DEVELOPMENT OF BRIDGE MANAGEMENT TOOLS FOR PREDICTING BRIDGE REPLACEMENT PROJECTS

by

Aidan Nicholai Alar

A thesis submitted to the faculty of
The University of North Carolina at Charlotte
in partial fulfillment of the requirements
for the degree of Masters of Science in
Civil and Environmental Engineering

Charlotte

2018

Approved by:

____

Dr. Matthew J. Whelan

____

Dr. Tara L. Cavalline

____

Dr. Brett Tempest

ABSTRACT

AIDAN NICHOLAI ALAR. Development of Bridge Management Tools for Predicting Bridge Replacement Projects. (Under the direction of DR. MATTHEW J. WHELAN)

One of the responsibilities of the NCDOT is the prioritization of highway bridge replacement projects throughout the North Carolina transportation system. The Priority Rating Index (PRI) is a multi-criteria formula that is currently used to provide a score based on condition and functional data in the Bridge Management System (BMS) for each of the highway bridges in North Carolina to aid in ranking the priority of potential replacement projects. The PRI comprehensively utilizes many of the performance measures considered to be important by the NCDOT Structures Management Unit, the group responsible for maintaining North Carolinas highway bridge network. However, anecdotal evidence from Structures Management Unit personnel, supported by an analysis of PRI score distributions among bridges selected and not selected for replacement, suggests that the PRI is a poor indicator of whether a bridge will actually be scheduled for replacement. In addition, the PRI double counts some performance measures, uses nonlinear and case-based formulas that do not produce a transparent link between measures and priority, and neglects some important maintenance related considerations that influence priority for replacement. Therefore, the purpose of this study is to develop an objective decision-support tool for prioritizing bridge replacement candidates that accounts for the multiple goals and preferences of the Structures Management Unit. Critical criteria and performance measures are proposed through a review of BMS improvements from other states as well as discussions with the Structures Management Unit. Additionally, new performance measures are introduced to incorporate historical maintenance burden and current maintenance needs. The trade-off preferences of the Structure Management Unit for each of the performance measures are modeled with value functions through a developed Excel VBA macro. Data driven prioritization formulas are created through statistical regression of binary logistic and constrained linear least squares models using the statewide bridge inventory to result in utility functions that provide a priority score for each of the replacement projects. The statistical models provide insight on the performance measures that have been statistically linked to bridge replacement projects as well as their relative importance. Analysis of the predictive accuracy for binary classification of projects, distributions of prioritization scores, and odds ratios computed from the predicted prioritization scores are used to compare the performance of the models and arrive at a recommended best model. Risk attitudes are incorporated with logistic regression and constrained linear least squares, resulting in a utility function that provides a priority score for each of the replacement projects. The results of this study seek to provide transparent, defensible ratings for bridge replacement projects that can be used in future budget planning and can be adjusted if the goals of the NCDOT change.

ACKNOWLEDGMENTS

I would like to thank my advisor Dr. Matthew J. Whelan for his inspirational support during this writing of this thesis and Dr. Tara L. Cavalline for always being a positive supportive person throughout this research effort. Furthermore, I would like to thank Dr. Brett Q. Tempest for accepting to be in my committee and NCDOT for funding this research.

I would also like to thank:

- My lab mates: Jon Moss, Steven Dulin, Timothy Kernicky, Yihong Cai, and Adam Howe.
- My fellow graduate NCDOT research students: Kelsey Lane and Patrick Phillips.
- My parents: Nari Delariarte-Stevens and Justin Stevens.
- My friends from the Venture outdoor adventure group, my improv group Sunday Supper, Charlotte Comedy Theater, Charlotte Storytellers, YesPlus meditation club, and UNCC Steel Bridge Team 2014-2015.

TABLE OF CONTENTS

LIST OF FIGURES viii

LIST OF TABLES xi

CHAPTER 1: INTRODUCTION 1

1.1. Introduction 1

1.2. Anticipated Contribution of the Research Effort 4

1.3. Organization of Thesis 5

CHAPTER 2: LITERATURE REVIEW 7

2.1. Current Prioritization Index Utilized by NCDOT 7

2.1.1. Reasons for the Shortcomings in the PRI 13

2.2. NCHRP 590 Overview 18

2.2.1. Indiana 25

2.2.2. California 29

2.2.3. Virginia 31

2.3. Other Examples of Structural Project Prioritization 37

CHAPTER 3: PERFORMANCE CRITERIA AND MEASURES AND DEVELOPMENT OF VALUE FUNCTIONS 42

3.1. Introduction 42

3.2. Proposed Performance Criteria and Measures 42

3.3. Newly Introduced Performance Measures 48

3.4. Performance Measure Data Sources 54

3.5. VBA Design 55

3.6. Value Function Development 59

CHAPTER 4: DEVELOPMENT OF STATISTICAL MODELS FOR PREDICTING BRIDGE REPLACEMENTS 77

4.1. Development of the List of Bridges Selected for Replacement 77

4.2. Overview of Constrained Linear Regression 79

4.3. Overview of Binary Logistic Regression 82

4.4. Development of Matrix of Statistical Models Investigated 85

4.5. Summary of Statistical Models Developed 86

4.5.1. CLLS Regression Models 87

4.5.2. Binary Logistic Regression Models 88

CHAPTER 5: ASSESSMENT OF STATISTICAL MODELS 95

5.1. Introduction 95

5.2. Method for Developing of Classification Threshold 95

5.3. Analysis of Predictive Accuracy of Models 97

5.4. Analysis of Distribution of Prioritization Scores 99

5.5. Summary and Recommended Prediction Model 109

CHAPTER 6: CONCLUSION 113

6.1. Summary of Research and Key Findings 113

6.2. Recommendations for Future Work 115

REFERENCES 117

Appendix A: Excel VBA Macro Script for Performance Measure Value Function Development 119

LIST OF FIGURES

FIGURE 2.1: Histogram of PRI Scores for Bridges Selected for Replacement 11

FIGURE 2.2: Histogram of PRI Scores for Bridges Not Selected for Replacement 11

FIGURE 2.3: Classification of Bridges with the Ten Highest PRI Values Currently in the State Inventory 12

FIGURE 2.4: Maximum Point Effect of Performance Measures in the PRI 15

FIGURE 2.5: Example of a Single-Criteria Value Function, adapted from Patidar, 2007. 21

FIGURE 2.6: Relationship Between Certain Equivalent and Expected Value, adapted from Skinner, 2009. 22

FIGURE 2.7: Effects of Risk on Value Based on Risk Attitudes, adapted from Skinner, 2009. 23

FIGURE 2.8: Example of a Single-Criterion Value and Utility Function, adapted from Patidar, 2007. 24

FIGURE 2.9: IBMS Performance Measure Disutility Functions, adapted from Sinha et al., 2009. 28

FIGURE 2.10: IBMS Disutility Standard Shapes, adapted from Sinha et al., 2009 29

FIGURE 2.11: VDOT Index Value Function for ADT/LN 35

FIGURE 3.1: PRI Proposed Performance Criteria and Measures 44

FIGURE 3.2: Relationship Between Maintenance Burden and Maintenance Needs In Terms of Yearly Maintenance Expenditures 50

FIGURE 3.3: VBA Program Details 57

FIGURE 3.4: Distribution of ADT ECDF-Derived Value Function Values and Linear Value Function Values 61

FIGURE 3.5: Example of Scaling of Values from NCDOT Prioritization 4.0 62

FIGURE 3.6: Graphical Illustration of the Development of Value Functions Based on the Empirical Cumulative Distribution: (a) the ECDF for bridge posting values, (b) the average value for each bridge posting value (c) the normalized ECDF-derived value function 65

FIGURE 3.7: Substructure condition rating as (a) raw unscaled rating system, (b) a linear value function and (c) an ECDF-based value function 66

FIGURE 3.8: Linear and ECDF Value Functions for Infrastructure Condition 67

FIGURE 3.9: Linear and ECDF Value Functions for Vulnerability 68

FIGURE 3.10: Linear and ECDF Value Functions for Mobility 69

FIGURE 3.11: Linear and ECDF Value Functions for Functionality 70

FIGURE 3.12: Linear and ECDF Value Functions for Maintenance Needs 71

FIGURE 3.13: Linear and ECDF Value Functions for Maintenance Burden 72

FIGURE 3.14: Linear and ECDF Value Functions for Reoccurring Maintenance Burden 73

FIGURE 3.15: Linear and ECDF Value Functions for Maintenance Needs to Replacement Cost Ratio 74

FIGURE 3.16: Linear and ECDF Value Functions for Maintenance Burden to Replacement Cost Ratio 75

FIGURE 3.17: Linear and ECDF Value Functions for Reoccurring Maintenance Burden to Replacement Cost Ratio 76

FIGURE 4.1: Relationship of Odds and Probability 84

FIGURE 4.2: Logistic Regression Overview 85

FIGURE 4.3: Substructure Condition Value Function with Example Odds Ratios 92

FIGURE 4.4: ADT Value Function with Example Odds Ratios 93

FIGURE 5.1: Model ETC: (a) Bridges Selected For Replacement (b) Bridges Not Selected For Replacement (c) Bridges with the Top 10 Scores 100

FIGURE 5.2: Model ERC: (a) Bridges Selected For Replacement (b) Bridges Not Selected For Replacement (c) Bridges with the Top 10 Scores 101

FIGURE 5.3: Model LTC: (a) Bridges Selected For Replacement (b) Bridges Not Selected For Replacement (c) Bridges with the Top 10 Scores 102

FIGURE 5.4: Model LRC: (a) Bridges Selected For Replacement (b) Bridges Not Selected For Replacement (c) Bridges with the Top 10 Scores 103

FIGURE 5.5: Model ETB: (a) Bridges Selected For Replacement (b) Bridges Not Selected For Replacement (c) Bridges with the Top 10 Scores 104

FIGURE 5.6: Model ERB: (a) Bridges Selected For Replacement (b) Bridges Not Selected For Replacement (c) Bridges with the Top 10 Scores 105

FIGURE 5.7: Model LTB: (a) Bridges Selected For Replacement (b) Bridges Not Selected For Replacement (c) Bridges with the Top 10 Scores 106

FIGURE 5.8: Model LRB: (a) Bridges Selected For Replacement (b) Bridges Not Selected For Replacement (c) Bridges with the Top 10 Scores 107

LIST OF TABLES

TABLE 2.1: Comparison of the Number of Currently Selected and Not Selected Bridge Replacement Projects by PRI Range 13

TABLE 2.2: Double Counted NBI Performance Measures in the PRI 14

TABLE 2.3: Criteria, Performance Measures, and Relative Weights Developed for the NCHRP Report 590 BMS Framework. 20

TABLE 2.4: Criteria, Performance Measures, and Relative Weights Developed for the IBMS Ranking Module (Adapted from Sinha et al. 2009.) 26

TABLE 2.5: Goals and Performance Measures Developed for the NCHRP Report 590 BMS Framework 30

TABLE 2.6: VDOT IF Score Explanatory Variables and Performance Measures 32

TABLE 2.7: Relative Weights Developed for VDOT IF Score Based on Different Methods 36

TABLE 2.8: Example of a Confusion Matrix, adapted from Salman and Salem, 2012 40

TABLE 2.9: Confusion Matrix for a Binary Logistic Regression Model for Predicting Critical Pipe Sections, adapted from Salman and Salem 2012 41

TABLE 3.1: Infrastructure Condition Ratings and Descriptions, adapted from Weseman, 1995 45

TABLE 3.2: NBI Scour Critical Ratings and Descriptions, adapted from Weseman 1995 47

TABLE 3.3: Proposed Scour Codes and Descriptions 48

TABLE 3.4: Definitions for Bridge Posting Scores 48

TABLE 3.5: Maintenance Action Classification Definitions and Examples 53

TABLE 3.6: Data Sources and Associated Performance Measures 56

TABLE 3.7: Assumed Proportionality for Bridge Replacement Priority 64

TABLE 4.1: Filtering Guidelines Used to Develop List of Bridges Selected For Replacement 80

TABLE 4.2: Summary of Statistical Regression Test Cases 86

TABLE 4.3: Relative Weights for Constrained Linear Least Squares Models (Relative Rank Shown in Parentheses) 88

TABLE 4.4: P-Values for Logistic Regression Models 89

TABLE 4.5: Odds Ratios of Binary Logistic Regression Models (Relative Rank Shown in Parentheses) 90

TABLE 4.6: Model Summarization of Logistic Regression Models 94

TABLE 5.1: Predictive Accuracy Test Scores for Each Model (Relative Ranking in Parentheses) 98

TABLE 5.2: Distribution of Bridges Among Three Thresholds For Logistic Regression Models 111

TABLE 5.3: Distribution of Bridges Among Three Thresholds For CLLS Models 112

TABLE 5.4: Distribution of Bridges Among Three Thresholds For PRI 112

CHAPTER 1: INTRODUCTION

1.1 Introduction

According to the latest ASCE report card, the current state of infrastructure, including as bridges, dams, and roads, in the U.S. is D+. North Carolina's infrastructure has an overall grade of C [ASCE, 2016]. North Carolina Department of Transportation (NCDOT) is responsible for maintaining the second largest state-maintained road network in the United States with highway assets value of $575 billion [NCDOT, 2016]. The highway assets consist of culverts, pavement, and bridges among other transportation structures. There are approximately 13,500 bridges in North Carolina with an estimated asset value of $60 billion. The Bridge Maintenance Improvement Plan (BMIP), the funding program for bridge replacement and rehabilitation projects has a funding need of $250 million per year [NCDOT, 2016]. The NCDOT goal for 2030 is to reduce the number of structurally deficient (SD) bridges, or bridges that are in relatively "poor condition," to 10% or less.

An important responsibility of NCDOT is to ensure that the transportation network of North Carolina safely and effectively meets the needs of the traveling public. Planning for preservation, maintenance, and replacement projects of bridges falls under the responsibilities of the Structures Management Unit (SMU) of NCDOT. Specifically, the SMU is responsible for prioritizing bridge replacement projects for all of North Carolina. The tool that the SMU engineers use to assist with this process is a scoring index called the Priority Replacement Index (PRI). The PRI takes into account multiple National Bridge Inventory (NBI) performance measures that are collected by NCDOT, such as inspection condition ratings, usage of statistics like the Average Daily Traffic (ADT), and potential vulnerabilities such as scour criticality, in order to assess the operational condition of a bridge and, ultimately, providing a score that indicates how well a bridge would be a candidate for replacement. The SMU bridge engineers have found that the PRI scores have not consistently provided prioritization scores that align with the bridges that the bridge engineers choose for replacement. An initial analysis of the PRI index has revealed, among other issues: double-counting of performance ratings, poor spread in the distribution of scores, and an observation that the calculation of the PRI is difficult to follow due to multiple sub-calculations that are specific to each performance measure. Ideally, a bridge replacement prioritization index should accurately reflect the preferences of the decision makers and follow a clear calculation methodology for transforming performance measure ratings to an overall priority score, neither of which are provided by the current prioritization index. Therefore, the objective of this thesis is investigate and develop an improved bridge replacement prioritization index that has the following characteristics:

- Transparent – Clear method of how a performance measure is converted into a overall replacement prioritization score.

- Data-Driven – Utilization of the NCDOT Bridge Management System (BMS) databases as well as the use of maintenance records.

- Normalized – Reduction of clustering of prioritization scores of bridges to allow bridges to be ranked effectively.

- Accurate – Results that reflect the engineers preferences in regards to bridge replacement selections.

- Comprehensive – Contains all significant performance measures that drive the decisions of bridge engineers for bridge selection.

One method to address the issue of developing clear and objective bridge project prioritization indices is through the use of value and utility functions within decision analysis. A value function a mathematical model that reflects a decision makers preferences and provides a conversion of performance measure ratings into a normalized value [Patidar et al., 2007]. The value function for each of the performance measures are then combined into a single function, called a multi-criteria utility function. This approach has been adopted in other bridge management systems (BMS) of other states including Indiana, California, and Virgina where multi-criteria utility functions have been developed and implemented to determine the criticality of bridges conditions [Sinha et al., 2009, Johnson, 2008, Moruza et al., 2016]. Additionally, the National Cooperative Highway Research Program (NCHRP) has published a report discussing guidelines for developing multi-objective BMS optimization strategies, which recommends the use of value and utility functions for quantifying decision-maker trade-offs for changes of performance measures [Patidar et al., 2007].

Development of performance criteria and measures for a new bridge replacement priority index for North Carolina was completed as part of an earlier effort [Lane, 2016]. However the development of value functions for these measures as well as the weighting and aggregation of these performance measures into a single prioritization formula has not yet been addressed. Additionally, while new performance measures to incorporate the impact of maintenance burden and maintenance needs have been proposed, the optimal approach for calculating the value functions for these performance measures to produce the best correlation with replacement priority has not yet been investigated.

1.2 Anticipated Contribution of the Research Effort

The main objective of this thesis is the development of a data-driven statistical tool to assist NCDOT bridge engineers in identifying and ranking the priority of bridge replacement projects throughout North Carolina. This thesis explores the current prioritization index and identifies shortcomings to be addressed in the revised prioritization method. Following guidance obtained through literature review, an improved model for prioritization of replacement projects is developed using utility theory and statistical regressions. Specifically, value functions are developed to model the preference structure of NCDOT engineers for each performance measure using readily available data collected in the NCDOT Bridge Management System (BMS). This data, which is derived from multiple sources, is cleaned, organized, and compiled into a single common database via an Excel VBA macro. Individual maintenance actions are defined and classified based on the effect of the action on prolonging the life of a bridge. Different functional forms of value functions, including linear models and empirical cumulative distribution (ECDF) models, are introduced and evaluated through the statistical analysis. Multiple-criterion utility functions are developed using both constrained linear least squares and binary logistic regression models. The models are then compared relative to each other and the existing PRI scores based on their predictive accuracy, distribution of scores, and classification odds ratios.

1.3 Organization of Thesis

The outline of this thesis is as follows:

- Chapter 2 presents a discussion of the current prioritization index used by the NCDOT, Priority Replacement Index (PRI), as well as identifies shortcomings of PRI in terms of ranking bridge replacement projects. Also, a literature review of the different prioritization techniques developed for bridge management systems (BMS) across of different state DOTs is presented. Furthermore, studies involving probabilistic methods for determining priority for replacement of critical sections of municipal sewer pipe networks based on historical data is reviewed.

- Chapter 3 introduces performance measures and criteria that were developed by concurrent research [Lane, 2016], as well as detailing the development of new maintenance burden and maintenance needs performance measures to incorporate historical maintenance actions and quantifying the severity of bridge conditions the element level, respectively. A description of the Excel macro that automates the process of database development and creation of the value functions for the performance measures is provided along with the models for the actual value functions used in this studies.

- Chapter 4 discusses the development of a matrix of statistical models for predicting bridge replacements to evaluate the best functional forms of value functions, maintenance cost performance measures, and the statistical model used to create the utility function. Background on binary logistic regression and constrained linear least squares (CLLS) regression methods is provided. The developed statistical models are presented alongside discussion of the relative importance of individual performance measures found to be statistically significant in each model.

- Chapter 5 provides a discussion and analysis of the matrix of statistical models from the previous chapter, starting with optimal threshold development to measure model effectiveness based on predictive accuracy and priority score distribution.

- Chapter 6 summarizes conclusions of the research effort and provides recommendations for future work in this area.

CHAPTER 2: LITERATURE REVIEW

2.1 Current Prioritization Index Utilized by NCDOT

The current system used by NCDOT for prioritization of bridge replacement projects is called the Priority Replacement Index (PRI). The PRI is a ranking system developed by the NCDOT and uses a combination of two previously used prioritization formulas, the Federal Highway Administration (FHWA) Sufficiency Rating and Deficiency Points, in conjunction with additional bridge infrastructure measures. The performance measures that are used for calculating the PRI are nationally utilized metrics that are indexed in the National Bridge Inventory (NBI) [Weseman, 1995]. The NCDOT maintains a database of the NBI measures as well as other bridge records offering data from recent inspections and maintenance in a commercial BMS called AgileAssets. This system integrates multiple databases maintained for North Carolina transportation structures, including bridges, culverts, and traffic signal structures, to provide a means of projecting future maintenance, repair, and replacement needs to optimize decision making for infrastructure investments. The performance measures for the current instance in time that are referenced by the PRI are stored in the Network Master database within the AgileAssets BMS.

The current PRI is computed on a 120 point scale, where the higher the amount of points a bridge is assigned on the scale, the more likely that the given bridge is a good candidate for replacement. Ideally, the PRI ranking is supposed to serve as an objective and actionable method for clearly distinguishing bridges requiring replacement rather than repair or rehabilitation and sorting the projects in order of priority. While there are no fixed thresholds used to identify replacement candidates, a general guideline has been suggested to separate the PRI scale into three ranges for replacement. Under this guideline, bridges with a PRI score from zero up to 30 are considered "poor candidates" for replacements, bridges with a score of 30 up to 50 are considered "good" candidates, and bridges with a score of 50 or higher are considered "very good" candidates for replacement. The PRI equation is

$$\begin{aligned}\text{PRI} = {} & .45(\text{Deficiency Points}) + .45(100 - \text{Sufficiency Rating}) \\ & + 1.25[28 - \text{Deck Condition} - \text{Superstructure Condition} \\ & - 2(\text{Substructure Condition})] + 10(\text{Temporary Shoring}) \qquad (2.1)\end{aligned}$$

Deficiency Points is a collection of performance measures developed through prior NCDOT-sponsored research that represent the level of inadequacy of a bridge in terms of expected functionality and is computed on a 100 point scale [Johnston and Zia, 1984]. Deficiency Points is an index designed to quantify the likelihood and urgency for a bridge replacement with higher point totals being associated with greater priority. There are four main performance criteria that are addressed in the Deficiency Points calculation: single vehicle load capacity, vertical roadway under/over clearances, estimated remaining life, and clear deck width. The performance criteria in the Deficiency Points calculation focus heavily on vehicle to bridge posting weight ratios, functionality appraisal ratings, geometry, Average Daily Traffic (ADT), and estimated remaining life.

The Sufficiency Rating is a federal rating that was previously used to determine eligibility for federal funding to repair or replace each bridge [Weseman, 1995]. It is an overall rating of structural adequacy, functionality, and essentiality of use and is computed on a 100 point scale. Since this rating evaluates sufficiency rather than deficiency, bridges with lower sufficiency ratings are often considered to be more suitable candidates for replacement. The sufficiency rating is based on four performance criteria: structural adequacy and safety, serviceability and functional obsolescence, essentiality for public use, and special reductions. The performance criteria are calculated through a number of both linear and nonlinear equations that utilize 19 different performance measures sourced from the NBI data. Further information on the sufficiency rating can be found in the NBI Recording and Coding Guide [Weseman, 1995].

The remaining components of the PRI are general condition ratings (deck, superstructure, and substructure) and an additional binary assignment of points that are incorporated if the structure has been provided with temporary shoring. The condition ratings are overall ratings of the three principal components of the bridge assigned by a bridge inspector using a 0 to 9 scale. Along with the additional points assigned to bridges with temporary shoring, the additional points provided to the PRI by these condition ratings are designed to provide greater priority for replacement to structures in an advanced state of deterioration with potentially significant reductions in load carrying capacity.

NCDOT bridge engineers believe that the PRI does a poor job of indicating which bridges should be replaced based on anecdotal evidence obtained through current and prior practice. Early in this study, an analysis was performed to investigate if the PRI is in fact a poor indicator for bridge replacement projects by evaluating the PRI scores of bridges that have been selected for replacement relative to the remaining bridges in the state inventory that are not scheduled to be replaced. This was based on bridge data sourced from the 2016 Network Master database along with a list of all bridges either currently being replaced or scheduled for replacement that was provided by the NCDOT. This analysis consisted of records for 13,826 bridges of which 1,249 or 9.03% were currently scheduled for replacement. The distributions of PRI scores among bridges selected for replacement and those not selected for replacement were used to evaluate the performance of the PRI as a means of classifying bridge replacement projects and to postulate reasons for shortcomings in the performance of the index.

The histograms of PRI scores for bridges that are selected for replacement and those not currently selected for replacement are shown in Figure 2.1 and Figure 2.2, respectively. On average, bridges selected for replacement do tend to have higher PRI scores than those not selected for replacement, however a closer examination of the data reveals issues within the index. First, the histograms reveal that bridges selected for replacement exhibit a bimodal distribution of scores that has mean and median values in the lower half of the index range and a spread that encompasses a large portion of the index range. Ideally, a prioritization index should clearly distinguish replacement projects from all other bridges with a distribution that is skewed toward the higher end of the index range. To illustrate, the dashed red lines in Figures 2.2 and 2.1 delineate the ideal distribution shapes.

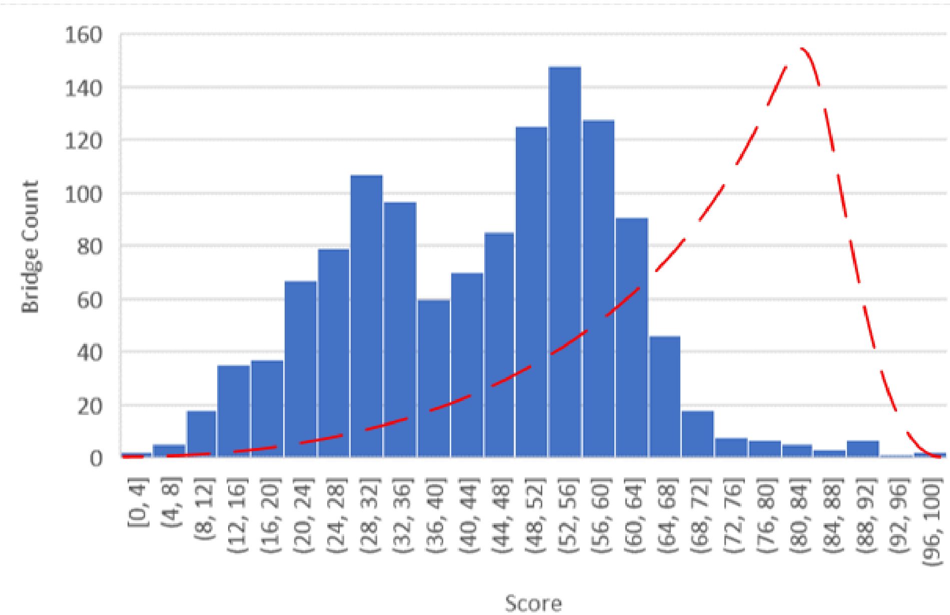


FIGURE 2.1: Histogram of PRI Scores for Bridges Selected for Replacement

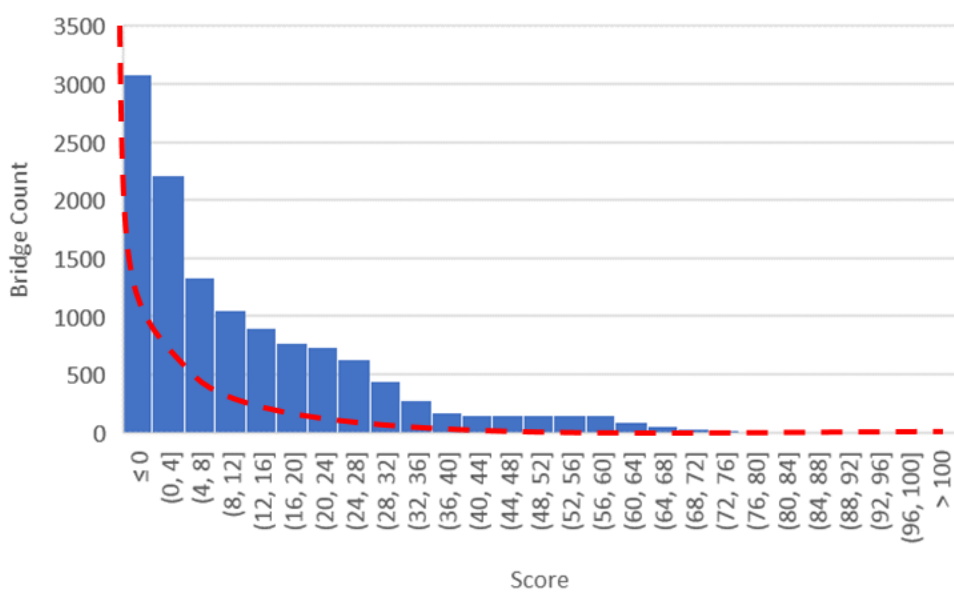


FIGURE 2.2: Histogram of PRI Scores for Bridges Not Selected for Replacement

An analysis of the PRI score at the extents of the distributions reveals that only two of the bridges within the ten highest PRI scores are currently selected for replacement, as shown in Figure 2.3. This analysis also reveals that the PRI does not utilize the full 120 point range when implemented in practice and that the majority of bridges are clustered at PRI scores below 30. While this would be acceptable if this skew was simply the result of proper classification of bridges not requiring replacement (which is the majority of bridges in the state) as low PRI structures, a comparative analysis of the number of bridges within the recommended ranges for qualitative classification of projects (Table 2.1) reveals classification issues. As evidenced by this table, there are more bridges that are not currently selected for replacement than those selected for replacement in both of the PRI ranges associated with "Very Good" and "Good" candidates. Collectively, this analysis supports the conclusion developed by NCDOT engineers that the PRI is an imperfect index for classification and prioritization of bridge replacement projects with numerous shortcomings.

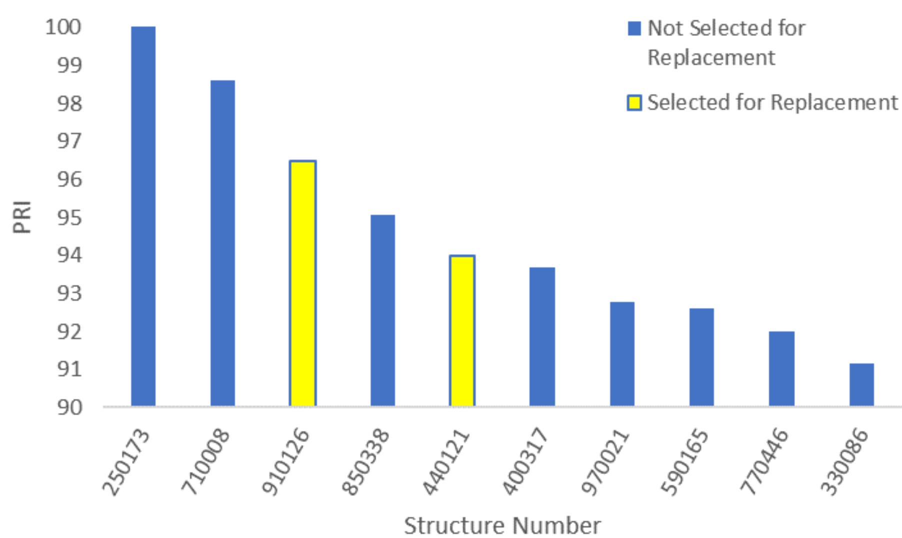


FIGURE 2.3: Classification of Bridges with the Ten Highest PRI Values Currently in the State Inventory

Table 2.1: Comparison of the Number of Currently Selected and Not Selected Bridge Replacement Projects by PRI Range

Replacement Candidacy	PRI Range	Selected	Not Selected
Very Good	PRI$\geq$ 50	537	588
Good	30$\leq$ PRI $<$ 50	409	1023
Poor	PRI $<$ 30	303	10966

2.1.1 Reasons for the Shortcomings in the PRI

One suspected reason for the apparent shortcomings in the PRI is that it was developed as an ad-hoc index collecting several prior indices that were not originally designed to be used together. As a result, the PRI suffers from significant double-counting of performance measures and does not present a clearly transparent view of how performance measures are individually weighted in computing the priority score. A count of the number of uses of each of the performance measures and the maximum potential contribution to the PRI by each measure was performed to illustrate these issues in the current index. Out of the 21 performance measures that are used for calculating the current PRI, seven suffer from either double, triple, or quadruple counting, as shown in Table 2.2.

An analysis of the maximum potential number of points that each performance measure can contribute to the PRI is presented in Figure 2.4 which is adapted from research from Lane (2016). Since several equations in the PRI are either conditional or nonlinear, the actual contribution of each measure to the index is dependent on the individual bridge characteristics (which may be viewed as another shortcoming of

TABLE 2.2: Double Counted NBI Performance Measures in the PRI

NBI Item	Performance Measure	Count
29	ADT Over	4
19	Detour Length	3
51	Clear Deck Width	2
53	Vertical Clearance Over	2
58	Deck Condition	2
60	Substructure Rating	2
59	Superstructure Rating	2

the index). The analysis reveals that the ADT carried by the bridge has the largest potential impact on the PRI score by affecting as many as 72.62 points, while the Defense Highway Designation has the smallest potential effect with only a maximum effect on 1.76 points. Overall, this analysis indicates that the PRI score is dominated by the ADT and general condition ratings, while a large number of the 21 performance measures have relatively little impact on the PRI score.

In addition to double counting of performance measures, another potential reason for the shortcomings of the PRI is that it incorporates general condition ratings rather than element-level inspection data. General condition ratings aggregate the inspector ratings to form a single condition rating for each primary component of the bridge, but do not offer the same granularity of information on the location and extent of structural deterioration that element-level ratings do. Inspection and rating of bridges at the element-level has been mandated by the FHWA as part of the MAP-21 legislation [MAP-21, 2012]. A report on the improvements to bridge inspections nationally showed that the NBI served as the main reporting system but did not include condition ratings that are granular enough for maintenance prioritization

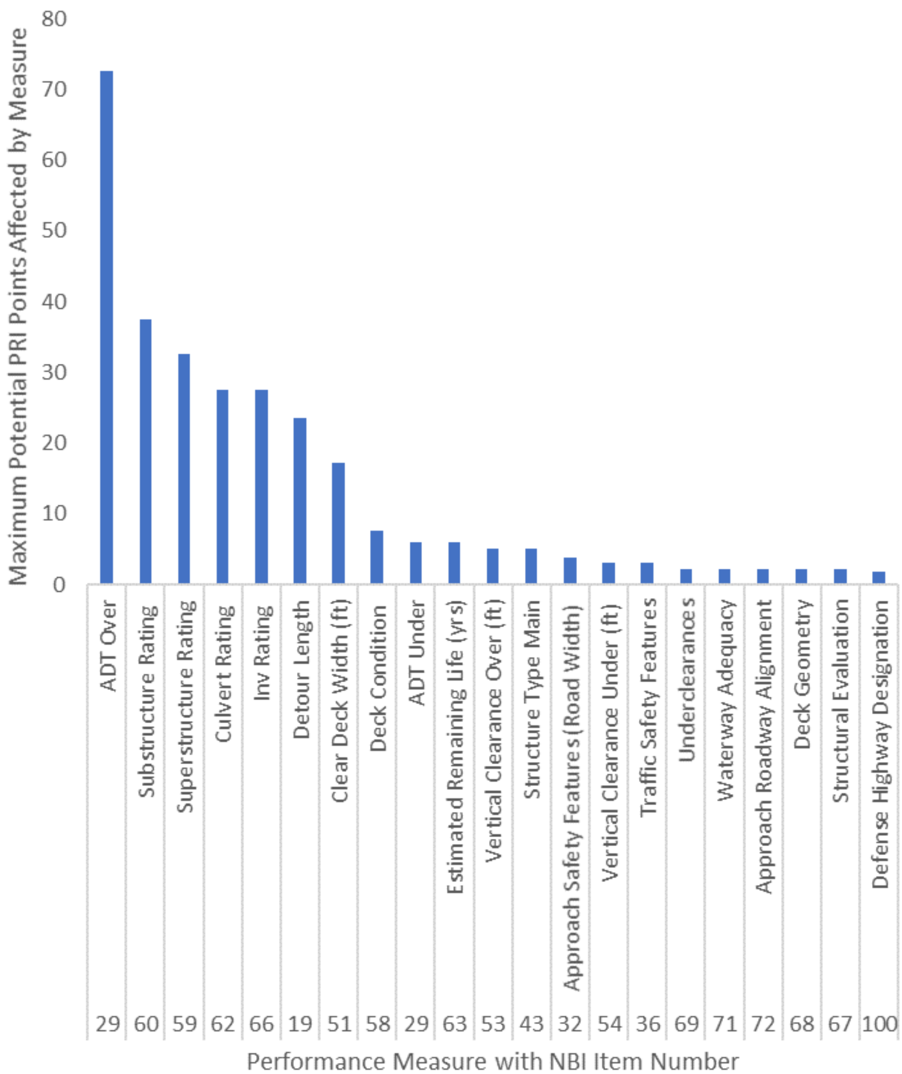


Figure 2.4: Maximum Point Effect of Performance Measures in the PRI

[Sobanjo and Thompson, 2016]. The PRI only has three condition ratings that address the overall state of large parts of a bridge and do not give inspectors the ability to record the localized deterioration across the elements of the structure and substructure of a bridge. Initially, the AASHTO Commonly Recognized Bridge Elements (CoRe) guidance was introduced as a system for element-level condition rating. In 2013, the AASHTO Manual for Bridge Element Inspection (MBEI) was released to provide a national standard for element inspection and recording [Farrar and Newton, 2014].

Since 2014, NCDOT bridge inspectors have recorded element-level health ratings for each bridge to stay compliant with the FHWA inspection and recording requirements. Furthermore, using the element-level health ratings, a list of inspector recommended maintenance needs is developed for each structure. These tasks range from insignificant actions, such as removing deck debris and maintaining handrails, to major rehabilitation, such as replacing timber piles and repairing modular bridge joints. Ideally, a bridge with low maintenance needs should be considered a candidate for repair or rehabilitation instead of replacement. NCDOT has recently introduced the storage of new element-level condition rating data in their BMS as well as integrated the database of Inspector Recommended Maintenance Needs required to correct low element condition ratings for each structure. In this database, the element-level condition ratings for each structure are associated with specific maintenance actions for repair of each element and aggregated into total counts. Furthermore, based on the condition rating of the elements, the corrective actions and counts are designated as either priority maintenance needs or recommended maintenance needs. Collectively this information allows for more detailed accounting of the type and number of elements requiring corrective action than the general condition ratings do and, further, provides a means for estimating the total cost of repairs.

A third shortcoming of the current PRI is that it does not consider the effects of maintenance history on the decision to prioritize replacement of a structure. Maintenance history is defined in this study as the maintenance actions that have been completed on each bridge and their associated costs. Maintenance history is likely to influence prioritization of bridge replacements in two ways. First, if the condition and rate of deterioration of a bridge has caused NCDOT to repeatedly perform maintenance year after year on the structure to maintain an acceptable level of service, then it is more likely to be a candidate for replacement. Such bridges present a maintenance burden that requires above average use of state resources and the cost-benefit ratio for such burdensome maintenance is unlikely to be a better economic decision than replacement. In contrast, bridges that have recently received major investments in either rehabilitation or preservation are less likely to be priority candidates for replacement, as returns on these investments in terms of increased service life are expected. For a period of approximately ten years, NCDOT engineers have maintained a digital record of the number of maintenance actions performed on each bridge as well as the cost of each action. There are 12,299 bridges in North Carolina with recorded maintenance history that occurred within the past ten years. The average number of actions per bridge is 4.78 actions over ten years with a standard deviation of 4.51 actions. Overall, there are 58,832 recorded actions that were performed in the past ten years at an average of 5,454 actions per year. However, this maintenance history is neither explicitly incorporated in the PRI or implicitly captured by any of the performance measures currently incorporated in the index.

The goal of a bridge prioritization index is to determine a preference order of candidate bridge replacement projects to facilitate data-driven project selection. The preference order ranks bridge replacement projects that have a higher priority than other project selections with a specific point value. This value should reflect both the current condition of the bridge as well as other criteria and considerations that the engineers and decision-makers use when actually selecting and prioritizing structures for replacement. When there are many criteria that have diverse or conflicting goals, one of the important factors is determining the trade-off in value between one selection candidate relative to another. In the following section, a literature review is presented to summarize recent studies that outline methods of ranking structures based on performance measures.

2.2 NCHRP 590 Overview

The National Cooperation for Highway Research Program (NCHRP) Report 590 set forth guidelines for state DOT bridge managers to incorporate multiple performance criteria for decision making related to bridge improvement projects . In addition to the current practice of choosing improvements with a focus on long term economic budgeting, the guidelines provide methods to include the effects of other criteria that are valuable for decision making. For each of the criteria, performance measures were identified to allow for a quantitative comparison of alternative bridge improvement projects. The criteria and performance measures were quantified and combined using value and utility theory, allowing preferences and risk attitudes of the decision maker to be objectively integrated when ranking improvement projects [Patidar et al., 2007].

Combining different performance measures presents some challenges, since the levels of each measure do not have a common scale. For example, NBI infrastructure condition ratings have an integer scale from 0 to 9, while the Health Index is on a 0 to 100 scale. Furthermore, the relative contributions, or weighting, of the individual performance measures to the combined index needs to reflect both preference and risk. The NCHRP 590 Report used utility theory in order to convert all performance measures to a common scale in a way that can be clearly understood and changed in the future as the needs of the bridge agency change. Utility theory provides a method of capturing and representing the preferences of decision makers in terms of trade-offs and how those preferences are affected by risk attitudes [Patidar et al., 2007]. The effects of bridge improvement project candidates on the criteria are calculated with a utility function, which is a mathematical representation of the preference structure.

The process that was used to develop utility functions in the NCHRP 590 Report consisted of three main steps: weighting, scaling, and amalgamation. The weighting step consisted of developing relative weights for the criteria and performance measures by using the results of a practitioner survey. The scaling step involved the development of single-criterion utility functions that represent the practitioner preference structure for individual performance measures. The final step, amalgamation, is the combination of the single-criterion utility functions into a single utility function that provides a single prioritization score for a bridge improvement project.

The five criteria and 12 performance measures that are considered in the NCHRP 590 illustrative decision model are summarized in Table 2.3. Each of the criteria are associated with a set of performance measures to provide a way of quantifying how each bridge project candidate contributes to the criteria. For example, there are three performance measures that are associated with the Preservation of Bridge Condition criterion, which are NBI Ratings, Health Index, and Sufficiency Rating. Each of the performance measures have different levels of importance to the decision maker and are assigned corresponding values, called relative weights, to quantify the importance preference of the decision maker. The Health Index with a value of 0.507 will have a larger impact for the Preservation of Bridge Condition criteria than either the sufficiency rating (0.222) or NBI ratings (0.271). Likewise, each of the individual criterion are assigned relative weight values to reflect their contribution to the overall index. For example, the Preservation of Bridge Condition criterion with a relative weight of 0.360 will impact the overall value of a bridge improvement project more significantly than User Cost Minimization with a weight of 0.110. These relative weights were determined using practitioner surveys.

Table 2.3: Criteria, Performance Measures, and Relative Weights Developed for the NCHRP Report 590 BMS Framework.

Criteria	Performance Measures
Preservation of Bridge Condition (0.360)	NBI Ratings (0.271) Health Index (0.507) Sufficiency Rating (0.222)
Traffic Safety Enhancement (0.205)	Geometric Rating (0.570) Inventory Rating (0.430)
Protection from Extreme Events (0.150)	Scour Vulnerability Rating (0.385) Fatigue/Fracture Criticality Rating (0.265) Earthquake Vulnerability Rating (0.205) Other Disaster Vulnerability Rating (0.145)
Agency Cost Minimization (0.175)	Initial Cost (N/A) Life-Cycle Agency Cost (N/A)
User Cost Minimization (0.110)	Life-Cycle User Cost (1)

The development of single-criterion utility functions in the scaling step first involves the creation of single-criterion value functions. A single-criterion value function provides a real number scalar representation of preference, known as value, of a decision maker for all levels of the criterion [Patidar et al., 2007]. A single-criterion value function was developed for each performance measure, with the exception of the cost related measures, such as life-cycle agency cost. For initial costs, life-cycle agency cost, and life-cycle user costs, a net present value analysis was performed. An example of a value function is shown in Figure 2.5, where the value to the bridge manager is shown for any level of the deck condition rating performance measure.

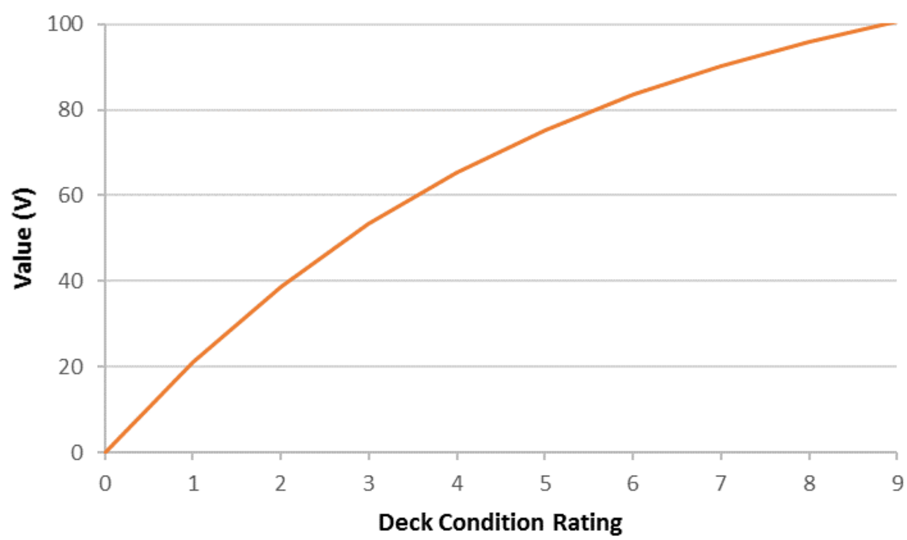


FIGURE 2.5: Example of a Single-Criteria Value Function, adapted from Patidar, 2007.

A single-criterion utility function takes the value function and adjusts the form based on the risk preferences of the decision maker. In utility theory, when the stakes are increased, the value of an alternative changes accordingly [Skinner, 2009]. For a risk averse or conservative decision maker, the value of a very risky alternative will be reduced. Utility theory allows an estimation of the value affected by risk, or utility, using a method called certain equivalence. Certain equivalence is a measure of value that a decision maker would place on the certainty of a potential outcome. The value difference between the expected value and the certain equivalence is known as the risk premium and is shown in Figure 2.6.

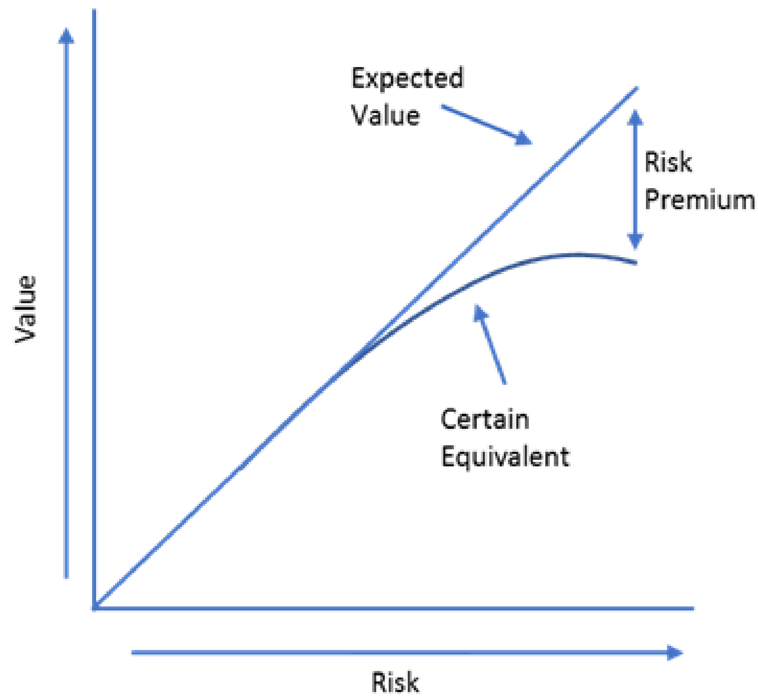


FIGURE 2.6: Relationship Between Certain Equivalent and Expected Value, adapted from Skinner, 2009.

In the NCHRP 590 Report, three types of risk attitudes were assessed for each of the performance measures: risk seeking, risk neutral, and risk averse. These risk attitudes can be modeled, respectively, by

$$u(z) \sim -e^{-cv(z)}, c > 0 \tag{2.2}$$

$$u(z) \sim v(z) \tag{2.3}$$

$$u(z) \sim e^{cv(z)}, c > 0 \tag{2.4}$$

where $u(z)$ is a single-criterion utility function, $v(z)$ is a single-criterion value function, z represents the level of a given performance measure, and c is a constant used to model the effect of risk on the utility. The effects of the different types of risk attitudes on the value of the decision are depicted in Figure 2.7. In the NCHRP 590 study, the type of risk for each performance measure was determined using the average of certainty equivalents from the gamble method portion of the practitioner surveys. Similar to the individual value functions, the single-criterion utility function is scaled from a range of lowest utility (0) to highest utility (100).

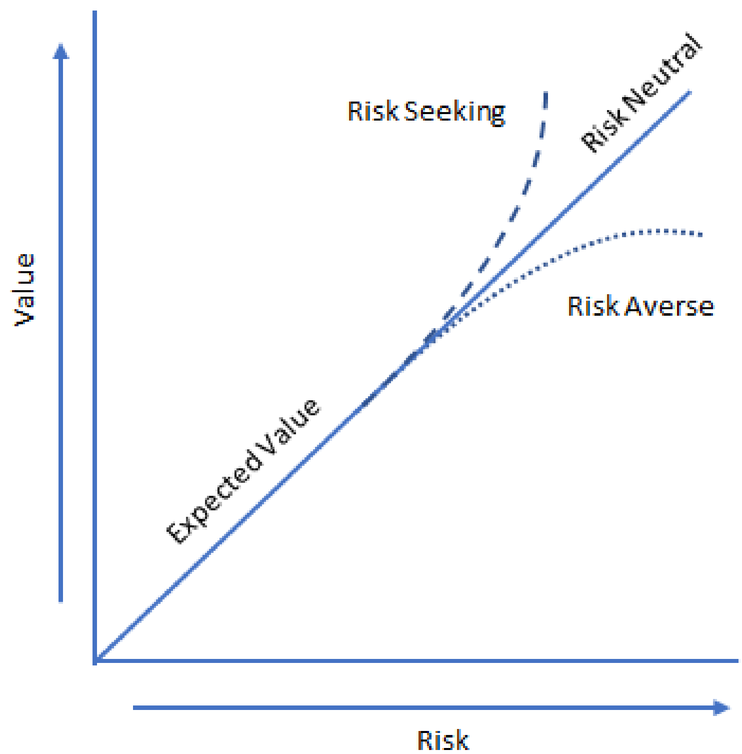


FIGURE 2.7: Effects of Risk on Value Based on Risk Attitudes, adapted from Skinner, 2009.

As an example, the single-criterion utility function and corresponding single-criterion value function for deck condition rating are shown in Figure 2.8. This performance measure was found to have an average certainty equivalent that correlated with the risk averse form, which is evident in the figure as all utility values are lower than the expected values. In the NCHRP 590 Report, the risk averse form of single-criterion utility functions was simplified as a linear function. The remainder of single-criterion utility functions were found to be risk neutral and thus modeled the same as the single-criterion value function.

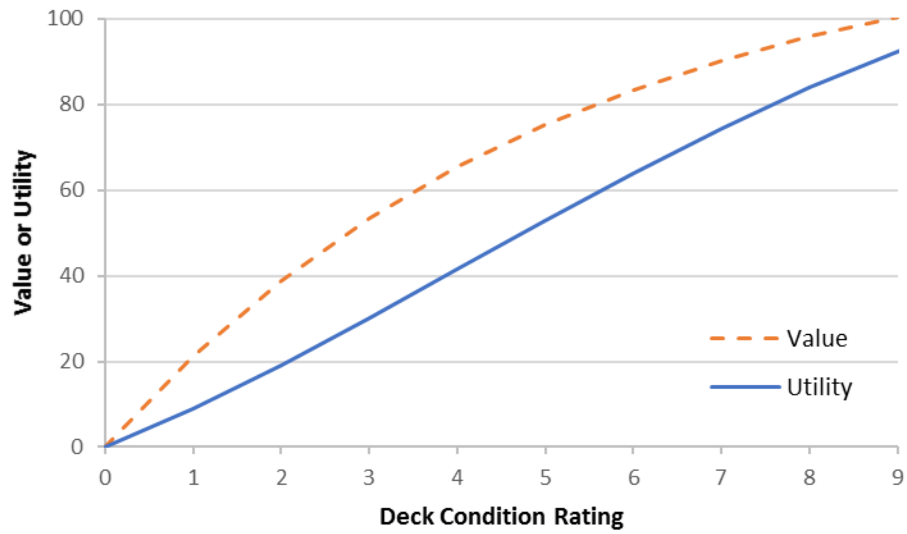


FIGURE 2.8: Example of a Single-Criterion Value and Utility Function, adapted from Patidar, 2007.

The final step, amalgamation involves the combination of the single criterion utility functions into a single, multi-criteria utility function. Scaling constants, which were derived from the practitioner survey, were used to determine the functional form used to create the utility function. Most commonly, multi-criteria utility functions are developed using either a multiplicative form

$$ku(z_1, z_2, ..., z_p) + 1 = \prod_{i=1}^{p} [kk_i u(z_i) + 1] \tag{2.5}$$

or an additive form of

$$u(z_1, z_2, ..., z_p) = \sum_{p}^{i=1} k_1 u_i(z_i) \tag{2.6}$$

where k, k_i are relative weighting scaling constants and $u_i(z_i)$ is a single criterion utility function. NCHRP 590 used the multiplicative utility form to develop their multi-criteria utility functions. Since the multi-criteria utility function aggregates all of the performance measures and criteria into a single score, it can represent the importance of a bridge improvement action based on the performance measures. With such an approach, projects can be directly compared under the assumption that if

$$u(z') > u(z'') \tag{2.7}$$

the bridge improvement candidate with the set of performance measures z' is preferable to another bridge improvement candidate with the set of performance measures z''.

2.2.1 Indiana

The Indiana Bridge Management System (IBMS) prioritizes bridge improvement actions using disutility change [Sinha et al., 2009]. Disutility represents the level of undesirability of the condition of a bridge based on the preferences of the bridge manager. A disutility function is the inverse of a utility function, where criteria with poorer levels are given a higher value. This results in bridges with the most need for improvement having the highest value. An important distinction in the approach used in this study relative to the NCHRP 590 report is that the disutility change used to develop the ranking value, is the difference in disutility value of the bridge with improvement and without improvement. The disutility change is based on "delta values" calculated in the Decision Tree (DTREE) module of the IBMS. Delta values are the projected increase of condition ratings and other performance measures expected to be caused by an improvement action. The performance measures, criteria and relative weights for this study were based on previous survey results obtained from an expert panel of bridge managers [Saito and Sinha, 1989] and are summarized in Table 2.4.

TABLE 2.4: Criteria, Performance Measures, and Relative Weights Developed for the IBMS Ranking Module (Adapted from Sinha et al. 2009.)

Criteria	Performance Measures
Economic Efficiency (10 Points)	Agency Cost (50%)
	User Cost (50%)
Bridge Condition Preservation (50 Points)	Structure Condition (40%)
	Remaining Service Life (40%)
	Wearing Surface (20%)
Bridge Safety Disutility (30 Points)	Clear Deck Width (30%)
	Vertical Clearance Over (10%)
	Horizontal Clearance Under (10%)
	Vertical Clearance Under (10%)
	Inventory Rating (40%)
Community Impact Disutility (10 Points)	Detour Length (100%)

The intent of the prioritization process was to improve economic efficiency, preserve bridge condition states, improve bridge traffic safety, and reduce community impact. The measure of economic efficiency was based on bridge life-cycle agency and user costs, which are calculated in the Life Cycle Cost (LCCOST) module of the IBMS. Preservation of bridge condition refers to maintaining the structural integrity and physical condition of a bridge and was measured by the minimum structure condition rating, remaining service life, and wearing surface. Bridge traffic safety describes the spatial adequacy and geometric design of a bridge and was based on the ratio of current levels to desirable levels for clear deck width, vertical clearance over, vertical clearance under, and horizontal clearance. The desirability levels for each measure are calculated in the DTREE module. The community impact criterion reflects the safety risk to commuters that use the bridge as well as the increase in delivery costs for nearby businesses and is measured by detour length. Specifically, the detour disutility function for the With Improvement scenario is calculated with

$$U_{DLB} = 100 - \frac{100 * (g_1 - DL)^n}{g_1^1} \tag{2.8}$$

where U_{DLB} is the disutility value without improvement; g_1 is the minimum detour length required for a disutility of 100; DL is the detour length, and n is a constant. The detour utility for the Without Improvement scenario is

$$U_{DLA} = \frac{(dl - dy) * U_{DLB}}{dl} \tag{2.9}$$

where U_{DLA} is disutility value with improvement, dl is the design life of the bridge, and dy is the number of years until replacement.

The disutility functions developed for each of the performance measures are reproduced in Figure 2.9. There are three main forms that the disutility functions take: linear, concave, and convex. These variations of the functional shape allow the model to reflect the risk attitudes of the panel of bridge experts (risk seeking, risk neutral, or risk averse). These standard shapes of disutility functions are shown in Figure 2.10. The inflection point at value 2 is the break point where disutility of a bridge is reduced and the second inflection point at 9 is the break point where the bridge is in perfect condition.

The disutility for an individual criterion, U_z, was calculated from the disutility

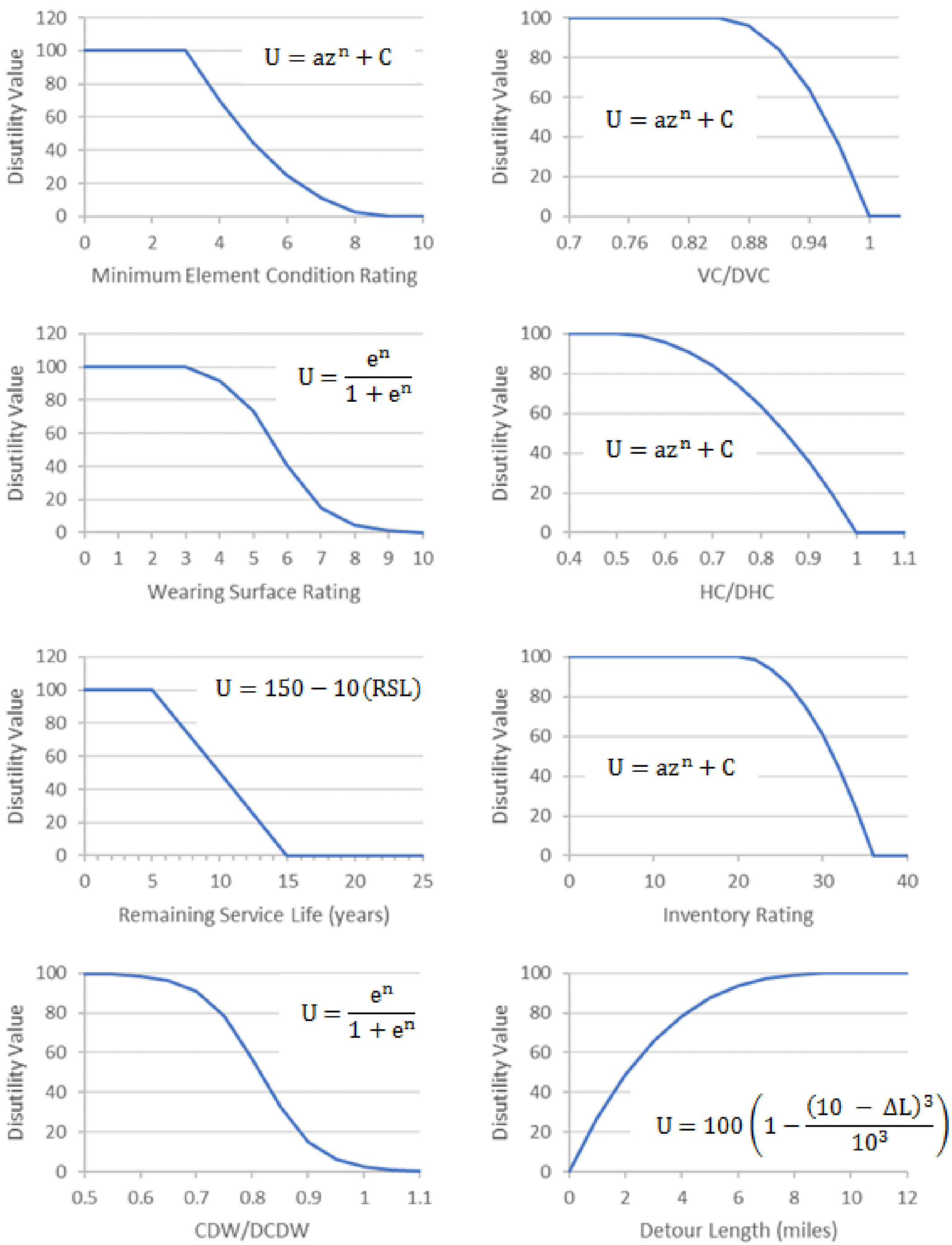


Figure 2.9: IBMS Performance Measure Disutility Functions, adapted from Sinha et al., 2009.

functions of associated performance measures with the following equation

$$U_z = \sum_{p}^{i=1} W_i U_i \tag{2.10}$$

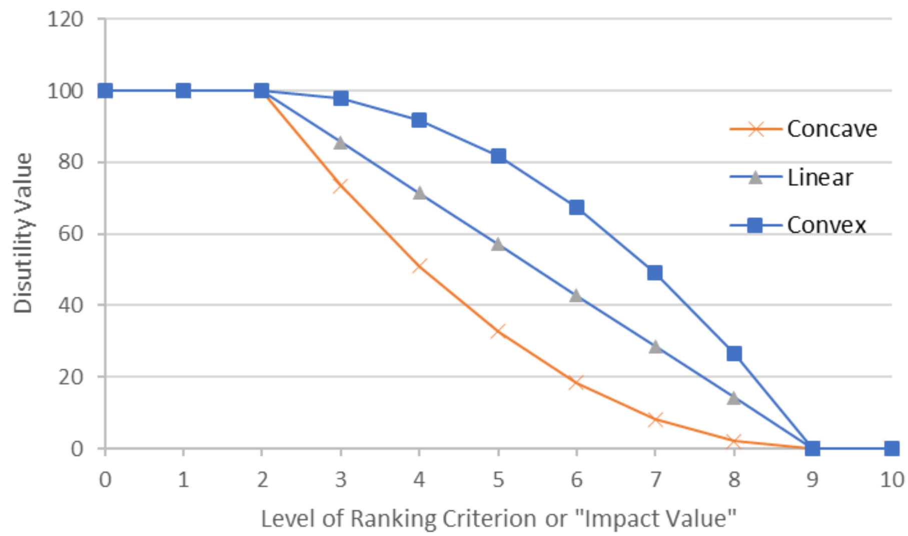


FIGURE 2.10: IBMS Disutility Standard Shapes, adapted from Sinha et al., 2009

where U_i is the disutility for an individual performance measure i and p is the number of performance measures associated with the criterion. Additionally, the composite disutility function U has the same additive functional form

$$U = \sum_{p*}^{z=1} W_z U_z \tag{2.11}$$

where p^* is the number of performance criteria.

2.2.2 California

The California Department of Transportation (Caltrans) combines benefit-cost ratios and utility functions to prioritize preservation projects among the 13,000 bridges in the California transportation system [Johnson, 2008]. One of the main benefits of using utility functions noted by the author is that it allows a clear way of combining criteria with different scales since the utility for each criteria are evaluated on a common scale from 0 to 1. Additionally, utility functions allow risk-associated criteria that were not previously included in previous project prioritization, such as scour potential and seismic risk, to be included. In this study, the utility benefit-cost ratio was calculated as

$$\text{Project Utility B/C Ratio} = U_t(TEV)/\text{Project Cost} \tag{2.12}$$

where U_t is the net project utility and TEV is the total element value of a bridge. The TEV is a quantitative method to describe the value of the bridge structure and allows the utility to be scaled by the size a bridge [Shepard and Johnson, 2001].

The utility functions developed in this study were based on five criteria: rehabilitation and replacement needs, scour needs, bridge rail upgrade needs, seismic retrofit needs, and mobility needs. The associated measures and relative weights for these measures that were developed in the study are shown in Table 2.5. In contrast to the prior studies, the performance measures do not have explicit relative weights. Instead, the utility functions were developed only for the criteria and not for each individual performance measure.

TABLE 2.5: Goals and Performance Measures Developed for the NCHRP Report 590 BMS Framework

Utility Component	Key Parameters
Rehabilitation and Replacement Needs (25 Points)	BHI, ADT, Repair Urgency (U), and DL
Scour (20 Points)	NBI SC, ADT, and DL
Bridge Rail Upgrade Needs (10 Points)	Caltrans seismic priority (S_v), ADT, and DL
Seismic Retrofit Needs (25 Points)	Caltrans rail upgrade score (RS)
Mobility Needs (20 Points)	Pontis improvement benefit (P)

The utility function for each of the criterion were developed with the logit form

$$X_i = \frac{1}{1 + e^{-C_i}} \tag{2.13}$$

where X_i is the utility function for each criteria and C_i is a function of the significant decision parameters for each component. The C functions were calculated in a previous iteration of the Caltrans BMS to calculate the value of each criteria. The net utility, or multi-criteria utility function, for a project is calculated using the additive functional form

$$U_t = \sum \alpha_i \beta_i X_i \tag{2.14}$$

where U_t is the utility, X_i is the i^{th} single-criterion utility function, α_i is a binary variable indicating whether representing the single-criterion utility function applied to a given structure, and β_i is the relative weight of a given single-criterion utility function.

2.2.3 Virginia

The Virginia DOT (VDOT) ranks all state-maintained transportation structures by transportation network importance using a cumulative score called the Importance Factor (IF) score in order to assist in the decision process of determining structure should have priority for maintenance, replacement, and rehabilitation expenditures [Moruza et al., 2016]. The IF score consists of nine explanatory variables, five of which are modeled using index value functions and the remaining four are binary variables that indicate if a structure is part of a defined highway system. These explanatory variables in the IF score are shown in Figure 2.6. The relative weights for each of the explanatory variables were determined using input from both the expert panel opinions as well as regression analysis. A process called backcalculated nonstandardized normalized coefficients (BNN) was applied to create the final relative weights in a manner that utilized both the results from the practitioner surveys and the statistical regression. An additive form of the multi-criteria utility function was used to combine the explanatory variable values into a single value.

TABLE 2.6: VDOT IF Score Explanatory Variables and Performance Measures

Variable	Name	Associated Performance Measures
A	ADT/LN	ADT, Number of Lanes
B	ADTT/LN	ADTT, Number of Lanes
C	AGR(ADT)	FADT, ADT, YFADT, YADT
D	Bypass Impact	Detour Length, ADT
E	Access Impact	Bypass Impact, POI, PROX
F	Base Highway	BHN
G	Strategic Highways	STRAHNET
H	Surface Transportation Action Agreement	STAA
I	Virginia Highway System	VSYS

Unlike most bridge prioritization formulas, the IF score does not have explanatory variables that use physical condition inventory items, such as geometric ratings or structural condition scores (substructure, superstructure, and deck). Instead, the IF score uses inventory items that measure the current and future use of a structure, bypass impact, access imact, and association with designated highway networks. The current use of a structure is measured by ADT per lane (ADT/LN) and ADTT per lane (ADTT/LN) of the structure. The inclusion of lane data allows a measure of usage relative to the capacity of a structure. The Annualized Growth Rate, $AGR(ADT)$, is a measure of the estimated increase of usage for a structure each year and is calculated with the equation.

$$AGR(ADT) = \Big[\frac{FADT}{ADT}\Big]^{1/(YFADT-YADT+1)} - 1 \tag{2.15}$$

where :

$FADT$ = Future Average Daily Traffic;

$YADT$ = Year of Average Daily Traffic;

$YFADT$ = Year of Future Average Daily Traffic;

This formula was used instead of FADT since the base year and future year to calculate FADT for each structure is not consistent among transportation structures. Bypass Impact is the combined effect of ADT and detour length if a structure was closed. The Access Impact variable represents the importance of a transportation structure based on the number of critical facilities, also referred to as Points of Interest (POI), in close proximity to the transportation structure. Schools, police and fire departments, and hospitals are examples of POIs. Information about critical facility locations were derived from the VDOT Geographic Information System (GIS) department. POIs within a three mile radius of a transportation structure were considered for the Access Impact calculation, where POIs closer to a transportation structure were assigned a higher value. The remainder of explanatory variables are binary indicators that show if a structure is a component of a designated highway network as defined by VDOT.

The preference structure for each of the explanatory variables, except for the highway network indicators, are modeled with value functions. In the context of the IF score, the value functions are referred to as index value functions. Each index value function uses raw data from the VDOT bridge inventory, however there were a variety of methods that were implemented to develop the final forms of the index value functions. The index value functions representing the preference structure of ADT/LN and $ADTT/LN$ were developed using empirical cumulative value functions (ECDF) and simplified using a step function. The index value function for $AGR(ADT)$ was created with an ECDF step function of the values developed with Equation 2.15. The Bypass Impact index value function was developed as the sum of the ECDFs for ADT and Bypass Detour Length (BYP). The Access Impact index is calculated using the BYP index value and proximity index value function with the equation

$$E_k = v(BYP) \times \Sigma_j [n_j \times v(PROX)] \tag{2.16}$$

$$\begin{aligned}
\text{where}: \ E_k &= \text{Index Value for Access Impact;} \\
v(BYP) &= \text{Index Value of Bypass Impact;} \\
n &= \text{Count of key locations;} \\
j &= \text{Distance interval a key location is in;} \\
v(PROX) &= \text{Distance from transportation structure to key location}
\end{aligned}$$

To provide an example of the index value functions, the ADT/LN index value function is shown in Figure 2.11. An index value of 1 was assigned to structures with an ADT/LN of 8500 and higher since these structures represented about 10% of the overall structure population. A similar method was applied for structures with an ADT/LN of 23 or lower. The index value function is not the actual ECDF, but a trendline that approximates the ECDF.

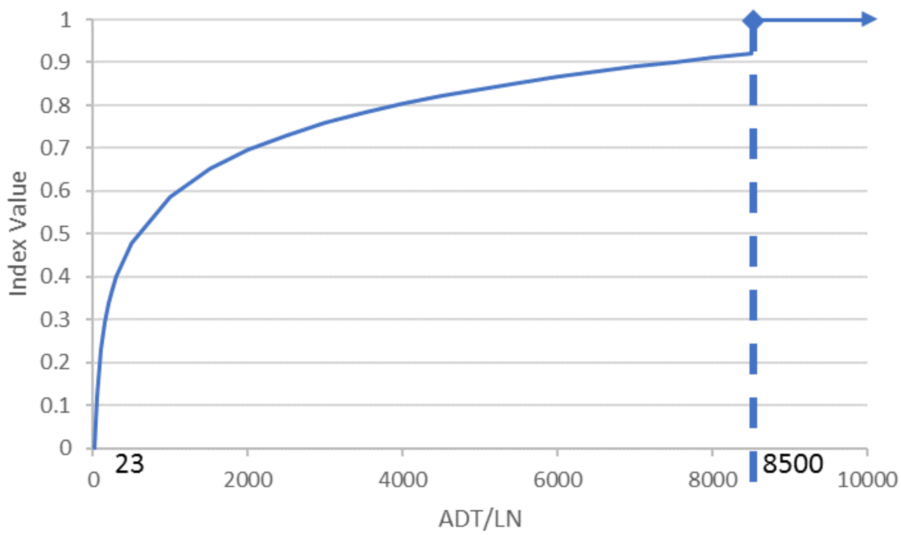


FIGURE 2.11: VDOT Index Value Function for ADT/LN

The relative weights for the explanatory variables were developed using BNN coefficients that reflect the intended and actual impacts of each variable. The intended relative weights were developed by the VDOT expert panel, while the actual relative weights were developed using linear regression analysis. Development of BNN coefficients, which is a method of adjusting actual impacts of a model to better reflect the desired impact, was applied to better align the performance of the model with the intentions of the expert panel using the equation

$$\beta_i^* = \beta_i \times (S_{X_i} \div S_Y) \tag{2.17}$$

where :

β_i^* is the standardized coefficient for standardized index values of variable i

β_i is the nonstandardized coefficient for nonstandardized index values of variable i

X_i is the set of nonstandardized index values for variable i

Y is the IF score

S is the standard deviation statistic

The relative weights developed by the expert panel, linear regression, and BNN coefficients are shown in Table 2.7.

TABLE 2.7: Relative Weights Developed for VDOT IF Score Based on Different Methods

Variable	Name	Panel	Regression	BNN
A	ADT/LN	0.20	0.345	0.147
B	ADTT/LN	0.10	0.161	0.079
C	AGR(ADT)	0.15	0.199	0.143
D	Bypass Impact	0.25	0.281	0.282
E	Access Impact	0.10	0.054	0.233
F	Base Highway	0.05	0.129	0.024
G	Strategic Highways	0.05	0.099	0.032
H	Surface Transportation Authorization Act Network	0.05	0.108	0.029
I	Virginia Highway System	0.05	0.102	0.031

The IF score of a transportation structure was developed using the additive multi-criteria utility function

$$\text{IF Score} = \sum \beta_i X_i \tag{2.18}$$

where X_i is the i^{th} index value and β_i is the BNN relative weight of a given index value. The calculation of the IF score for each of the VDOT structures was automated by an Excel VBA macro.

2.3 Other Examples of Structural Project Prioritization

Multi-criteria utility theory is implemented in other asset management prioritization practices, with significant prior work related to municipal sewer systems. The rising costs of emergency sewer pipe section repairs associated with the previous practice of random pipe section structural inspections motivated studies aimed at developing probabilistic models to predict if a pipe section is in a deficient state [Ariaratnam et al., 2001, Davies et al., 2001, Salman and Salem, 2012]. This way, inspection planning can be focused on potentially critical locations and in turn reduce the number of emergency repairs.

One of the promising methods identified in studies for developing an accurate probabilistic model is binary logistic regression, which allows the use of multiple predictor variables to estimate the probability of an event occurring. Pipe section data found in typical sewer system historical database records, including age, material, diameter, and waste type, were typically used as predictor variables. The event, or response variable, was typically either identified as the pipe section structural condition [Ariaratnam et al., 2001] or severity of potential failure modes [Davies et al., 2001]. The response variables had to be converted to binary response variables in order to meet the requirements of the binary logistic regression method. To accomplish this, one study simplified a pipe section structural deterioration integer scale of 1-5 to sections with a level of 5 would be assigned a binary rating of 1 and sections with any other condition level would be assigned a binary rating of 0 [Ariaratnam et al., 2001]. A similar simplification was applied in the study that used failure mode ratings [Davies et al., 2001].

The logit function used in the logistic regression can be utilized to calculate the estimated probability of an event occurring, $f(z)$ [Ariaratnam et al., 2001]. Through this approach, the estimated probability is

$$f(z) = \frac{1}{1 + e^{-z}} \tag{2.19}$$

where z is defined as

$$z = \beta_0 + \sum_{i=i}^{k} \beta_i X_i \tag{2.20}$$

where β_0 is a constant, X_i is the i^{th} predictor variable and β_i is the regression coefficient associated with X_i.

The remaining processes for creating the final regression model involve simplification and validation. First, the model was iteratively simplified based on the premise of reducing the number of insignificant predictor variables and improving the overall model based on Akaike Information Criterion (AIC) values, a process that is adopted in this thesis and will be discussed in a later chapter. One of the methods to validate the probabilistic models in these studies was to calculate the sensitivity, specificity, and predicted value of a positive result of a model (PV+) [Salman and Salem, 2012]. These statistical measures allow for the comparison between observed events and events predicted by the model. Each of these tests provides a percentage score that can be used to compare the predictive accuracy of different predictive models. In the Salman and Salem (2012) study, two types of logistic regression models were developed and the model with the best percentage scores among the three statistical measures discussed here was considered to be the best model. Additionally, 80% of the overall dataset, the calibration set, was used to develop the logistic model and the remaining 20% of the overall dataset, the validation set, was tested with the logistic model to determine if results between the two groups would be similar. This grouping test can be used to determine if the logistic model over-fits the dataset used to develop the original model, or if it is expected to perform well on other datasets.

Each of the validation methods are based on post-test terminology, which are referred to as: true positive (TP), true negative (TN), false positive (FP), and false negative (FN) [Glasser, 2008]. A true positive is when an event is observed and a model also predicted that the event would occur. A true negative is when an event does not occur and the model correctly predicted that the event would not occur. A false positive is when an occurrence of an event was not observed, but the model incorrectly predicted that the event would occur. Likewise, a false negative is when an occurrence of an event was observed, but the model predicted that the event would not occur. These statistics can be assembled into a table, called a confusion matrix. An example of the format of a confusion matrix is shown in Table 2.8, where a value of 0 represents an event not occurring, while a value of 1 represents an event occurring.

Sensitivity, specificity, and the predicted value of a positive result are computed from the values in the confusion matrix. Sensitivity is the percentage of correctly

Table 2.8: Example of a Confusion Matrix, adapted from Salman and Salem, 2012

Observed Condition	Predicted Condition 0	1
0	TN	FP
1	FN	TP

predicted event occurrences among the total instances of actual occurrences and is determined as

$$\frac{TP}{TP + FN} \tag{2.21}$$

Specificity is the percentage of correctly predicted event non-occurrences among the total instances of actual non-occurrences and is determined as

$$\frac{TN}{TN + FP} \tag{2.22}$$

The predicted value of a positive result (PV+) is the percentage of correctly predicted event occurrences among all predicted event occurrences and is determined as

$$\frac{TP}{TP + FP} \tag{2.23}$$

For all of these statistical measures, a greater percentage score reflects stronger predictive accuracy of the statistical model. However, by examining statistical measures, such as sensitivity and specificity, individually, one may examine how well the binary logistic model performs for the event occurrence and non-occurrence separately. To illustrate the process of comparing predictive models, take for example the confusion matrix developed for logistic binary regression model for predicting the deterioration states of pipe sections in Table 2.9.

Table 2.9: Confusion Matrix for a Binary Logistic Regression Model for Predicting Critical Pipe Sections, adapted from Salman and Salem 2012

Observed Condition	Predicted Condition		Correct Prediction (%)
	0	1	
0	4,683	1,187	79.8
1	1,612	1,616	50.1

Sensitivity of the binary logistic model is

$$\frac{1616}{1616 + 1612} = 50.1\% \tag{2.24}$$

while the specificity is

$$\frac{4,683}{4,683 + 1,187} = 79.8\% \tag{2.25}$$

and PV+ is

$$\frac{1,616}{1,616 + 1,187} = 57.7\% \tag{2.26}$$

The results of the binary logistic regression predictive accuracy tests were compared with test results of the multinomial logistic regression model, and the predictive accuracy for the binary regression were higher in two of the three tests (specificity and PV+). Therefore it was concluded that binary regression was a better model than multinomial regression for predicting sewer pipe conditions.

CHAPTER 3: PERFORMANCE CRITERIA AND MEASURES AND DEVELOPMENT OF VALUE FUNCTIONS

3.1 Introduction

This chapter presents the proposed performance measures and criteria that will be tested for significance for bridge replacement prioritization as well as the methodology for developing performance measure value functions. Specifically, this chapter will include descriptions of each criterion and associated performance measures, the use of bridge maintenance records as performance measures, performance measure data sources, details on value function development, and computed value functions for each performance measure. The performance measure value functions in this chapter are later used for developing bridge replacement prediction models in a subsequent chapter.

3.2 Proposed Performance Criteria and Measures

As a part of a larger overall research effort, Lane (2016) investigated infrastructure goals reflected in recent federal and state legislation and proposed a set of performance criteria and performance measures for prioritizing bridge replacement projects in North Carolina. A tree diagram detailing all of the performance criteria and associated measures considered in this study is shown in Figure **??**. The six performance criteria are: Infrastructure Condition, Vulnerability, Mobility, Functionality, Maintenance Needs, and Maintenance Burden. In general, each of the proposed criteria are merely overarching goals that are assessed by performance measures based on standard NBI measures, with the exception of the performance measures under the Maintenance Needs and Maintenance Burden that are based on bridge maintenance cost records. The relationship between criteria and performance measures for the proposed approach is similar to the relationship found in the IBMS, where instead of having four goals assessed by one to five performance measures, there are six criteria that are assessed with a similar range of performance measures [Sinha et al., 2009]. Each of performance measures for Infrastructure Condition, Vulnerability, Mobility, and Functionality will be described in this section, while the maintenance-based criteria and measures will be discussed separately in the following section due to the uniqueness of these measures and the data sources to compute them.

The Infrastructure Condition criterion is computed using the deck, superstructure, and substructure condition ratings, which are based on the general structural conditions of a bridge. These condition ratings are developed during biennial inspection and are rated on a scale of 0-9. The condition descriptions associated with each rating are prescribed by the NBI recording and coding guide [Weseman, 1995] and are reproduced in Table 3.1.

The Vulnerability criterion is based on bridge structural failure risk related to fracture and scour criticality. Fracture critical vulnerability is a binary indicator of the presence of a non-redundant tensile component of a bridge for which failure would result in collapse of the bridge. These components are known as fracture-critical members (FCMs) [Weseman, 1995]. The scour vulnerability rating indicates the risk of structural failure at bridge piers due to potential hydraulic erosion events.

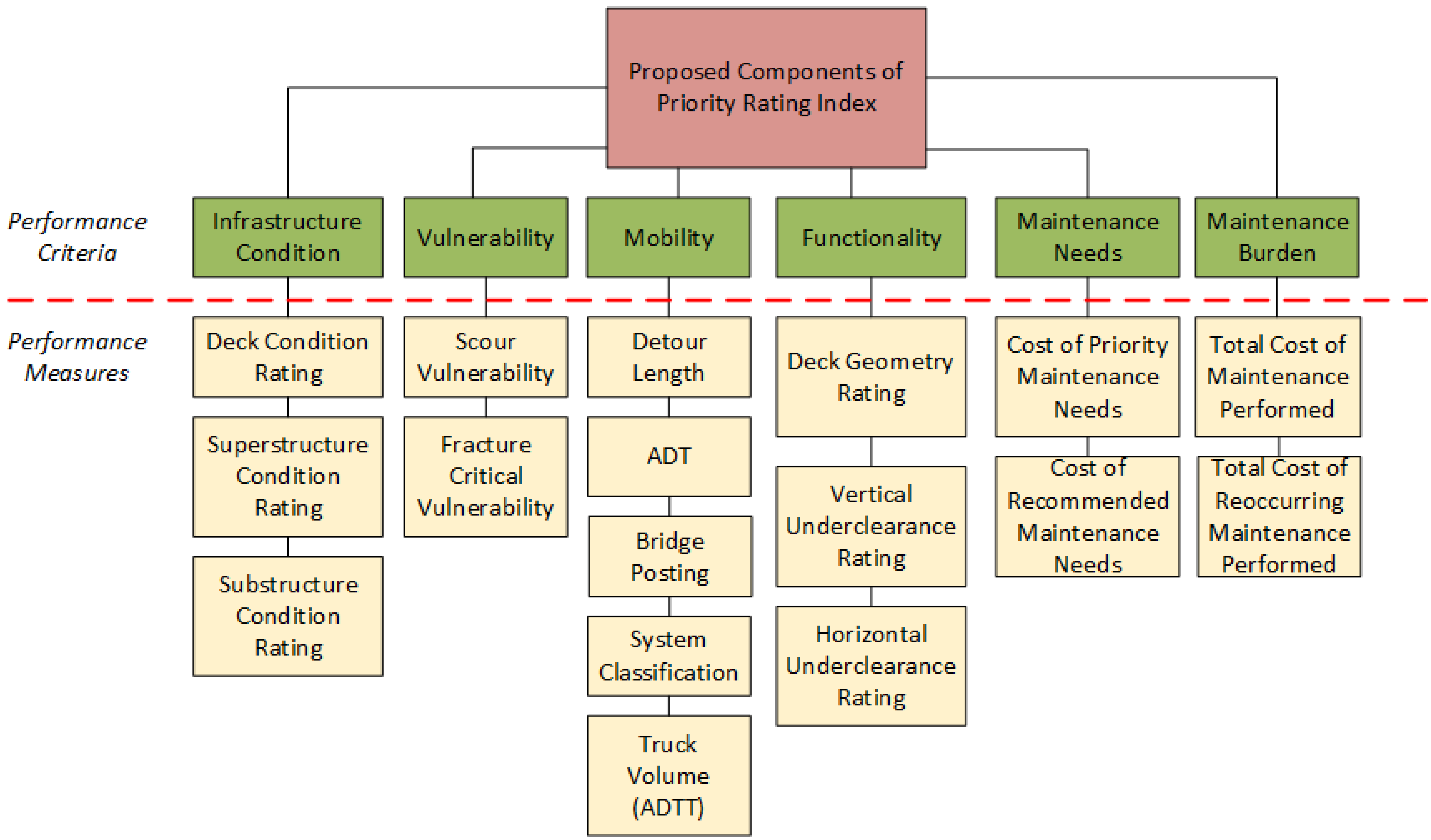


Figure 3.1: PRI Proposed Performance Criteria and Measures

TABLE 3.1: Infrastructure Condition Ratings and Descriptions, adapted from Weseman, 1995

Code	Description	Comments
N	Not Applicable	None
9	Excellent	None
8	Very Good	No problems noted.
7	Good	Some minor problems.
6	Satisfactory	Structural elements show some minor deterioration.
5	Fair	All primary structural elements are sound but may have minor section loss, cracking, spalling or scour.
4	Poor	Advanced section loss, deterioration, spalling or scour.
3	Serious	Loss of section, deterioration, spalling or scour have seriously affected primary structural components. Local failures are possible. Fatigue cracks in steel or shear cracks in concrete may be present.
2	Critical	Advanced deterioration of primary structural elements. Fatigue cracks in steel or shear cracks in concrete may be present or scour may have removed substructure support. Unless closely monitoredit may be necessary to close the bridge until corrective action is taken.
1	Imminent Failure	Major deterioration or section loss present in critical structural components or obvious vertical or horizontal movement affecting structure stability. Bridge is closed to traffic but corrective action may put back in light service.
0	Failed	Out of service - beyond corrective action.

The scour vulnerability rating scale is prescribed by the NBI recording and coding guide [Weseman, 1995] and ranges from 0 to 9, where a higher rating represents a lower structural risk due to scour (Table 3.2). Additionally, a bridge can be assigned a non-numerical rating to either indicate that the bridge is not over water (N) or over an unknown foundation or tidal water and has not been not evaluated for scour (U or T). In order to use scour ratings for regression and prediction, the non-numerical ratings were converted into numerical values and values were adjusted to better reflect severity of scour potential. In this reassignment of ratings, bridges not evaluated for scour were assigned to carry a higher potential vulnerability to scour than bridges that have been determined to be stable for scour conditions and bridges not subject to scour. The adjusted scour critical codes are summarized in Table 3.3.

The Mobility criterion is intended to reflect the usage of a bridge as well as the potential impact of closure on the community. Mobility is measured by: Detour Length, Average Daily Traffic (ADT), Volume of Average Daily Truck Traffic (ADTT), System Classification, and Bridge Posting. Detour Length reflects the additional distance a vehicle must travel in the event that the bridge is closed for repair or replacement. The ADT is the volume of all vehicles that are estimated to cross a bridge on a typical day. ADTT is a percentage of ADT that is estimated to be freight vehicles. System Classification indicates the type of route a bridge is on, which is either a secondary, primary, or interstate route. Lastly, the Bridge Posting rating indicates the vehicle to maximum bridge weight capacity relative to the maximum posting given by the state. This rating is based on a scale of 0-5 determined by the ratio of the bridge posting to the maximum state legal posting and is reproduced from the NBI Recording and

Table 3.2: NBI Scour Critical Ratings and Descriptions, adapted from Weseman 1995

Code	Description
N	Bridge not over waterway.
U	Bridge with "unknown" foundation that has not been evaluated for scour. Since risk cannot be determined, flag for monitoring during flood events and, if appropriate, closure.
T	Bridge over "tidal" waters that has not been evaluated for scour, but considered low risk. Bridge will be monitored with regular inspection cycle and with appropriate underwater inspections.
9	Bridge foundations (including piles) on dry land well above flood water elevations.
8	Bridge foundations determined to be stable for assessed or calculated scour conditions; calculated scour is above top of footing.
7	Countermeasures have been installed to correct a previously existing problem with scour. Bridge is no longer scour critical.
6	Scour calculation/evaluation has not been made.
5	Bridge foundations determined to be stable for calculated scour conditions; scour within limits of footing or piles.
4	Bridge foundation determined to be stable for calculated scour conditions; field review indicates action is required to protect exposed foundations from effects of additional erosion and corrosion.
3	Bridge is scour critical; bridge foundations determined to be unstable for calculated scour conditions
2	Bridge is scour critical; field review indicates that extensive scour has occurred at bridge foundations. Immediate action is required to provide scour countermeasures.
1	Bridge is scour critical; field review indicates that failure of piers/abutments is imminent. Bridge is closed to traffic.
0	Bridge is scour critical. Bridge has failed and is closed to traffic.

Table 3.3: Proposed Scour Codes and Descriptions

New Code (NBI code)	Description
7 (N)	Bridge not over waterway.
6 (9,8,5)	Bridge foundations stable for calculated scour conditions.
5 (U, T, 6)	Bridge not evaluated for scour.
4 (4)	Same as NBI description
3 (3)	”
2 (2)	”
1 (1)	”
0 (0)	”

Coding Guide in Table 3.4.

Table 3.4: Definitions for Bridge Posting Scores

Score	Relationship of Operating Rating to Maximum Legal Load
5	Equal to or above legal loads
4	0.1 - 9.9% below
3	10.0 - 19.9% below
2	20.0 - 29.9% below
1	30 - 39.9% below
0	> 39.9% below

The Functionality criterion is an indicator of the adequacy of the geometric design for vehicles and is assessed with: deck geometry rating, vertical underclearance rating, and horizontal clearance rating. Inadequate geometric design can lead to traffic safety and mobility issues causing a bridge to be functionally obsolete for modern use and design standards.

3.3 Newly Introduced Performance Measures

To address specific shortcomings of the PRI discussed in the prior chapter, Maintenance Needs and Maintenance Burden criteria have been introdued with performance measures derived from maintenance records that are maintained by the NCDOT.

These records, in the Maintenance Management System (MMS), use a list of over 200 standard maintenance actions, each of which includes a description of the action, the affected bridge element, unit cost, and number of units of a maintenance action. The maintenance record data set that is used for this study includes records from the past ten years to quantify the extent of Maintenance Burden historically associated with each structure. The most current set of bridge inspection records, which includes proposed future maintenance actions identified by bridge inspectors, is used to quantify the extent of Maintenance Needs for each structure. An advantage of utilizing maintenance record data is that element bridge health can be determined since maintenance actions include which elements need attention and the severity of deterioration of the bridge elements can be determined since the unit cost and number of units (elements) are included with the records. An assumption used for each of the criteria is that the higher the cost is for a given maintenance action, the more critical that the maintenance action is for the prioritization of bridge replacements. The remainder of this section will discuss the details of the individual criteria and the associated performance measures.

The main difference between Maintenance Burden and Maintenance Needs is the timing of when an action is applied to the bridge. Maintenance Burden is a measure of the amount of resources that have already been invested recently in the structure to maintain an operational state and is therefore computed using records of the maintenance actions that have been applied during the past ten years relative to the most current bridge inspection report. Maintenance Needs is meant to assess the current structural condition of a bridge in terms of identified cost of maintenance actions that would need to be done in the future to restore bridge element to adequate condition states. As an illustration, Figure 3.2 shows the relationship between Maintenance Needs and Maintenance Burden data.

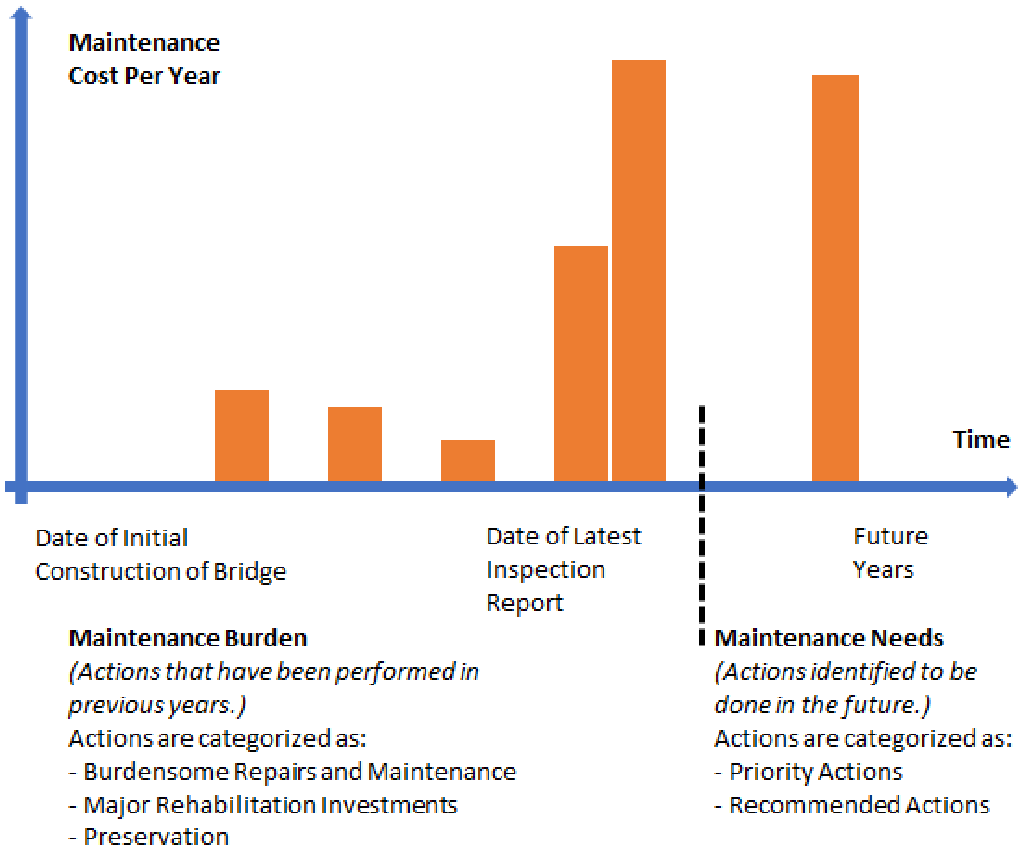


FIGURE 3.2: Relationship Between Maintenance Burden and Maintenance Needs In Terms of Yearly Maintenance Expenditures

The Maintenance Needs performace measures rely on element-level maintenance actions identified by inspectors in order to keep the conditions of the bridge at acceptable levels in the future and are measured by the cost of priority maintenance needs and cost of recommended maintenance needs. These performance measures are developed using the Inspector Recommended Maintenance Needs database (IRMN), which is based on element-level condition ratings. Each of the maintenance actions proposed for a bridge are classified at one of three priority levels: critical, priority, and recommended. Critical maintenance needs are required to be addressed immediately, so they are not reflected in the IRMN and are not appropriate for prioritization of bridge replacement projects. Priority needs are the next level of urgent need for a maintenance action, while recommended needs represents actions with even less urgency but are still required in order to restore or maintain bridge element condition ratings at acceptable targets. Through use of the aggregated quantities of individual maintenance needs and their estimated unit costs, the total cost of priority and recommended needs for each structure can be computed to reflect the total scale and urgency of all current maintenance needs.

Maintenance Burden performance measures assess the costs of completed maintenance actions on a bridge to determine if previous expenditures indicate that the replacement would alleviate burdensome and potentially costly maintenance actions that have been ineffective or only partially effective in prolonging the service life of the structure. The Maintenance Burden criterion consists of two measures: total cost of maintenance actions performed on the bridge and total cost of reoccurring maintenance actions performed on the bridge. Reoccurring maintenance actions are instances when the same type of maintenance was performed at separate times over the ten year period, which may better reflect the burden presented to divisions than single instance maintenance actions. These two measures are further divided by action classifications: burdensome repairs and maintenance, major rehabilitation investments, and preservation treatments. Each of the maintenance actions were classified into these categories in order to distinguish actions that could increase the likelihood of bridge replacement (burdensome repairs and maintenance), and actions that are done that could lead to prolonged use of a bridge (major rehabilitation investments and preservation actions). The assumption made is that bridges that have received major rehabilitation or preservation treatments are less likely to be suitable candidates for replacement since the state has likely not yet received the service life benefits of these investments. In contrast, bridges that have required significant maintenance for repairs that do not significantly improve the overall condition of the bridge are more suitable for replacement, particularly if the maintenance has been performed year after year. In addition to these classifications, there are some maintenance actions that have no effect on the decision to replace a bridge, such as removal of graffiti or beaver control, and such actions were identified in a separate category to be removed from all subsequent analysis. The action classifications were determined by individually classifying actions using engineering judgment. Maintenance Burden classifications were created by the research team using a survey and group discussion. The maintenance action classification survey queried the opinions of the research team members about the perceived classification of each action. There were 244 unique actions in the Maintenance Burden Database. The survey results from each of the researchers were compared and actions that that were classified as two or three different types were collected and discussed individually in a group meeting. During this meeting, each researcher discussed why he or she chose a certain classification. At the end of the discussion, the group decided what the action classifications were together. Examples of maintenance actions for each classification are shown in Table 3.5. These classifications were then reviewed by NCDOT personnel in an interim project meeting and deemed to be appropriate for the purposes of this research. Only maintenance actions that were performed within the last ten years were included in the analysis.

Table 3.5: Maintenance Action Classification Definitions and Examples

Classification	Maintenance Action Examples
Burdensome Repairs and Maintenance (Actions that increase the likelihood of bridge replacement.)	Pothole Patching; Maintenance of Cracks and Joints in Pavement Repair Concrete Wings and Walls
Major Rehabilitation Investments (Actions that significantly prolong the service life of a bridge)	Replacement of Bridge Expansion Joints Replacement of Timber Bridge Flooring Replacement of Steel Columns and Piles Replacement of Superstructure
Preservation (Actions that prolong the service life of a bridge)	Cleaning and Painting of Structural Steel Deck Washing Maintain Drainage System Hot Mix Asphalt Overlay

As previously mentioned, reoccurring maintenance actions are expected to be more strongly correlated with preference to replace bridges due to the burden and costs associated with routinely performing the same actions on the same structure to keep it operational. Therefore, actions that have been reoccurring over the years for a given bridge were analyzed using a PivotTable in Excel. This was done by organizing the rows of the Maintenance Burden database first by structure ID and then by the total cost for each of the maintenance burden actions by year. If the action repeats for a subsequent year, then the action is considered to be a reoccurring action. Additionally, the maintenance classifications are applied to each of the reoccurring actions, and the output of reoccurring burdensome repairs and maintenance, insignificant actions, major rehabilitation, and preservation are computed and are indexed back to the main bridge database using the unique structure ID.

One aspect of this research effort is to determine the most statistically significant form that the Maintenance Needs and Maintenance Burden performance measures should be quantified. Specifically, either the total costs could be used to reflect the total scale of the need and emphasize larger or more costly bridges in the prioritization, or the costs could be normalized by the structure size. For this second approach, the total Maintenance Needs and Maintenance Burden costs were divided by the estimated replacement cost of the structure, which is a value already estimated by NCDOT using the deck area and the route type. In this way, the Maintenance Needs and Maintenance Burden performance measure scores are proportional to a fraction of the replacement cost.

3.4 Performance Measure Data Sources

The raw data used to develop a centralized database for statistical analysis are: the Network Master, the 2016 NBI, Inspector Recommended Maintenance Needs database (IRMN), and the Maintenance Management System (MMS) history database. The Network Master is a web-accessible database in the AgileAssets Asset Management System used by NCDOT that contains design, functional, geographic, and inventory data as well as other bridge information collected by the state. The Network Master provides a snapshot of the entire state inventory of structures at the current instant in time and, unlike historical databases, is routinely updated with new inspection data. The Network Master contains most of the NBI items necessary to compute the proposed set of performance measures, with the exception of a few that needed to be sourced directly from the corresponding annual NBI file submitted to the Federal Highway Administration. Specifically, ADTT, fracture criticality, and bridge posting data was extracted for each bridge from this source. The IRMN contains information about maintenance actions recommended by bridge inspectors in order to preserve or improve the condition of a bridge component or overall condition of the bridge. This data source contains a list of "treatments" for each bridge with an associated priority classification, quantity, and unit cost for each treatment. There are two levels of priority, which are "priority" and "recommended". These classifications were assigned by the inspectors based on element-condition ratings originally recorded in NCDOT Wearable Inspection and Grading Information Network System (WIGINS) used on site bridge inspectors. The MMS history database contains records of maintenance actions performed for each structure with the associated description, start and ending dates, amount, and cost. The data sources and their use for calculating the individual associated performance measures are summarized in Table 3.6. For the purposes of this study, all of the databases were sourced concurrently in July of 2016.

3.5 VBA Design

A program was created in order to automate the process of assembling the separate databases, filtering the data, computing the performance measures, and, ultimately creating value functions for each performance measure. Additionally, the program was developed to index the value function data to individual bridge structures so that

Table 3.6: Data Sources and Associated Performance Measures

Source	Performance Measure	NBI Item #
Network Master	Deck Condition	58
	Superstructure Condition	59
	Substructure Condition	60
	Scour Critical Bridge	113
	Detour Length	19
	ADT	29
	Bridge System	NA
	Deck Geometry Appraisal	68
	Underclearance Appraisal	69
NBI	ADTT	109
	Fracture Critical (Critical Feature Inspection)	92
	Bridge Posting	70
IRMN Database	Priority Maintenance Total	NA
	Recommended Maintenance Total	NA
MMS History	Major Rehab Total	NA
	Preservation Total	NA
	Burdensome Repairs and Maintenance Total	NA
	Reoccurring Major Rehab	NA
	Reoccurring Preservation	NA
	Reoccurring Burdensome Repairs and Maintenance	NA

prioritization scores can be computed for every bridge. The program was created using Microsoft Excel VBA macro routine and the source code is provided in Appendix A. This section discusses the different components of the VBA script in more detail. The overall VBA process employed by the VBA program is shown in Figure 3.3.

The VBA Import Data script allows the data sources to be used for value function creation to be specified by allowing the user to import selected files. In this process,

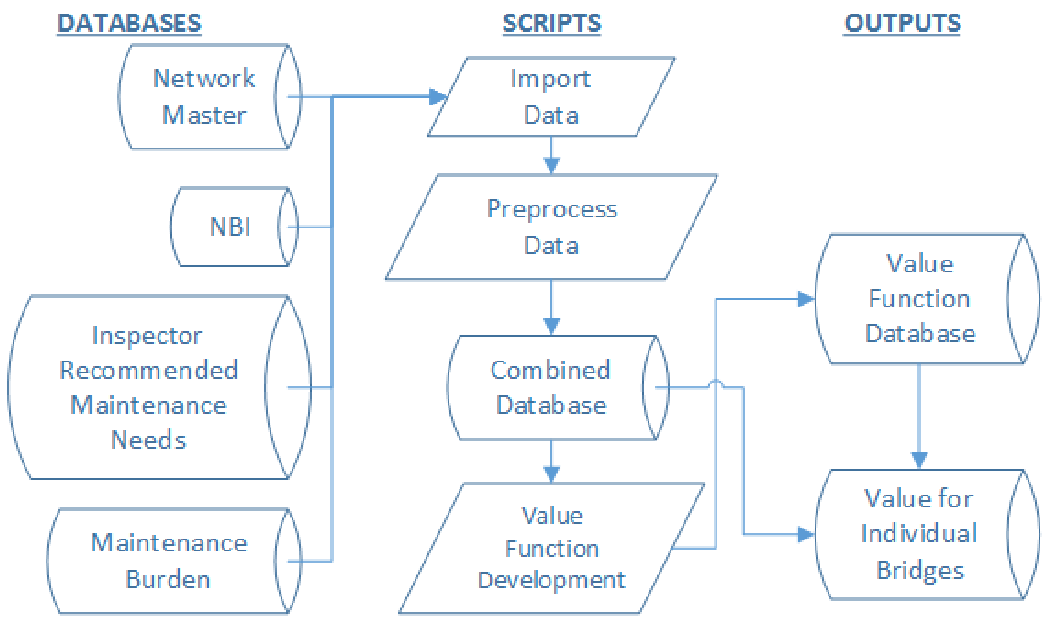


FIGURE 3.3: VBA Program Details

the database files are all imported as individual Excel worksheets. This script is mainly an automation of the copy and paste operations used to append data in of a worksheet from one workbook to another. However, this script imports all databases into one workbook based on the structure ID as a unique field allowing for association of data to specific structures.

The Preprocessing Data script performs filtering, transformations, and other operations on the source data necessary for it to be used for the creation of value functions. The specific actions taken for each data source will be discussed. On the Network Master data, filtering was applied to extract only the data related only to bridges, as this database includes other transportation structures such as road signs and traffic lights. Preprocessing of the NBI database involved conversions of text strings to numbers and migration of the processed and filtered data to the common database. The Maintenance Needs database was transformed by separately summarizing the total priority action costs and total recommended action cost for each bridge. This was accomplished using a Pivot Table, which is a data summarization tool within the Microsoft Excel application. For the MMS history database used for the Maintenance Burden performance measures, the bridge identification numbers were converted to structure IDs, as the BMS and MMS systems use different identifiers for each structure. Then the date of each maintenance action was determined using the end dates of the actions, or if not available, the start date. Maintenance actions that were performed earlier than ten years before the most current bridge inspection were assumed to no longer be significant toward the predicting the prioritization of bridge replacement projects and, therefore, all instances of these actions were removed from the database. Records with incorrect structure IDs, completion years, and negative costs were treated as anomalies and were also were removed from the database by the macro. Automated analysis steps were programmed into the macro to summarize the historical maintenance action costs for each bridge by the different types of maintenance categories as previously defined. This calculation of burdensome repair costs, major rehabilitation costs, and preservation action costs for single instance and reoccurring actions was also completed with the use of PivotTables.

The preprocessing data script results in a common database using the indexing feature, VLOOKUP, in Microsoft Excel to associate data from each preprocessed source using the structure ID. The data from the common database was then operated on by the Value Function Development script. This script performs the operations necessary to develop a data-driven value function specific to the snapshot of the bridge inventory created by the sourced database files. The specifics of these calculations are discussed in detail in the following section. Lastly, the VBA was developed to leverage the created value functions to convert all of the raw performance measures for each structure over to values associated with either linear or ECDF derived value functions. This conversion was performed using the VLOOKUP command to find the value from each value function associated with the set of specific raw performance measures for each bridge. The structure-specific value scores for each of the performance measures were used for the statistical analysis and prediction model development.

3.6 Value Function Development

The preferences structure of NCDOT bridge engineers for each performance measure were modeled using value functions. There are different modeling methods as well as perspectives of observing the measures that could potentially be used, therefore all logical models were developed in order to be tested for statistical significance when estimating the probability of bridge replacement through regression later in the study. This section will introduce the background of the value function development used for this study.

Each of the performance measure value functions were alternatively modeled as a linear function as well as an empirical cumulative distribution function (ECDF) to determine which function structure was more significant for estimating the probability of replacement candidacy. A value function can be modeled as a linear function if the value trade-off ratio, which is represented by the slope of the function, is the same for all performance measure values. Since the ECDF-derived value function is an empirically derived probability distribution, its use should also result in a more uniform distribution of scores than a linear scales, which may address issues with clustering of scores observed in the current PRI. As an example, the distribution of ADT ECDF-derived value function values and linear value function values are shown in Figure 3.4. ECDFs were used with VDOT index value functions, developed using the 10th percentiles of the distribution a trend line [Moruza et al., 2016]. ECDFs were also used in the NCDOT Prioritization 4.0, as shown in Figure 3.5, which is a reason why this format was pursued as a means to ensure consistency with an approach that NCDOT engineers are familiar with and incorporate in other prioritization systems. Both the linear and ECDF forms of the value functions were designed to range in scale from 0 to 100.

The linear value functions were developed with the equation

$$\text{V} = (\text{Individual Bridge Performance Measure Value} - A) \times \frac{100}{B - A} \tag{3.1}$$

where A and B is the minimum and maximum bridge performance measure values observed across the entire bridge population. This equation shifts the performance measure scale to an origin of zero and scales the magnitude of the performance measure scale to 100. The form of the value function shown in equation 3.1 assumes that the performance measures are proportional with the associated priority for replacement. While this may be the case for some performance measures, such as ADT and detour length, other measures, such as condition and appraisal ratings, are expected to be inversely proportional with the associated priority for replacement. For these

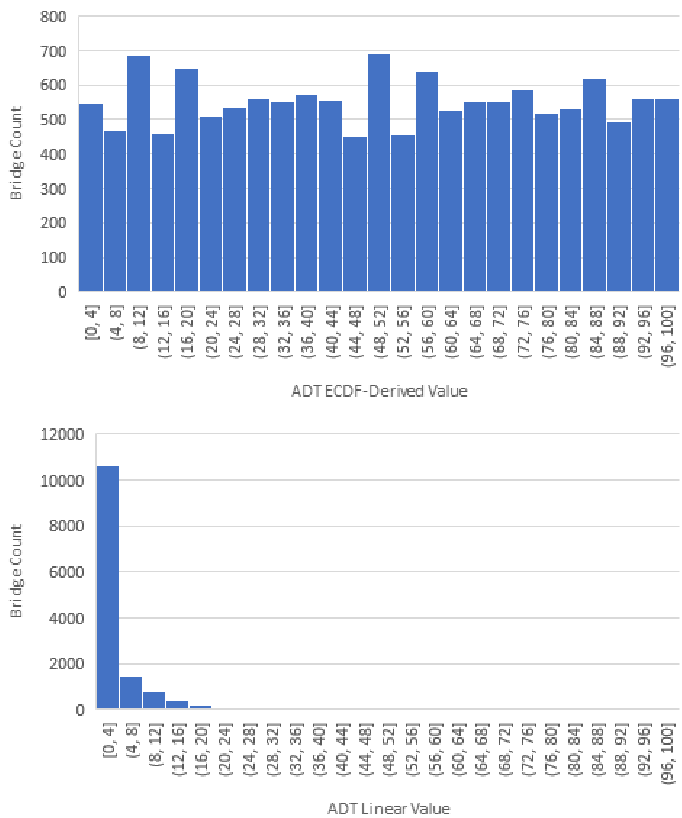


FIGURE 3.4: Distribution of ADT ECDF-Derived Value Function Values and Linear Value Function Values

measures, the linear forms of the value functions were developed as

$$\text{V} = (\text{Individual Bridge Performance Measure Value} - B) \times \frac{100}{A - B} \tag{3.2}$$

The ECDF-derived value functions were developed using the equation

$$\text{V} = \sum_{i=0}^{n} \frac{100 \times i}{n - 1} \tag{3.3}$$

where n is the full population of performance measure scores and i is the iteration in the series or if the performance measure ratings are expected to be inversely propor-

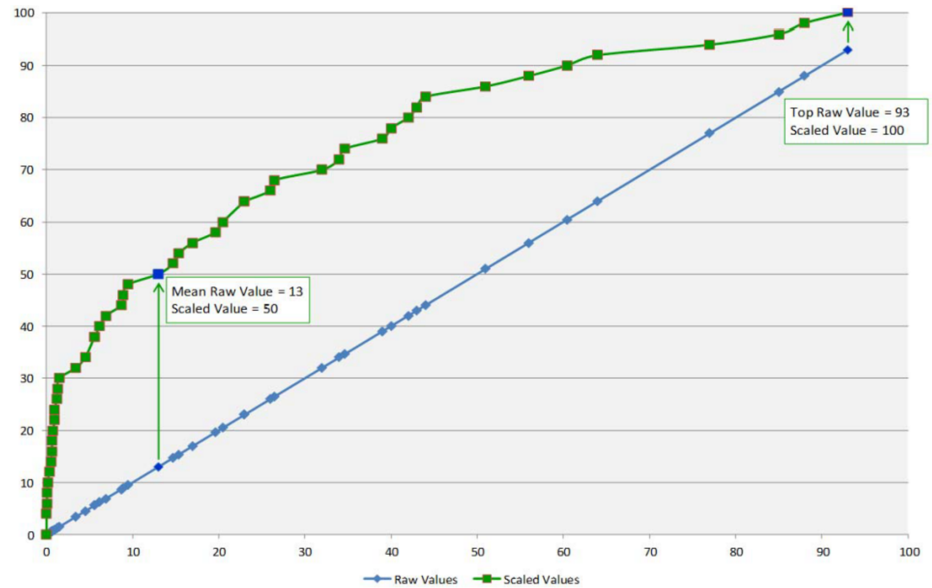


Figure 3.5: Example of Scaling of Values from NCDOT Prioritization 4.0

tional to the associated priority of replacement is

$$V = 100 - \sum_{i=0}^{n} \left(\frac{100i}{n-1}\right) \tag{3.4}$$

The ECDF values averaged by the performance measure ratings then were normalized to a 0 to 100 scale.

For an example of the development of an ECDF-derived value function, consider the bridge posting performance measure . First, the empirical cumulative distribution using Equation 3.3 was computed for bridge posting ratings 0 to 5, as shown in Figure 3.6(a). Second, the average value for each bridge posting rating is calculated and shown in Figure 3.6(b). Using the average value allows a single value to be associated to each posting rating. Note that the averaging of results in a value function does not utilize the full range of values, with a maximum of 92.28 and minimum of 35.57. To compensate for this effect, the final step is the normalization of the function by scaling values to a range of 0 to 100, as shown in Figure 3.6(c).

The assumed proportionality associated with each performance measure is summarized in Table 3.7. As an example of the development of the value functions from the individual performance measures, consider the substructure condition rating. The substructure condition rating provides a value from 0 to 9 as shown in Figure 3.7 (a) the corresponding linear scale value functions for this performance measure is shown, Figure 3.7 (b), which reflects the inverted proportionality to assume that lower ratings correlate with increased odds of selection for bridge replacement. Additionally, the linear value function scales the performance measure to a range of 0 to 100. The ECDF derived form of the value function for substructure condition rating is shown in Figure 3.7 (c) along with the underlying ECDF upon which the rolling average was computed. Note that the size of the steps on the ECDF reflect the relative number of bridges, so this value function indicates that a significant portion of the bridge population has a substructure rating of 5, 6, or 7. Bridges within the population that have a lower rating receive larger and nearly equal scores since there are relatively few of them and bridges with higher substructure condition scores receive very low scores for prioritization. Linear and ECDF derived value functions for the infrastructure condition performance measures are provided in Figure 3.8, for the vulnerability performance measures in Figure 3.9, for mobility performance measures in Figure 3.10, and for the functionality performance measures in Figure 3.11. The linear and ECDF derived value functions for the Maintenance Needs and Maintenance Burden performance measures computed with total cost are presented in Figures 3.12, 3.13, and 3.14. Likewise, the linear and ECDF derived value function for the Maintenance Needs and Maintenance Burden performance measures computed as a fraction of the estimate replacement cost are presented in Figures 3.15, 3.16, 3.17.

TABLE 3.7: Assumed Proportionality for Bridge Replacement Priority

Performance Measure	Replacement Probability Proportionality
ADT	Proportional
ADTT	Proportional
Bridge System	Proportional
Burdensome Repairs and Maintenance Total	Proportional
Detour Length	Proportional
Fracture Critical	Proportional
Priority Maintenance Total	Proportional
Recommended Maintenance Total	Proportional
Reocc. Burdensome Repairs and Maint.	Proportional
Truck Volume/Capacity	Proportional
Bridge Posting	Inversely Proportional
Deck Condition	Inversely Proportional
Deck Geometry Appraisal	Inversely Proportional
Major Rehab Total	Inversely Proportional
Preservation Total	Inversely Proportional
Reoccurring Major Rehab	Inversely Proportional
Reoccurring Preservation	Inversely Proportional
Scour Critical Bridge	Inversely Proportional
Substructure Condition	Inversely Proportional
Superstructure Condition	Inversely Proportional
Underclearance Appraisal	Inversely Proportional

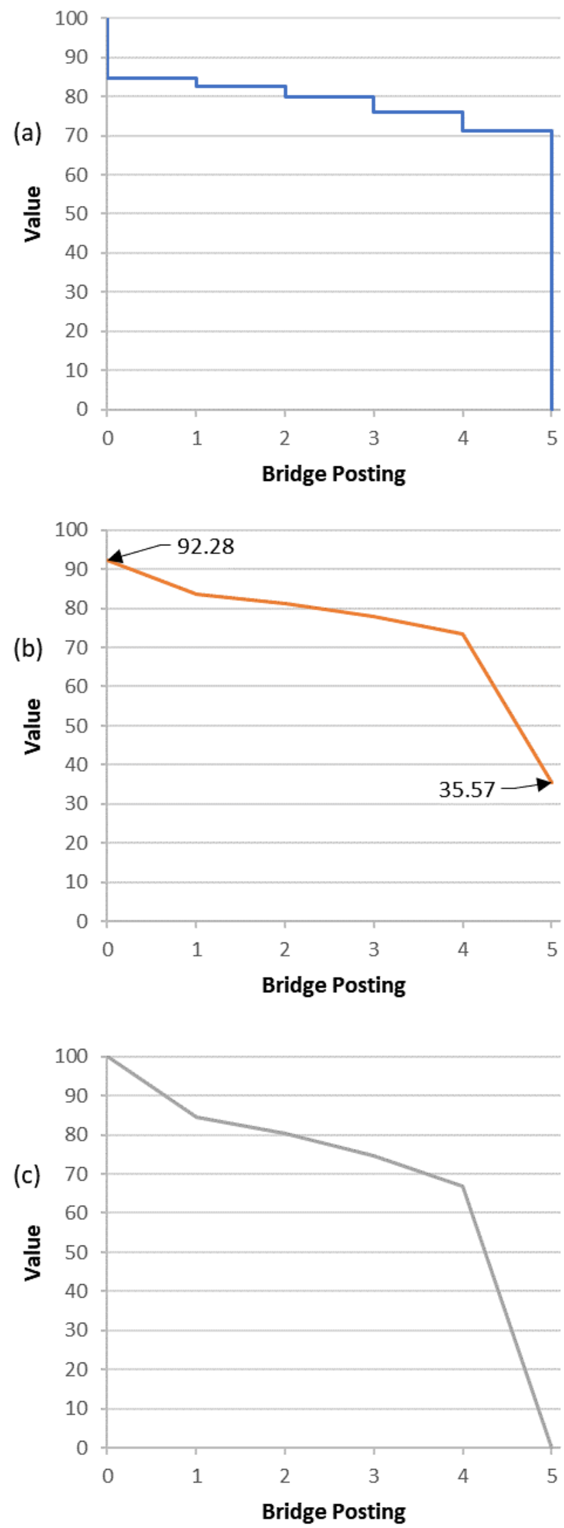


FIGURE 3.6: Graphical Illustration of the Development of Value Functions Based on the Empirical Cumulative Distribution: (a) the ECDF for bridge posting values, (b) the average value for each bridge posting value (c) the normalized ECDF-derived value function

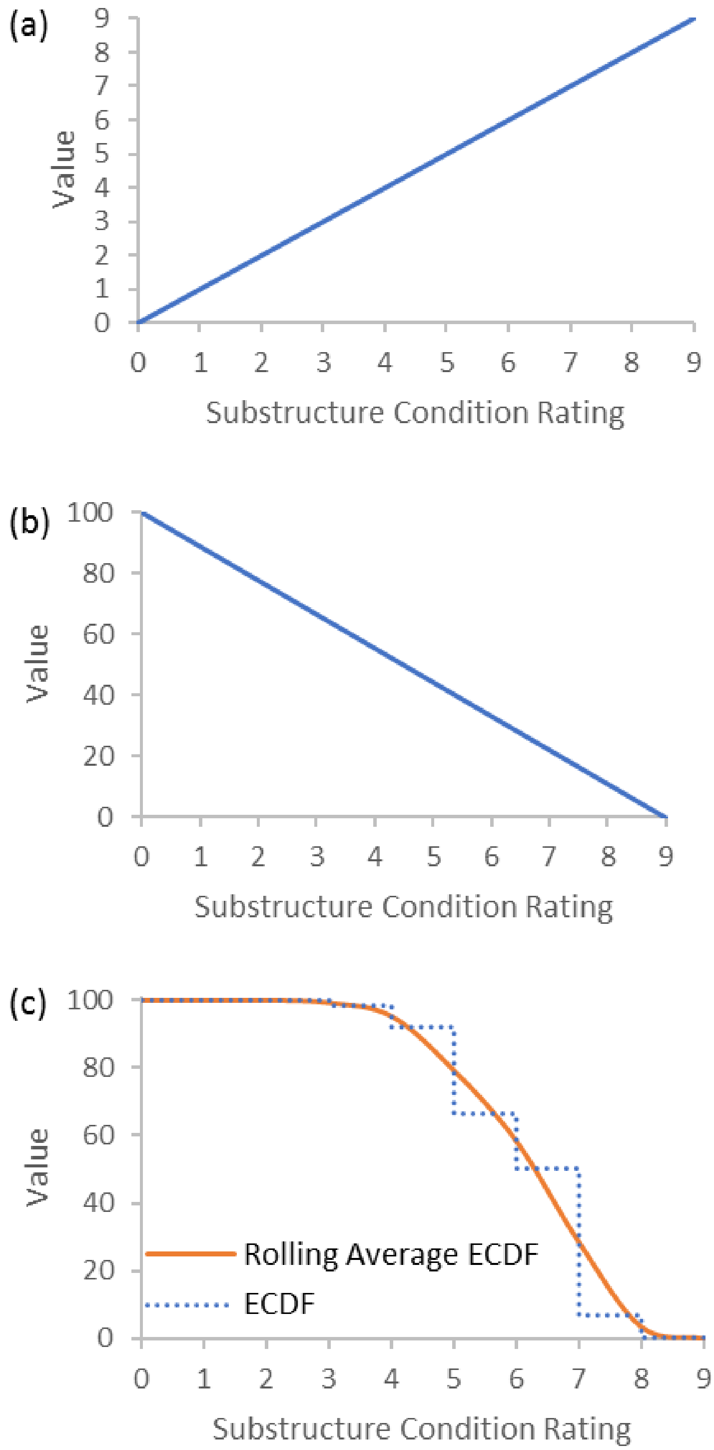


Figure 3.7: Substructure condition rating as (a) raw unscaled rating system, (b) a linear value function and (c) an ECDF-based value function

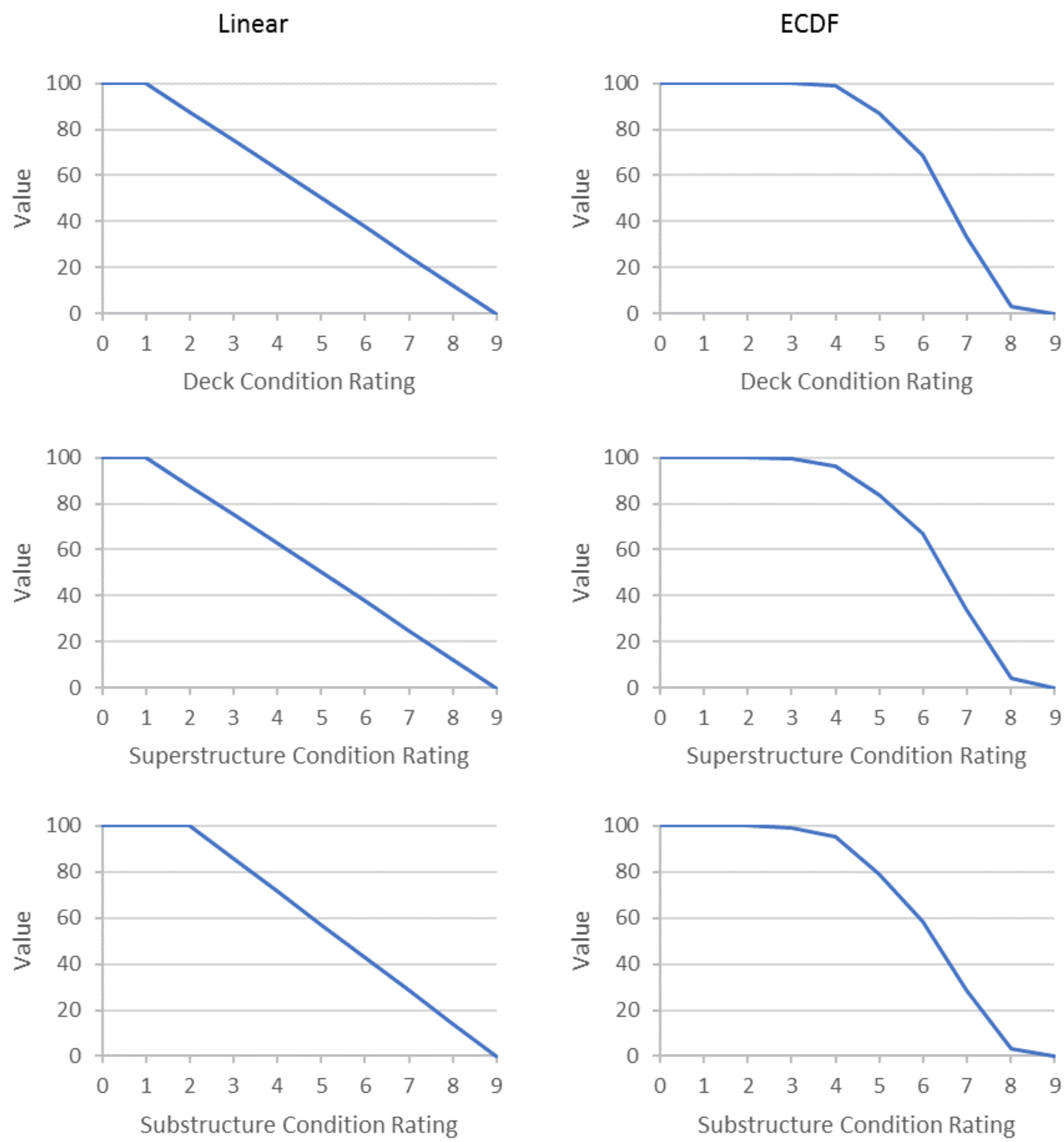


FIGURE 3.8: Linear and ECDF Value Functions for Infrastructure Condition

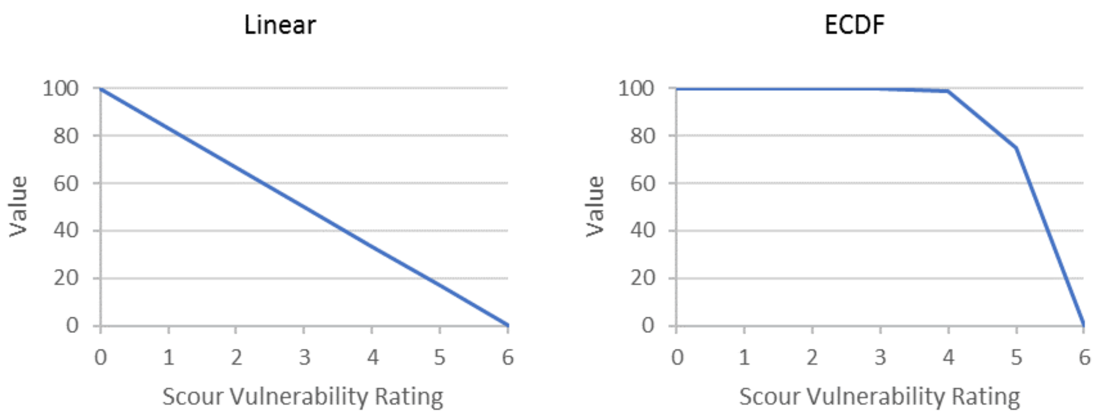


FIGURE 3.9: Linear and ECDF Value Functions for Vulnerability

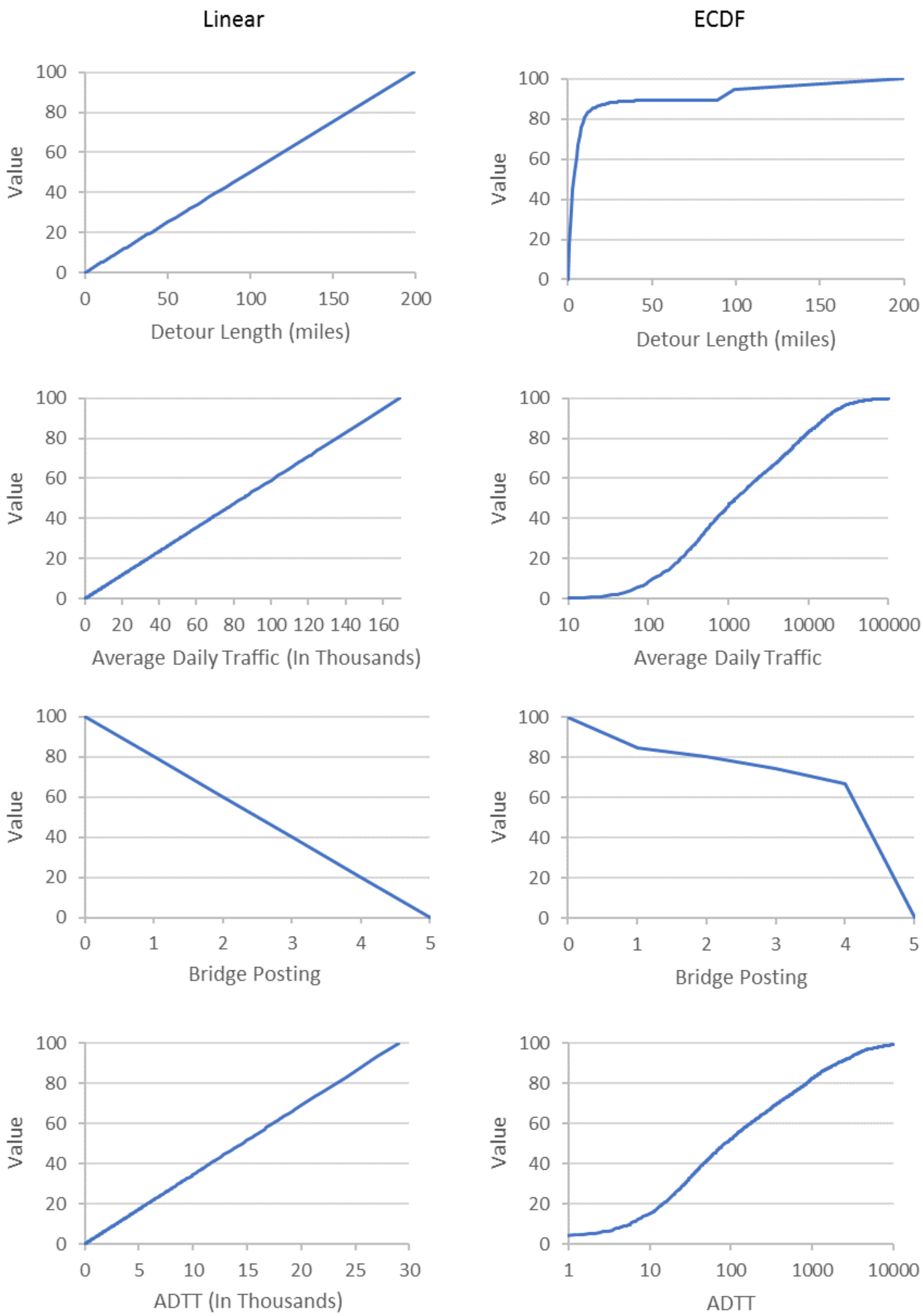


FIGURE 3.10: Linear and ECDF Value Functions for Mobility

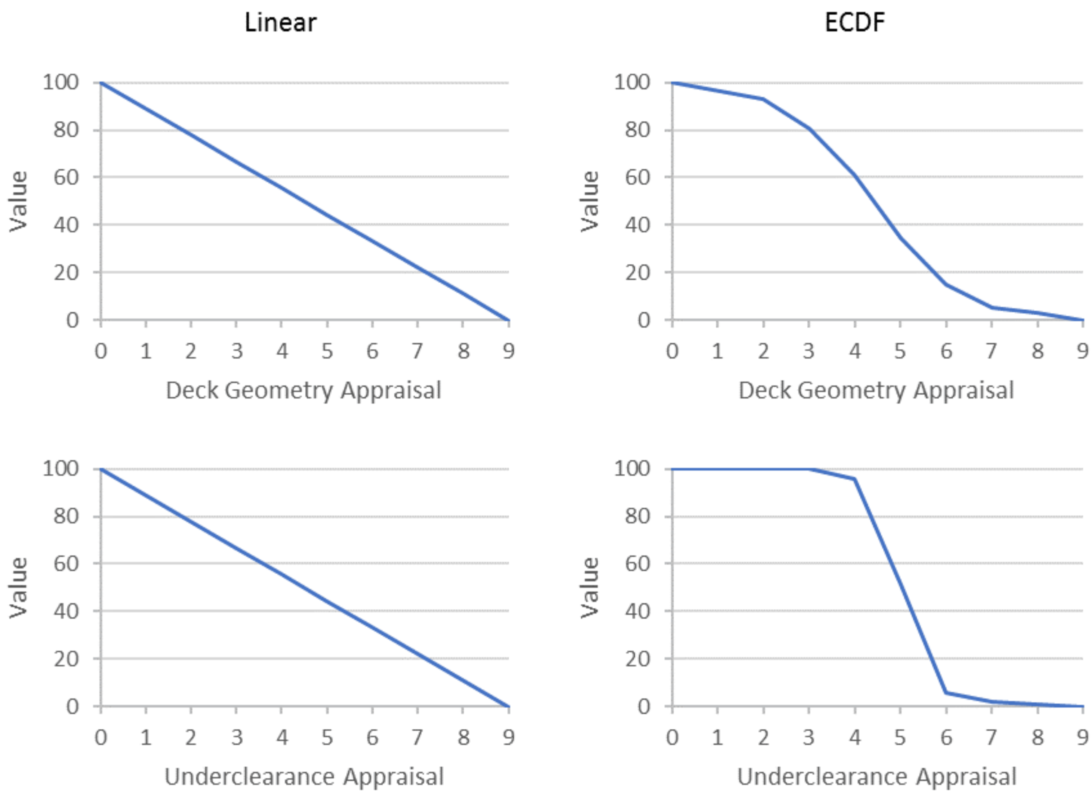


Figure 3.11: Linear and ECDF Value Functions for Functionality

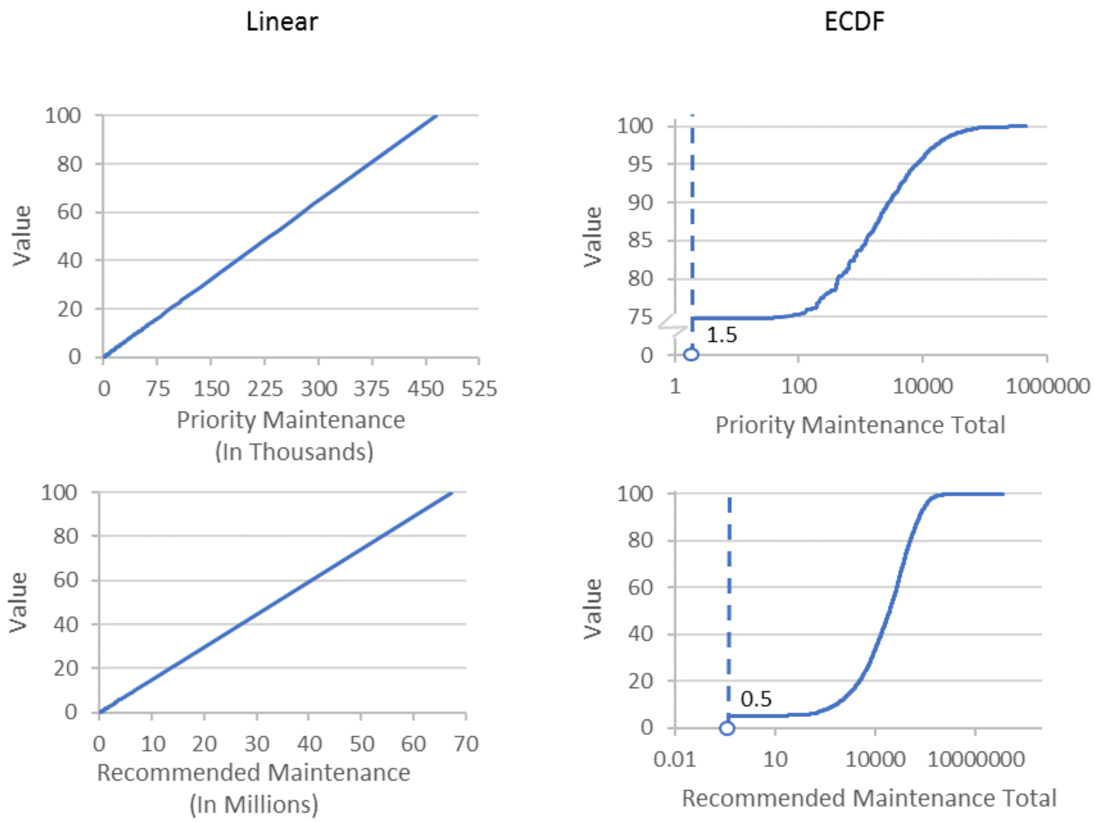


Figure 3.12: Linear and ECDF Value Functions for Maintenance Needs

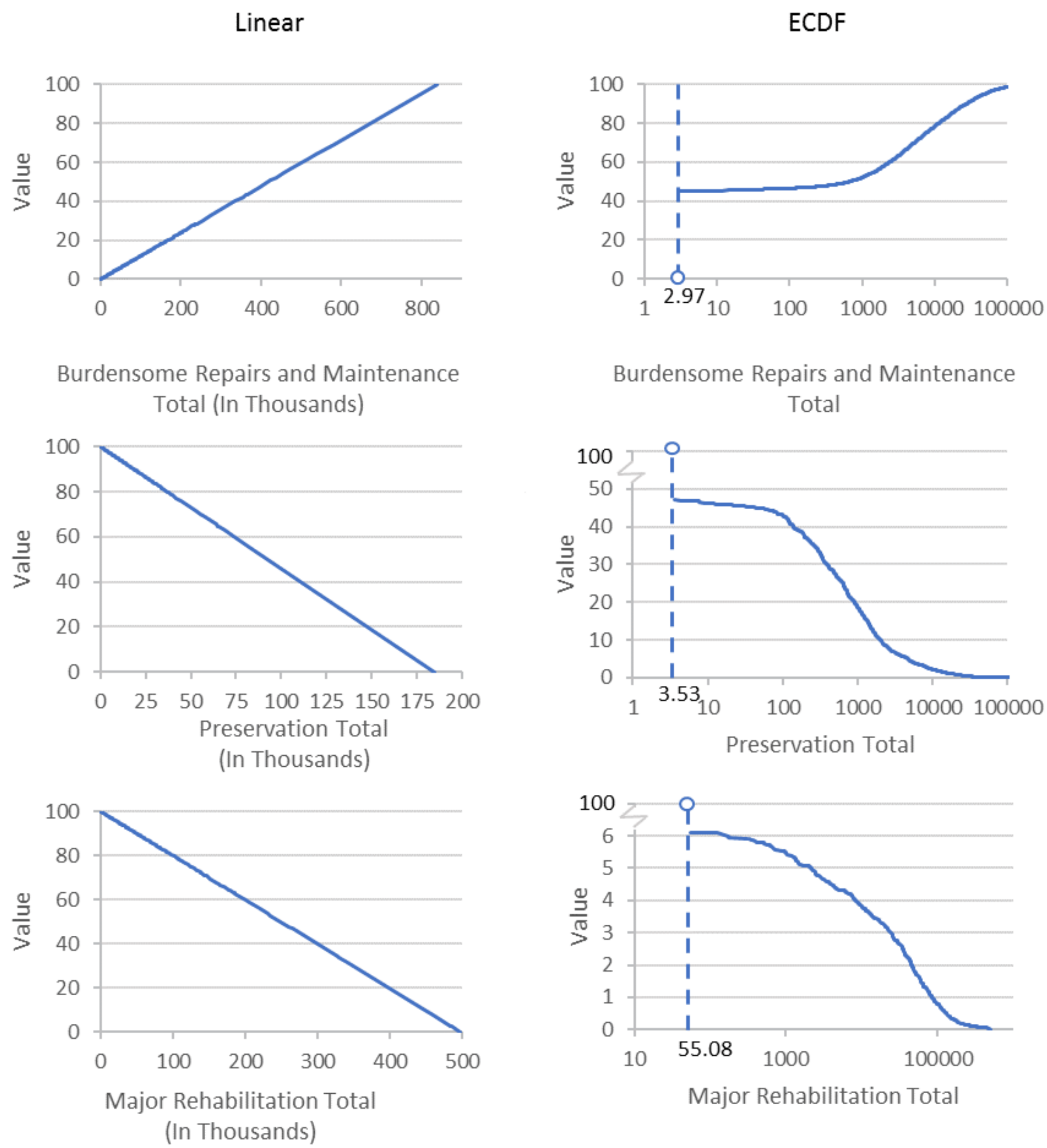


Figure 3.13: Linear and ECDF Value Functions for Maintenance Burden

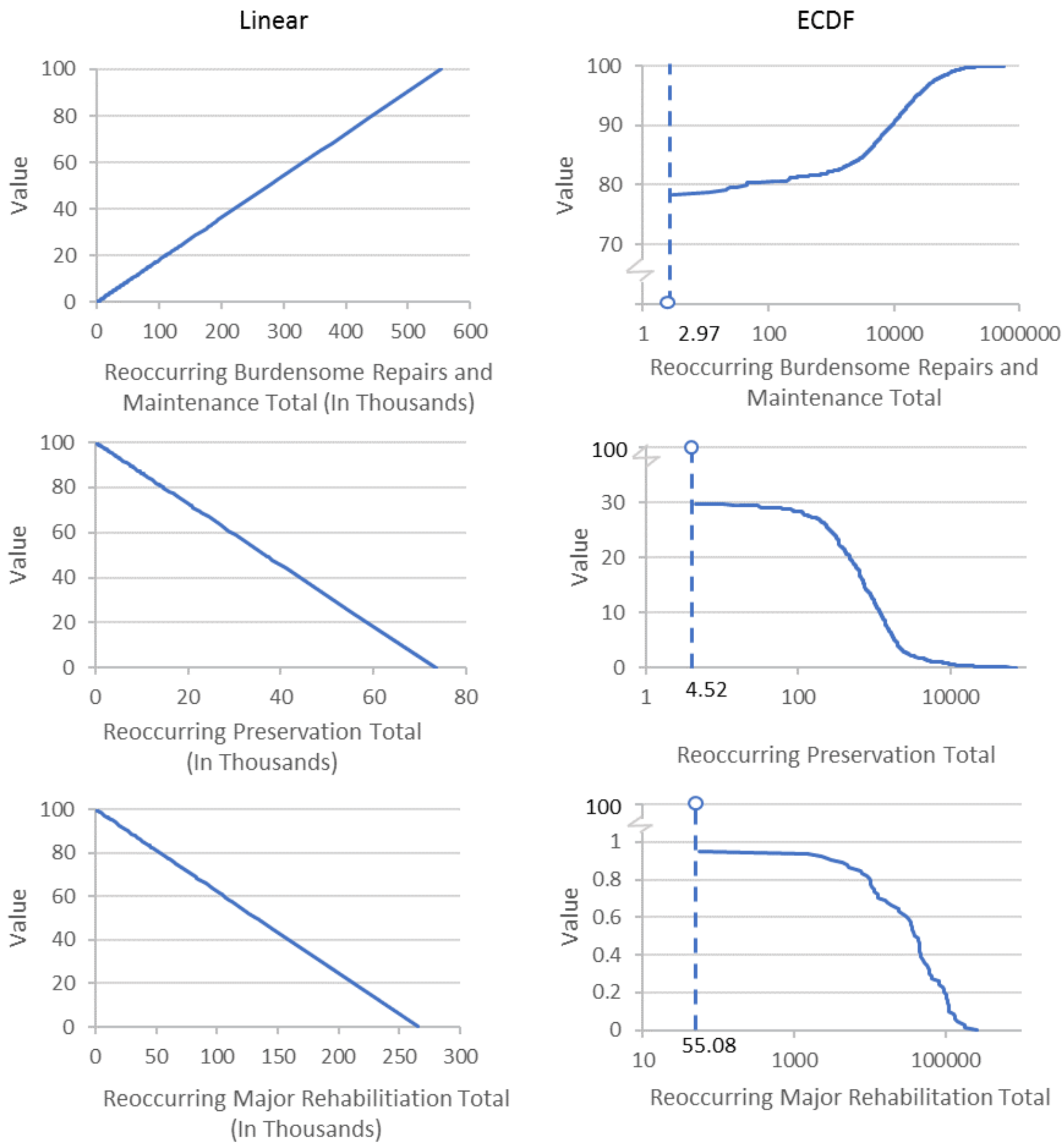


Figure 3.14: Linear and ECDF Value Functions for Reoccurring Maintenance Burden

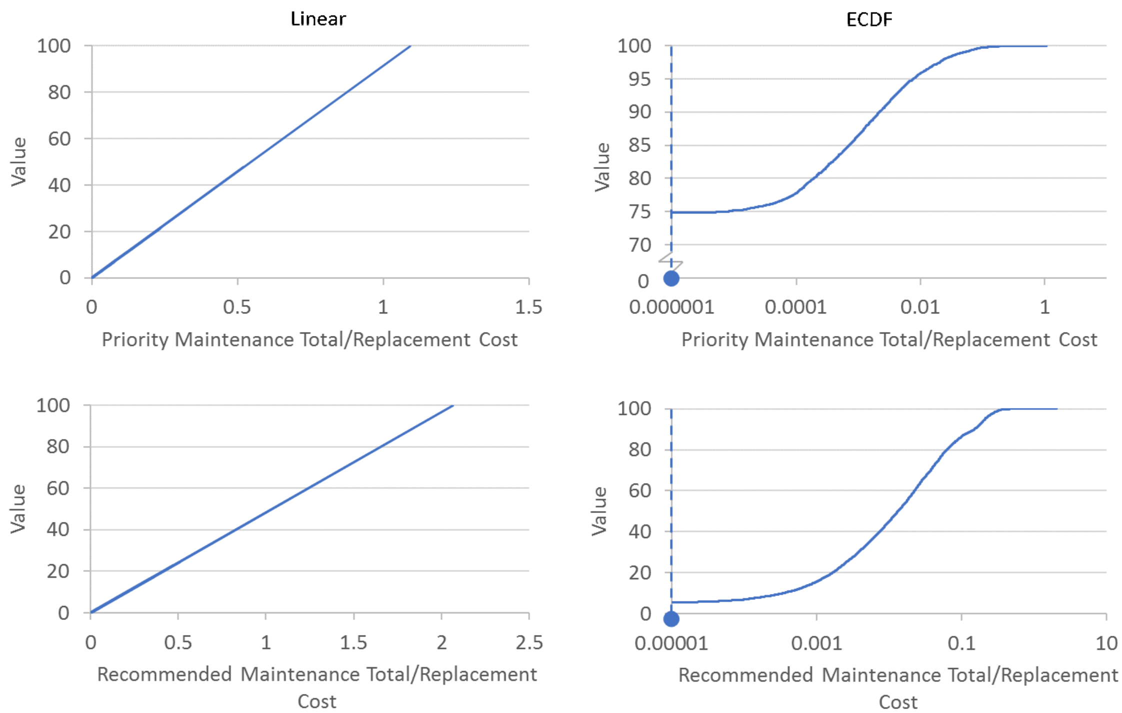


FIGURE 3.15: Linear and ECDF Value Functions for Maintenance Needs to Replacement Cost Ratio

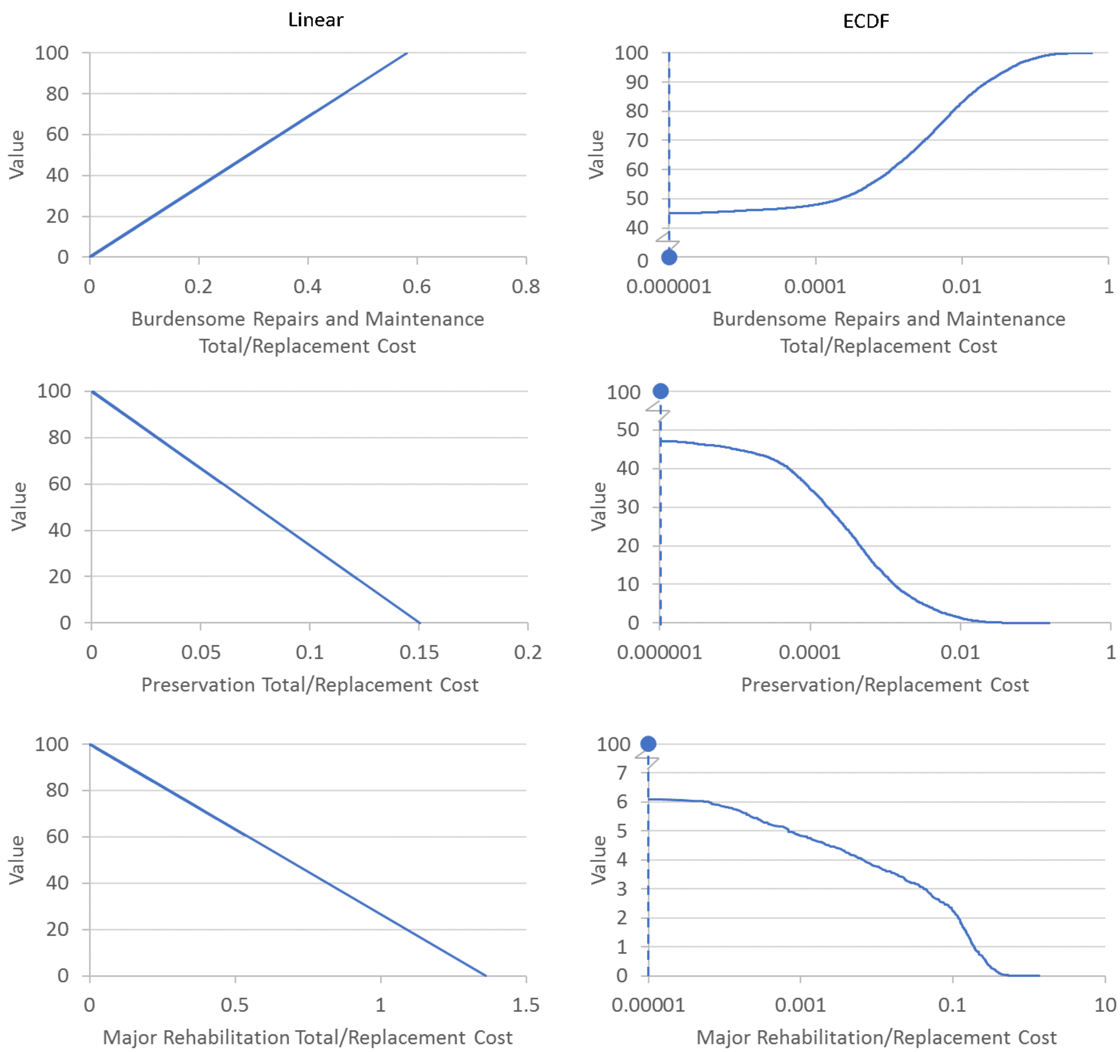


Figure 3.16: Linear and ECDF Value Functions for Maintenance Burden to Replacement Cost Ratio

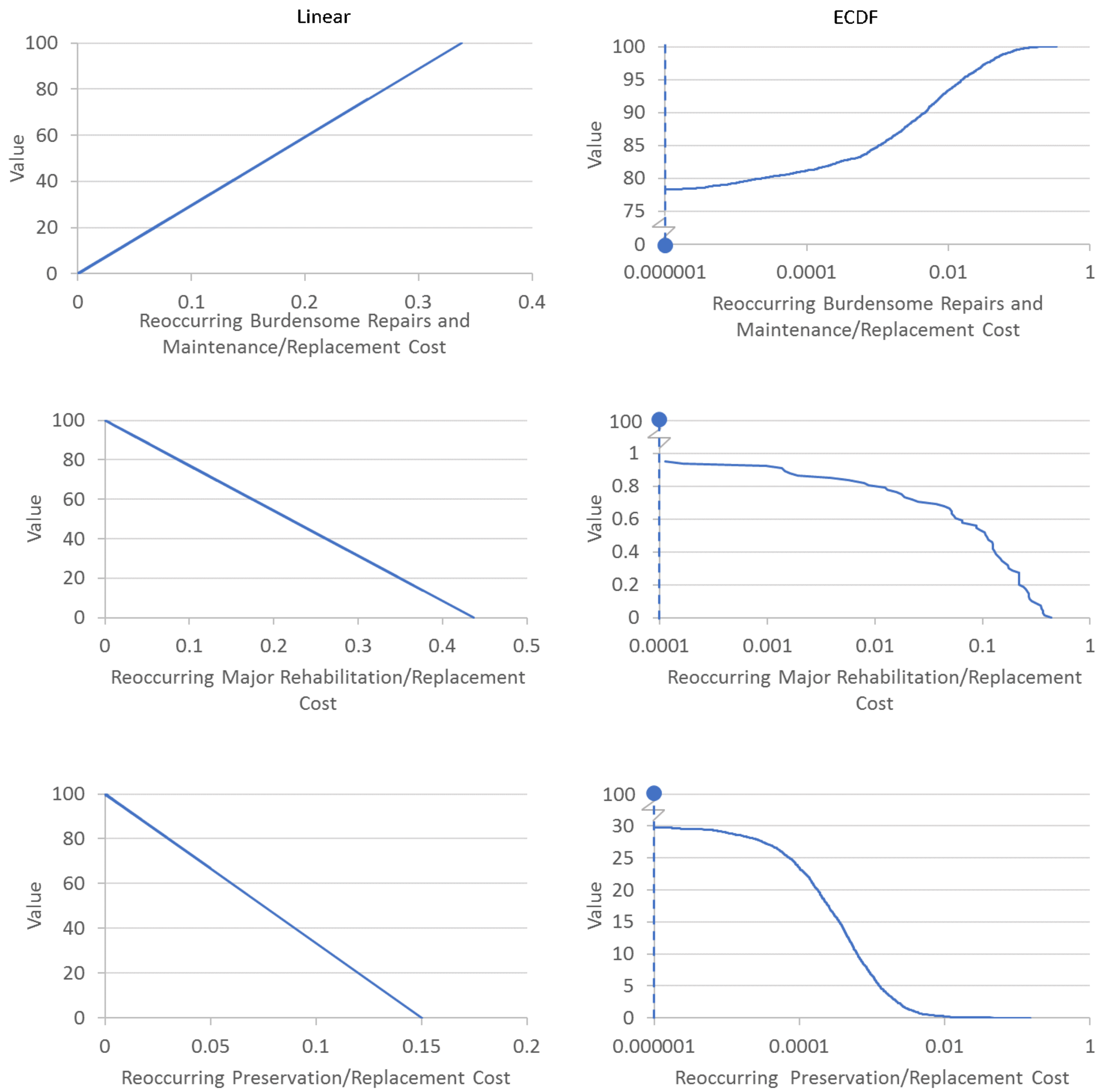


Figure 3.17: Linear and ECDF Value Functions for Reoccurring Maintenance Burden to Replacement Cost Ratio

CHAPTER 4: DEVELOPMENT OF STATISTICAL MODELS FOR PREDICTING BRIDGE REPLACEMENTS

This chapter includes the discussion of data-driven methodologies to develop statistical models for predicting the likelihood that a bridge will be selected for replacement. The development of the binary response variable indicating selection of a bridge for replacement by NCDOT engineers is discussed to compliment the development of the predictor variables in the prior chapter. Then, a matrix of different statistical models that will be tested with varying functional forms of value functions and statistical analysis methods will be introduced to examine the best approach for data-driven development of the prioritization index. To support the development of this matrix, background information on constrained linear least squares regression and binary logistic regression will be provided. Statistical models developed from the approaches outlined in this chapter will be analyzed and compared in the subsequent chapter.

4.1 Development of the List of Bridges Selected for Replacement

In order to propose a revised prioritization index based on historical selection of bridge replacement projects through statistical analysis, the set of computed performance measure value functions were applied to the database of bridge records developed in the previous chapter. The set of computed value function values for the state inventory serves as the predictor variables for all subsequent statistical regressions performed. The next component necessary for developing the statistical model is the response variable. For this study, the response variable event is the binary classification of whether or not a bridge has been selected for replacement by NCDOT engineers. NCDOT does not maintain a singular list that contains all of the bridges selected for replacement, but instead has two data sources called the Baseline Plan (BMIP) and a list of Active Bridge Projects (ABP). The BMIP contains a list of bridges that have been identified for future replacement, whereas the ABP contains all bridge projects that were occurring at the time that the study was performed. It was assumed that all bridge replacement projects were within these two lists. The BMIP records contain structure identification numbers, thus requiring few indexing procedures to build the response variable from this list of future projects. However, the ABP records did not include the structure identification numbers or sufficiently detailed descriptions of the scope of work, which requires additional preprocessing and verification of this list.

The ABP records contained three identifiers, which were the project contract number, Transportation Improvement Program (TIP) number, and WBS element number. The ABP records included all active bridge projects, so contracts that involve rehabilitation or repair actions were included in the list as well as bridge replacement projects, since the prioritization index is intended to rank only bridge replacement projects. The first step in the preprocessing of the list was to isolate only the replacement projects. Instead of checking each of the records individually, keywords that indicated that a bridge replacement was associated with a project contract were searched for in the contract description and location columns of each record to filter only the bridge projects associated with replacements. If an ABP record contained a keyword, the associated project contract documents were then manually searched for on the NCDOT Connect bidding and letting document database. The associated project contract documents were then manually reviewed to both confirm that the project was a replacement project and to identify the corresponding structure number. This manual review of contract documents was also effective in identifying instances where more than one bridge was replaced under the same contract.

Additional checks for both data sources include: duplicate checks; bridge structure checks; and a check of infrastructure condition ratings. An assumption was made that if none of the infrastructure condition ratings were lower than 7, then the data listed in the Network Master contained performance measure data was based on the newly replaced bridge and not the bridge that was being replaced. In these few instances, the associated response variable was assigned to indicate that the bridge is no longer selected for replacement. The filtering rules applied for both the BMIP and ABP databases are summarized in Table 4.1. In total, 258 active bridge replacement projects were identified from the ABP list. An additional 991 bridge replacement projects were identified from the BMIP database, resulting in a total 1249 bridges identified as bridges selected for replacement. This represents 9.028% of the 13,834 total number of bridge records assembled in the database.

4.2 Overview of Constrained Linear Regression

The constrained linear least squares (CLLS) method is a method for fitting a linear equation to measured data while satisfying a given set of criteria established with either constraint equations or bounds on the regression coefficients [Gavin, 2015]. This

TABLE 4.1: Filtering Guidelines Used to Develop List of Bridges Selected For Replacement

Database	Guidelines
ABP	Must have either a TIP or WBS contract number OR Must have a project contract number that matches with a project contract document in the NCDOT Connect portal AND Project contract explicitly states that the bridge will be replaced. AND Contract description contains the one of the words “Structure” “Str” ”bridge” or “replacement” AND Does not contain the words ”Preservation” or “Rehabilitation”
BMIP	Improvement type must be “Replacement” Let year is 2016 or later
Both	Duplicate check Condition ratings check Bridge structure check

method minimizes the value of the following sum-of-squares-of-errors (SSE) equation

$$\min \sum_{i=1}^{N} |C_i x - d_i|^2 \text{ where } Cx \leq d \text{ and } lb < x < ub \tag{4.1}$$

where N is the number of data observations included in the regression, C_i represents the set of predictor variables for the i^{th} data set, x represents linear regression coefficients, d_i the response variable, and lb and ub are the lower and upper bounds that represent the constraints for the potential solutions. In the application of the CLLS method to the bridge replacement dataset, the response variable is the binary response variable that indicates if each bridge was found to be selected for replacement or not. Additionally, the values of lb and ub are set as 0 and 1, respectively. This was done to ensure that all of the coefficients would be positive and reflect the assumed proportionality of the performance measures established in Table 3.7. The portion of the equation $Cx = d$ represents multiple predictor variables, and can be expanded as

$$C_1x_1 + C_2x_2 + ...C_nX_n = d \tag{4.2}$$

where $C_1 + C_2 + ...C_n$ are associated with each of the different performance measures and n represents the total number of predictor variables. The left hand side of this equation takes the same form as the additive multi-criteria utility function defined in the NCHRP 590 Report.

CLLS was performed within Excel by developing a spreadsheet to compute the SSE for a given set of regression coefficients and using the solver application within the Excel Analysis ToolPak. Of the three options that were available, the GRG Nonlinear solver was used. This solver iteratively adjust the values of the regression coefficients until a solution that minimizes the SSE is found. The solver application allows for bounds on the regression coefficients be established by encoding them as constraint equations.

Since the form of the linear regression model matches that of the additive multi-criteria utility function, the regression coefficients can be directly interpreted as relative weights. For some of the performance measures, the CLLS returned very small relative weights that suggest that the performance measure is not significant and can be removed from the prediction model. A threshold value of 0.0001 was used to identify such measures and the relative weights for each were forced to zero. Lastly, the relative weights were scaled such that the sum of relative weight totaled one using

$$\hat{x} = \frac{x}{\sum x} \tag{4.3}$$

where $\hat{x}_i$ is the scaled relative weight for performance measure i and n is the number of performance measures included in the regression. Since the value functions were developed on a scale of 0-100, the relative weights were scaled so that the prioritization score from the CLLS model scales to 0-100.

4.3 Overview of Binary Logistic Regression

Logistic regression is a statistical method of determining effects of predictors on the probability of responses [Pardoe et al., 2017]. This approach was used in several studies aimed at developing municipal sewer inspection prioritization methods that were discussed in the literature review. There are three types of logistic regression, which are binary, nominal, and ordinal [Pardoe et al., 2017]. These types are based on availability and classification of the possible responses. If there are only two outcomes possible, a binary logistic regression is used. If there are multiple outcomes that are possible, then a nominal logistic regression can be used. Additionally, if there is a hierarchy among the outcomes, ordinal logistic regression is used. Since the response variable indicating if a bridge has been selected for replacement can only have two outcomes, binary logistic regression was chosen for this study. Also, binary logistic regression has been found to be preferable for other infrastructure analyses as summarized in the literature review [Salman and Salem, 2012].

Binary logistic regression allows for estimation of the probability, p, of a binary event occurring x, [Pardoe et al., 2017] which can be defined as

$$p = \frac{\text{Outcome of Interest}}{\text{All Possible Outcomes}} \tag{4.4}$$

The odds of an event occurring is the ratio probability that an event will occur, $p(\text{event})$, to the probability of the event not occurring, $1 - p(x)$ [Foltz, 2015]

$$Odds = \frac{p(x)}{1 - p(x)} = \frac{p}{1 - p} \tag{4.5}$$

The relationship between odds and probability is illustrated in Figure 4.1.

Given that $p(x) = p$, the natural log of the odds of an event occurring is known as the logit of p

$$ln(\frac{p}{1 - p}) = logit(p) \tag{4.6}$$

The inverse logit provides a functional form that provides an s-shaped curve [Foltz, 2015] that represents the typical value function shape used in previous studies discussed in Section 2.3 of the literature review. This function also restrains the probability between 0 and 1, which is necessary for a binary response variable and is shown as

$$logit^{-1}(\alpha) = \frac{e^{\alpha}}{1 + e^{\alpha}} \tag{4.7}$$

Assuming that a prediction model has the shape of a logit function, the equation

$$ln(\frac{p}{1 - p}) = \beta_0 + \beta_1 x_1 + ...\beta_n x_n \tag{4.8}$$

where $\beta_0, \beta_1...\beta_n$ are regression coefficients, can be used to solve for p, which would

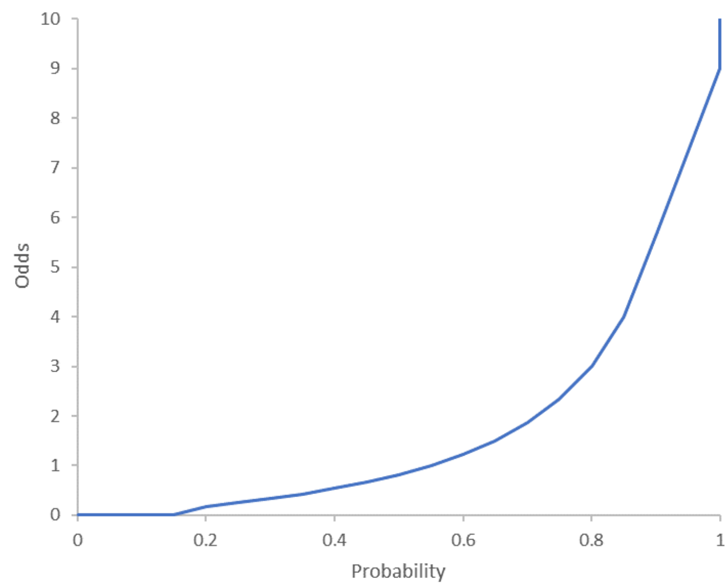


FIGURE 4.1: Relationship of Odds and Probability

result in the estimated probability from the regression model being

$$p = \frac{e^{\beta_0+\beta_1 x_1+\ldots\beta_n x_n}}{1 + e^{\beta_0+\beta_1 x_1+\ldots\beta_n x_n}} \tag{4.9}$$

Binary logistic modeling was completed using the Minitab statistical software package. The regression coefficients are calculated using the Maximum Likelihood Estimation (MLE) method [Pardoe et al., 2017]. Additionally, the predictors were reduced to the most significant performance measures using a backward stepwise elimination of variables. This is an iterative process that begins by computing the p-values for all potential predictor variables. For each step the predictor variable with that is found to be the least significant, or has the greatest p-value is removed until all variables have a p-value less than the specified alpha-to-remove, α [Pardoe et al., 2017], resulting in a final model. For this process, α=0.05. The logistic regression process is summarized in Figure 4.2.

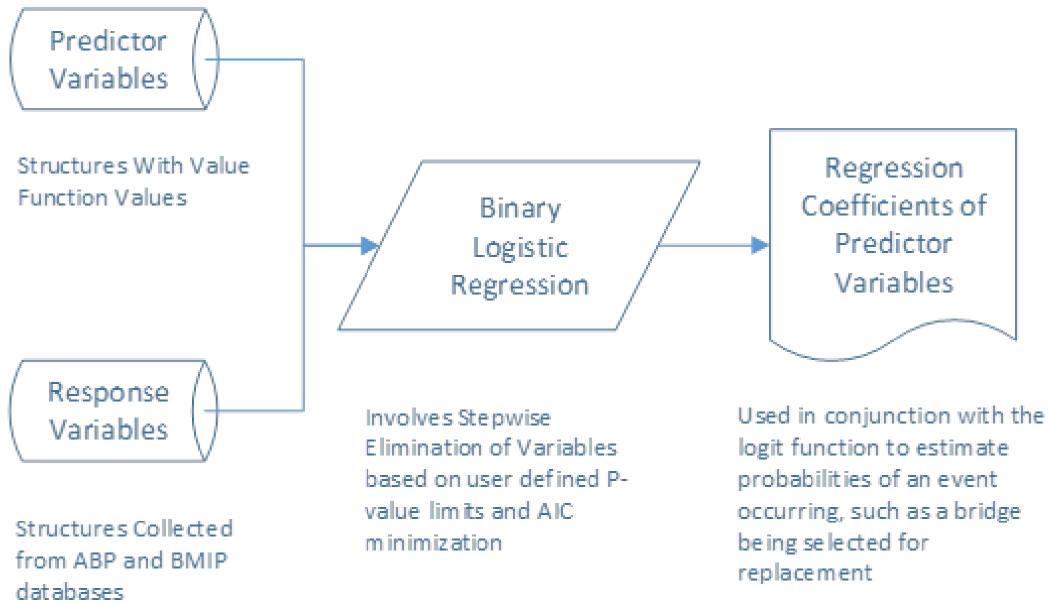


FIGURE 4.2: Logistic Regression Overview

4.4 Development of Matrix of Statistical Models Investigated

In the previous chapter, there are two distinct functional forms of the value functions that were defined, which are linear and derived from the ECDF. Additionally, there were two approaches proposed for calculating costs within the maintenance performance measures. One approach is to base the value functions on the total cost of the maintenance category, such as recommended maintenance total cost, and the other form is use the ratio of the total cost of a maintenance category to the replacement cost of the bridge. By performing separate statistical regressions using the different combinations of the forms of the value functions and maintenance costs, the optimal forms can be identified by assessing the goodness of fit for the various statistical models. In addition to investigating the functional form of the value functions and maintenance cost measures, the form of the statistical model used will also be assessed. One approach will be to fit the data to a form consistent with additive multi-criteria utility function using constrained linear least squares regression. The other approach will be to use binary logistic regression to predict the binary replacement classification following successful use of the statistical approach in infrastructure prioritization studies involving sewers and pipe systems that were identified in the literature review. Therefore, in order to determine the best combination of functional forms that will provide the best prediction models, different cases of all of the potential variants are defined in Table 4.2. These models will be evaluated and compared in the next chapter.

TABLE 4.2: Summary of Statistical Regression Test Cases

Case	Form of Value Function	Form of Maintenance Burden	Form of Statistical Model
LTC	Linear	Total	CLLS
ETC	ECDF	Total	CLLS
LRC	Linear	Replacement Ratio	CLLS
ERC	ECDF	Replacement Ratio	CLLS
LTB	Linear	Total	Binary Logistic Regression
ETB	ECDF	Total	Binary Logistic Regression
LRB	Linear	Replacement Ratio	Binary Logistic Regression
ERB	ECDF	Replacement Ratio	Binary Logistic Regression

Key to Case Letters:
First letter, value function form: L = Linear, E = ECDF-Derived
Second letter, maintenance form: T = Total, R = Replacement Ratio
Third letter, statistical model form: C = CLLS, B = Binary Logistic Regression

4.5 Summary of Statistical Models Developed

The statistical models developed over the matrix of cases presented in Section 4.2 are introduced in this Section. The significance of individual performance measures, as quantified by the regression coefficients, are examined across the models developed.

4.5.1 CLLS Regression Models

The linear regression coefficients developed for each of the CLLS models are shown in Table 4.3. Only significant performance measures with a relative weight greater than 0.0001 are shown in this table, which consists of 9 of the original 21 proposed performance measures. The ranking for each of the performance measures is shown in parentheses, where a rank of one represents the most significant variable, and a rank of two is the second most significant variable, etc. The LRC model retained the most variables (8), while the ETC and ERC retained the least variables (6). Priority maintenance needs, bridge posting, reoccurring burdensome maintenance, substructure condition rating, and scour criticality were shown to be significant among all four CLLS models. Priority maintenance maintained a ranking among the top three of all performance measures across all models. Recommended maintenance needs and fracture criticality were found to be significant for only one model, LRC. Scour criticality was found to be significant in each of the models, but remained in the lower half of rankings. The set of performance measures that were not used in any of the CLLS statistical models were: deck geometry, ADT, ADTT, interstate classification, secondary classification, detour length, preservation maintenance burden, deck condition rating, underclearance appraisal, major rehabilitation maintenance burden, reoccurring major rehabilitation burden, and reoccurring preservation maintenance burden.

Significant differences in performance measures included in the models and the associated relative weights suggests that the form of the value function is significant to the performance of the statistical model. In contrast, models that use the same functional format of the value function typically retained the same variables and coefficient values. In other words, the maintenance cost relative to replacement cost models were not found to have as significant effect on the regression coefficients and performance measures included in each statistical model. All four models retained at least three maintenance variables, with the LRC model retaining the highest amount of four.

TABLE 4.3: Relative Weights for Constrained Linear Least Squares Models (Relative Rank Shown in Parentheses)

Variable	ETC		ERC		LTC		LRC	
Priority Maint.	0.172	(3)	0.139	(3)	0.254	(2)	0.260	(1)
Bridge Posting	0.358	(1)	0.371	(1)	0.128	(4)	0.120	(4)
Reoc. Burd. Maint.	0.114	(4)	0.107	(4)	0.327	(1)	0.217	(2)
Substr. Cond.	0.272	(2)	0.274	(2)	0.129	(3)	0.075	(6)
Scour Criticality	0.043	(5)	0.043	(6)	0.041	(6)	0.032	(7)
Reccomended Maint.	-		-		-		0.199	(3)
Burdensome Maint.	0.041	(6)	0.065	(5)	0.110	(5)	-	
Fracture Critical	-		-		-		0.087	(5)
Superstr. Cond.	-		-		0.012	(7)	0.009	(8)

4.5.2 Binary Logistic Regression Models

The p-values for each of the variables found to be significant in the set of four models are shown in Table 4.4. P-values indicate the probability that a performance measure is statistically significant [Pardoe et al., 2017]. Performance measures with a p-value below a threshold value of 0.05 were used in the final binary logistic regression models developed. In contrast to the CLLS regression models, a larger number of performance measures were found to be statistically significant in the binary logistic regression models, with 16 measures included in the models developed with ECDF-derived value functions. There were 10 performance measures that were significant in all four models: substructure condition rating, superstructure condition rating, deck geometry rating, ADT, ADTT, bridge posting, interstate classification, secondary classification, scour criticality, and burdensome maintenance. Performance Measures that were not included in any of the models are: deck condition rating, underclearance appraisal, reoccurring major rehabilitation, reoccurring preservation maintenance, and fracture criticality.

Table 4.4: P-Values for Logistic Regression Models

Variable	ETB	ERB	LTB	LRB
Substr. Con.	0	0	0	0
Superstr. Con.	0	0	0	0
Deck Geometry	0	0	0	0
ADT	0.018	0.018	0.002	0.004
Bridge Posting	0	0	0	0
ADTT	0.007	0.005	0	0
Interstate Class	0.008	0.008	0.006	0.006
Scour Criticality	0	0	0	0
Burdensome Maint.	0	0	0.036	0.016
Reoc. Burd. Maint.	0.007	0.008	-	0
Reccomended Maint.	0	0	-	-
Major Rehab. Maint.	-	-	-	0.001
Secondary Class	0.004	0.007	0.002	0.004
Detour Length	-	-	0.005	0.009
Priority Maint.	0.004	0.005	-	-
Preservation Maint.	0.011	0.012	-	-

The odds ratios of the performance measures developed from the regression coefficients of each binary logistic regression model are listed in Table 4.5. Rankings of the odds ratios were developed by calculating the absolute difference of the odds ratio to 1, since the higher the difference, the greater the effect a change in value of a performance measure is predicted to have [IDRE, 2017]. The substructure condition was found to be the highest ranking variable for all models except the LRB model where major rehabilitation maintenance burden was found to be the highest ranking variable. As with the CLLS models, the models using the same functional form of the value function had similar variable rankings and a very strong correlation amongst their odds ratios. This further suggests that there is no significant difference between using total maintenance costs or maintenance costs relative to replacement cost for the maintenance-related performance measures. Typically, models with the same functional form of value functions retain the same performance measures, except for the instances of reoccurring burdensome maintenance and major rehabilitation between LTB and LRB. Major rehabilitation maintenance burden is also unique since this was the highest ranking variable for LRB, but found to be insignificant among the other three models.

Table 4.5: Odds Ratios of Binary Logistic Regression Models (Relative Rank Shown in Parentheses)

Variable	ETB		ERB		LTB		LRB	
Substr. Con.	1.044	(1)	1.044	(1)	1.062	(1)	1.063	(2)
Superstr. Con.	1.016	(2)	1.016	(2)	1.036	(3)	1.037	(4)
Deck Geometry	1.012	(3)	1.012	(3)	1.024	(4)	1.024	(6)
ADT	1.005	(8)	1.005	(8)	0.958	(2)	0.960	(3)
Bridge Posting	1.012	(4)	1.012	(4)	1.011	(7)	1.011	(9)
ADTT	1.006	(7)	1.006	(7)	1.018	(6)	1.017	(8)
Interstate Class	0.993	(5)	0.993	(5)	0.991	(8)	0.991	(11)
Scour Criticality	1.007	(6)	1.007	(6)	1.008	(9)	1.009	(10)
Burdensome Maint.	1.004	(10)	1.004	(10)	1.019	(5)	0.980	(7)
Reoc. Burd. Maint.	1.003	(12)	1.003	(12)	1	-	1.035	(5)
Reccomended Maint.	1.005	(9)	1.005	(8)	1	-	1	-
Major Rehab. Maint.	1	-	1	-	1	-	1.064	(1)
Secondary Class	1.004	(11)	1.003	(11)	1.004	(11)	1.003	(13)
Detour Length	1	-	1	-	0.994	(10)	0.994	(12)
Priority Maint.	1.003	(13)	1.003	(13)	1	-	1	-
Preservation Maint.	0.998	(14)	0.998	(14)	1	-	1	-

The reason that odds ratios are used to quantify the impacts of the performance measure rather than logistic regression coefficients is because the odds are directly proportional to changes in the performance measure values. This is illustrated in the following example. Consider a bridge with a current probability of replacement of 0.286, which corresponds to an odds of 0.400 for replacement selection. The odds ratio for each performance measure indicates the proportional change in odds for a unit change in value associated with the measure. However, since the raw values of the performance measures are converted to value functions, the value functions must be used to determine the effect of a unit change in a performance measure on the odds of replacement. For instance, a decrease in substructure from 7 to 6 would result in a value change of Δx_1=29.82 using the ECDF-based value function, as shown in Figure 4.3. Assuming that the conditions of the other performance measures remain the same, the new odds can be calculated using the odds ratio as shown

$$\text{(New Odds)} = \text{(Old Odds)} \times (\text{Odds Ratio}_{sub}^{\Delta_{sub}}) = (0.4) * (1.0436)^{(29.82)} = 1.428 \quad (4.10)$$

In other words, this change in condition rating would increase the odds of replacement by $(1.0436)^{29.82}$ or 357%. This increase in odds of replacement would be the same for any other bridge experiencing a change in substructure condition rating from 7 to 6 regardless of the original probability or odds associated with the structure. Furthermore, consider another scenario with a similar bridge with the same probability of replacement but with a substructure condition rating change from 5 to 4. As shown in 4.3, a lower change in value $\Delta X_2 = 16.01$ results.Such a change is less significant on the odds, as the calculated new odds of replacement would be

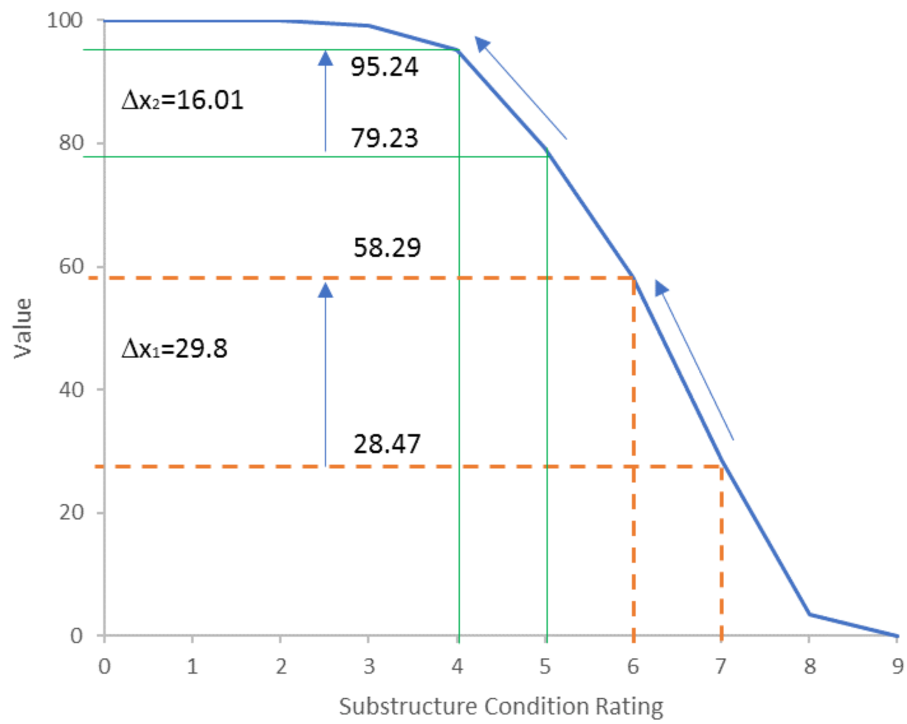


Figure 4.3: Substructure Condition Value Function with Example Odds Ratios

$(0.4) * (1.0436)^{(16.01)} = 0.792$ resulting in an increase of odds by only 198%.

Thus, while the change of the substructure condition rating score changed by one point for both scenarios, the change in value from the value functions were different. The value change in scenario 1 had a larger value change difference, resulting in a greater increase in odds, and ultimately an increase in the probability that the bridge will be replaced. Using logistic regression coefficients results in more difficulty in determining the effect of a change in performance measures on the probability and odds of replacement when ECDF based value functions are used since they are nonlinear. When linear value functions are used, the effect of changes in performance measures on the odds is uniform across the range of the performance measure.

The direct relationship between value changes and odds ratios can be expanded to

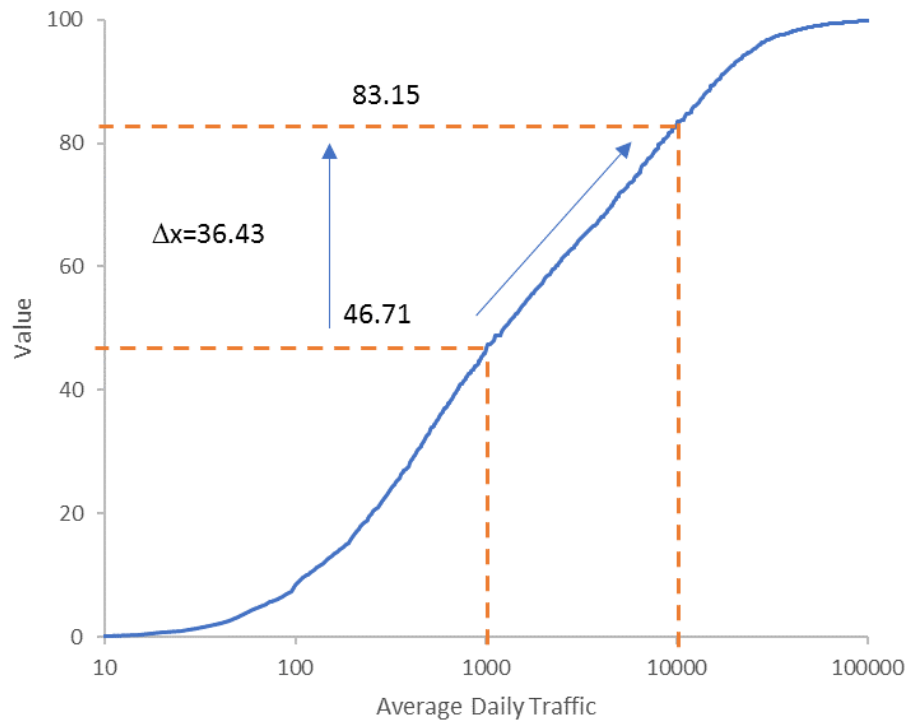


FIGURE 4.4: ADT Value Function with Example Odds Ratios

include multiple performance measures using the general equation

$$(\text{New Odds}) = (\text{Old Odds}) \prod_{i=1}^{n} (\text{Odds Ratio}_i^{\Delta_{x_i}}) \tag{4.11}$$

The use of this equation is shown in the following example. If the same bridge from before was affected with the substructure value change ΔX_1 and an increase of ADT from 1,000 to 10,000, which would result in a value change of 36.43 using the ECDF based ADT value function as shown in Figure 4.4, the net result would be an increase in odds of $(1.0436)^{29.82}(1.005)^{36.43} = 4.28$ or 428.6%. Of course, since the original odds of replacement for this hypothetical structure were low, even a large percentage increase in odds does not necessarily mean that these changes would make the structure a likely candidate for replacement.

Model summary statistics, including the R^2, adjusted R^2, and AIC for each of the models, are shown in Table 4.6. The R^2 statistic quantifies how well a model fits a given set of data while the adjusted R^2 measures the data fitting with inclusion of a penalty for use of more predictor variables [Pardoe et al., 2017]. The AIC determines the relative goodness of a model using a similar principle of measuring goodness of fit while penalizing additional model complexity [Pardoe et al., 2017]. The ETB has the best summary statistics of the four binary logistic regression models, although all models exhibit similar statistics. The predictive fidelity of the statistical models will be examined in more detail in the following chapter.

Table 4.6: Model Summarization of Logistic Regression Models

Model	Deviance R-Sq	Deviance R-Sq(adj)	AIC
ETB	0.3285	0.3268	5661.95
ERB	0.3282	0.3265	5665.80
LTB	0.3207	0.3194	5722.30
LRB	0.3232	0.3217	5705.52

CHAPTER 5: ASSESSMENT OF STATISTICAL MODELS

5.1 Introduction

The developed prediction models are compared for usefulness based on prediction accuracy and distribution of prioritization scores in this chapter. Thresholds for replacement status classification based on the model predictions are established by optimizing the predictive values. Prediction accuracy is assessed using sensitivity, specificity, and predictive value of a positive result. Analysis of the distribution of prioritization scores is based on visual assessment of histogram modality and characteristics, established PRI replacement candidacy thresholds, and the top ten scores assigned to bridge in the inventory compared with observed replacement selections. Based on the analysis of the models, recommendations for implementation are formulated.

5.2 Method for Developing of Classification Threshold

All of the developed statistical models assign a score to each structure on a scale of 0 to 100, which is assumed to reflect the priority for replacement based on the nature of the development of the statistical models. Since the actual priority rank for bridges in the current inventory is unknown, the predictive accuracy of the models can only be assessed by establishing a threshold value for each model to convert the scores developed by the statistical models to a binary classification of replacement status.

The model threshold value represents the minimum cut-off score that indicates if a bridge is predicted to be selected for replacement. It is important to develop a model that will provide the highest rate of correct predictions. Since each model has a potential score ranging from 0 to 100, a threshold of zero would mean that all bridges would be predicted to be replaced, while on other extreme of a threshold of 100, none of the bridges would selected for replacement. Thus, the ideal threshold for a model would have the highest rates of both the correctly predicted bridges selected for replacement and correctly predicted bridges not selected for replacement. These values are also known as predictive values, specifically the predictive value of a positive result (PV+) and predictive value of a negative result (PV-), respectively [Glasser, 2008]. The equations determining PV+ and PV-, from the true positives (TP), false positives (FP), true negatives (TN), and false negatives (FN) are

$$\text{PV+} = \frac{\text{TP}}{\text{TP} + \text{FP}} \tag{5.1}$$

$$\text{PV-} = \frac{\text{TN}}{\text{TN} + \text{FN}} \tag{5.2}$$

For each model, the PV+ and PV- test scores were calculated using the model prediction scores for every bridge in the state inventory against the list of observed bridges selected for replacement. Bridges predicted to be replaced by each statistical model were based on comparing the model prediction score to a variable threshold value. The data sets for observed predicted bridges selected for replacement were compared to compute TP, TN, FN, FP rates and the PV+ and PV- values for the variable threshold value. Through iterative adjustment of the threshold value to optimize the sum of the predictive values, a model threshold value was established for each statistical model.

5.3 Analysis of Predictive Accuracy of Models

The three tests that were used for determining the predictive accuracy for each of the statistical models developed in the prior chapter are sensitivity, specificity, and PV+. Recall that the equations for sensitivity and specificity are

$$\text{Sens.} = \frac{\text{TP}}{\text{TP} + \text{FN}} \tag{5.3}$$

$$\text{Spec.} = \frac{\text{TN}}{\text{TN} + \text{FP}} \tag{5.4}$$

in the context of the bridge replacement problem, sensitivity measures the percentage of the time that a bridge actually selected for replacement will be classified as selected for replacement based on the prediction score and assigned threshold value. Likewise, specificity is the percentage of cases where a bridge not selected for replacement is correctly predicted as not being a replacement candidate based on the prediction score and assigned threshold value. In addition to these standard statistical tests, an analysis of the accuracy of classification for the sets of bridges with the ten highest scores developed from each statistical model, similar to that performed for the PRI in Chapter 2, was performed. The values that were used for these calculations were determined using the optimal threshold values for each model based on maximizing the predictive values. The model that has the best average rank among the tests is considered to be the best model in terms of predictive accuracy. The results including relative rankings for each test are shown in Table 5.1.

Table 5.1: Predictive Accuracy Test Scores for Each Model (Relative Ranking in Parentheses)

Model	Sens.		Spec.		PV+		Top 10		Avg. Rank
ERB	91.03 %	(6)	75.84 %	(2)	27.22 %	(3)	5	(2)	**3.25**
LRB	89.67 %	(8)	76.94 %	(1)	27.85 %	(1)	4	(3)	**3.25**
ETB	91.91 %	(3)	74.41 %	(5)	26.28 %	(5)	6	(1)	3.50
LTB	91.91 %	(3)	74.45 %	(4)	26.31 %	(4)	3	(4)	3.75
PRI	91.67 %	(5)	75.80 %	(3)	27.32 %	(2)	2	(5)	3.75
ERC	94.08 %	(1)	64.57 %	(9)	20.86 %	(9)	2	(5)	6.00
LTC	92.95 %	(2)	68.61 %	(8)	22.72 %	(7)	1	(8)	6.25
ETC	89.91 %	(7)	68.65 %	(7)	22.17 %	(8)	2	(5)	6.75
LRC	89.03 %	(9)	71.34 %	(6)	23.56 %	(6)	0	(9)	7.50

The ERC model was found to have the highest sensitivity among all models and PRI, while the LRB model was found to have the highest specificity and PV+. The ERC model, while performing the best for sensitivity, was found to have the worst specificity and PV+ with a significantly lower value than any of the other models without such a significant difference for the sensitivity. Similarly, the LRB had the second lowest sensitivity while performing the best for the other two tests. The LRC performed the worst among all models for sensitivity. The ETB model was found to have the most correctly classified bridges selected for replacement in the list of the top ten prioritization scores (6), while only the LRC model had no bridges actually selected for replacement in its ranking of the top ten. Generally, the logistic regression models had the most bridges selected for replacement in the top ten bridges compared to the CLLS models. Compared to the PRI, all models except for LTC and LRC performed just as well or better for this performance test. Analysis of the classification success across the bridges receiving the top ten scores for each of the models are also shown in Figures 5.1 to 5.8. Overall, the ERB and LRB models performed the best across all four tests based on the average test ranking, while the LRB model would be considered the best model if based only on the number of best scores developed across all tests. Based on the relative rankings, the models developed with the binary logistic regression outperformed those developed using constrained linear least squares regression. Interestingly, models with maintenance costs relative to replacement costs outperformed those using total replacement costs for the binary logistic regression models.

5.4 Analysis of Distribution of Prioritization Scores

An important additional consideration when determining the usefulness of an index is the distribution of scores for the bridges across the statewide inventory. Ideally, the distribution of scores should maximize the range of the index, distinguish replacement candidates from those not considered for replacement, and avoid clustering of assigned prioritization scores. The PRI was found to have a bimodal distribution of scores among the observed bridges selected for replacement, meaning that the value of the scores do not properly reflect the intended effect of ranking the bridges for replacement. A visual analysis of the distribution of prioritization scores developed by each of the statistical models based on the modality, skew, and range was utilized to assess how well each model produced a desirable distribution of scores. An ideal distribution of scores for bridges not selected for replacement would be heavily skewed toward the lower scores since the majority of the state inventory is not in urgent need of replacement. The scores of bridges that are selected for replacement would ideally

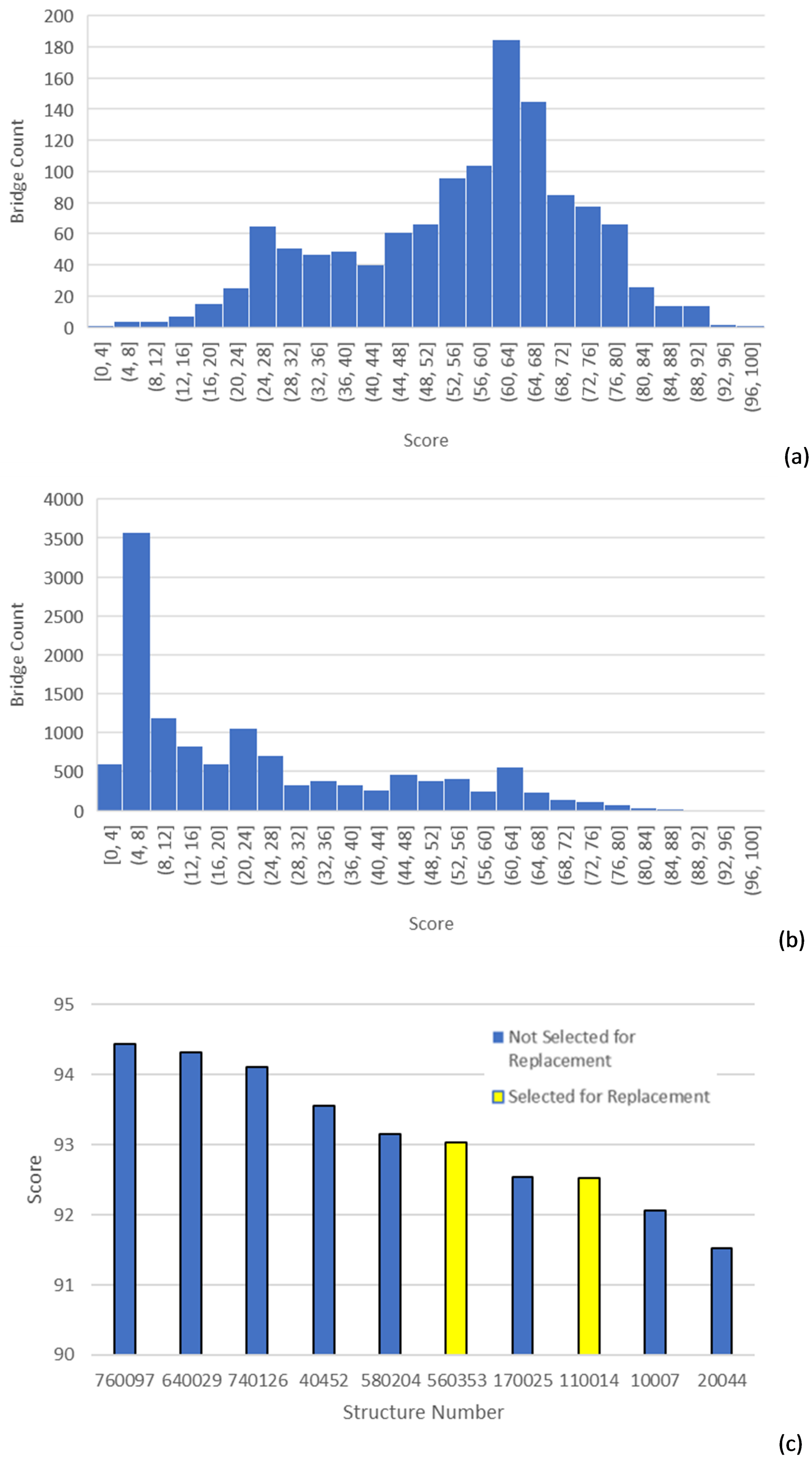


Figure 5.1: Model ETC: (a) Bridges Selected For Replacement (b) Bridges Not Selected For Replacement (c) Bridges with the Top 10 Scores

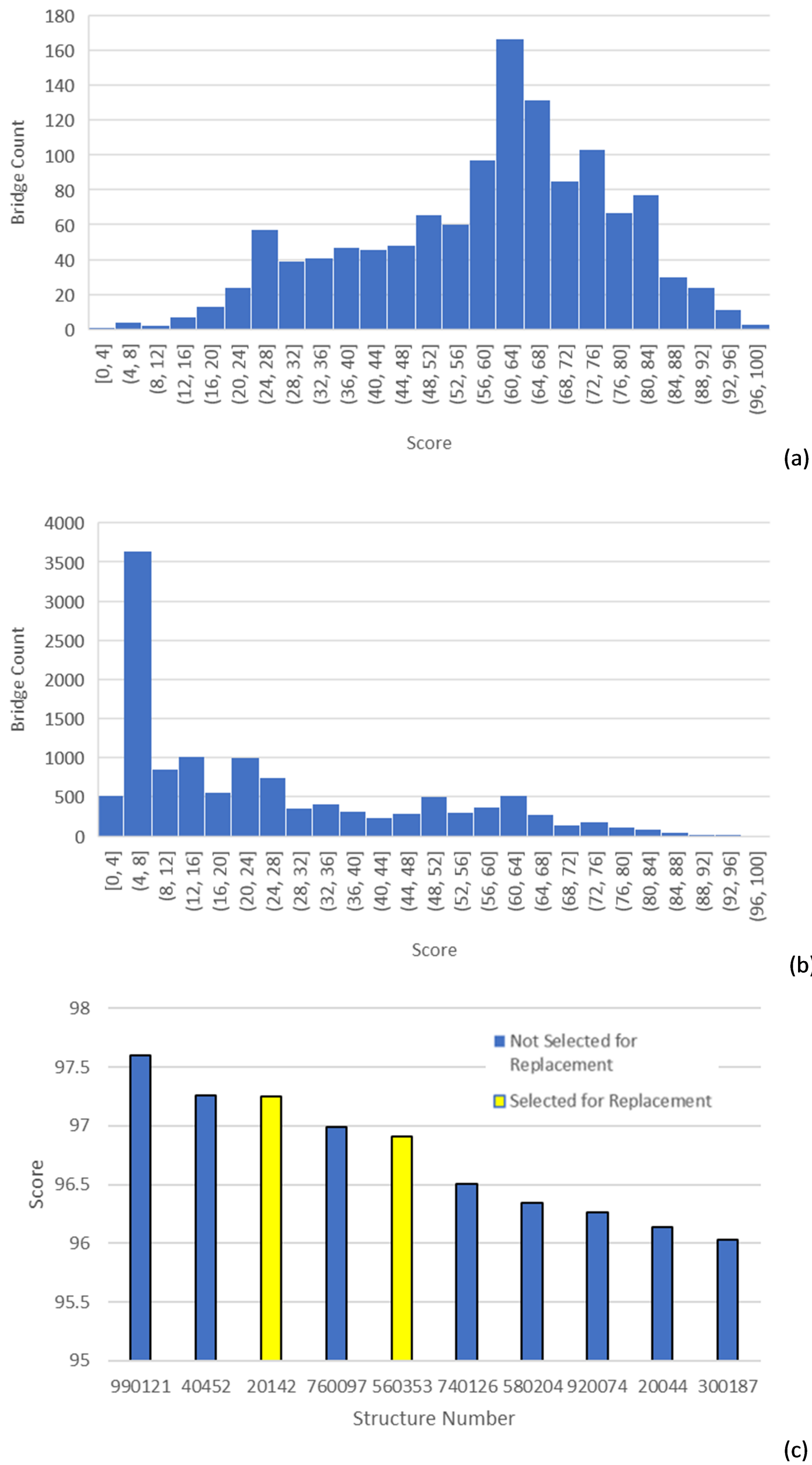


Figure 5.2: Model ERC: (a) Bridges Selected For Replacement (b) Bridges Not Selected For Replacement (c) Bridges with the Top 10 Scores

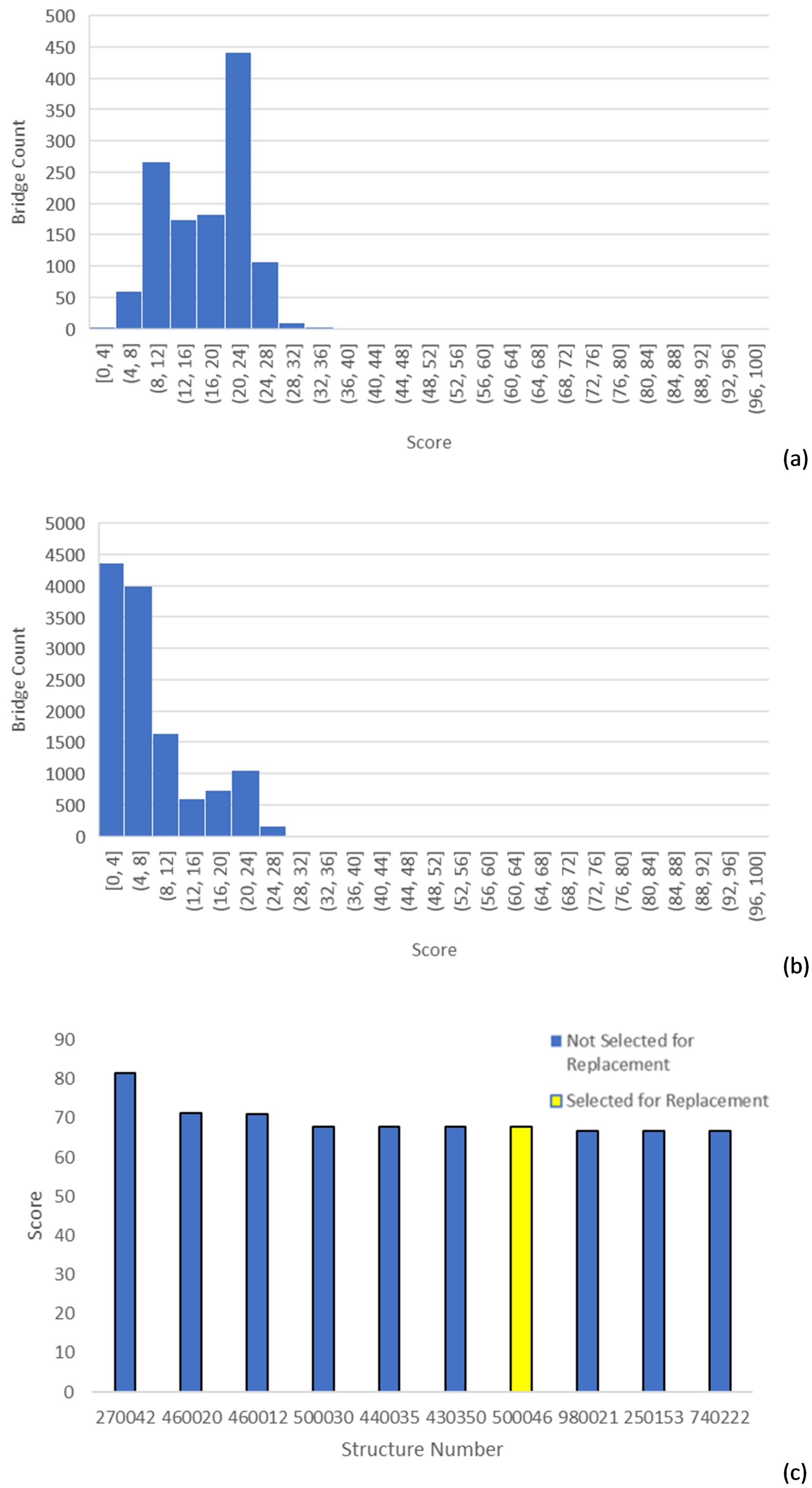
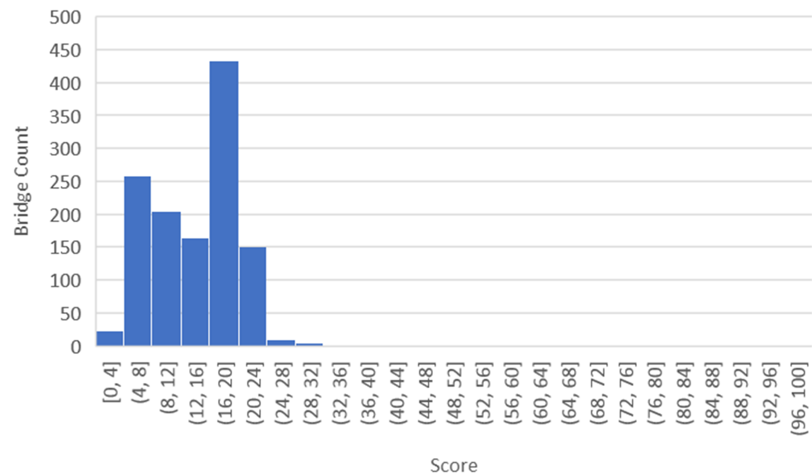
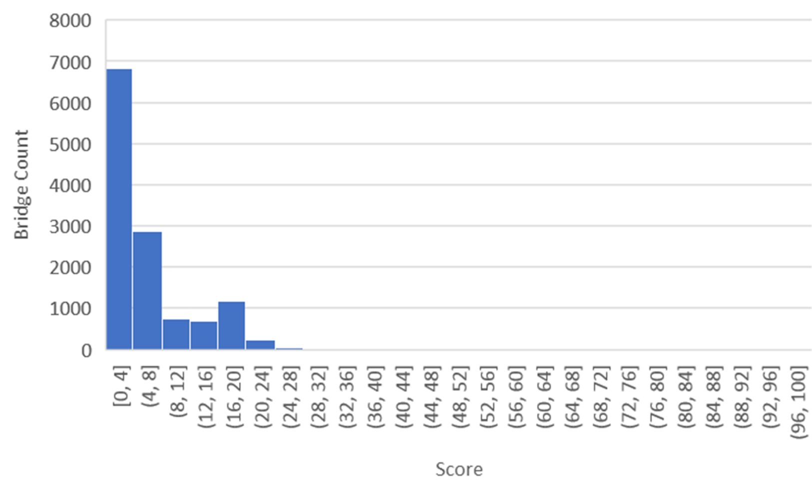


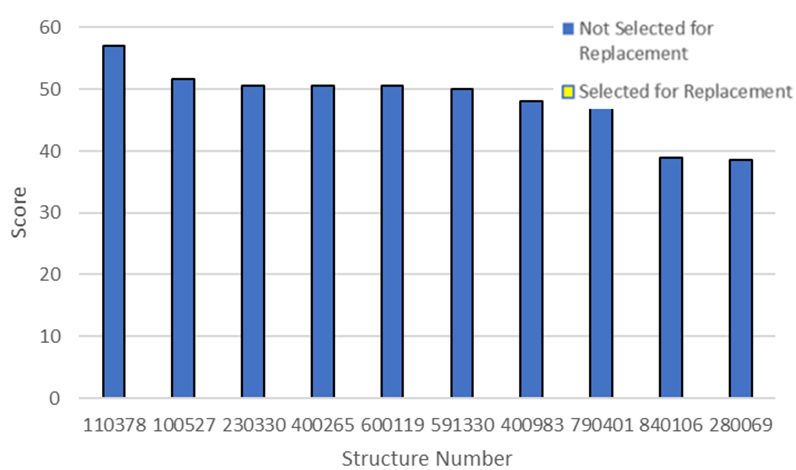
Figure 5.3: Model LTC: (a) Bridges Selected For Replacement (b) Bridges Not Selected For Replacement (c) Bridges with the Top 10 Scores



(a)



(b)



(c)

FIGURE 5.4: Model LRC: (a) Bridges Selected For Replacement (b) Bridges Not Selected For Replacement (c) Bridges with the Top 10 Scores

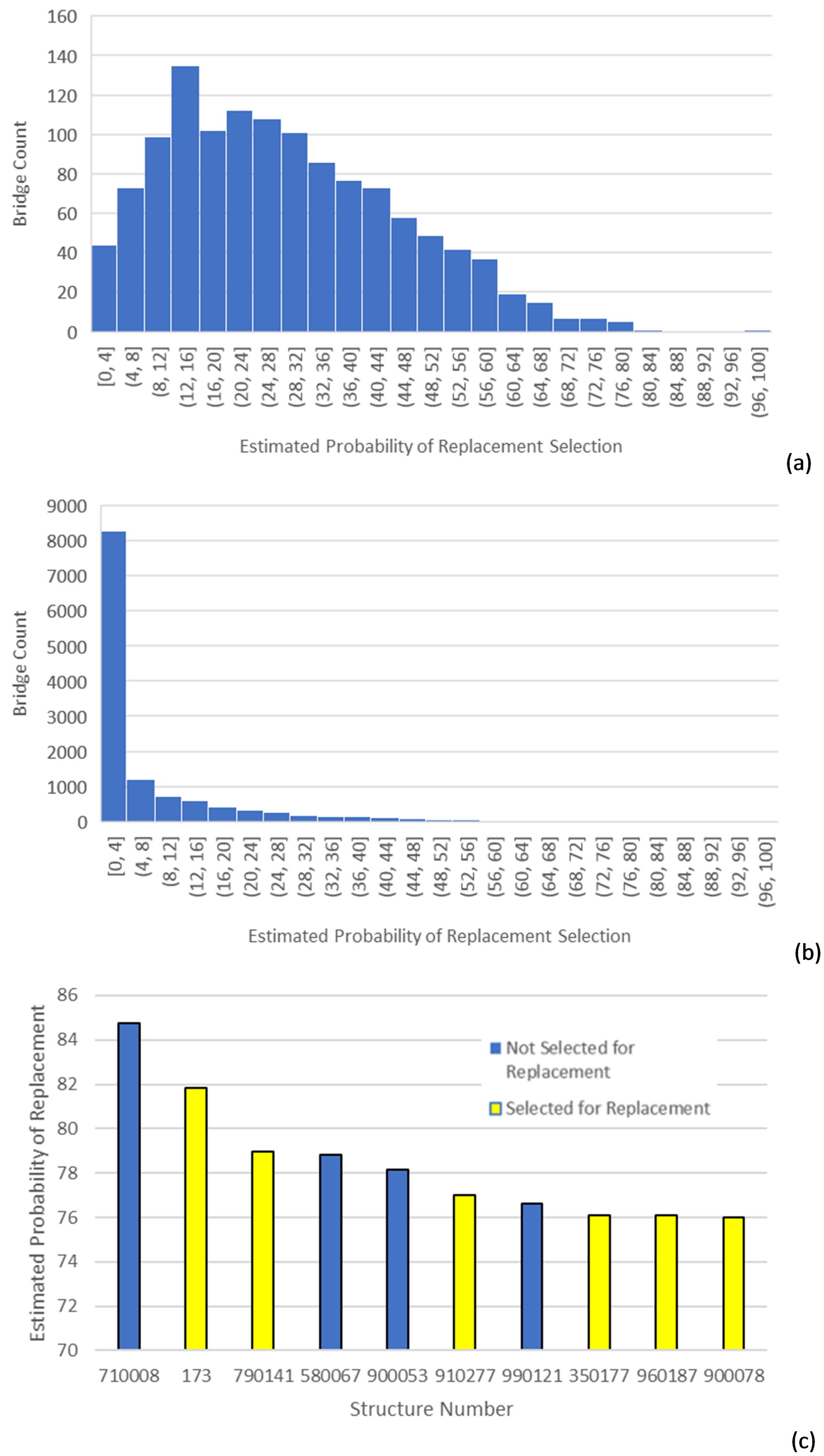


Figure 5.5: Model ETB: (a) Bridges Selected For Replacement (b) Bridges Not Selected For Replacement (c) Bridges with the Top 10 Scores

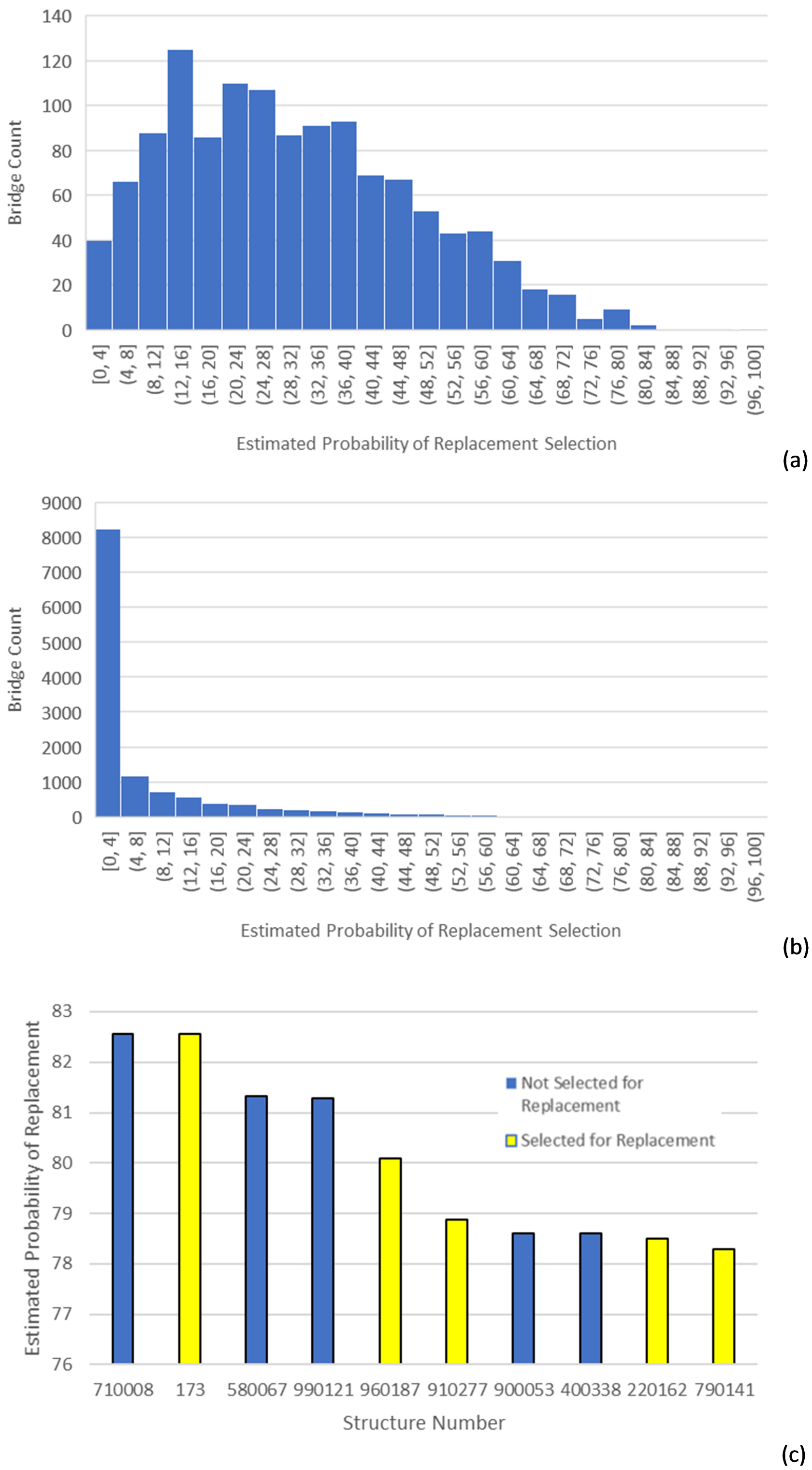


FIGURE 5.6: Model ERB: (a) Bridges Selected For Replacement (b) Bridges Not Selected For Replacement (c) Bridges with the Top 10 Scores

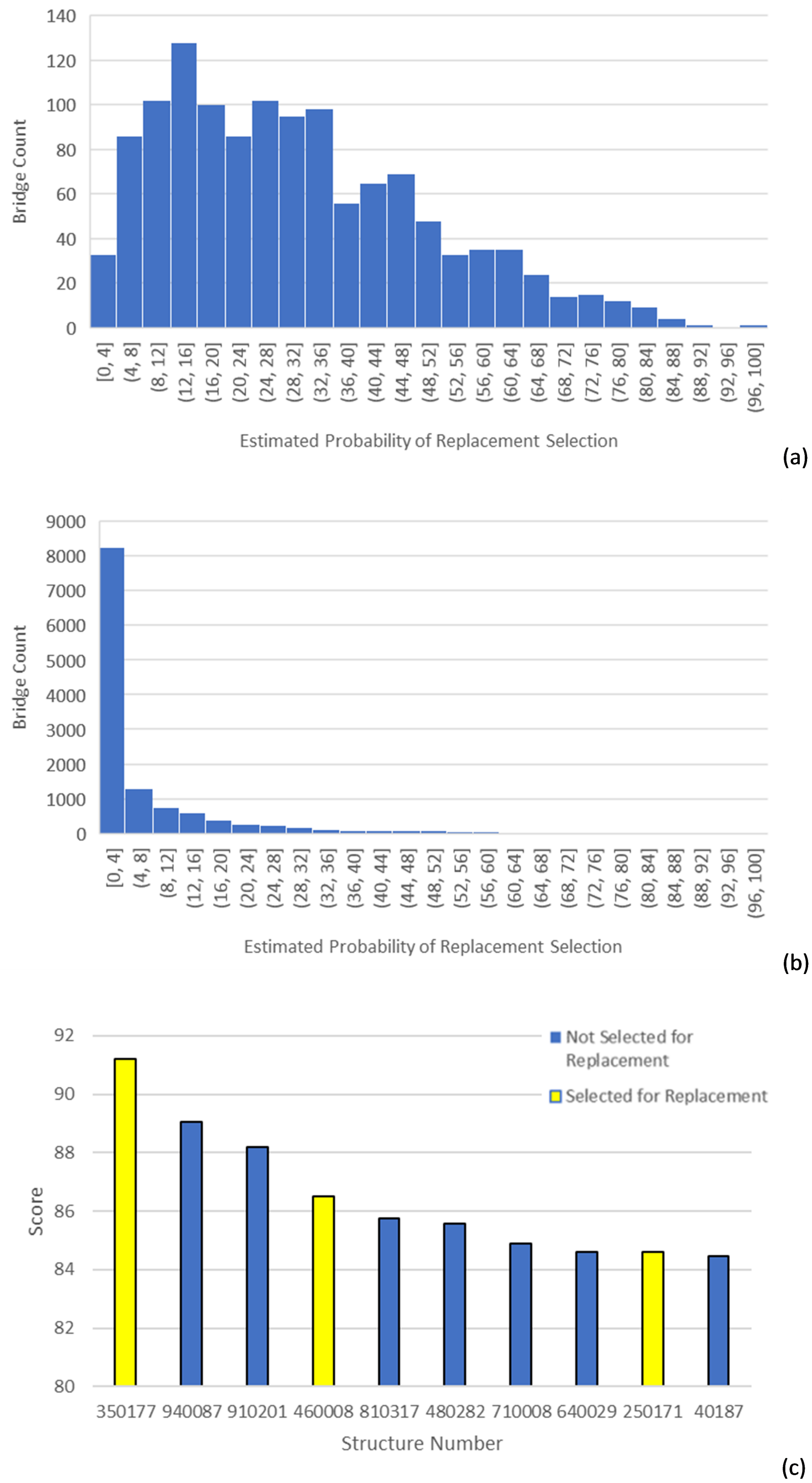


FIGURE 5.7: Model LTB: (a) Bridges Selected For Replacement (b) Bridges Not Selected For Replacement (c) Bridges with the Top 10 Scores

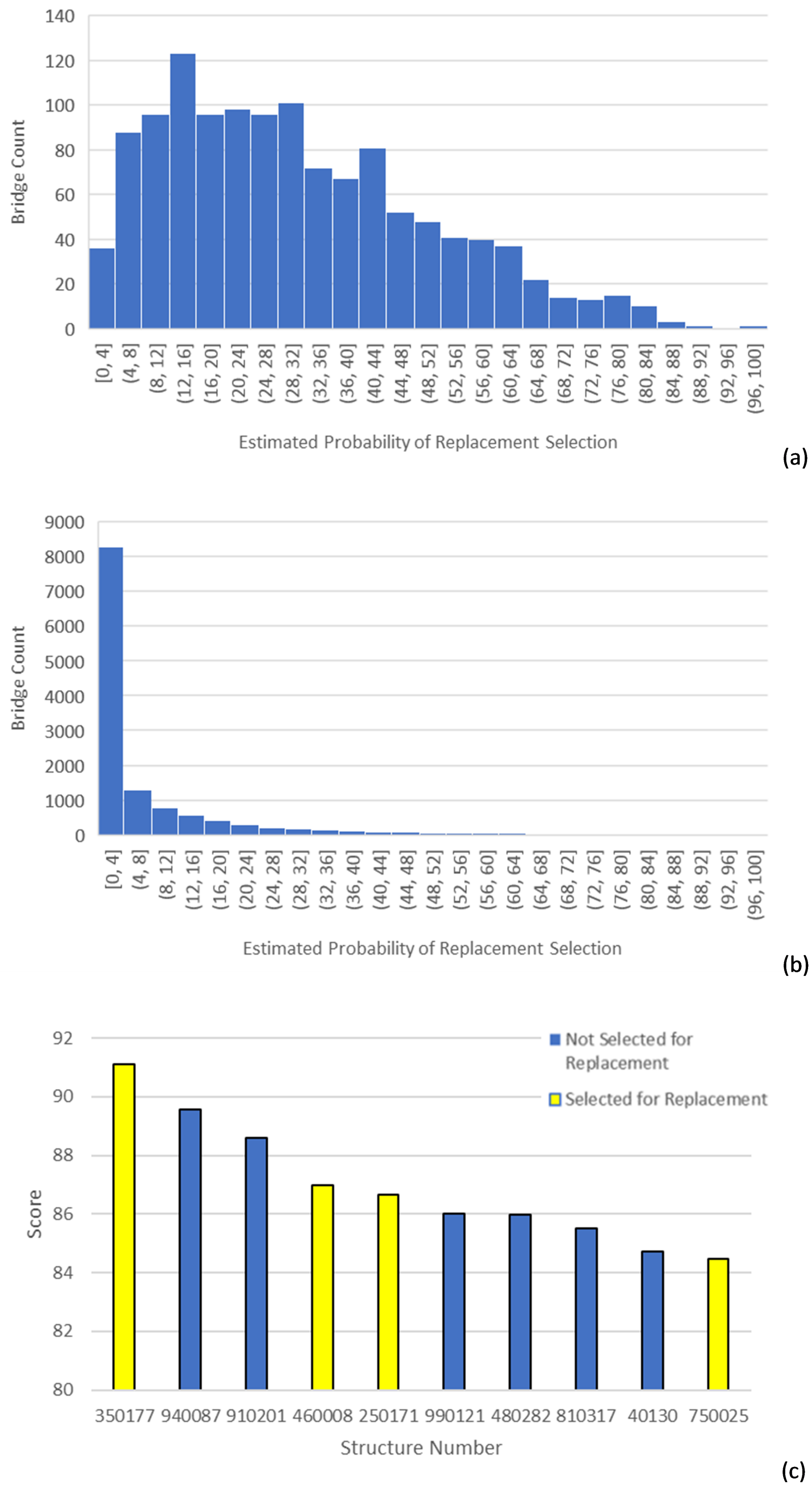


FIGURE 5.8: Model LRB: (a) Bridges Selected For Replacement (b) Bridges Not Selected For Replacement (c) Bridges with the Top 10 Scores

be a uniform distribution across the higher range scores to provide clear separation of priority amongst different candidates. Additionally, the combination of the two distributions would utilize the total range of potential scores. The distribution of the bridges selected and not selected for replacement for each of the developed statistical models are shown in Figures 5.1 to 5.8. Each of the histograms were consistently developed across a range of 0 to 100 with 25 uniformly sized bins each having a size of 4 points.

Overall, three of the CLLS models (ETC, LTC, LRC) contained a slightly bimodal distribution among the bridges selected for replacement, while the distribution for the ERC model was found to be generally normal. The LTC and LRC models had a clustering of prioritization scores and did not fully utilize the full range of the index. The binary logistic regression models have generally normal distributions for the scores assigned to bridges selected for replacement, although the scores are mildly skewed toward lower values. The ETB model has the smoothest distribution of scores across the bridges selected for replacement. For the bridges not selected for replacement, the binary logistic regression models do a superior job of assigning low scores to these bridges with only a small fraction of the bridges in the inventory not selected for replacement receiving higher prioritization scores.

The overall PRI replacement candidacy thresholds introduced in Chapter 2, where a score within the range of 0 to 30 is considered a poor replacement candidate, 30 to 50 range is considered a good candidate, and beyond 50 is considered a very good candidate, were also applied to the developed statistical models to evaluate the general distribution of scores relative to the actual classification status. Odds ratios were calculated for each model over each of the candidacy threshold categories to produce a score that normalizes the classification success based on the relative number of bridges predicted to receive a score within each category. Ideally, the odds ratio for bridges selected for replacement should increase across the candidacy threshold categories. These odds ratios as well as the distribution of observed bridges selected or not selected for replacement are summarized in Tables 5.2 to 5.4. The logistic regression models, in particular the ETB model, has the best odds ratio distribution amongst the three thresholds. The constrained linear least squares regression models perform especially poor, as the odds ratios for event the highest score category are significantly below one.

5.5 Summary and Recommended Prediction Model

The statistical models developed in the previous chapter were subjected to a number of tests to analyze their relative predictive accuracy and distribution of prioritization scores. Overall, the logistic regression models were found to have a better distribution of scores and improved predictive accuracy than CLLS models. Among the predictive accuracy tests and top ten bridge score replacement classification test, the LRB model was found to have the best ranking among the two of the four tests, while overall, the ERB and LRB model were found to have the best average relative rankings. However, compared to the LRB model, the ERB model was found to have more observed bridges selected for replacement among the top ten prioritization scores, a smoother distribution of scores for the bridges selected for replacement, as well as a more ideal odds ratio for bridges to be selected rather not selected among the "very good" replacement candidate range. Consequently, the recommended statistical model for the analysis performed is the ERB model, which was based on binary logistic regression using ECDF-derived value functions and maintenance performance measures that use maintenance costs relative to the replacement cost. It should be noted that, due to similar performance between the ERB and LRB models, it is likely that modest improvement may be possible by combining some performance measures with ECDF-derived value functions with others defined using linear value functions with total maintenance costs and binary logistic regression. Additionally, these models can be validated with future inspection and maintenance needs data. Overall, the ERB model results in some improvements in the predictive values, but does not provide significant difference in performance relative to the PRI in terms of accuracy based only on binary classification using a fixed threshold. However, the ERB model produces significant improvements in the distribution of prioritization scores. Furthermore, the ERB model provides an index that avoids double counting of variables, incorporates element-level condition ratings, reflects maintenance history, and allows for direct interpretation of how individual performance measures contribute to the prioritization score using odds ratios.

Table 5.2: Distribution of Bridges Among Three Thresholds For Logistic Regression Models

ETB				
Score	Not Selected	Selected	Grand Total	Odds Ratio
[0,30)	11885	717	12602	0.060
[50,30)	560	377	937	0.673
[50,100]	140	155	295	1.107
Grand Total	12585	1249	13834	
ERB				
Score	Not Selected	Selected	Grand Total	Odds Ratio
[0,30)	11773	671	12444	0.057
[30,50)	614	383	997	0.624
[50,100]	198	195	393	0.985
Grand Total	12585	1249	13834	
LTB				
Score	Not Replaced	Replaced	Grand Total	Odds Ratio
[0,30)	11847	680	12527	0.057
[30,50)	498	358	856	0.719
[50,100]	240	211	451	0.879
Grand Total	12585	1249	13834	
LRB				
Score	Not Replaced	Replaced	Grand Total	Odds Ratio
[0,30)	11843	673	12516	0.057
[30,50)	495	358	853	0.723
[50,100]	247	218	465	0.883
Grand Total	12585	1249	13834	

Table 5.3: Distribution of Bridges Among Three Thresholds For CLLS Models

ERC				
Score	Not Selected	Selected	Grand Total	Odds Ratio
[0,30)	8526	125	8651	0.015
[50,30)	1837	245	2082	0.133
[50,100]	2222	879	3101	0.396
Grand Total	12585	1249	13834	

ETC				
Score	Not Selected	Selected	Grand Total	Odds Ratio
[0,30)	8730	149	8879	0.017
[30,50)	1804	258	2062	0.143
[50,100]	2042	842	2884	0.412
Grand Total	12576	1249	13825	

LTC				
Score	Not Selected	Selected	Grand Total	Odds Ratio
[0,30)	3		3	0.000
[30,50)	571	58	629	0.102
[50,100]	12011	1191	13202	0.099
Grand Total	12585	1249	13834	

LRC				
Score	Not Selected	Selected	Grand Total	Odds Ratio
[0,30)	12570	1246	13816	0.099
[30,50)	9	3	12	0.333
[50,100]	6		6	0.000
Grand Total	12585	1249	13834	

Table 5.4: Distribution of Bridges Among Three Thresholds For PRI

PRI				
Score	Not Selected	Selected	Sum	Odds Ratio
[0,30)	10966	303	11269	0.028
[30,50)	1023	409	1432	0.400
[50,100]	588	537	1125	0.913
Grand Total	12577	1249	13826	

CHAPTER 6: CONCLUSION

6.1 Summary of Research and Key Findings

This thesis seeks to develop a data-driven prediction tool that can be used to assist NCDOT Structures Management Unit (SMU) engineers during the decision making process used in prioritizing highway bridge replacement projects. A literature review was performed to summarize current approaches for prioritizing bridge replacement projects and developing risk-based priority indices using decision analysis and utility theory. In addition, prioritization methods for other infrastructure systems that use statistical regression to create prediction models for prioritization were reviewed. Shortcoming in the current PRI prioritization index were identified and a new set of performance measures and criteria were developed for an improved index based on analysis of historical data. Additionally, performance measures were introduced to quantify the impact of prior maintenance actions, including burdensome reoccurring maintenance, on the prioritization of replacement. Strategies for analyzing historical maintenance data were developed and automated analysis of multiple bridge and maintenance databases was performed by preparing a macro. New performance measures were also introduced to incorporate the element level conditions of each of the bridges by using inspector recommended maintenance lists that indicate the severity of an element condition and the cost associated with the maintenance action. Multiple methods were assessed to determine how maintenance would best correlate with the values of the bridge engineers. The performance measures were converted into value functions using two different functional forms that describe the value trade-offs for each of the bridges, ECDFs and linear functions. The response variable used for statistical models, the binary indication from the list of bridges selected for replacement, was defined and compiled using data sources that consist of active bridge projects and bridges planned for replacement in the future. A matrix of multiple regression models, consisting of variations of the functional forms of value functions, maintenance costs, and regression approaches, was developed. The statistically significant performance measures identified within each regression models were discussed to examine their relative importance and comparisons were made among the models in the matrix. A threshold of the prioritization scores was defined for each of the statistical models to produce a binary classification of bridge replacement status by optimizing the amount of correct predictions of bridges selected and not selected for replacement. A methodology for determining the best statistical model for predicting bridge replacement selections based on predictive accuracy and distribution of priority replacement scores was defined, using established processes used for validating models identified in the literature review. Ultimately, resulting in a recommended statistical model for predicting bridge project replacements was identified that used ECDF-based value functions, maintenance costs computed using total sums, and the binary logistic regression models.

6.2 Recommendations for Future Work

The following is a summary of potential improvements and recommendations for future research:

- The statistical models developed are based on practitioner preference and risk as reflected in historical decision making, however, no direct practitioner input was incorporated. Surveying methods, such as those outlined in NCHRP 590 could be used to acquire such feedback. Survey data from bridge engineers and the statistical regression results can be combined using the backcalculated nonstandardized normal (BNN) coefficient method as discussed in the literature review [Moruza et al., 2016].

- The significance of other performance measures not currently considered in the models could be assessed, including, such as: distance to critical facilities, correlation of individual maintenance actions and probability of bridge replacement selection, and future ADT. These are measures that have not been identified in the initial set of performance measures investigate by NCDOT.

- The use of mixed functional forms of value functions, such as using linear and ECDF value functions together with the same model, to potentially improve predictive accuracy should be explored.

- The use of value difference to prioritize maintenance actions among different bridges similar to INDOT's use of disutility difference between a do-nothing approach and a list of given maintenance actions with standard disutility effects

[Sinha et al., 2009] could also be evaluated as a potential improvement to the methodology.

- Currently, estimated replacement costs are calculated in the BMS using a simple formula based on the system of the route that the bridge is on and the total deck area. The use of maintenance performance measures normalized to the estimated replacement cost should be revisited after research is performed to improve the replacement cost estimation models. Research on this topic is currently being investigated [Phillips, 2017].

- The prediction model development process can be generalized to other assets, such as culverts, sewer pipelines, dams, and other critical infrastructure systems. Additionally, the methodology could potentially be extended to include prioritization ratings for rehabilitation and preservation projects.

REFERENCES

[Ariaratnam et al., 2001] Ariaratnam, S. T., El-Assaly, A., and Yang, Y. (2001). Assessment of infrastructure inspection needs using logistic models. *Journal of Infrastructure Systems*, 7(4):160–165.

[ASCE, 2016] ASCE (2016). Failure to act: Closing the infrastructure investment gap for america's economic future.

[Davies et al., 2001] Davies, J., Clarke, B., Whiter, J., Cunningham, R., and Leidi, A. (2001). The structural condition of rigid sewer pipes: a statistical investigation. *Urban Water*, 3(4):277–286.

[Farrar and Newton, 2014] Farrar, M. M. and Newton, B. (2014). Perspective: The AASHTO manual for bridge element inspection. *ASPIRE*.

[Foltz, 2015] Foltz, B. (2015). Statistics 101: Logistic regression, an introduction.

[Gavin, 2015] Gavin, H. (2015). Constrained linear least squares.

[Glasser, 2008] Glasser, S. P. (2008). Research methodology for studies of diagnostic tests. In *Essentials of Clinical Research*, pages 245–257. Springer.

[IDRE, 2017] IDRE (2017). How do i interpret odds ratios in logistic regression? — stata faq.

[Johnson, 2008] Johnson, M. B. (2008). Project prioritization using multiobjective utility functions. *Transportation Research Circular E-C127*, pages 189–194.

[Johnston and Zia, 1984] Johnston, D. W. and Zia, P. (1984). *Level-of-service system for bridge evaluation*. Number 962.

[Lane, 2016] Lane, K. (2016). *Performance criteria and measures for prioritization of bridge replacement projects*. MS Thesis. University of North Carolina at Charlotte. Charlotte, North Carolina.

[MAP-21, 2012] MAP-21 (2012). P.l. 112-141, the moving ahead for progress in the 21st century act (MAP-21).

[Moruza et al., 2016] Moruza, A. K., Matteo, A. D., Mallard, J. C., Milton, J. L., Nallapaneni, P. L., and Pearce, R. L. (2016). Methodology for ranking relative importance of structures to Virginias roadway network. Technical report.

[NCDOT, 2016] NCDOT (2016). Maintenance operations and performance analysis report (MOPAR). Technical report.

[Pardoe et al., 2017] Pardoe, I., Simon, L., and Young, D. (2017). STAT 501: Regression methods.

[Patidar et al., 2007] Patidar, V., Labi, S., Sinha, K., and Thompson, P. (2007). NCHRP report 590. *Multiobjective Optimization for Bridge Management Systems*.

[Phillips, 2017] Phillips, P. (2017). *Predicting Costs for Bridge Replacement Projects*. MS Thesis. University of North Carolina at Charlotte. Charlotte, North Carolina.

[Saito and Sinha, 1989] Saito, M. and Sinha, K. (1989). The development of optimal strategies for maintenance, rehabilitation and replacement of highway bridges, final report vol 5: priority ranking method.

[Salman and Salem, 2012] Salman, B. and Salem, O. (2012). Modeling failure of wastewater collection lines using various section-level regression models. *Journal of Infrastructure Systems*, 18(2).

[Shepard and Johnson, 2001] Shepard, R. W. and Johnson, M. B. (2001). California bridge health index: A diagnostic tool to maximize bridge longevity, investment. *TR News*, (215).

[Sinha et al., 2009] Sinha, K. C., Labi, S., McCullouch, B. G., Bhargava, A., and Bai, Q. (2009). Updating and enhancing the Indiana bridge management system (BMS).

[Skinner, 2009] Skinner, D. C. (2009). *Introduction to decision analysis: a practitioner's guide to improving decision quality*. Probabilistic Pub.

[Sobanjo and Thompson, 2016] Sobanjo, J. O. and Thompson, P. D. (2016). Implementation of the 2013 AASHTO manual for bridge element inspection.

[Weseman, 1995] Weseman, W. (1995). Recording and coding guide for the structure inventory and appraisal of the nation's bridges. *United States Department of Transportation (Ed.), Federal Highway Administration, USA*.

Appendix A: Excel VBA Macro Script for Performance Measure Value Function Development

```
Option Explicit

Private Sub Import_Data()

'***************************
'   Variable Type Definitions
'***************************

Dim wb1 As Workbook              'This workbook.
Dim wb2 As Workbook              'The workbook with data to be imported.
Dim FileToOpen As Variant        'Directory location with data.
Dim Sheet As Worksheet           'Sheets in wb2 to be imported.
Dim PasteStart As Range          'Location of where data will be pasted.
Dim c As Range                   'Counter variable.
Dim firstAddress As String       'cell that meets find criteria.
Dim arr() As Variant
Dim aCell As Range
Dim LastRow As Long
Dim lastRowA As Long
Dim lastRowB As Long
Dim bCell As Range
Dim PNCell As Range
Dim RNCell As Range
Dim rng1 As Range
Dim i As Integer
Dim RowMatch As Integer


'""""""""""""""""""""""""""""
' DEFINE the priority level system
'""""""""""""""""""""""""""""
Dim target As Double           'Value to index

Dim arr1 As Variant
Dim arr2 As Variant
Dim arr3 As Variant

Dim arrayCollection() As Variant
Dim arrayWorksheets() As Variant

Dim boo As Variant           ' "True" value if target isin a given array.
Dim x As Variant             'counter variable.
Dim r As Variant             'row counter variable.
Dim Col_Val As Variant       'location of first empty column.
Dim Var As Variant           'priority of a structure.
Dim repl As String           'location of column
Dim priorityLvl As String

Dim ws_Network As Worksheet  'identify worksheet.
Dim ws_BMIPbase As Worksheet
Dim ws_BMIPdyn As Worksheet
Dim ws_Needs As Worksheet
Dim ws_NeedsPivot As Worksheet
Dim ws_BaselineWithPRI As Worksheet

Set wb1 = ThisWorkbook
Set ws_Network = wb1.Sheets("bridges")
Set ws_Needs = wb1.Sheets("Needs")
Set ws_NeedsPivot = wb1.Sheets("NeedsPivot")
Set ws_BaselineWithPRI = wb1.Sheets("BaselineWithPRI")
```

```
''//*****************************************************************
''// Import the bridge database file to the main spreadsheet.
''//*****************************************************************
'Define this workbook.
'Define the pasting location.
'Clear out previous data on worksheet.
Set wb1 = Workbooks("CDMCommonDatabaseMaker.xlsm")
Set PasteStart = wb1.Sheets("Bridges").Range("A1")

With wb1
.Sheets("Bridges").Cells.Clear
.Sheets("needs").Cells.Clear
.Sheets("needsPivot").Cells.Clear
.Sheets("BaselineWithPRI").Cells.Clear
End With

'""""""""""""""""""""""""""""
' IMPORT Structure database
'""""""""""""""""""""""""""""
'Define workbook with data to import.
'If no file is selected, end the program.
'Define the workbook to be imported and opened.
'Copy all of the data into an array.
'Increase the paste location to fit the whole array, and paste array.

FileToOpen = Application.GetOpenFilename _
(Title:="Select the Bridge AgileAssets file to import", _
FileFilter:=".xls (*.xls*),")

If FileToOpen = False Then
MsgBox "No file specified", vbExclamation, "ERROR"
Exit Sub

Else

Set wb2 = Workbooks.Open(Filename:=FileToOpen)

For Each Sheet In wb2.Sheets
With Sheet.UsedRange
arr = .Value
End With
Next Sheet
wb2.Close
PasteStart.Resize(UBound(arr, 1), UBound(arr, 2)).Value = arr
End If

'"""""""""""""""""""""""""""""
' IMPORT BMIP Databases
'"""""""""""""""""""""""""""""
'Define the pasting location.
'Copy all of the data into an array.
'Increase the paste location to fit the whole array,
'and paste array.
Set PasteStart = wb1.Sheets("BaselineWithPRI").Range("A1")
FileToOpen = Application.GetOpenFilename _
(Title:="Select BMIP Baseline Database file to import", _
FileFilter:=".xls (*.xls*),")

If FileToOpen = False Then
MsgBox "No file specified", vbExclamation, "ERROR"
Exit Sub
Else
Set wb2 = Workbooks.Open(Filename:=FileToOpen)
For Each Sheet In wb2.Sheets
With Sheet.UsedRange
arr = .Value
End With
Next Sheet
```

```
wb2.Close
PasteStart.Resize(UBound(arr, 1), UBound(arr, 2)).Value = arr
End If

Sheets("Bridges").Activate

'""""""""""""""""""""""""
' FLAG Non-Criteria Structures.
'""""""""""""""""""""""""
'find the column with target data
'find last row.
'if the column is found.
With ws_Network
Set aCell = .Rows(1).Find("structure no.")
If Not aCell Is Nothing Then
LastRow = .Range(Split(.Cells(, aCell.Column).Address, _
"$")(1) & .Rows.Count).End(xlUp).Row
End If
End With

For i = 2 To LastRow

Select Case Cells(i, 8).Value
Case "O"
Case "P"
Case "C"
Case Else
Cells(i, 8).Value = "ZZ"
End Select

Next i

'"""""""""""""""""""""""""
' SORT by structure type.
'"""""""""""""""""""""""""

Sheets("bridges").Activate
With wb1.Sheets("bridges").UsedRange
.Sort _
key1:=Cells(1, 8), _
order1:=xlAscending, _
Header:=xlYes
End With

'""""""""""""""""""""""""""
' REMOVE the structures not within the given criteria
'""""""""""""""""""""""""""

With Sheets("bridges").Columns(8)
Set c = .Find("ZZ", LookIn:=xlValues)
If c <> False Then
firstAddress = c.Address
End If
End With
Sheets("bridges").Range(firstAddress, Cells(i, 8)) _
.EntireRow.Delete




'""""""""""""""""""""""""""""""""""
' POPULATE arrays with range data
'""""""""""""""""""""""""""""""""""

With ws_BaselineWithPRI
Set aCell = .Rows(1).Find("number")
If Not aCell Is Nothing Then
```

```
LastRow = .Range(Split(.Cells(, aCell.Column) _
.Address, "$")(1) & .Rows.Count).End(xlUp).Row
arr1 = .Range(aCell.Offset(1), _
.Cells(LastRow, aCell.Column))
End If
End With




With ws_Network

'""""""""""""""""""""""""""""""""""
' FIND first empty column and create priority level column.
'""""""""""""""""""""""""""""""""""
'Go all the way to the right then go to first col on left.
Col_Val = .Cells(1, .Columns.Count) _
.End(xlToLeft).Column
If Col_Val > 1 Then
Col_Val = Col_Val + 1
End If

.Cells(1, Col_Val).Value = "Priority Level"

'"""""""""""""""""""""""""""""""""""
' FIND last row of given column header
'"""""""""""""""""""""""""""""""""""
'find the column with target data
'if the column is found.
'find last row.
Set aCell = .Rows(1).Find("structure no.")
If Not aCell Is Nothing Then
LastRow = .Range(Split(.Cells(, aCell.Column). _
Address, "$")(1) & .Rows.Count).End(xlUp).Row


Set bCell = .Rows(1).Find("Consider Replacement?")

End If

'"""""""""""""""""""""""""""""""""""
' FLAG priority level for the bridges.
'"""""""""""""""""""""""""""""""""""

For x = 2 To LastRow

target = .Cells(x, aCell.Column).Value
repl = .Cells(x, bCell.Column).Value
Var = .Application.Match(target, arr1, 0)
If IsError(Var) Then
If repl = "Yes" Then
Var = 0
Else
Var = -100
End If

End If

'"""""""""""""""""""""""""""""""""""
' ASSIGN priority level
'"""""""""""""""""""""""""""""""""""

Select Case Var
Case Is > 0
priorityLvl = "BaselineWithPRI"
Case Is = 0
priorityLvl = "Consider"
Case Is = -100
```

```
priorityLvl = "No Consideration"

End Select
.Cells(x, Col_Val).Value = priorityLvl

Next x

End With


''//*******************************************
''// INSPECTOR RECOMMENDED MAINTENANCE NEEDS
''//*******************************************
'Define the pasting location.
'Copy all of the data into an array.
'Increase the paste location to fit the whole array,
'and paste array.

Set PasteStart = wb1.Sheets("Needs").Range("A1")
FileToOpen = Application.GetOpenFilename _
(Title:="Select Inspector Recommended Maintenance Needs" _
& "file to import", _
FileFilter:=".xls (*.xls*),")

If FileToOpen = False Then
MsgBox "No file specified", vbExclamation, "ERROR"
Exit Sub

Else
Set wb2 = Workbooks.Open(Filename:=FileToOpen)
For Each Sheet In wb2.Sheets
With Sheet.UsedRange
arr = .Value
End With
Next Sheet
wb2.Close
PasteStart.Resize(UBound(arr, 1), UBound(arr, 2)).Value = arr
End If

Dim PivotSheet As Worksheet
Dim PivotName As String
Dim nTargetCol1 As Integer
Dim nTargetCol2 As Integer
Dim LastColumn As Integer
Dim aHeaderList As Variant
Dim pivotData As Range
PivotName = "NeedsPivot"
Set PivotSheet = Sheets("NeedsPivot")

''//******************************************************************
''// CREATE TOTAL COST COLUMN
''//******************************************************************
With ws_Needs
.Activate
'Define boundaries.
LastColumn = fFindLastColumn
LastRow = fFindLastRow
aHeaderList = fCreateHeaderList
'Return the column of the queried attributes.
nTargetCol1 = Application.Match("Quantity", aHeaderList, 0)
nTargetCol2 = Application.Match("Actual Unit Cost", aHeaderList, 0)
.Cells(1, LastColumn + 1) = "Total Cost"
'Return the product of quantity and actual unit cost in the last
'row.
For x = 2 To LastRow
.Cells(x, LastColumn + 1) = .Cells(x, nTargetCol1) * _
.Cells(x, nTargetCol2)
Next x
```

```
''//******************************************************************
''// DEFINE source data for pivot table.
''//******************************************************************
'Redefine the table to include the new Total Cost Column
LastColumn = LastColumn + 1
'Create a regular table with the maintenance needs data.
fCreateTable ("NeedsTable")
'Define source data for the pivot table.
ActiveWorkbook.PivotCaches.Create(SourceType:=xlDatabase, SourceData:= _
"NeedsTable", Version:=xlPivotTableVersion14).CreatePivotTable _
TableDestination:="NeedsPivot!R2C2", TableName:=PivotName, _
DefaultVersion:=xlPivotTableVersion14
Sheets("NeedsPivot").Select
Cells(2, 2).Select
ActiveWorkbook.ShowPivotTableFieldList = True
''//******************************************************************
''// DEFINE attributes to populate pivot table.
''//******************************************************************

With PivotSheet.PivotTables(PivotName).PivotFields("Structure No.")
.Orientation = xlRowField
.Position = 1
End With

With PivotSheet.PivotTables(PivotName).PivotFields("Priority Type")
.Orientation = xlColumnField
.Position = 1
End With

With PivotSheet.PivotTables(PivotName).PivotFields("Maintenance Code")
.Orientation = xlColumnField
.Position = 2
End With

PivotSheet.PivotTables(PivotName).AddDataField PivotSheet.PivotTables( _
PivotName).PivotFields("Total Cost"), "Sum of Cost", xlSum

'Format the Pivot Table
With PivotSheet.PivotTables(PivotName)
.ColumnGrand = False
.RowGrand = False
.RowAxisLayout xlTabularRow
End With
End With

''//******************************************************************
''// IMPORT maintenance needs cost data to the Network Master sheet.
''//******************************************************************

'Define location for importing data.
With ws_Network
.Activate
lastRowA = fFindLastRow

Set aCell = .Rows(1).Find("structure no.")
End With

'Find locations of wanted attributes in the pivot table.
With ws_NeedsPivot
.Activate
LastRow = fFindLastRow
Set bCell = .Rows(4).Find("structure no.")
Set PNCell = .Rows(3).Find("Priority Maintenance Total")
Set RNCell = .Rows(3).Find("Recommended Maintenance Total")
LastColumn = RNCell.Column
End With
```

```
''//*****************************************************************
''// MATCH attribute data to global key.
''//*****************************************************************

'Define the Global Key table

Dim rGlobalKeys As Range
Dim rLocalDatabase As Range
Dim rPriStart As Range
Dim rRecStart As Range
Dim lastColumnA As Integer

'Define Search parameter ranges.
ws_Network.Activate
Set rGlobalKeys = ws_Network.Range(Cells(2, aCell.Column), _
Cells(lastRowA, aCell.Column))

ws_NeedsPivot.Activate
Set rLocalDatabase = ws_NeedsPivot.Range(Cells(bCell.Row + 1, _
bCell.Column), Cells(LastRow, LastColumn))

'Write name of column to be populated in the global database.
With ws_Network
.Activate
lastColumnA = fFindLastColumn
.Cells(1, lastColumnA + 1).Value = "Priority Maintenance Total"
.Cells(1, lastColumnA + 2).Value = "Recommended " _
& "Maintenance Total"

'Define starting points of column to populate.
Set rPriStart = .Cells(2, lastColumnA + 1)
Set rRecStart = .Cells(2, lastColumnA + 2)

Dim rPopRange As Range

'Populate columns.
Set rPopRange = Range(rPriStart, Cells(lastRowA, rPriStart.Column))
rPopRange = Application.WorksheetFunction.VLookup(rGlobalKeys, _
rLocalDatabase, 32, False)

Set rPopRange = Range(rRecStart, Cells(lastRowA, rRecStart.Column))
End With

End Sub

Function fFindLastColumn() As Integer
With ActiveSheet
fFindLastColumn = .Cells(1, .Columns.Count).End(xlToLeft).Column
End With
End Function

Function fFindLastRow() As Double
With ActiveSheet
If Application.WorksheetFunction.CountA(.Cells) <> 0 Then
fFindLastRow = .Cells.Find(what:=".", _
after:=.Range("a1"), _
lookat:=xlPart, _
LookIn:=xlFormulas, _
searchorder:=xlByRows, _
searchdirection:=xlPrevious, _
MatchCase:=False).Row
Else
fFindLastRow = 1
End If
End With
End Function
```

```
Public Function fCreateHeaderList() As Variant
'Requires nLastCol
Dim nLastCol As Integer
Dim origin As Range

nLastCol = fFindLastColumn

With ActiveSheet
Set origin = .Cells(1, 1)
fCreateHeaderList = .Range(origin, .Cells(1, nLastCol)).Value
End With

End Function

Function fCreateValueArray(sAttName As String) As Variant
'assumes data is on activesheet.
'finds attribute name in header and creates an
'array of the associated data.
Dim nLastCol As Integer
Dim nLastRow As Double
Dim rOrigin As Range
Dim aHeaderList As Variant
Dim nTargetCol As Variant
Dim rng As Range

'define parameters.
nLastCol = fFindLastColumn
nLastRow = fFindLastRow
Set rOrigin = ActiveSheet.Cells(1, 1)

'create header array.
aHeaderList = fCreateHeaderList

'Return the column of the queried attribute.
nTargetCol = Application.Match(sAttName, aHeaderList, 0)

'Create array of the attribute values.
Set rng = ActiveSheet.Range(Cells(rOrigin.Row + 1, _
nTargetCol), Cells(nLastRow, nTargetCol))
fCreateValueArray = rng.Value


End Function

Function fCreateTable(sTableName As String)
'Monday, 2/13/2017 ANA
'Creates a table with all range data on active sheet.
Dim rData As Range
With ActiveSheet
Set rData = .UsedRange
.ListObjects.Add(xlSrcRange, rData, , xlYes).Name = sTableName
End With
End Function

'""""""""""
'Part 2 of Macro
'""""""""""
Sub MaintBurden()

Dim targetSheet As Worksheet
Dim PasteStart As Range
Dim lastCol As Long
Dim LastRow As Long
Dim sTarget As String
Dim nTargetCol As Long
Dim arr As Variant
Dim i As Long
Dim x As String
```

```
Dim c As Range
Dim firstAddress As String
Dim sTableName As String
Dim PivotName As String

'""""""""""""""""""""""""""""
' IMPORT Maintenance Burden Database
'""""""""""""""""""""""""""""
'Define worksheet to import data.
Set targetSheet = Sheets("MaintRaw")

With targetSheet
.Activate
.Cells.Clear

Set PasteStart = .Cells(1, 1)
End With

Call mImportFile(PasteStart)

'""""""""""""""""""""""""""""
' CREATE a structure ID for each structure in
' the maintenance burden raw data.
'""""""""""""""""""""""""""""

'Define parameters of used range in active sheet.
lastCol = fFindLastColumn
LastRow = fFindLastRow
sTarget = "INV_ELEM_NAME"

'Create a title for column


Cells(1, lastCol + 1) = "Structure ID"
nTargetCol = fFindCol(sTarget)
arr = Range(Cells(2, nTargetCol), Cells(LastRow, nTargetCol)).Value

'convert array strings into numbers.
'must subtract 10000 to convert from 99 county numbers to 100.
For i = LBound(arr) To UBound(arr)
x = arr(i, 1)
x = Val(Mid(x, 2, 3) & Right(x, 4)) - 10000
arr(i, 1) = x
Next i

'Paste array into empty column.

Set PasteStart = Range(Cells(2, lastCol + 1), _
Cells(LastRow, lastCol + 1))

PasteStart.Value = arr

'""""""""""""""""""""""""""""
' CREATE year column for each structure in the maintenance burden
' raw data sheet.
'""""""""""""""""""""""""""""

'Create a title for column
lastCol = fFindLastColumn
sTarget = "End_Date"
Cells(1, lastCol + 1) = "Year"

'Find column with target string.
nTargetCol = fFindCol(sTarget)

'Create an array of target column data.
arr = Range(Cells(2, nTargetCol), Cells(LastRow, nTargetCol)).Value
```

```
'convert array strings into years.
For i = LBound(arr) To UBound(arr)
x = arr(i, 1)
x = Val(Mid(x, 7, 4))
arr(i, 1) = x
Next i

'Paste array into empty column.

Set PasteStart = Range(Cells(2, lastCol + 1), _
Cells(LastRow, lastCol + 1))

PasteStart.Value = arr

'"""""""""""""""""""""""""""
' Remove maintenance actions originating before 2007
'"""""""""""""""""""""""""""

'"""""""""""""""""""""""""
' SORT by structure type.
'"""""""""""""""""""""""""
lastCol = fFindLastColumn
'Sort column in ascending order.
With ActiveSheet.UsedRange
.Sort _
key1:=Cells(1, lastCol), _
order1:=xlDescending, _
Header:=xlYes
End With

'""""""""""""""""""""""""""
' REMOVE the structures not within the given criteria
'""""""""""""""""""""""""""

With ActiveSheet.Columns(lastCol)
Set c = .Find(2006, , xlValues, xlPart, xlByRows, xlNext, _
False, , False)
If c Is Nothing Then

Else

firstAddress = c.Address
ActiveSheet.Range(firstAddress, Cells(LastRow, lastCol _
)).EntireRow.Delete
End If
End With


'""""""""""""""""""""""""""
' REMOVE potential year typos --- Check
'""""""""""""""""""""""""""
With ActiveSheet.UsedRange
'Sort column in ascending order.
.Sort _
key1:=Cells(1, lastCol), _
order1:=xlAscending, _
Header:=xlYes
End With

With ActiveSheet.Columns(lastCol)
Set c = .Find(2077, , xlValues, xlPart, xlByRows, xlNext, False, , False)
'hard coded in year.
If c <> False Then
firstAddress = c.Address
End If
End With

ActiveSheet.Range(firstAddress, Cells(LastRow, lastCol)).EntireRow.Delete
```

```
'""""""""""""""""""""""""""
' REMOVE structures with strange IDs
'""""""""""""""""""""""""""
sTarget = "Structure ID"
nTargetCol = fFindCol(sTarget)

With ActiveSheet.UsedRange
'Sort column in ascending order.
.Sort _
key1:=Cells(1, nTargetCol), _
order1:=xlDescending, _
Header:=xlYes
End With

With ActiveSheet.Columns(nTargetCol)
Set c = .Find(-10000, , xlValues, xlPart, xlByRows, xlNext, False, , False)
'hard coded in strID.
If c <> False Then
firstAddress = c.Address
End If
End With

ActiveSheet.Range(firstAddress, Cells(LastRow, lastCol)).EntireRow.Delete


'""""""""""""""""""""""""""
' REMOVE actions with negative costs.
'""""""""""""""""""""""""""

sTarget = "TOTAL_COST"
nTargetCol = fFindCol(sTarget)

With ActiveSheet.UsedRange
'Sort column in ascending order.
.Sort _
key1:=Cells(1, nTargetCol), _
order1:=xlDescending, _
Header:=xlYes
End With

With ActiveSheet.Columns(nTargetCol)
.Select
Set c = .Find("(", , xlValues, xlPart, xlByRows, _
xlNext, False, , False)

If c Is Nothing Then

Else
firstAddress = c.Address
c.Select

End If
End With

ActiveSheet.Range(firstAddress, Cells(LastRow, lastCol)) _
.EntireRow.Delete

'""""""""""""""""""""""""""
' RE-SORT by year
'""""""""""""""""""""""""""

sTarget = "Year"
nTargetCol = fFindCol(sTarget)

With ActiveSheet.UsedRange
'Sort column in ascending order.
.Sort _
```

```
key1:=Cells(1, nTargetCol), _
order1:=xlAscending, _
Header:=xlYes
End With


'"""""""""""""""""""""""""""
' CREATE Table
'"""""""""""""""""""""""""""

Sheets("maintraw").Activate
sTableName = "MaintRawTable"
fCreateTable (sTableName)

'remove formulas for speed


'"""""""""""""""""""""""""""
' CREATE Action Classification Column
'"""""""""""""""""""""""""""

lastCol = fFindLastColumn
Cells(1, lastCol + 1) = "Action Classification"
Cells(2, lastCol + 1).Formula = _
"=VLOOKUP([@[ACTIVITY_NAME]],ActionClassTable,2)"

'"""""""""""""""""""""""""""
' CREATE PivotTable for Maintenance Burden Action Data
'"""""""""""""""""""""""""""
PivotName = "MaintRawPivot"

Sheets("pivot").Cells.Clear

'Define source data for the pivot table.
ActiveWorkbook.PivotCaches.Create _
(SourceType:=xlDatabase, SourceData:= _
sTableName, Version:=xlPivotTableVersion14).CreatePivotTable _
TableDestination:="Pivot!R1C1", TableName:=PivotName, _
DefaultVersion:=xlPivotTableVersion14
Sheets("Pivot").Activate
Cells(1, 1).Select
ActiveWorkbook.ShowPivotTableFieldList = True


With ActiveSheet.PivotTables(PivotName)

With .PivotFields("ACTIVITY_NAME")
.Orientation = xlRowField
.Position = 1
End With

With .PivotFields("TOTAL_COST")
.Orientation = xlDataField
.Position = 1
End With

With .PivotFields("AMOUNT")
.Orientation = xlDataField
.Position = 2
End With

End With



'Format the Pivot Table
With ActiveSheet.PivotTables(PivotName)
.ColumnGrand = False
```

```
.RowGrand = False
.RowAxisLayout xlTabularRow
.PivotFields("ACTIVITY_NAME").AutoSort _
xlDescending, "sum of total_cost"
End With


'"""""""""""""""""""""""""""
' COPY PivotTable Maintenance Burden data
'"""""""""""""""""""""""""""
With Sheets("TotalCostTable")
.Cells.Clear
.Activate
Sheets("Pivot").UsedRange.Copy
.Cells(1, 1).PasteSpecial Paste:=xlPasteValues

LastRow = fFindLastRow
lastCol = fFindLastColumn

'Set Title

Cells(1, lastCol + 1) = "Average Cost Per Unit"
fCreateTable ("TotalCostTable")
Cells(2, lastCol + 1).Formula = _
"=iferror([@[Sum of TOTAL_COST]] / [@[Count of AMOUNT]],0)"

End With

With Sheets("AverageCostTable")
.Cells.Clear
.Activate
Sheets("TotalCostTable").UsedRange.Copy
.Cells(1, 1).PasteSpecial Paste:=xlPasteValues

.UsedRange.Sort _
key1:=Cells(1, 4), _
order1:=xlDescending, _
Header:=xlYes
fCreateTable ("AverageCostTable")
End With

'"""""""""""""""""""""""""""
' CREATE PivotTable for Individual Bridge Burden Total Cost
'"""""""""""""""""""""""""""

PivotName = "BridgeBurdenTotCostTable"
wsDest = "BridgeBurdenTotCost"
Sheets(wsDest).Cells.Clear

'Define source data for the pivot table.
ActiveWorkbook.PivotCaches.Create(SourceType:=xlDatabase, SourceData:= _
sTableName, Version:=xlPivotTableVersion14).CreatePivotTable _
TableDestination:=wsDest & "!R1C1", TableName:=PivotName, _
DefaultVersion:=xlPivotTableVersion14
Sheets("BridgeBurdenTotCost").Activate
Cells(1, 1).Select
ActiveWorkbook.ShowPivotTableFieldList = True


With ActiveSheet.PivotTables(PivotName)

With .PivotFields("Structure ID")
.Orientation = xlRowField
.Position = 1
End With

With .PivotFields("TOTAL_COST")
.Orientation = xlDataField
```

```
.Position = 1
End With

'Include cost for each burden classification.
With .PivotFields("Action Classification")
.Orientation = xlColumnField
.Position = 1
End With
End With

'Format the Pivot Table
With ActiveSheet.PivotTables(PivotName)
.ColumnGrand = False
.RowGrand = True
.GrandTotalName = "Total Burden Cost"
.RowAxisLayout xlTabularRow
.PivotFields("structure ID").AutoSort _
xlAscending, "structure ID"
End With

'""""""""""""""""""""""""""
' CREATE PivotTable for Individual Bridge Burden REOCCURING Cost
'""""""""""""""""""""""""""


PivotName = "BridgeBurdenReocCostTable"
wsDest = "BridgeBurdenReocCost"
Sheets(wsDest).Cells.Clear

'Define source data for the pivot table.
ActiveWorkbook.PivotCaches.Create(SourceType:=xlDatabase, SourceData:= _
sTableName, Version:=xlPivotTableVersion14).CreatePivotTable _
TableDestination:=wsDest & "!R1C1", TableName:=PivotName, _
DefaultVersion:=xlPivotTableVersion14
Sheets(wsDest).Activate
'Cells(1, 1).Select
ActiveWorkbook.ShowPivotTableFieldList = True
With ActiveSheet.PivotTables(PivotName)

With .PivotFields("Structure ID")
.Orientation = xlRowField
.Position = 1
End With

With .PivotFields("Activity_Name")
.Orientation = xlRowField
.Position = 2
End With

With .PivotFields("Activity_Name")
.Orientation = xlDataField
.Position = 1
.Name = "Count of Activity Name"
.Function = xlCount
End With

With .PivotFields("TOTAL_COST")
.Orientation = xlDataField
.Name = "Reoccuring Cost"
.Position = 2
End With
'Remove any non reoccuring actions
.PivotFields("ACTIVITY_NAME").PivotFilters.Add _
Type:=xlValueDoesNotEqual, _
DataField:=.PivotFields("Count of Activity Name"), _
Value1:=1
.PivotFields("Structure ID").ShowDetail = False
```

```
With .PivotFields("Action Classification")
.Orientation = xlColumnLabels
.Position = 2
End With
End With

'""""""""""""""""""""""""""""
' Extract data from the maintenance burden total cost
' to the bridges sheet
'""""""""""""""""""""""""""""

'Identify the attributes that are wanted from the local dataset.
Sheets("bridges").Activate


sTarget = "Recommended Maintenance Total"
nCol = fFindCol(sTarget) + 1
nLrow = fFindLastRow
nlcol = fFindLastColumn

Set rDel = Range(Cells(1, nCol), Cells(nLrow, nlcol))
rDel.Clear

Application.DisplayAlerts = True

Application.DisplayAlerts = False

'Define the global key for indexing as the structure ID's.
Set globalkey = Range(Cells(2, 7), Cells(fFindLastRow, 7))

Set destRange = Range(Cells(2, fFindLastColumn2 + 1), _
Cells(fFindLastRow, fFindLastColumn2 + 1))
Set colTitle = Cells(1, fFindLastColumn + 1)

Sheets("BridgeBurdenTotCost").Activate


Set localKey = Sheets("BridgeBurdenTotCost").UsedRange
Set indexRange = localKey.Offset(0, 1)

Sheets("bridges").Activate


'Identify the structure numbers that are in the global dataset.
'Determine if there is a match between the structure numbers
'between the global and local dataset.
'Return the local matching row number.
'Create an array to collect returned data.
'Return data from given columns and matching row number to array.
'If no match, then print "0".
'print array to spreadsheet.

For y = 2 To 6
Set colTitle = Cells(1, fFindLastColumn + 1)

destRange.Value = WorksheetFunction.VLookup(globalkey, _
localKey, y, False)

If y <> 6 Then
colTitle.Value = localKey.Cells(2, y) & " Total"
Else
colTitle.Value = localKey.Cells(2, y)
End If

destRange.Replace "#N/A", "0"
destRange.Replace "", "0"

Set destRange = destRange.Offset(0, 1)
```

```
Next y

Sheets("BridgeBurdenReocCost").Activate

Set localKey = Sheets("BridgeBurdenReocCost").UsedRange

Sheets("bridges").Activate

For y = 6 To 11
Set colTitle = Cells(1, fFindLastColumn + 1)
colTitle.Value = "Reoccurring " & localKey.Cells(3, y)
If colTitle.Value = "Reoccurring " Then
colTitle.Value = localKey.Cells(2, y)
End If

destRange.Value = WorksheetFunction.VLookup(globalkey, _
localKey, y, False)
destRange.Replace "#N/A", "0"
destRange.Replace "", "0"

Set destRange = destRange.Offset(0, 1)

Next y

Application.DisplayAlerts = True

Debug.Print "Results: There are " & Sheets("maintRaw") _
.UsedRange.Rows.Count _
; " maintenance burden actions remaining in the raw data."

'Import the Bridge Posting, Fracture Critical,
'and Percent ADTT from the NBI database to the common database.

Application.DisplayAlerts = False

'Define attributes to import from NBI.
Dim coll As New Collection
coll.Add "Bridge Posting"
coll.Add "Fracture Critical"
coll.Add "Percent ADTT"

'Define parameters
With Sheets("bridges")
.Activate
nLastRow = fFindLastRow
nDestCol = fFindLastColumn + 1
End With

Set localKey = Sheets("NBIval").UsedRange

'start 'y' loop
For y = 1 To coll.Count
Set colTitle = Cells(1, nDestCol)
colTitle.Value = coll(y)

Sheets("NBIval").Activate
Z = fFindCol(coll(y))


Sheets("Bridges").Activate
With destRange
.Value = WorksheetFunction.VLookup _
(globalkey, localKey, Z, False)
.Replace "#N/A", "0"
.Replace "", "0"
End With

Set destRange = destRange.Offset(0, 1)
```

```
nDestCol = nDestCol + 1

'End 'y' loop
Next y

Application.DisplayAlerts = True

Sheets("Bridges").Activate
destCol = fFindLastColumn + 1
Cells(1, destCol).Value = "ADTT"

Set rDest = Range(Cells(2, destCol), Cells(fFindLastRow, destCol))

nADT = fFindCol("ADT")
nPADTT = fFindCol("Percent ADTT")

For Each cell In rDest

cell.Value = Cells(cell.Row, nADT) * Cells(cell.Row, nPADTT) / 100


Next cell

'Calculate the Truck Volume/Capacity ratio.
Sheets("Bridges").Activate
destCol = fFindLastColumn + 1
Cells(1, destCol).Value = "Truck Volume/Capacity"

Set rDest = Range(Cells(2, destCol), _
Cells(fFindLastRow, destCol))

nSpans = fFindCol("Through Lanes On")
nADTT = fFindCol("ADTT")

For Each cell In rDest
If Cells(cell.Row, nSpans).Value = 0 Then
cell.Value = 0
Else
cell.Value = Cells(cell.Row, nADTT) _
/ Cells(cell.Row, nSpans)
End If

Next cell




''\\****************************************
''\\ Produce the ECDFs
''\\****************************************

With Sheets("ECDFtargets")
.Activate
rAttributelist = ActiveSheet.ListObjects("attributeECDFrequest") _
.ListColumns(1).DataBodyRange
End With

'Delete previous value function data.
Sheets("valuefcns").Cells.Delete
Sheets("valuePivot").Cells.Delete

For Each cell In rAttributelist
sTarget = cell

With Sheets("Bridges")
.Activate
LastRow = fFindLastRow
nTargetCol = fFindCol(sTarget)
```

```
Set rTargetData = .Range(Cells(2, nTargetCol), _
Cells(LastRow, nTargetCol))
End With

With Sheets("Valuefcns")
.Activate
lastCol = fFindLastColumn + 1
Cells(1, lastCol).Value = sTarget
Cells(1, lastCol + 1).Value = sTarget & " Value Function"

Set rDest = .Range(Cells(2, lastCol), Cells(LastRow, lastCol))
rDest.Value = rTargetData.Value
End With
rDest.Sort key1:=Cells(2, lastCol), order1:=xlAscending, Header:=xlNo
'Remove N assuming that N does not equal 0.
lastRow2 = WorksheetFunction.CountA(rDest)
Set rDest = rDest.Offset(, 1)
'"""""""""""
'vLookUp if the attribute needs to be inverted
'"""""""""""
Sheets("ECDFtargets").Activate
'set local key as the data related to the attribute ECDF request table.
Set localKey = ActiveSheet.ListObjects("attributeECDFrequest").DataBodyRange
Sheets("bridges").Activate
sECDFopt = WorksheetFunction.VLookup(sTarget, localKey, 2, False)
Sheets("valuefcns").Activate
If sECDFopt = "Invert" Then
'Inverted ECDF loop
For i = 2 To lastRow2 + 1
Sheets("ValueFcns").Cells(i, lastCol + 1).Value = 100 - ((i - 1) / (lastRow2) *
    100)
Next i
Else
'Regular ECDF loop
For i = 2 To lastRow2 + 1
Sheets("ValueFcns").Cells(i, lastCol + 1).Value = (i - 1) / (lastRow2) * 100
Next i
End If
Next cell
''''''''''''
'Make tables for each attribute utility function
'This is for the ValuePivot tab.

'Counter for table names
q = 0
r = -2


' i is for the spacing of the tables

'Start 'i' loop.
For i = 2 To fFindLastColumn Step 2

'Create Pivot table
'r is the spacing between pivot tables in "valuePivot".
q = q + 1
r = r + 4

Sheets("ValueFcns").Activate
'identify attribute location.
'Select Range to make table
'Make table
Set rAttribute = Cells(1, i)
Set rVal = Cells(1, i + 1)
Set rUtility = Range(rAttribute, rVal)
Set rUtility = Range(rUtility, rUtility.End(xlDown))
Set rDest = Sheets("ValuePivot").Cells(1, r)
```

```
ActiveSheet.ListObjects.Add(xlSrcRange, rUtility, , xlYes) _
.Name = "ValTable" & q

sTableSource = "ValTable" & q
sPivotName = "ValPivot" & q

Sheets("ValuePivot").Activate

ActiveWorkbook.PivotCaches.Create(SourceType:=xlDatabase, SourceData:= _
sTableSource, Version:=6).CreatePivotTable TableDestination:=rDest _
, TableName:=sPivotName, DefaultVersion:=6


'Adjust Settings for pivot Table


With ActiveSheet.PivotTables(sPivotName)
With .PivotFields(rAttribute.Value)
.Orientation = xlRowField
.Position = 1
End With

.AddDataField ActiveSheet.PivotTables( _
sPivotName).PivotFields( _
rVal.Value), _
"Average Utility Value", xlAverage
.ColumnGrand = False
.RowGrand = False
End With


ActiveSheet.PivotTables(sPivotName).CompactLayoutRowHeader _
= rAttribute.Value

'Calculate the Scaled Value for each value function.
LastCol1 = fFindLastColumn
Cells(2, LastCol1).Select
Range(Selection, Selection.End(xlDown)).Select
nLastRow = Selection.Count + 1
A = WorksheetFunction.Min(Range(Cells(2, LastCol1), _
Cells(nLastRow, LastCol1)))
B = WorksheetFunction.Max(Range(Cells(2, LastCol1), _
Cells(nLastRow, LastCol1)))
Cells(1, LastCol1 + 1).Value = "Scaled Value"

'Input the scaled calculation.
For u = 2 To nLastRow
Cells(u, LastCol1 + 1).Value = _
(Cells(u, LastCol1) - A) / (B - A) * 100
Next u

'End 'i' loop
Next i



'Put attribute names in the common database.

Sheets("ecdfTargets").Activate
Sheets("ecdfTargets").Cells(2, 4).Select

Set rNames = Range(Selection, Selection.End(xlDown))
rNames.Select
Selection.Copy

Sheets("bridges").Activate
Cells(1, MaintBurdenModule.fFindLastColumn + 1).Select
Selection.PasteSpecial Paste:=xlPasteValues, _
```

```
Operation:=xlNone, SkipBlanks:= _
False, transpose:=True

Application.CutCopyMode = False

'names of attributes for ECDF creation.
Sheets("ecdfTargets").Activate
Sheets("ecdfTargets").Cells(2, 1).Select

actualnames = Range(Selection, Selection.End(xlDown))

'""""""""""""
'vLookUp the value for each attribute
'""""""""""""

Sheets("bridges").Activate
nBridgeLastRow = fFindLastRow


For i = 1 To UBound(actualnames)


'define common database references.
Sheets("bridges").Activate

nTargetCol = fFindCol(actualnames(i, 1))

Set globalkey = Range(Cells(2, nTargetCol), _
Cells(nBridgeLastRow, nTargetCol))

'destRange is the first unused column.
Set destRange = Range(Cells(2, fFindLastColumn2 + 1), _
Cells(fFindLastRow, fFindLastColumn2 + 1))


'define local database for indexing.
Sheets("valuepivot").Activate

sTarget = actualnames(i, 1)
nTargetCol = fFindCol(sTarget)
Range(Cells(2, nTargetCol), Cells(2, nTargetCol + 2)).Select

Set localKey = Range(Selection, Selection.End(xlDown))
Sheets("bridges").Activate

destRange.Value = WorksheetFunction.VLookup _
(globalkey, localKey, 3, False)



Next i




End Sub
Sub mImportFile(PasteStart)
'""""""""""""""""""""""""
' Imports a data sheet that the user chooses to given worksheet.
'""""""""""""""""""""""""
'Define workbook with data to import. Using the Agile Assets
'network master data.
'If no file is selected, end the program.
'Define the workbook to be imported and opened.
'Increase the paste location to fit the whole array, and paste array.
FileToOpen = Application.GetOpenFilename _
(Title:="Select the file to import", _
FileFilter:=".xls (*.xls*),")
```

```
If FileToOpen = False Then
MsgBox "No file specified", vbExclamation , "ERROR"
Exit Sub

Else

Set wb2 = Workbooks.Open(Filename:=FileToOpen)
arr = wb2.Sheets(1).UsedRange.Value

wb2.Close
PasteStart.Resize(UBound(arr, 1), UBound(arr, 2)).Value = arr

End If

End Sub

Function fFindLastColumn() As Integer
'""""""""""""""""""""""""
' FINDS last column number of used range in active sheet.
'""""""""""""""""""""""""
With ActiveSheet
fFindLastColumn = .Cells(1, .Columns.Count).End(xlToLeft).Column
End With
End Function

Function fFindLastRow() As Double
'""""""""""""""""""""""""
' FINDS last row number of used range in active sheet.
'""""""""""""""""""""""""
With ActiveSheet
If Application.WorksheetFunction.CountA(.Cells) <> 0 Then
fFindLastRow = .Cells.Find(what:=".", _
after:=.Range("a1"), _
lookat:=xlPart, _
LookIn:=xlFormulas, _
searchorder:=xlByRows, _
searchdirection:=xlPrevious, _
MatchCase:=False).Row
Else
fFindLastRow = 1
End If
End With
End Function

Function fFindCol(sTarget) As Integer
'""""""""""""""""""""""""
' FINDS column number of target string item.
'""""""""""""""""""""""""

With ActiveSheet
Set origin = .Cells(1, 1)
LastColumn = fFindLastColumn
fFindCol = .Range(origin, Cells(1, LastColumn)).Find _
(sTarget, lookat:=xlWhole).Column
End With

End Function


Function fFindLastColumn2() As Integer
'""""""""""""""""""""""""
' FINDS last column number of used range in active sheet.
' Looks at second column. Quick fix for final data collection
'""""""""""""""""""""""""
With ActiveSheet
fFindLastColumn2 = .Cells(2, .Columns.Count).End(xlToLeft).Column
```

```
End With
End Function
Function fCreateTable(sTableName As String)
'Monday, 2/13/2017 ANA
'EDIT Tuesday, 3/14/2017 ANA
'Creates a table with all range data on active sheet.
'The sTableName will be the name of the table for future reference.
Dim rData As Range


With ActiveSheet
Set rData = .UsedRange
.ListObjects.Add(xlSrcRange, rData, , xlYes).Name = sTableName
End With



End Function

Sub MakeOneColumn()
'Convert the different columns of the bridges that are currently
'being replaced into one column.

Dim vaCells As Variant
Dim vOutput() As Variant
Dim i As Long, j As Long
Dim lRow As Long

If TypeName(Selection) = "Range" Then
If Selection.Count > 1 Then
If Selection.Count <= Selection.Parent.Rows.Count Then
vaCells = Selection.Value

ReDim vOutput(1 To UBound(vaCells, 1) * _
UBound(vaCells, 2), 1 To 1)

For j = LBound(vaCells, 2) To UBound(vaCells, 2)
For i = LBound(vaCells, 1) To UBound(vaCells, 1)
If Len(vaCells(i, j)) > 0 Then
lRow = lRow + 1
vOutput(lRow, 1) = vaCells(i, j)
End If
Next i
Next j

Selection.ClearContents
Selection.Cells(1).Resize(lRow).Value = vOutput
End If
End If
End If

End Sub
```