A DATA-DRIVEN APPROACH TO SUPPORT BRIDGE

ASSET MANAGEMENT

By

CRISTIAN CONTRERAS-NIETO

Civil Engineer
Escuela Colombiana de Ingenieria Julio Garavito
Bogota D.C., Cundinamarca
2003

Master of Science in Civil Engineering
Oklahoma State University
Stillwater, Oklahoma
2014

Submitted to the Faculty of the
Graduate College of the
Oklahoma State University
in partial fulfillment of
the requirements for
the Degree of
DOCTOR OF PHILOSOPHY
December, 2017

A DATA DRIVEN APPROACH TO SUPPORT BRIDGE

ASSET MANAGEMENT

Dissertation Approved:

Dr. Yongwei Shan

Dissertation Adviser

Dr. M. Phil Lewis

Dissertation Co-Adviser

Dr. John N. Veenstra

Dr. Julie Ann Hartell

Dr. Ricki G. Ingalls

ACKNOWLEDGEMENTS

Without the support, time, and advice of several people, I would not have been able to complete this dissertation. First of all, I would like to acknowledge my advisor, Dr. Yongwei Shan, for his endless patience, support, advice, kindness, friendship, and generosity. His commitment and hard working character have demonstrated to me once again that success is the result of commitment and love for what you do. Also, I would like to acknowledge my co-advisor, Dr. Phil Lewis, for all the guidance and support over the past six years at OSU (masters and Ph.D.). I would not be the person, engineer, or professional who I am without of having the fortune to have these two amazing mentors. My next acknowledgement is to my doctoral committee members, Dr. John N. Veenstra (Head), Dr. Julie Ann Hartell, and Dr. Ricki G. Ingalls because of all the constructive advice to my research and their great classes. In addition,

I am truly grateful for the friendship, support, feedback, and data that Matt Mitchell, Area Maintenance Engineer of ODOT – Division 4, gave to me. He always took time to help me with any question that I had related to bridge maintenance guides in his division.

Through this academic adventure, several people supported me and gave me strength to continue working on and pursuing my goals. I would like to acknowledge my parents, my sisters, and my uncle (Shifu) for being in my life and not allowing me to give up in most difficult part of this journey.

Finally, but not the least, I would like to thank Diana, my lovely and wise wife, for all her love, patience, advice, and time. Thank you for teaching me that "with love everything can be done" and providing me your hand every time I fell and did not how to get up again.

Thank you God for blessing all of us!

Acknowledgements reflect the views of the author and are not endorsed by committee members or Oklahoma State University.

Name: CRISTIAN CONTRERAS-NIETO

Date of Degree: DECEMBER, 2017

Title of Study: A DATA-DRIVEN APPROACH TO SUPPORT BRIDGE ASSET MANAGEMENT

Major Field: CIVIL ENGINEERING

Abstract: Economic growth and reduction of poverty lies in a well-planned, constructed, and maintained infrastructure that includes water and sanitation networks, airports, schools, health facilities, and highways systems. As bridges are an integral component of the nation's highway system infrastructure, deficient bridges without timely maintenance may endanger the public and affect the economy on a broader scale. Currently, more than the 9.0% of the bridges in the U.S. are graded as structurally deficient, and the new estimate to address these bridges is $123 billion. Thus, in order to keep the level of safety and serviceability of these infrastructure assets, efforts in an accurate prediction of condition ratings, a better characterization of deficient bridges, and a focus on prioritization of deficient bridges can help. Currently, bridge stakeholders face budget constraints; thus, they need a systematic approach to better estimate maintenance budgets, make informed decisions in bridge design, and prioritize bridge maintenance. This dissertation research has two major objectives. The first objective is to provide a framework to predict and characterize superstructure deficiency. The second objective is to present a methodology to prioritize bridge maintenance. This dissertation used NBI databases as the main data source and utilized data mining techniques, multi-criteria decision analysis, and GIS to achieve the objectives of the study. Moreover, this dissertation follows a three-journal paper format. The first paper addresses the development of a framework to create predictive models of superstructure ratings for steel and prestressed concrete bridges. The second paper identifies a framework to characterize superstructure deficiency of steel bridges. The third paper presents a decision-making framework to prioritize bridge maintenance through using aggregate bridge ratings and average daily traffic (ADT). This dissertation contributes to the overall body of knowledge by establishing frameworks to develop reliable models to predict superstructure ratings, identify factors that accelerate superstructure deficiency, and prioritize bridge maintenance. The results of this dissertation can be used by any bridge stakeholder to complement their current bridge management programs.

TABLE OF CONTENTS

Chapter Page

Chapter Page

LIST OF TABLES

Table Page

LIST OF FIGURES

Figure Page

CHAPTER I

INTRODUCTION AND BACKGROUND

Due to the importance of infrastructure, stakeholders have been focused on tasks such as predicting infrastructure condition and managing infrastructure assets in order to optimize limited budgets and satisfy the community's needs. Therefore, academia has concentrated its effort to analyze and understand infrastructure lifespan and the effects of maintenance, materials, environment, and so on. However, every component of infrastructure is affected by different factors and may be seen as unique. For example, water systems have a different deterioration rate than highways. While a water line is affected by chemical characteristics of the water it carries, a road is impacted by the number of vehicles that travel on it every day. Furthermore, the same infrastructure component may have so many subgroups that have to be studied independently because of the same reason mentioned above (i.e. roads: highly transited vs. slightly transited). This research is focused on one component of infrastructure, bridges, and it considers different subgroups of bridges based on characteristics such as superstructure material or year built.

PROBLEM STATEMENT

Economic growth and reduction of poverty lies in a well-planned, constructed, and maintained infrastructure that includes water and sanitation systems, airports, schools, health facilities, and highways and bridges networks. The effectiveness and efficiency of domestic movement of goods and services are highly dependent on the highway system. Bridges are an integral component of the nation's highway system; therefore, their condition must be well-maintained. However, the overall bridge conditions in the state of Oklahoma received a rating of D^+, which is lower than the national average of C^+, according to the American Society of Civil Engineers, ASCE - 2013 Report Card for America's Infrastructure (ASCE, 2013). Furthermore, because of Oklahoma's percentage of structurally deficient bridges (21%), Oklahoma is one of the worst states in the nation. Moreover, although the spending on bridges has increased over the last decade, those investments have not been enough. The most recent estimate of bridge rehabilitation is $123 billion ASCE (2017). However, this estimate does not include the funds required for the preservation of bridges that have not reached deficiency yet. For this reason, Federal Government and Departments of Transportation (DOTs) must rely on effective maintenance and replacement decisions tools that can optimize the utility of the limited budgets while fulfilling the serviceability and safety levels required.

In order to find solutions to these issues, DOTs have started to develop and implement plans to improve their transportation systems. For example, the Oklahoma Department of Transportation (ODOT) is working on the Eight-Year Construction Work (Oklahoma Department of Transportation - ODOT, 2015). This plan includes more than 1,800 projects and more than $6.5 billion in improvements. Thus, the state of Oklahoma aims to eliminate structurally deficient bridges in the state by the next decade. However, the plan did not mention the strategies regarding how to address those bridges that are becoming structurally deficient. Because of this, it is important to provide the DOTs and other bridge stakeholders with reliable tools to forecast bridge deficiencies. Then they can predict when a bridge reaches deficiency levels, and so that preventive maintenance in lieu of corrective maintenance can be scheduled in advance. As a result, the number of deficient bridges could be minimized and maintenance budgets could be optimized.

In quest of boosting bridge conditions, academia has been involved in analyzing and understanding bridge deficiencies. Previous studies have approached the prediction of deficiency for the three main components of a bridge (deck, substructure, superstructure). As Veshosky, Beidleman, Buetow, and Demir (1994), Huang (2010), and Contreras-Nieto (2014) found, the bridge age is one of the most significant predictors of bridge deficiency. Nevertheless, some of these authors do not coincide in other bridge deficiency predictors, such as structure length, number of spans, or type of service under the bridge. Therefore, there is still no consensus on which and how other factors influence bridge deficiency other than age and average daily traffic (ADT).

In order to address and find optimal solutions for bridge deficiency, collaboration between bridge stakeholders (DOTs and bridge designers) and academia is integral. The findings of the collaborative research can be utilized in the decision making process of planning the maintenance and reconstruction of bridges. Consequently, all parties may achieve better results because of continuous dialog between the collaborators throughout the course of the research. Also, the fostered collaborative relationship between academia and industry could speed up the data sharing process and increase the quantity and quality of data as well. This collaborative relationship would guarantee the success of tools developed by academia and implemented by bridge stakeholders.

To consolidate the bridge inspections nationwide, The Federal Highway Administration (U.S. Department of Transportation) manages the inventory and inspection information collected by DOTs in the National Bridge Inventory (NBI) databases. Right now, digital data files are available from 1992 to 2016. NBI is used as an asset management database that has been the main source of information for many studies related to bridge deficiency. As a reference, the 2013 NBI database contains more than 607,000 structures and 110 variables.

To find a way to predict the causes of bridge deficiency, applying advanced data mining concepts and developing a variety of models using bridge inspections records can provide a better prediction than the current available models. In addition to the use of traditional techniques to build models, such as multiple regression, current data miners also employ new techniques, such as gradient boosting (Seni & Elder, 2010). These new techniques often outperform traditional ones; however, sometimes traditional ones outperform the new ones depending on the research goal, interpretability, and available data. Regardless of the implementation of techniques, data miners analyze and choose the best model according to their scope, model performance, and expectations. Data mining software like SAS Enterprise Miner MT gives researchers the opportunity to select the model with the best performance by comparing and analyzing the results of fit statistics, such as misclassification rate or average squared error (Christie, 2011). Comparing the models developed and choosing the best model benefits the research in terms of accuracy and stability of predictions.

In addition, bridge management is facing another problem: how to prioritize bridge maintenance with limited budgets. For example, bridge engineers must examine the state of their bridges to select which structures would be considered for maintenance or other types of services. This process is done every year in ODOT – Bridge Divisions. In addition to the annual budget constraints, ODOT – Bridge Divisions do not have adequate methodology or tools to assist bridge engineers with the analysis. Therefore, this process may take weeks of the valuable time of those professionals and their assistants.

Nevertheless, bridge inspections include a lot of information. For example, the bridge inspectors evaluate the three main components of bridges: deck, substructure, and superstructure in detail (FHWA, 1995). Each of the components is given a proper rating, ranging from 0 being the worst condition to 9 being the best condition. Deck rating reports the complete condition of the deck. Substructure rating describes the physical condition of abutments, fenders, footings, piles, piers, and other elements. The physical condition of all structural members (girders/beams, superstructure joints, protective coating, etc.) is reported in the superstructure rating. The superstructure is the section that receives and supports loads (traffic) from the deck. Also, the superstructure transfers the reactions of the load to the substructure. The following two chapters are focused on the significance of the superstructure and the variety of the materials and construction connections used in the superstructure. In Chapter 4 the three ratings are combined in order to prioritize bridge maintenance.

Inevitably, applicable and reliable models to forecast bridge ratings are essential to determine bridge life-span, plan bridge maintenance in advance, and choose building materials. Also, bridge engineers need a methodology to analyze bridge condition and prioritize bridge maintenance efficiently. Therefore, this dissertation aims to develop models to predict and understand the superstructure rating by applying data mining techniques. Moreover, the methodology implemented to create the models in this study can be utilized to produce models for the other two bridge ratings (deck and substructure ratings). In order to complement the condition prediction and characterization of bridges, this study also plans to address the prioritization of bridge maintenance by applying a multi-criteria decision making system integrated with Geography Information System to visualize and streamline the process.

DISSERTATION STRUCTURE

This dissertation includes five chapters. The first chapter introduces the problem statement, objectives, structure of the dissertation, research objectives, scope of the research, and literature review. Chapter 2, 3, and 4 follow in the format of journal papers. Any of these chapters can be served as a journal paper, and they will be submitted to peer-reviewed academic journals. Therefore, each chapter from Chapter 2 to 4 consists of an abstract, introduction, background and literature review, research methodology, results, conclusions and recommendations, and references. While each paper answers a research question and can be considered as a standalone paper, some overlaps may be observed between Chapter 2 and Chapter 3. Then, conclusions that state the general findings of the dissertation research, contributions to the overall body of knowledge, limitations of the research, and recommendations for future studies form Chapter 5. Finally, the appendices include additional material such as scatter plots, histograms, tables with predicted superstructure ratings, maintenance prioritization of bridges (ranking), and some completed surveys related to Chapter 4

RESEARCH SCOPE

This doctoral research is divided into two major parts: predicting/characterizing superstructure deficiency, and prioritizing bridge maintenance. The first part has the purpose of developing a methodology that can be used by bridge stakeholders to predict/understand bridge deficiency. Also, the methodology could be generalized to any of the superstructure materials or bridge datasets.

As a source of data, the National Bridge Inventory - NBI database (U. S. Department of Transportation. Federal Highway Administration, 2015) was used. The NBI database assembles reliable information of bridge inspections across the nation and contains the condition of more than 607,000 bridges as of 2013 (U. S. Department of Transportation. Federal Highway Administration, 2015). Furthermore, several previous studies have been based on this database, such as Veshosky et al. (1994), Tang, Kanaan, Wnag, Oh, and Kwigizile (2012), and Lee (2012). In addition, more than 100 items are kept in the NBI files, which consist of characteristics, conditions, and ratings of each bridge. Three main NBI ratings reveal the condition of the bridge at the moment of the inspection; these ratings are deck, substructure, and superstructure. Because superstructure has been the interest of previous research of the author (Contreras-Nieto, 2014), this study is focused on superstructure rating as well. However, the methodology developed could be applied for deck or substructure ratings. Although bridge inspections from 1992 to 2016 are available in digital form, this research is based on the 2013-NBI and 2014-NBI databases. This is due to the fact that 2013-NBI was the latest database at the time when the author started his research For predicting superstructure rating, Chapter 2, the subset was created based on the following parameters: year built (≥ 1955), state (Oklahoma), kind of material-design (prestressed concrete and steel), type of design (stringer/multibeam or girder), and deck type (concrete cast-in-place). On the other hand, in characterizing steel bridge deterioration (Chapter 3) the subset was created based on the following parameters: year built (all steel bridges), state (nationwide), kind of material-design (steel and steel continuous), type of design (stringer/multibeam or girder), and deck type (concrete cast-in-place).

In order to develop the models, different techniques were implemented. Nevertheless, the techniques used for prediction will not be the same for characterization. The reason for this is that while predicting the superstructure rating is focused on a continuous variable (values are from 0 to 9), characterizing the superstructure is focused on a binary variable (values are deficient and non-deficient). Thus, multiple regression, regression trees, and artificial neural networks were implemented for predicting the superstructure rating. On the other hand, logistic regression, decision trees, artificial neural networks, gradient boosting, and support vector machine were applied for characterizing the deterioration of steel bridges.

Also, the second major part intends to optimize the process of prioritizing bridge maintenance. Therefore, the author has been working with the Division 4 Bridge Engineer and County Programs of the Oklahoma Department of Transportation, in order to develop a new approach to improve the bridge maintenance selection process in the ODOT Division 4. Although, the methodology proposed was implemented to rate the entire population of deficient bridges contained in 2014-NBI, it can also be used to prioritize deficient bridges in any of the eight ODOT divisions by employing their own databases. The information and bridges contained in the ODOT database are partially different from the NBI database. The main difference is that the ODOT database is updated almost daily with information collected during the bridge inspections done in previous days by ODOT inspectors. This means that the bridge condition recorded in this database is more accurate than the condition recorded in the NBI databases.

RESEARCH OBJECTIVES

The conceptual overview of the research objectives is presented in Figure 1-1. The primary objective of this research is to develop new methodologies and tools to support bridge asset management by applying data mining concepts and techniques as well as decision making analyses. This main objective can be divided into three secondary objectives, as follows:

1. Develop predictive models of superstructure ratings for steel and prestressed concrete.
2. Characterize steel bridge deterioration in order to understand what factors and how they influence superstructure deficiency.
3. Develop a decision-making process with the assistance of GIS to prioritize bridge maintenance

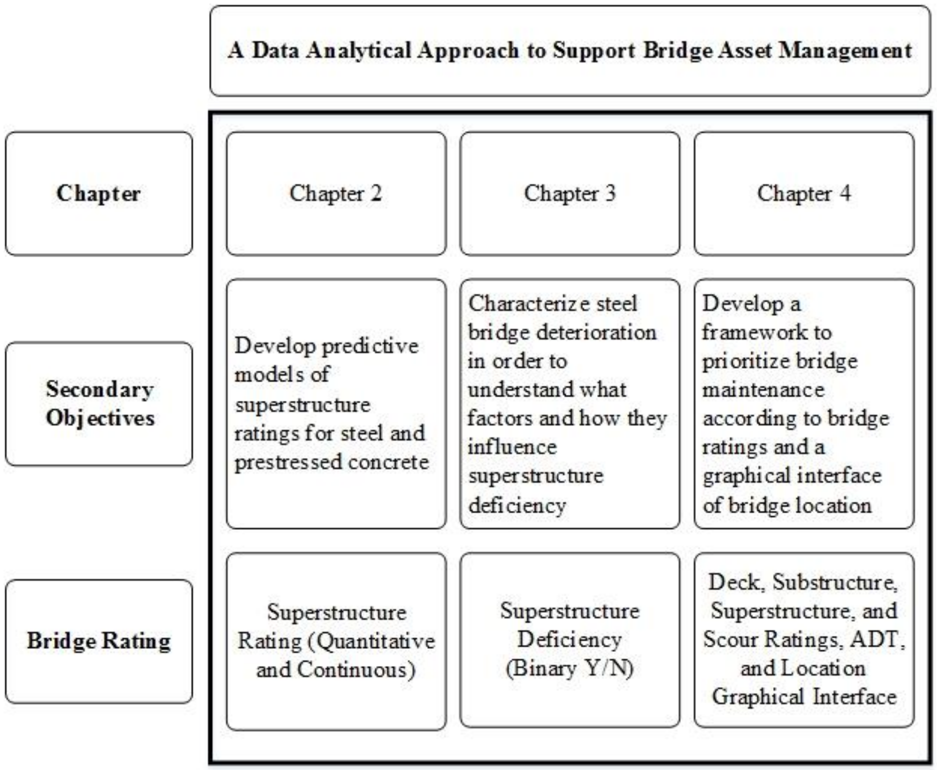

Figure 1-1 Conceptual Overview of Research Questions and Dissertation Structure

REFERENCES

America Society of Civil Engineers - ASCE (2013). "2013 Report Card of America's Infrastructure." <http://www.infrastructurereportcard.org/>. (July 2015).

America Society of Civil Engineers - ASCE (2017). "2017 Report Card of America's Infrastructure." <https://www.infrastructurereportcard.org/cat-item/bridges/>. (June 2017).

Berry, M. J. A., and Linoff, G. (2011). *Data mining techniques: for marketing, sales, and customer relationship management*, Wiley Pub, Indianapolis, Ind.

Christie, P. G., Jim; thompson, Jeff; and Well, Chip (2011). "Applied Analytics Using SAS® Enterprise MinerTM Course Notes." SAS Institute Inc.

Contreras-Nieto, C. (2014). "Development of linear models to predict superstructure ratings of steel and prestressed concrete bridges." Dissertation/Thesis, ProQuest, UMI Dissertations Publishing.

DeVille, B., and Neville, P. (2013). *Decision Trees for Analytics Using SAS® Enterprise Miner™*, SAS Institute Inc.

Federal Highway Administration - FHWA (1995). "Recording and Coding Guide for the Structure Inventory and Appraisal of the Nation's Bridges." U.S. Dept. of Transportation, Federal Highway Administration, Washington, D.C.

Huang, Y. (2010). "Artificial Neural Network Model of Bridge Deterioration." *Journal of Performance of Constructed Facilities*, 24(6), 597-602.

Kim, B.-H. (2009). "Support Vector Machine & Classification Using Weka (Brain, Computation, and Neural Learning)." Seoul Nation University.

Lee, S. (2012). "Current state of bridge deterioration in the US-part1." *Materials performance*, 51(1), 62-67.

Oklahoma Department of Transportation - ODOT (2015). "ODOT Inveils Eigth-Year Construction Work Plan." <http://www.ok.gov/odot/Eight-Year_Construction_Work_Plan.html>. (September 2015).

Seni, G., and Elder, J. F. (2010). "Ensemble Methods in Data Mining: Improving Accuracy Through Combining Predictions." *Synthesis Lectures on Data Mining and Knowledge Discovery*, 2(1), 1-126.

Tang, P., Kanaan, O., Wnag, J., Oh, J., and Kwigizile, V. (2012). "Statistical Identification of Items Importants for Evaluating Bridge Conditions." *Transportation Research Board.*

U. S. Department of Transportation. Federal Highway Administration (2015). <https://www.fhwa.dot.gov/bridge/nbi/ascii.cfm>. (September, 2015).

U.S. Department of Transportation (2013). "Federal Highway Administration National Bridge Inventory." <https://www.fhwa.dot.gov/bridge/nbi/ascii.cfm?year=2013>. (2013).

Veshosky, D., Beidleman, C., Buetow, G., and Demir, M. (1994). "Comparative Analysis of Bridge Superstructure Deterioration." *Journal of Structural Engineering*, 120(7), 2123-2136.

CHAPTER II

DEVELOPING PREDICTIVE MODELS OF SUPERSTRUCTURE RATINGS FOR STEEL AND PRESTRESSED CONCRETE BRIDGES

ABSTRACT

A large number of deficient bridges may endanger the public and affect the economy on a broader scale. Bridge superstructure rating is a critical element that affects the overall sufficiency rating of a bridge. Accurately predicting the superstructure performance of a bridge may help agencies better prioritize their resources for maintenance and repairs. The main objective of this paper is to utilize data mining techniques to develop reliable models to predict the superstructure rating of bridges. This research utilizes the National Bridge Inventory (NBI) database as the main source of information. A focused subset was created based on the defined scope of the research: year built ($\geq$ 1955), superstructure material (prestressed concrete and steel), type of design (stringer/multi-beam or girder), and deck type (concrete cast-in-place). This paper takes three approaches for model development, including linear regression, decision tree, and neural network. The best model was identified for each superstructure material through comparisons among different models. In addition, a discussion of individual variables and their contributions to predicting superstructure rating was presented. The identified models provide values in helping to determine the timing for a bridge superstructure's maintenance and reconstruction.

INTRODUCTION

Bridges are a critical component of the national infrastructure. According to the 2017 American Society of Civil Engineers Infrastructure Report Card, 9.1% of bridges in the nation are rated as structurally deficient (ASCE 2017). The Moving Ahead for Progress in the 21st Century (MAP-21) Act mandates every state to invest resources to meet the national goal of collectively improving transportation asset performance (FHWA 2015). In response to this mandate, many state departments of transportation (DOTs) have made asset management their priority. However, the gap between desired and available funds requires state DOTs to be sophisticated in prioritizing maintenance and repair schedules. As a result, it is necessary for transportation agencies to understand and predict bridge deterioration. Then, transportation agencies can schedule maintenance and allocate proper funds in an optimal manner.

Superstructure is a critical element of bridges since it directly bears the load of traffic and deck components. In addition, the rating of the superstructure affects the overall sufficiency rating of a bridge, which is the combination of condition ratings (superstructure, deck, substructure) and other bridge scores. The deck, superstructure, and substructure of a bridge function as a whole and each component is vital for fulfilling levels of safety and serviceability. Bridge superstructure condition is a topic of interest for government agencies, concrete and steel companies, and academia (FHWA 2014).

Also, prediction of infrastructure assets' conditions, such as bridge ratings, is a critical methodical process for infrastructure asset management. This process is based on maintenance, improvements, and operation of the assets with a cost-effective approach (FHWA and ASSHTO, 2000). Thus, any agency responsible for bridge management is required to implement methodologies to extract valuable information from the bridge inspections, which finally will be translated into future bridge condition ratings. Then asset management decisions can be made based on the results obtained.

In order to build prediction models, this study applied data mining techniques to the 2013 National Bridge Inventory data (NBI-2013). This data mining approach spans from data partitioning and missing value imputations to comparing candidate models and selecting the final model. An interpretation of the models is provided, and significant predictors of superstructure ratings are discussed. In addition, each model with the best performance is implemented to two NBI datasets (2013 and 2014) in order to score the corresponding bridges (observations) and predict the superstructure ratings for these two years. Therefore, it is possible to determine how close the predicted ratings are to the real (observed) superstructure ratings contained in the NBI databases. Limitations and advantages of the developed models are presented. As a final step, validation and implementation of the models was performed, which is distinct from the majority of previous studies whose results of predicting models are not disclosed (Moomen, 2016; Contreras-Nieto, 2014; Huang, 2010; Veshosky et al., 1994).

BACKGROUND AND LITERATURE REVIEW

The prediction of bridge ratings has been an interest of researchers for decades, especially since the 1990s. Therefore, various approaches have been applied to available data at the time. Some of the frequently used techniques are linear regression, Markov models (stochastic model), artificial neural networks (ANN), and regression and classification decision trees. The focus of those studies has been the three main NBI ratings - deck, substructure, and superstructure. This section provides a background of the studies, and it is divided according to the NBI ratings of interest (target variable).

Some studies have been centered on the analysis of the three NBI ratings. For example, Tang et al. (2012) analyzed which bridge parameters are significant for predicting NBI ratings through Pearson's correlation, multiple regression, and Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC). Nevertheless, this study did not involve model development. Another study conducted by Carriquiry et al (2013) presented an approach of applying decision trees to the Commonly – Recognized Structural elements (Core elements) and NBI databases. Although the authors found that the decision tree technique is better for predicting the ratings than the NBI translator software, it is not clear which Core elements are significant. Moreover, Bu et al. (2015) compared a standard Markovian model and an Elman Neural Network and Back Prediction model in order to predict transition probabilities of the condition ratings of bridges without maintenance. However, the Markov models are based on probability distributions, and their results are expressed in a transition probability matrix (Thompson and Johnson 2005). Therefore, it is not possible to determine the effect of inputs on the target variable as other techniques can; this can be noticed in the results of studies such as Micevski et al. (2002), Thompson and Johnson (2005), and Bu et al. (2015). More recently, Moomen (2016) developed an analysis to predict bridge condition ratings using Indiana's NBI databases. In this study, the author proposed and implemented both a deterministic (regression model) and a probabilistic (ordered probit model) approach to different families of bridges. It was found that deterministic models predict better bridge ratings than the binary probit models developed in this study. Although the author mentioned the access to NBI databases from 1992 to 2016, he did not clarify what year of NBI database was used. Also, the author could have used all 24 years of data, but he did not explain how the independence assumption of data was addressed, which is very important for any regression analysis.

Regarding deck-rating prediction, some of the most relevant studies are presented. First, Kim and Yoon (2010) used Pearson's correlation and multiple linear regression techniques to identify elements of deck deterioration in NBI inspections. Another analysis was peformed by Huang (2010). The author investigated 942 decks in the state of Wisconsin and his data sources were inspection, maintenance, and inventory databases. Artificial Neural Networks – Multilayer perceptron was the technique chosen for developing the model. Elbehairy et al. (2006) also presented an approach of implementing a Markovian model to predict deck ratings in order to prioritize deck maintenance. However, the probability transition matrix used in the study was developed in 1988 by Jiang et al. (1988) and was based on bridge deck materials, type of highways, and bridge age.

A few studies have focused on superstructure-rating prediction. For instance, Veshosky et al. (1994) applied regression analysis to the NBI (1990) database in order to predict superstructure ratings. They built two models to predict the superstructure rating: one model for prestressed concrete bridges and one for steel bridges. Similarly, Contreras-Nieto (2014) utilized Pearson's correlation and multiple regression to predict bridge superstructure ratings of prestressed concrete and steel bridges. Although the R-square of the two models exceeded those obtained by Veshosky et al. (1994), the models by Contreras-Nieto accounted for about 30% of the variation in the data.

The prediction of superstructure rating has been addressed in different ways since the 1990s. However, all studies have aimed to develop predictive models and to understand the parameters that are significant to superstructure deterioration. There has not been a single study that evaluates the performance of an array of analytic techniques to select the most reliable prediction model. Therefore, this research departs from the current body of knowledge by applying data mining concepts and three families of techniques (linear regression, decision trees, and artificial neural networks) to develop models to predict superstructure ratings by material type (prestressed concrete and steel). The model with the best performance was selected as the best model for each material. As a result of applying this methodology, the author believes that the most reliable models for the prediction of the superstructure ratings for each material type can be obtained.

Techniques implemented in this research have also been successfully implemented in other fields such as marketing, healthcare/medicine, geography, criminology, and other engineering areas. For example, Maryam and Marzieh (2012) used neural networks to predict customers’ response to advertisement using hybrid databases. It was proved that better prediction rates were obtained in their studies compared to previous studies. Using logistic regression, decision trees, artificial neural networks, and support vector machines, Ang and Goh (2013), using data from Asia, developed models in order to predict delinquent behavior and distinguish factors that characterize juvenile delinquents. The results were considered promising because of the high accuracy rates of some of the models; artificial neural networks and decision trees came up with accuracy rates higher than 95%. To predict urban changes, Boulila et al. (2011) implemented decision trees to estimate the change of urban areas identified from satellite image databases. The results showed that fuzzy decision trees provided help to map and interpret environmental changes. This also can be implemented in prevention and monitoring disasters. The promising results obtained in the previous studies lends a strong argument to apply some of these techniques to the NBI databases.

METHODOLOGY

In this section, the data source and research parameters are introduced. Additionally, a brief description of the data mining methodology and techniques implemented in this study through the use of SAS® Enterprise Miner are explained.

Data Source

The data for this study were collected from NBI-2013 (USDOT 2013). Although this database contains all 50 states, the author focused on the state of Oklahoma. Oklahoma’s dataset has more than 25,000 bridges, but the study dataset was created according to the following criteria: year built (from 1955 to present), superstructure material (prestressed concrete and steel), and type of design (stringer/multi-beam, girder with a deck type of concrete cast-in-place). These criteria were selected because they characterized the majority of critical bridges in Oklahoma. Also, these bridges are of the interest to prestressed concrete and steel fabricators as well as researchers

(Contreras-Nieto, 2014; Akgül & Frangopol, 2003). As a result of filtering the database with these criteria, the data subset included 8,257 bridges in total, with 3,586 prestressed concrete and 4,671 steel bridges.

Furthermore, there are over 45 items/variables for a bridge record included in the NBI database, which reflects information from bridge inspection reports based on the National Bridge Inspection Standard (FHWA 1995). Some of the items contain information, such as deck rating (Item 58), substructure rating (Item 60), latitude (Item 16) and longitude (Item 17):) that are not of the interest to this study and thus excluded from model developement. In addition, several derived variables were created to capture information of interest. For example, 'age' is the result of subtracting the year built from 2013 since NBI-2013 was used in this study. Table 2-1 provides a list of variables for analysis and their descriptions. The selection of the variables was based on the findings of previous studies (Moomen, 2016; Contreras-Nieto, 2014; Tang et al. (2012), and Veshosky et al.,1994).

Table 2-1. Variable Description

Variable Name	Description	Role	Level
Age	Age of the Bridge in 2013	Input	Interval
Length_Class	Class of Structure Length (Item 49)	Input	Nominal
Item 3	County Code	Input	Nominal
Item 22	Owner	Input	Nominal
Item 26	Functional Classification of Inventory Route	Input	Nominal
Item 29	Average Daily Traffic (ADT)	Input	Interval
Item 31	Design Load	Input	Nominal
Item 42b	Type of Service under Bridge	Input	Nominal
Item 45	Number of Spans in Main Unit	Input	Interval
Item 46	Number of Approach Spans	Input	Interval
Item 48	Length of Maximum Span	Input	Interval
Item 49	Structure Length	Input	Interval

Variable Name	Description	Role	Level
Item 51	Bridge Roadway Width, Curb-to-Curb	Input	Interval
Item 52	Deck Width, Out-to-Out	Input	Interval
Item 68	Deck Geometry	Input	Nominal
Item 69	Underclearances, Vertical and Horizontal	Input	Nominal
Item 71	Waterway Adequacy	Input	Nominal
Item 72	Approach Roadway Alignment	Input	Nominal
Ssrating	Superstructure Rating	Target	Interval

Direct Data Mining Technique

As Berry and Linoff (2011) elaborated, statisticians have invented many data mining techniques, and those techniques are now combined into statistical software such as Statistical Analysis System (SAS). Also, these authors stated that differences exist between data mining and statistics, but they are not fundamental differences. For example, data miners' challenge is to understand anything from voluminous data rather than explain the results of a sample related to the whole dataset. Another example of the difference between statistics and data mining is the time dependency of the data used in data mining. It differs from scientific experiments because usually statisticians consider recurrent observations to be independent observations. Nevertheless, data miners and statisticians look for solutions to similar problems by implementing similar techniques.

Due to the big amount of data available in NBI databases, the author applied the Direct Data Mining Methodology (DDMM) suggested by Berry and Linoff (2011) to fulfill the objective of this research. DDMM is implemented when model development has a specific goal. In this study, the goal was to understand and predict bridge superstructure ratings. This methodology is composed of 10 steps, ranging from translating the business problem into a data mining problem to deploying models and assessing results. In this research, the principal step is to 'assess models' because it allows the author to select the best model according to the comparison of different statistical indicators. Because the target variable is considered continuous as in previous studies such as Bu et al. (2015) and Elbehairy et al. (2006), the Average Square Error (ASE) was used for model comparison.

Addressing Problems with the Data

In order to apply DDMM and develop reliable predictive models, SAS® Enterprise Miner (EM) was used because of its capability of managing big data and a modern set of modeling techniques. Also, EM's interface simplifies the process of creating descriptive and predictive models by allowing the user to build the process flow from a variety of step-specific nodes (toolbar). Data are often "dirty" and are not friendly to data analysis because it includes factors such as categorical variables (e.g. States) with too many levels, quantitative variables with skewed distributions and outliers, and missing values. The following section describes how these problems were addressed.

The data includes some categorical variables that take many values, and most data mining algorithms cannot directly handle this type of data according to Berry and Linoff (2011). One method of solving this issue is to group the values by finding relationships among them and the target variable. As a result, groups with similar target values are created. The author applied the decision trees algorithm in order to group and reduce the number of levels. Then Skewness and Kurtosis values, which are measures of the shape of data, are calculated in combination with histograms in order to assess the normality of the data. The desirable values of Skewness and Kurtosis are close to zero (0) as Meyers (2009) stated. Also, this analysis permits researches to identify and delete outliers. In this case, EM is used to find the most appropriate transformation (logarithm, square, etc) to the variables with issues with normality shape and discard possible outliers. Finally, when records do not have information in all variables, some algorithms have difficulty in handling them. One solution of addressing this issue is to throw them out of the data, but this brings bias into the data and reduces the number of records. Another solution is to impute the missing values by substituting the missing values based on metrics (average, median, or mode) or using an algorithm such as decision trees. The advantage of this approach is that the number of records included in the analysis can be maximized.

Building the Models

The model set was created through a data partition node in which the dataset was divided into training and validation sets. Of the total data, 70% were used for training and 30% for validation. Then the data treatment was addressed in this stage as described in the section above, so that the dataset is ready for creating the models. Figure 2-1 presents the creation of models and the comparison of those models, which are considered the most important components of the research. Although Figure 2-1 shows a partial flow chart of the process to develop the prestressed concrete models, the complete chart includes other steps. Some of the steps are analysis of variable distribution, transformation variable, inputting missing values, and reduction of levels for categorical variables. The model development step is highlighted in Figure 2-1. Moreover, regression, decision trees, and artificial neural network techniques were chosen to create the models.

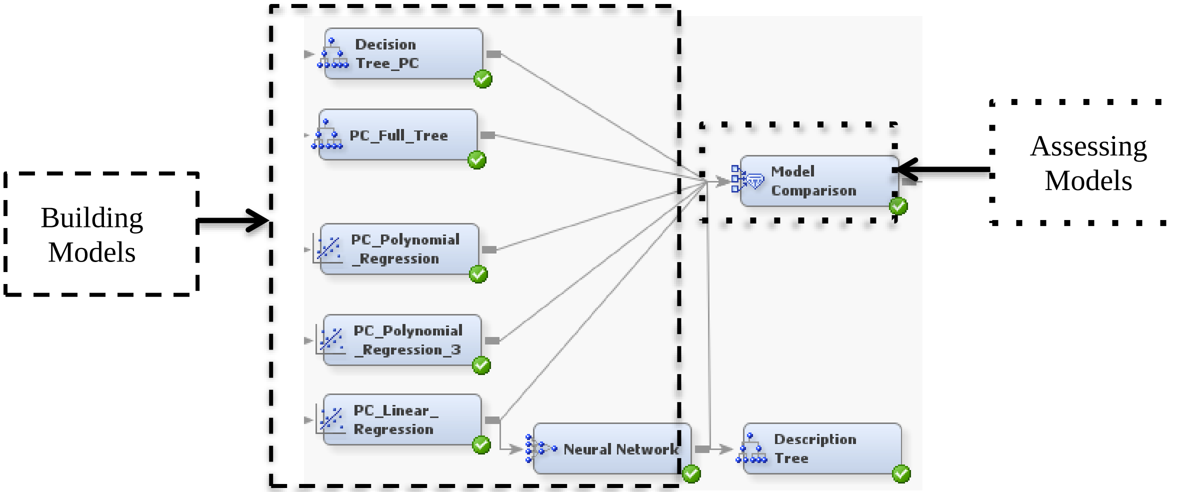

Figure 2-1. Partial Research Flow in SAS® Enterprise Miner

Regression models are expressed mathematically in the form of $Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \ldots + \beta_n X_n$, where Y is the target variable (superstructure rating, see Table 2-1); X are predictor variables (for example, age and owner); and β are coefficients. However, this equation may be modified by increasing the order of the variables' predictors (second or third degree). Three different regression models were used: one linear model and two polynomial models with two and three degrees. In addition, stepwise selection was used as the approach for variable selection (Walsh, 2005). The validation error was used for selecting the best model among three regression models.

Decision tree is another modeling technique and is considered powerful for DDMM (Walsh, 2005). This method partitions the data into 'pure' groups (leaves). Each record is assigned to a group that has similar target values. As a result, the final model is a series of rules for dividing the dataset into groups using the most important variables, which are chosen by the decision tree algorithm (Berry & Linoff, 2011). Those rules are known as English Rules, and anyone can understand them. With this technique, two decision tree models were developed. One is performed interactively, and the other one is the completed decision tree. The author created the interactive model through the selection of variables and their order according to his knowledge and preferences. In the other case, the full decision tree model was generated by EM automatically in accordance with default preferences.

The last technique implemented was artificial neural networks (ANN). Although this approach is recognized as a powerful and flexible technique in supervised prediction analysis, but it is often known as 'black box' because of the difficulty in model interpretation (Berry & Linoff, 2011). An ANN model is formed by the inputs, hidden units, and activation functions. In this study, one ANN model was generated with six hidden units. It should be noted that the predicted values for superstructure rating are values between zero and nine for all models developed.

Finally, EM allows one to compare different models based on the result of a single statistic (misclassification rate, average profit, or average square error) through the Comparison Node (SAS® Enterprise Miner 14.1, 2015). In this study, the objective was to predict a numerical variable; therefore, the average squared error (ASE) is used as the selection criterion (see Equation 1). According to the ASE criterion, the lowest value is the best because the model is less biased than a model with a higher value (Christie, 2011). As a result, the best model was chosen by comparing the ASE values of the six models developed.

$$\mathrm{ASE} = \frac{1}{n}\sum_{i=0}^{n}\left(\tilde{f}(i) - f(i)\right)^2 \qquad (1)$$

where n is the number of observations on the dataset; $\tilde{f}(x)$ is an estimate of the observation i; $f(x)$ is the true value of the observation i.

RESULTS

The results and key findings of developing predictive models of superstructure ratings for both superstructure materials are presented in this section. The selection of the best model for predicting the target variable is shown by superstructure material type. Also, a brief description of the best model is provided.

Basic Descriptive Statistics

In order to understand the data, Table 2-2 and Table 2-4 summarize the basic descriptive statistics of the quantitative variables, Prestressed Concrete and Steel datasets, respectively. Then Table 2-3 and Table 2-5 present the number of levels, number of missing values, and type of variable input of the qualitative variables for the two datasets, PC and S, respectively.

Table 2-2. Prestressed Concrete Dataset Descriptive Statistics

Variable	Description	Missing	N	Minimum	Maximum	Mean	Standard Deviation
Age	Age of the Bridge in 2013	0	3,586	0	56	20.07	12.24
Item_29	Average Daily Traffic (ADT)	74	3,512	1	32,500	2859.86	4,824.84
Item_45	Number of Spans in Main Unit	0	3,586	1	57	2.93	3.43
Item_46	Number of Approach Spans	0	3,586	0	48	0.04	0.86
Item_49	Structure Length	0	3,586	79	17,008	702.15	1,016.61
Item_51	Bridge Roadway Width, Curb-to-Curb	0	3,586	40	543	109.89	45.65
Item_52	Deck Width, Out-to-Out	0	3,586	41	869	119.05	51.08
Ssrating	Superstructure Rating	0	3,586	3	9	7.84	0.75

Table 2-3. Prestressed Concrete Dataset Description of Qualitative Variables

Variable	Description	Type	Number of Levels	Missing
Item_22	Owner	Nominal	6	373
Item_26	Functional Classification of Inventory Route	Nominal	12	0
Item_3	County Code	Nominal	77	0
Item_31	Design Load	Nominal	8	382
Item_42b	Type of Service under Bridge	Nominal	9	0
Item_68	Deck Geometry	Nominal	8	0
Item_69	Underclearances, Vertical and Horizontal	Nominal	8	0
Item_71	Waterway Adequacy	Nominal	8	0
Item_72	Approach Roadway Alignment	Nominal	7	0

Variable	Description	Type	Number of Levels	Missing
Length_Classif	Class of Structure Length (Item 49)	Nominal	4	0

Table 2-4. Steel Dataset Descriptive Statistics

Variable	Description	Missing	N	Minimum	Maximum	Mean	Standard Deviation
Age	Age of the Bridge in 2013	0	4,671	1	60	28.09	17.8
Item_29	Average Daily Traffic (ADT)	154	4,517	10	32,000	2,354.05	5,378.54
Item_45	Number of Spans in Main Unit	0	4,671	1	37	2.11	2.2
Item_46	Number of Approach Spans	0	4,671	0	37	0.15	1.19
Item_49	Structure Length	0	4,671	61	16,538	398.64	712.34
Item_51	Bridge Roadway Width, Curb-to-Curb	0	4,671	43	439	89.96	41.45
Item_52	Deck Width, Out-to-Out	0	4,671	43	2,195	96.99	61.5
Ssrating	Superstructure Rating	0	4,671	0	9	6.26	1.09

Table 2-5 Steel Dataset Description of Qualitative Variables

Variable	Description	Type	Number of Levels	Missing
Item_22	Owner	Nominal	8	220
Item_26	Functional Classification of Inventory Route	Nominal	11	0
Item_3	County Code	Nominal	76	0
Item_31	Design Load	Nominal	8	41
Item_42b	Type of Service under Bridge	Nominal	10	0
Item_68	Deck Geometry	Nominal	9	0
Item_69	Underclearances, Vertical and Horizontal	Nominal	9	0
Item_71	Waterway Adequacy	Nominal	9	0
Item_72	Approach Roadway Alignment	Nominal	7	0
Length_Classif	Class of Structure Length (Item 49)	Nominal	4	0

Reducing Levels of Categorical Variables

Eight different categorical variables were adjusted because of the large number of categories associated with each variable. These variables are: Item 3 (County Code), Item 22 (Owner), Item 26 (Functional Classification of Inventory Route), Item 31 (Design Load), Item 42b (Type of Service under Bridge), Item 68 (Deck Geometry), Item 71 (Waterway Adequacy), and Item 72 (Approach Roadway Alignment). For example, Item 3 had 76 levels (number of counties) before reducing the levels; then by implementing decision trees algorithm seven groups were created. Figure 2-2 presents the results of Steel dataset through the of English Rules for the first two groups developed.

```
*------------------------------------------------------------*
 Node = 6
*------------------------------------------------------------*
if Item_3 IS ONE OF: 39, 149, 141, 137, 43, 145, 129, 75
then
 Tree Node Identifier   = 6
 Number of Observations = 421
 Predicted: Ssrating = 6.9857482185

*------------------------------------------------------------*
 Node = 8
*------------------------------------------------------------*
if Item_3 IS ONE OF: 109, 71, 7, 51, 151, 131, 17, 115, 61, 19, 107, 81, 125, 89, 97, 147, 133, 41, 121, 85 or MISSING
then
 Tree Node Identifier   = 8
 Number of Observations = 1029
 Predicted: Ssrating = 6.0796890185
```

Figure 2-2. Level Reduction for Steel Bridges Item 3 - County Code

Variables with High Values of Skewness and Kurtosis

Based on the Skewness and Kurtosis values, seven quantitative variables need to be transformed in order to reduce their values, and so improve the shape of their distributions. Thus, the data would be as close as possible to a normal distribution that is preferable for the required analysis in this study. Table 2-6 summarizes the Skewness and Kurtosis values for the qualitative variables of the prestressed concrete bridge dataset. Although some transformed variables did improve these values, others did not do much for lowering the values close to zero (0). For example, Item 29 (ADT) went from 2.79 and 9.30 to 0.11 and -1.33 for Skewness and Kurtosis, respectively, after applying the logarithm transformation. Nevertheless, Item 46 (Number of Approach Spans) went from 15.25 and 277.12 to 10.90 and 128.39 for Skewness and Kurtosis, respectively, after applying the logarithm transformation. It is important to mention that EM recommends which transformation is the best for each variable according to the data distribution, but it is up to the researcher to choose the transformation to be used in the analysis.

Table 2-6. Skewness and Kurtosis Values of PC Bridge Quantitative Variables

Method	Variable Name	Description	Formula	Skewness	Kurtosis
Original	Item_29	Average Daily Traffic (ADT)		2.79	9.30
Original	Item_45	Number of Spans in Main Unit		6.80	67.30
Original	Item_46	Number of Approach Spans		15.25	277.12
Original	Item_48	Length of Maximum Span		0.51	4.85
Original	Item_49	Structure Length		7.40	75.56
Original	Item_51	Bridge Roadway Width, Curb-to-Curb		2.98	13.91
Original	Item_52	Deck Width, Out-to-Out		3.67	26.07
Computed	LOG_Item_29	Log transformation of Average Daily Traffic (ADT)	log(Item_29 + 1)	0.11	-1.33
Computed	LOG_Item_45	Log transformation of Number of Spans in Main Unit	log(Item_45 + 1)	1.24	3.01
Computed	LOG_Item_46	Log transformation of Number of Approach Spans	log(Item_46 + 1)	10.90	128.39
Computed	LOG_Item_49	Log transformation of Structure Length	log(Item_49 + 1)	1.21	2.41
Computed	LOG_Item_51	Log transformation of Bridge Roadway Width, Curb-to-Curb	log(Item_51 + 1)	1.25	2.11
Computed	LOG_Item_52	Log transformation of Deck Width, Out-to-Out	log(Item_52 + 1)	1.41	2.55
Computed	SQRT_Item_48	Square-root transformation of Length of Maximum Span	Sqrt(Item_48 + 1)	-0.19	0.72

In addition to the Kurtosis and Skewness values, the histograms of the distribution for the variables are studied as well. As shown in Figure 2-3, the shape of the distribution of Item 49 (ADT) after the logarithm transformation has improved compared to the distribution before transformation.

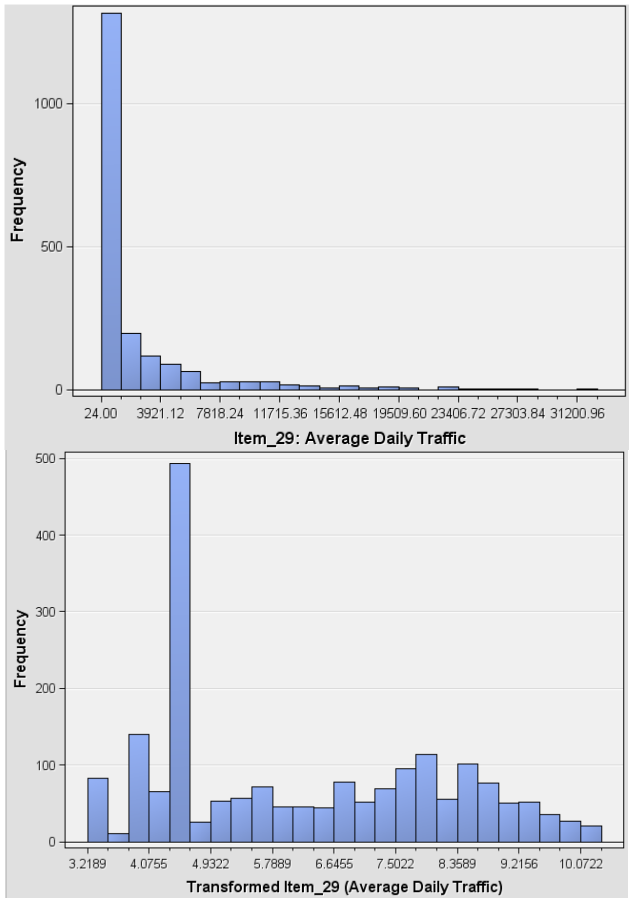

Figure 2-3. Before and After Transformation of Item 29: ADT for PC Bridges

Impute Values

It was determined that just three variables had missing values for both superstructure materials, prestressed concrete (PC) and steel (S). Table 2-7 presents the number of missing values for these three variables categorized by the superstructure material. Decision trees were implemented to impute these missing values. As a result the sample size was not reduced.

Table 2-7. Missing Values in the Training Datasets

Variable Name	Description	Impute Method	Type of Variable	Steel - Number of Missing (Train)	PC - Number of Missing (Train)
Item_22	Owner	TREE	NOMINAL	166	264
Item_31	Design Load	TREE	NOMINAL	28	259
LOG_Item_29	Log transformation of Average Daily Traffic (ADT)	TREE	INTERVAL	117	56

Prestressed Concrete Superstructure

According to the comparison of ASE statistics in the validation dataset in Table 2-8, Neural2 (ANN) is the model with the best performance with the least ASE value of 0.32. The model with the second lowest ASE value is Tree9, which is the decision tree developed interactively. Finally, the models with the highest ASE values are the regression models with values over 0.35. This means that this technique is the least preferable method for predicting superstructure rating for prestressed concrete bridges.

Table 2-8. SAS® Enterprise Miner – PC Comparison Node Results

Model	Model Description	Validation Average Squared Error	Training Average Squared Error
Neural2	Neural Network[1]	0.32	0.29
Tree9	Decision Tree_PC[2]	0.33	0.27
Tree15	PC_Full_Tree[3]	0.35	0.32
Reg3	PC_Polynomial_Regression[4]	0.36	0.33
Reg4	PC_Polynomial_Regression_3[5]	0.36	0.33
Reg2	PC_Linear_Regression[6]	0.37	0.33

Note:
[1] Neural Network – Multilayer perceptron with 6 hidden units
[2] Decision Tree model created by the author using logworth values
[3] Decision Tree model created by SAS EM automatically based on properties defined
[4] Polynomial regression model – Stepwise approach – Quadratic model
[5] Polynomial regression model – Stepwise approach – Cubic model
[6] Linear regression model – Stepwise approach

Another approach to determine how well a model is working on the dataset is by comparing the score ranking matrix, which presents the mean predicted values and the mean target values through the data. For example, in Figure 2-4 superior chart, while the mean target value of the top 40% of the data is 8.07, the mean predicted value is 7.97. The chart with smaller gaps between the two lines (mean predicted value and mean target value) is the model that performs better at predicting the values of the target variable. Comparing the neural network model (Neural2) and the linear regression model, Figure 2-4 and Figure 2-5, respectively, it is seen that Neural2 model better predicted the mean superstructure rating for both datasets, training and validation (see Figure 2-4). In contrast, the linear regression model low performance is confirmed by looking at the second half of Figure 2-5, where the mean predicted and mean target lines deviate from each other as those shown in Figure 2-4.

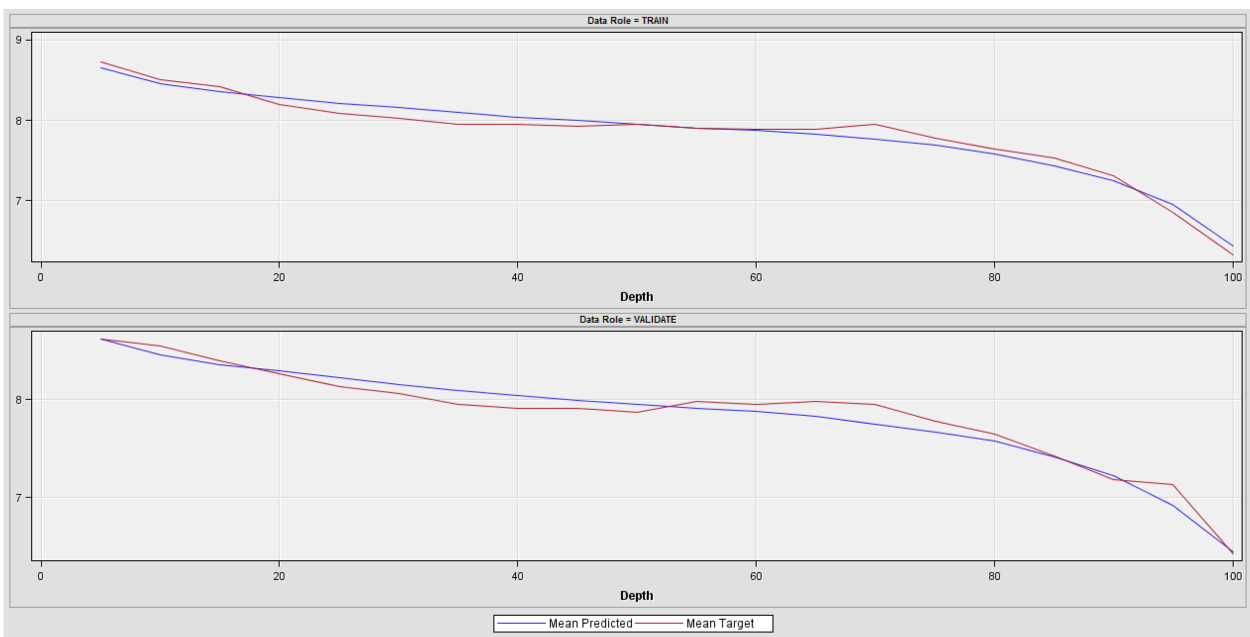

Figure 2-4. Score Ranking Matrix - PC Neural Network Model (Model Neural2)

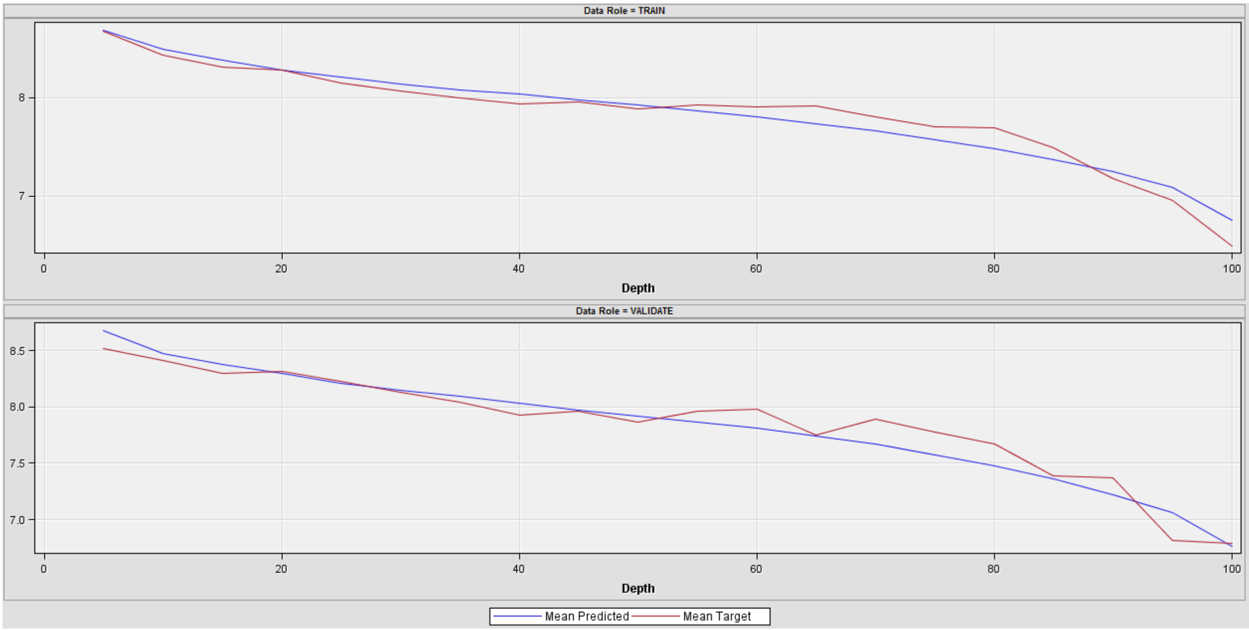

Figure 2-5. Score Ranking Matrix - PC Linear Regression Model (Model Reg2)

On the other hand, due to the complexity of explaining an ANN model (best model Neural2), a decision tree node was used in order to obtain the relevancy of each predictor in the Neural2 model. Table 2-9 shows the variables that are considered important for the model. As previous studies have found, age is significant in predicting superstructure ratings. Age is the most critical predictor in this study since its importance coefficient is 1.0 in the validation dataset. Also, Item 49 (structure length) and Item 22 (owner) are the second and the third most important variables, respectively. It should be noted that the impact of other variables is weighted in relation to the most important variable in EM. Because of this, Item 49 has a value of 0.35, which means this predictor received an importance of 35% compared to age. Similarly, Item 22 has 29% importance for predicting superstructure ratings.

Table 2-9. Variable Importance Neural2 Model for PC

Variable Name	Label	Description	Training Importance	Validation Importance
Age		Bridge age in 2013	1.00	1.00
LOG_Item_49	Transformed Item_49	Log transformation of Structure Length	0.31	0.35

Variable Name	Label	Description	Training Importance	Validation Importance
REP_IMP_Item_22	Replacement: Imputed Item_22	Owner	0.22	0.29

Steel Superstructure

For steel superstructure rating prediction models, model Tree11 was selected as the best model with an ASE of 0.67 in the validation dataset shown in Table 2-10. Model Tree11 is the decision tree model developed automatically by EM. The ANN model (Neural) ranked second according to performance with an ASE value of 0.68.

Table 2-10. SAS® Enterprise Miner – Steel Comparison Node Results

Model	Model Description	Validation Average Squared Error	Train Average Squared Error
Tree11	Steel Full Tree[1]	0.67	0.62
Neural	Neural Network[2]	0.68	0.60
Reg8	Steel Polynomial Regression[3]	0.74	0.70
Reg9	Steel Polynomial Regression[4]	0.74	0.69
Tree10	Steel Decision Tree[5]	0.76	0.72
Reg7	Steel Linear Regression[6]	0.81	0.75

Note:
[1] Decision Tree model created by SAS EM automatically based on properties defined
[2] Neural Network – Multilayer perceptron with 6 hidden units
[3] Polynomial regression model – Stepwise approach – Cubic model
[4] Polynomial regression model – Stepwise approach – Quadratic model
[5] Decision Tree model created by the author using logworth values
[6] Linear regression model – Stepwise approach

A similar score ranking matrix comparison was performed on steel models and Figure 2-6 and Figure 2-7 show the full tree (best model) and the linear regression models performed to predict the mean superstructure rating for both training and validation datasets. The mean predicted and mean target lines of the full tree model matched perfectly in the training dataset, and those lines are very close in the validation dataset chart (see Figure 2-6). In contrast, those lines are slightly further away from each other for the linear regression model as can be observed in Figure 2-7.

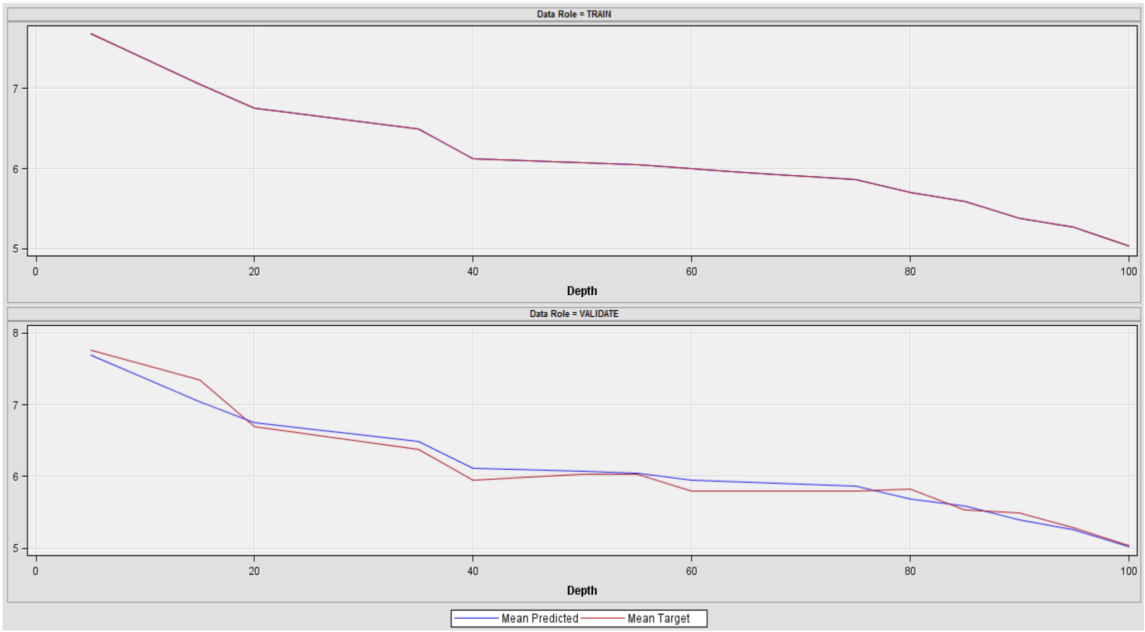

Figure 2-6. Score Ranking Matrix – Steel Full Tree Model (Model Tree11)

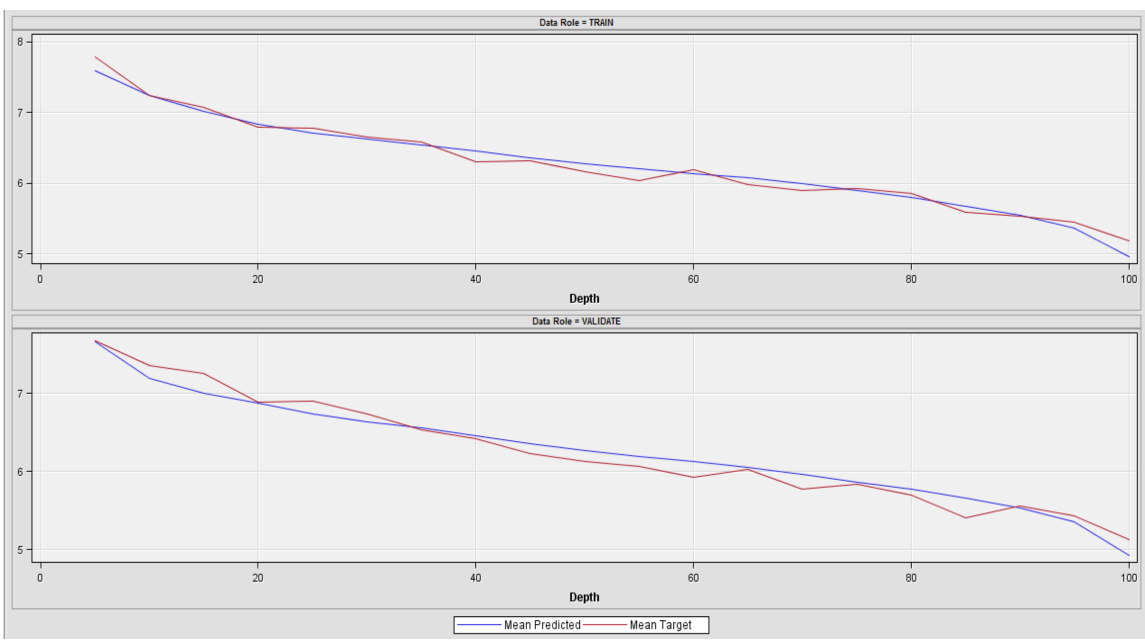

Figure 2-7. Score Ranking Matrix - Steel Linear Regression Model (Model Reg7)

Model Tree 11 had more than 20 leaves. Therefore, the author used English Rules shown in Figure 2-8 to explain the model in a more straightforward way. For example, for bridges located in county 109, with an ADT less than 1,700 and ages less than 8.5 years, the predicted mean superstructure rating was 7.6 - close to 8 which indicates a very good condition. In contrast, bridges older than 8.5 years but with the same characteristics described in the previous example obtained a predicted mean value of 6.2, which indicates a satisfactory condition.

English Rules

```
*------------------------------------------------------------*
 Node = 8
*------------------------------------------------------------*
if Item_3 IS ONE OF: 109, 143, 71, 73, 39, 149, 141, 3, 53, 43, 145, 37, 103, 47, 129, 17, 75, 59, 87 or MISSING
AND Item_29 < 1700
AND Age < 8.5
then
 Tree Node Identifier   = 8
 Number of Observations = 340
 Predicted: Ssrating = 7.6176470588
```

Figure 2-8. SAS® Enterprise Miner – Partial English Rules

Also, English Rules shows that steel bridges with deck width less than 7.8 meters (it should be noted that the lengths are expressed in decimeters in the NBI databases, 78.5 dm), design load different than MS 18/HS 20, age older than 11 years, and located in counties 73 or 119 were more likely to have the lowest mean superstructure ratings with a predicted value of 5.0 – Fair Condition, as shown in Figure 2-9. On the other hand, the highest predicted average rating (9.0 – Excellent Condition) was shown among bridges that are younger than 8 years old, with ADT higher than 1,700, and located in counties such as 109 or 143, which is presented in Figure 2-10.

```
if Item_52 < 78.5 or MISSING
AND Item_31 IS ONE OF: OTHER OR UNKNOWN, M 18 / H 20, M 9 / H 10, M 13.5 / H 15 or MISSING
AND Item_3 IS ONE OF: 73, 7, 119, 69, 53, 93, 47, 59, 95, 105, 139
AND Age >= 11.5 or MISSING
then
 Tree Node Identifier   = 98
 Number of Observations = 161
 Predicted: Ssrating = 5.0248447205
```

Figure 2-9. English Rules - Lowest Average Predicted Rating

```
if Item_3 IS ONE OF: 109, 143, 71, 73, 39, 149, 141, 3, 53, 43, 145, 37, 103, 47, 129, 17, 75, 59, 87 or MISSING
AND Item_29 >= 1700 or MISSING
AND Age < 8.5
then
 Tree Node Identifier   = 9
 Number of Observations = 23
 Predicted: Ssrating = 8.6086956522
```

Figure 2-10. English Rules - Highest Average Predicted Rating

Finally, variable importance is shown in Table 2-11. Once again, age is the most significant variable for superstructure rating prediction. Furthermore, Item 3 (County) is the second most influential predictor with an importance value of 0.67 with respect to age. Before this study, the County Code (which means the location of a bridge in a county) has never been identified as a significant predictor for bridge superstructure rating. Item 42b (Type of Service under the Bridge) and Item 31 (Design Load) rank third and fourth with importance values of 0.25 and 0.15, respectively. The other four variables have some importance in the model, but their values are less than 0.13. These variables are Item 29 (ADT), Item 52 (Deck Width), Item 49 (Structure Length), and Item 68 (Deck Geometry).

Table 2-11. Variable Importance - Full Decision Tree for Steel Bridges

Variable Name	Description	Training Importance	Validation Importance
Age	Age of the Bridge in 2013	1.00	1.00
Item_3	County Code	0.83	0.67
Item_42b	Type of Service under Bridge	0.22	0.25
Item_31	Design Load	0.09	0.15
Item_29	Average Daily Traffic (ADT)	0.15	0.13
Item_52	Deck Width, Out-to-Out	0.15	0.11
Item_49	Structure Length	0.13	0.05
Item_68	Deck Geometry	0.06	0.04

Revised Models for Steel Bridges

The best model for predicting superstructure ratings of prestressed concrete bridges is an ANN model (Neural2) with an ASE value of 0.32 in the validation dataset (see Table 2-8). In contrast, a decision tree model (Tree11) is the best model for predicting superstructure ratings of steel bridges. The ASE value for Tree11 is 0.68 in the validation dataset (see Table 2-10). If one compares the ASE values of the two best models (steel and prestressed concrete), a substantial variation between the values is perceived. This shows that the prestressed concrete model is less biased than the steel model because the prestressed concrete model has a lower value of ASE.

Consequently, the steel model lacks flexibility which may lead to under-fitting the predicted ratings (Walsh 2005). Due to this, the author decided to find a way to reduce the ASE value of steel models.

According to Connor, Dexter, & Mahmoud (2005), improvements on material, design (fatigue design provision), production, and inspection (fabrication and in-service) for steel bridges occurred in the 1970s. As a result of those improvements and special precautions for fracture-critical members, bridges designed and built since 1980 are less susceptible to face fatigue and fracture. In addition, because of the impact of corrosion in steel bridges, coating methods were refined, which may positively affect the condition of steel bridges. Given the improvements in steel bridges, the authors decided to divide the steel bridges into three more homogenous groups based on the following age ranges for further study:

1. Steel bridges built before 1975
2. Steel bridges designed between 1976 to 1985
3. Steel bridges designed after 1985

In recent years, research on steel bridges has progressed. In fact, the introduction of High Performance Steel (HPS) for U.S.A. bridges started in 1996 (Wilson, 2000). The study of this new material started in 1992 when the FHWA, the American Iron and Steel Institute (AISI), and the U.S. Navy jointly developed a superior steel for bridges. As a result of this partnership, three new steel grades were developed: HPS-50W, HPS-70W, and HPS-100W. The letter ‘W’ stands for weathering capability, which means that these steel grades work under normal atmospheric circumstances with no painting-protection layer. The ASTM A709/A709M – 16a is the specification of carbon and high-strength low alloy steel that covers the grades mentioned above. Because of the HPS development another group of steel bridges can be differentiated from the previous three groups stated by Wilson (2000), which corresponds to bridges built since 2000 when HPS began to be frequently used as an outcome of the Transportation Equity Act for the 21st Century (TEA-21).

Based on the findings mentioned above and with the objective of reducing the variability found in the steel bridges dataset, this dataset was divided into four groups according to the age built. They are summarized in Table 2-12. Then the same methodology implemented for developing predictive models of superstructure ratings was applied to the four sub datasets in order to produce the respective models.

Table 2-12. Groups of Steel Bridges

Group	Year Built Range	Age Range in 2013	Number of Bridges	Deficient Bridges	Average Age	Average Superstructure Rating
1	1955 – 1975	>37	1,712	93	48.7	6.0
2	1975 – 1985	28 – 37	434	11	31.5	5.9
3	1986 – 1999	14 – 27	1,226	4	20.3	6.0
4	2000 – 2013	<14	1,299	1	7.1	6.9

Sub Dataset Models

According to the methodology established in this study, data issues within the sub datasets have to be addressed. Therefore, reducing levels of categorical variables, transformation of variables with high values of skewness and kurtosis, and imputing values of missing values were performed in each group dataset. Then these four sub-datasets were randomly divided into training and validation sets, 70% and 30% respectively. Finally, predicting models were created by implementing regression, regression trees, and neural networks techniques. Table 2-13, Table 2-14, Table 2-15, and Table 2-16 present the comparisons of the best six models for each steel group dataset.

Table 2-13. SAS® Enterprise Miner – Steel Group 1 – Comparison Node Results

Model	Model Description	Validation Average Squared Error	Train Average Squared Error
Neural	Neural Network[1]	0.90	0.74
Tree9	Full Tree[2]	0.92	0.63

Model	Model Description	Validation Average Squared Error	Train Average Squared Error
Reg3	Polynomial Regression[3]	0.92	0.78
Reg4	Polynomial Regression[4]	0.92	0.78
Reg2	Linear Regression[5]	0.93	0.78
Tree8	Decision Tree[6]	1.08	0.89

Note:
[1] Neural Network – Multilayer perceptron with 6 hidden units
[2] Decision Tree model created by SAS EM automatically based on properties defined
[3] Polynomial regression model – Stepwise approach – Cubic model
[4] Polynomial regression model – Stepwise approach – Quadratic model
[5] Linear regression model – Stepwise approach
[6] Decision Tree model created by the author using logworth values

Table 2-14. SAS® Enterprise Miner – Steel Group 2 – Comparison Node Results

Model	Model Description	Validation Average Squared Error	Train Average Squared Error
Tree9	Full Tree[1]	0.45	0.29
Neural	Neural Network[2]	0.46	0.33
Tree8	Decision Tree[3]	0.46	0.28
Reg4	Polynomial Regression[4]	0.47	0.35
Reg2	Polynomial Regression[5]	0.47	0.35
Reg2	Linear Regression [6]	0.48	0.35

Note:
[1] Decision Tree model created by SAS EM automatically based on properties defined
[2] Neural Network – Multilayer perceptron with 6 hidden units
[3] Decision Tree model created by the author using logworth values
[4] Polynomial regression model – Stepwise approach – Cubic model
[5] Polynomial regression model – Stepwise approach – Quadratic model
[6] Linear regression model – Stepwise approach

Table 2-15. SAS® Enterprise Miner – Steel Group 3 – Comparison Node Results

Model	Model Description	Validation Average Squared Error	Train Average Squared Error
Tree29	Full Tree[1]	0.39	0.33
Reg15	Polynomial Regression[2]	0.44	0.37
Neural4	Neural Network[3]	0.44	0.27
Reg14	Polynomial Regression [4]	0.44	0.37
Tree28	Decision Tree[5]	0.44	0.35
Reg13	Linear Regression [6]	0.44	0.37

Note:

[1] Decision Tree model created by SAS EM automatically based on properties defined
[2] Polynomial regression model – Stepwise approach – Cubic model
[3] Neural Network – Multilayer perceptron with 6 hidden units
[4] Polynomial regression model – Stepwise approach – Quadratic model
[5] Decision Tree model created by the author using logworth values
[6] Linear regression model – Stepwise approach

Table 2-16. SAS® Enterprise Miner – Steel Group 4 – Comparison Node Results

Model	Model Description	Validation Average Squared Error	Train Average Squared Error
Tree9	Full Tree[1]	0.57	0.48
Neural	Neural Network[2]	0.61	0.48
Reg3	Linear Regression [3]	0.65	0.53
Reg4	Polynomial Regression[4]	0.66	0.53
Reg2	Polynomial Regression[5]	0.66	0.53
Tree8	Decision Tree[6]	0.84	0.55

Note:
[1] Decision Tree model created by SAS EM automatically based on properties defined
[2] Neural Network – Multilayer perceptron with 6 hidden units
[3] Linear regression model – Stepwise approach
[4] Polynomial regression model – Stepwise approach – Cubic model
[5] Polynomial regression model – Stepwise approach – Quadratic model
[6] Decision Tree model created by the author using logworth values

After analyzing the results, some variations from the initial Steel dataset were noticed in both the best models and ASEs. First of all, the ASE of each group improved significantly except in group 1. The validation ASE for the complete dataset was 0.67 (see Table 2-10), and it was reduced to 0.45, 0.39, and 0.57 for groups 2 (Table 2-14), 3 (Table 2-15), and 4 (Table 2-16), respectively. This indicates that creating bridge groups based on year built according to improvement on steel as a material, construction specifications, and bridge design reduced some variability within the group. In contrast, the ASE of group 1, which contains older bridges (>37 years old), is 0.90 as Table 2-13 shows; it went higher than the initial ASE of 0.67. This might be due to the fact that Group 1 contains bridges that are too dissimilar and its big range of age, so it is not possible to develop a good model to predict superstructure ratings. A solution for this issue may be to divide this group into smaller and homogenous groups. Also, this group contains the oldest bridges and some of them may have received major maintenance/repairs. As a result, those bridges are in much better condition than other bridges within the same group. In addition, group 1 contains the most deficient superstructure steel bridges (93 deficient bridges – Table 2-12), which may also influence the development of a robust predicting model.

Another finding based on the division of steel bridges and the development of the new predicting models for Groups 2 to 4 was that the models with best performance are decision trees (see Table 2-14, Table 2-15, and Table 2-16). However, these models are complex to visualize because they are formed of many branches. Their English Rules are long and tedious to understand in some cases, but the objective of this study is to predict the superstructure rating rather to characterize the data. For Group 1, a neural network model slightly outperformed the decision tree (Table 2-13); however, the overall model performance for Group 1 is not very desirable due to its relatively larger ASE.

In order to understand the predicting models for the four steel groups, a variable importance table was created for each of them. Table 2-17, Table 2-18, Table 2-19, and Table 2-20 present the input variables that formed the best models as well as the percentage of importance of those variables relative to the most relevant input variable. For example, Item 3 (County) is the most important parameter to predict the superstructure ratings for steel bridges in group 1. Then Item 26 (Functional Classification Inventory Route) receives an importance value of 57% when compared to Item 3, and it is the second most influential predictor (see Table 2-1).

Table 2-17. Variable Importance – Group 1 – Neural Network Model

Variable Name	Description	Training Importance	Validation Importance
Item_3	County Code	1.00	1.00
Item_26	Functional Classification Inventory Route	0.57	0.34
Item_42b	Type of Service under Bridge	0.30	0.39

Table 2-18. Variable Importance – Group 2 – Decision Tree Model

Variable Name	Description	Training Importance	Validation Importance
Item_3	County Code	1.00	1.00
Item_51	Bridge Roadway Width, Curb-to-Curb	0.31	0.34
Item_26	Functional Classification Inventory Route	0.26	0.25

Table 2-19. Variable Importance – Group 3 – Decision Tree Model

Variable Name	Description	Training Importance	Validation Importance
Item_3	County Code	1.00	1.00
Item_22	Owner	0.24	0.38
Item_48	Length of Maximum Span	0.21	0.11
Age	Age of the Bridge in 2013	0.12	0.20
Item_49	Structure Length	0.10	0.11
Item_52	Deck Width, Curb-to-Curb	0.08	0.04

Table 2-20. Variable Importance – Group 4 – Decision Tree Model

Variable Name	Description	Training Importance	Validation Importance
Age	Age of the Bridge in 2013	1.00	1.00
Item_3	County Code	0.83	0.72
Item_49	Structure Length	0.33	0.39
Item_51	Bridge Roadway Width, Curb-to-Curb	0.16	0.09
Item_31	Design Load	0.13	0.16
Item_29	Average Daily Traffic (ADT)	0.06	0.07
Item_48	Length of Maximum Span	0.04	0.03

It is observed that age is no longer the most significant variable for superstructure prediction among the bridges built before 2000 (Table 2-17, Table 2-18, and Table 2-19). Instead of this parameter, the location of the bridge (Item 3 – County) was selected as the main input parameter to predict superstructure rating in this period. Item 3 was also identified as an explanatory data item by Tang et al. (2012). It is also equivalent to ‘Region’, a variable used by

Morcous et al. (2002). Although they developed ANOVA tests to evaluate the association between deck bridge deterioration rate and qualitative variables, it was found that there is a significant difference in the mean of deck bridge deterioration rates among the Region classes. Furthermore, the importance of the predictor that followed Item 3 (County Code) is less than 60% for all three groups of bridges built before 2000. This importance of the location indicated in the predicting model could signify that the difference in environmental conditions may exist throughout Oklahoma, which deserves future study.

In addition, Item 26 (Functional Classification Inventory Route) became a relevant factor to predict superstructure rating for bridges built before 1986 (see Table 2-17 and Table 2-18). Functional Classification Inventory Route categorizes bridges according to the type of road on which they served. Some of the codes that this parameter has are: rural principal arterial – interstate; urban minor arterial; and urban local. Morcous et al. (2002) also found that there is a significant difference among means of the different levels of Highway Class variable. This corroborates that Item 26 is a significant factor for bridge condition rating prediction. Nevertheless, Item 26 is more important for group 1 than for group 2 as shown by the percentage of importance, 57% and 26%, respectively.

In contrast to results of previous studies (Tang et al., 2012; Veshosky et al., 1994; Kim and Yoon, 2010) in which either traffic (Item 29 – Average Daily Traffic) or design load (Item 31) were identified as relevant predictors of bridge deterioration, predicting models for group 1, group 2, and group 3 do not contain any of these variables (Table 2-17, Table 2-18, and Table 2-19). Yet, it is different for group 4 whose model includes both input variables, Item 31 and Item 29 as it is shown in Table 2-20. Nevertheless, these parameters appear at the bottom of the table because of their low percentage of importance, 13% and 6%, respectively.

Finally, the most similar model to the initial decision tree model for the complete steel dataset (Table 2-11) is the model of group 4 (Table 2-20). Although model of group 4 does not contain all predictors of the previous models, it is formed by five out of the eight parameters in the initial model. Also, age and Item 3 (County) are two of the most significant variables to predict superstructure ratings, with percentages of importance 100% and 83%, respectively.

VALIDATION

After developing different predicting models and selecting the best model for each material, the validation of the model performance was completed. The author adopted the validation method in which the predicted outcome is compared to the real result. It means a comparison between the superstructure ratings obtained (outcome of the model by implementing the models) and the 'real' superstructure rating (NBI database). Both 2013 and 2014 NBI databases were used for validation.

In order to determine the precision and accuracy of the models for both databases, observed vs. predicted scatter plots were developed as Piñeiro et al. (2008) demonstrated (observed values should be placed in y-axis, and the predicted values should be placed in x-axis). Then fitting trendlines (regression model) to the data, the slope (accuracy) and the precision (R^2) were determined. In addition, histograms of the residual values were assembled to visualize the distribution of difference between these values.

Prestressed Concrete

Table 2-21 presents the results of four prestressed concrete bridges by implementing the neural network (Neural2). The whole prestressed concrete dataset was scored and the results were the source to create Figure 2-11 with the trendline whose equation and R^2 are in the figure as well. From the R^2 (0.39) and the slope (0.85), it can be stated that 39% of the linear variation in the observed values can be explained by the variation in the predicted values, and the model is 85% accurate. Moreover, the R^2 (0.38) and the slope (0.86) of scoring NBI-2014 (see Table 2-22) are almost identical to NBI-2013. NBI-2014 scatter plot can be seen in Appendix A. Nevertheless, these models have a limitation with respect to predicting superstructure ratings lower than six. The reason is that the number of bridges with low superstructure ratings (<6) are just 70 of 3,586, which is only 2% of the dataset.

Table 2-21. Scoring NBI-2013 - PC Model

Key[1]	Age[2]	ADT[3]	Observed[4]	Predicted[5]	Round Predicted[6]	Residuals[7]
Ok-198350000000000	33	5800	6	6.32	6	0
Ok-190710000000000	38	100	6	6.28	6	0
Ok-181350000000000	42	16400	6	6.03	6	0
Ok-158650000000000	50	9450	7	5.72	6	1

Note:
[1] Code that identifies the bridge in the dataset
[2] Age of the bridge in 2013
[3] Average daily traffic in NBI-2013
[4] Superstructure rating in NBI-2013
[5] Predicted superstructure rating obtained by implementing the model
[6] Rounded predicted superstructure rating
[7] Residuals: observed value - predicted value

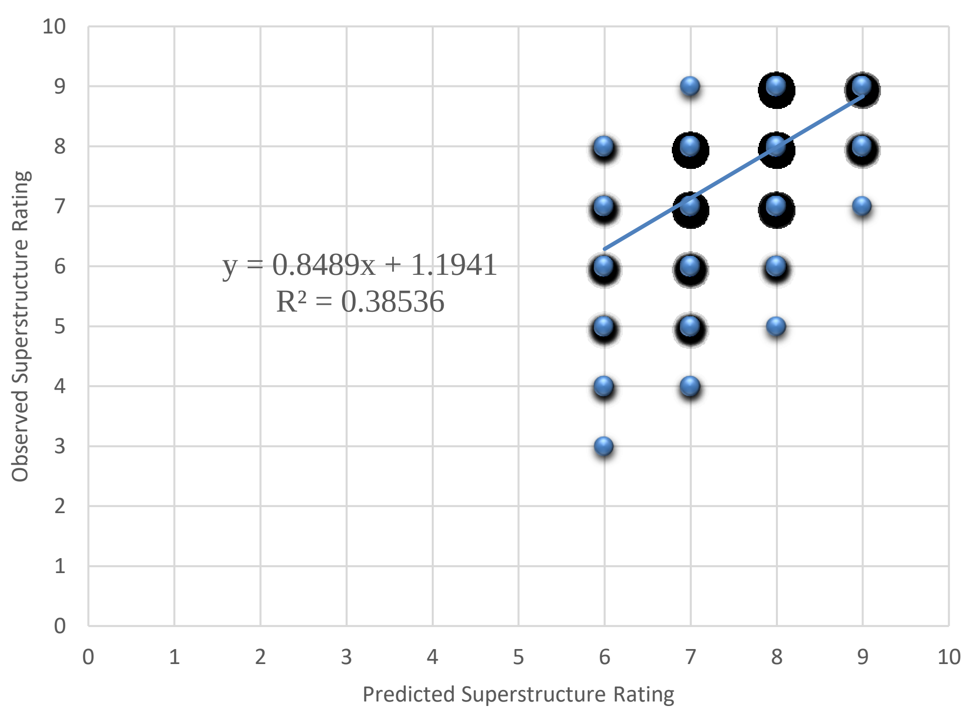

Figure 2-11. Superstructure Rating Residuals - PC NBI-2013

Table 2-22. Validation – Coefficients of Trendlines

Database	Slope	R^2
NBI-2013	0.85	0.39
NBI-2014	0.86	0.38

Also, by looking at the histogram of the residuals (Figure 2-12), 2,604 superstructure ratings were predicted correctly in NBI-2013 while 2,557 were predicted correctly in NBI-2014. The variation of one unit in the superstructure rating could be the distinction between a deficient bridge and a non-deficient bridge; however, it can also be an acceptable error in predicting superstructure ratings if the subjectivity of bridge inspectors is taken into consideration (Veshosky et al., 1994). Thus, just 2.2% (NBI-2013) and 2.3% (NBI-2014) of the predicted values have a substantial difference (>1) from the observed values; it means that almost 98% of the prestressed concrete superstructure ratings are predicted accurately.

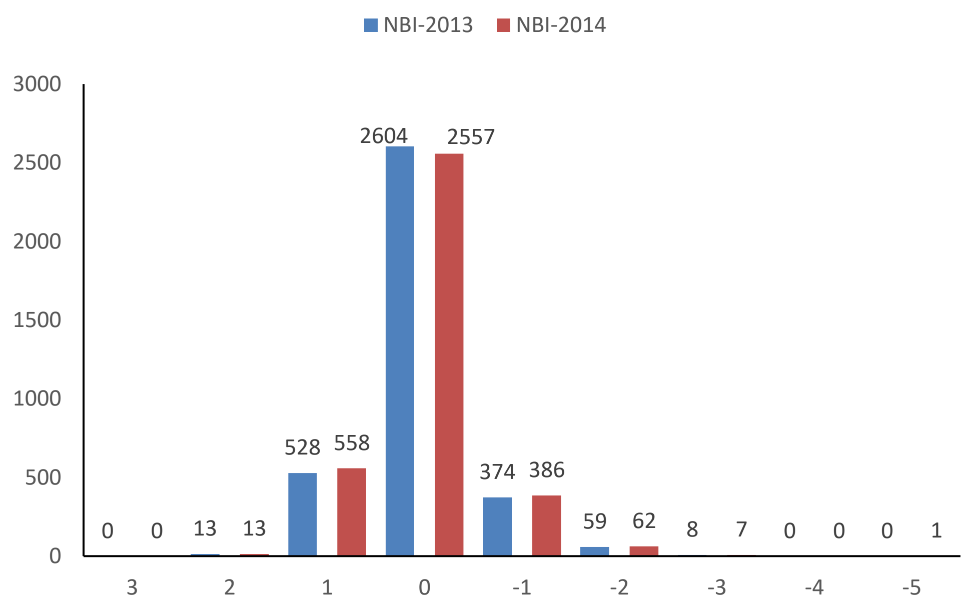

Figure 2-12. Frequency of Superstructure Rating Residuals

Steel

Similar to the procedures performed on prestressed concrete bridges, the steel sub-datasets were scored with the best model for each steel group. Thus, the predicted superstructure ratings were obtained, and then a comparison between observed values and predicted values was performed. Because of the considerable size of the tables, the author decided to present the shortest table, which corresponds to Group 2, in Appendix A. The scatter plots with trendlines and histograms (Superstructure Rating Residuals) for each group were plotted. Figure 2-13 and Figure 2-14 show the results for Group 4 only; the rest of scatter plots and histograms are in Appendix A as well.

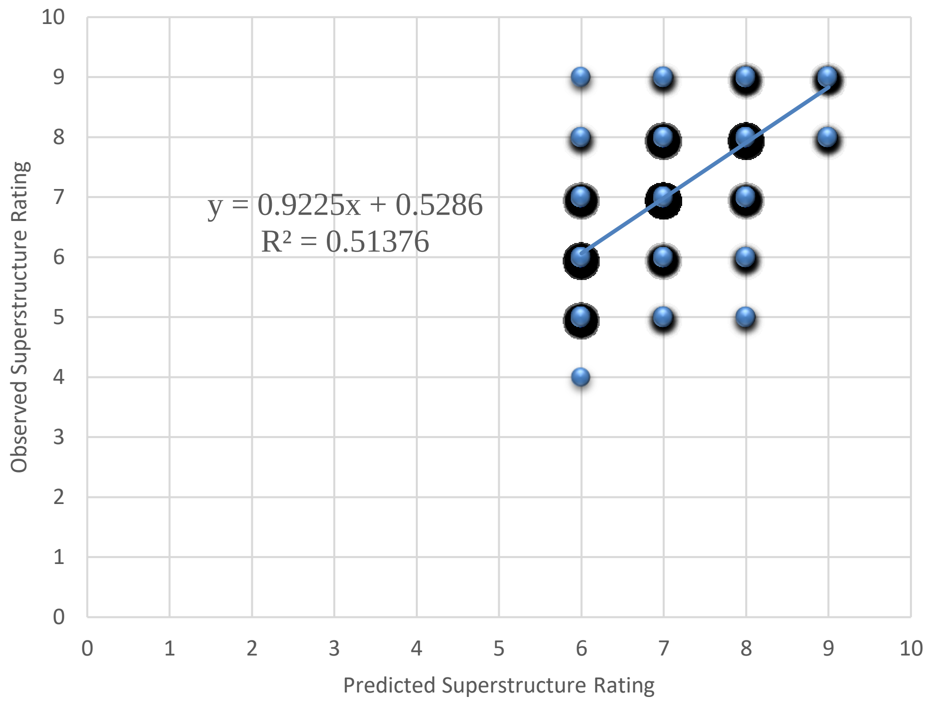

Figure 2-13. Superstructure Rating Residuals - Steel Group 4 NBI-2013

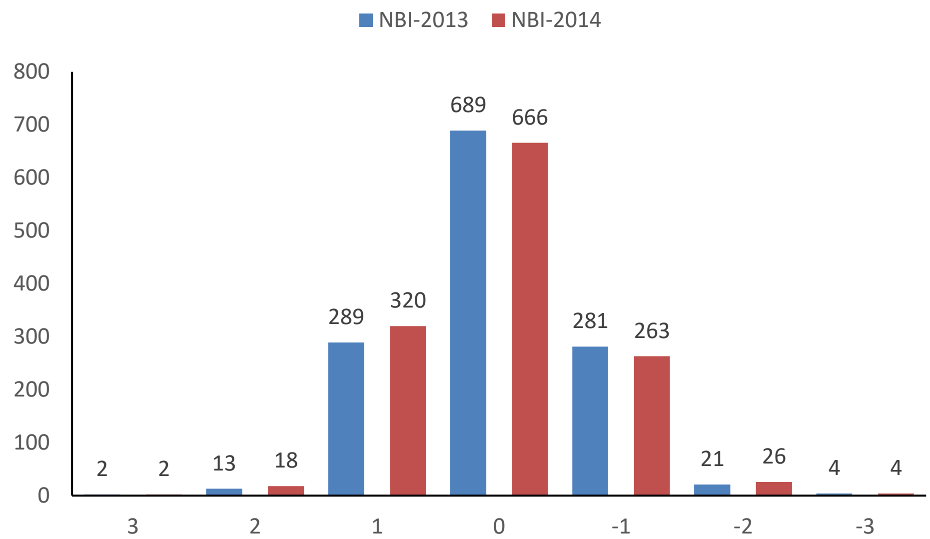

Figure 2-14. Frequency of Superstructure Rating Residuals - Steel Group 4

Table 2-23. Validation Summary - Coefficients of Parameters for Steel Groups

	NBI-2013					NBI-2014		
Group	**Slope**	**R^2**	**Lowest Predicted Rating**	**Number of Deficient Bridges**	**Percentage of Deficient Bridges**	**Slope**	**R^2**	**Lowest Predicted Rating**
1	0.83	0.15	5	93	5.4	0.83	0.15	5
2	0.84	0.45	5	11	2.5	0.84	0.44	5
3	0.77	0.45	5	7	0.3	0.75	0.42	5
4	0.92	0.51	6	1	0.1	0.81	0.44	6

Also, Table 2-23 presents the summary of the results of all groups. It is noticed that the accuracy of the models goes from 77% (Group 3) to 92 % (Group 4). While just 15% of the linear variation in the observed values is explained by the variation in the predicted values in group 1, this percentage increases to 51 in Group 4. Although it can be concluded that the Group 4 model performs better than the other three models based on the validation results, all four models have the same limitation as the prestressed concrete, not being able to predict superstructure ratings lower than 5 (6 in the case of group 4). However, the number of deficient bridges in each group is very low (see Table 2-23) and is even almost zero in group 4. Therefore, the models cannot perform better than this because there are not enough observations of deficient bridges to train the models.

Similarly, as it was done for the prestressed concrete results, the relative frequency of the residuals greater than 1 (Table 2-24) was calculated based on the information shown in the histograms. Based on the results, it can be concluded that more than 92% of the superstructure ratings were predicted within a difference of one unit, so it is promising to use these models to determine minor maintenance/repairs.

Table 2-24. Percentage of Considerable Differences (>1) of Residuals

Database	Group 1	Group 2	Group 3	Group 4
NBI-2013	7.9%	3.7%	2.6%	3.1%
NBI-2014	7.7%	3.7%	3.2%	3.8%

CONCLUSION AND RECOMMENDATIONS

This paper presents the most reliable models for predicting superstructure ratings for both steel and prestressed concrete bridges in Oklahoma. Although a considerable amount of models has been developed in the last two decades, most studies approach the problem with a single statistical method. Moreover, the author has not found another study that simultaneously examined multiple techniques to select best-fit models nor did a study score two complete NBI datasets to examine and validate the accuracy of the model. It is recommended that this methodology can be used to predict deck and substructure ratings in Oklahoma as well as in any other state. Predicting NBI ratings may help DOTs schedule and estimate bridge maintenance. Furthermore, government agencies can develop better budgets and plan for bridge repair and maintenance. Since the scope of the research is limited to the state of Oklahoma, the results can only be applied to prestressed concrete and steel bridges located in Oklahoma. Nevertheless, the methodology developed in this research can be implemented to obtain predictive models for other bridge ratings, bridge material/design, and in other states. In addition, if cost records of bridges maintenance are available, this information can be complementary for the model development. Moreover, data about natural hazards such as earthquakes, flood, and tornados may be incorporated in future studies.

To predict superstructure ratings of prestressed concrete bridges, a neural network model Neural2) was selected because it outperformed the other techniques with a ASE value of 0.32 in the validation dataset. Conversely, a decision tree is the best model to predict superstructure ratings of steel bridges. However, its ASE is 0.67, which is more than double than the ASE of prestressed concrete bridges. To identify better models, the steel bridge dataset was segregated into four groups per the built year as changes of steel as material, bridge design, and steel bridge inspection occurred over time. Then the same methodology was implemented to develop predicting models for steel sub datasets. It was found that the best model for the oldest bridges (group 1) is a neural network model with an ASE of 0.90 higher than the initially developed model, which suggests that this group is still very heterogeneous and deserves a further study in the future. Yet for the other three groups, decision tree models outperformed other models with ASE values of 0.45, 0.37, and 0.57, which means an improvement from the initial steel dataset. This again confirms that the steel dataset contains more variability than the prestressed concrete bridge dataset. It can also be concluded that the prestressed concrete model is less biased than the four steel models because of a lower ASE value.

Although the prestressed concrete model and steel bridge model for Group 4 have age as the main predictor of superstructure ratings, which has been a consistent finding from previous studies. However, age is not necessarily the main predictor for other steel bridges models (Group1 and Group 2, and Group 3). These models take the location of the bridge (Item 3 County Code) as the most significant input variable rather than age. This could be a result of the segregation of steel dataset into four different groups according to year built. Moreover, the prestressed concrete models is formed by other two parameters: structure length (Item 49) and owner (Item 22), but their percentages of importance with respect to age are less than 32%.

Based on the results of the validation, the prestressed concrete model has an accuracy of 85% and the 39% of the linear variation in the observed values is explained by the variation in the predicted values. As to steel bridges, the model has improved with an accuracy of 77% before data segregation and 92% after segregation, and the same trend was observed on precision, 15% before data segregation and 51% after segregation. In addition, by scoring the complete datasets of 2013 and 2014 according to the scope of this study, it can be concluded that the percentage of superstructure ratings predicted with a tolerance of ± 1 rating is around 98% for prestressed concrete and more than 92% for steel sub-datasets. However, none of the models is capable of predicting superstructure ratings lower than 5. The reason of this limitation is that very few bridges have a superstructure rating lower than 5 compared to the overall population, so they might be treated as outliers by the model. This outcome may restrict the use of these models for determining the timing for a bridge superstructure's reconstruction; nevertheless, the models are accurate enough to be implemented for the timing of minor maintenance and repairs. However, this limitation must be addressed in future research in order to develop models capable to predict all ranges of superstructure ratings.

REFERENCES

Akgül, F., and Frangopol, D. M. (2003). "Rating and Reliability of Existing Bridges in a Network." *Journal of Bridge Engineering*, 8(6), 383-393.

America Society of Civil Engineers - ASCE (2017). "2017 Report Card of America's Infrastructure." <https://www.infrastructurereportcard.org/cat-item/bridges/>. (June 2017).

Ang, R. P., and Goh, D. H. (2013). "Predicting Juvenile Offending." *International Journal of Offender Therapy and Comparative Criminology*, 57(2), 191-207.

Bektas, B. A., Carriquiry, A., and Smadi, O. (2013). "Using Classification Trees for Predicting National Bridge Inventory Condition Ratings." *Journal of Infrastructure Systems*, 19(4), 425-433.

Berry, M. J. A., and Linoff, G. (2011). *Data mining techniques: for marketing, sales, and customer relationship management*, Wiley Pub, Indianapolis, Ind.

Boulila, W., Farah, I. R., Ettabaa, K. S., Solaiman, B., and Ghézala, H. B. (2011). "A data mining based approach to predict spatiotemporal changes in satellite images." *International Journal of Applied Earth Observations and Geoinformation*, 13(3), 386-395.

Bu, G. L., J.; Guan, H.; Loo, Y.; and Blumenstein, M. (2015). "Prediction of Long-Term Bridge Performance: Integrated Deterioration Approach with Case Studies." ASCE, 29(3), 04014089.

Christie, P. G., Jim; thompson, Jeff; and Well, Chip (2011). "Applied Analytics Using SAS® Enterprise MinerTM Course Notes." SAS Institute Inc.

Connor, R. J., Dexter, R., and Mahmoud, H. (2005). "Inspection and Management of Bridges with Fracture-Critical Details." *Transportation Research Board of the National Academies*.

Contreras-Nieto, C. (2014). "Development of linear models to predict superstructure ratings of steel and prestressed concrete bridges." Dissertation/Thesis, ProQuest, UMI Dissertations Publishing.

Elbehairy, H., Elbeltagi, E., Hegazy, T., and Soudki, K. (2006). "Comparison of Two Evolutionary Algorithms for Optimization of Bridge Deck Repairs." *Computer-Aided Civil and Infrastructure Engineering*, 21(8), 561-572.

Federal Highway Administration - FHWA (1995). "Recording and Coding Guide for the Structure Inventory and Appraisal of the Nation's Bridges." U.S. Dept. of Transportation, Federal Highway Administration, Washington, D.C.

FHWA, F. H. A., and ASSHTO (2000). *Asset management primer*, Washington, D.C. : U.S. Dept. of Transportation, Federal Highway Administration, Office of Asset Management, Washington, D.C.

Huang, Y. (2010). "Artificial Neural Network Model of Bridge Deterioration." *Journal of Performance of Constructed Facilities*, 24(6), 597-602.

Kim, Y. J., and Yoon, D. K. (2010). "Identifying Critical Sources of Bridge Deterioration in Cold Regions through the Constructed Bridges in North Dakota." *Journal of Bridge Engineering*, 15(5), 542-552.

Jiang, Y., Saito, M. & Sinha, K. (1988). *Bridge performance prediction model using Markov chain, in Transportation Re- search Record* (No. 1180).

Maryam, D., and Marzieh, A. (2012). "A Hybrid Data Mining Model to Improve Customer Response Modeling in Direct Marketing." *Indian Journal of Computer Science and Engineering*, 3(6), 844-856.

Meyers, L. S. (2009). Data analysis using SAS Enterprise guide, Cambridge New York : Cambridge University Press, Cambridge. New York.

Micevski, T., Kuczera, G., and Coombes, P. (2002). "Markov Model for Storm Water Pipe Deterioration." *Journal of Infrastructure Systems*, 8(2), 49-56.

Moomen, M. (2016). "Deterioration modeling of highway bridge components using deterministic and stochastic methods." S. Labi, J. Fricker, and K. Sinha, eds., ProQuest Dissertations Publishing.

Morcous, G., Rivard, H., and Hanna, A. (2002). "Case-Based Reasoning System for Modeling Infrastructure Deterioration." *Journal of Computing in Civil Engineering*, 16(1), 104-114.

Piñeiro, G., Perelman, S., Guerschman, J. P., and Paruelo, J. M. (2008). "How to evaluate models: Observed vs. predicted or predicted vs. observed?" *Ecological Modelling*, 216(3), 316-322.

SAS Institute INC (2015). "Getting Started with SAS® Enterprise Miner 14.1."

Tang, P., Kanaan, O., Wnag, J., Oh, J., and Kwigizile, V. (2012). "Statistical Identification of Items Importants for Evaluating Bridge Conditions." *Transportation Research Board*.

Thompson, P. D., and Johnson, M. B. (2005). "Markovian bridge deterioration: developing models from historical data." *Structure and Infrastructure Engineering*, 1(1), 85-91.

U.S. Department of Transportation (2013). "Federal Highway Administration National Bridge Inventory." <https://www.fhwa.dot.gov/bridge/nbi/ascii.cfm?year=2013>. (2013).

Veshosky, D., Beidleman, C., Buetow, G., and Demir, M. (1994). "Comparative Analysis of Bridge Superstructure Deterioration." *Journal of Structural Engineering*, 120(7), 2123-2136.

Walsh, S. (2005). "Applying Data Mining Techniques Using SAS® Enterprise Miner." SAS Institute Inc.

Wilson, A. D. (2000). "Production of High Performance Steel for U.S.A Bridges." National Bridge Research Organization, University of Nebraska-Lincoln, Nebraska Library Commission.

CHAPTER III

CHARACTERIZATION OF STEEL BRIDGE SUPERSTRUCTURE DETERIORATION THROUGH DATA MINING TECHNIQUES

ABSTRACT

As a significant number of steel bridges are approaching the end of their service life, understanding deterioration characteristics will help bridge stakeholders better prioritize bridge maintenance, repairs, and rehabilitation. This paper applies data mining techniques including logistic regression, decision trees, neural networks, gradient boosting, and support vector machine to the 2013 National Bridge Inventory to classify steel bridges as deficient or non-deficient through the estimation of the probability of steel bridge superstructures reaching deficiency. A focused subset was created based on the defined scope of the research: design material (steel and steel continuous), type of design (stringer/multi-beam or girder), and deck type (cast-in-place concrete). Deterioration factors considered included age, average daily traffic, design load, maximum span length, and structure length. The impacts of these factors affecting steel bridge superstructure deterioration were identified. Outcomes of the analysis afford bridge stakeholders the opportunity to better understand factors that relate to steel bridge deterioration as well as provide a means to assess other risks associated with bridge maintenance, repair, and rehabilitation.

INTRODUCTION

According to the American Society of Civil Engineers'(ASCE) 2017 Report Card for America's Infrastructure, one in nine bridges in the United States were structurally deficient in 2016 (ASCE 2017). This report also presents the new estimate of bridge rehabilitation, which is $123 billion. Although the classification of structurally deficient does not mean that a bridge is unsafe or is not accepted to be transited, neither does it imply that a structurally deficient bridge is not a risk if it does not get considerable maintenance. In general, steel and concrete are the types of materials predominantly used for bridge superstructure. Typically, a particular bridge is referred to as 'steel bridge' or 'concrete bridge' based on its superstructure material type. Steel bridges have a long history with an average age older than concrete bridges (U.S. Department of Transportation, 2013). As shown in Figure 3-1, superstructure deficient steel bridges comprised over one-half of all deficient bridges in 2013 (U.S. Department of Transportation, 2013). Moreover, as a significant number of steel bridges are approaching the end of their service life, deteriorated bridges may endanger the traveling public and eventually impact the economy if unaddressed. In order to keep steel bridges serviceable and safe for the public, it is crucial to maintain these bridges properly; however, bridge maintenance and rehabilitation activities are constantly constrained by tight budgets and other competing projects. Although the superstructure rating is not the only parameter to determine whether or not a bridge is structurally deficient, according to Federal Highway Administration, the superstructure has been identified as a source of bridge deterioration, especially steel superstructures in cold regions (Kim and Yoon, 2010). Moreover, researchers, such as Farhey (2014), have recommended studying bridge characteristics of low structural performances in order to improve bridge management/maintenance and understand deterioration.

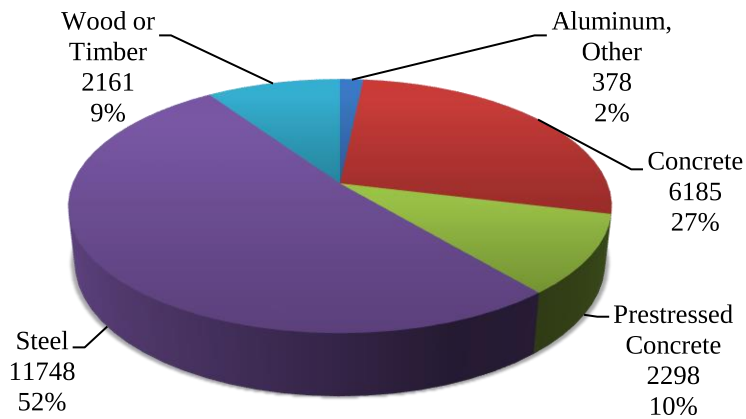

Figure 3-1. Distribution of Deficient Bridges by Superstructure Material (2013 NBI)

In order to make wise decisions concerning bridge design and bridge maintenance prioritization, determining when a bridge reaches deficiency and what characteristics accelerate deficiency have become topics of interest. These areas of study have also been recommended by ASCE (2017) in its last report. Moreover, as Schulz et al. (2017) stated, Metropolitan Planning Organizations and Department of Transportation are facing limited funds for transportation projects, which including bridges, these agencies are obligated to select and prioritize potential projects (i.e. maintenance or new construction).

This study focuses on using data analytics to characterize steel bridge deterioration by implementing new data mining approaches, which have outperformed regression techniques in some cases, to steel bridge inspection data from the U.S. nationwide. Thus, the author studies the connection between characteristics of steel bridges and superstructure deficiency. Some of the characteristics of steel bridges involved in this research are age, structure length, location, design load, and average daily traffic (ADT) that are recorded in the National Bridge Inventory (NBI). In addition, it is possible to predict the probability of a steel bridge superstructure reaching deficiency based on its characteristics. The objective of this study is to identify critical factors associated with steel bridge superstructure deficiency as well as a model capable of predicting a steel bridge superstructure's deficiency status based on the probability obtained from the model.

Thus, this paper contributes to the overall knowledge by providing a managerial tool for DOTs and other bridge stakeholders to understand the superstructure of steel deficient bridges and make future decisions on steel bridge design.

BACKGROUND AND LITERATURE REVIEW

National Bridge Inventory (NBI)

Historical data is a good source to gain more understanding about deterioration of bridges. Both NBI and Commonly Recognized (CoRe) Bridge Elements are used in bridge deterioration studies. NBI can be accessed online publicly (U.S. Department of Transportation, 2013) while CoRe data are not publicly accessible. NBI archives bridge condition data reported by each state Department of Transportation (DOT) across the nation on an annual basis. The NBI database contains not only various condition ratings of bridges but also basic design information, current service, and other exhaustive information about every bridge in the nation. NBI has been extensively used for many studies related to bridges. Twenty-four years of bridge condition data from 1992 to 2016 are available for public use (U.S. Department of Transportation, 2013).

Insights from Previous Related Studies

Previous studies predominantly focused on prediction of bridge ratings, such as superstructure ratings and deck ratings. Moreover, a variety of techniques have been applied, including regression, artificial neural networks, Markov models, and regression trees. Veshosky et al. (1994) used NBI-1990 data and developed two groups of regression models to predict superstructure ratings: one for prestressed concrete bridges and the other for steel bridges. However, they specifically examined bridges in seven states and only included bridge age and average daily traffic (ADT) as predictors for model development. Contreras-Nieto (2014) conducted a similar study for bridges in the state of Oklahoma. He employed multiple regression and included more predictors for model development and confirmed that age was a significant predictor of bridge superstructure ratings. In addition, the results showed that design load was another significant variable for predicting superstructure ratings, but ADT was not. Both models developed by Veshosky and Contreras-Nieto had low coefficients of determination (R^2) of 20% and 31%, respectively. Bektas et al. (2012) applied Classification And Regression Trees (CART) algorithms to the CoRe element condition data from three states to predict NBI bridge ratings. Their results did not explicitly state which variables were significant to the prediction of superstructure ratings. Bu et al. (2015) compared the performance of a standard Markovian model and an Elman Neural Network and Back Prediction model in predicting transition probabilities of condition ratings of bridges irrespective of maintenance. However, their model provided limited knowledge of specific bridge characteristics that correlate with deterioration.

More recently, Bektaş (2017) predicted NBI ratings by implementing a recursive partition method and decision trees technique. In this study, the main source of information is field-collected data of national bridge elements with their condition ratings extracted from the NBI-2016 database. Although the results looked promising because the percentage of correct matches varied between 42% and 62%, which are higher than previous studies, the importance of the contributors (predictors) of each model was not presented. In another study developed by Saeed et al. (2017), deterioration models were created for different bridge design types and superstructure materials (concrete and prestressed concrete) by using NBI databases and polynomial and exponential functions. In addition, the models developed were implemented to determine the life span of the bridge studied. In other words, the models were used to predict when a bridge reaches a superstructure rating equal to four. The results confirmed the significance of factors such as age, number of spans in the main unit, and average daily traffic of trucks in bridge deterioration as previous studies have stated. Also, climate-related variables (freeze index, number of cold days, number of freeze thaw cycles, and precipitation) were also found to be significant factors.

New Machine Learning Techniques

In addition, Gradient Boosting Machine (GBM) and Support Vector Machine (SVM) are state-of-the-art data mining techniques that have been successfully applied in other areas of the civil engineering field. For example, Tarefder et al. (2014) performed a sensitivity analysis of the Mechanistic-Empirical Pavement Design Guide (MEPDG) by using GBM, and they were able to identify high-, moderate-, and minimal-sensitivity inputs with respect to the MEPDG outputs of the flexible pavement design. Also, Chou et al. (2011) and Omran et al. (2016) used GBM and SVM techniques with a variety of other techniques in order to predict concrete compressive strength. By comparing the implemented techniques, both studies conclude that advanced data mining techniques obtained high accuracy predictions. However, it takes more time to develop and train these models than other techniques. In another study, Zhu and Brilakis (2010) employed SVM and artificial neural network (ANN) techniques to identify concrete material sections. Then the results were compared with previous outcomes obtained from manual classification methods. Although it was found that ANN performed better than SVM, SVM results are considered decent because accuracy of the datasets (training and test) was higher than 96%.

Significance and Uniqueness of the Study

Although different studies used NBI as a data source to develop prediction models for NBI ratings, these previous studies were somewhat limited. For example, the models were developed based on a small number of states or regional representations. Also, the number of predictors considered in the models were very limited. The analysis presented here aims to address these limitations by: 1) Including a large number of bridges from various states; 2) Including a large variety of bridge characteristics including numerical, categorical and ordinal variable types; and 3) Incorporating new machine learning techniques (ANN, decision trees, GBM, and SVM) into classical statistics techniques (regression).

METHODOLOGY

This section presents the steps and techniques used to achieve the objective of the study.

1. Develop a research dataset from the NBI-2013 database;
2. Characterize the data through exploratory analysis;
3. Perform quality assurance on the data in order to detect outliers, impute missing values, and transform data to meet the underlying assumptions of the specified data analysis techniques;
4. Develop models to characterize deficiency by implementing five different techniques: logistic regression, decision trees, neural network, gradient boosting, and support vector machine;
5. Validate the models in order to select the best-fit model;
6. Assess the relative impact of the model parameters on steel bridge superstructure deficiency; and
7. Implement the best model using two NBI databases (2013 and 2014) in order to assess the accuracy of the classification model selected in Number 5.

Dataset

The analysis only addresses bridges with steel superstructures using NBI-2013 data. A data subset was created based on the following criteria:

1. Material types including steel and steel continuous;
2. Design type including stringer/multibeam or girder; and
3. Deck type was cast-in-place concrete.

These steel bridge characteristics represent the majority of bridges in the steel bridge population. These characteristics enabled the investigators to develop more focused models. Based on the above criteria, 90,420 steel bridges were selected for data analysis.

Model Parameters

NBI database stores over 100 variables that fulfill the National Bridge Inspection Standards (23 CFR 650.3); however, not all of them are useful for model development (FHWA, 1995). Through a review of previous research (Contreras-Nieto, 2014; FHWA, 1995; Tang et al., 2012), relevant variables were selected. Table 3-1 presents a list of the variables included in the model development process along with a brief description of the variable. A detailed explanation of the variables can be obtained from the NBI data dictionary (http://nationalbridges.com/nbiDesc.html). These variables depict the factors regarding structure characteristics, traffic, and agencies. Some of the variables included in Table 3-1 are derived from NBI data. For example, age was derived from the year that a bridge was built. The dependent variable (target variable) for this analysis was a binary variable (deficient or non-deficient). This variable was derived from the NBI superstructure condition rating, ranging from 0 to 9 with '0' indicating a failed condition and '9' indicating an excellent condition. Accordingly, the ratings equal to or less than four were classified as deficient and ratings above four were classified as non-deficient. The remaining variables in Table 3-1 were used as independent (input) variables. However, this analysis also attempted to identify whether or not these independent variables are correlated (redundant) and are really significant to predict superstructure ratings.

Table 3-1. Summary of Model Variables

Variable Name	Description	Role	Type
Age	Age of the Bridge in 2013	Input	Interval
Item 1	State Code	Input	Nominal
Item 22	Owner	Input	Nominal
Item 26	Functional Classification of Inventory Route	Input	Nominal
Item 29	Average Daily Traffic (ADT)	Input	Interval
Item 31	Design Load	Input	Nominal
Item 42b	Type of Service under Bridge	Input	Nominal
Item 45	Number of Spans in Main Unit	Input	Interval
Item 46	Number of Approach Spans	Input	Interval
Item 48	Length of Maximum Span	Input	Interval
Item 49	Structure Length	Input	Interval
Item 51	Bridge Roadway Width, Curb-to-Curb	Input	Interval
Item 52	Deck Width, Out-to-Out	Input	Interval
Sup_def	Superstructure classified as deficient or not deficient	Target	Binary

Data Mining Techniques

Different data mining methods were employed to develop the models, including logistic regression, decision trees, neural network, Gradient Boosting Machine, and Support Vector Machine. Some techniques may have several varied models depending on the variation of model variable treatment or variable selection methods. Using multiple techniques instead of a single technique may yield the best-fit models. Among these data mining techniques, Gradient Boosting Machine (GBM) and Support Vector Machine (SVM) are state-of-the-art data mining techniques that have been successfully applied in other areas of civil engineering as presented in the literature review.

The tool used for the analysis was SAS® Enterprise Miner (EM). All of the techniques mentioned above were packaged in the software. In addition, its capability in handling a variety of data types makes it a viable tool for this study. EM also permits comparison of the performance of the developed models and choosing the best model. In the following section, these different data mining techniques are described in detail.

Logistic Regression

A logistic regression model can be expressed mathematically by Equation (1).

$$ln\left(\frac{p}{1-p}\right) = logit(p) = \beta_0 + \beta_1 X_1 + \cdots + \beta_K X_k \quad (1)$$

Where p is the probability of an event occurring. In this study, p is the probability of a steel bridge having a deficient superstructure. X*'s* are the predictors (independent variables) and β's are the coefficients. Since the NBI database involves categorical variables, these variables were recoded into multiple variables to accommodate different levels of values associated with the categorical variables. Two different groups of models were developed. One group allowed the quadratic transformation of the original variables and the other group did not.

Decision Trees

Decision trees split the data into smaller groups that have similar values of the dependent variable (target). In order to divide the data into "pure" groups, the algorithm selects an input and calculates a fixed split point; thus, branch-like segments are created. The quality of the breakup is measured by Equation (2); the highest value of $logworth$ is the best partition (Christie et al., 2011). As a result of this process, it is a collection of rules to segment the data. The rules are known as English Rules because they are easy to interpret; therefore, decision trees are widely used in data mining. Figure 3-2 presents a simple example of decision trees that has two nodes and three leaves. In addition, this technique is not sensitive to missing values and outliers, which is an advantage over other techniques.

$$logworth = -log(chi\ squared\ p - value) \qquad (2)$$

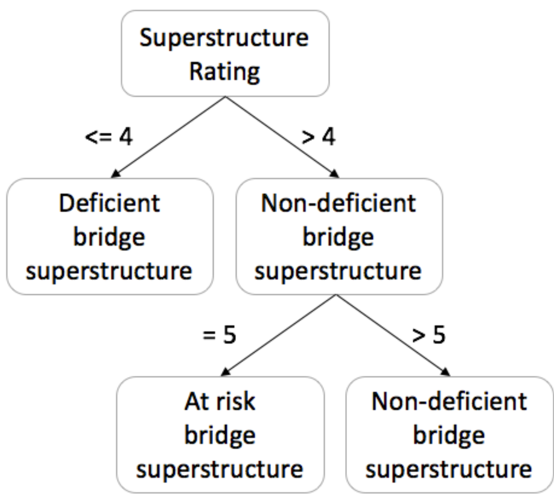

Figure 3-2. Bridge Superstructure Classification

Neural Network

Neural networks are commonly used in the field of machine learning and cognitive science. The neural network mimics the way human and animal brains' perform computation when dealing with multiple complex inputs (Hagan et al., 1996). Neural networks are superior in handling problems with non-linear functions. Figure 3-3 shows a high-level architecture of the neural network. A neural network model is formed by layers (input, hidden, and output) and activation functions (combination and transfer functions). Although neural networks are flexible and powerful in solving many practical problems, it is not the first choice when model interpretability is required for the solution. Nevertheless, it can be interpreted by a sensitivity analysis (Berry & Linoff, 2011). As a result, it is still possible to obtain the relative importance of input variables compared to each other.

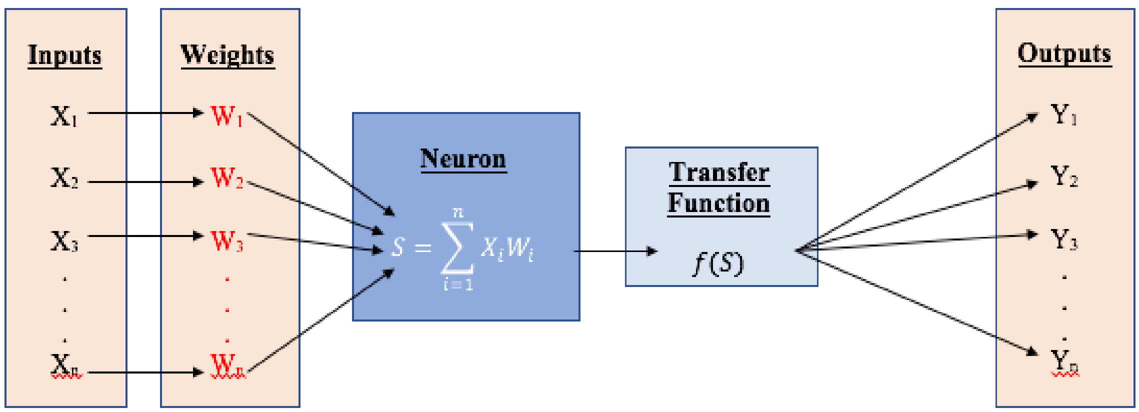

Figure 3-3. Processing Information in a Neural Network

Gradient Boosting

Gradient boosting is a sub-technique of decision trees. Gradient boosting models are a sequential assembly of different decision trees. Prediction is based on several subtrees which makes these models more robust; however, the interpretability of the models is reduced because of the complexity of the gradient boosting models. Therefore, the results cannot be summarized with English rules as with the decision tree technique. Nevertheless, the high complexity of gradient boosting models gives the models good prediction power, and they are useful for variable selection (Yuan, 2015).

Support Vector Machine

According to Berry & Linoff (2011), as well as DeVille and Neville (2013), Support Vector Machine (SVM) is a relatively new technique that aims to divide data into classes. This classification is accomplished by applying a kernel function in order to bring the data from a lower dimension to a higher dimension. The algorithm divides the data with a new plane; however, the main limitation of this technique is to implement the appropriate kernel function among linear, sigmoid, polynomial, KMOD, radial basis function (RBF), or exponential RBF kernels. For example, Figure 3-4 presents a particular case of a two-feature space where a kernel function (ϕ) is implemented to map the data to a three-feature space. This method is an alternative to neural networks that strives to obtain better results.

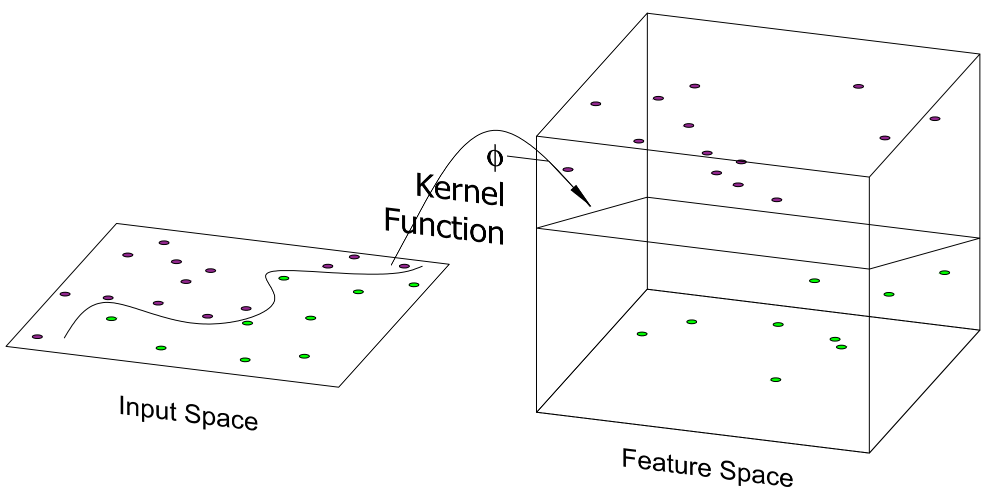

Figure 3-4. Support Vector Machine

All of the aforementioned techniques were used to develop models to predict the probability of a specific steel bridge having a deficient superstructure. If an obtained probability is less than the determined threshold (see in Rare Case section), the outcome of the prediction is 'non-deficient'. Conversely, the output of the prediction is considered 'deficient' if the computed probability is equal or greater than the threshold.

Level Reduction of Nonnumeric Inputs

Categorical, or nonnumeric, inputs with excessive levels may lead to overfitting of the regression and neural network models (Berry & Linoff, 2011). For example, the state code of the bridge location has 50 values; therefore, variable values with a large number of levels were addressed by using decision trees algorithm to group levels based on their relationship with the target variable in order to reduce the number of the levels of qualitative variables.

Rare Case

According to NBI-2013 data, 2,950 of 90,420 bridge superstructures were categorized as deficient, accounting for 3.3% of the population. In other words, 96.7% of the bridges were not deficient. If a given model always predicts ‘non-deficient’ regardless of the model input and this model is used for prediction based on the data, approximately 97 of 100 times the model would have the right prediction. This does not necessarily mean that this model is a good predictive model. Since deficient bridges were the study of interest, and few bridges had deficient superstructures, this problem was treated a ‘rare case’ (Milley, 1998).

Creating a balanced sample is one of the approaches to manage ‘rare case’ problems for model development (Berry & Linoff, 2011; Milley, 1998). The balanced sample guarantees that the rare events included in the sample are in the same proportion as the non-rare events; therefore, a balanced sample was created with 2,950 deficient bridges and 2,950 non-deficient bridges. The non-deficient bridges were selected randomly from all non-deficient steel bridges in order to avoid the ‘rare case’ problem. The probabilities obtained from a model that is developed from a balanced sample must be corrected because of the balanced sample effect (Veshosky et al., 1994). According to Potts and Patetta (2000), Equation 3 can be used to correct the probabilities obtained with a balanced sample.

$$\hat{p}_i = \frac{\hat{p}_i^{\,*}\rho_0\pi_1}{(1-\hat{p}_i^{\,*})\rho_1\pi_0+\hat{p}_i^{\,*}\rho_0\pi_1} \qquad (3)$$

Where $\hat{p}_i$ is the corrected probability; $\hat{p}_i{}^*$is the probability obtained with the balanced sample model; π_0 and $\pi_1$are the proportions of the (0) and (1) events in the population, respectively; and ρ_0 and ρ_1 are the proportions of the (0) and (1) events in the sample, respectively.

In addition to the correction of the probabilities, an adjustment of the classification threshold has to be made. As Pozzolo et al. (2015) explained, the reason for this adjustment is to conserve the predictive accuracy obtained in the balanced datasets after correcting the probabilities based on the proportion of the event in the population. Therefore, Equation 4 is applied to determine the threshold used for the decision-making process (classification of bridge superstructures).

$$\tau' = \frac{\beta\tau_s}{(\beta-1)\tau_s+1} \qquad (4)$$

Where $\tau'$is the threshold for the unbiased probability $\hat{p}_i$ (see Equation 3); $\beta = \frac{N^+}{N^-}$ (N^+denotes the number of deficient bridges, and N^- denotes the number of non-deficient bridges); and $\tau_s$is the threshold defined in the balanced sample, which is 0.5 (50%).

Model Training and Validation

After a balanced sample was created, the sample was further divided into two datasets. Through a random drawing process, 70% of the data were used for training and 30% of the data were used for model validation.

Model Implementation

The best model is fed with independent variables of the of NBI database in order to obtain the probability of a bridge having a deficient superstructure and consequently determine its classification (deficient/non-deficient). Thus, to compare the classification results with the NBI superstructure ratings, a confusion matrix is created (see Table 3-2). Confusion matrix or error matrix are used to represent the consistency of two classification outcomes (Thoonen et al., 2012). From the cross tabulation (confusion matrix), indices to measure the goodness of the predicting model can be calculated.

Table 3-2. Confusion Matrix

		Predicted Rating	
		Non-deficient	Deficient
Actual Rating	Non-deficient	a	b
	Deficient	c	d

The following key indicators (Matthiesen 2010) that can be calculated considering the values of the class counters (a, b, c, d – Table 3-2) are:

1. Accuracy (AC): the proportion of bridge ratings predicted correctly.

$$AC = \frac{a+d}{a+b+c+d}$$

2. Sensitivity/Recall/True-positive rate (TP): the proportion of deficient bridges that were predicted correctly.

$$TP = \frac{d}{c+d}$$

3. Specificity/True-negative rate (TN): the proportion of non-deficient bridges that were predicted correctly.

$$TN = \frac{a}{a+b}$$

4. False-alarm/False-positive rate (FP): the proportion of non-deficient bridges that were predicted incorrectly.

$$FP = \frac{b}{a+b}$$

5. Miss-rate/False-negative rate (FN): the proportion of deficient bridges that were predicted incorrectly.

$$FN = \frac{c}{c+d}$$

6. Precision (P): the proportion of predicted deficient bridges that were predicted correctly.

$$P = \frac{d}{b+d}$$

In addition, as Haibo He and Garcia (2009) stated, two other metrics are also analyzed in problem with rare events, which are described in the following.

7. F-Measure (f-score): combines recall (TP) and precision in order to evaluate the success of the classification.

$$F - score = \frac{(1+\beta)^2 * Recall * Precision}{\beta^2 * Recall + Precision}$$

8. G-mean: assesses the balance between classification capacity on the predominant and rare classes.

$$G - mean = \sqrt{TP * TN}$$

However, not all the previous measures are meaningful for all problems; it depends on the nature of the problem itself. In this case, the author chose the following indicators: accuracy (AC), sensitivity (TP), and false-negative (FN) to measure how good the model is in terms of properly classifying deficient, non-deficient bridges, or both. In addition, F-measure and G-mean were adopted because they are useful measures when the problem is involved with rare cases. It is critical to develop a robust model that can correctly predict the deficiency state of the bridges in order to complement bridge inspections to estimate and schedule their maintenance/repairs or reconstruction in a proactive manner. If these bridges are misclassified, no action can be taken in advance. Therefore, the consequences of a bridge collapse are mainly fatalities and injuries, and also travel times are increased or the path can be disrupted, depending on the case. Finally, economic impacts are faced in any scenario, but the amount depends on the magnitude of the disaster.

RESULTS

Figure 3-5 is a particular view of the regression models that presents a partial view of the project flow developed in EM, which contains three major parts, including: 1) Building of the models (some) ; 2) Assessment of the models; and 3) Score of the full dataset (calculating the probability of a deficient superstructure for the entire dataset). Different models were developed and the model with the best performance was selected. The best model was used to estimate the probability of a bridge having a deficient superstructure.

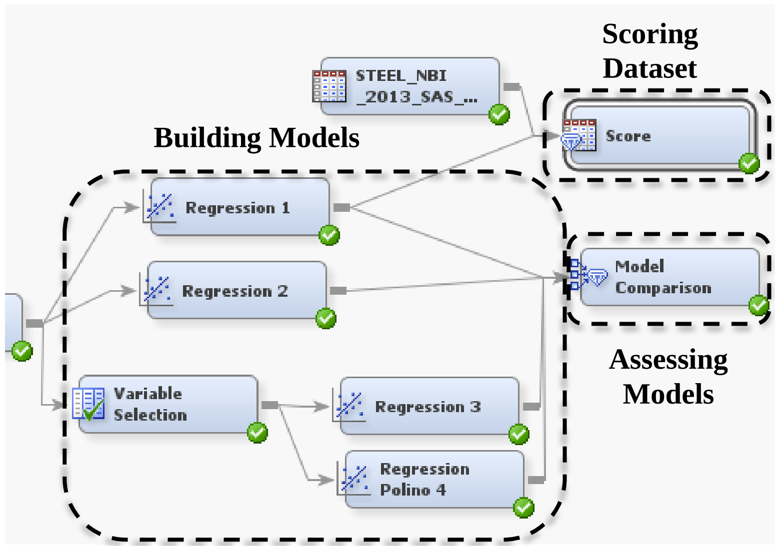

Figure 3-5. SAS® Enterprise Miner Partial Project Flow

Model Comparison

Misclassification rates were used as a metric to evaluate the effectiveness of the models. The misclassification rate was defined as a ratio of the number of predictions with false negative and false positive to the total number of predictions. Table 3-3 shows the misclassification rates for both the validation and training sets by using various data mining approaches. The model variation under each data mining technique is explained in the footnotes of Table 3-3. Validation misclassification rates were used as the primary model performance indicator. The models were ranked per the validation misclassification rate in ascending order. All of the models had a very close misclassification rate, ranging from 0.210 to 0.254. The model with the best performance was Regression 1 that used a stepwise approach to select model parameters, which had the lowest misclassification rate of 0.210. This means that the probability of the model having a wrong prediction was 21.0% in the validation dataset.

Table 3-3. Results of Model Comparison Node Results

Model	Validation Misclassification Rate	Training Misclassification Rate
Regression 1[1]	0.210	0.213
Regression Polynomial 4[2]	0.213	0.220
Regression 3[3]	0.214	0.221
Gradient Boosting	0.215	0.218
Regression 2[4]	0.216	0.218
Neural Network[5]	0.218	0.217
SVM[6]	0.227	0.224
Full Decision Tree[7]	0.229	0.198
Decision Tree[8]	0.232	0.222
Decision Tree (Interactive)[9]	0.234	0.229
SVM (2)[10]	0.254	0.236

Note:
[1] Linear model – Stepwise approach
[2] Polynomial model – Variable selection approach
[3] Linear model – Variable selection approach
[4] Polynomial model – Stepwise approach
[5] Neural Network – Multilayer perceptron with 6 hidden units
[6] SVM model- Kernel function: linear
[7] Decision Tree created by SAS EM automatically based on properties defined
[8] Decision Tree created by the authors using logworth values
[9] Decision Tree created by the authors using preferred inputs and its logworth values
[10] SVM model- Kernel function: polynomial

Best Model

The best model was Regression 1; Table 3-4 summarizes its predictors and statistical results. The coefficient estimates of the model predictors correspond to the β values in Equation 1. Standard estimates are an indicator of the effects of predictors on the model, and they are meaningful when numeric variables are compared. According to Table 3-4, age was the most influential quantitative predictor because the absolute value of its standard estimate (0.69) was the highest. This finding agrees with previous studies where bridge age has been identified as a significant input of bridge deterioration. The relationship between the probability of a bridge having a deficient superstructure and its age was a positive relationship. In other words, as the bridge age increases, the probability of superstructure deficiency increases. In addition, the logarithm of the length of the maximum span is the second most significant numerical predictor; this variable also was found to be a significant predictor by Tang et al. (2012). The length of the maximum span and the probability of a bridge having a deficient superstructure were in an inverse relationship. This signifies that the probability of a superstructure reaching deficiency increases as the length of the maximum span decreases.

Table 3-4. Statistical Summary of the Best Model Using Maximum Likelihood Estimates

Variable	DF	Coeff. Estimate	Error	Chi-Square	Pr > Chi-Sq.	Standard Estimate
Intercept	1	0.25	0.61	0.17	0.68	
ADT	1	10^{-4}	0.00	17.52	<.0001	0.12
LOG_Structure Length	1	0.49	0.08	37.11	<.0001	0.25
LOG_Length of Maximum Span	1	-1.02	0.13	61.79	<.0001	-0.33
PWR_Bridge Roadway Width	1	-2.07	0.56	13.71	0.00	-0.12
State Group 1	1	0.29	0.09	10.72	0.00	
State Group 2	1	0.63	0.10	40.76	<.0001	
State Group 3	1	1.21	0.13	83.07	<.0001	
State Group 4	1	1.02	0.10	95.30	<.0001	
State Group 5	1	-0.43	0.13	10.66	0.00	
State Group 6	1	-0.04	0.10	0.17	0.68	
State Group 7	1	-1.11	0.13	78.65	<.0001	
Owner – G1	1	0.06	0.17	0.12	0.73	
Owner – G2	1	0.21	0.18	1.37	0.24	
Owner – G3	1	0.58	0.19	9.05	0.00	
Owner – G4	1	-0.04	0.57	0.01	0.94	
Age	1	0.05	0.00	472.47	<.0001	0.69

The odds ratio is the only measure of association directly estimated from the logistic regression model (Ziegel, 2003). The odds ratio represents the change of odds in the principal outcome in relation to one unit of change in a predictor (see Equation 5). Table 3-5 presents the odds ratio estimates for the predictors of the best model (Regression 1). It should be noted that the increment for ADT is in thousands and the increment for length/width is 100 mm as defined by NBI. For an increase of one thousand vehicles in ADT, the odds of a bridge having a deficient superstructure increases by 10%. Similarly, the probability of a bridge having a deficient superstructure increases by 5% each year afterwards when comparedwith the preceding year.

$$Odds\ ratio = \frac{P(X)}{1-P(X)} \qquad (5)$$

Since the state location and the owner of the bridges are categorical variables, the interpretation for qualitative variables is slightly different than quantitative variables. State Group 8 (Alaska, Oregon, Texas, Utah, Vermont, and Wyoming) was used as the base case. The odds ratios in other states groups were calculated as the ratio of the probability of having a deficient superstructure when compared to State Group 8. For example, bridges located in State Group 3 (Michigan, Pennsylvania, and Rhode Island) had a probability of a deficient superstructure 16 times as high as in State Group 8 when other predictors were held the same. Similarly, with the ownership predictor, Owner Group 5 (State Toll Authority, U.S. Forest Service, and National Park Service) was used as the base case for odds ratio estimates. The probability of a bridge superstructure being deficient in Owner Group 3 (Town Highway Agency, City or Municipal Highway Agency, Other State Agency, and Rail Road) is about four times as high as that in Owner Group 5.

Table 3-5. Odds Ratio Estimates

<table>
<tr><th>Effect</th><th>Units</th><th>Estimate</th><th>Remark</th></tr>
<tr><td>ADT</td><td>1000 vehicles</td><td>1.10</td><td rowspan="10">State Group 1: IA, IL, MA, ND, SD, SC, VA, WV
State Group 2: CT, DE, IN, MN, MS, OK
State Group 3: MI, PA, RI
State Group 4: KY, LA, NC, PR, TN, WI
State Group 5: ID, KS, NH, OH, WA
State Group 6:AR, AL, AZ, CA, GA, FL, ME, NM, NY
State Group 7: CO, MD, MO, MT, NE, NJ
State Group 8: AK, OR, TX, UT, VT, WY

Owner G1: State Highway Agency, Other Local Agencies, and Bureau of Indian Affairs.
Owner G2: County Highway Agency.</td></tr>
<tr><td>LOG_Structure Length</td><td>0.1 meters</td><td>1.64</td></tr>
<tr><td>LOG_Length of Maximum Span</td><td>0.1 meters</td><td>0.36</td></tr>
<tr><td>PWR_Bridge Roadway Width</td><td>0.1 meters</td><td>0.13</td></tr>
<tr><td>State Group 1 vs 8</td><td></td><td>6.31</td></tr>
<tr><td>State Group 2 vs 8</td><td></td><td>8.91</td></tr>
<tr><td>State Group 3 vs 8</td><td></td><td>15.84</td></tr>
<tr><td>State Group 4 vs 8</td><td></td><td>13.14</td></tr>
<tr><td>State Group 5 vs 8</td><td></td><td>3.08</td></tr>
<tr><td>State Group 6 vs 8</td><td></td><td>4.56</td></tr>
</table>

Effect	Units	Estimate	Remark
State Group 7 vs 8		1.56	Owner G3: Town Highway Agency, City or Municipal Highway Agency, Other State Agency, and Rail Road Owner G4: State Park, Forest, or Reservation Agency, Local Toll Authority, Corps of Civil Engineers, and Army. Owner G5: State Toll Authority, U.S. Forest Service, and National Park Service
Owner G1 vs G5		2.38	
Owner G2 vs G5		2.75	
Owner G3 vs G5		4.01	
Owner G4 vs G5		2.15	
Age	1 year	1.05	

In order to gain insights into the model and its misclassification rate, other relevant indicators, such as sensitivity (TP) and false negative (TN), were examined as well. The ratios were 83% and 19% for sensitivity and false negative, respectively

Implementation of Best Model

In order to test the model with the best performance (Table 3-3), two NBI databases (2013 and 2014) were scored. As a result, the probability of a bridge having a deficient superstructure was predicted, and then those probabilities were corrected using Equation 3. In addition, the classification threshold for the corrected probabilities was determined (Equation 4); its value was 0.0409 in the case of the NBI-2013. Thus, bridges were classified as deficient when the corrected probability exceeded 0.0409 and otherwise were classified as non-deficient. A comparison between the observed bridge classification based on the superstructure rating (NBI) and the predicted classification was performed, and it is presented as a confusion matrix.

NBI-2013

After scoring the full dataset of 90,420 steel bridges, the corrected probability of a bridge having a deficient superstructure given by the best model ranged from almost 0 to 0.99 with a mean of 0.05 and a standard deviation of 0.007. Figure 3-6 presents the histogram of the corrected probabilities of a bridge superstructure being classified as deficient. The distribution was skewed to the right; thus, the majority of bridges obtained a low probability of superstructure deficiency. It means bridges are more likely to be classified as non-deficient, which is the expected outcome.

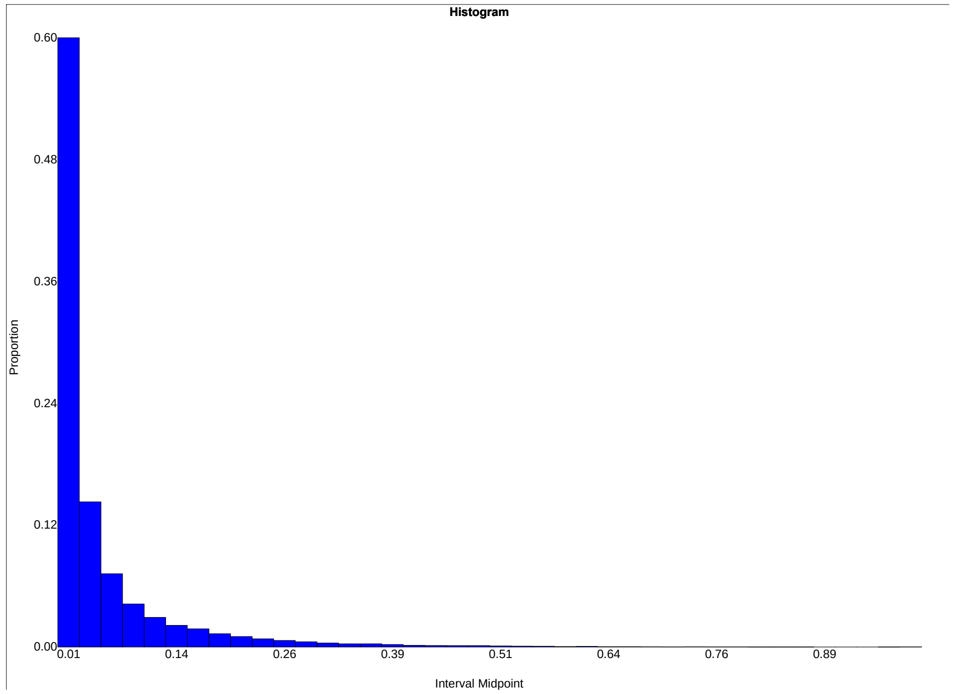

Figure 3-6. Histogram of Probability of a Steel Bridge Superstructure Being Classified as Deficient – NBI-2013

NBI-2014

Similarly, steel bridges in NBI-2014 were scored using the best model selected previously. The corrected probability ranged from almost 0 to 0.98 with a mean of 0.04 and a standard deviation of 0.005. Similar to the trend observed in NBI-2013 database, the distribution of the probability was skewed to the right (see Figure 3-7).

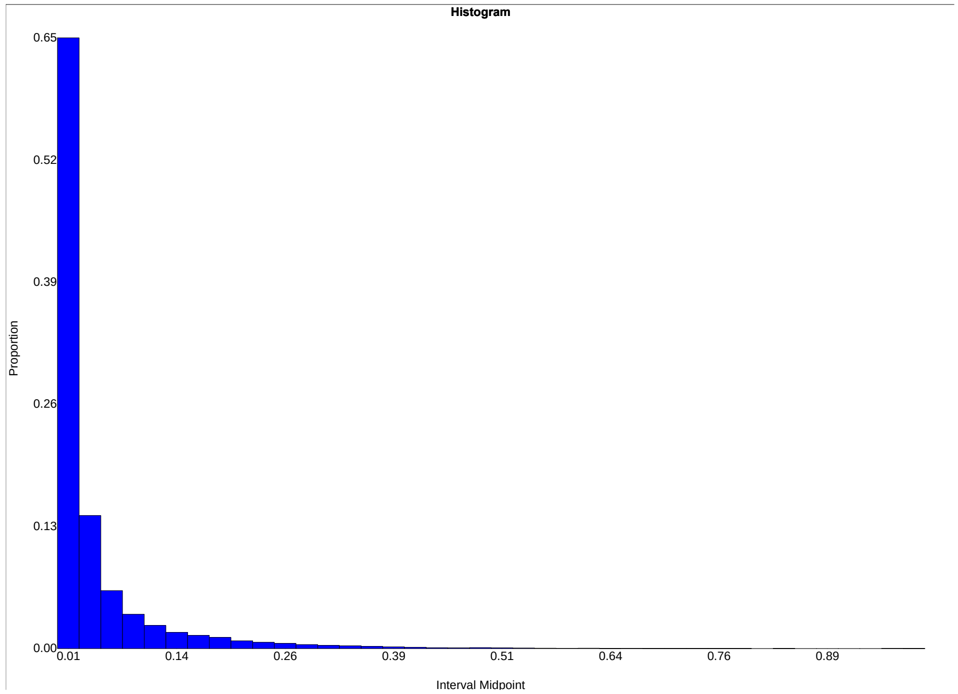

Figure 3-7. Histogram of Probability of a Steel Bridge Superstructure Being Classified as Deficient – NBI-2014

Confusion Matrix

In order to summarize the outcome of the scoring of the two databases and to evaluate the performance of the model, confusion matrices (Table 3-6) and some measures (Table 3-7) were created for each database.

Table 3-6. Confusion Matrix of the Model

		Predicted Rating NBI 2013		**Predicted Rating NBI 2014**	
		Non-deficient	Deficient	Non-deficient	Deficient
Actual Rating	Non-deficient	63,587	23,337	63,809	18,520
	Deficient	672	2,824	774	2,476

Table 3-7. Summary of the Performance Measure of the Model Tested on NBI 2013 and 2014

Measure	NBI 2013	NBI 2014
Accuracy (AC)	73.4%	77.5%
Sensitivity (TP)	80.8%	76.2%
Miss-rate (FN)	19.2%	23.8%
Precision	10.8%	11.8%
F-measure	19.0%	20.4%
G-mean	76.9%	76.8%

According to Table 3-7, all measures are very similar and desirable for both databases. The 80.8% of superstructure deficient bridges were classified correctly (TP), which is a significant outcome of the research, even though the miss-rate (FN) is 19.2%. In addition, the ideal results for this type of problem is to obtain high sensitivity (TP) and high precision, but precision and TP are conversely related. This means that precision will be adversely affected when TP is high and vice-versa (Pozzolo et al., 2015). In general, the accuracy of the classification is higher than 73% for both databases which is also confirmed by the G-mean (76%), another indicator of effective performance for unbalanced databases. However, the F-measure is low (19%), but it is caused by the same inverse relationship between TP and precision.

CONCLUSIONS AND RECOMMENDATIONS

This analysis utilized a variety of data mining techniques to characterize deterioration rates of steel bridge superstructures. Logistic regression was an effective approach to predict the probability of superstructure deficiency, as well as identify the impacts of bridge factors. The results showed that the best model obtained a misclassification rate of 21% on the balanced sample, which was considered satisfactory. Critical factors included the state, owner, age, ADT, maximum span length, structure length, and bridge roadway width. The most important predictor was age, which is consistent with findings identified by previous researchers.

According to the odds ratio estimates on the predictors, several findings were considered significant. Bridges characterized for having longer maximum spans and wider road widths obtained low probabilities of having deficient superstructures. On the other hand, a high probability of having a deficient superstructure was found for older, longer, and higher traffic bridges. Both the owner and location of the bridge were influencing factors that correlated with the probability of superstructure deficiency. The odds ratio estimates on those two predictors provide a benchmark for steel bridge performance among different state agencies and owner groups.

In addition, the outcome of the implementation of the model to the steel bridges (NBI-2013) showed that it is possible to classify more than 80% of the deficient bridges and the accuracy is 73.4%. These measures are considered promising, and the author is optimistic about implementing the model for forecasting the probability of a steel bridge having a deficient superstructure throughout its life span; thus, it is possible to determine when the bridge superstructure will reach deficiency levels. Although, the model has a limitation with respect to the miss-rate (FN), one bridge collapse is too many. Therefore, the model is considered conservative, yet the model is safe to use for the intended purpose of classifying deficient bridges. On the other hand, this study does not aim to substitute bridge inspections but complement bridge management.

This analysis contributes to the overall body of knowledge by providing bridge stakeholders with new insights into bridge characteristics and other critical factors that are associated with steel bridge superstructure deficiency. The results provide transportation agencies with several new decision making criteria when considering bridge design alternatives, including the length of bridge span and bridge roadway width. Also, although the main objective of this study is to characterize steel bridge superstructure deficiency, this model may also be used for other purposes such as predicting the probability of a bridge being deficient through its expected life span. Thus, bridge stakeholders may use these probabilities to prioritize bridge maintenance, repairs, and rehabilitation schedules as part of their asset management plan. The best model may also be used to predict the age at which a steel bridge will reach its superstructure deficiency based on the stakeholder's current maintenance plan and traffic level. In addition, the best model may be used by the Federal Highway Administration to forecast future conditions of steel bridge superstructures. This information may serve as guidance to determine funding levels for appropriate agencies.

The scope of this analysis was on steel bridge superstructures. A similar approach can be applied to other bridge material types and other bridge component ratings. The best model was developed based on the NBI-2013 database, and both NBI-2013 and NBI-2014 were used to validate the model. Moreover, future NBI databases may be used to validate the best model. Although the NBI provides a wealth of information related to bridges, it does not contain all the factors that impact bridge deterioration. Other non-included factors are climate data, hydrological data, and bridge maintenance records. Future research must integrate these data with NBI data to have a thorough understanding of bridge deterioration rates.

REFERENCES

America Society of Civil Engineers - ASCE (2017). "2017 Report Card of America's Infrastructure." <https://www.infrastructurereportcard.org/cat-item/bridges/>. (June 2017).

Bektaş, B. A. (2017). "Use of Recursive Partitioning to Predict National Bridge Inventory Condition Ratings from National Bridge Elements Condition Data." *Transportation Research Record: Journal of the Transportation Research Board*(2612), 29-38.

Bektas, B. A., Carriquiry, A., and Smadi, O. (2013). "Using Classification Trees for Predicting National Bridge Inventory Condition Ratings." *Journal of Infrastructure Systems*, 19(4), 425-433.

Berry, M. J. A., and Linoff, G. (2011). *Data mining techniques: for marketing, sales, and customer relationship management*, Wiley Pub, Indianapolis, Ind.

Bu, G. L., J.; Guan, H.; Loo, Y.; and Blumenstein, M. (2015). "Prediction of Long-Term Bridge Performance: Integrated Deterioration Approach with Case Studies." *ASCE*, 29(3), 04014089.

Chou, J.-S., Chiu, C.-K., Farfoura, M., and Al-Taharwa, I. (2011). "Optimizing the Prediction Accuracy of Concrete Compressive Strength Based on a Comparison of Data-Mining Techniques." *Journal of Computing in Civil Engineering*, 25(3), 242-253.

Christie, P. G., Jim; thompson, Jeff; and Well, Chip (2011). "Applied Analytics Using SAS® Enterprise MinerTM Course Notes." SAS Institute Inc.

Contreras-Nieto, C. (2014). "Development of linear models to predict superstructure ratings of steel and prestressed concrete bridges." Dissertation/Thesis, ProQuest, UMI Dissertations Publishing.

DeVille, B., and Neville, P. (2013). *Decision Trees for Analytics Using SAS® Enterprise Miner™*, SAS Institute Inc.

Dreyfus, G. (2005). *Neural networks : methodology and applications*, Berlin. New York: Springer, Berlin. New York.

Farhey, D. (2015). "Operational Structural Performances of Bridge Types by Deterioration Trends." *Journal Of Performance Of Constructed Facilities*, 29(2).

Federal Highway Administration - FHWA (1995). "Recording and Coding Guide for the Structure Inventory and Appraisal of the Nation's Bridges." U.S. Dept. of Transportation, Federal Highway Administration, Washington, D.C.

Hagan, M. T., H. B. Demuth, M. H. Beale, and O. De Jesús. (1996). *Neural network design*. PWS publishing company Boston.

Huang, Y.-H. (2010). "Artificial neural network model of bridge deterioration." *Journal of Performance of Constructed Facilities*, 24(6), 597-602.

Kim, B.-H. (2009). "Support Vector Machine & Classification Using Weka (Brain, Computation, and Neural Learning)." Seoul Nation University.

Kim, Y. J., and Yoon, D. K. (2010). "Identifying Critical Sources of Bridge Deterioration in Cold Regions through the Constructed Bridges in North Dakota." *Journal of Bridge Engineering*, 15(5), 542-552.

King, G., and L. Zeng. (2001). "Logistic regression in rare events data. " *Political analysis*, 9(2), 137-163.

Matthiesen, R. (2010). *Bioinformatics Methods in Clinical Research*, Humana Press, Totowa, NJ, Totowa, NJ.

Milley, A. H. S., James D.;Williams, John S. (1998). "Data Mining and the Case for Sampling." SAS Institute Inc.

Omran, B. A., Chen, Q., and Jin, R. (2016). "Comparison of Data Mining Techniques for Predicting Compressive Strength of Environmentally Friendly Concrete." *Journal of Computing in Civil Engineering*, 0(0), 04016029.

Potts, W. P., Michael J. (2002). "Predictive Modeling Using Logistic Regression Course Notes." S. I. Inc., ed., SAS Institute Inc., Cary, NC.

Pozzolo, A. D., Caelen, O., Johnson, R. A., and Bontempi, G. "Calibrating Probability with Undersampling for Unbalanced Classification." *Proc., 2015 IEEE Symposium Series on Computational Intelligence*, 159-166.

Saeed TU, Moomen M, Ahmed A, Murillo-Hoyos J, and Volovski M, L. S. (2017). "Performance Evaluation and Life Prediction of Highway Concrete Bridge Superstructure across Design Types." *Journal of Performance of Constructed Facilities*.

Schulz, A., Zia, A., and Koliba, C. (2017). "Adapting bridge infrastructure to climate change: institutionalizing resilience in intergovernmental transportation planning processes in the Northeastern USA." *Mitigation and Adaptation Strategies for Global Change*, 22(1), 175-198.

Tang, P., Kanaan, O., Wnag, J., Oh, J., and Kwigizile, V. (2012). "Statistical Identification of Items Importants for Evaluating Bridge Conditions." *Transportation Research Board*.

Tarefder, R. A., Sumee, N., and Storlie, C. (2014). "Study of MEPDG Sensitivity Using Nonparametric Regression Procedures." *Journal of Computing in Civil Engineering*, 28(1), 134-144.

Thoonen, G., Hufkens, K., Borre, J. V., Spanhove, T., and Scheunders, P. (2012). "Accuracy assessment of contextual classification results for vegetation mapping." *International Journal of Applied Earth Observations and Geoinformation*, 15, 7-15.

U.S. Department of Transportation (2013). "Federal Highway Administration National Bridge Inventory." <https://www.fhwa.dot.gov/bridge/nbi/ascii.cfm?year=2013>. (2013).

Veshosky, D., Beidleman, C., Buetow, G., and Demir, M. (1994). "Comparative Analysis of Bridge Superstructure Deterioration." *Journal of Structural Engineering*, 120(7), 2123-2136.

Yuan, S. (2015). "Random gradient boosting for predicting conditional quantiles." *Journal of Statistical Computation and Simulation*, 85(18), 3716-3726.

Zhu, Z., and Brilakis, I. (2010). "Parameter optimization for automated concrete detection in image data." *Automation in Construction*, 19(7), 944-953.

Ziegel, E. R. (2003). Logistic Regression: A Self-Learning Text. *Technometrics*, Vol. 45, No. 1, 2003, p. 109.

CHAPTER IV

PRIORIZATION OF BRIDGE MAINTENANCE BASED ON BRIDGE RATINGS

ABSTRACT

Due to budget constraints of Departments of Transportation (DOTs) and a significant number of deficient bridges around the U.S., there is a need for a systematic approach to more efficiently and optimally allocate limited resources for the bridge maintenance effort. This paper presents a decision-making framework to prioritize bridge maintenance through using aggregated bridge ratings and average daily traffic (ADT). The aggregated bridge ratings were developed by summing weighted deck, substructure, superstructure, and scour ratings; the weights were determined by analyzing a group of focused bridge experts' opinions on the relative importance of deck, substructure, superstructure, and scour with respect to bridge resiliency, riding comfort, safety, and serviceability using the Analytic Hierarchy Process (AHP). In addition, a geographical information system (GIS) user interface that integrated GoogleTM Fusion Tables and Google Maps and incorporated the developed decision-making framework was created to visualize the priorities of the bridges for maintenance. Through case studies and validation with a division bridge engineer at Oklahoma DOT (ODOT), the developed framework was proven to be a robust and reliable approach. This study contributes to the industry practice by providing a systematic approach to facilitate state DOTs' decision-making for bridge maintenance.

INTRODUCTION

Departments of transportation (DOTs) rely on the results of bridge inspections in order to identify the current state of bridges and determine the proper measures to take in order to keep the level of serviceability and safety of those infrastructure assets. As a result, DOTs have developed an assortment of tools for keeping information safe and supporting the decision making process. Maintenance management systems, bridge management systems, and bridge inspections reporting systems are some of the tools implemented by DOTs as Hearn and Johnson (2011) presented.

However, bridge maintenance funds are not enough to repair all the current deficient bridges that include both structurally deficient and functionally obsolete bridges. According to the 2017 ASCE Report Card (ASCE 2017), around 56,000 bridges in the U.S. are classified as structurally deficient. According to the new federal estimate, $123 billion are needed to fix deficient bridges. Nevertheless, budget reductions of DOTs have impacted their available funds. Although spending on bridge projects increased from $11.5 billion in 2006 to $17.5 billion in 2012, it was not adequate to address the rehabilitation of all bridges. Moreover, Oklahoma Department of Transportation's (ODOT) budget has faced more than $190 million in reductions since 2010, and Oklahoma possessed the fifth highest percentage of structurally deficient bridges nationally in 2015 (TRIP 2016). Therefore, the DOTs' need to be more efficient at optimally utilizing those limited resources.

In order to better utilize the fund, prioritization of bridge repairs is needed. This paper presents the development of a new rating system that aggregates the current bridge condition ratings and one appraisal rating by applying the Analytic Hierarchy Process (AHP). The prioritization process takes the new rating and average daily traffic (ADT) in to consideration to rank the urgency of maintenance. In order to facilitate the process of prioritization, geographical information system (GIS) was developed to assist the visualization. As a result of implementing this proposed approach, a ranked list of bridges can be generated, which can be used by the DOTs to determine the priority of maintenance schedule.

BACKGROUND AND LITERATURE REVIEW

The increasing interest in asset management and its applications has captured the attention of transportation agencies around the world because the infrastructure of a nation should be kept in good condition in order to support the economic development, meet recreational and social necessities, boost the public health and safety, and provide for sustainability (Robak et al. 2015). Also, the rate of economic growth and development tends to slow down when reconstruction is required (Munasinghe et al., 1995). As bridges and viaducts are critical to the transportation infrastructure, their maintenance is essential in order to keep their desired level of serviceability and safety. Therefore, some DOTs have established their own bridge priority measures, identifying and tracking bridge maintenance, etc., even though those approaches could be similar among DOTs (Hearn and Johnson 2011). For example, Delaware DOT has a formula to rank bridge services candidates. This formula gives values from 0 to 100, and it integrates 13 inputs (structural condition rating, structural deficiency, benefit-cost ratio, health index, load capacity, width capacity, vertical clearances, waterway adequacy, functional class, detour length, average daily traffic, historical significance, and structural vulnerabilities). However, other DOTs (i.e. Michigan and New York) just use the NBI condition ratings to determine the type of service or to classify bridges, as Hearn and Johnson (2011) stated.

Moreover, researchers have focused on many different topics but mainly in predicting bridge conditions, selecting bridge maintenances by minimizing cost and maximizing benefits, integrating bridge data and geographical information, monitoring bridge condition in real time, and optimizing allocation of maintenance resources. However, although the prioritization of bridge maintenance has been identified as an important and interesting topic, there are not many studies of it. Therefore, more research on prioritizing bridge maintenance should be done in order to find solutions for this problem.

Usually a bridge is considered for maintenance when it reaches a deficient condition. This excludes any car collision with a bridge because, in those cases, bridge repairs are conducted under different regulations. Therefore, several research studies have investigated the prediction of bridge condition ratings, especially deck and superstructure ratings. For example, Morcous et al. (2002) found potential in predicting bridge conditions by implementing a case-based reasoning (CBR) approach using inspection data of bridge decks in Quebec. In another study, artificial neural network models were built with data from concrete bridge decks located in Wisconsin (Huang 2010) to forecast deck ratings. Also, Bektas et al. (2013) developed models to predict deck ratings as well as also substructure and superstructure ratings. It was conducted by using classification and regression trees method and three transportation agencies' bridge inspection databases. Recently, Contreras-Nieto et al. (2016) combined data mining techniques and the National Bridge Inventory (NBI) databases in order to predict superstructure rating for steel and prestressed concrete bridges.

Another issue with bridge management is the limited budgets assigned for infrastructure maintenance. Furthermore, bridge maintenance has to compete for resources with other transportation infrastructure elements, such as pavements, rail lines, and ports. Thus, allocating those restricted funds has become a challenge for almost every transportation agency. It is crucial to optimize the budget for the best interests of the general public. Therefore, studies such as Zhang and Gao (2012) have focused on determining better maintenance alternatives for bridge decks in order to extend expected lifespan and decrease the expected life cycle cost of bridges. They applied the Weibull distribution to model the lifetime of bridge decks, in addition to comparing three different maintenance scenarios. Similarly, Miyamoto et al. (2000) developed a bridge management system which provides a variety of maintenance strategies according to bridge age (i.e. glass cloth + epoxy injection + recovery of cross section at 65 years) by maximizing maintenance results (quality) and minimizing maintenance costs. The authors used Genetic Algorithms in order to solve the optimization problem. Another approach suggested by Elbehairy et al. (2006) presents an optimal solution to select bridge deck repairs with combination of life cycle cost by implementing Genetic Algorithms and Shuffled Frog Leaping. All of these previous efforts are commendable because they all attempted to maximize the benefit and minimize the cost.

Capturing infrastructure data with geographical information and monitoring infrastructure conditions in real time are two other emerging research directions. Although developed countries have created systems to manage bridge inspections (i.e. National Bridge Inventory, Pontis, and Bridgit in the United States), some countries such as Canada are still facing problems in integrating their different bridge system, as Hammad et al. (2007) stated. Therefore, transportation agencies must have suitable systems and applications to maintain all basic information about bridges and bridge inspection records and other supplementary information, such as bridge as-built drawings. Thus, bridge stakeholders can have data to predict and understand bridge deterioration, and optimize maintenance and repairs of bridges. For instance, She and Sarshar (1999) proposed an approach in order to combine geographical information systems (GIS) and bridge inspections in Malaysia. They used five different software packages in order to achieve their objective: Mapinfo Professional, Microsoft Visual Basic, Microsoft Word, AutoCAD, and Microsoft Windows Media Player. In addition, monitoring infrastructure in real time is considered as a breakthrough of management systems. Hu et al. (2013) developed a prototype system that is capable of evaluating the condition of a bridge as well as the network performance by analyzing the data collected by wireless sensor network platform. Some of the sensors included are temperature, strain gauges, and accelerometers.

Overall, studies have concentrated on predicting bridge conditions by using a variety of mathematical and statistical methodologies, selecting maintenance by optimizing funds, implementing GIS to acquire bridge inspection information, and monitoring bridge state in real time. However, two different studies proposed a new approach to prioritize maintenance and rehabilitation of bridges by creating new ratings. The first research developed by Valenzuela et al. (2010) integrated factors such as structural condition, hydraulic vulnerability, seismic risk, and strategic importance of bridges. As a result, an integrated bridge index (IBI) was proposed in order to explain the bridge maintenance relevance in Chile. The other study incorporates the rate and pattern performances with longevity, condition, and durability measures in order to generate an equivalent structural performance of bridges (Farhey 2015). Thus, the results help identify bridges with serious deterioration in order to be scheduled for maintenance. This current study differentiates itself from Valenzuela et al. (2010) and Farhey (2015) because the bridge maintenance prioritization is based on local bridge experts' preferences, current deficient bridge condition, and number of served customers (ADT) at a state label (Oklahoma).

SCOPE

This paper presents a decision-making framework to prioritize bridge maintenance through using aggregated bridge ratings and average daily traffic. The bridge maintenance prioritization considers all deficient bridges in Oklahoma, and they are divided in two types of bridges: water-crossing and non-water-crossing. Water-crossing bridges are built over rivers, lakes, or unstable channels while non-water-crossing bridges are used to cross other types of obstacles such as roads. Therefore, a weighted rating system was created for each bridge group. The difference is that non-water-crossing bridge weighted rating considers the three main condition ratings (deck, superstructure, and substructure) while the water-crossing weighted rating includes scour rating as a complement of the three condition ratings. In addition, to prioritize bridge maintenance, the results of this study also show the relative importance of the different NBI ratings from the point of view of experts in regards to decision-making for the two bridge types.

METHODOLOGY

In order to achieve the objective of this study, the framework of this research can be divided into the following steps:

1. Prepare a bridge dataset that contains information such as bridge ratings (deck, substructure, superstructure, and scour), average daily traffic (ADT), and geographic location.
2. Establish a weighted bridge rating system. The weights were obtained by using an analytic hierarchy process (AHP) to analyze the survey response from bridge experts.
3. Develop a criticality chart of the risk level using the weighted ratings and ADT levels (very high, high, medium, medium-low, and low). Thus, bridge maintenance can be prioritized.
4. Develop a GIS system using Google™ Fusion Tables to assist with visualization.
5. Analyze the maintenance prioritization results (obtained in step 3) in combination of the geographical information to propose optimal bridge maintenance schedules.
6. Validate the appropriateness of this tool and its results by asking a bridge engineer at ODOT to use it and compare the outcome of the previous method and this method.

Data Source

NBI database is used to report the condition of the all nation's bridges over 20 feet to the Congress (FHWA, 1995). Also, this database is considered the most complete database (Farhey, 2015). NBI is one of the main sources of information for many studies related to prediction of bridge ratings, bridge deterioration, bridge maintenance, and bridge management systems as presented in the literature review. Since ODOT's recent goal is to address all deficient bridges by 2020, the focus of the study is all deficient bridges with any of the four ratings lower than 5. The NBI-2014 dataset was used for this study.

Analytic Hierarchy Process

The Analytic Hierarchy Process (AHP) was created by Thomas L. Saaty in the 1980's (Lee et al., 2011). It is based on a set of pairwise comparisons that depend on experts' discernment to evaluate the dominance of preferences among a collection of more than two options. As Brunelli (2014) mentioned, AHP is implemented especially with intangible criteria and alternatives. Therefore, it is especially used to solve choice problems where options are assessed regarding to multiple criteria. Those problems are known as multi-criteria decision making problems. The following three steps described the process flow of implementing the AHP.

1. Determine the weight vector.
2. Determine the pairwise comparison matrices.
3. Rank the alternatives.

Determining the Weight Vector

One goal and a finite collection of options are part of a decision process in which a decision maker chooses the best alternative. The finite collection of alternatives is represented by $X = \{x_1, \dots, x_n\}$. The simple approach to obtain the weight vector is that the decision maker assigns an importance value to each option. These values form the weight vector, $\boldsymbol{w} = (w_1, \dots, w_n)^T$, where w_1 corresponds to the relative importance of option x_1. The best option (ith) is the one whose value w_i is the highest.

However, a better alternative to estimate the priority vector is by pairwise comparisons (matrix **A**) whose final result is a weight vector (***w***) (Brunelli, 2014). The pairwise comparisons are based on the scale used in Saaty (2008) that is summarized in Table 4-1. In this way, numbers represent how important one component is over another with regard to the criterion used to compare them. For example, how much does a person prefer water to a soda while eating lunch? If the number entered is 1, it means that the preference for water and soda during lunch is equal. In contrast, if the number entered is 9, it shows that water consumption is extremely preferred over soda consumption. Although these comparisons form the matrix **A** (shown in Equation 1), it is necessary to perform an additional step to obtain the weight vector (***w***). Moreover, matrix **A** has an important condition, multiplicative reciprocity $a_{ij} = 1/a_{ij} \forall i$. It means that the ratios expressed in Equation (1) can be represented as Equation (2).

Table 4-1. The Fundamental Scale of Absolute Numbers

Intensity of Importance	Definition	Explanation
1	Equal Importance	Two activities contribute equally to the objective
2	Weak or slight	
3	Moderate importance	Experience and judgement slightly favor one activity over another
4	Moderate plus	
5	Strong importance	Experience and judgement strongly favor one activity over another
6	Strong plus	
7	Very strong or demonstrated importance	An activity is favored very strongly over another, its dominance demonstrated in practice
8	Very, very strong	
9	Extreme importance	The evidence favoring one activity over another is of the highest possible order of affirmation

$$\boldsymbol{A} = \left(w_i/w_j\right)_{n\times n} = \begin{pmatrix} w_1/w_1 & w_1/w_2 \cdots & w_1/w_n \\ w_2/w_1 & w_2/w_2 \cdots & w_2/w_n \\ \cdots & \cdots \ \ddots & \vdots \\ w_n/w_1 & w_n/w_2 \ldots & w_n/w_n \end{pmatrix} \quad (1)$$

$$\boldsymbol{A} = \begin{pmatrix} 1 & a_{12} & \cdots & a_{1n} \\ 1/a_{12} & 1 & \cdots & a_{2n} \\ \cdots & \cdots & \ddots & \vdots \\ 1/a_{1n} & 1/a_{2n} & \ldots & 1 \end{pmatrix} \quad (2)$$

Even though different approaches have been developed, the *Geometric Mean Method* is widely used to obtain the weight vector (priority vector), ***w***. This method was proposed by Crawford and Williams (1985) and uses the matrix **A** that contains the ratios between the pairwise comparisons among the alternatives, X. This approach proposes that the components of **w** are determined as the geometric mean of the elements on the specific row divided by a normalization term (Equation 3). As a result, the sum of the components of ***w*** is 1.

$$w_i = {\left(\prod_{j=1}^{n} a_{ij}\right)^{\frac{1}{n}}} \Big/ {\sum_{i=1}^{n}\left(\prod_{j=1}^{n} a_{ij}\right)^{\frac{1}{n}}} \qquad (3)$$

However, in cases where the decision maker cannot develop the matrix **A** due to lack of knowledge or complexity of the case, a technique based on collective intelligence (Surowiecki, 2004) can be implemented. The AHP for group decisions has been successfully used in a variety of fields because of its applicability to solve conflicts. The process to implement this approach is that every member of a group of experts on the topic is asked to perform a pairwise comparison. Then the weight vectors of each pairwise comparison are summarized in one matrix with which the final weight vector is obtained. In this study, the weighted geometric mean formula (Equation 4) was implemented in order to determine the final weight vector, $\boldsymbol{w^G}$ (Brunelli, 2014).

$$w^G = \left(\prod_{h=1}^{m} {w_i}^{(h)^{\lambda_i}}\right) \qquad (4)$$

Where *h*th represents the decision maker, and λ_h is the importance of the *h*th decision maker. This approach aims to aggregate the experts' individual priorities instead of aggregating their individual judgments.

Determining the Pairwise Comparison Matrices

When the preference of the decision maker is influenced by other factors (criteria), the problem becomes a Multi Criteria Decision Making (MCDM) problem. It means that the experts have to evaluate the alternatives based on a set of criteria, which influences the preference of one alternative over the others. Thus, every alternative has to be evaluated under each criterion. As a result, the number of pairwise comparison matrices ($\mathbf{B}^{zi}$) is equal to the number of criteria that compose the set (*zi*). Then, the geometric mean equation is implemented (Equation 3). Finally, the weight geometric mean formula is applied in order to obtain final vectors, $\boldsymbol{w^{G,zi}}$, for each criterion *zi*. This methodology is widely use in a variety of problems, especially when a large number of alternatives or interconnected objectives are involved (Kazakis et al., 2015).

Consistency Ratio

In order to examine whether the assignment of preference by each member of the group of bridge experts is consistently performed, the author evaluated the consistency of the pairwise comparisons by measuring inconsistency. Although various consistency indices have been developed (Brunelli, 2014), the consistency index (Equation 5) and consistency ratio (Equation 4.6) were implemented in this study.

$$CI = \frac{\lambda_{max} - n}{n-1} \qquad (5)$$

Where λ_{max} is the maximum eigenvalue of the pairwise comparison matrix **A**, and n stands for the number of attributes of matrix **A**. As Lee et al. (2000) explained, the consistency index (*CI*) measures how close λ_{max} is to n in proportion to the number of attributes.

$$CR = \frac{CI}{RI} \qquad (6)$$

However, *CI* must be rescaled by dividing it by a real number *RI*, which stands for random index. *RI* values are presented in Table 4-2; those values were created by Saaty by simulating the random allocation of the weights 1-9 to matrices of different size and calculating the average value of the *CI* when no endeavor is made at consistency (Lee et al., 2000). The result of *CR* should be less than 0.1 (Brunelli, 2014). *CR* represents the percentage of inconsistency in the judgements as if they had been given randomly.

Table 4-2. Values of *RI*

n	***3***	***4***	***5***	***6***	***7***
RI	0.525	0.882	1.109	1.248	1.342

Ranking the Alternatives

Following the development and synthesis of matrices **A** and matrices $\mathbf{B}^{zi}$, and calculating the final vectors for both alternatives ($\boldsymbol{w}^{G}$) and alternatives regarding the set of criteria ($\boldsymbol{w}^{G,zi}$), the vector of global scores, $\boldsymbol{v}$, is obtained by multiplying $\boldsymbol{w}^{G}$ and $\boldsymbol{w}^{G,zi}$.

$$v = w^{G} \cdot w^{G,zi} \qquad (7)$$

Where the ith element of the vector, $\boldsymbol{v}$, denotes the weights designated by the AHP to the ith alternative. The best alternative is the one whose score is the highest.

Alternatives

The NBI databases contain three main condition ratings: deck, substructure, and superstructure ratings (FHWA, 1995). These ratings show the current condition of the existing bridges in comparison to the as-built condition. Furthermore, NBI databases include appraisal ratings such as Item 113: Scour Critical Bridges, which records the current state of the bridge regarding its vulnerability to scour. Table 4-3 presents the codes used to categorize the current condition of the deck, substructure, and superstructure in the NBI databases. A similar scale is used for assessing scour rating where 4 or less is considered deficient. The four proposed ratings provide a comprehensive characterization of the whole module. This makes the deck, substructure, superstructure, and scour ratings ideal to develop a new rating that aggregates them and represents the overall current condition of the bridge, which is totally different to the overall sufficiency rating contained in the NBI databases. The sufficiency rating (see Equation 8) consists of four different components: structurally adequacy and safety (S1); serviceability and functional obsolescence (S2); essentiality for public use (S3); and special reductions (S4). The maximum percentage of the components are 55%, 30%, 15%, and 13%, respectively.

Table 4-3. Condition Ratings Code - NBI Databases

Code	Description
N	Not applicable
9	Excellent condition
8	Very good condition
7	Good condition
6	Satisfactory condition
5	Fair condition
4	Poor condition
3	Serious condition
2	Critical condition
1	“Imminent” failure condition
0	Failed condition

$$Sufficiency\ Rating = S1 + S2 + S3 - S4 \qquad (8)$$

Criteria

Because of the type of goal and the complexity of the problem, a set of criteria was established in order for a group of experts to evaluate the alternatives according to each criterion included in this set. The author used the following criteria to compare the importance of the four ratings of a bridge. The definitions of the four criteria are as follows.

1. Safety
   It is related to whether or not users can use the bridge without putting their life at risk
2. Serviceability
   It is understood as the bridge meets the objective to what it was built.
3. Comfort
   It stands as the customers’ satisfaction of using /riding on the bridge.
4. Resiliency
   It refers to the ability of the bridge to absorb catastrophic impacts (Natural: tornado, earthquakes, etc. Man-made: collisions) with timely returns to normalcy.

Bridge Experts Selection and Survey

Deciding on the bridge experts to make the pairwise comparison was as critical as selecting the approach to achieve this study. Therefore, the author relied on the experience and mastery of bridge and maintenance engineers and also bridge inspectors of ODOT. Moreover, in order to complement the group of experts, the researcher invited professors who teach material courses and structural design. Thus, this study can consider and merge a variety of priorities (preferences) from each member of the expert group. Although 19 experts were invited to complete the survey, just 11 questionnaires were filled out, which corresponds to a 60% answer rate.

Due to the long extension of the survey, a partial view of the questions (comparisons) is presented in Figure 4-1 and Figure 4-2. The respondents were asked to mark their priorities according to Saaty's scale (Table 4-1) for the set of criteria (Figure 4-1) and the set of alternatives based on the four criteria. The comparisons among the four NBI ratings based on Safety can be seen in Figure 4-2. The whole survey document and some complete surveys have been attached in the appendix.

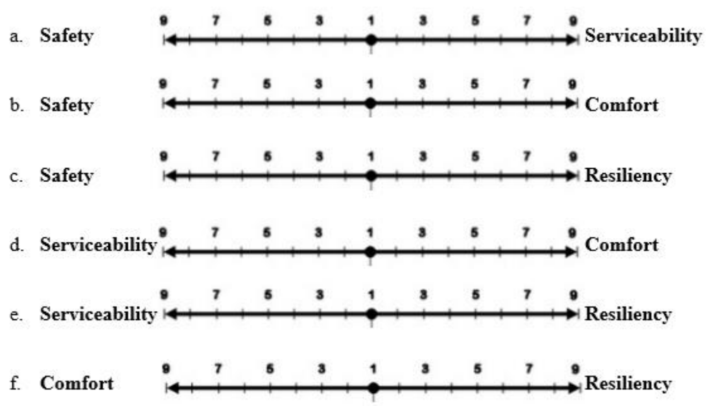

Figure 4-1. Criteria Comparisons

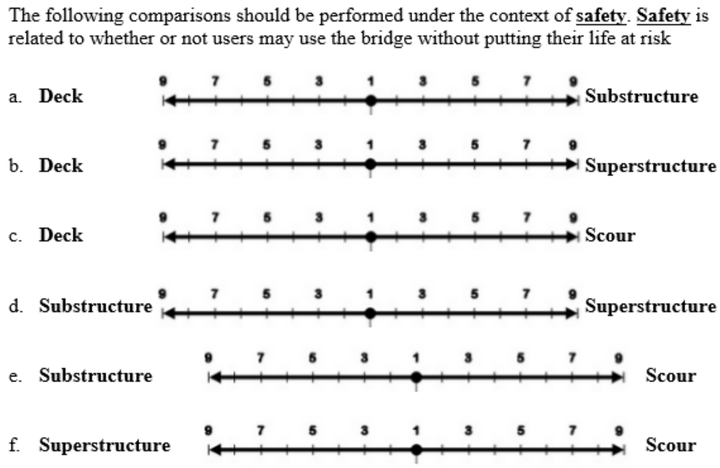

Figure 4-2. Bridge Ratings Comparisons based on Safety

AHP Representation

Figure 4-3 presents the AHP for water-crossing bridges. The top level represents the goal of implementing this technique, intermediate level indicates the criteria, and the bottom level shows different alternatives.

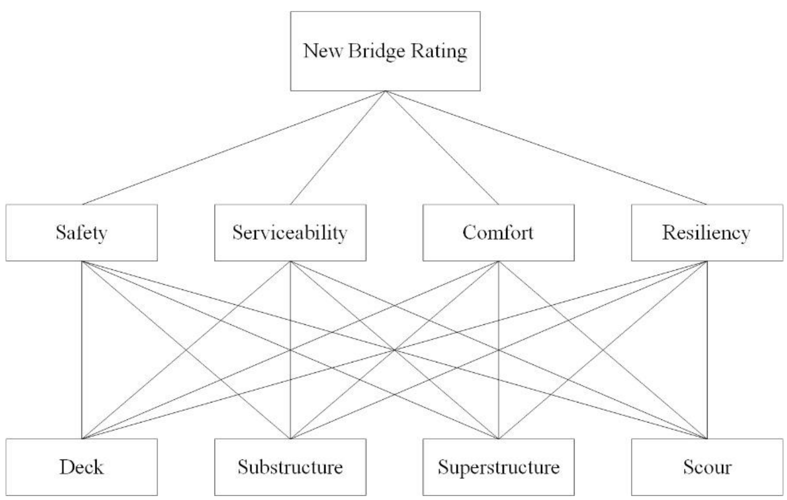

Figure 4-3. Analytic Hierarchy Process of the Research

As a result of applying the AHP methodology, the author’s expectation was an equation with the weight of each of the current bridge condition ratings as Equation 9 represents, which is the equation for water-crossing bridges. For non-water-crossing bridges, the equation does not have the scour component. The range of the weighted rating is from zero (0) to nine (9) for all bridges.

$$Merged\ Rating = \alpha(Deck) + \beta(Substructure) + \gamma(Superstructure) + \delta(Scour) \quad (9)$$

Geographical Interface

Based on the idea of GIS (Geographical Information System) that aims at the integration of large amounts of data and geospatial information (Romps 2008), GIS applications have been implemented in many fields, including asset management and operations. As Cerkas (2008) states, GIS is evolving fast with good results for utility management. For example, one application is the visual age analysis of field facilities in which the age or the life span of utilities can be displayed by colors or symbols.

Because of GIS' capability, it can be implemented in order to keep a great inventory of all the infrastructure assets, in this case, bridges. Also, the database could record information about the maintenance/repairs, such as labor hours, costs, and quantities. Therefore, bridge stakeholders (i.e. DOT's personnel) can communicate facts of each bridge among one another; similarly, the natural gas industry (NGI) uses GIS to keep the inventory of all their assets combined with other information such as tax rates so that essential financial information can be determined (Romps 2008). In addition, Ekawati and Suharjito (2016) stated that using GIS could be considered as a thematic map due to two main reasons: first, there is a visualization of the data with respect to a geographical location, and second, users can analyze and differentiate data with another location or record.

Although many methods have been developed to implement data visualization using GIS data, Google™ created a service pack called *Fusion Tables*. Fusion tables was introduced in summer 2009 (Signore 2016). This service aims to manage and integrate databases saved on the cloud in order to improve the collaboration among parts (Gonzalez et al. 2010). Also, it is possible to develop a thematic map and share data by creating links; therefore, it could be considered as an effective tool for publishing data.

In order to illustrate a simple representation and the capability of implementing Fusion Tables and Google Maps, Figure 4-4 shows bridges located in the state of New York. In addition, the bridges displayed in the figure have been filtered by their average daily traffic (ADT); just bridges with ADT values between 10,000 and 32,200 are exhibited. Furthermore, a summary of data related to a bridge is displayed in a call-out box shown in Figure 4-4. The developer of the thematic map can customize the information that is dispplayed for the users. However, this is just a fragment of what can be developed using these applications and what this research is aiming to accomplish.

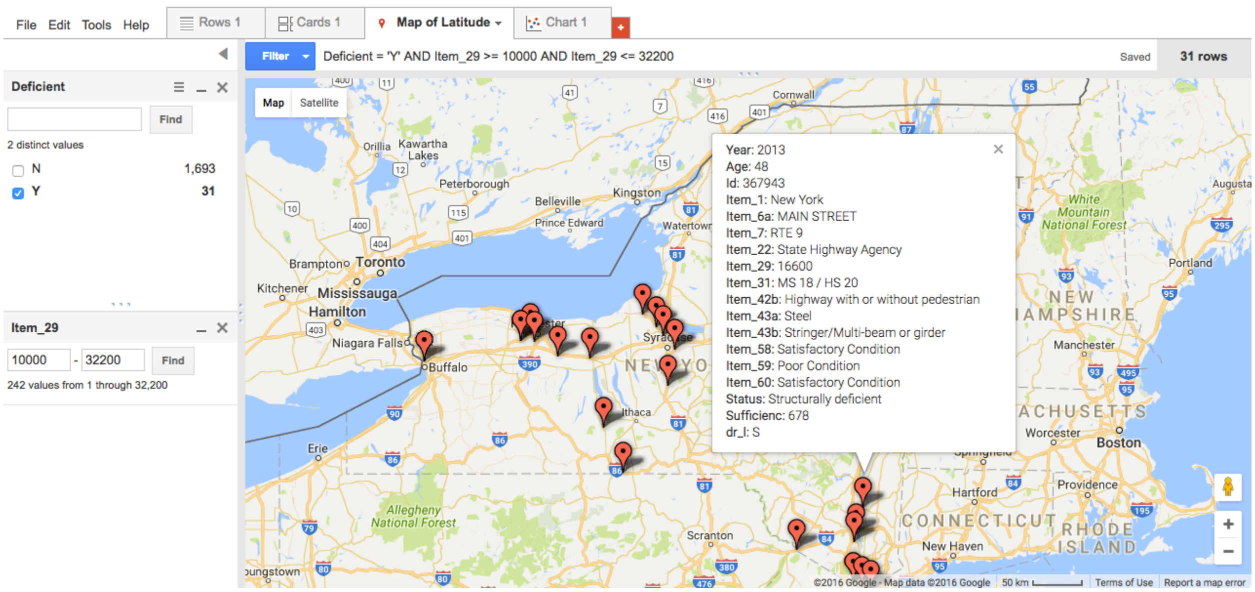

Figure 4-4. Fusion Table and Google Map Representation of New York Bridge

RESULTS

This section presents the results of the analysis and detailed description about both AHP methodology and bridge maintenance prioritization involve in the analysis. Using Fusion Tables, some figures were developed in order to represent the outcome of the research.

Group AHP Weighting Method

Based on the 11 experts' responses, the AHP methodology was implemented in order to obtain the rank of alternatives. The first step was to create the comparison matrices and calculate the weight vectors for each expert as it is presented in Table 4-4. A total of 99 matrices were built.

Table 4-4. Expert 1 – Comparison Matrix of NBI Condition Ratings by Resiliency Criteria

	Deck	Superstructure	Substructure	Scour	Weight Vector
Deck	1	1/3	1/5	1/3	0.084
Superstructure	3	1	1	1/3	0.217
Substructure	5	1	1	1	0.324
Scour	3	3	1	1	0.375

In addition, the consistency ratio, CR (Equation 6), was determined for each comparison matrix. As a result, some comparisons were discarded because their CR values were higher than 0.1, which is the threshold recommended and used in previous studies (Dong and Cooper, 2016; Lee et al., 2011). Table 4-5 summarizes the CR values for each matrix by the expert's identifier (ID) for water-crossing bridges. The values followed by an asterisk (*) are the discarded comparisons. Therefore, between 7 and 9 comparisons were used to obtain the priority vector for each criterion. To clarify, while 7 comparisons were used to determine the priority vector of comparisons based on *resiliency*, 9 comparisons were used for *safety*.

Table 4-5. Consistency Ratios (CR)

ID	CRITERIA	SAFETY	SERVICIABILITY	COMFORT	RESILENCY
1	0.10	0.00	0.00	0.00	0.00
2	0.10	0.13*	0.13*	0.24*	0.08
3	0.19*	0.09	0.00	0.12*	0.17*
4	0.04	0.22*	0.05	0.00	0.13*
5	0.00	0.06	0.02	0.08	0.02
6	0.00	0.05	0.03	0.09	0.02
7	0.13*	0.07	0.00	0.00	0.00
8	0.08	0.05	0.01	0.08	0.08
9	0.01	0.00	0.02	0.08	0.09

ID	CRITERIA	SAFETY	SERVICIABILITY	COMFORT	RESILENCY
10	0.13*	0.08	0.06	0.00	0.70*
11	0.26*	0.00	0.13*	0.13*	0.13*

Then using the weighted geometric mean (Equation 4), the final weight vector, $\boldsymbol{w}^G$, was calculated for each criterion (safety, serviceability, comfort, and resiliency). To illustrate the process, Table 4-6 presents the experts' individual priorities, w_i, and the final weight vector ($\boldsymbol{w}^G$) for the *serviceability* criterion. Because scour rating obtained the highest coefficient (0.289), it is possible to conclude the group of experts considered scour rating as the most important rating regarding the *serviceability* criterion. However, superstructure and substructure ratings were ranked as second and third with 0.274 and 0.248 as weights, respectively. Finally, the group of experts do believe that deck rating is the least important rating regarding *serviceability*; and its weight is 0.189.

Table 4-6. Final Weight Vector - Serviceability

Expert / Rating	w_1	w_2	w_3	w_4	w_5	w_6	w_7	w_8	w_9	w^G
Deck	0.084	0.250	0.087	0.139	0.570	0.250	0.128	0.431	0.035	**0.189**
Superstructure	0.217	0.250	0.154	0.166	0.266	0.250	0.142	0.380	0.413	**0.274**
Substructure	0.324	0.250	0.180	0.197	0.090	0.250	0.347	0.117	0.314	**0.248**
Scour	0.375	0.250	0.579	0.498	0.073	0.250	0.384	0.073	0.238	**0.289**

Similarly, final weight vectors ($\boldsymbol{w}^G$) for the comparisons of each criterion were calculated (see Table 4-7). For non-water-crossing bridges, it is observed that deck was the most important rating for *safety*, *serviceability*, and *comfort* after merging all the experts' priorities as it is presented in Table 4-7. To highlight, considering *comfort* criterion, deck rating obtained 77.1% of importance while superstructure and substructure ratings importance are very low, 15.9% and 7.0%, respectively. However, a drastic change of this trend is noticed when *resiliency* criterion is analyzed. Deck rating descended to the last position, and substructure rating, which was the least important rating in the other three criteria, jumped to the top becoming the most relevant rating when *resiliency* is the factor of comparison. Finally, superstructure rating was ranked as the second more important rating throughout all four criteria.

Table 4-7. Final Weight Vectors of the Alternatives Set

	Non-water-crossing Bridges				Water-crossing Bridges			
Final Weight Vector / Rating	$w^{G,Safet}$	$w^{G,Serviceabil}$	$w^{G,Comfor}$	$w^{G,Resiliency}$	$w^{G,Safet}$	$w^{G,Serviceabilit}$	$w^{G,Comf}$	$w^{G,Resiliency}$
Deck	**0.401**	**0.389**	**0.771**	0.146	0.160	0.189	**0.548**	0.074
Superstructure	0.354	0.364	0.159	0.409	0.261	0.274	0.188	0.299
Substructure	0.245	0.247	0.070	**0.445**	0.287	0.248	0.120	**0.381**
Scour	N/A	N/A	N/A	N/A	**0.292**	**0.289**	0.145	0.246

On the other hand, for water-crossing bridges, the relative importance of the ratings appears different when compared to non-water-crossing bridges. When *safety* and *serviceability* criteria are considered, scour rating became the most important and deck rating ranked the least, which is the opposite for non-water-crossing bridges. Nevertheless, deck and substructure ratings continued being the most important ratings when *comfort* and *resiliency* are valued, respectively. In general, the importance of superstructure rating did not change across both bridge types.

Then, final weight vectors ($\boldsymbol{w}^{\boldsymbol{G}}$) for comparisons among the criteria set (safety, serviceability, comfort, and resiliency) were determined, and the results are presented in Table 4-8. It is observed that *safety* criterion is considered the most important among the other three criteria; its importance coefficient is 47.5%. In order of importance, the other three criteria were ranked as follows: *resiliency* (25.1%), *serviceability* (20.5%), and *comfort* (7.0%).

Table 4-8. The Weight Vector of the Criteria Set

Criterion	$\boldsymbol{w}^{\boldsymbol{G}}$
Safety	0.475
Serviceability	0.205
Comfort	0.070
Resiliency	0.251

Finally, the weights (v) were calculated according to Equation 7, and the results are summarized in Table 4-7 and Table 4-8.Thus, Figure 4-5 summarizes the weights of the NBI ratings based on experts' opinions, and these priorities can be ranked by analyzing their coefficients. With regard to non-water-crossing bridges, it is possible to say that deck and superstructure ratings hold position one and two, even though their weights are very close. Substructure is in the final position with the lowest weight (28.3%). However, water-crossing bridges priorities differ from the other type of bridges. The main difference is that deck rating ranked on the bottom with the lowest weight of 0.171. Although substructure rating became the most relevant rating for prioritizing water-crossing bridges, their weights were almost the same at 0.291 and 0.283. In addition, it is possible to say that scour and superstructure ratings have the same importance for prioritizing water-crossing bridges.

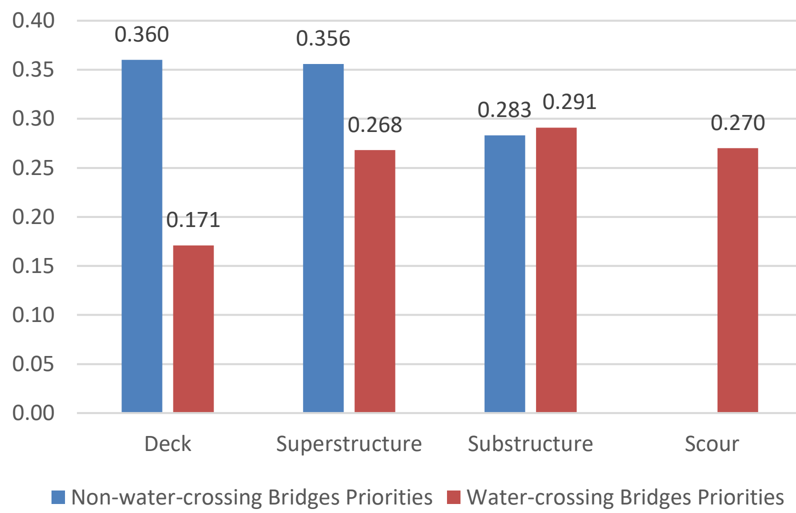

Figure 4-5. Final Weights for the NBI Ratings

Aggregated Rating

By using Equation 9 with the three/four weights (Figure 4-5) and the four NBI ratings, a new rating was calculated for each deficient bridge contained in the Oklahoma's NBI-2014 database. Table 4-9 summarizes the number of deficient bridges by interval of the new rating, weighted rating (WR). The highest numbers of deficient bridges in Oklahoma have WR values between 4 and 7.

Table 4-9. Number of Bridges by Aggregated Rating (WR) Interval

WR Interval	Frequency	Percentage	Cum. Percentage
0-1	10	0.3%	0.3%
1-2	24	0.8%	1.1%
2-3	144	4.8%	6.0%
3-4	410	13.7%	19.7%
4-5	1119	37.4%	57.1%
5-6	1173	39.3%	96.4%
6-7	108	3.6%	100.0%

Average Daily Traffic (ADT) Classes

ADT is a typical factor used by DOTs (Delawere, Oregon, and Washington) to prioritize maintenance (Hearn and Johnson, 2011). Based on the three ADT classes defined in Item – 67 – Structural Evaluation: 1) 0-500, 2) 501-5,000, and 3) >5,000 vehicles per day (FHWA, 1995), the author proposed five ADT classes instead (Low, Low-Medium, Medium, High, and Very High) as can be seen in Table 4-10. The main reason of increasing the number of ADT classes is that ranges used in Item 67 are so wide, and thus more classes needed to be defined. In addition, the priority of the ADT classes proposed is based on the idea of serving as many customers as possible.

Table 4-10. ADT Classes

ADT Class	ADT Range	Priority
Low	0 – 500	5
Low-Medium	501 – 2,000	4
Medium	2,001 – 5,000	3
High	5,001 – 15,000	2
Very High	>15,000	1

Proposed Bridge Maintenance Prioritization

Because this proposed prioritization is a combination between WR and ADT, the highest priority for maintenance is for bridges whose WR is low and ADT is very high. However, it was necessary to create groups and subgroups to achieve the research objective. Figure 4-6 contains the proposed groups and subgroups that are used to prioritize deficient bridges in Oklahoma; this is similar to Alroomi et al. (2012) who presented their criticality matrix. Group I has the highest priority over the rest of the four groups (II, III, IV, and V). In addition, each group has been further segmented into three or more subgroups (i.e. a, b, c) in order to recognize the relevance of bridges with higher ADTs and lower WR within the same group. For example, bridges that fall into subgroup II-a will have higher priority than bridges in group II-c. However, in order to serve as many customers (taxpayers) as possible, the groups and subgroups were defined considering the ADT level. As a result, subgroup IV-c was defined to be in group IV instead of in group V. Similarly, although subgroup IV-c has higher WR than subgroup IV-d, the customers that can be served in subgroup IV-c are substantially higher than customers in subgroup IV-d.

Moreover, an additional group named ‘Replacement’ was suggested. It contains bridges with WR equal to or lower than 3, and they may be considered for replacement or reconstruction. According Jebreen and Johnston (1995), 3 may considered as the lowest level of service of bridges. It means that lower conditions (<3) could not be improved to a satisfactory state. The reason of this suggestion is that it could be more economical to replace the whole bridge than repair the current structure in order to bring it to good condition as suggested by Higuchi and Macke (2008). Nevertheless, this group should be studied in depth in future research.

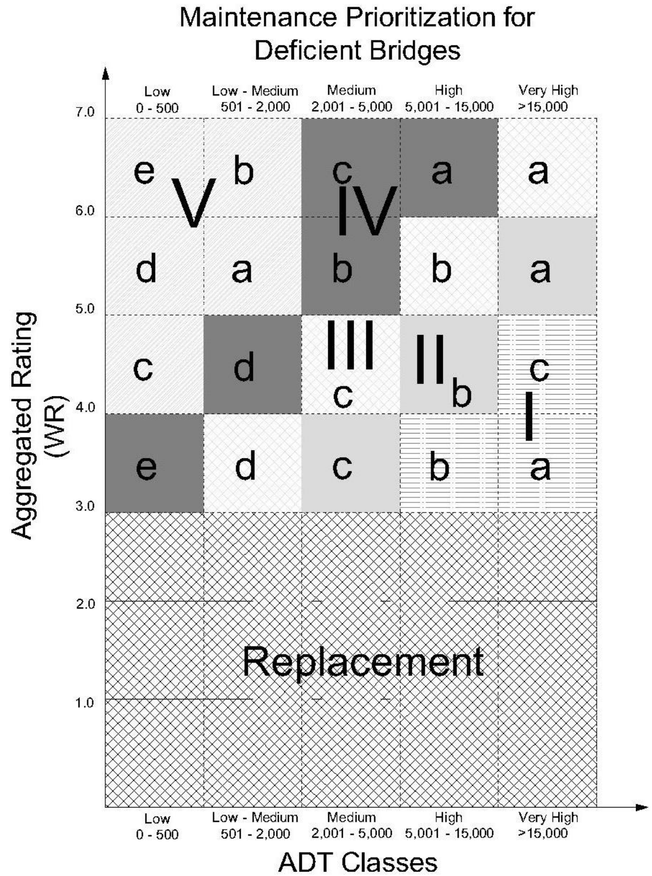

Figure 4-6. Maintenance Prioritization for Deficient Bridges

As a result of implementing this approach, the deficient bridges in Oklahoma were ranked. Table 4-11 contains the top 5 deficient bridges in the state. These five bridges have some similarities in their characteristics. For example, ODOT is responsible of their maintenance because they are on the state highway system. Also, the superstructure material of all five bridges is steel. In addition, the range of year built of these five bridges is between 1958 to 1964. Finally, the top three bridges are non-water-crossing bridges while the other two bridges are water-crossing.

Table 4-11. Top 5 of Deficient Bridges in Oklahoma

Structure number	County	Year Built	ADT	Deck	Sup	Sub	Scour	WR	Group	Sub group	Rank
151790000000000	109	1960	39,000	3	4	4	N	3.649	I	a	1

Structure number	County	Year Built	ADT	Deck	Sup	Sub	Scour	WR	Group	Sub group	Rank
151160000000000	125	1960	17,450	3	5	4	N	3.994	I	a	2
153620000000000	143	1961	10,800	3	5	3	N	3.690	I	b	3
141110000000000	143	1958	10,450	5	4	3	3	3.610	I	b	4
160360000000000	143	1964	10,300	5	5	3	3	3.878	I	b	5

On the other hand, all bridges classified into the category ‘replacement’ are not in the state highway system; therefore, they are the responsibility of County Highway Agency or City/Municipal Highway Agency instead of ODOT. To highlight, the range of ADT of these bridges goes from 1 to 750 vehicles per day, which could be classified as low or low-medium categories (according to Table 4-10).

Geographical Interface – Google™ Fusion Tables

To better assist the prioritization process, a visualization interface was created using Google™ Fusion Tables. Thus, it is possible to visualize the location of all deficient bridges. In addition, Fusion Tables allows the user to create filters based on the information contained in the table. Figure 4-7 shows the location of the top 50 deficient bridges. Using this interface, bridge stakeholders are able to quickly make decisions based on the geographical location and the aggregated ratings of the deficient bridges. Also, bridge maintenance engineers can group maintenance schedules for bridges that are adjacent to each other. For example, if ODOT’s decision is to schedule maintenance for all bridges located on I-35, it is as simple as filtering deficient bridges by location (I-35) as can be seen in Figure 4-8.

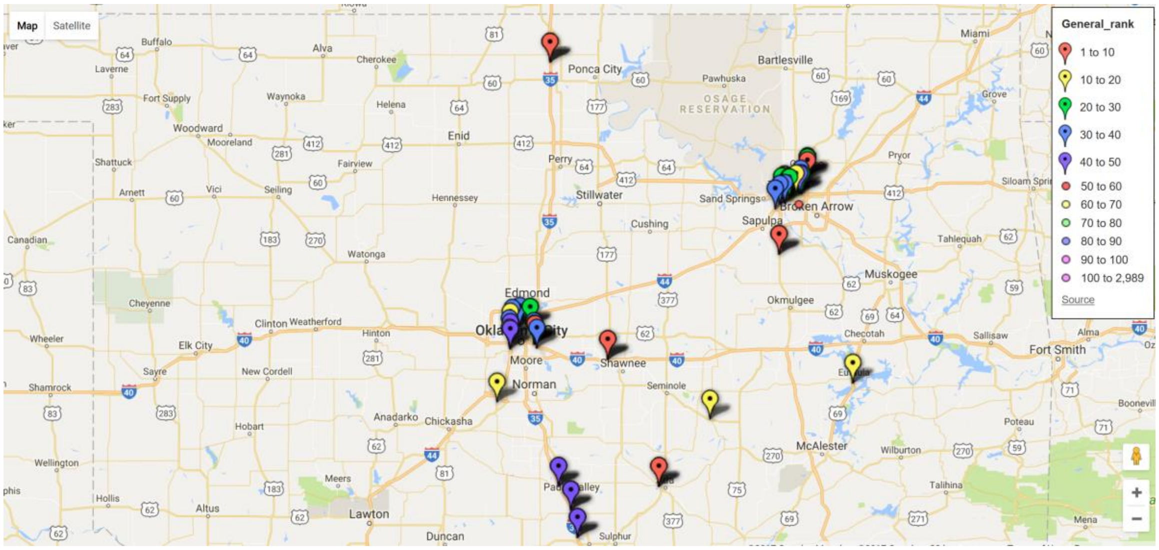

Figure 4-7. Deficient Bridges Top 50 – Google™ Fusion Tables

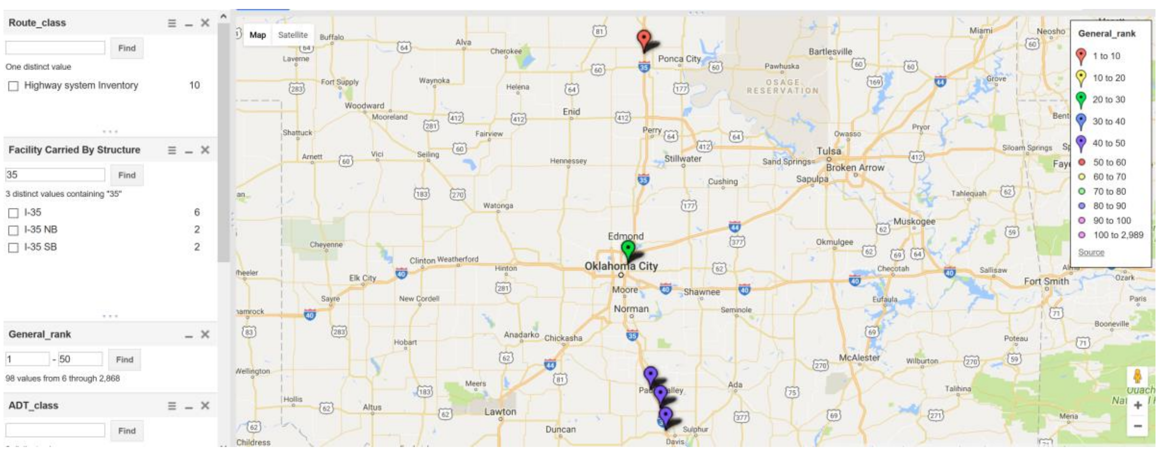

Figure 4-8. Deficient Bridges Located on I-35 – Google™ Fusion Tables

Geographical Interface – Google™ Fusion Table

Furthermore, as an additional validation, a case study was conducted. The bridges of this case study are located in Division 4 of ODOT (Canadian, Garfield, Grant, Kay, Kingfisher, Logan, Noble, Oklahoma, and Payne – Figure 4-9). The current condition of the bridges was provided by ODOT because it is the institution in charge of reporting the state of Oklahoma's bridges to the FHWA.

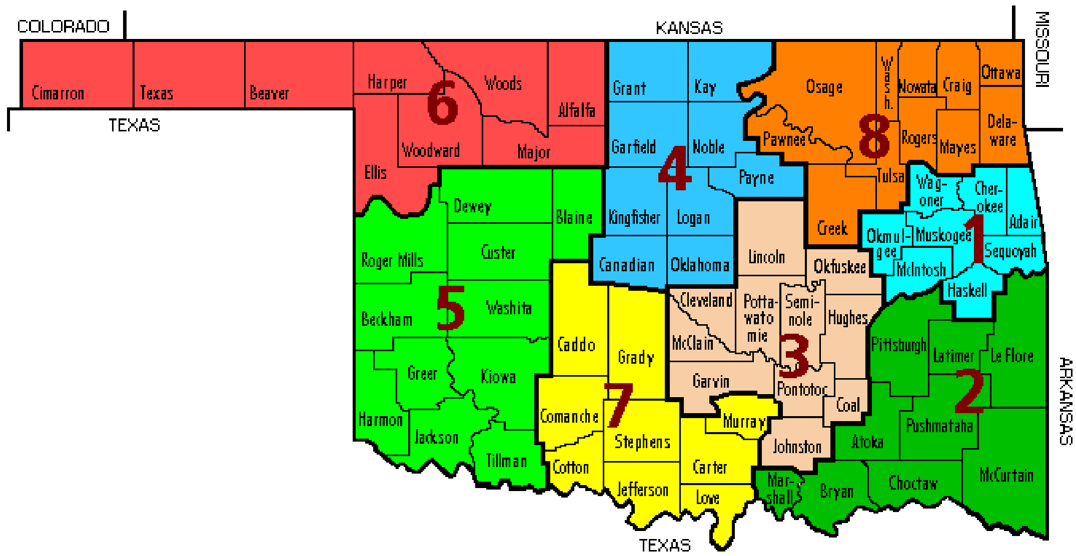

Figure 4-9. ODOT Divisions. ODOT, 2017 (https://www.ok.gov/odot/About_ODOT/Contact_ODOT_Divisions/Field_Divisions.html)

By comparing the Division 4 bridge maintenance ranking with the results of this study, it was determined that the all top 5 bridges (Table 4-11) are in the ODOT Division 4's top 10 bridges, as it is summarized in Table 4-12. Moreover, both rankings match the bridges ranked as 1 and 5. The difference of both ratings is based on the following considerations.

1. ADT is taken into consideration in this study while it is not used by ODOT.
2. ODOT gives high priority to deficient bridge superstructures because those bridges can be fixed fast.
3. When a deficient bridge with high traffic (high ADT) needs maintenance, ODOT look for the best time to schedule for maintenance in order to not affect the traffic.
4. Political importance is another factor that can change the maintenance priority of bridges at ODOT.

Table 4-12. Comparison of Rankings

ODOT's Rank	Proposed Rank	NBI	County	Structure#	Facility Carried	Group	Sub-group
1	1	15179	Oklahoma	5568 0245SX	I-40EB	I	a
2	4	15123	Oklahoma	5568 0226SX	I-40EB	I	c
3	2	10566	Canadian	0904 0690X	I-40 Business	I	b
4	25	5047	Logan	4206 1442X	SH-33	II	b

ODOT's Rank	Proposed Rank	NBI	County	Structure#	Facility Carried	Group	Sub-group
5	5	18610	Oklahoma	5507 0347NXR	I-44	I	c
6	30	13685	Kingfisher	3716 1138X	SH-51	III	d
7	3	13932	Garfield	2405 0092X	US 81	I	b
8	24	14408	Kay	3625 0698WX	I-35	II	b
9	16	21356	Oklahoma	5515 0566EX	I-35NB	II	a
10	41	4085	Canadian	0902 0000X	US 281	V	a

In addition, a meeting with the Division 4 Bridge Engineer was held in which the methodology, the ranking, and the graphical interface were presented. The comments were all positive and encouraging. Because of the simplicity and completeness of the methodology, the Bridge Engineer thinks that it could be useful for his division, and he is interested in implementing it for the bridge maintenance prioritization for this coming year.

CONCLUSIONS AND RECOMMENDATIONS

As a considerable number of bridges are reaching deficiency in combination with restricted funds available for bridge maintenance, DOTs have to invest their limited budget efficiently. Moreover, as Oklahoma is one of the top five states with the highest percentage of structurally deficient bridges in the nation, Oklahoma's goal is to eliminate those bridges by 2020. Prioritizing bridge maintenance based on NBI ratings and numbers of customers severed (ADT) is one solution to battle the shortage of maintenance funds. Departing from the current body of knowledge in bridge maintenance prioritization, this study examined the preferences of a group of bridge experts on the relative importance of bridge ratings (deck, superstructure, substructure, and scour) with respect to a set of criteria (bridge resiliency, riding comfort, safety, and serviceability) by implementing AHP. As a result, a weighted rating was calculated for each deficient bridge. This measure was used in combination with the level of ADT to rank the priority of bridge maintenance. The author used NBI-2014 database as the main source of bridge ratings. Weights of the four ratings were obtained for the two groups of bridges: non-water-crossing and water-crossing.

Through the analysis of the survey of bridge experts in Oklahoma, it was found that deck is the most critical rating when *safety* (0.401), *serviceability* (0.389), and *comfort* (0.771) are concerned for non-water-crossing bridges. With respect to the *resiliency* criterion, substructure (0.445) was given the highest importance. In contrast, scour was the most important rating for *safety* (0.292) and *serviceability* (0.289) for water-crossing bridges. Similar to the results of non-water-crossing bridges, higher preferences were demonstrated for the *comfort* and the *resiliency* criteria for the deck and substructure of water-crossing bridges; deck received the *comfort* weight of 0.548, and substructure received the *resiliency* weight of 0.381. In addition, the four criteria ranked by the magnitude of relevance from strongest to weakest were: (1) *safety*, (2) *resiliency*, (3) *serviceability*, and (4) *comfort*. Finally, the weights of the four ratings were obtained for the two types of bridges. Water-crossing bridges received the following: 0.171 (deck), 0.268 (superstructure), 0.291 (substructure), and 0.270 (scour). Non-water-crossing bridges received the following: 0.360 (deck), 0.356 (superstructure), and 0.283 (substructure). This paper contributes to the overall body of knowledge by determining the importance of the four ratings and creating an equation to calculate a weighted rating for each type of bridge. Combining the ADT classes and the weighted rating, deficient bridges were ranked. Thus, ODOT can implement the rank developed in order to prioritize bridge maintenance of bridges in the state highway system. In addition, County Highway Agency or City/Municipal Highway Agency could also use this prioritization process to schedule the maintenance of deficient bridges under their responsibility. The developed GIS interface can easily visualize the information and facilitate the decision-making process.

Although the scope of this study is focused on deficient bridges in Oklahoma, this decision-making framework could be replicated in any other state. The following recommendations are suggested for future studies. First, the four bridge ratings may be complemented with other appraisal ratings than scour. Future studies can examine whether or not these appraisal ratings are considered relevant to prioritize bridge maintenance. Second, this framework can be replicated in other states. The results of relative importance of the ratings can be compared by states to reveal any discrepancy among groups of local bridge experts. Last, future research can examine the ‘replacement’ group proposed in this study, so the threshold value for the decision of replacement can be determined.

REFERENCES

Alroomi, A., Jeong, D. H. S., and Oberlender, G. D. (2012). "Analysis of Cost-Estimating Competencies Using Criticality Matrix and Factor Analysis." *Journal of Construction Engineering and Management*, 138(11): 1270-1280.

America Society of Civil Engineers - ASCE (2017). "2017 Report Card of America's Infrastructure." <https://www.infrastructurereportcard.org/cat-item/bridges/>. (June 2017).

Bektas, B. A., Carriquiry, A., and Smadi, O. (2013). "Using Classification Trees for Predicting National Bridge Inventory Condition Ratings." *Journal of Infrastructure Systems*, 19(4), 425-433.

Brunelli, M. (2014). *Introduction to the analytic hierarchy process*, Cham : Springer.

Cerkas, M. (2008). "GIS: The New Utility Infrastructure." *Transmission & Distribution World*, 60(3), 20.

Contreras-Nieto, C., Lewis, P., and Shan, Y. (2016). "Developing Predictive Models of Superstructure Ratings for Steel and Prestressed Concrete Bridges." *Construction Research Congress 2016*, 859-868.

Crawford, G., and Williams, C. (1985). "A note on the analysis of subjective judgment matrices." *Journal of Mathematical Psychology*, 29(4), 387-405.

Dong, Q. and Cooper, O. (2016). "A peer-to-peer dynamic adaptive consensus reaching model for the group AHP decision making." *European Journal of Operational Research*, 250(2): 521-530.

Ekawati, R., and Suharjito (2016). "API Fusion Tables and Google Maps Integration for GIS Thematic Mapping Visualization." *International Journal of Software and its Applications*, 10(2), 77-92.

Elbehairy, H., Elbeltagi, E., Hegazy, T., and Soudki, K. (2006). "Comparison of Two Evolutionary Algorithms for Optimization of Bridge Deck Repairs." *Computer-Aided Civil and Infrastructure Engineering*, 21(8), 561-572.

Farhey, D. (2015). "Operational Structural Performances of Bridge Types by Deterioration Trends." *Journal Of Performance Of Constructed Facilities*, 29(2).

Federal Highway administration (FHWA). (1995). *Recording and Coding Guide for the Structure Inventory and Appraisal of the Nation's Bridges*. (Publication FHWA-PD-96-001). Washington, D.C: U.S. Dept. of Transportation, Federal Highway Administration.

Gonzalez, H., Halevy, A., Langen, A., Madhavan, J., McChesney, R., Shapley, R., Shen, W., and Goldberg-Kidon, J. (2010). "Socialising Data with Google Fusino Tables." IEEE computer Society, 25-32.

Hammad, A., Yan, J., and Mostofi, B. (2007). "Recent Development of Bridge Management Systems in Canada " *Transportation Association of Canada.* Saskatoon, Saskatchewa.

Hearn, G., and Johnson, B. (2011). "Identification and Tracking of Highway Bridge Maintenance." *Transportation Research Record: Journal of the Transportation Research Board*, 2220, 12-20.

Higuchi, S., and Macke, M. (2008). "COST-BENEFIT BASED OPTIMIZATION OF MAINTENANCE INTERVENTIONS FOR DETERIORATING STRUCTURES." *JSCE Journal of Structural and Earthquake Engineering*, 63(4), 727-743.

Hu, X., Wang, B., and Ji, H. (2013). "A Wireless Sensor Network-Based Structural Health Monitoring System for Highway Bridges." *Computer-Aided Civil and Infrastructure Engineering*, 28(3), 193-209.

Huang, Y. (2010). "Artificial Neural Network Model of Bridge Deterioration." *Journal of Performance of Constructed Facilities*, 24(6), 597-602.

Jebreen, M. I., and Johnston, D. W. (1995). "Bridge Maintenance Level of Service Optimization Based on an Economic Analysis Approach." North Carolina State University, Raleigh, NC.

Kazakis, N., Kougias, I., and Patsialis, T. (2015). "Assessment of flood hazard areas at a regional scale using an index-based approach and Analytical Hierarchy Process: Application in Rhodope–Evros region, Greece." *Science of the Total Environment*, 538, 555-563.

Lee, D., McCool, J., & Napieralski, L. (2000). Assessing adult learning preferences using the analytic hierarchy process. *International Journal of Lifelong Education*, 19(6), 548-560. doi:10.1080/02601370050209078

Lee, L.-S., Chang, L.-T., Lai, C.-C., & Lin, K.-Y. (2011). "Using The Analytical Hierarchy Process To Construct Performance Indicators For Comprehensive High Schools In Taiwan." *Social Behavior and Personality* 39(5): 615-626.

Miyamoto, A., Kawamura, K., and Nakamura, H. (2000). "Bridge Management System and Maintenance Optimization for Existing Bridges." *Computer-Aided Civil and Infrastructure Engineering*, 15(1), 45-55.

Morcous, G., Rivard, H., and Hanna, A. (2002). "Case-Based Reasoning System for Modeling Infrastructure Deterioration." *Journal of Computing in Civil Engineering*, 16(1), 104-114.

Munasinghe, M., Clarke, C. L., World, B., and International Decade for Natural Disaster, R. (1995). *Disaster prevention for sustainable development : economic and policy issues*, Washington, DC : International Decade for Natural Disaster Reduction and World Bank, Washington, DC.

Oklahoma Department of Transportation - ODOT (2016). "Field Division 4 - Headquarters." <https://www.ok.gov/odot/About_ODOT/Contact_ODOT_Divisions/Division_4.html>. (2016).

Robak, A., Monkman, D., Gilberston, T., Bush, S., O'Connor, B., Blakemore, R., Garlik, S., Hughes, J., Gordon, M., Sharman, B., Smith, M., Maughan, A., Onley, D., Belgrove, T., Waugh, R., Holland, G., Roberts, L., Way, P., Dever-Tod, K., Lutchman, R., and Holdt, C. v. (2015). *IIMM International Infrastructure Management Manual 2015*, Institute of Public Works Engineering Australasia (IPWEA).

Romps, N. G. (2008). "GIS: Mapping the Industry's Future." *American Gas*, 90(2), 15-17.

Saaty, T. L. (2008). "Decision Making with the Analytic Hierarchy Process." *Int. J. of Services Sciences*, 1(1), 83-98.

She, T. H., and Sarshar, M. (1999). "A Geographic Information System (GIS)–Based Bridge Management System." *Computer-Aided Civil and Infrastructure Engineering*, 14(6), 417-427.

Signore, A. (2016). "Mapping and sharing agro-biodiversity using Open Data Kit and Google Fusion Tables." *Computers and Electronics in Agriculture*, 127, 87-91.

Surowiecki, J. (2004). "Mass Intelligence." *Forbes Global*, 7(9), 19.

TRIP (2016). "Oklahoma Transportation by the Numbers." TRIP a national transportation reserach group, Washington, DC.

Valenzuela, S., Solminihac, H. d., and Echaveguren, T. (2010). "Proposal of an Integrated Index for Prioritization of Bridge Maintenance." *Journal of Bridge Engineering*, 15(3), 337-343.

Zhang, X., and Gao, H. (2012). "Determining an Optimal Maintenance Period for Infrastructure Systems." *Computer-Aided Civil and Infrastructure Engineering*, 27(7), 543-554.

CHAPTER V

CONCLUSIONS

As shown in the conceptual overview of the research (Figure 5-1), two major parts were investigated in this dissertation: superstructure deficiency; and bridge maintenance prioritization. The first interest of this study was to validate the development of a framework for predicting and characterizing superstructure deficiency by implementing data mining techniques such as regression (multiple or logistic), regression trees, artificial neural networks, gradient busting, or support vector machine on the NBI-2013 database. The second interest was to develop a decision-making framework to prioritize bridge maintenance through using aggregated bridge ratings and average daily traffic (ADT).

Although the three parts of a bridge (deck, superstructure, and substructure) are examined in bridge inspections, superstructure has been the interest of several previous studies such as Veshosky et al. (1994), Dilger (1998), Zhou et al. (2004), Gangone et al. (2011), Jiao et al. (2013), and Menkulasi and Kuruppuarachchi (2017). In addition, the author has also studied the superstructure previously (Contreras-Nieto, 2014). Due to this, the first part of this study focused on superstructure rating. Moreover, the study of the superstructure rating was also approached in two different ways: prediction and characterization. In particular, a framework for creating models to predict the superstructure ratings of steel and prestressed concrete bridges in the state of Oklahoma was developed in Chapter 2. In Chapter 3 a framework was developed for characterizing steel bridge deterioration in order to understand which factors influence superstructure deficiency and how they do so. Chapter 2 and Chapter 3 used NBI-2013 database as the main source of information and SAS Enterprise Miner ™ as model development software. While Chapter 2 considered the superstructure rating as a continuous variable with a scale from 0 to 9, Chapter 3 studied the superstructure rating as a binary target with possible outcome of deficient and non-deficient. The second part of this dissertation was presented in Chapter 4 where a decision-making framework to prioritize maintenance through using aggregated bridge ratings and average daily traffic (ADT). In addition, by incorporating the developed decision-making framework into a geographical information system (GIS), a user interface was created using Google TM Fusion Tables and Google Maps. Thus, the bridge maintenance prioritization can be visualized.

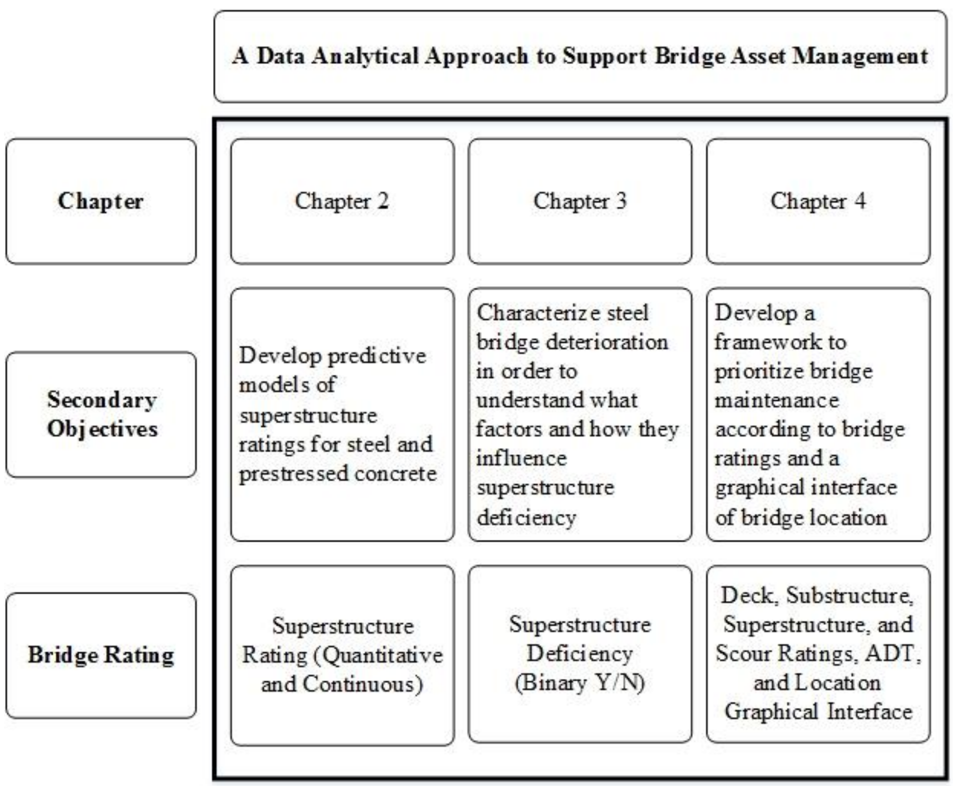

Figure 5-1. Conceptual Overview of Dissertation

MAIN CONCLUSIONS AND CONTRIBUTIONS TO THEORY AND PRACTICE

This dissertation is a practical research. Although the first part of this dissertation focused on one specific bridge condition rating, superstructure rating, the applicability of the developed methodologies in Chapter 2 and Chapter 3 can be implemented to predict other condition bridge ratings (deck or substructure ratings). In addition, the decision-making framework to prioritize bridges can also be implemented regardless of the location of bridges and maintenance responsible parties. Therefore, the applicability of this dissertation is not limited to the examined bridge rating nor bridge parameters such as location; the essence of this study is its capability and the developed methodologies themselves. This study directly addresses the interest of different bridge stakeholders such as Departments of Transportation and City/Municipal Highway Agencies.

Departments of Transportation (DOTs) or City/Municipal Highway Agencies can use the developed framework in different ways. First, they could implement the concept of the framework in order to predict the bridge condition ratings through time by creating models with their current databases. As a result, new maintenance strategies could be developed in order to keep desired safety and serviceability levels. Similarly, the second use is that these bridge stakeholders can characterize any bridge condition rating by building models. Therefore, factors with significant impact on bridge deterioration could be determined, and bridge owners could create strategies to address them when possible. Third, this framework could be used to estimate maintenance budgets. By implementing their developed models through time and analyzing the outcome, bridge stakeholders could determine when a bridge should be scheduled for maintenance according to number of units the analyzed condition rating has dropped. Thus, a maintenance estimate can be prepared by using previous unit costs and the bridge characteristics (i.e. road width and length).

In addition, the combination of bridge maintenance prioritization and the geographical information system user interface presented in this research showed its potential to solve the decision-making process of selecting deficient bridges to be maintained. This methodology is an alternative to the current way of prioritizing and visualizing deficient bridges. Also, this approach addressed issues such as the subjectivity of the current process in selecting deficient bridges. Moreover, it integrates opinions of experts in order to find a better solution to bridge maintenance prioritization. Regardless of the bridge stakeholder, number of deficient bridges, or budget, this methodology could be used as a managerial tool to better select deficient bridges and schedule their maintenance, and also speed up this process.

Although the contributions mentioned above are generalizable since the framework could be implemented to model any condition rating or prioritize bridge maintenance in any state, the author narrowed his research because of the relevance of the superstructure rating and the accessibility of the data from Oklahoma Department of Transportation (ODOT) – Division 4. Besides the general contributions of this study, the contribution associated with the three objectives examined in this dissertation could interest bridge stakeholders and researches.

In Chapter 2, a framework to predict superstructure ratings for steel and prestressed concrete bridges was developed. It was achieved by implementing data mining techniques to develop prediction models based on bridge inspection data in Oklahoma. The main source of information for this objective was NBI-2013. The following conclusions can be taken from the results obtained.

1) Although age is the main predictor of superstructure ratings for the prestressed concrete and the steel Group 4 bridges, which is consistent with previous studies, it is not the main predictor for other steel bridge groups (Group 1, Group 2, and Group 3). These steel bridge groups have location (Item 3 County) as the most significant predictor of superstructure rating.

2) The percentage of superstructure ratings predicted with a tolerance of ± 1 rating is around 98% for presetessed concrete and more than 92% for steel groups. However, all models have the limitation of not being able to predict superstructure ratings lower than 5.

Chapter 3 presented a framework that is capable of characterizing superstructure deficiency of steel bridges. Bridge inspection data from the entire nation (NBI-2013) was used to accomplish this goal. The results revealed that logistic regression performed better than machine learning techniques. Nevertheless, because every year of the NBI databases is unique, it is possible that if the same techniques used in this study were used in another year, a machine learning technique could perform better than logistic regression. The following conclusions can be determined from this chapter.

1) Bridges characterized for having longer maximum spans and wider road widths obtained low probabilities of having deficient superstructures.
2) Older, longer, and higher trafficked bridges obtained the highest probabilities of having deficient superstructures.
3) The owner and the location of the bridge were identified as significant factors of steel superstructure deficiency.

Chapter 4 contains the framework to prioritize bridge maintenance in Oklahoma. Moreover, this chapter also presents a geographical interface to improve the visualization of deficient bridge priorities and support the scheduling of bridge maintenance. As a validation case, data from ODOT – Division 4 was used to prioritize deficient bridges. By presenting and explaining the methodology and the results to the bridge engineer of ODOT – Division 4, it is possible to say that the approach is effective and help the task of prioritizing bridge maintenance. Also, the graphical interface supports the visualization of the results in order to make better decisions on bridge maintenance prioritization. The major findings of this chapter are as follows:

1) When safety, serviceability, and comfort are concerned, deck was the most critical rating for non-water-crossing bridges. Substructure rating obtained the highest importance with respect to the resiliency criterion.

2) In regards to water-crossing bridges, scour was the most important rating for safety and serviceability. Both deck and substructure ratings were given higher preference when comfort and resiliency are concerned.
3) By the magnitude of relevance from strongest to weakest, the four criteria were ranked: safety, resiliency, serviceability, and comfort.
4) The weights that form the aggregated rating for water-crossing bridges are: 0.171 (deck), 0.268 (superstructure), 0.291 (substructure), and 0.270 (scour). Non-water-crossing bridges have the following weights: 0.360 (deck), 0.356 (superstructure), and 0.283 (substructure).

SUGGESTIONS FOR FUTURE RESEARCH

The prediction and characterization models developed in Chapter 2 and Chapter 3, respectively, were built using the bridge information recorded in the NBI-2013 database. However, NBI databases do not contain information that may be helpful for the prediction of bridge condition ratings or to the characterization of bridge deficiencies. Data of previous bridge maintenance, climate and hydrological data, and CoRe data are examples of information that could complement the NBI databases. Integrating any of these examples of additional data with NBI databases may be considered for use in future research.

The approaches implemented in Chapter 2 to predict the superstructure rating have an assumption of data independency, which means that no observation could be repeated. Therefore, it was not possible to use multiple years of the NBI databases, and it can be seen as a limitation of the study. Using Time Series Analysis, this limitation can be addressed. Thus, future research could be focused on developing time series analysis models to predict bridge condition by using all NBI databases as sources of information.

Chapter 3 was focused on the characterization of steel bridge superstructure deterioration. Similar studies can be conducted on other superstructure materials, such as concrete, and prestressed concrete. In addition to superstructure rating, deck and substructure ratings in NBI contain valuable information about the state of a bridge as well. Future studies can attempt to characterize the deterioration of decks and substructure.

Finally, the process of the prioritization of bridge maintenance described in Chapter 4 can be automated through the development of a software application, which allows stakeholders to load a file with the bridges and their inspection records and then a list of bridges with the respective priority can be generated and visualized in the map. In terms of the weights for the ratings, the application could have the flexibility of either using user-defined weights or the weights obtained in this study. Similarly, the application could have options to personalize ADT classes, include other ratings, etc.

REFERENCES

Contreras-Nieto, C. (2014). "Development of linear models to predict superstructure ratings of steel and prestressed concrete bridges." Dissertation/Thesis, ProQuest, UMI Dissertations Publishing.

Dilger, W. (1998). Bridge superstructure: New developments. *Canadian Journal of Civil Engineering, 25*(2), 389.

Gangone, M., Whelan, M., and Janoyan, K. (2011). "Wireless Monitoring of a Multispan Bridge Superstructure for Diagnostic Load Testing and System Identification." *Computer-Aided Civil and Infrastructure Engineering*, *26*(7), 560-579.

Jiao, Y., Liu, H., Zhang, P., Wang, X., and Wei, H. (2013). "Unsupervised Performance Evaluation Strategy for Bridge Superstructure Based on Fuzzy Clustering and Field Data." *The Scientific World Journal, 2013*, 6.

Menkulasi, F., and Kuruppuarachchi, D. (2017). "Development of alternative concrete bridge superstructure systems for short and medium span bridges." *Engineering Structures*, 143, 274-294.

Veshosky, D., Beidleman, C., Buetow, G., and Demir, M. (1994). "Comparative Analysis of Bridge Superstructure Deterioration." *Journal of Structural Engineering*, 120(7), 2123-2136.

Zhou, S., Rizos, D.C., and Petrou, M. F. (2004). Effects of superstructure flexibility on strength of reinforced concrete bridge decks. *Computers and Structures, 82*(1), 13-23.

REFERENCES

Alroomi, A., Jeong, D. H. S., and Oberlender, G. D. (2012). "Analysis of Cost-Estimating Competencies Using Criticality Matrix and Factor Analysis." *Journal of Construction Engineering and Management*, 138(11): 1270-1280.

Akgül, F., and Frangopol, D. M. (2003). "Rating and Reliability of Existing Bridges in a Network." *Journal of Bridge Engineering*, 8(6), 383-393.

America Society of Civil Engineers - ASCE (2013). "2013 Report Card of America's Infrastructure." <http://www.infrastructurereportcard.org/>. (July 2015).

America Society of Civil Engineers - ASCE (2017). "2017 Report Card of America's Infrastructure." <https://www.infrastructurereportcard.org/cat-item/bridges/>. (June 2017).

Ang, R. P., and Goh, D. H. (2013). "Predicting Juvenile Offending." *International Journal of Offender Therapy and Comparative Criminology*, 57(2), 191-207.

Bektaş, B. A. (2017). "Use of Recursive Partitioning to Predict National Bridge Inventory Condition Ratings from National Bridge Elements Condition Data." *Transportation Research Record: Journal of the Transportation Research Board*(2612), 29-38.

Bektas, B. A., Carriquiry, A., and Smadi, O. (2013). "Using Classification Trees for Predicting National Bridge Inventory Condition Ratings." *Journal of Infrastructure Systems*, 19(4), 425-433.

Berry, M. J. A., and Linoff, G. (2011). *Data mining techniques: for marketing, sales, and customer relationship management*, Wiley Pub, Indianapolis, Ind.

Boulila, W., Farah, I. R., Ettabaa, K. S., Solaiman, B., and Ghézala, H. B. (2011). "A data mining based approach to predict spatiotemporal changes in satellite images." *International Journal of Applied Earth Observations and Geoinformation*, 13(3), 386-395.

Brunelli, M. (2014). *Introduction to the analytic hierarchy process*, Cham : Springer.

Bu, G. L., J.; Guan, H.; Loo, Y.; and Blumenstein, M. (2015). "Prediction of Long-Term Bridge Performance: Integrated Deterioration Approach with Case Studies." *ASCE*, 29(3), 04014089.

Cerkas, M. (2008). "GIS: The New Utility Infrastructure." *Transmission & Distribution World*, 60(3), 20.

Chou, J.-S., Chiu, C.-K., Farfoura, M., and Al-Taharwa, I. (2011). "Optimizing the Prediction Accuracy of Concrete Compressive Strength Based on a Comparison of Data-Mining Techniques." *Journal of Computing in Civil Engineering*, 25(3), 242-253.

Christie, P. G., Jim; thompson, Jeff; and Well, Chip (2011). "Applied Analytics Using SAS® Enterprise MinerTM Course Notes." SAS Institute Inc.

Connor, R. J., Dexter, R., and Mahmoud, H. (2005). "Inspection and Management of Bridges with Fracture-Critical Details." *Transportation Research Board of the National Academies*.

Contreras-Nieto, C. (2014). "Development of linear models to predict superstructure ratings of steel and prestressed concrete bridges." Dissertation/Thesis, ProQuest, UMI Dissertations Publishing.

Contreras-Nieto, C., Lewis, P., and Shan, Y. (2016). "Developing Predictive Models of Superstructure Ratings for Steel and Prestressed Concrete Bridges." *Construction Research Congress 2016*, 859-868.

Crawford, G., and Williams, C. (1985). "A note on the analysis of subjective judgment matrices." *Journal of Mathematical Psychology*, 29(4), 387-405.

DeVille, B., and Neville, P. (2013). *Decision Trees for Analytics Using SAS® Enterprise Miner™*, SAS Institute Inc.

Dilger, W. (1998). Bridge superstructure: New developments. *Canadian Journal of Civil Engineering, 25*(2), 389.

Dong, Q. and Cooper, O. (2016). "A peer-to-peer dynamic adaptive consensus reaching model for the group AHP decision making." *European Journal of Operational Research*, 250(2): 521-530.

Dreyfus, G. (2005). *Neural networks : methodology and applications*, Berlin. New York: Springer, Berlin. New York.

Ekawati, R., and Suharjito (2016). "API Fusion Tables and Google Maps Integration for GIS Thematic Mapping Visualization." *International Journal of Software and its Applications*, 10(2), 77-92.

Elbehairy, H., Elbeltagi, E., Hegazy, T., and Soudki, K. (2006). "Comparison of Two Evolutionary Algorithms for Optimization of Bridge Deck Repairs." *Computer-Aided Civil and Infrastructure Engineering*, 21(8), 561-572.

Farhey, D. (2015). "Operational Structural Performances of Bridge Types by Deterioration Trends." *Journal Of Performance Of Constructed Facilities*, 29(2).

Federal Highway administration (FHWA). (1995). *Recording and Coding Guide for the Structure Inventory and Appraisal of the Nation's Bridges*. (Publication FHWA-PD-96-001). Washington, D.C: U.S. Dept. of Transportation, Federal Highway Administration.

FHWA, F. H. A., and ASSHTO (2000). *Asset management primer*, Washington, D.C. : U.S. Dept. of Transportation, Federal Highway Administration, Office of Asset Management, Washington, D.C.

Gangone, M., Whelan, M., and Janoyan, K. (2011). "Wireless Monitoring of a Multispan Bridge Superstructure for Diagnostic Load Testing and System Identification." *Computer-Aided Civil and Infrastructure Engineering*, *26*(7), 560-579.

Gonzalez, H., Halevy, A., Langen, A., Madhavan, J., McChesney, R., Shapley, R., Shen, W., and Goldberg-Kidon, J. (2010). "Socialising Data with Google Fusino Tables." IEEE computer Society, 25-32.

Hagan, M. T., H. B. Demuth, M. H. Beale, and O. De Jesús. (1996). *Neural network design*. PWS publishing company Boston.

Hammad, A., Yan, J., and Mostofi, B. (2007). "Recent Development of Bridge Management Systems in Canada " *Transportation Association of Canada.* Saskatoon, Saskatchewa.

Hearn, G., and Johnson, B. (2011). "Identification and Tracking of Highway Bridge Maintenance." *Transportation Research Record: Journal of the Transportation Research Board*, 2220, 12-20.

Higuchi, S., and Macke, M. (2008). "COST-BENEFIT BASED OPTIMIZATION OF MAINTENANCE INTERVENTIONS FOR DETERIORATING STRUCTURES." *JSCE Journal of Structural and Earthquake Engineering*, 63(4), 727-743.

Hu, X., Wang, B., and Ji, H. (2013). "A Wireless Sensor Network-Based Structural Health Monitoring System for Highway Bridges." *Computer-Aided Civil and Infrastructure Engineering*, 28(3), 193-209.

Huang, Y. (2010). "Artificial Neural Network Model of Bridge Deterioration." *Journal of Performance of Constructed Facilities*, 24(6), 597-602.

Jebreen, M. I., and Johnston, D. W. (1995). "Bridge Maintenance Level of Service Optimization Based on an Economic Analysis Approach." North Carolina State University, Raleigh, NC.

Kazakis, N., Kougias, I., and Patsialis, T. (2015). "Assessment of flood hazard areas at a regional scale using an index-based approach and Analytical Hierarchy Process: Application in Rhodope–Evros region, Greece." *Science of the Total Environment*, 538, 555-563.

Kim, B.-H. (2009). "Support Vector Machine & Classification Using Weka (Brain, Computation, and Neural Learning)." Seoul Nation University.

Kim, Y. J., and Yoon, D. K. (2010). "Identifying Critical Sources of Bridge Deterioration in Cold Regions through the Constructed Bridges in North Dakota." *Journal of Bridge Engineering*, 15(5), 542-552.

King, G., and L. Zeng. (2001). "Logistic regression in rare events data. " *Political analysis,* 9(2), 137-163.

Jiao, Y., Liu, H., Zhang, P., Wang, X., and Wei, H. (2013). "Unsupervised Performance Evaluation Strategy for Bridge Superstructure Based on Fuzzy Clustering and Field Data." *The Scientific World Journal, 2013*, 6.

Jiang, Y., Saito, M. & Sinha, K. (1988). *Bridge performance prediction model using Markov chain, in Transportation Re- search Record* (No. 1180).

Lee, D., McCool, J., & Napieralski, L. (2000). Assessing adult learning preferences using the analytic hierarchy process. *International Journal of Lifelong Education*, 19(6), 548-560. doi:10.1080/02601370050209078

Lee, L.-S., Chang, L.-T., Lai, C.-C., & Lin, K.-Y. (2011). "Using The Analytical Hierarchy Process To Construct Performance Indicators For Comprehensive High Schools In Taiwan." *Social Behavior and Personality* 39(5): 615-626.

Lee, S. (2012). "Current state of bridge deterioration in the US-part1." *Materials performance*, 51(1), 62-67.

Maryam, D., and Marzieh, A. (2012). "A Hybrid Data Mining Model to Improve Customer Response Modeling in Direct Marketing." *Indian Journal of Computer Science and Engineering*, 3(6), 844-856.

Matthiesen, R. (2010). *Bioinformatics Methods in Clinical Research*, Humana Press, Totowa, NJ, Totowa, NJ.

Menkulasi, F., and Kuruppuarachchi, D. (2017). "Development of alternative concrete bridge superstructure systems for short and medium span bridges." *Engineering Structures*, 143, 274-294.

Meyers, L. S. (2009). Data analysis using SAS Enterprise guide, Cambridge New York : Cambridge University Press, Cambridge. New York.

Micevski, T., Kuczera, G., and Coombes, P. (2002). "Markov Model for Storm Water Pipe Deterioration." *Journal of Infrastructure Systems*, 8(2), 49-56.

Milley, A. H. S., James D.;Williams, John S. (1998). "Data Mining and the Case for Sampling." SAS Institute Inc.

Miyamoto, A., Kawamura, K., and Nakamura, H. (2000). "Bridge Management System and Maintenance Optimization for Existing Bridges." *Computer-Aided Civil and Infrastructure Engineering*, 15(1), 45-55.

Moomen, M. (2016). "Deterioration modeling of highway bridge components using deterministic and stochastic methods." S. Labi, J. Fricker, and K. Sinha, eds., ProQuest Dissertations Publishing.

Morcous, G., Rivard, H., and Hanna, A. (2002). "Case-Based Reasoning System for Modeling Infrastructure Deterioration." *Journal of Computing in Civil Engineering*, 16(1), 104-114.

Munasinghe, M., Clarke, C. L., World, B., and International Decade for Natural Disaster, R. (1995). *Disaster prevention for sustainable development : economic and policy issues*, Washington, DC : International Decade for Natural Disaster Reduction and World Bank, Washington, DC.

Oklahoma Department of Transportation - ODOT (2015). "ODOT Inveils Eigth-Year Construction Work Plan." <http://www.ok.gov/odot/Eight-Year_Construction_Work_Plan.html>. (September 2015).

Oklahoma Department of Transportation - ODOT (2016). "Field Division 4 - Headquarters." <https://www.ok.gov/odot/About_ODOT/Contact_ODOT_Divisions/Division_4.html>. (2016).

Omran, B. A., Chen, Q., and Jin, R. (2016). "Comparison of Data Mining Techniques for Predicting Compressive Strength of Environmentally Friendly Concrete." *Journal of Computing in Civil Engineering*, 0(0), 04016029.

Piñeiro, G., Perelman, S., Guerschman, J. P., and Paruelo, J. M. (2008). "How to evaluate models: Observed vs. predicted or predicted vs. observed?" *Ecological Modelling*, 216(3), 316-322.

Potts, W. P., Michael J. (2002). "Predictive Modeling Using Logistic Regression Course Notes." S. I. Inc., ed., SAS Institute Inc., Cary, NC.

Pozzolo, A. D., Caelen, O., Johnson, R. A., and Bontempi, G. "Calibrating Probability with Undersampling for Unbalanced Classification." *Proc., 2015 IEEE Symposium Series on Computational Intelligence*, 159-166.

Robak, A., Monkman, D., Gilberston, T., Bush, S., O'Connor, B., Blakemore, R., Garlik, S., Hughes, J., Gordon, M., Sharman, B., Smith, M., Maughan, A., Onley, D., Belgrove, T., Waugh, R., Holland, G., Roberts, L., Way, P., Dever-Tod, K., Lutchman, R., and Holdt, C. v. (2015). *IIMM International Infrastructure Management Manual 2015*, Institute of Public Works Engineering Australasia (IPWEA).

Romps, N. G. (2008). "GIS: Mapping the Industry's Future." *American Gas*, 90(2), 15-17.

Saaty, T. L. (2008). "Decision Making with the Analytic Hierarchy Process." *Int. J. of Services Sciences*, 1(1), 83-98.

Saeed TU, Moomen M, Ahmed A, Murillo-Hoyos J, and Volovski M, L. S. (2017). "Performance Evaluation and Life Prediction of Highway Concrete Bridge Superstructure across Design Types." *Journal of Performance of Constructed Facilities*.

SAS Institute INC (2015). "Getting Started with SAS® Enterprise Miner 14.1."

Schulz, A., Zia, A., and Koliba, C. (2017). "Adapting bridge infrastructure to climate change: institutionalizing resilience in intergovernmental transportation planning processes in the Northeastern USA." *Mitigation and Adaptation Strategies for Global Change*, 22(1), 175-198.

Seni, G., and Elder, J. F. (2010). "Ensemble Methods in Data Mining: Improving Accuracy Through Combining Predictions." *Synthesis Lectures on Data Mining and Knowledge Discovery*, 2(1), 1-126.

She, T. H., and Sarshar, M. (1999). "A Geographic Information System (GIS)–Based Bridge Management System." *Computer-Aided Civil and Infrastructure Engineering*, 14(6), 417-427.

Signore, A. (2016). "Mapping and sharing agro-biodiversity using Open Data Kit and Google Fusion Tables." *Computers and Electronics in Agriculture*, 127, 87-91.

Surowiecki, J. (2004). "Mass Intelligence." *Forbes Global*, 7(9), 19.

Tang, P., Kanaan, O., Wnag, J., Oh, J., and Kwigizile, V. (2012). "Statistical Identification of Items Importants for Evaluating Bridge Conditions." *Transportation Research Board*.

Tarefder, R. A., Sumee, N., and Storlie, C. (2014). "Study of MEPDG Sensitivity Using Nonparametric Regression Procedures." *Journal of Computing in Civil Engineering*, 28(1), 134-144.

Thoonen, G., Hufkens, K., Borre, J. V., Spanhove, T., and Scheunders, P. (2012). "Accuracy assessment of contextual classification results for vegetation mapping." *International Journal of Applied Earth Observations and Geoinformation*, 15, 7-15.

Thompson, P. D., and Johnson, M. B. (2005). "Markovian bridge deterioration: developing models from historical data." *Structure and Infrastructure Engineering*, 1(1), 85-91.

TRIP (2016). "Oklahoma Transportation by the Numbers." TRIP a national transportation reserach group, Washington, DC.

U. S. Department of Transportation. Federal Highway Administration (2015). <https://www.fhwa.dot.gov/bridge/nbi/ascii.cfm>. (September, 2015).

U.S. Department of Transportation (2013). "Federal Highway Administration National Bridge Inventory." <https://www.fhwa.dot.gov/bridge/nbi/ascii.cfm?year=2013>. (2013).

Valenzuela, S., Solminihac, H. d., and Echaveguren, T. (2010). "Proposal of an Integrated Index for Prioritization of Bridge Maintenance." *Journal of Bridge Engineering*, 15(3), 337-343.

Veshosky, D., Beidleman, C., Buetow, G., and Demir, M. (1994). "Comparative Analysis of Bridge Superstructure Deterioration." *Journal of Structural Engineering*, 120(7), 2123-2136.

Walsh, S. (2005). "Applying Data Mining Techniques Using SAS® Enterprise Miner." SAS Institute Inc.

Wilson, A. D. (2000). "Production of High Performance Steel for U.S.A Bridges." National Bridge Research Organization, University of Nebraska-Lincoln, Nebraska Library Commission.Yuan, S. (2015). "Random gradient boosting for predicting conditional quantiles." *Journal of Statistical Computation and Simulation*, 85(18), 3716-3726.

Zhang, X., and Gao, H. (2012). "Determining an Optimal Maintenance Period for Infrastructure Systems." *Computer-Aided Civil and Infrastructure Engineering*, 27(7), 543-554

Zhou, S., Rizos, D.C., and Petrou, M. F. (2004). Effects of superstructure flexibility on strength of reinforced concrete bridge decks. *Computers and Structures*, *82*(1), 13-23.

Zhu, Z., and Brilakis, I. (2010). "Parameter optimization for automated concrete detection in image data." *Automation in Construction*, 19(7), 944-953.

Ziegel, E. R. (2003). Logistic Regression: A Self-Learning Text. *Technometrics*, Vol. 45, No. 1, 2003, p. 109.

APPENDICES

APPENDIX A

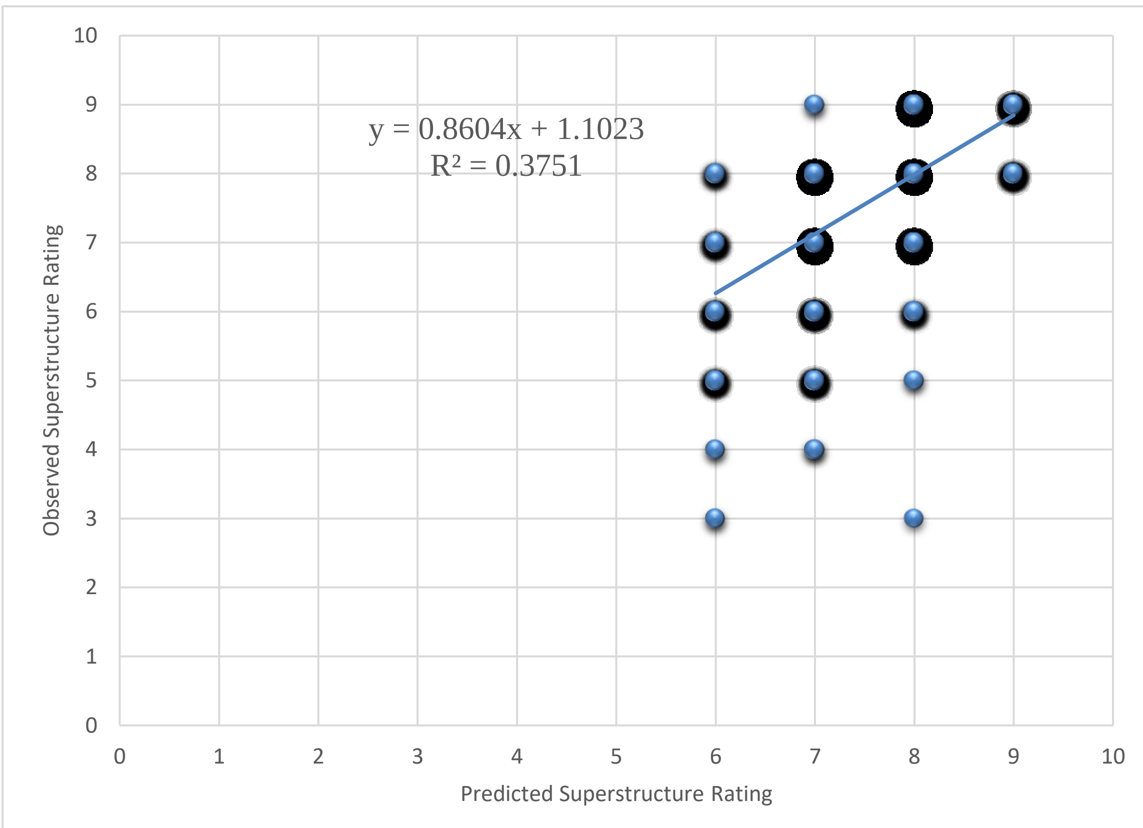

Figure A.1. Superstructure Rating Residuals - PC NBI-2014

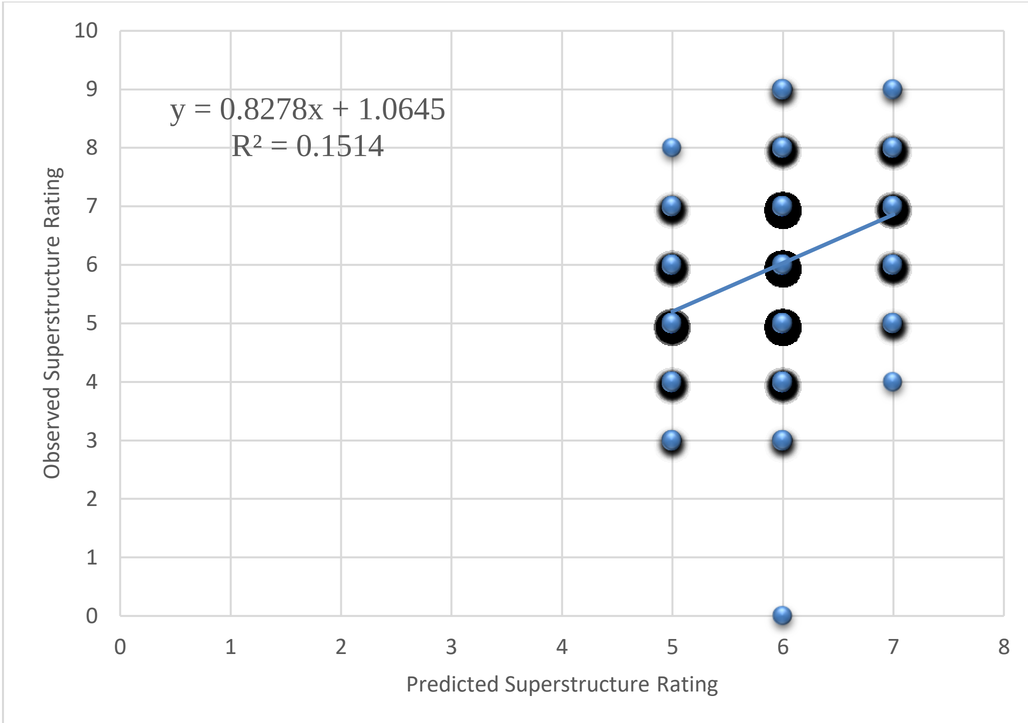

Figure A.2. Superstructure Rating Residuals – Steel Group 1 – NBI-2013

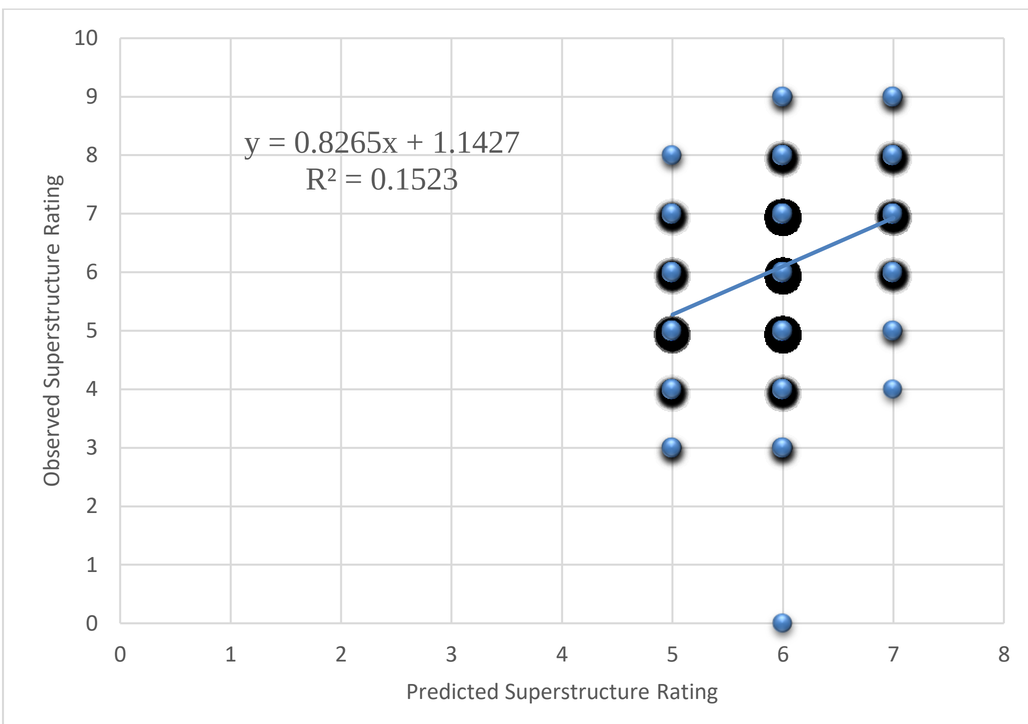

Figure A.3. Superstructure Rating Residuals – Steel Group 1 – NBI-2014

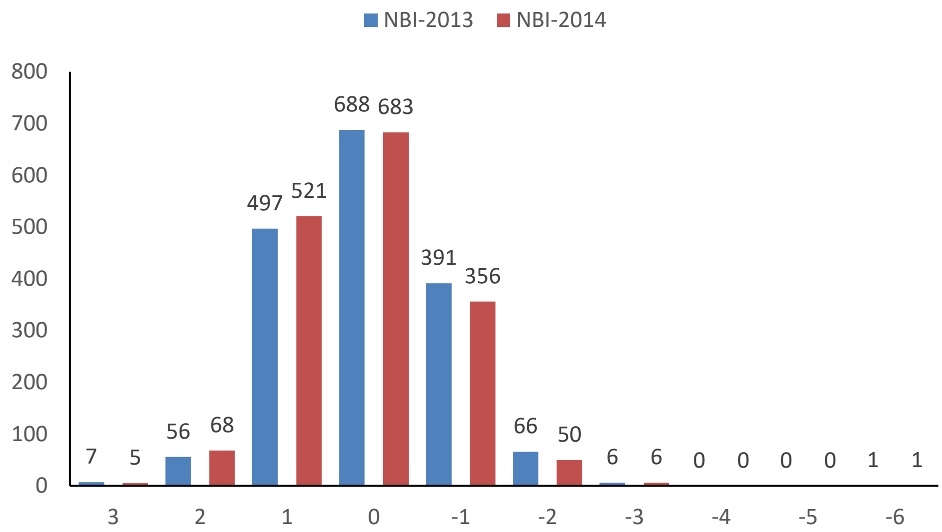

Figure A.4. Frequency of Superstructure Rating Residuals – Steel Group 1

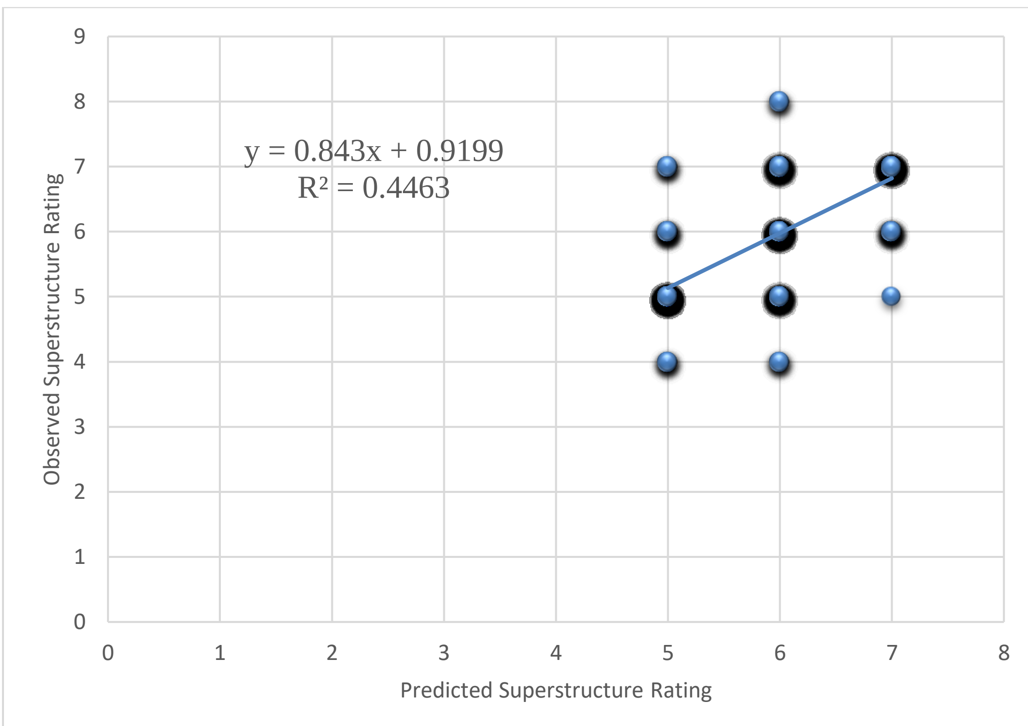

Figure A.5. Superstructure Rating Residuals – Steel Group 2 – NBI-2013

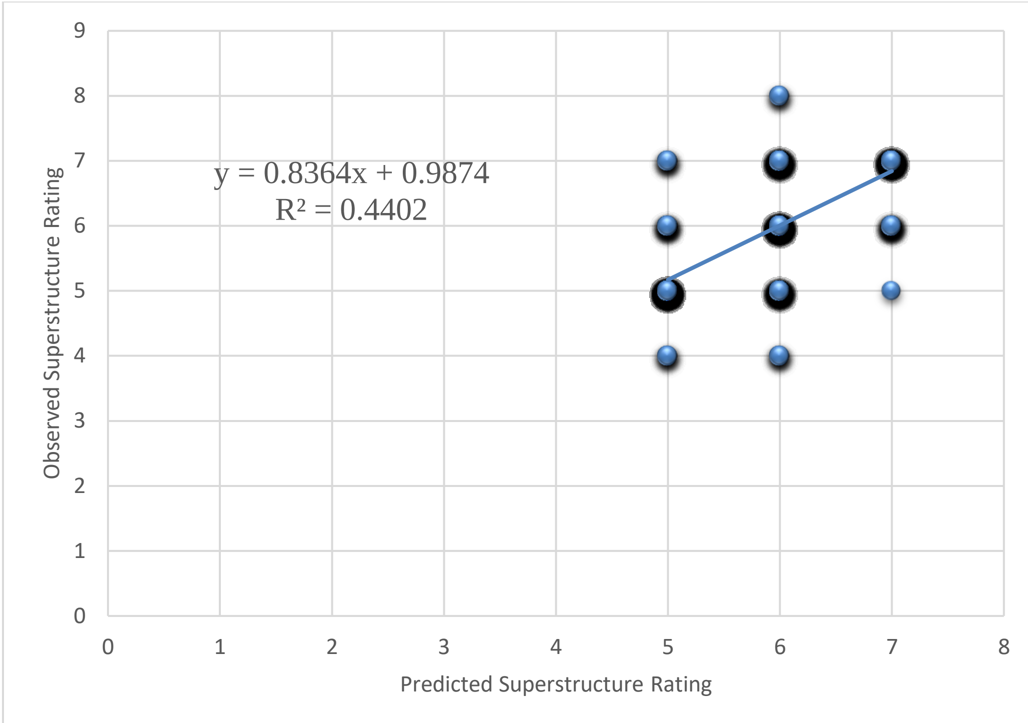

Figure A.6. Superstructure Rating Residuals – Steel Group 2 – NBI-2014

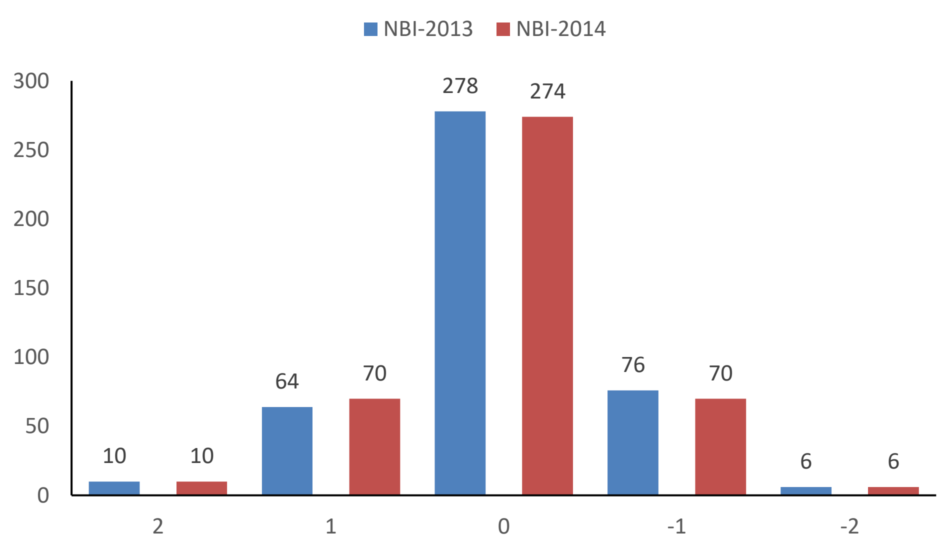

Figure A.7. Frequency of Superstructure Rating Residuals – Steel Group 2

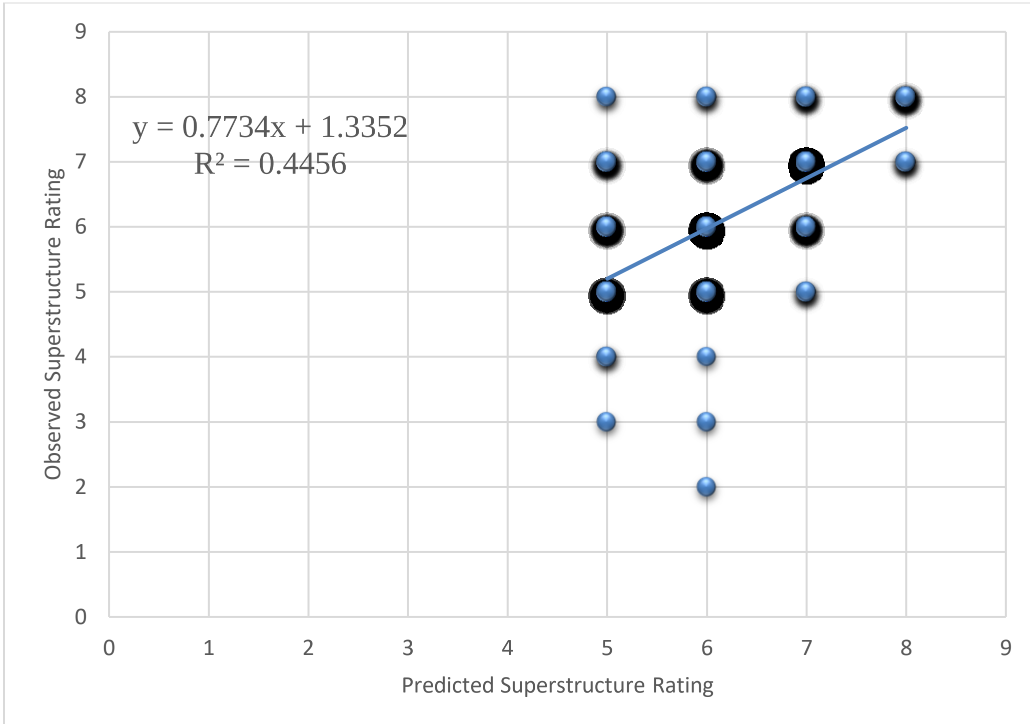

Figure A.8. Superstructure Rating Residuals – Steel Group 3 – NBI-2013

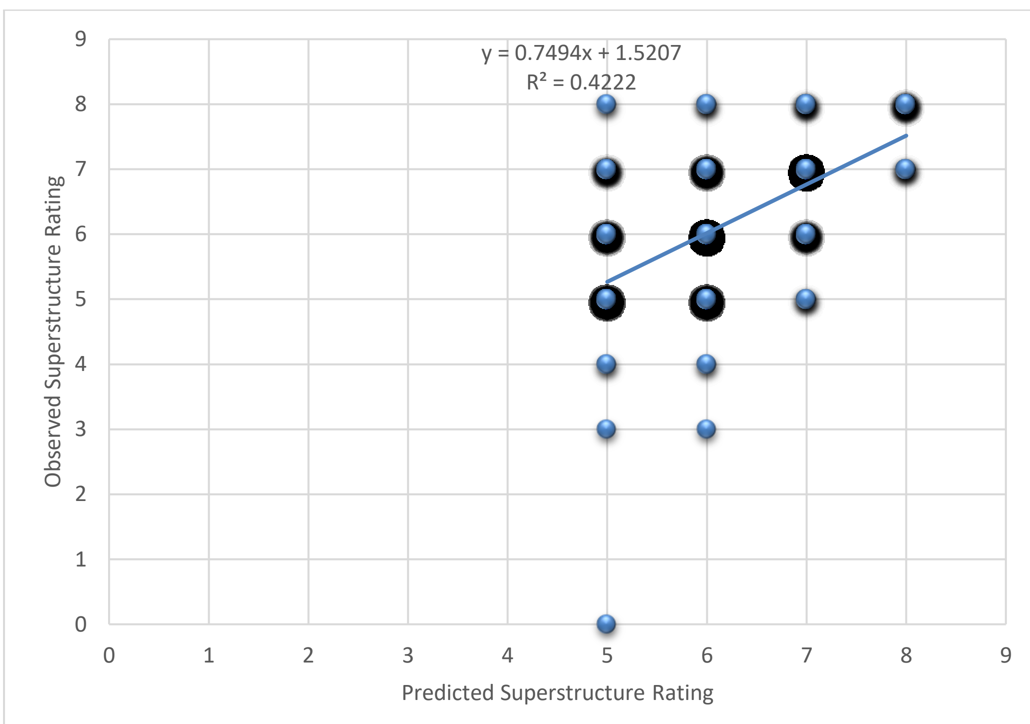

Figure A.9. Superstructure Rating Residuals – Steel Group 3 – NBI-2014

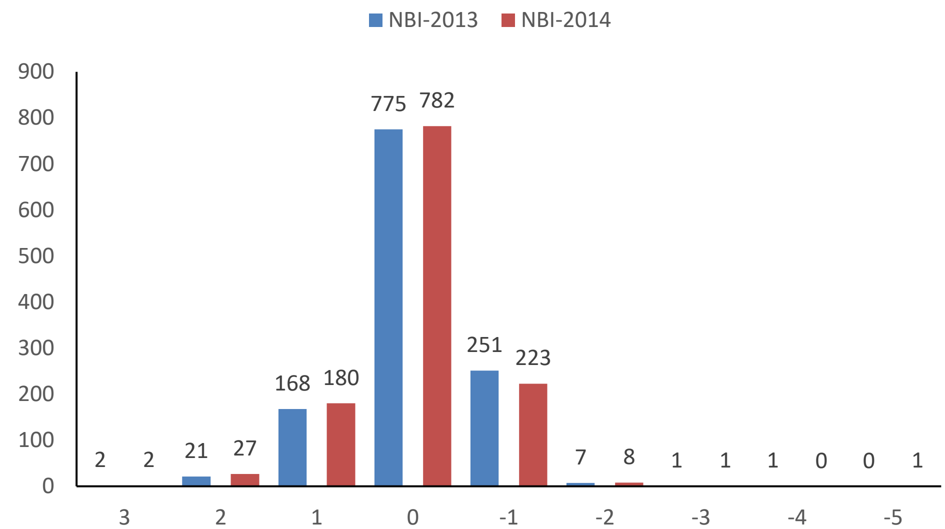

Figure A.10. Frequency of Superstructure Rating Residuals – Steel Group 3

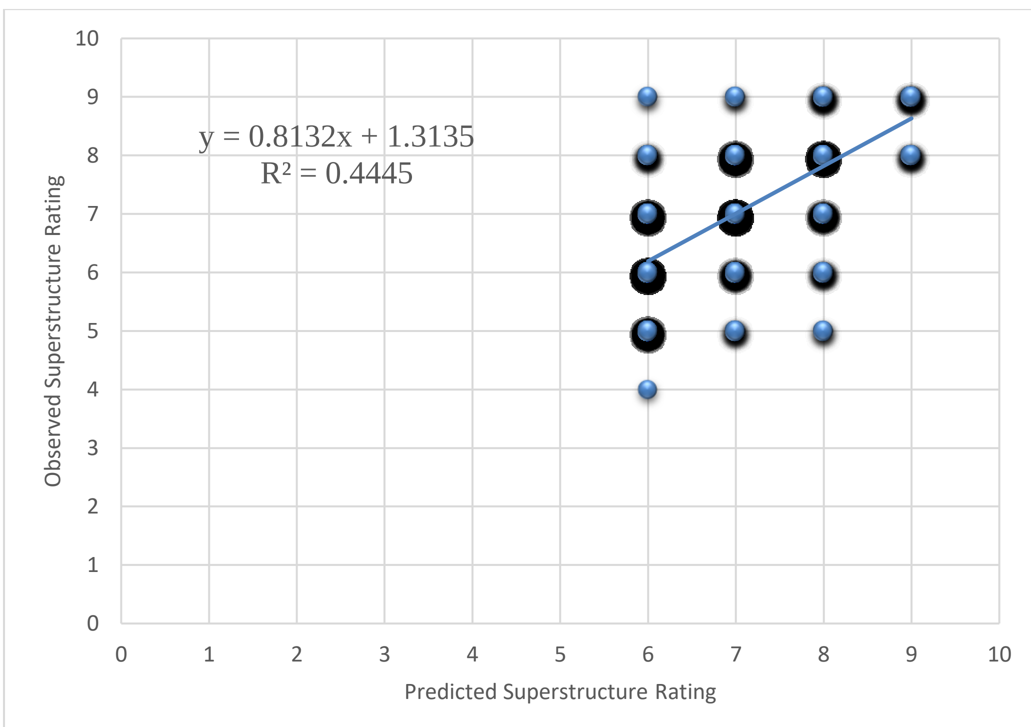

Figure A.11. Superstructure Rating Residuals – Steel Group 4 – NBI-2014

Table A.1 Predicted Superstructure Ratings – Steel Group 2 NBI-2013

Key	Age	ADT	Observed	Predicted	Round Predicted	Residual
Ok-119820000000000	28	30	7	6.83	7	0
Ok-119930000000000	28	50	7	6.32	6	1
Ok-190170000000000	28	100	5	4.95	5	0
Ok-208670000000000	28	100	5	5.09	5	0
Ok-208680000000000	28	25	6	5.33	5	1
Ok-208690000000000	28	100	6	5.15	5	1
Ok-208700000000000	28	24	6	5.34	5	1
Ok-208740000000000	28	100	5	5.06	5	0
Ok-208760000000000	28	30	5	5.87	6	-1
Ok-208820000000000	28	50	7	6.29	6	1
Ok-208830000000000	28	100	6	6.21	6	0
Ok-208860000000000	28	200	5	6.11	6	-1
Ok-208870000000000	28	24	7	5.87	6	1
Ok-208910000000000	28	100	5	5.08	5	0
Ok-208920000000000	28	100	5	5.14	5	0
Ok-208930000000000	28	100	5	4.94	5	0
Ok-208940000000000	28	100	5	4.77	5	0
Ok-208960000000000	28	100	5	5.98	6	-1
Ok-209060000000000	28	25	7	6.62	7	0
Ok-209080000000000	28	25	7	7.04	7	0
Ok-209090000000000	28	25	7	7.04	7	0
Ok-209100000000000	28	75	6	6.53	7	-1
Ok-209110000000000	28	100	6	6.23	6	0
Ok-209120000000000	28	50	7	6.53	7	0
Ok-209130000000000	28	100	6	6.08	6	0
Ok-209170000000000	28	50	5	5.87	6	-1
Ok-209190000000000	28	24	5	5.15	5	0
Ok-209200000000000	28	24	7	6.79	7	0
Ok-209220000000000	28	100	5	5.00	5	0
Ok-209230000000000	28	100	5	5.47	5	0
Ok-209240000000000	28	441	6	5.77	6	0
Ok-209250000000000	28	50	5	5.11	5	0
Ok-209280000000000	28	100	7	6.18	6	1
Ok-209320000000000	28	50	6	5.87	6	0
Ok-209330000000000	28	60	5	5.87	6	-1
Ok-209450000000000	28	50	7	6.87	7	0
Ok-209460000000000	28	100	6	5.87	6	0
Ok-209480000000000	28	100	6	6.20	6	0
Ok-209490000000000	28	44	6	5.93	6	0
Ok-209500000000000	28	100	6	6.16	6	0
Ok-209550000000000	28	100	5	4.99	5	0
Ok-209560000000000	28	25	7	6.99	7	0
Ok-209570000000000	28	30	7	6.83	7	0
Ok-209630000000000	28	100	7	6.36	6	1
Ok-209650000000000	28	100	7	6.27	6	1
Ok-209760000000000	28	100	6	5.57	6	0
Ok-209790000000000	28	100	7	5.87	6	1
Ok-209800000000000	28	100	7	6.25	6	1
Ok-209810000000000	28	44	6	6.18	6	0
Ok-209850000000000	28	25	7	6.86	7	0
Ok-209860000000000	28	25	7	6.65	7	0
Ok-209870000000000	28	50	6	6.11	6	0
Ok-209900000000000	28	100	5	4.91	5	0
Ok-209960000000000	28	100	6	5.87	6	0

Key	Age	ADT	Observed	Predicted	Round Predicted	Residual
Ok-209970000000000	28	200	5	5.87	6	-1
Ok-210010000000000	28	100	6	5.71	6	0
Ok-210020000000000	28	100	4	5.49	5	-1
Ok-210030000000000	28	100	6	6.11	6	0
Ok-210040000000000	28	25	7	6.74	7	0
Ok-210060000000000	28	100	5	5.00	5	0
Ok-210070000000000	28	100	6	6.16	6	0
Ok-210080000000000	28	50	6	5.87	6	0
Ok-210260000000000	28	50	6	6.16	6	0
Ok-210270000000000	28	100	6	5.87	6	0
Ok-210300000000000	28	100	6	5.87	6	0
Ok-210380000000000	28	100	6	6.16	6	0
Ok-210420000000000	28	24	6	6.46	6	0
Ok-210450000000000	28	100	6	6.31	6	0
Ok-210490000000000	28	50	6	6.32	6	0
Ok-210830000000000	28	7000	7	5.87	6	1
Ok-211270000000000	28	2800	7	5.87	6	1
Ok-232490000000000	28	100	7	6.26	6	1
Ok-235250000000000	28	1448	5	5.87	6	-1
Ok-263450000000000	28	24	5	5.11	5	0
Ok-270750000000000	28	25	6	5.20	5	1
Ok-112250000000000	29	30	7	6.83	7	0
Ok-119790000000000	29	50	7	6.76	7	0
Ok-119960000000000	29	25	7	6.55	7	0
Ok-121520000000000	29	25	7	6.41	6	1
Ok-146420000000000	29	100	7	6.68	7	0
Ok-205840000000000	29	100	5	4.93	5	0
Ok-205860000000000	29	200	7	6.66	7	0
Ok-205870000000000	29	25	7	6.85	7	0
Ok-205880000000000	29	50	7	6.58	7	0
Ok-205890000000000	29	30	7	6.84	7	0
Ok-205900000000000	29	50	6	6.52	7	-1
Ok-206010000000000	29	100	6	5.76	6	0
Ok-206020000000000	29	25	6	6.71	7	-1
Ok-206030000000000	29	25	7	6.93	7	0
Ok-206110000000000	29	100	6	6.29	6	0
Ok-206230000000000	29	100	6	5.87	6	0
Ok-206350000000000	29	100	5	4.94	5	0
Ok-206360000000000	29	100	6	5.73	6	0
Ok-206370000000000	29	50	7	6.89	7	0
Ok-206390000000000	29	150	5	5.29	5	0
Ok-206400000000000	29	60	5	5.08	5	0
Ok-206410000000000	29	160	5	5.37	5	0
Ok-206420000000000	29	24	7	6.66	7	0
Ok-206550000000000	29	25	7	6.54	7	0
Ok-206570000000000	29	25	7	6.57	7	0
Ok-206660000000000	29	100	5	5.57	6	-1
Ok-206690000000000	29	30	5	6.48	6	-1
Ok-206800000000000	29	75	5	5.04	5	0
Ok-206820000000000	29	60	5	5.03	5	0
Ok-206830000000000	29	100	7	6.38	6	1
Ok-206940000000000	29	25	7	6.65	7	0
Ok-206950000000000	29	100	5	4.94	5	0
Ok-206960000000000	29	100	6	6.40	6	0
Ok-206970000000000	29	25	7	6.85	7	0
Ok-206980000000000	29	600	5	5.66	6	-1
Ok-207060000000000	29	100	6	5.53	6	0
Ok-207070000000000	29	400	6	5.35	5	1
Ok-207210000000000	29	25	5	5.16	5	0

Key	Age	ADT	Observed	Predicted	Round Predicted	Residual
Ok-207230000000000	29	8300	5	6.21	6	-1
Ok-207290000000000	29	25	7	7.14	7	0
Ok-207330000000000	29	25	6	6.64	7	-1
Ok-207350000000000	29	50	7	6.47	6	1
Ok-207360000000000	29	100	5	5.02	5	0
Ok-207380000000000	29	25	7	6.46	6	1
Ok-207400000000000	29	100	6	5.67	6	0
Ok-207500000000000	29	25	7	6.86	7	0
Ok-207510000000000	29	25	7	6.55	7	0
Ok-207520000000000	29	25	7	6.59	7	0
Ok-207550000000000	29	100	5	5.00	5	0
Ok-207580000000000	29	100	5	5.71	6	-1
Ok-207720000000000	29	485	7	5.54	6	1
Ok-207800000000000	29	100	6	5.87	6	0
Ok-207930000000000	29	50	5	5.11	5	0
Ok-208240000000000	29	2000	6	5.56	6	0
Ok-208310000000000	29	250	5	6.24	6	-1
Ok-208610000000000	29	1000	7	6.19	6	1
Ok-208620000000000	29	1000	7	6.19	6	1
Ok-208630000000000	29	7850	7	5.87	6	1
Ok-208650000000000	29	5465	7	5.87	6	1
Ok-230680000000000	29	100	5	5.65	6	-1
Ok-241430000000000	29	100	7	6.23	6	1
Ok-139640000000000	30	100	5	5.78	6	-1
Ok-148900000000000	30	100	5	5.38	5	0
Ok-203290000000000	30	25	5	4.95	5	0
Ok-203310000000000	30	25	7	6.88	7	0
Ok-203320000000000	30	25	7	6.94	7	0
Ok-203330000000000	30	25	5	5.20	5	0
Ok-203460000000000	30	25	7	7.02	7	0
Ok-203490000000000	30	50	5	5.23	5	0
Ok-203540000000000	30	50	7	6.70	7	0
Ok-203610000000000	30	25	5	5.06	5	0
Ok-203620000000000	30	100	5	4.88	5	0
Ok-203850000000000	30	100	5	5.87	6	-1
Ok-203870000000000	30	100	5	4.82	5	0
Ok-203900000000000	30	25	7	6.99	7	0
Ok-203920000000000	30	100	7	6.61	7	0
Ok-203940000000000	30	503	6	5.87	6	0
Ok-204110000000000	30	24	5	5.28	5	0
Ok-204120000000000	30	25	7	6.82	7	0
Ok-204140000000000	30	43	5	5.11	5	0
Ok-204150000000000	30	100	5	4.94	5	0
Ok-204190000000000	30	100	5	4.73	5	0
Ok-204200000000000	30	200	5	5.34	5	0
Ok-204320000000000	30	75	5	5.00	5	0
Ok-204350000000000	30	50	7	6.26	6	1
Ok-204370000000000	30	100	7	5.87	6	1
Ok-204390000000000	30	25	7	6.59	7	0
Ok-204450000000000	30	100	5	5.87	6	-1
Ok-204610000000000	30	100	5	5.87	6	-1
Ok-204670000000000	30	125	7	6.27	6	1
Ok-204730000000000	30	100	5	5.87	6	-1
Ok-204750000000000	30	25	7	6.60	7	0
Ok-204760000000000	30	100	6	6.05	6	0
Ok-204830000000000	30	71	4	5.62	6	-2
Ok-204850000000000	30	100	5	5.87	6	-1
Ok-204860000000000	30	40	7	6.73	7	0
Ok-204880000000000	30	100	7	6.43	6	1

Key	Age	ADT	Observed	Predicted	Round Predicted	Residual
Ok-204940000000000	30	500	6	6.57	7	-1
Ok-204960000000000	30	100	7	6.66	7	0
Ok-204980000000000	30	249	6	5.55	6	0
Ok-205120000000000	30	25	5	5.19	5	0
Ok-205220000000000	30	50	5	6.42	6	-1
Ok-205250000000000	30	24	5	5.28	5	0
Ok-205280000000000	30	100	5	4.87	5	0
Ok-205450000000000	30	130	4	4.99	5	-1
Ok-205480000000000	30	100	7	5.87	6	1
Ok-232530000000000	30	100	5	4.94	5	0
Ok-121840000000000	31	100	7	6.56	7	0
Ok-149230000000000	31	50	5	5.87	6	-1
Ok-186350000000000	31	85	5	5.09	5	0
Ok-200110000000000	31	150	5	6.33	6	-1
Ok-200210000000000	31	25	7	7.01	7	0
Ok-200250000000000	31	1500	6	6.54	7	-1
Ok-200280000000000	31	25	7	6.95	7	0
Ok-200320000000000	31	100	5	4.87	5	0
Ok-200520000000000	31	75	7	5.87	6	1
Ok-200570000000000	31	30	6	6.86	7	-1
Ok-200790000000000	31	50	5	6.20	6	-1
Ok-200820000000000	31	25	5	5.20	5	0
Ok-200880000000000	31	75	7	5.91	6	1
Ok-201080000000000	31	100	5	5.74	6	-1
Ok-201100000000000	31	35	5	5.10	5	0
Ok-201140000000000	31	100	7	6.70	7	0
Ok-201150000000000	31	2000	6	5.92	6	0
Ok-201160000000000	31	25	6	6.49	6	0
Ok-201250000000000	31	24	5	5.33	5	0
Ok-201260000000000	31	25	7	6.49	6	1
Ok-201310000000000	31	63	5	5.25	5	0
Ok-201340000000000	31	25	5	5.87	6	-1
Ok-201350000000000	31	52	5	5.16	5	0
Ok-201370000000000	31	2600	6	5.89	6	0
Ok-201500000000000	31	100	7	5.87	6	1
Ok-201600000000000	31	24	5	5.28	5	0
Ok-201640000000000	31	150	7	6.13	6	1
Ok-201660000000000	31	25	7	6.65	7	0
Ok-201670000000000	31	632	6	5.87	6	0
Ok-201740000000000	31	200	7	6.95	7	0
Ok-201750000000000	31	25	7	6.30	6	1
Ok-201760000000000	31	25	7	6.65	7	0
Ok-201780000000000	31	100	6	5.97	6	0
Ok-201910000000000	31	100	7	5.87	6	1
Ok-201930000000000	31	50	5	5.87	6	-1
Ok-201940000000000	31	100	7	5.13	5	2
Ok-201950000000000	31	100	7	5.13	5	2
Ok-201970000000000	31	50	7	6.71	7	0
Ok-201980000000000	31	25	6	5.14	5	1
Ok-201990000000000	31	96	5	5.03	5	0
Ok-202010000000000	31	24	7	6.86	7	0
Ok-202030000000000	31	50	7	5.87	6	1
Ok-202050000000000	31	100	6	5.33	5	1
Ok-202060000000000	31	100	7	6.70	7	0
Ok-202130000000000	31	100	6	5.87	6	0
Ok-202210000000000	31	100	7	6.45	6	1
Ok-202260000000000	31	100	7	6.52	7	0
Ok-202410000000000	31	100	5	5.87	6	-1
Ok-202560000000000	31	100	7	6.60	7	0

Key	Age	ADT	Observed	Predicted	Round Predicted	Residual
Ok-202620000000000	31	100	6	6.55	7	-1
Ok-202650000000000	31	100	5	5.25	5	0
Ok-202790000000000	31	100	5	5.87	6	-1
Ok-202870000000000	31	375	7	6.01	6	1
Ok-202890000000000	31	50	5	5.87	6	-1
Ok-203040000000000	31	50	6	6.05	6	0
Ok-203210000000000	31	100	7	6.70	7	0
Ok-204640000000000	31	75	6	6.27	6	0
Ok-182090000000000	32	50	7	6.71	7	0
Ok-198420000000000	32	50	5	5.22	5	0
Ok-198430000000000	32	70	5	5.19	5	0
Ok-198440000000000	32	24	6	5.22	5	1
Ok-198450000000000	32	100	6	5.73	6	0
Ok-198460000000000	32	50	6	6.82	7	-1
Ok-198490000000000	32	150	7	6.77	7	0
Ok-198790000000000	32	100	5	5.87	6	-1
Ok-198830000000000	32	50	5	5.10	5	0
Ok-198850000000000	32	25	7	7.02	7	0
Ok-198880000000000	32	25	7	7.02	7	0
Ok-198900000000000	32	50	7	6.92	7	0
Ok-198970000000000	32	50	5	5.05	5	0
Ok-199160000000000	32	100	6	6.24	6	0
Ok-199190000000000	32	100	7	6.83	7	0
Ok-199230000000000	32	150	7	6.77	7	0
Ok-199280000000000	32	500	7	5.97	6	1
Ok-199290000000000	32	100	6	5.87	6	0
Ok-199360000000000	32	95	5	5.08	5	0
Ok-199380000000000	32	100	6	5.87	6	0
Ok-199400000000000	32	550	6	6.14	6	0
Ok-199460000000000	32	30	6	5.87	6	0
Ok-199480000000000	32	25	5	5.20	5	0
Ok-199500000000000	32	25	7	6.94	7	0
Ok-199520000000000	32	25	6	5.51	6	0
Ok-199540000000000	32	63	5	5.25	5	0
Ok-199640000000000	32	100	6	5.87	6	0
Ok-199700000000000	32	175	6	5.87	6	0
Ok-199790000000000	32	12200	6	6.56	7	-1
Ok-199930000000000	32	70	5	5.13	5	0
Ok-200040000000000	32	5250	7	4.76	5	2
Ok-088130000000000	33	100	6	5.87	6	0
Ok-173980000000000	33	100	6	5.47	5	1
Ok-178700000000000	33	100	4	4.77	5	-1
Ok-197260000000000	33	200	7	6.66	7	0
Ok-197270000000000	33	200	7	6.65	7	0
Ok-197360000000000	33	50	7	6.38	6	1
Ok-197370000000000	33	85	5	4.98	5	0
Ok-197380000000000	33	1500	5	5.13	5	0
Ok-197400000000000	33	11100	6	5.89	6	0
Ok-197430000000000	33	50	5	6.16	6	-1
Ok-197500000000000	33	100	6	6.19	6	0
Ok-197510000000000	33	100	5	5.05	5	0
Ok-197560000000000	33	100	6	5.87	6	0
Ok-197590000000000	33	800	6	5.49	5	1
Ok-197630000000000	33	200	6	5.87	6	0
Ok-197710000000000	33	500	7	5.60	6	1
Ok-197750000000000	33	8750	5	5.88	6	-1
Ok-197760000000000	33	920	6	6.06	6	0
Ok-197780000000000	33	5000	7	7.14	7	0
Ok-197800000000000	33	200	7	5.73	6	1

Key	Age	ADT	Observed	Predicted	Round Predicted	Residual
Ok-197840000000000	33	4000	6	5.87	6	0
Ok-197860000000000	33	.	5	6.09	6	-1
Ok-197890000000000	33	4692	7	6.44	6	1
Ok-197940000000000	33	.	7	6.02	6	1
Ok-198000000000000	33	7050	8	6.48	6	2
Ok-198050000000000	33	100	7	6.67	7	0
Ok-198120000000000	33	5000	7	6.68	7	0
Ok-196410000000000	33	50	7	6.71	7	0
Ok-196420000000000	33	100	7	5.90	6	1
Ok-196490000000000	33	150	6	6.56	7	-1
Ok-196500000000000	33	50	7	6.92	7	0
Ok-196600000000000	33	100	5	5.70	6	-1
Ok-196770000000000	33	50	7	6.92	7	0
Ok-196780000000000	33	50	7	6.93	7	0
Ok-196790000000000	33	500	4	5.17	5	-1
Ok-196800000000000	33	25	7	7.01	7	0
Ok-196850000000000	33	24	5	5.21	5	0
Ok-196860000000000	33	100	5	4.99	5	0
Ok-196880000000000	33	25	5	5.23	5	0
Ok-196970000000000	33	100	6	6.24	6	0
Ok-196990000000000	33	1056	7	6.56	7	0
Ok-197060000000000	33	50	5	5.87	6	-1
Ok-197070000000000	33	100	5	5.87	6	-1
Ok-197090000000000	33	50	5	5.05	5	0
Ok-197120000000000	33	100	6	6.19	6	0
Ok-197150000000000	33	25	6	6.39	6	0
Ok-197190000000000	33	117	6	6.01	6	0
Ok-197200000000000	33	50	5	5.11	5	0
Ok-197210000000000	33	125	5	4.93	5	0
Ok-198190000000000	33	.	7	6.02	6	1
Ok-198230000000000	33	4206	6	6.31	6	0
Ok-198310000000000	33	1700	7	5.87	6	1
Ok-198390000000000	33	5000	6	6.05	6	0
Ok-232080000000000	33	100	6	5.87	6	0
Ok-235450000000000	33	100	5	4.94	5	0
Ok-235460000000000	33	100	5	5.07	5	0
Ok-251710000000000	33	3300	5	5.07	5	0
Ok-260720000000000	33	24	5	5.56	6	-1
Ok-260730000000000	33	50	5	4.78	5	0
Ok-261320000000000	33	100	5	5.75	6	-1
Ok-291200000000000	33	100	7	5.33	5	2
Ok-195350000000000	34	100	7	6.43	6	1
Ok-195480000000000	34	25	5	5.00	5	0
Ok-195500000000000	34	25	7	6.43	6	1
Ok-195530000000000	34	400	5	5.20	5	0
Ok-195560000000000	34	100	5	5.87	6	-1
Ok-195700000000000	34	100	7	6.83	7	0
Ok-195910000000000	34	25	5	4.93	5	0
Ok-195930000000000	34	100	5	6.19	6	-1
Ok-195990000000000	34	10500	5	5.96	6	-1
Ok-196020000000000	34	2517	6	5.83	6	0
Ok-196030000000000	34	2517	5	5.65	6	-1
Ok-196060000000000	34	100	5	4.88	5	0
Ok-196080000000000	34	100	5	5.74	6	-1
Ok-196140000000000	34	.	6	5.96	6	0
Ok-196170000000000	34	100	6	5.57	6	0
Ok-196250000000000	34	800	6	5.87	6	0
Ok-196320000000000	34	10204	6	6.08	6	0
Ok-196360000000000	34	32000	6	5.36	5	1

Key	Age	ADT	Observed	Predicted	Round Predicted	Residual
Ok-249910000000000	34	24	5	4.99	5	0
Ok-011590000000000	35	100	7	6.62	7	0
Ok-193770000000000	35	100	7	6.83	7	0
Ok-193780000000000	35	100	5	4.96	5	0
Ok-193820000000000	35	50	5	4.96	5	0
Ok-193930000000000	35	100	5	5.18	5	0
Ok-193960000000000	35	25	5	5.13	5	0
Ok-193970000000000	35	100	5	5.25	5	0
Ok-194030000000000	35	100	5	5.87	6	-1
Ok-194100000000000	35	50	7	6.28	6	1
Ok-194140000000000	35	100	5	4.79	5	0
Ok-194270000000000	35	100	5	4.77	5	0
Ok-194310000000000	35	100	7	6.55	7	0
Ok-194340000000000	35	100	5	5.00	5	0
Ok-194360000000000	35	9100	6	6.17	6	0
Ok-194430000000000	35	100	4	4.78	5	-1
Ok-194500000000000	35	100	5	6.62	7	-2
Ok-194530000000000	35	17000	7	5.46	5	2
Ok-194540000000000	35	17000	7	5.46	5	2
Ok-194550000000000	35	300	6	6.95	7	-1
Ok-194630000000000	35	203	5	4.87	5	0
Ok-194700000000000	35	23650	5	5.41	5	0
Ok-194710000000000	35	23650	5	5.54	6	-1
Ok-194790000000000	35	21200	4	5.50	6	-2
Ok-194840000000000	35	13150	6	5.87	6	0
Ok-194860000000000	35	5000	5	6.29	6	-1
Ok-194900000000000	35	100	7	5.87	6	1
Ok-194930000000000	35	200	6	6.73	7	-1
Ok-194940000000000	35	5000	5	6.27	6	-1
Ok-194950000000000	35	100	6	5.87	6	0
Ok-194960000000000	35	100	6	5.87	6	0
Ok-194970000000000	35	100	7	5.87	6	1
Ok-194980000000000	35	100	7	5.87	6	1
Ok-194990000000000	35	2517	6	5.69	6	0
Ok-195030000000000	35	20000	6	5.51	6	0
Ok-195040000000000	35	19800	6	5.51	6	0
Ok-195070000000000	35	3200	8	5.67	6	2
Ok-195080000000000	35	3200	6	5.75	6	0
Ok-195130000000000	35	.	6	6.09	6	0
Ok-195140000000000	35	.	6	6.14	6	0
Ok-285770000000000	35	.	7	6.01	6	1
Ok-291530000000000	35	3200	8	5.66	6	2
Ok-CEPSWTOKCOPANSP	35	450	6	5.87	6	0
Ok-018410000000000	36	100	5	5.87	6	-1
Ok-184120000000000	36	100	5	4.87	5	0
Ok-192800000000000	36	100	5	4.69	5	0
Ok-193020000000000	36	100	6	5.77	6	0
Ok-193030000000000	36	25	5	6.37	6	-1
Ok-193040000000000	36	24	5	5.15	5	0
Ok-193090000000000	36	100	5	5.87	6	-1
Ok-193120000000000	36	25	5	5.32	5	0
Ok-193130000000000	36	50	5	5.09	5	0
Ok-193140000000000	36	50	7	6.60	7	0
Ok-193150000000000	36	25	6	6.57	7	-1
Ok-193160000000000	36	25	6	6.61	7	-1
Ok-193180000000000	36	25	5	5.12	5	0
Ok-193190000000000	36	24	5	5.28	5	0
Ok-193210000000000	36	25	5	4.93	5	0
Ok-193270000000000	36	100	5	5.87	6	-1

Key	Age	ADT	Observed	Predicted	Round Predicted	Residual
Ok-193310000000000	36	25	6	6.30	6	0
Ok-193380000000000	36	55	5	5.04	5	0
Ok-193450000000000	36	5134	5	5.52	6	-1
Ok-232540000000000	36	100	5	5.00	5	0
Ok-191160000000000	37	70	4	5.87	6	-2
Ok-191400000000000	37	75	5	6.22	6	-1
Ok-191470000000000	37	24	5	5.21	5	0
Ok-191640000000000	37	100	4	5.87	6	-2
Ok-191660000000000	37	100	4	5.00	5	-1
Ok-191670000000000	37	100	5	5.11	5	0
Ok-191680000000000	37	25	5	5.00	5	0
Ok-191690000000000	37	85	5	5.04	5	0
Ok-191710000000000	37	200	6	6.02	6	0
Ok-191810000000000	37	24	5	5.16	5	0
Ok-191830000000000	37	250	6	6.34	6	0
Ok-191840000000000	37	75	6	6.23	6	0
Ok-191850000000000	37	75	6	6.23	6	0
Ok-191860000000000	37	100	5	5.00	5	0
Ok-191990000000000	37	350	5	4.91	5	0
Ok-192020000000000	37	161	6	5.87	6	0
Ok-192070000000000	37	256	7	6.53	7	0
Ok-192080000000000	37	100	5	4.75	5	0
Ok-192480000000000	37	5000	5	6.13	6	-1
Ok-192530000000000	37	6200	4	6.22	6	-2
Ok-237840000000000	37	24	6	5.22	5	1
Ok-289830000000000	37	5000	8	5.63	6	2

Figure A.12. Bridge Expert Survey

Dear respondent,

The goal of this survey is to develop a decision-making system for the purpose of prioritizing bridge maintenance/repairs in Oklahoma. The objective of this survey is to understand the relative importance of the National Bridge Inventory (NBI) condition ratings with respect to safety, serviceability, comfort, and resiliency. By implementing a group decision-making analysis technique on the survey results of a group of experts in the areas of bridge design, management, and materials, the weights of the three NBI condition ratings will be developed. Eventually, a comprehensive rating system that combines the three condition ratings will be formulated. Therefore, your expert opinion will be invaluable for a successful research outcome.

It should take you less than 10 minutes to complete the survey. Thank you for taking your time to participate in this survey.

Respondent Information
Job Title: ____
Year of experience in this position: __________________________
Organization: ____

You are about to answer some questions regarding two different types of bridges: non-water-crossing and water-crossing bridges. Water-crossing bridges are built over rivers, lakes, or unstable channels while non-water-crossing is used to cross other types of obstacles such as roads.

In order to answer the following questions, you should use the scale presented in Table 1. Below is an example demonstrating the use of the scale.

<u>Example</u>

A survey participant believes that the deck rating has a strong importance in comparison to the substructure rating. Its preference is shown in the following scale by marking “X” on the number 5 on the left side.

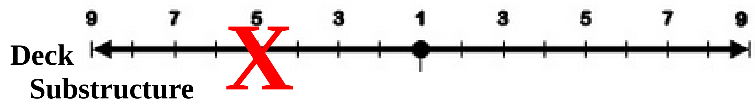

Table 1. The Fundamental Scale of Absolute Numbers

Intensity of Importance	Definition	Explanation
1	Equal Importance	Two ratings contribute equally to the objective
2	Weak or slight	
3	Moderate importance	Experience and judgement slightly favor one rating over another
4	Moderate plus	
5	Strong importance	Experience and judgement strongly favor one rating over another
6	Strong plus	
7	Very strong or demonstrated importance	A rating is favored very strongly over another, its dominance demonstrated in practice
8	Very, very strong	
9	Extreme importance	The evidence favoring one rating over another is of the highest possible order of affirmation

Non-Water-Crossing Bridges

1. In your opinion, please indicate which rating is more important than the other in the following three pairwise comparisons. Please show your result by marking an X on the numerical scale. (See the explanation of the scale in Table 1 and example in page 2)

   a. **Deck** 9 7 5 3 1 3 5 7 9 **Substructure**

   b. **Deck** 9 7 5 3 1 3 5 7 9 **Superstructure**

   c. **Substructure** 9 7 5 3 1 3 5 7 9 **Superstructure**

2. The following comparisons should be performed under the context of **safety**. **Safety** is related to whether or not users may use the bridge without putting their life at risk.

   a. **Deck** 9 7 5 3 1 3 5 7 9 **Substructure**

   b. **Deck** 9 7 5 3 1 3 5 7 9 **Superstructure**

   c. **Substructure** 9 7 5 3 1 3 5 7 9 **Superstructure**

3. The following comparisons should be performed under the context of **serviceability**. **Serviceability** is defined as the bridge meeting the objectives of its intended use.

   a. **Deck** 9 7 5 3 1 3 5 7 9 **Substructure**

   b. **Deck** 9 7 5 3 1 3 5 7 9 **Superstructure**

   c. **Substructure** 9 7 5 3 1 3 5 7 9 **Superstructure**

4. The following comparisons should be performed under the context of **comfort**. **Comfort** is defined as the customers' satisfaction of using/riding on the bridge.

		9	7	5	3	1	3	5	7	9	
a.	**Deck**	9	7	5	3	1	3	5	7	9	**Substructure**
b.	**Deck**	9	7	5	3	1	3	5	7	9	**Superstructure**
c.	**Substructure**	9	7	5	3	1	3	5	7	9	**Superstructure**

5. The following comparisons should be performed under the context of **resiliency**. **Resiliency** is referred to the ability of the bridge to absorb catastrophic impacts (Natural: tornado, earthquakes, etc. Man-made: collisions) with timely returns to normalcy.

a.	**Deck**	9	7	5	3	1	3	5	7	9	**Substructure**
b.	**Deck**	9	7	5	3	1	3	5	7	9	**Superstructure**
c.	**Substructure**	9	7	5	3	1	3	5	7	9	**Superstructure**

6. If you have a total of 100 points to allocate among the three main bridge condition ratings (deck, superstructure and substructure) used in the NBI, how would you allocate the points based on their relative importance?

RATING	WEIGHT
Deck	
Superstructure	
Substructure	
Total	**100**

7. Is there any other rating or factor that you consider to be of importance in bridge maintenance/repairs decision-making? If yes, please provide a list of all ratings or factors.

Water-Crossing Bridges

1. In your opinion, please indicate which rating is more important than the other in the following pairwise comparisons. Please show your result by marking an X on the numerical scale. (See the explanation of the scale in Table 1 and example in page 2)

   a. **Deck** 9 7 5 3 1 3 5 7 9 **Substructure**

   b. **Deck** 9 7 5 3 1 3 5 7 9 **Superstructure**

   c. **Deck** 9 7 5 3 1 3 5 7 9 **Scour**

   d. **Substructure** 9 7 5 3 1 3 5 7 9 **Superstructure**

   e. **Substructure** 9 7 5 3 1 3 5 7 9 **Scour**

   f. **Superstructure** 9 7 5 3 1 3 5 7 9 **Scour**

2. The following comparisons should be performed under the context of **<u>safety</u>**. **<u>Safety</u>** is related to whether or not users may use the bridge without putting their life at risk

   a. **Deck** 9 7 5 3 1 3 5 7 9 **Substructure**

   b. **Deck** 9 7 5 3 1 3 5 7 9 **Superstructure**

   c. **Deck** 9 7 5 3 1 3 5 7 9 **Scour**

   d. **Substructure** 9 7 5 3 1 3 5 7 9 **Superstructure**

   e. **Substructure** 9 7 5 3 1 3 5 7 9 **Scour**

   f. **Superstructure** 9 7 5 3 1 3 5 7 9 **Scour**

3. The following comparisons should be performed under the context of **<u>serviceability</u>**. **<u>Serviceability</u>** is defined as the bridge meeting the objectives of its intended use.

   a. **Deck** 9 7 5 3 1 3 5 7 9 **Substructure**

   b. **Deck** 9 7 5 3 1 3 5 7 9 **Superstructure**

   c. **Deck** 9 7 5 3 1 3 5 7 9 **Scour**

   d. **Substructure** 9 7 5 3 1 3 5 7 9 **Superstructure**

   e. **Substructure** 9 7 5 3 1 3 5 7 9 **Scour**

   f. **Superstructure** 9 7 5 3 1 3 5 7 9 **Scour**

4. The following comparisons should be performed under the context of **comfort**. **Comfort** is defined as the customers' satisfaction of using/riding on the bridge.

   a. **Deck** 9 7 5 3 1 3 5 7 9 **Substructure**

   b. **Deck** 9 7 5 3 1 3 5 7 9 **Superstructure**

   c. **Deck** 9 7 5 3 1 3 5 7 9 **Scour**

   d. **Substructure** 9 7 5 3 1 3 5 7 9 **Superstructure**

   e. **Substructure** 9 7 5 3 1 3 5 7 9 **Scour**

   f. **Superstructure** 9 7 5 3 1 3 5 7 9 **Scour**

5. The following comparisons should be performed under the context of **resiliency**. **Resiliency** is referred to the ability of the bridge to absorb catastrophic impacts (Natural: tornado, earthquakes, etc. Man-made: collisions) with timely returns to normalcy.

   a. **Deck** 9 7 5 3 1 3 5 7 9 **Substructure**

   b. **Deck** 9 7 5 3 1 3 5 7 9 **Superstructure**

   c. **Deck** 9 7 5 3 1 3 5 7 9 **Scour**

   d. **Substructure** 9 7 5 3 1 3 5 7 9 **Superstructure**

   e. **Substructure** 9 7 5 3 1 3 5 7 9 **Scour**

   f. **Superstructure** 9 7 5 3 1 3 5 7 9 **Scour**

6. If you have a total of 100 points to allocate among the four main bridge condition ratings (Deck, superstructure, substructure, and scour) used in the NBI, how would you allocate based on their relative importance?

RATING	WEIGHT
Deck	
Superstructure	
Substructure	
Scour	
Total	**100**

7. Is there any other rating or factor that you consider to be of importance in bridge maintenance/repairs decision-making? If yes, please provide a list of all ratings or factors.

Table A.2 Rank of Highway System Deficient Bridges in Oklahoma

Structure Number	County (Parish) Code	latitude	longitude	Year Built	Material	Deck	Sup	Sub	Scour	Rating	ADT class	Group	Subgroup	Highway Rank
151790000000000	109	35.44985833	-97.43229167	1960	Steel continuous	3	4	4	N	3.649	A	1	a	1
151160000000000	125	35.38202222	-97.00955000	1960	Steel	3	5	4	N	3.994	A	1	a	2
153620000000000	143	36.22062778	-95.85138889	1961	Steel	3	5	3	N	3.690	B	1	b	3
141110000000000	143	35.87655833	-96.01558889	1958	Steel	5	4	3	3	3.610	B	1	b	4
160360000000000	143	35.87648333	-96.01527778	1964	Steel	5	5	3	3	3.878	B	1	b	5
151250000000000	71	36.77360556	-97.34614167	1960	Steel	3	4	3	N	3.345	B	1	b	6
151240000000000	71	36.77376944	-97.34577222	1960	Steel	4	4	3	N	3.696	B	1	b	7
170280000000000	123	34.77952222	-96.71041111	1967	Steel continuous	4	4	3	3	3.439	B	1	b	8
170290000000000	123	34.77943889	-96.71045556	1967	Steel continuous	4	4	3	3	3.439	B	1	b	9
155330000000000	63	35.09713889	-96.41582222	1962	Steel	3	4	5	N	3.953	B	1	b	10
186100000000000	109	35.50958333	-97.57625000	1973	Steel continuous	5	4	5	N	4.655	A	1	c	11
170410000000000	143	36.16017500	-95.91487500	1967	Steel continuous	4	4	4	N	4.000	A	1	c	12
180500000000000	143	36.14521667	-95.99825556	1971	Steel	5	5	4	N	4.696	A	1	c	13
182820000000000	143	36.15256389	-95.98008333	1972	Steel	4	4	4	N	4.000	A	1	c	14
172840000000000	143	36.14262222	-95.95849444	1968	Steel	4	4	7	N	4.912	A	1	c	15
173020000000000	143	36.14074444	-95.95488056	1968	Steel	4	4	7	N	4.912	A	1	c	16
172830000000000	143	36.14301667	-95.95849444	1968	Steel	4	4	7	N	4.912	A	1	c	17
151230000000000	109	35.45133611	-97.43504722	1960	Steel	4	5	4	4	4.268	A	1	c	18
151220000000000	109	35.45167778	-97.43502222	1960	Steel	5	5	4	5	4.709	A	1	c	19
142030000000000	109	35.53316667	-97.45972222	1958	Steel	5	5	3	U	4.392	A	1	c	20
283950000000000	109	35.53320556	-97.45974722	1975	Steel	5	5	4	U	4.696	A	1	c	21
153430000000000	143	36.23482500	-95.84825000	1961	Concrete continuous	4	6	5	N	4.994	A	1	c	22
152010000000000	143	36.23858056	-95.84827222	1960	Steel	5	4	4	7	4.981	A	1	c	23

Structure Number	County (Parish) Code	latitude	longitude	Year Built	Material	Deck	Sup	Sub	Scour	Rating	ADT class	Group	Subgroup	Highway Rank
152020000000000	143	36.23868333	-95.84795278	1960	Steel	5	4	4	7	4.981	A	1	c	24
183420000000000	143	36.11084444	-96.01158333	1972	Steel continuous	4	6	5	N	4.994	A	1	c	25
194790000000000	143	36.08869167	-96.03635833	1978	Steel	7	4	3	N	4.749	A	1	c	26
167420000000000	143	36.18262222	-95.88887778	1966	Steel	5	5	4	N	4.696	A	1	c	27
151150000000000	125	35.38225278	-97.00998333	1960	Steel	4	6	4	N	4.690	A	1	c	28
172260000000000	49	34.66559167	-97.22272778	1968	Steel	4	4	5	5	4.561	A	1	c	29
172270000000000	49	34.66559722	-97.22330278	1968	Steel	5	4	5	5	4.732	A	1	c	30
175990000000000	49	34.77894444	-97.29220000	1969	Steel continuous	4	5	4	4	4.268	A	1	c	31
175900000000000	49	34.53275833	-97.18693333	1969	Steel continuous	6	6	3	3	4.317	A	1	c	32
175980000000000	49	34.77877500	-97.29272222	1969	Steel continuous	4	5	4	4	4.268	A	1	c	33
175910000000000	49	34.53266111	-97.18747778	1969	Steel continuous	7	6	3	3	4.488	A	1	c	34
169610000000000	143	36.16182222	-95.89584444	1967	Steel	4	6	6	N	5.298	A	2	a	35
170480000000000	109	35.46055556	-97.57622222	1967	Steel continuous	6	4	6	N	5.310	A	2	a	36
187730000000000	109	35.42751111	-97.57456389	1974	Steel	6	7	4	N	5.737	A	2	a	37
165540000000000	143	36.11570000	-95.90481389	1965	Concrete continuous	5	6	4	N	5.041	A	2	a	38
126240000000000	109	35.52754722	-97.52793056	1951	Steel	7	7	4	4	5.317	A	2	a	39
126230000000000	109	35.52733611	-97.52739167	1951	Steel	7	7	4	4	5.317	A	2	a	40
195140000000000	109	35.52078333	-97.54276667	1978	Steel continuous	5	6	4	8	5.787	A	2	a	41
195130000000000	109	35.52154167	-97.54276944	1978	Steel continuous	6	6	4	8	5.958	A	2	a	42
285790000000000	109	35.52920556	-97.51393056	1962	Steel	6	5	4	N	5.047	A	2	a	43
173480000000000	143	36.14298611	-95.95675833	1968	Steel	6	4	5	N	5.006	A	2	a	44
155690000000000	109	35.52919444	-97.51420278	1962	Steel	6	5	4	N	5.047	A	2	a	45
165550000000000	143	36.11531667	-95.90481111	1965	Concrete continuous	6	6	4	N	5.392	A	2	a	46
151110000000000	109	35.44523611	-97.42383333	1960	Steel continuous	6	6	4	N	5.392	A	2	a	47

Structure Number	County (Parish) Code	latitude	longitude	Year Built	Material	Deck	Sup	Sub	Scour	Rating	ADT class	Group	Subgroup	Highway Rank
128270000000000	143	36.08970556	-95.99554167	1952	Concrete continuous	6	6	4	N	5.392	A	2	a	48
184670000000000	109	35.51158056	-97.57513333	1973	Steel	4	7	5	N	5.339	A	2	a	49
167530000000000	17	35.46452222	-97.71879167	1966	Steel continuous	4	7	6	N	5.643	A	2	a	50
158400000000000	143	36.08875833	-96.00659722	1963	Concrete continuous	5	6	4	N	5.041	A	2	a	51
158390000000000	143	36.08876111	-96.00684444	1963	Concrete continuous	6	6	4	N	5.392	A	2	a	52
172240000000000	143	36.13826111	-96.10152500	1968	Steel	6	7	4	N	5.737	A	2	a	53
194710000000000	143	36.08921111	-96.03593056	1978	Steel	7	5	4	N	5.398	A	2	a	54
124060000000000	147	36.74751667	-95.93528889	1950	Concrete	5	4	5	8	5.542	A	2	a	55
167430000000000	143	36.18258056	-95.88940000	1966	Steel	5	6	4	N	5.041	A	2	a	56
192600000000000	143	36.12027500	-96.11645556	1976	Prestressed concrete *	7	5	4	8	5.398	A	2	a	57
192790000000000	143	36.12261667	-96.11643056	1976	Prestressed concrete *	7	4	5	8	5.884	A	2	a	58
169400000000000	87	35.02980556	-97.37573889	1967	Steel	4	6	5	N	4.994	B	2	b	59
139320000000000	47	36.40423611	-97.89006944	1957	Steel continuous	5	4	5	N	4.655	B	2	b	60
142040000000000	135	35.39794167	-94.44119722	1958	Steel	4	5	6	N	4.953	B	2	b	61
139220000000000	121	34.89183611	-95.78079444	1957	Steel	5	4	4	N	4.351	B	2	b	62
182710000000000	109	35.39348889	-97.33573333	1972	Steel	5	5	4	N	4.696	B	2	b	63
144090000000000	71	36.69513611	-97.34564444	1959	Concrete continuous	4	5	4	N	4.345	B	2	b	64
144080000000000	71	36.69513333	-97.34601389	1959	Concrete continuous	6	4	4	N	4.702	B	2	b	65
161750000000000	135	35.44796389	-94.78085556	1964	Steel	3	7	5	N	4.988	B	2	b	66
168100000000000	17	35.46773611	-97.72468889	1966	Steel continuous	4	5	4	N	4.345	B	2	b	67
157670000000000	143	36.15805000	-96.24683056	1963	Steel	3	7	4	N	4.684	B	2	b	68
157680000000000	143	36.15815556	-96.24653889	1963	Steel	4	7	3	N	4.731	B	2	b	69
128480000000000	73	35.87005833	-97.93257222	1952	Steel continuous	5	5	4	4	4.439	B	2	b	70

Structure Number	County (Parish) Code	latitude	longitude	Year Built	Material	Deck	Sup	Sub	Scour	Rating	ADT class	Group	Subgroup	Highway Rank
155340000000000	117	36.30113611	-96.46380278	1962	Steel	4	5	5	N	4.649	B	2	b	71
151020000000000	71	36.80383889	-97.34358889	1960	Steel	5	5	4	N	4.696	B	2	b	72
157700000000000	111	35.43300000	-95.98832778	1963	Steel	4	5	5	N	4.649	B	2	b	73
157710000000000	111	35.43265000	-95.98865278	1963	Steel	4	5	6	N	4.953	B	2	b	74
157300000000000	107	35.42211111	-96.29986944	1963	Concrete continuous	4	4	5	N	4.304	B	2	b	75
005500000000000	143	36.36562500	-95.90468333	1918	Concrete continuous	4	4	3	8	4.789	B	2	b	76
183000000000000	147	36.72796667	-95.82231667	1972	Prestressed concrete *	5	6	3	3	4.146	B	2	b	77
005510000000000	143	36.36562500	-95.90059444	1918	Concrete continuous	4	3	3	8	4.521	B	2	b	78
126220000000000	71	36.68079167	-97.11295000	1951	Steel	3	3	3	3	3.000	C	2	c	79
050190000000000	113	36.56491667	-96.31631389	1936	Steel	3	3	3	N	3.000	C	2	c	80
100760000000000	69	34.21986667	-96.70162222	1943	Steel	4	5	3	3	3.707	C	2	c	81
181450000000000	143	36.15969444	-95.98283611	1971	Steel continuous	8	6	4	N	6.094	A	3	a	82
181460000000000	143	36.16074167	-95.98480833	1971	Steel continuous	8	6	4	N	6.094	A	3	a	83
151890000000000	143	36.24175278	-95.84826944	1960	Steel	7	6	4	8	6.129	A	3	a	84
174990000000000	49	34.85196389	-97.32294167	1969	Steel	4	7	6	7	6.196	A	3	a	85
172250000000000	143	36.13793333	-96.10152500	1968	Steel	5	7	4	N	5.386	B	3	b	86
050470000000000	83	35.88006389	-97.43052222	1936	Steel	4	5	4	8	5.348	B	3	b	87
145210000000000	17	35.52741944	-98.28790556	1959	Steel continuous	5	5	4	8	5.519	B	3	b	88
040570000000000	37	35.98491944	-96.11348611	1933	Steel	4	5	5	8	5.639	B	3	b	89
180440000000000	143	36.15215278	-96.20834444	1971	Steel	7	6	4	N	5.743	B	3	b	90
161520000000000	135	35.45096667	-94.75456111	1964	Steel	4	5	5	8	5.639	B	3	b	91
182720000000000	109	35.39307500	-97.33572778	1972	Steel	6	5	4	N	5.047	B	3	b	92
161530000000000	135	35.45068889	-94.75573611	1964	Steel	3	4	6	8	5.491	B	3	b	93
145220000000000	17	35.52762778	-98.28809167	1959	Steel continuous	5	5	4	8	5.519	B	3	b	94

Structure Number	County (Parish) Code	latitude	longitude	Year Built	Material	Deck	Sup	Sub	Scour	Rating	ADT class	Group	Subgroup	Highway Rank
138860000000000	109	35.52716389	-97.50383889	1957	Steel	7	7	4	4	5.317	B	3	b	95
160800000000000	135	35.45045278	-94.75822222	1964	Steel	4	6	7	N	5.602	B	3	b	96
166150000000000	121	34.86061389	-95.59633889	1965	Steel continuous	5	3	6	8	5.565	B	3	b	97
143650000000000	103	36.20234722	-97.32821111	1959	Concrete continuous	7	6	4	N	5.743	B	3	b	98
143660000000000	103	36.20234444	-97.32784444	1959	Concrete continuous	7	6	4	N	5.743	B	3	b	99
192610000000000	133	35.21843889	-96.67143889	1976	Prestressed concrete *	7	5	4	8	5.861	B	3	b	100
168130000000000	111	35.55111111	-95.95166667	1966	Concrete	6	5	4	8	5.690	B	3	b	101
005260000000000	37	35.98838611	-96.59703889	1918	Concrete	5	4	4	8	5.251	B	3	b	102
151010000000000	71	36.80383889	-97.34321667	1960	Steel	6	5	4	N	5.047	B	3	b	103
099080000000000	115	36.80456111	-94.72762222	1942	Concrete	5	5	4	8	5.519	B	3	b	104
130790000000000	133	35.22110556	-96.64361667	1953	Concrete	4	6	5	5	5.097	B	3	b	105
130360000000000	133	35.22500278	-96.65890556	1953	Concrete	4	4	5	8	5.371	B	3	b	106
180370000000000	101	35.48245556	-95.27401111	1971	Concrete continuous	4	5	7	N	5.257	B	3	b	107
050170000000000	115	36.69686944	-94.95667778	1936	Steel	3	4	5	8	5.200	B	3	b	108
131200000000000	49	34.73508056	-97.21430556	1953	Steel continuous	5	4	5	8	5.542	B	3	b	109
195150000000000	33	34.10341111	-98.54159444	1978	Prestressed concrete *	7	7	6	4	5.899	B	3	b	110
139250000000000	133	35.17199167	-96.57850833	1957	Steel continuous	4	5	6	8	5.930	B	3	b	111
105630000000000	145	35.97917222	-95.38103889	1946	Steel	6	4	5	N	5.006	B	3	b	112
065890000000000	89	33.94075000	-94.75898611	1938	Steel	5	3	5	8	5.274	B	3	b	113
109640000000000	89	34.04189167	-94.62193056	1948	Steel	4	4	5	8	5.371	B	3	b	114
161250000000000	125	35.26103056	-96.92207778	1964	Steel	7	6	4	N	5.743	B	3	b	115
145200000000000	123	34.96626389	-96.92982500	1959	Steel	4	3	5	5	4.293	C	3	c	116
101200000000000	115	36.68503889	-94.91334167	1944	Steel	3	3	4	8	4.641	C	3	c	117

Structure Number	County (Parish) Code	latitude	longitude	Year Built	Material	Deck	Sup	Sub	Scour	Rating	ADT class	Group	Subgroup	Highway Rank
136910000000000	93	36.36955000	-98.45477222	1956	Steel	5	6	4	5	4.977	C	3	c	118
153750000000000	37	36.10817778	-96.42842222	1961	Steel	4	5	4	5	4.538	C	3	c	119
165740000000000	121	34.94876944	-95.84444444	1965	Prestressed concrete *	4	5	5	N	4.649	C	3	c	120
165730000000000	121	34.94853333	-95.84436111	1965	Prestressed concrete *	4	5	5	N	4.649	C	3	c	121
021350000000000	103	36.28984444	-97.14893056	1928	Steel	5	4	5	5	4.732	C	3	c	122
073420000000000	49	34.76018056	-97.19861111	1939	Steel	4	4	5	5	4.561	C	3	c	123
040400000000000	113	36.75668889	-96.23512778	1933	Steel	3	3	5	7	4.662	C	3	c	124
005200000000000	35	36.87382500	-95.01560000	1918	Concrete	5	4	3	8	4.960	C	3	c	125
132370000000000	105	36.69923611	-95.56093333	1954	Steel	6	4	5	5	4.903	C	3	c	126
174800000000000	71	36.69349167	-97.30083333	1969	Steel continuous	5	4	5	N	4.655	C	3	c	127
174790000000000	71	36.69388889	-97.30095833	1969	Steel continuous	5	4	4	N	4.351	C	3	c	128
155840000000000	37	36.15836389	-96.40048611	1962	Steel continuous	4	4	5	5	4.561	C	3	c	129
054780000000000	73	35.84199722	-97.72549167	1937	Steel	5	4	3	3	3.610	D	3	d	130
045920000000000	3	36.80677222	-98.24028611	1935	Steel	4	3	3	3	3.171	D	3	d	131
136900000000000	151	36.51429167	-98.88000556	1956	Steel	3	4	3	3	3.268	D	3	d	132
046030000000000	117	36.50411944	-96.72830556	1935	Steel	3	3	3	3	3.000	D	3	d	133
040390000000000	53	36.81048333	-97.62689167	1926	Steel	6	4	3	3	3.781	D	3	d	134
017370000000000	53	36.88570000	-97.67969167	1926	Steel	4	3	3	3	3.171	D	3	d	135
195010000000000	143	36.12033056	-96.11605278	1978	Prestressed concrete *	7	8	4	8	6.433	B	4	a	136
192210000000000	119	36.11608056	-97.01355278	1976	Prestressed concrete *	7	6	4	8	6.129	B	4	a	137
192340000000000	119	36.11605556	-97.00786944	1976	Prestressed concrete *	7	7	4	8	6.397	B	4	a	138
157720000000000	111	35.43290278	-95.98501944	1963	Steel	4	7	6	8	6.466	B	4	a	139
157730000000000	111	35.43254167	-95.98528333	1963	Steel	4	7	6	8	6.466	B	4	a	140

Structure Number	County (Parish) Code	latitude	longitude	Year Built	Material	Deck	Sup	Sub	Scour	Rating	ADT class	Group	Subgroup	Highway Rank
216560000000000	97	36.16624444	-95.34652222	1987	Prestressed concrete *	7	7	4	N	6.088	B	4	a	141
170180000000000	105	36.89338056	-95.62890278	1967	Steel continuous	4	4	7	8	5.953	C	4	c	142
167440000000000	27	35.01487500	-97.23470278	1966	Prestressed concrete *	5	7	4	8	6.055	C	4	c	143
073270000000000	89	34.03151111	-94.75533056	1939	Steel	4	5	7	8	6.221	C	4	c	144
192750000000000	119	35.98571389	-96.91508611	1976	Prestressed concrete *	7	4	5	8	5.884	C	4	c	145
065560000000000	13	34.00735833	-96.18300556	1938	Steel	4	7	7	8	6.757	C	4	c	146
141990000000000	113	36.66413333	-96.34747778	1958	Steel continuous	4	4	5	8	5.371	C	4	c	147
160480000000000	47	36.39687500	-97.68240278	1964	Steel	5	6	3	8	5.496	C	4	c	148
141820000000000	113	36.66507222	-96.35152500	1958	Steel continuous	4	5	6	8	5.930	C	4	c	149
105370000000000	153	36.39941667	-99.57356667	1946	Steel	4	5	4	8	5.348	C	4	c	150
192230000000000	139	36.77755833	-101.33102500	1976	Prestressed concrete *	4	4	4	8	5.080	C	4	c	151
037610000000000	29	34.44541389	-96.20658333	1932	Steel	5	4	6	8	5.833	C	4	c	152
169810000000000	119	36.11428611	-96.73911944	1967	Prestressed concrete *	6	7	4	8	6.226	C	4	c	153
185310000000000	77	34.89530833	-95.43724444	1973	Prestressed concrete *	4	8	6	8	6.734	C	4	c	154
179530000000000	71	36.81166111	-97.25578056	1970	Concrete	4	6	5	8	5.907	C	4	c	155
165230000000000	153	36.45090000	-99.39055000	1965	Steel	5	5	4	8	5.519	C	4	c	156
054560000000000	49	34.79654444	-97.13540000	1937	Steel	5	4	5	7	5.272	C	4	c	157
034190000000000	5	34.25873056	-95.83377778	1931	Steel	4	3	7	8	5.685	C	4	c	158
034220000000000	5	34.26113889	-95.85630278	1931	Steel	5	3	7	8	5.856	C	4	c	159
168050000000000	113	36.68867500	-96.71753611	1966	Steel continuous	4	7	6	8	6.466	C	4	c	160
037880000000000	153	36.35952500	-99.35376944	1932	Steel	4	5	4	8	5.348	C	4	c	161
110970000000000	73	35.86063611	-97.91951111	1949	Concrete	5	5	4	8	5.519	C	4	c	162
040090000000000	113	36.75760278	-96.24215556	1933	Steel	5	4	5	8	5.542	C	4	c	163

Structure Number	County (Parish) Code	latitude	longitude	Year Built	Material	Deck	Sup	Sub	Scour	Rating	ADT class	Group	Subgroup	Highway Rank
065570000000000	113	36.57446944	-96.31119167	1938	Steel	3	5	6	8	5.759	C	4	c	164
045550000000000	151	36.79778056	-98.70445556	1935	Steel	6	5	4	N	5.047	C	4	c	165
054840000000000	11	35.84103056	-98.22906389	1937	Steel	4	4	4	8	5.080	C	4	c	166
055140000000000	75	35.01566111	-99.11870278	1937	Steel	4	4	5	8	5.371	C	4	c	167
055190000000000	75	35.01547500	-99.12514444	1937	Steel	4	4	5	8	5.371	C	4	c	168
055090000000000	75	35.01446667	-99.13048611	1937	Steel	5	4	6	8	5.833	C	4	c	169
192720000000000	139	36.64393611	-101.21362222	1976	Prestressed concrete *	6	5	4	8	5.690	C	4	c	170
161880000000000	121	35.21921111	-95.59375278	1964	Steel continuous	3	5	6	8	5.759	C	4	c	171
037250000000000	67	34.07880278	-97.95904167	1932	Steel	4	6	6	8	6.198	C	4	c	172
161930000000000	93	36.18208611	-98.92086389	1964	Steel continuous	5	5	4	4	4.439	D	4	d	173
034290000000000	35	36.89059444	-95.09456944	1931	Steel	4	4	5	N	4.304	D	4	d	174
050250000000000	119	36.15390278	-96.78423333	1936	Steel	3	3	4	7	4.371	D	4	d	175
095290000000000	113	36.53119444	-96.71935000	1940	Steel	5	4	4	4	4.171	D	4	d	176
169790000000000	53	36.72805000	-97.79023611	1967	Steel continuous	6	5	4	5	4.880	D	4	d	177
073500000000000	67	34.13143333	-98.09820278	1939	Steel	4	4	5	5	4.561	D	4	d	178
095230000000000	53	36.81057500	-97.84258611	1940	Steel	4	3	5	6	4.563	D	4	d	179
040300000000000	113	36.82636111	-96.25304167	1933	Steel	3	3	3	8	4.350	D	4	d	180
186120000000000	65	34.63758333	-99.41075000	1973	Prestressed concrete *	6	5	4	8	5.690	D	5	a	181
045790000000000	3	36.80344444	-98.26592500	1935	Steel	4	4	4	8	5.080	D	5	a	182
124540000000000	3	36.80410278	-98.26088611	1950	Steel	4	5	4	8	5.348	D	5	a	183
034260000000000	153	36.22946944	-99.32857778	1931	Steel	4	5	5	8	5.639	D	5	a	184
034440000000000	153	36.26810000	-99.33573333	1931	Steel	4	5	5	8	5.639	D	5	a	185
054810000000000	51	35.04420000	-97.74391944	1937	Steel	6	4	4	8	5.422	D	5	a	186
037640000000000	117	36.29333611	-96.78224722	1932	Steel	4	5	5	8	5.639	D	5	a	187

Structure Number	County (Parish) Code	latitude	longitude	Year Built	Material	Deck	Sup	Sub	Scour	Rating	ADT class	Group	Subgroup	Highway Rank
049650000000000	35	36.69897778	-95.42155278	1936	Steel	3	4	7	8	5.782	D	5	a	188
055180000000000	79	34.77821944	-94.63994722	1937	Steel	4	5	6	8	5.930	D	5	a	189
165350000000000	67	33.98859444	-97.94090000	1965	Steel	5	5	4	7	5.249	D	5	a	190
144970000000000	89	34.39123333	-94.69400278	1959	Steel continuous	4	5	6	8	5.930	D	5	a	191
045950000000000	3	36.80571111	-98.24844722	1935	Steel	4	4	4	8	5.080	D	5	a	192
034240000000000	35	36.98534722	-95.08303056	1931	Steel	3	4	5	8	5.200	D	5	a	193
033600000000000	69	34.32935278	-96.65788611	1931	Steel	4	5	6	7	5.660	D	5	a	194
169780000000000	53	36.72286111	-97.80118333	1967	Steel continuous	6	5	4	8	5.690	D	5	a	195
040010000000000	53	36.69500833	-97.53807500	1933	Steel	3	5	4	8	5.177	D	5	a	196
040720000000000	151	36.81208889	-99.07800278	1933	Steel	4	4	5	8	5.371	D	5	a	197
042330000000000	53	36.68140278	-97.56314722	1934	Steel	5	4	5	8	5.542	D	5	a	198
040520000000000	151	36.80953333	-99.02085278	1933	Steel	4	5	5	8	5.639	D	5	a	199
175470000000000	71	36.68349167	-97.11986944	1969	Steel	4	3	5	8	5.103	D	5	a	200
175480000000000	71	36.68319444	-97.11997222	1969	Steel	4	3	5	8	5.103	D	5	a	201
094390000000000	53	36.81060278	-97.78092500	1940	Steel	3	4	5	8	5.200	D	5	a	202
040850000000000	17	35.54027778	-98.32277778	1933	Steel	5	4	5	7	5.272	D	5	a	203
105600000000000	53	36.81103056	-97.98430278	1946	Steel	5	4	5	8	5.542	D	5	a	204
095120000000000	53	36.81056667	-97.78461944	1940	Steel	7	4	4	8	5.593	D	5	a	205
105590000000000	53	36.81118889	-98.00083611	1946	Steel	5	4	6	8	5.833	D	5	a	206
042190000000000	151	36.81217500	-99.11845833	1934	Steel	4	4	6	8	5.662	D	5	a	207
040480000000000	79	34.71487500	-94.55565278	1933	Steel	4	4	6	8	5.662	D	5	a	208
017140000000000	53	36.89604167	-97.67274167	1926	Steel	4	4	6	8	5.662	D	5	a	209
107340000000000	97	36.38823611	-95.05906389	1947	Steel continuous	3	4	5	8	5.200	D	5	a	210
135310000000000	25	36.62917500	-102.68186389	1955	Steel continuous	4	5	4	8	5.348	D	5	a	211

Structure Number	County (Parish) Code	latitude	longitude	Year Built	Material	Deck	Sup	Sub	Scour	Rating	ADT class	Group	Subgroup	Highway Rank
174810000000000	113	36.62975000	-96.69070278	1969	Prestressed concrete *	4	7	5	8	6.175	D	5	b	212
128490000000000	59	36.77005278	-99.36696389	1952	Steel	3	4	4	8	4.909	E	5	c	213
019180000000000	149	35.44845278	-99.16861111	1927	Steel	6	6	3	3	4.317	E	5	c	214
046010000000000	113	36.96944722	-96.19534444	1935	Steel	4	4	5	7	5.101	E	5	d	215
045930000000000	113	36.93871944	-96.20479167	1935	Steel	5	4	4	8	5.251	E	5	d	216
037240000000000	113	36.90584444	-96.20473333	1932	Steel	4	5	5	8	5.639	E	5	d	217

VITA

Cristian Contreras-Nieto

Candidate for the Degree of

Doctor of Philosophy/Education

Thesis: A DATA DRIVEN APPROACH TO SUPPORT BRIDGE ASSET MANAGEMENT

Major Field: Civil Engineering

Biographical:

Education:

Completed the requirements for the Doctor of Philosophy in Civil Engineering (Construction Engineering and Management) at Oklahoma State University, Stillwater, Oklahoma in December, 2017.

Completed the requirements for the Master of Science in Civil Engineering (Construction Engineering and Management) at Oklahoma State University, Stillwater, Oklahoma in May 2014.

Completed the requirements for the Bachelor of Science in Civil Engineering at Escuela Colombiana de Ingenieria, Bogota D.C., Colombia in 2003.

Experience:

Teacher Assistant, Oklahoma State University, School of Civil & Environmental Engineering, Stillwater, OK, from August to August 2017.

Research Assistant, Oklahoma State University, School of Civil & Environmental Engineering, Stillwater, OK, from January 2013 to August 2017.

Construction Auditor of Warehouse, Industrias Japan S.A., Tocancipá, Cund, Colombia, from October 2010 to July 2011.

Part-time Professor, Escuela Colombiana de Ingenieria, Bogota D.C., Colombia, from January 2010 to May 2010 and January 2005 to December 2007.

Professional Memberships:

American Society of Civil Engineers (ASCE), Student Member
American Institute of Steel Construction, Student Member
Chi Epsilon National Civil Engineering Honor Society
The Honor Society of Phi Kappa Phi