng = ~Longitude, lat = ~Latitude,
color = 'yellow', opacity = 1, radius = as.numeric(input$radius) * 1609.344) %>%
addLegend(color = c('green', 'blue'), title = 'Map Legend', labels = c('Hazards','Crossings'),
position = 'bottomright', opacity = 0.9)}}}
else if(length(input$df) == 1){if( as.numeric(input$df) == 1) {
Datasets$crashdata <- crashes %>% filter(County %in% input$county) %>%
filter(Accident_Date > input$dateRange[1] & Accident_Date < input$dateRange[2]) %>%
filter(Accident_Time > as.numeric(input$time[1]) * 100 & Accident_Time <
as.numeric(input$time[2]) * 100)
```

```
if ( nrow(Datasets$crashdata) == 0){ErrorMap()}
else{CrashMap(Datasets$crashdata) %>% addLegend(color = 'red',
title = 'Map Legend', labels = 'Crashes', position = 'bottomright', opacity = 0.9)}}
else if (as.numeric(input$df) == 2) {
Datasets$data <- inventory %>% filter(County %in% input$county)
Datasets$data <- CrashesInCircle(input$radius, Datasets$crashdata, Datasets$data)
Datasets$data <- HazmatsInCircles(input$radius, Datasets$hazmatdata,Datasets$data)
if (nrow(Datasets$data) == 0){ErrorMap()}
else{CrossingsMap(crossings = Datasets$data, radius = input$radius) %>%
addCircles(data = Datasets$data, lng = ~Longitude, lat = ~Latitude,
color = 'yellow', opacity = 1, radius = as.numeric(input$radius) * 1609.344)}}
else if (as.numeric(input$df) == 3){
Datasets$hazmatdata <- hazmat %>% filter(County %in% input$county)
if ( nrow(Datasets$hazmatdata) == 0){ErrorMap()}
else{HazmatMap( Datasets$hazmatdata ) %>% addLegend(color = 'green',
title = 'Map Legend', labels = 'Hazards', position = 'bottomright', opacity = 0.9)}}}
else{ErrorMap()}}, ignoreNULL = TRUE)

#Add implementation for two and three datasets. Should take an hour or so. Implement legends.
(Create function here) #Also change opacity/think about opacity for crossing
observeEvent(input$apsCalculator, {#Lines
merger <- left_join(Datasets$data, aps_data, by = "CrossingID") %>%
filter(!is.na(Risk)) %>% arrange(-Risk)
```

```
Datasets$data <- merger
if(nrow(Datasets$data) == 0){ErrorMap()}
else{pal <- colorNumeric(palette = "Yellows", domain = Datasets$data$Risk)
APSMap <- CrossingsMap(Datasets$data , input$radius) %>% clearMarkers() %>%
clearShapes() %>% addCircles(data = Datasets$data, lat = ~Latitude, lng = ~Longitude,
label = ~CrossingID, color = ~pal(Datasets$data$Risk), fillColor = ~pal(Datasets$data$Risk),
fillOpacity = 1, radius = as.numeric(input$radius) * 1609.344, popup = ~paste0(
"Overall Risk Score: ", Risk,  "<br>", "CrossingID: ", CrossingID, "<br>")) %>%
addPolylines(data = HighWay, color = 'gray', weight = .5, highlightOptions =
highlightOptions(color = 'white', weight = 3)) %>%
addPolylines(data = Railroad, color = 'black', weight = 5, opacity = 1, highlightOptions =
highlightOptions(color = 'white', weight = 3))
if(length(input$df) == 1){APSMap}
else if(length(input$df) == 2){APSMap %>% addMarkers(data = Datasets$crashdata,
lat = ~Latitude, lng = ~Longitude, label = ~AccidentID, icon = ~crash_triangle,
popup = ~paste0("Crash Day: ", Accident_Date, "<br>", "Time: ", Accident_Time, "<br>",
"Weather at Crash: ", Weather, "<br>", "Type of Crash: ", CrashType, "<br>",
"Accident Severity", AccSeverity, "<br>", "Total Vehicles: ", TotVeh, "<br>",
"Total Injured: ", TotInj, "<br>", "Total Fatalaties", TotFatal, "<br>"))}
else{APSMap %>% addMarkers(data = Datasets$crashdata, lat = ~Latitude, lng = ~Longitude,
label = ~AccidentID, icon = ~crash_triangle, popup = ~paste0(
"Crash Day: ", Accident_Date, "<br>", "Time: ", Accident_Time, "<br>",
"Weather at Crash: ", Weather, "<br>", "Type of Crash: ", CrashType, "<br>",
```

```
"Accident Severity", AccSeverity, "<br>", "Total Vehicles: ", TotVeh, "<br>",
"Total Injured: ", TotInj, "<br>", "Total Fatalaties", TotFatal, "<br>")) %>%
addMarkers(data = Datasets$hazmatdata, lat = ~Latitude, lng = ~Longitude, label = ~ReportID,
icon = ~square_green, popup = ~paste0("Overall Risk Score: ", "<br>",
"City: ", IncidentCity, "<br>", "Date: ", IncidentDate, "<br>",
"Time: ", IncidentTime, "<br>", "Quantity Released: ", QuantityReleased, "<br>",
"Commodity: ", `Commodity Long Name`, "<br>", "Result: ", IncidentResult, "<br>"))}}})
DataTable <- eventReactive( c(input$df , input$time ,  input$county, input$dateRange,
input$radius,input$dataview, input$apsCalculator), {
if( length(input$df) == 1){if( input$df == 2){if( nrow(Datasets$data) == 0){
stop(safeError(paste(input$df[1], "has filters that are outputting nothing!")))}
else if("Risk" %in% names(Datasets$data)){Datasets$data %>% arrange(-Risk)}
else{Datasets$data}}
else if (input$df == 1){if (nrow(Datasets$crashdata) == 0){
stop(safeError(paste(input$df[1], "has filters that are outputting nothing!")))}
else{Datasets$crashdata}}
else if(input$df == 3){if (nrow(Datasets$hazmatdata) == 0){
stop(safeError(paste(input$df[1], "has filters that are outputting nothing!")))}
else{Datasets$hazmatdata}}
else{stop(safeError(paste(input$df[1], "has filters that are outputting nothing!")))}}
else{if(input$dataview == 'Inventory'){
if(2 %in% input$df){if( input$df[[1]] == 1){
if("Risk" %in% names(Datasets$data)){Datasets$data %>% arrange(-Risk)}
```

```
Datasets$data}
else{if("Risk" %in% names(Datasets$data)){Datasets$data %>% arrange(-Risk)}
Datasets$data}}
else{stop(safeError(paste(input$dataview, "was not selected in dataframe input")))}}
else if(input$dataview == "Hazard"){
if(3 %in% input$df){Datasets$hazmatdata}
else{stop(safeError(paste(input$dataview, "was not selected in dataframe input")))}}
else{if(1 %in% input$df){Datasets$crashdata}
else{stop(safeError(paste(input$dataview, "was not selected in dataframe input")))}}}})

#Outputs table of our dataset
output$table = DT::renderDataTable({DataTable()})
output$download <- downloadHandler(filename = function(){
paste(input$dataview,".csv",sep="")},
content = function(fname){write.csv(DataTable(),fname)}})}

# Run the application
shinyApp(ui = ui, server = server)
```