Report No. K-TRAN: KU-22-2 ▪ FINAL REPORT ▪ January 2026

Development of Adjustment Factors and Load Ratings via Statistical Analyses of the National Bridge Inventory Database

Isaac McVey
Admin Husic, Ph.D.
William Collins, Ph.D., P.E.
Rémy Lequesne, Ph.D., P.E.
Joshua Roundy, Ph.D., P.E.

The University of Kansas

Kansas Department of Transportation

1 Report No. K-TRAN: KU-22-2	**2 Government Accession No.**	**3 Recipient Catalog No.**
4 Title and Subtitle Development of Adjustment Factors and Load Ratings via Statistical Analyses of the National Bridge Inventory Database		**5 Report Date** January 2026 **6 Performing Organization Code**
7 Author(s) Isaac McVey; Admin Husic, Ph.D.; William Collins, Ph.D., P.E.; Rémy Lequesne, Ph.D., P.E.; and Joshua Roundy, Ph.D., P.E.		**8 Performing Organization Report No.**
9 Performing Organization Name and Address The University of Kansas Department of Civil, Environmental & Architectural Engineering 1530 West 15th St. Lawrence, Kansas 66045-7609		**10 Work Unit No. (TRAIS)** **11 Contract or Grant No.** C2190
12 Sponsoring Agency Name and Address Kansas Department of Transportation Bureau of Research 2300 SW Van Buren Topeka, Kansas 66611-1195		**13 Type of Report and Period Covered** Final Report July 2021 – February 2024 **14 Sponsoring Agency Code** RE-0830-01
15 Supplementary Notes For more information write to address in block 9.		

16 Abstract

This study developed a new approach to establish baseline load ratings for bridges in Kansas without plans using data from the National Bridge Inventory (NBI). The approach is comprised of linear regression models to estimate load ratings for bridges with a condition rating of 8 or higher and adjustment factors to lower the estimated load rating to account for bridge condition ratings of 7 or lower. This approach beneficially establishes baseline load rating estimates for structures without prior ratings and secondary load ratings for bridges with prior load ratings to identify outliers and potential errors. The adjustment factors can be used to adjust load ratings obtained by any method to account for bridge condition if the condition was not specifically integrated into the analyses. Both the linear regression models and condition adjustment factors are designed to reflect trends among Kansas bridges within the NBI, not engineering judgment. This approach answers the following question for a given bridge: Knowing nothing more about the structure than what is available within the NBI, what is the expected rating based on similar bridges in similar condition within Kansas?

The proposed linear regression models include bridge age, modeled design load, structure kind (construction material), structure type (truss, girder, etc.) and deck width because, among variables reported in the NBI, these were most closely correlated with load rating. The adjustment factors were developed based on the median reported load rating for bridges with various condition ratings, and uncertainty was estimated using a bootstrapping simulation. The proposed models demonstrated satisfactory performance, capturing approximately half the variance observed in the data for the Inventory ($R^2 = 0.50$) and Operating ($R^2 = 0.49$) Ratings. Further validation and refinement, inclusion of additional predictors, and exploration of alternative methods are suggested to improve accuracy and applicability.

17 Key Words Bridges, Bridge performance, Adjustment factors, Structural health monitoring, Serviceability		**18 Distribution Statement** No restrictions. This document is available to the public through the National Technical Information Service www.ntis.gov.	
19 Security Classification (of this report) Unclassified	**20 Security Classification (of this page)** Unclassified	**21 No. of pages** 65	**22 Price**

Form DOT F 1700.7 (8-72)

This page intentionally left blank.

Development of Adjustment Factors and Load Ratings via Statistical Analyses of the National Bridge Inventory Database

Final Report

Prepared by

Isaac McVey
Admin Husic, Ph.D.
William Collins, Ph.D., P.E.
Rémy Lequesne, Ph.D., P.E.
Joshua Roundy, Ph.D., P.E.

University of Kansas

A Report on Research Sponsored by

THE KANSAS DEPARTMENT OF TRANSPORTATION
TOPEKA, KANSAS

and

UNIVERSITY OF KANSAS
LAWRENCE, KANSAS

January 2026

PREFACE

The Kansas Department of Transportation's (KDOT) Kansas Transportation Research and New-Developments (K-TRAN) Research Program funded this research project. It is an ongoing, cooperative and comprehensive research program addressing transportation needs of the state of Kansas utilizing academic and research resources from KDOT, Kansas State University and the University of Kansas. Transportation professionals in KDOT and the universities jointly develop the projects included in the research program.

NOTICE

DISCLAIMER

Abstract

This study developed a new approach to establish baseline load ratings for bridges in Kansas without plans using data from the National Bridge Inventory (NBI). The approach is comprised of linear regression models to estimate load ratings for bridges with a condition rating of 8 or higher and adjustment factors to lower the estimated load rating to account for bridge condition ratings of 7 or lower. This approach beneficially establishes baseline load rating estimates for structures without prior ratings and secondary load ratings for bridges with prior load ratings to identify outliers and potential errors. The adjustment factors can be used to adjust load ratings obtained by any method to account for bridge condition if the condition was not specifically integrated into the analyses. Both the linear regression models and condition adjustment factors are designed to reflect trends among Kansas bridges within the NBI, not engineering judgment. This approach answers the following question for a given bridge: Knowing nothing more about the structure than what is available within the NBI, what is the *expected* rating based on similar bridges in similar condition within Kansas?

The proposed linear regression models include bridge age, modeled design load, structure kind (construction material), structure type (truss, girder, etc.) and deck width because, among variables reported in the NBI, these were most closely correlated with load rating. The adjustment factors were developed based on the median reported load rating for bridges with various condition ratings, and uncertainty was estimated using a bootstrapping simulation. The proposed models demonstrated satisfactory performance, capturing approximately half the variance observed in the data for the Inventory (R^2 = 0.50) and Operating (R^2 = 0.49) Ratings. Further validation and refinement, inclusion of additional predictors, and exploration of alternative methods are suggested to improve accuracy and applicability.

Acknowledgments

Financial support was provided by the Kansas Department of Transportation. The opinions and conclusions presented herein are those of the writers and do not necessarily represent those of the sponsoring organization.

Table of Contents

List of Tables

List of Figures

Chapter 1: Introduction

The Kansas Department of Transportation (KDOT) is in the process of assigning load ratings to approximately 25,000 bridges within the state inventory, nearly 8,000 of which are concrete bridges with no record of design plans. However, assigning a reasonable load rating to a bridge can be challenging and costly when no reinforcement details are available, and the task is further complicated by the need to account for bridge condition, represented by a 0–9 rating scale based on inspectors' observations.

1.1 Problem Statement

A simple tool is needed to produce baseline load rating values for bridges in good or better condition. This tool should help establish expected load rating values for unrated bridges and identify erroneous load rating values that warrant further review. A method to simply account for bridge condition ratings is also needed.

1.2 Objectives

The primary objective of this study was to develop and evaluate a statistical model using the National Bridge Inventory (NBI) database to establish baseline load ratings for bridges in Kansas. A linear regression model was used to identify key variables, such as age and bridge condition, that correlate with recorded inventory and operating load ratings. This method can produce an expected load rating for most concrete bridges in Kansas, establish baseline load ratings for structures without prior ratings, designate adjustment factors to adjust load ratings to account for bridge condition degradation, and validate model predictions for recorded load ratings. This approach answers the following question for a given bridge: Knowing nothing more about the structure than what is available within the NBI, what is the expected rating based on similar bridges in similar condition within Kansas?

1.3 Methodology

To establish an expected load rating for bridges without design plans, this study constructed a workflow to describe the applied modeling approach (Figure 1.1), starting with a critical examination of NBI data by filtering out all non-applicable structures, such as culverts, and all non-concrete structures and identifying potential outliers and errors in the data to establish a well-constrained dataset. The remaining, applicable concrete structures included all slab, beam, reinforced-concrete, and post-tensioned concrete structures not classified as culverts. Multivariable linear regression was then used to establish a baseline load rating for pristine bridges with condition ratings of 8 or 9 (i.e., very good condition) by identifying key variables that correlated with reported inventory and operating load ratings. Adjustment factors were then developed as a function of superstructure condition rating to reduce the estimated load rating and account for bridge deterioration and distress. Finally, the multivariable linear regression model and adjustment factors were combined to produce an expected load rating for applicable concrete structures based on key variables (e.g., age, span length, structure type) and estimates of condition rating.

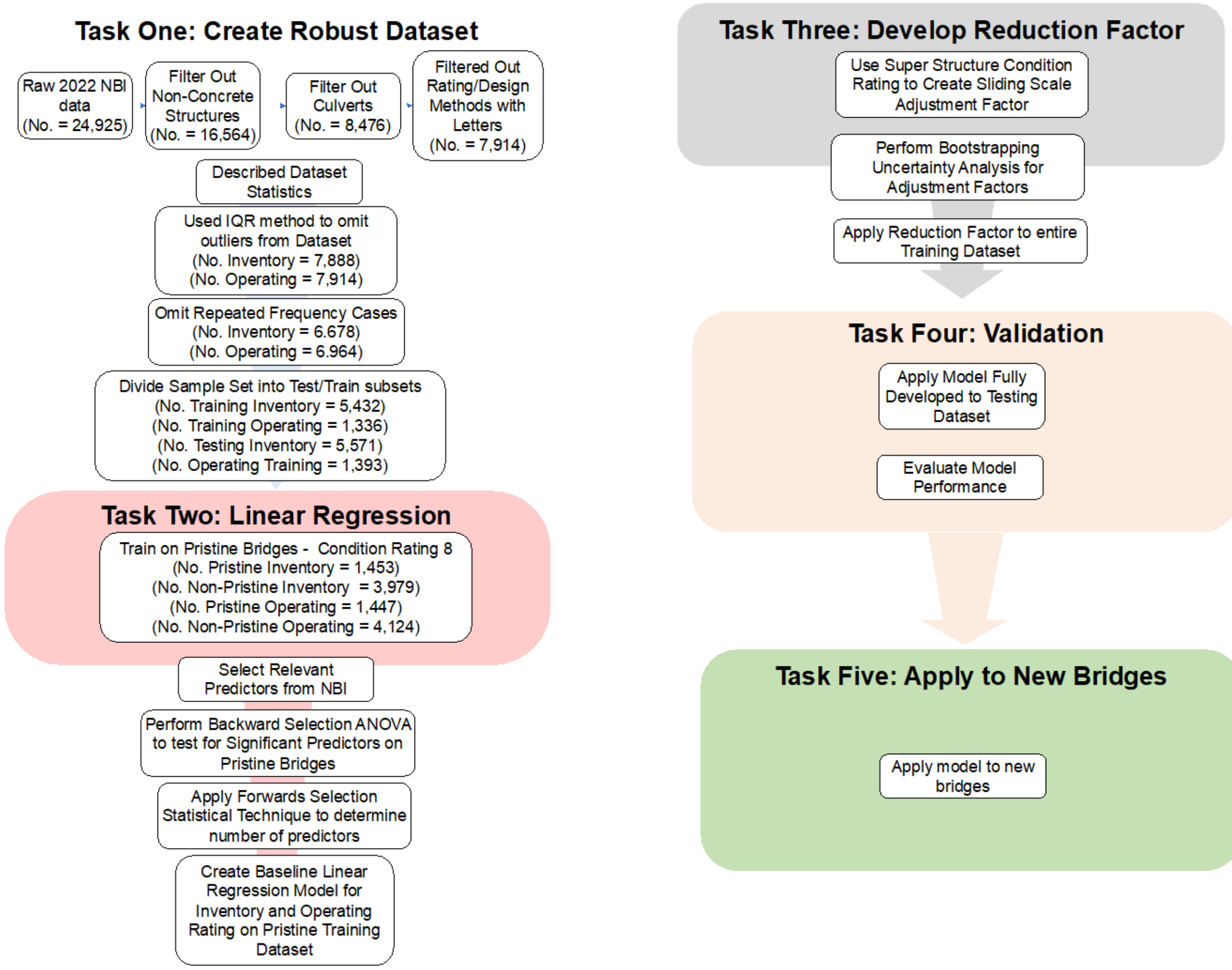

Figure 1.1: Workflow Framework to Develop Reduction Factors and a Linear Regression Model for Concrete Bridges with No Prior Plans

Chapter 2: Design Equations

Load rating is the process of determining the live load carrying capacity of a bridge (Gao, 2013; Ruiz, 2020). The bridge superstructure controls the live load carrying capacity, and the dominant load carrying members are the beams and slabs. Live load rating is computed at the inventory, operating, legal, and permitting levels. This analysis focuses on the two most prominent ratings:

- Inventory rating – the live load that a bridge can withstand for an indefinite amount of time without reducing the structural integrity of the bridge.
- Operating rating – the absolute maximum permissible live load that a bridge can be subjected to for a limited number of occurrences.

Load rating methodology has changed to reflect shifts in design over time. The allowable stress, load factor rating (LFR), and load and resistance factor rating (LRFR) approaches each utilize unique methodology.

2.1 Allowable Stress

The allowable stress method is based on allowable stress design (ASD), which was introduced in the 1930s to ensure that the stress applied load does not exceed the allowable stress of the design material that is assigned a safety factor. Allowable stress results rating factors are presented as:

$$\mathrm{RF} = \frac{F_a - f_d}{f_{(LL+I)}}$$

Equation 2.1

Where:

F_a = allowable stress of the material,

f_d = stress associated with the bridge weight (dead load), and

$f_{(LL+I)}$ = stress caused by live load and dynamic impact.

When determining inventory and operating load ratings, operating stress is 33% higher than inventory allowable stress (AASHTO, 2018; Ruiz, 2020).

2.2 Load Factor Rating Method

The LFR method, introduced in the 1970s, is based on load factor design (LFD). This method utilizes load and resistance factors calibrated by principles of reliability to obtain a more consistent probability of failure than determined by the ASD method. The load rating is reported in rating factor or tonnage (AASHTO, 2018) as:

$$\mathrm{RF} = \frac{\varphi R_n - \gamma_{DL} DL}{\gamma_{LL}(LL + I)}$$

Equation 2.2

Where:

R_n = nominal strength (or structural capacity),

DL = load effect from the dead load on the bridge structure,

$(LL + I)$ = lad effect from live load and dynamic impact,

φ = resistance factor,

γ_{DL} = dead load factor, and

γ_{LL} = live load factor.

When determining inventory and operating ratings, a live load factor of 2.17 is used for the inventory rating and a live load factor of 1.3 is used for the operating rating.

2.3 Load and Resistance Factor Rating

Introduced in the 1990s, the LRFR method, the current method for modern load rating, is based on load and resistance factor design (LRFD). Because LRFR is a relatively recent addition to load rating, fewer LRFR bridges lack design plans compared to bridges rated by other methods. LRFR utilizes load and resistance factors that are calibrated using the structural reliability theory to achieve reliability for strength limit. The LRFR method is presented as a rating factor and is the general equation applied for reinforced concrete bridges.

$$\mathrm{RF} = \frac{\varphi_c \varphi_s \varphi R_n - \gamma_{DC} DC - \gamma_{DW} DW}{\gamma_{LL}(LL + IM)}$$

Equation 2.3

Where:

R_n = nominal strength,

DC = load effect from the dead load of structural components,

DW = load affect from wearing surfaces on the bridge surface,

$LL + IM$ = live load effect caused by live load and dynamic impact,

φ_c = condition factor that accounts for uncertainty with bridge condition,

φ = resistance factor,

γ_{DC} = dead load factor for structural components,

γ_{DW} = dead load factor for wearing surface, and

γ_{LL} = live load factor.

The LRFR inventory and operating ratings are calculated using a γ_{LL} of 1.75 for inventory rating and 1.35 for operating rating (AASHTO, 2018; Ruiz, 2020).

Chapter 3: Statistical Methods

3.1 Task One: Create Robust Dataset

3.1.1 Filtering Criteria

The 2022 Kansas NBI database contains 24,925 bridges, but because the primary area of interest for this analysis was concrete bridge structures, all non-concrete bridge structures were removed from the sample dataset, reducing the sample size from 24,925 bridges to 16,564 bridges. In addition, all the 8,088 culverts within the Kansas bridge inventory of the NBI dataset were removed because they were outside the scope of this project, thereby reducing the sample size from 16,564 bridges to 8,476. Load rating methods were then used to determine inventory and operating load ratings (Tables 3.1 and 3.2). Load rating method 0 indicated field evaluation and documented engineering judgment, rating method 1 showed load factor, method 2 highlighted allowable stress, method 3 showed load and resistant factor, method 4 specified load testing, and method 5 indicated that no rating analysis or evaluation was performed. Load rating method 6 indicated load factor using MS18 loading, while method 7 showed allowable stress using MS18 loading, and method 8 showed load and resistance factor ratings using HL-93 loadings (FHWA, 1995; FHWA, 2011; AASHTO, 2018).

Although rating method D is not listed in the NBI Guide, load rating method C indicates that another load rating method was used to determine the load rating (FHWA, 2011). Therefore, bridges encoded as C or D in the dataset were removed from the sample size since they are not representative of the overall dataset. Pedestrian bridges and railroad bridges were also excluded from the sample dataset, and bridges that had no superstructure condition rating (denoted as N or 0, Table 3.3) were omitted from the training dataset because the adjustment factor is based on superstructure condition rating. Additional omissions further reduced the sample dataset from 8,476 bridges to 7,914 bridges.

Table 3.1: Common Rating Methods to Determine Inventory and Operating Ratings

Load Rating Method (Item 63)	1	2	3	4	5
-	6,461	982	125	168	243

Table 3.2: Updated Load Rating Methods to Determine Inventory and Operating Ratings

Load Rating Method (Item 64)	C	D	0	6	7	8
-	2	1	15	5	0	465

3.1.2 Superstructure Condition Rating Statistics

Figures 3.1, 3.2, and 3.3 depict the superstructure condition ratings (item 59 in NBI) for load factor, allowable stress, and load and resistance factor, respectively. As shown in the condition ratings and descriptions of overall structural integrity in Table 3.3, allowable stress is the load rating method that includes bridges with the lowest mean condition rating, with an average of 6.43 ± 0.96. Comparatively, the LFR method has an average condition rating of 6.96 ± 0.51. As shown in the table, the LRFR method has the highest condition rating, with an average condition rating of 7.76 ± 0.51. No bridges in Kansas have condition ratings of 9 or 1, and most have condition ratings of 7 or 8, indicating that most bridges in Kansas demonstrate satisfactory to very good condition.

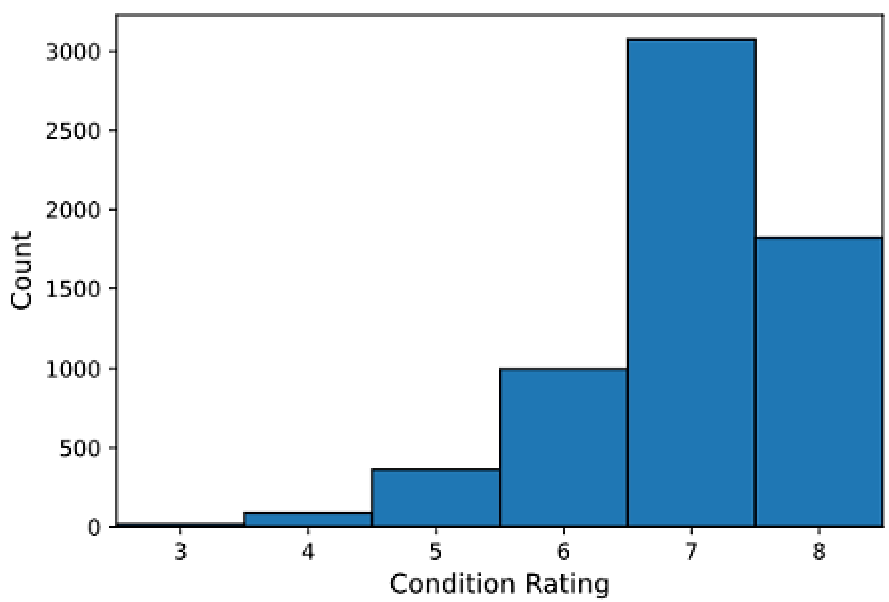

Figure 3.1: LFR Method Condition Rating Distribution

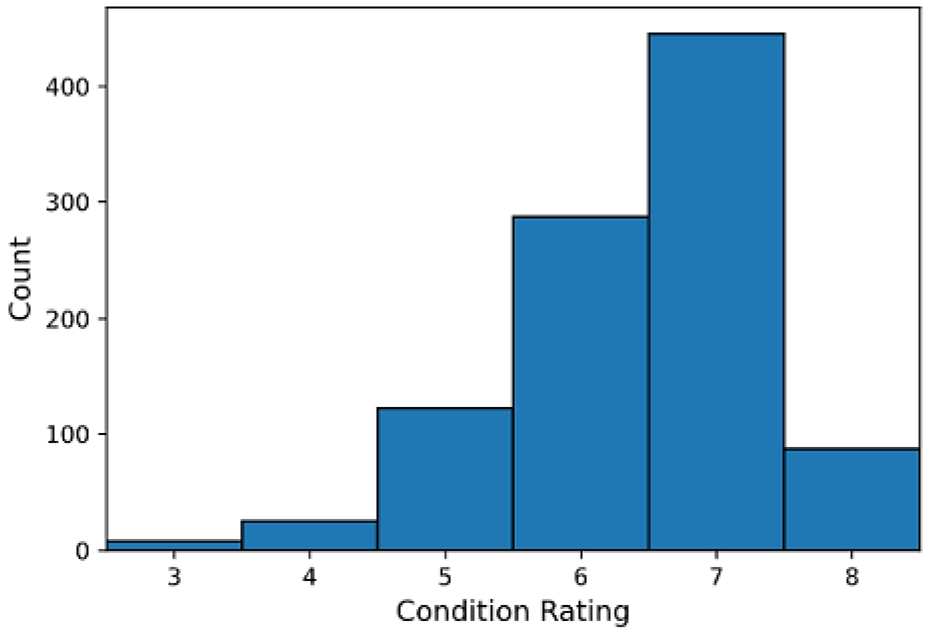

Figure 3.2: Allowable Stress Load Rating Method Condition Rating Distribution

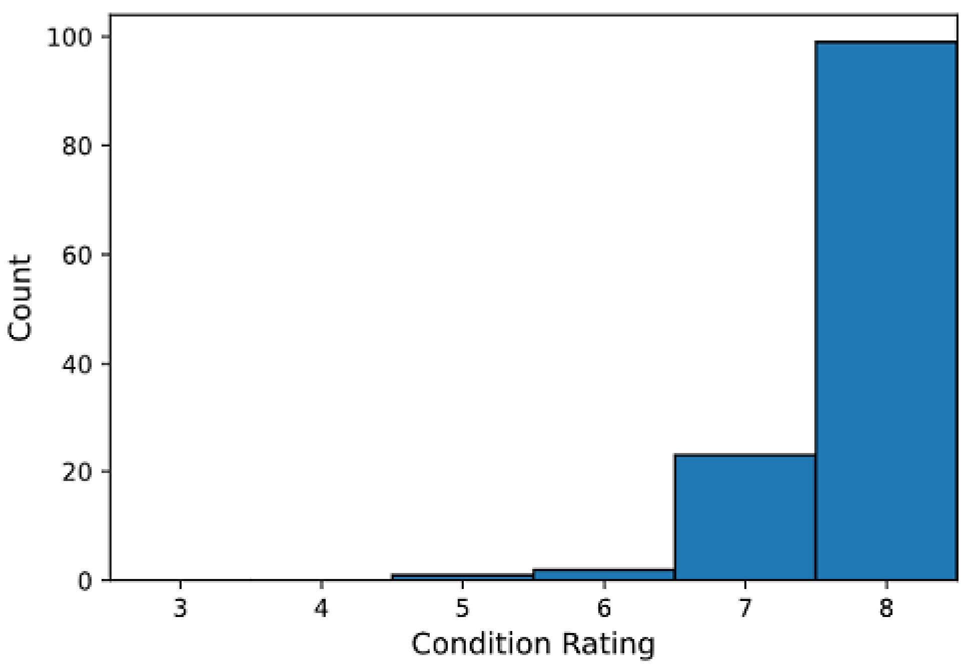

Figure 3.3: LRFR Method Condition Distribution

Table 3.3: Superstructure Condition and Structural Condition Ratings

Superstructure Condition Rating (item 59)	Condition Rating
N	Not Applicable
9	Excellent Condition
8	Very Good Condition – no problems
7	Good Condition – some minor issues
6	Satisfactory Condition – some deterioration
5	Fair Condition – all primary structural elements are sound but may have structural issues
4	Poor Condition – advanced section loss
3	Serious Condition – loss of section, structural integrity is compromised
2	Critical Condition – advanced structural integrity of primary structure
1	Imminent Failure – major deterioration
0	Failed Condition – bridge is out of service

3.1.3 Inventory and Operating Rating Dataset Statistics

The reported inventory and operating load ratings are depicted in Figures 3.4 and Figure 3.5. As shown, the inventory load rating has a median of 30.1 metric tons, while the operating rating has a median of 47.3 metric tons. Although both load rating distributions are approximately normal, Figure 3.6 shows that the inventory load rating has outliers above 62.3 metric tons. Outliers may decrease model predictions as they are not representative of the overall dataset.

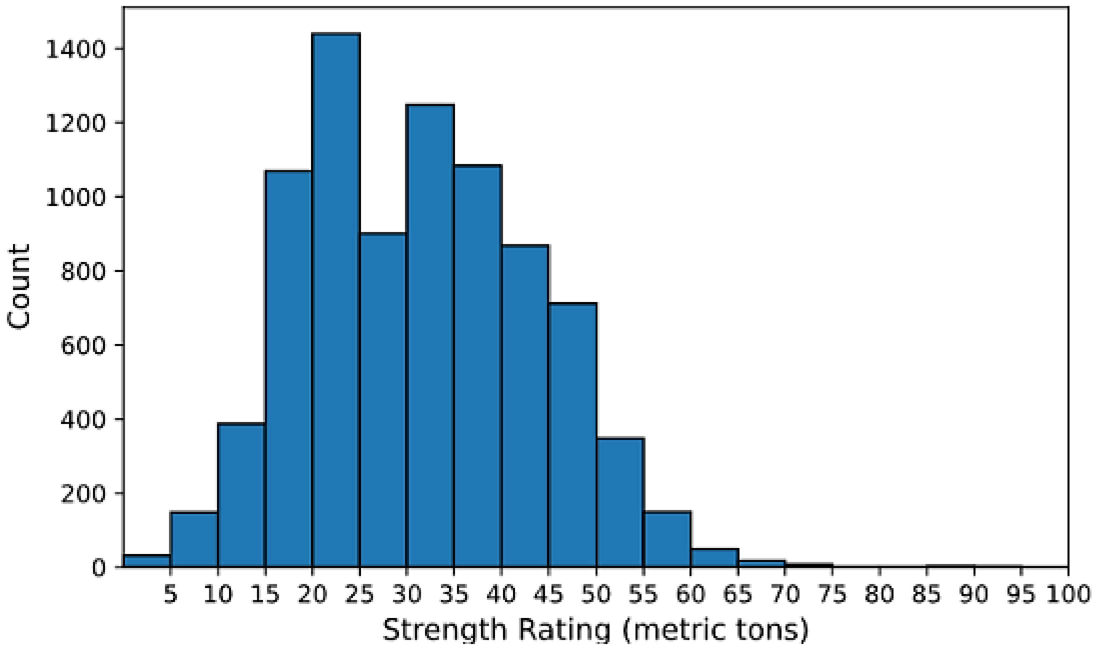

Figure 3.4: Inventory Strength Distribution

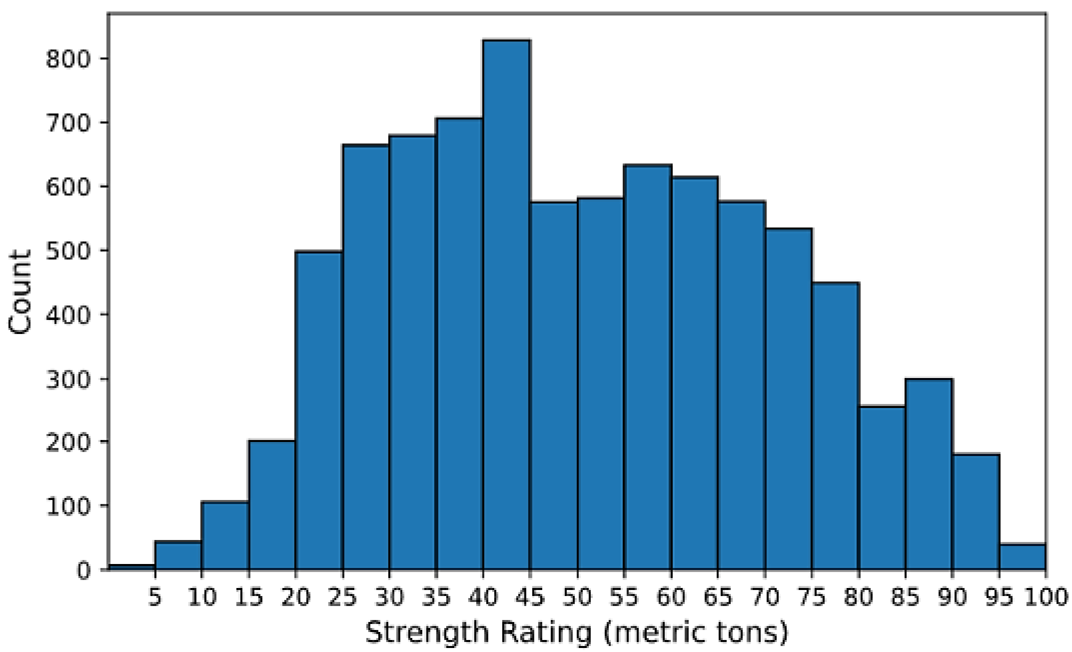

Figure 3.5: Operating Strength Distribution

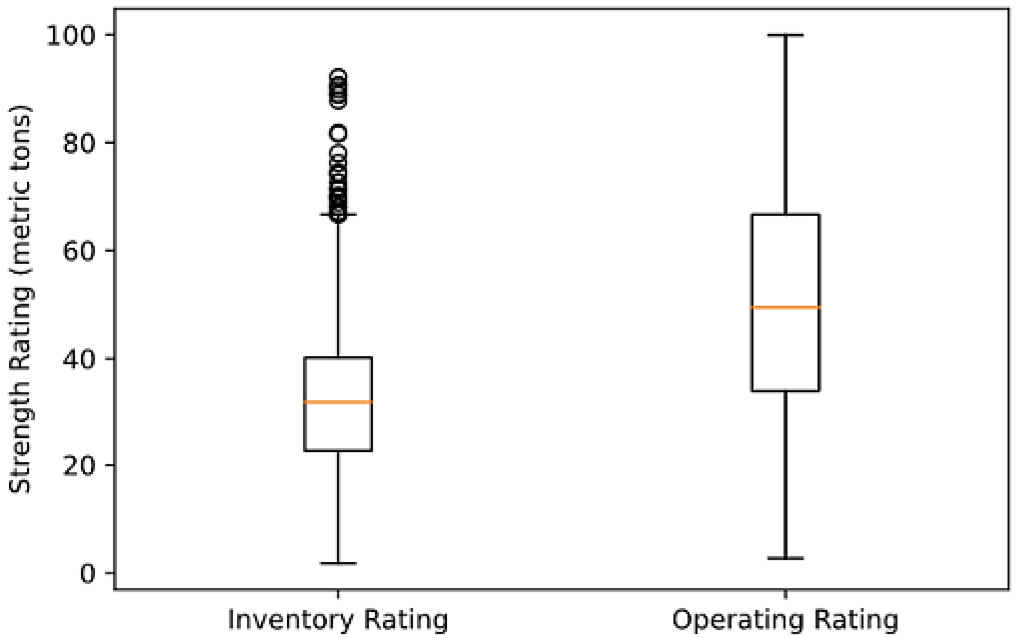

Figure 3.6: Operating versus Inventory Load Rating Distribution

3.1.4 Removal of Outliers and Erroneous Datapoints

Multiple regression models can be used to estimate the value of interest, such as inventory rating, based on known variables. However, because multiple regression models are dependent on the data used to develop them, they are sensitive to outliers and error. Therefore, this study conducted a thorough evaluation to remove potential outliers outside the general range of the dataset. Outliers are datapoints that deviate significantly from most datapoints within a dataset. Their presence can lead to unreliable multiple regression models, potentially altering the estimated mean and standard deviation of the sample dataset (Perez & Tah, 2020). Defining an outlier requires arbitrary selection of a reasonable range of values based on a lower and upper limit. The Interquartile Range (IQR) is often used to define reasonable bounds for the data. IQR is a measure of variability in a dataset, found by dividing the data into an ordered set of elements, or the difference between the 25th percentile and 75th percentile (Perez & Tah, 2020):

$$\text{IQR} = Q_{0.75} - Q_{0.25}$$

Equation 3.1

Where $Q_{0.25}$ is the 25th percentile, and $Q_{0.75}$ is the 75th percentile.

In this work, a point was considered an outlier if it exceeded 1.5 times the IQR with the upper limit and lower limit defined as

$$\text{Lower Limit} = Q_{0.25}\text{-}1.5\cdot\text{IQR}$$

Equation 3.2

$$\text{Upper Limit} = Q_{0.75}+1.5\cdot\text{IQR}$$

Equation 3.3

The application of this method to the inventory and operating rating datasets led to the omission of 26 outliers from the inventory dataset, but no outliers exceeded the upper or lower limits for the operating dataset. After omitting the outliers, 7,888 bridges remained in the inventory rating dataset and 7,915 bridges in the operating dataset. However, the inventory dataset was then only applicable to bridges with a maximum inventory load rating of 70.2 metric tons, which applies to 98.5% of bridges in Kansas.

Prior to the evaluation of independent predictors, certain values of the inventory and operating load ratings had an unusually high frequency of occurrence. For example, bridge owners commonly assign load ratings based on engineering judgment rather than by explicit calculation (Ruiz, 2020). These values, listed from highest to lowest repeating frequency, are 19.8, 24.3, 23.4, 32.4, and 32.7 metric tons for inventory ratings and 24.3, 32.4, 38.7, 40.5, and 90.6 metric tons for operating load ratings. These common values don't necessarily reflect an explicit calculation, but rather an expert judgement thus they may not be predictable by our mathematical model. Therefore, this study considered two cases for how to handle these high-frequency ratings. In the first case, no high frequency loads were assumed to be correct, (i.e., load rating entries with an unusually high frequency) were miscoded and have a different value than the one reported. However, this approach disregarded the common practice of using these values for load rating based on engineering judgment and accepted the inclusion of potentially incorrect cases. The second case removed the ten most frequently repeated inventory and operating load ratings and assumed that all the high frequency cases reported for the inventory and operating ratings were miscoded, prompting the removal of possible incorrectly coded cases but potentially removing cases with correct inventory or operating ratings. This study utilized the second case to remove all high frequency load ratings and prevent the removal of miscoded cases to improve the training dataset for model development, resulting in 1,210 repeated values for the inventory rating dataset and 951 repeated values for the operating rating dataset. Removal of all potentially miscoded high frequency load ratings resulted in 6,692 cases for inventory ratings and 6,964 bridges for operating ratings.

3.1.5 Final Training/Testing Dataset

The remaining 6,678 bridges and 6,964 bridges for the inventory and operating rating sample datasets, respectively, were divided so that 80% were designated for the training dataset to train the multivariable linear regression model and 20% were assigned to the testing dataset to confirm model performance (R^2) (van der Goot, 2021). The split resulted in 5,432 training bridges and 1,336 testing bridges for inventory ratings and 5,571 training bridges and 1,393 testing bridges for operating ratings.

3.1.6 Central Tendency and Variability Between Testing/Training Datasets

The development of a linear regression model to estimate a baseline load rating for bridges without design plans first required verification that the training and testing datasets exhibited the same decreasing central tendencies in load ratings with decreasing condition rating factor. As shown in Figures 3.7 and 3.8, both datasets demonstrated decreasing reported load ratings with decreasing condition factor. Therefore, the preliminary linear regression model was developed to establish baseline load ratings for bridges in very good condition.

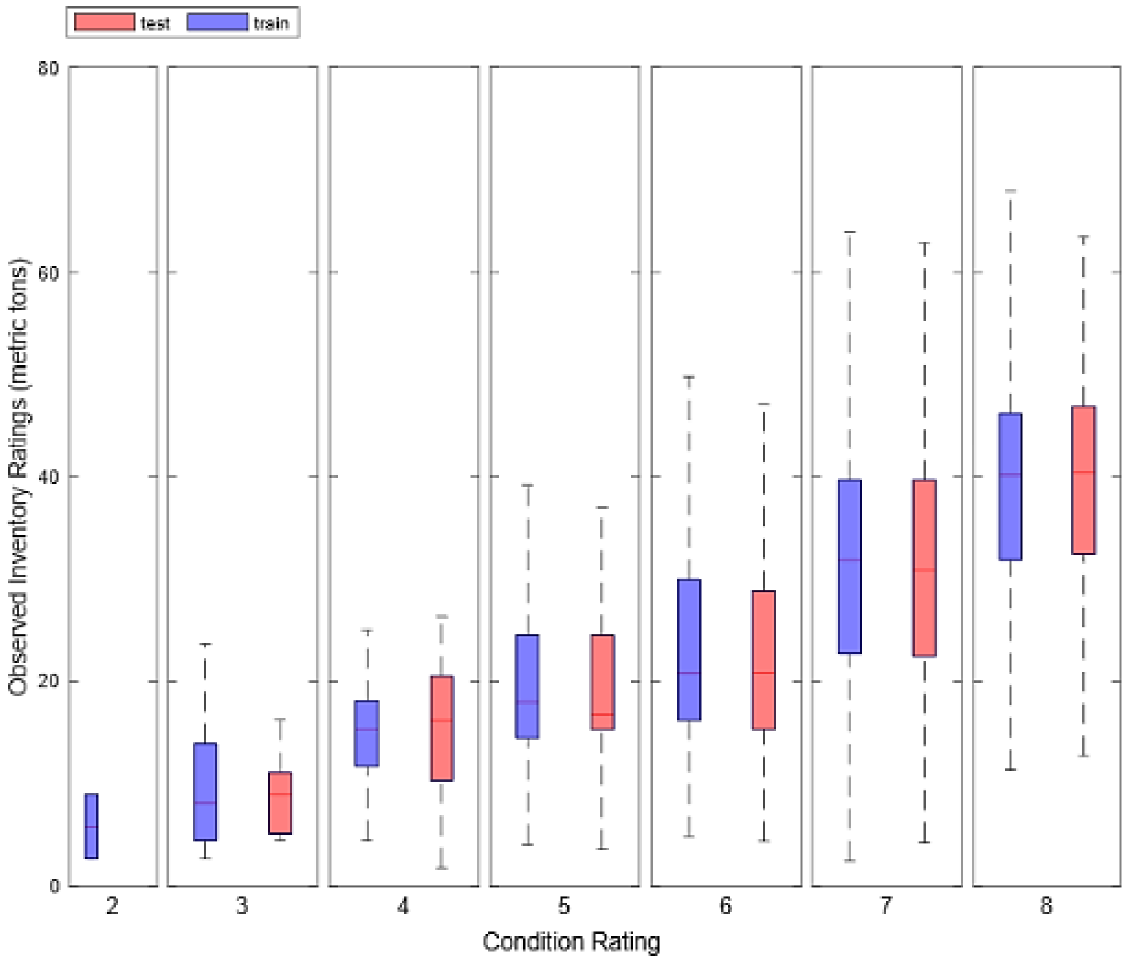

Figure 3.7: Condition Rating for Inventory Training Dataset (N = 5,342) and Testing Dataset (N = 1,336)

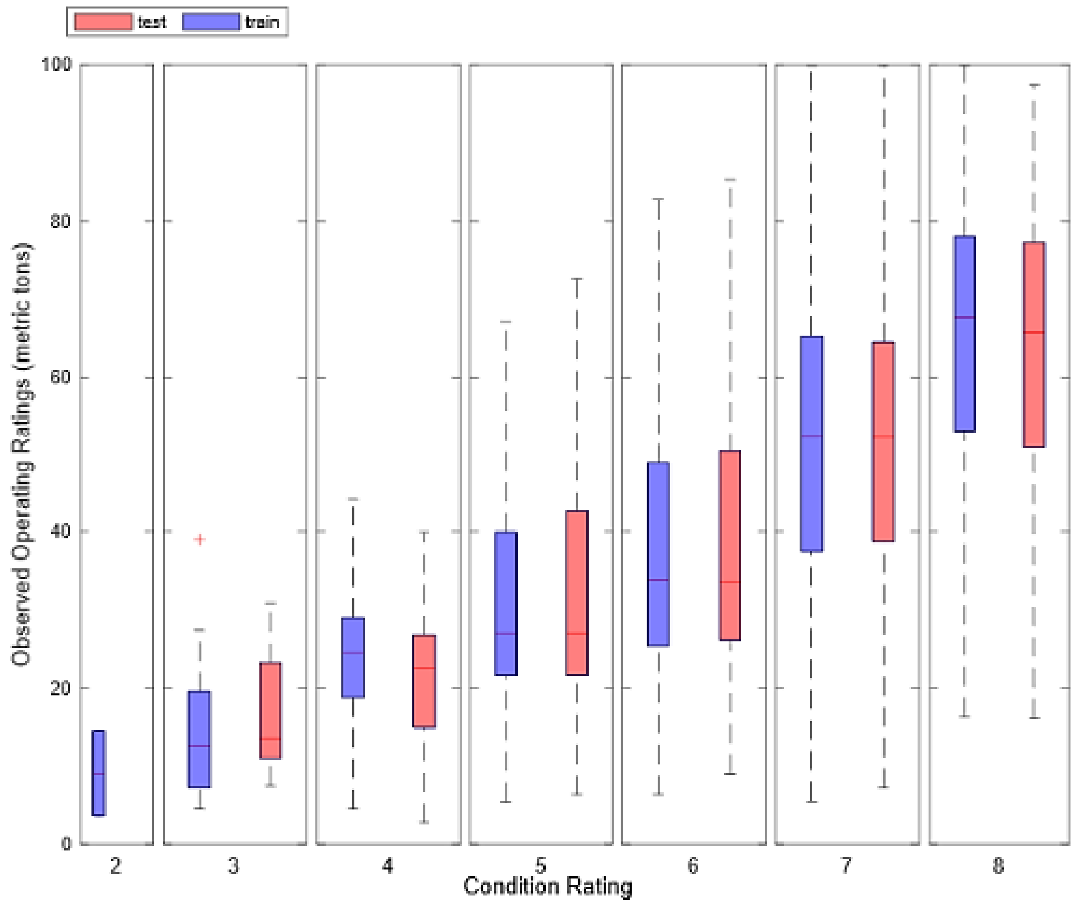

Figure 3.8: Condition Rating for Operating Training Dataset (N = 5,571) and Testing Dataset (N = 1,393)

3.2 Task Two: Develop Multivariable Linear Regression

To establish a baseline load rating for bridges without design plans, this study initially developed a multiple linear regression model for bridges with a superstructure condition rating of 8, meaning no detectable or notable structural deficiencies are present and the bridge is in excellent to superb condition. Consequently, a linear regression model for bridges in pristine condition established a baseline load rating for bridges with no deterioration in structural capacity/condition. This study used 1,453 bridges and 1,442 bridges to train the baseline load rating for inventory and operating ratings, respectively. After applying the 80%/20% training split previously described, 356 testing bridges and 360 bridges were reserved for the inventory and operating baselines, respectively, to evaluate load rating model performance.

3.2.1 Multiple Regression Model

This study also developed a multiple linear regression model to estimate the load rating of concrete bridges. A multiple linear regression model can estimate relationships between one dependent variable and one or more independent variables to create an equation to predict dependent variables based on observed data (Khainge et al., 2019). Development of a linear regression model requires an understanding of the assumptions and uncertainty associated with the model. To apply the multiple linear regression model, the ordinary least squared method was performed, including estimating the coefficients by minimizing the squared errors between the observations and the model. The least squared errors provide an estimation of the expected value of the dependent variable (i.e., inventory or operating rating). However, this predicted value differs from actual observation, a difference known as the residual or error. The resultant residuals can be positive or negative, and errors from well-fit multiple linear regression models are normally distributed, with a constant variance and a mean of zero (Uyanık & Güler, 2013).

3.2.2 Preliminary Predictor Selection

A usable multiple linear regression model must include input data that is readily available. Table 3.4 summarizes the NBI items considered for baseline load rating model development. The definitions of the NBI items are in the *Recording and Coding Guide for the Structure Inventory and Appraisal of the Nation's Bridges* (FHWA, 1995).

Table 3.4: NBI Items Considered as Potential Predictors

NBI Item Number	Qualification
Inventory rating (66)	Dependent variable
Operating rating (64)	Dependent variable
Year built (27)	Material properties, standard design practices
Design load (31)	Live load used for model
Structure kind (43A)	Construction material
Structure type (43B)	Structure type (i.e. truss, girder)
Deck width (52)	Design, load effects
Traffic lanes (28A)	Design, load effects from number of lanes on structure
Max span length (48)	Design, load effects
Main unit spans (45)	Design, load effects
Approach spans (46)	Design, load effects
Minimum lateral under clearance (55B)	Design, load effects
Average daily truck traffic (109)	Design, load effects
Total horizontal clearance (47)	Design, load effects
Left curb width (50A)	Design, load effects
Right curb width (50B)	Design, load effects
Approach width (32)	Design, load effects
Structure length (49)	Design, load effects
Traffic lanes under (42B)	Design, load effects
Degrees skew (34)	Design, load effects
Traffic direction (102)	Design, load effects
Average daily traffic (29)	Design, load effects

3.2.3 Backward Selection Analysis of Variance to Select Significant Predictors

The backwards selection technique of analysis of variance (ANOVA) can be used in multiple linear regression to select significant predictors. The ANOVA test, which is primarily used to find differences between group means, uses a student's t-test to determine whether a predictor is significantly different than the dependent variable (Kim, 2017). The general form of multiple linear regression is

$$\hat{y} = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \ldots \beta_k x_k + \varepsilon_i$$

Equation 3.4

Where:

$\hat{y}$ = dependent variable (inventory or operating rating),

x_i = independent variable,

k = number of independent variables,

β_0 = dependent variable sample mean, and

β_k = difference in the means between the independent variable and dependent variable *k* (Kim, 2017; Rouder et al., 2021; Yu et al., 2022).

When running the test, the ANOVA tests compares if there is a significant difference the dependent variable mean and predictor k, if the random errors, ε_i are independent of each other, normally distributed, and have equal variances. Thereafter, the F test can be used to assess the null hypothesis ($H_0 : \mu_1 = \mu_2 = \mu_3$) that there is no significant difference between the dependent variable sample mean, and predictor *k* sample mean at a 95% confidence interval (Kim, 2017). In the alternative hypothesis (H_A), if the p-value from the student's t-test is less than 0.05, then there is sufficient evidence to reject the null hypothesis and prove there is a significant difference between the predictor *k* and the dependent variable; therefore, the predictor *k* can be used to develop the multiple linear regression model (Kim, 2017; Yu et al., 2022). This study applied ANOVA by iteratively conducting the test with all the predictors and evaluating the p-value for each predictor *k* with the highest p-value, assuming it to be the least significant predictor. The test was repeated until only one predictor remained.

3.2.4 Number of Selected Predictors Used in Linear Regression Equation

Forward selection was used for the chosen number of predictors for the finalized linear regression. Forward selection is a stepwise procedure that includes predictors in the developed model if they improve model predictions for inventory or operating ratings. This procedure continues until no additional predictors improve model performance (Schneider et al., 2010). This study initiated forward selection with one predictor, followed by the coefficient of determination (R^2) and the mean squared error (MSE) (Equations 3.5 and 3.7) to assess model improvements. This process was iterated until the MSE was minimized for the testing dataset since the addition of variables always improves the R^2 for the training dataset. However, as the coefficient of determination improves, the MSE changes based on predictor combinations. The number of selected predictors for the developed model is complete when the MSE is minimized within the testing pristine bridge dataset.

$$\mathrm{MSE} = \frac{SSE}{n - (k + 1)}$$

Equation 3.5

Where:

SSE = the sum of squared error (Equation 3.6),

n = sample size, and

k + 1 = the number of predictors and the intercept.

$$\mathrm{SSE} = \sum_{i=1}^{n} (e_i - \bar{e})^2$$

Equation 3.6

Where:

e_i = residual errors, which is the difference between the linear regression and the observations ($e = \hat{y}_i - y_i$),

SSE = the measure of error, and

$\bar{e}$ = 0.

The coefficient of determination (R^2) is

$$R^2 = 1 - \frac{SSE}{SST}$$

Equation 3.7

Where SSE is the sum of squared error defined above, and SST is the sum of squares total.

The coefficient of determination is the proportion of variation within the dependent variable that is estimated by the developed multiple linear regression model.

$$SST = \sum_{i=1}^{n} (y_i - \bar{y})^2$$

Equation 3.8

Where y_i is the observation, and $\bar{y}$ is the average of the dependent variable. The sum of squares total is the summation of the squared deviation between the dependent variable and the average of the dependent variable. The SST is a measure of variance within the dependent variable (Wilks, 2020).

3.3 Task Three: Establish Adjustment Factor and Final Developed Model

3.3.1 Development of Adjustment Factors

Lequesne and Collins (2019) describe several methods for assigning load ratings to concrete bridges without design plans: rating based on historic design loads, rating based on bridge age and traffic loads, rating based on similar bridges with known load ratings, calculations with assumed measurements/properties, and load testing. Superstructure condition ratings describe the condition of a bridge and classify structural deficiencies, ranging from excellent condition to a threat of imminent failure (Table 3.5). Several state departments of transportation (DOTs) assign inventory and operating rating factors using engineering judgment typically only for structures with a condition rating of 5 or higher due to the limited guidance for load ratings with conditions less than 5.

Several DOTs (States F, L, and O) apply a sliding scale factor in which the load rating factors diminish based on condition ratings (Lequesne & Collins, 2019). This approach may be robust and cost effective for bridges without design plans due to the correlation between condition rating and reported strength load rating, but the factors differ among states. Therefore, robust adjustment factors must be systematically developed for bridge condition.

This study applied a systematic approach of sliding scale adjustment factor based on the median reported inventory and operating ratings. The applied adjustment factor (φ) to account for deterioration was calculated by relating the decrease in load rating relative to a pristine bridge as:

$$\varphi_i = L_i / L_8$$

Equation 3.9

Where:

φ_i = sliding scale adjustment factor (from 0 to 1),

L = median operating or inventory load rating for a given condition rating, and

i = index for condition rating (ranging from 2 to 8).

Table 3.5: Superstructure Condition Rating and Bridge Condition

Super Structure Condition	Condition Rating
1	Imminent Failure
2	Critical Condition
3	Serious Condition
4	Poor Condition
5	Fair Condition
6	Satisfactory Condition
7	Good Condition
8	Very Good Condition
9	Excellent

3.3.2 Bootstrapping Uncertainty Analysis for Adjustment Factors

This study employed a bootstrapping simulation without replacement to investigate the relationship between reported load ratings and sliding scale adjustment factors. This simulation is often used to model and analyze complex systems by incorporating random sampling and repeated iterations (Awang et al., 2015). The primary objective of the bootstrapping simulation was to verify the downward trend in reported load ratings while accounting for uncertainty associated with various adjustment factors from pristine condition to poor condition ratings. This approach captured the variability and distribution of load rating reductions for bridges in a variety of conditions. The uncertainty analysis applied the following procedure:

1. Subset sampling: Various condition rating subsets based on different condition factors were created from the training datasets for inventory and operating ratings to represent the range of reported bridge condition ratings.

2. Random sampling without replacement: Within each adjustment factor subset, load ratings were randomly sampled from the pristine condition bridges and the low-rated condition bridges. This random sampling process ensured that the simulated load ratings captured the inherent variability within each adjustment factor subset.
3. Iterative simulation: The bootstrapping simulation was performed by conducting 10,000 iterations of the random sampling process to generate an uncertainty distribution of load ratings adjustments for each adjustment factor subset.
4. Analysis and visualization: After completing the bootstrapping simulation, the adjustment factor uncertainty distributions were analyzed and the uncertainty range IQR associated with adjustment factor (φ_i) was assessed.

3.4 Task Four: Model Application and Validation

The general form the applied adjustment factors equation is defined as:

$$y = \varphi\hat{y}$$

Equation 3.10

Where:

y = predicted (deteriorated) load rating,

φ = sliding scale adjustment factor (Equation 3.10), and

$\hat{y}$ = pristine bridge rating (Equation 3.4).

Generalization of the final model (Equation 3.4) for inventory and operating was compared to the testing dataset. The testing/validation dataset consisted of 20% of the total data for the inventory and operating ratings. The 95% prediction interval for inventory and operating ratings for each set was calculated as:

$$\hat{y}_{95\%} = y \pm 1.96 * \text{SE}$$

Equation 3.11

Where:

$\hat{y}_{95\%}$ = inventory or operating estimate from Equation 3.4,

SE = standard error (metric tons) from the inventory or operating rating model (Equation 3.11), and

1.96 = z-score that corresponds to the two-tailed central limits of the central 95% probability of a normal distribution (p-values of 2.5% and 97.5%) (Aityan, 2022).

Equation 3.11 is an estimate of the observed trend of observation variation the model captures. The prediction interval was compared to the validation dataset once the 95% confidence intervals were calculated.

Chapter 4: Results

4.1 Preliminary Model Development

4.1.1 Selection of Significant Predictors for the Linear Regression Model

This study applied a backward selection approach to ANOVA on pristine bridges to determine which predictors can be used to estimate inventory or operating rating (Equation 3.4). The automated process iteratively eliminated the least significant predictor or the predictor with the highest p-value until only one significant predictor remained. As shown in Table 4.1 and Table 4.2, the test results showed nine significant predictors for inventory rating and 12 significant predictors with a p-value less than 0.05 for operating rating.

Some significant variables from the ANOVA test have the potential to improve the model more than others. For example, there are trade-offs with model improvement and complexity, so this study utilized forward selection on the significant predictors from the backward selection ANOVA test to minimize model complexity and error and determine which combination and number of predictors would minimize the mean squared error when Equation 3.4 was applied to the entire training dataset without the applied reduction factor (excluding pristine bridges). Results showed that four predictors were required to minimize the MSE for both inventory and operating ratings in the test dataset, as shown in Figures 4.1 and 4.2, respectively. Forward selection results also showed that five predictors were required to minimize the MSE for the inventory rating: YEAR_BUILT_027, DESIGN_LOAD_031, STRUCTURE_KIND_043A, DECK_WIDTH_MT_052, and STRUCTURE_TYPE_043B for an MSE of 62.95 and 66.61 for the testing and training datasets, respectively. Forward selection on the operating rating showed that four predictors were required to minimize the MSE for the testing dataset: YEAR_BUILT_027, DESIGN_LOAD_031, STRUCTURE_KIND_043A, and DECK_WIDTH_MT_052. The training operating rating MSE was 182.07, while the testing dataset MSE for operating rating was 182.27. Both models met the assumption of zero sum of model residuals, as shown in Figures 4.1 and 4.2. Combinations of predictors that improved model performance were then used without increasing model complexity for both models.

Table 4.1: Backward Selection ANOVA Test Results for Pristine Inventory Rating Bridges

Predictor Number	R^2	Name/NBI Code	P-Value
24	0.402	DECK_AREA	0.945
23	0.402	MEDIAN_CODE_033	0.719
22	0.402	TRAFFIC_DIRECTION_102	0.582
21	0.402	RIGHT_CURB_MT_050B	0.573
20	0.402	TRAFFIC_LANES_UND_028B	0.505
19	0.401	LEFT_CURB_MT_050A	0.427
18	0.401	LAT_UND_MT_055B	0.440
17	0.401	YEAR_ADT_030	0.368
16	0.401	ADT_029	0.367
15	0.400	STRUCTURE_LEN_MT_049	0.277
14	0.400	TRAFFIC_LANES_ON_028A	0.190
13	0.399	MAIN_UNIT_SPANS_045	0.212
12	0.398	DEGREES_SKEW_034	0.160
11	0.398	APPR_SPANS_046	0.059
10	0.396	SERVICE_UND_042B	0.081
9	0.395	HORR_CLR_MT_047	0.032
8	0.393	PERCENT_ADT_TRUCK_109	0.001
7	0.389	APPR_WIDTH_MT_032	0.002
6	0.384	MAX_SPAN_LEN_MT_048	0.000
5	0.378	STRUCTURE_TYPE_043B	0.000
4	0.368	DECK_WIDTH_MT_052	0.000
3	0.354	STRUCTURE_KIND_043A	0.000
2	0.321	DESIGN_LOAD_031	0.000
1	0.245	YEAR_BUILT_027	0.000

Table 4.2: Backward Selection ANOVA Test Results for Pristine Operating Rating Bridges

Predictor Number	R^2	Name/ NBI Code	P-Value
24	0.394	DEGREES_SKEW_034	0.951
23	0.394	RIGHT_CURB_MT_050B	0.762
22	0.394	TRAFFIC_LANES_UND_028B	0.583
21	0.394	ADT_029	0.604
20	0.394	SERVICE_UND_042B	0.497
19	0.393	APPR_WIDTH_MT_032	0.511
18	0.393	MEDIAN_CODE_033	0.361
17	0.393	STRUCTURE_LEN_MT_049	0.330
16	0.393	YEAR_ADT_030	0.322
15	0.392	PERCENT_ADT_TRUCK_109	0.218
14	0.391	LEFT_CURB_MT_050A	0.141
13	0.391	TRAFFIC_DIRECTION_102	0.134
12	0.390	DECK_AREA	0.026
11	0.387	MAIN_UNIT_SPANS_045	0.027
10	0.385	HORR_CLR_MT_047	0.019
9	0.383	TRAFFIC_LANES_ON_028A	0.004
8	0.379	APPR_SPANS_046	0.002
7	0.375	LAT_UND_MT_055B	0.000
6	0.366	STRUCTURE_TYPE_043B	0.000
5	0.357	MAX_SPAN_LEN_MT_048	0.000
4	0.344	STRUCTURE_KIND_043A	0.000
3	0.334	DECK_WIDTH_MT_052	0.000
2	0.323	DESIGN_LOAD_031	0.000
1	0.251	YEAR_BUILT_027	0.000

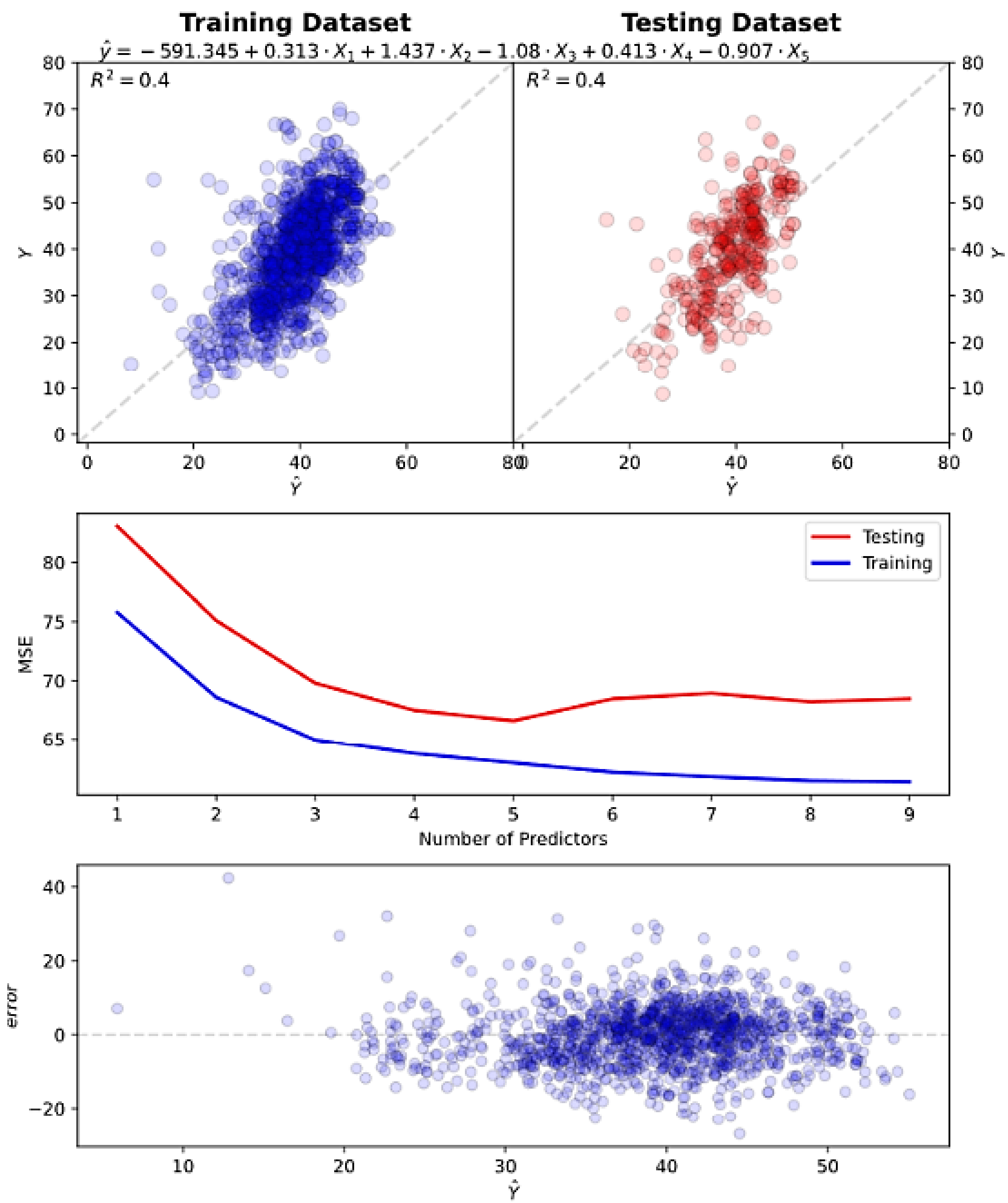

Figure 4.1: Forward Selection for Inventory Rating

The top left figure in Figure 4.1 plots the developed linear regression model for pristine testing bridges, while the top right figure shows the best predictor selection for training pristine bridges. The middle plot compares the MSE with additional predictors added via forward selection. The red line in the middle plot is the MSE for the testing dataset, and the blue line is the MSE for the training pristine bridges dataset. The bottom error plot shows that the error in the residuals is centered around 0 for the inventory rating preliminary model.

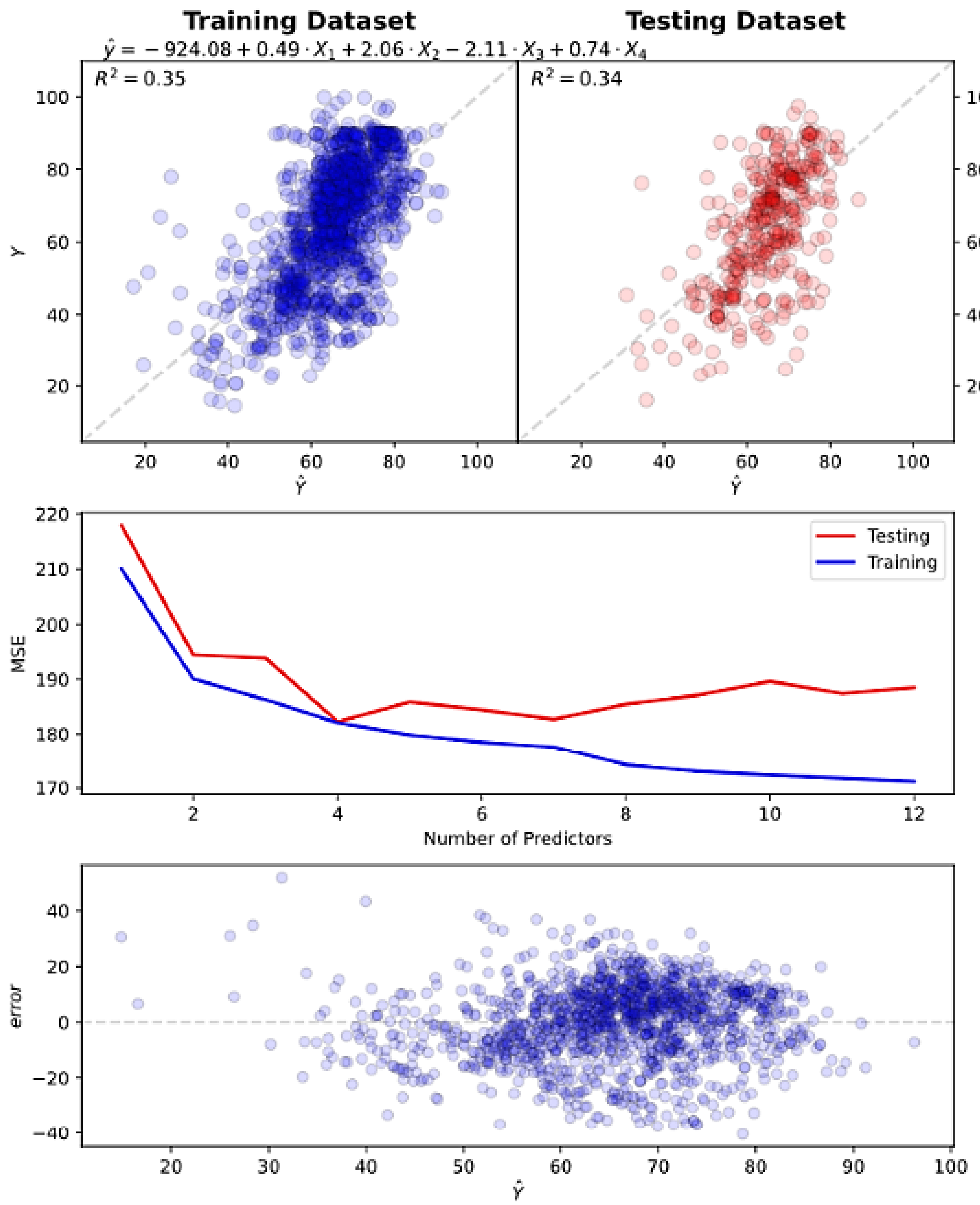

Figure 4.2: Forward Selection for Operating Rating

The top left figure in Figure 4.2 plots the developed linear regression model for pristine testing bridges, while the top right figure shows the best predictor selection for training pristine bridges. The middle plot compares the MSE with predictors added via forward selection. The red line in the middle plot is the MSE for the testing dataset, and the blue line is the MSE for the training pristine bridges dataset. The bottom error plot shows that the error in the residuals is centered around 0 for the operating rating preliminary model.

Figures 4.3 and 4.4 illustrate the correlation between the two most important predictor variables (YEAR_BUILT_027 and DESIGN_LOAD_031), respectively, and bridge rating. As shown, newer bridges and bridges designed for heavy truck loads have higher inventory and operating ratings. YEAR_BUILT_027 and DESIGN_LOAD_031 also correlate because newer bridges tend to be designed for heavy truck loads. Nevertheless, the previous analyses suggested that both predictors include relevant information. In other words, the model outputs improve if both YEAR_BUILT_027 and DESIGN_LOAD_031 are considered instead of just YEAR_BUILT_027. Results in the figures are shown as the mean plus-minus one standard deviation for all bridges.

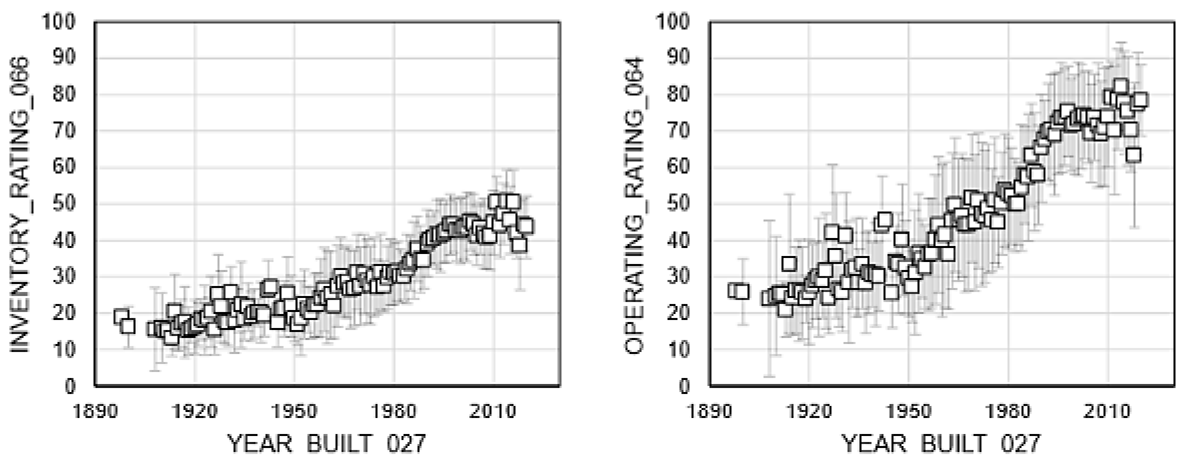

Figure 4.3: Inventory and Operating Ratings versus Year Built

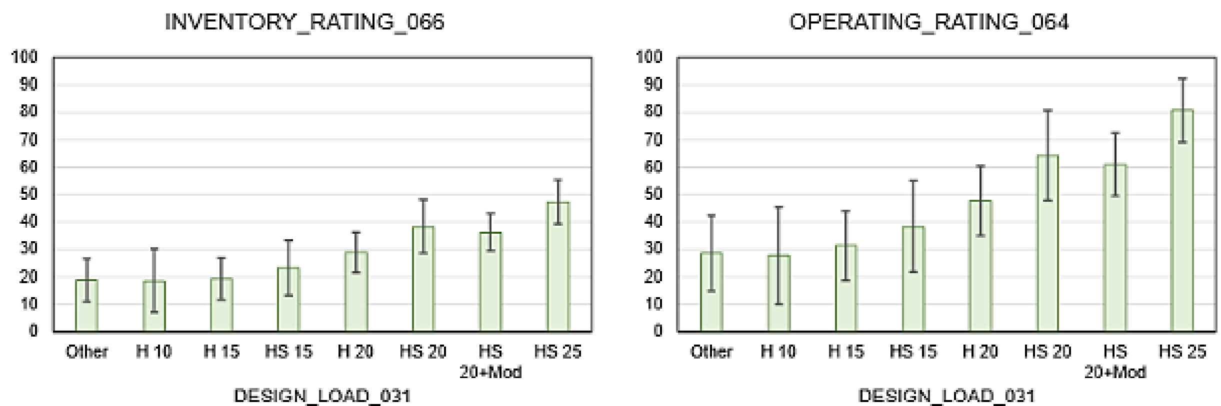

Figure 4.4: Inventory and Operating Ratings versus Designated Design Load

4.1.2 Preliminary Model Assessment

Study results indicated that the best overall performing model for inventory rating was year built (*Age*), design load (*DL*), structure kind (*Struc*), deck width (*W*), and structure type (*Type*). The best overall performing model for operating rating includes year built (*Age*), design load (*DL*), structure kind (*Struc*), and deck width (*W*). The general forms to estimate the baseline load ratings (both models in metric tons) are expressed as

$$\hat{y} = \beta_0 + \beta_1 Age + \beta_2 DL + \beta_3 Struc + \beta_4 W + \beta_5 Type$$

Equation 4.1

$$\hat{y} = \beta_0 + \beta_1 Age + \beta_2 DL + \beta_3 Struc + \beta_4 W$$

Equation 4.2

The significant max span length was not included in the preliminary baseline load rating models because including this variable did not minimize the MSE when the preliminary model was applied to the pristine test bridge datasets for both the inventory and operating ratings. A degree of uncertainty was present with the inclusion of the design live load predictor (*DL*) in Equations 4.1 and 4.2; however, because design loads encoded as 0 indicate that field evaluation and engineering judgment were used to determine the reported live load, meaning that 29 pristine training bridges and 36 training bridges without design plans were present in the inventory and operating rating training dataset, potentially skewing the model predictions for the baseline load rating (FHWA, 2011).

The preliminary models and their predictor β_k are described in Tables 4.3 and 4.4 for inventory and operating ratings, respectively. Figures 4.5 and 4.6 display the trained baseline load rating model and validated testing dataset for the inventory rating, while Figures 4.7 and 4.8 show the trained baseline load rating model and validated testing dataset for the operating rating. The preliminary models performed reasonably well, with a coefficient of determination of 0.40 and 0.35 for the inventory and operating ratings, respectively, and a standard error of 7.93 tons for the inventory rating and 13.49 metric tons for the operating rating. Both preliminary models also met the criteria for an approximate zero mean of the residuals. Therefore, the developed preliminary model was used to predict baseline load ratings for pristine condition bridges.

Table 4.3: Preliminary Baseline Load Rating Regression Model for Pristine Inventory Rating Bridges

Predictor	Name	Weight Term (β)
X_1	YEAR_BUILT_027	0.313
X_2	DESIGN_LOAD_031	1.437
X_3	STRUCTURE_KIND_043A	-1.08
X_4	DECK_WIDTH_MT_052	0.413
X_5	STRUCTURE_TYPE_043B	-0.907
Intercept	-	-591.345

Table 4.4: Preliminary Baseline Load Rating Regression Model for Pristine Operating Rating Bridges

Predictor	Name	Weight Term (β)
X_1	YEAR_BUILT_027	0.489
X_2	DESIGN_LOAD_031	2.060
X_3	STRUCTURE_KIND_043A	-2.109
X_4	DECK_WIDTH_MT_052	0.739
Intercept	-	-924.08

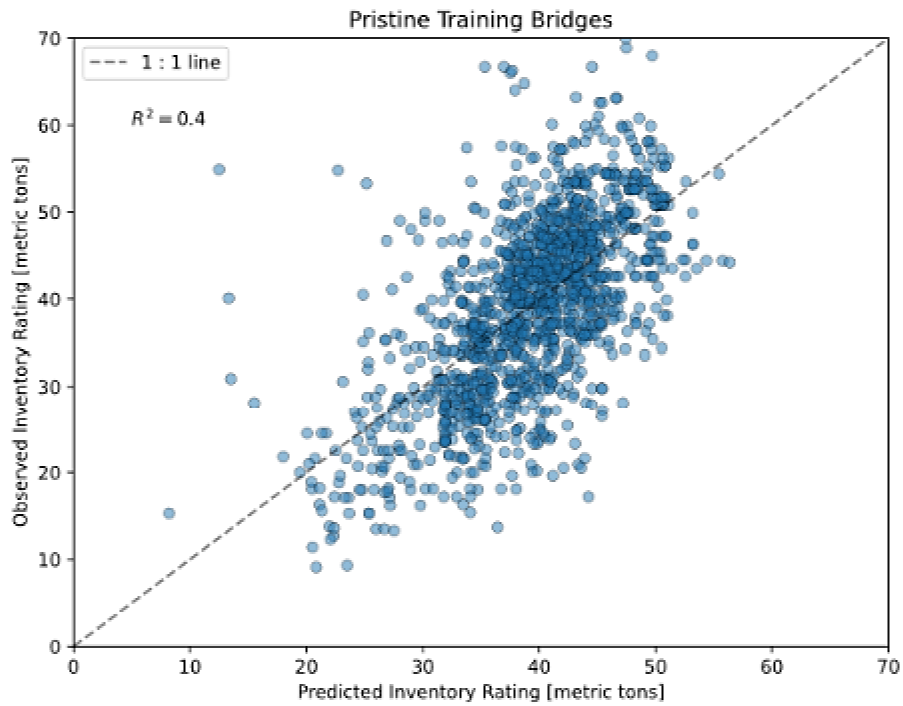

Figure 4.5: Preliminary Baseline Load Rating Estimation of Pristine Inventory Rating Training Bridges (N = 1,453)

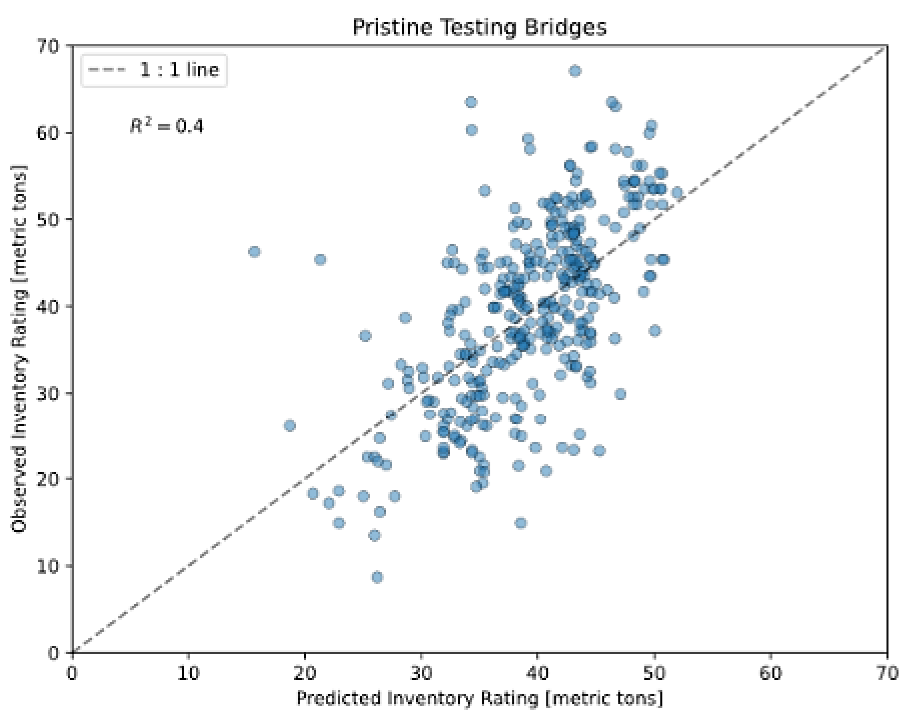

Figure 4.6: Preliminary Baseline Load Rating Estimation of Pristine Inventory Rating Testing Bridges (N = 356)

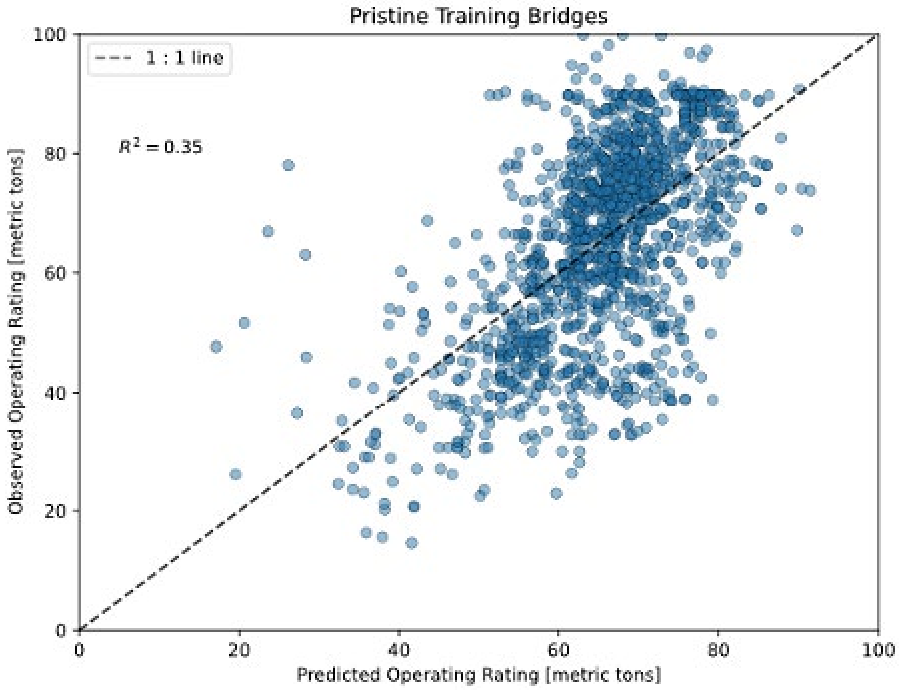

Figure 4.7: Preliminary Baseline Load Rating Estimation for Pristine Operating Rating Training Bridges (N = 1,442)

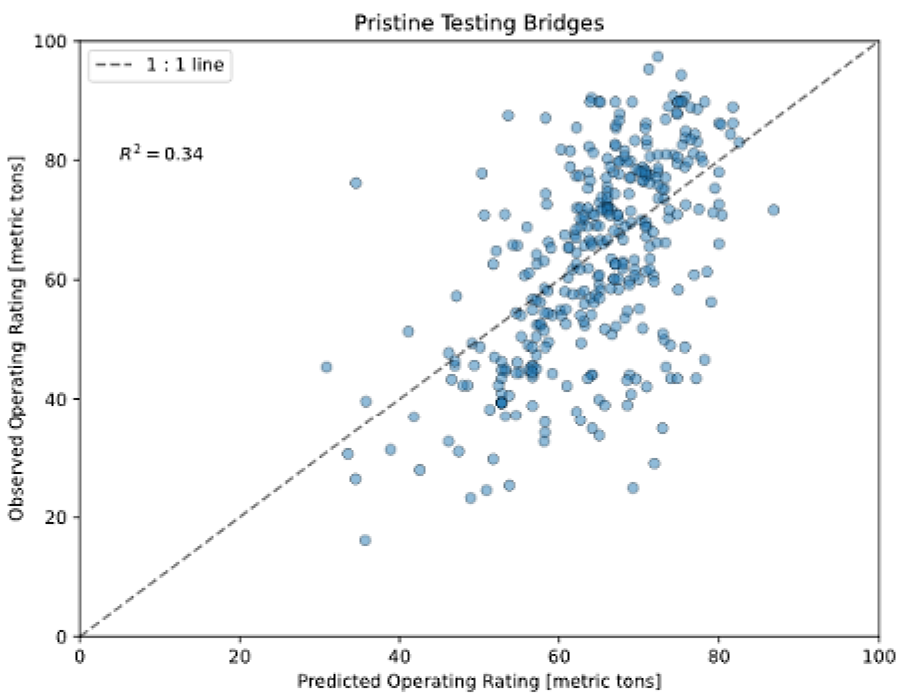

Figure 4.8: Preliminary Baseline Load Rating Estimation for Pristine Operating Rating Testing Bridges (N = 360)

4.2 Adjustment Factor Uncertainty Analysis

4.2.1 Applied Sliding Scale Adjustment Factors

The sliding scale adjustment factors (Equation 3.6) listed in Tables 4.5 and 4.6 are based on the median load rating of a given deteriorated bridge (super structure condition 2 to 7) relative to the median load rating of a pristine bridge (super structure condition 8). For example, the adjustment factor for condition factor 7 for the operating rating in Table 4.6 is 52.2 / 67.1 = 0.78, proving that bridges with a low condition factor are typically weaker than bridges with a high condition factor due to diminished structure condition. Section 3.1.5 verified the validity of the sliding approach because both the training and testing datasets had similarly diminished load ratings with decreasing condition rating. The adjustment factors (φ) were then applied to the model for the inventory and operating ratings to create Equation 3.10, which was applied to the entire training dataset.

Table 4.5: Median Inventory Rating with Condition Rating and Adjustment Factor (φ) for Training Dataset

Super Structure Condition	Inventory Rating Median	Adjustment Factor (φ_i)
2	5.9	0.15
3	7.7	0.19
4	16.2	0.40
5	17.5	0.44
6	20.9	0.52
7	31.5	0.79
8	40.1	1

Table 4.6: Median Operating Rating with Condition Rating and Adjustment Factor (φ) for Testing Dataset

Super Structure Condition	Operating Rating Median	Adjustment Factor (φ_i)
2	9.1	0.13
3	12.6	0.19
4	24.5	0.36
5	27.0	0.40
6	33.8	0.50
7	52.2	0.77
8	67.6	1

4.2.2 Reduction Factor Uncertainty Analysis

One limitation of applying a sliding scale adjustment factor (**φ**) based on the median is the increase of uncertainty for the reported load ratings with different condition factors. Normalization of the adjustment factor to the reported median rating only captures the downward trend of the central tendency; it does not encompass the variation of the reduction from pristine condition to bridges in poor condition. To verify this downward trend, this study performed a bootstrap simulation that randomly sampled the load ratings from different condition rating subsets 10,000 times to obtain a distribution of uncertainty for the sliding scale adjustment factor (Thomopoulos, 2013). Bootstrap simulation results shown in Figures 4.9 and 4.10 verify the downward trend with the uncertainty range with decreased condition rating. However, the upper range (75th percentile) of the IQR for the adjustment factors rating of “8” to “7” exceeded 1.0 because some bridges in pristine condition have a lower load rating than some bootstrap samples of bridges with a condition rating less than “7”. Nonetheless, a clear general reduction of the adjustment factor uncertainty range was reported (Tables 4.7 and 4.8). Thus, this systematic approach for the applied adjustment factor in Equation 3.11 is valid within a reasonable degree of uncertainty.

Table 4.7: Uncertainty IQR Range and Median for Inventory Rating Training Dataset

Condition Rating Adjustment (**φ_i**)	Lower End (**φ**) (25th percentile)	Median (**φ**) (50th percentile)	Upper End (**φ**) (75th percentile)
8 to 7	0.58	0.80	1.08
8 to 6	0.40	0.59	0.80
8 to 5	0.35	0.47	0.69
8 to 4	0.30	0.40	0.55
8 to 3	0.12	0.20	0.35
8 to 2	0.07	0.15	0.22

Table 4.8: Uncertainty IQR Range and Median for Operating Rating Training Dataset

Condition Rating Adjustment (**φ_i**)	Lower End (**φ**) (25th percentile)	Median (**φ**) (50th percentile)	Upper End (**φ**) (75th percentile)
8 to 7	0.59	0.78	1.07
8 to 6	0.41	0.58	0.81
8 to 5	0.36	0.49	0.70
8 to 4	0.30	0.41	0.57
8 to 3	0.12	0.20	0.35
8 to 2	0.07	0.16	0.22

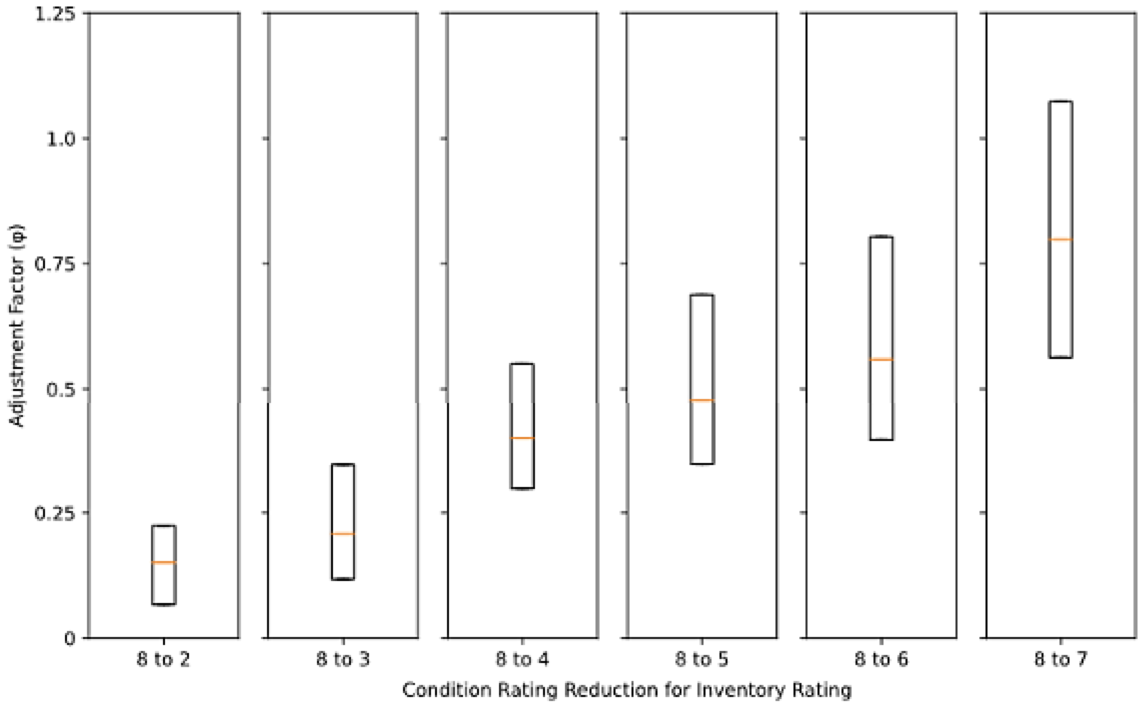

Figure 4.9: Inventory Rating Training Dataset Condition Rating, Adjustment Factor Uncertainty Analysis

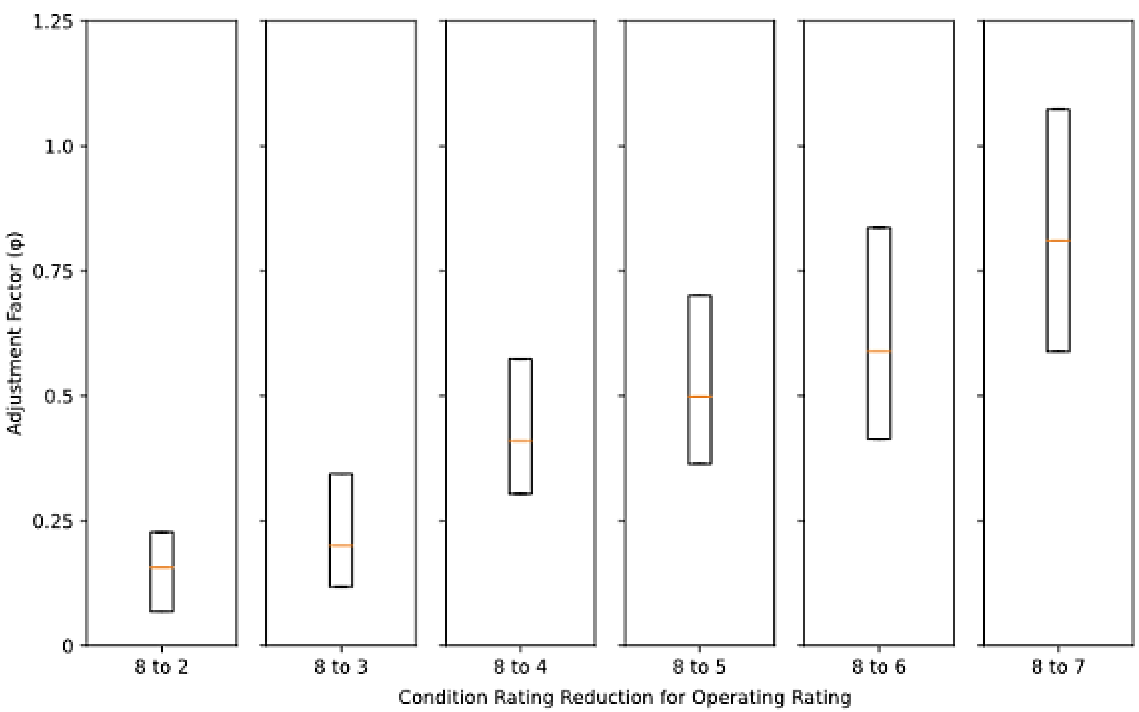

Figure 4.10: Operating Rating Training Dataset Condition Rating, Adjustment Factor Uncertainty Analysis

4.2.3 Adjustment Factor (φ) Applied to Full Training Dataset

Based on study results, the adjustment factors (φ) in Tables in 4.5 and 4.6 were applied to Equations 4.1 and 4.2 for the baseline load ratings models for inventory and operating ratings. The inventory final training model resulted in an R^2 of 0.50, an MSE of 128.06, and a standard error of 11.3 metric tons. The operating rating final training model resulted in an R^2 of 0.49, an MSE of 356.02, and a standard error of 18.89 metric tons. When the preliminary model was applied to the full training dataset with the applied reduction factor, the model performance metric, R^2 increased due to an increased sample size. However, when a linear regression model was applied to a larger dataset, the model was less sensitive to errors and more likely to capture the relationship between predictors and the inventory or operating ratings (Liu et al., 2022). As shown in Figures 4.11 and 4.12, the inventory rating model more effectively predicted the training datasets because of the lower MSE than the fully developed operating model, indicating that the inventory rating model is more accurate than the operating rating model, with less resultant error in model predictions.

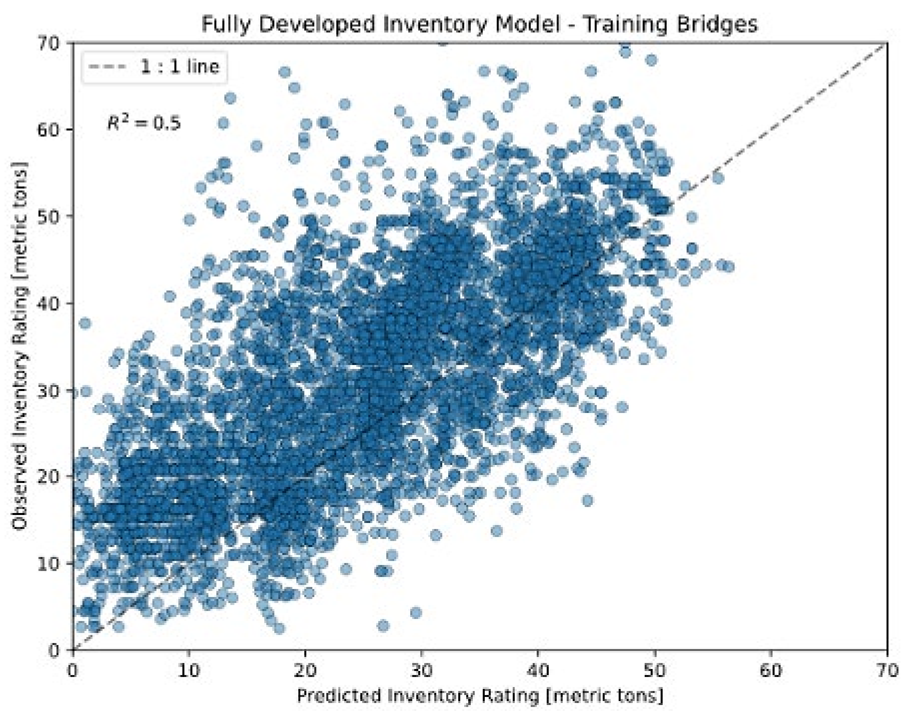

Figure 4.11: Fully Developed Linear Regression Model with the Adjustment Factors for Full Inventory Training Dataset with Applied Adjustment Factor (N = 5,342)

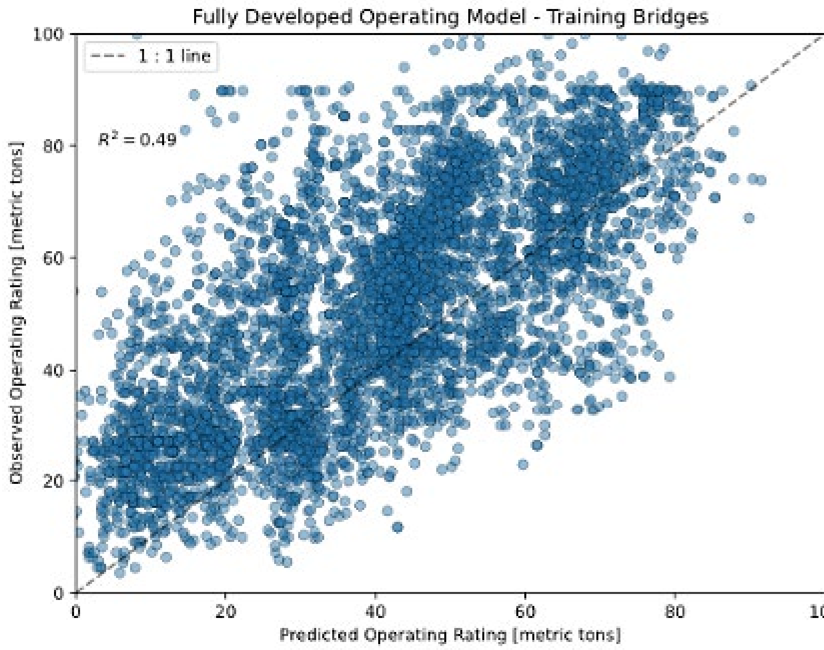

Figure 4.12: Fully Developed Linear Regression Model with the Adjustment Factors for Full Operating Training Dataset with Applied Adjustment Factor (N = 5,571)

4.3 Model Validation

This study then compared the developed inventory and operating rating models to the testing dataset, as shown in Figures 4.13 and 4.14. The validation dataset was comprised of 20% of the filtered dataset from section 4.1, totaling 1,336 bridges with inventory ratings and 1,393 bridges with operating ratings. The inventory rating validation model exhibited a slightly higher coefficient of determination of 0.51, capturing more variation in the validation dataset compared to 0.45 for the operating dataset. Additionally, the inventory rating model had less error for model predictions, with a standard error of 11.31 metric and an MSE of 128.07 compared to a standard error of 19.19 metric tons and an MSE of 368.29. Application of the fully developed model with the applied adjustment factors and standard error from the training datasets for inventory and operating ratings resulted in 94.9% of testing inventory rating bridges and 94.0% of testing operating rating bridges within the 95% prediction intervals. The remaining 5.1% of bridges not captured by the 95% prediction interval for the inventory rating model primarily were HS20 designs or designs with load ratings larger than 40 metric tons and superstructure in good or fair condition.

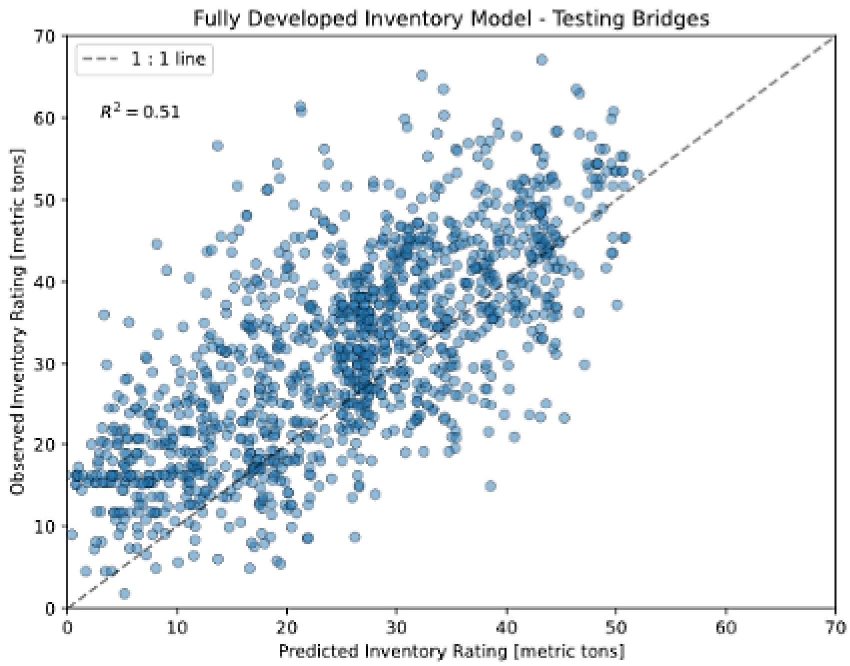

Figure 4.13: Linear Regression Model with Adjustment Factors for Inventory Applied to the Testing Dataset (N = 1,336)

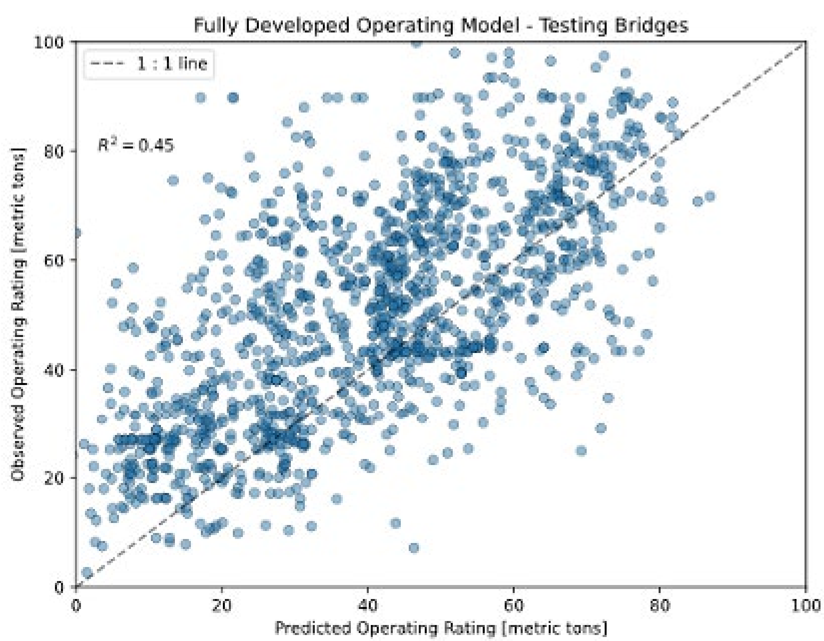

Figure 4.14: Linear Regression Model with Adjustment Factors for Operating Rating Applied to the Testing Dataset (N = 1393)

Chapter 5: Discussion

5.1 Baseline Load Ratings for Bridges with No Design Plans

To establish baseline inventory and operating load ratings for bridges without design plans, this study initially omitted non-applicable bridges from the dataset to develop the linear regression model. All non-concrete bridge structures, outliers, and bridges with unusually high reported strength frequencies were omitted, thereby reducing the dataset from 24,925 bridges to 6,692 inventory rating bridges and 6,964 operating rating bridges. The sample datasets for the inventory and operating ratings were then split, with 80% for training and 20% for model validation. The training and testing split resulted in 5,432 training and 1,336 testing bridges for inventory rating and 5,571 training bridges and 1,393 testing bridges for operating rating model development. Then the dataset was filtered to select only bridges with a condition rating of 8 (pristine condition bridges) to estimate baseline load rating, the median load rating was tabulated, and the distributions for each condition rating were graphed for the inventory and operating ratings (Tables 4.1 and 4.2) to confirm the validity of this approach. Although results showed that the approach was valid, uncertainty remained, as described in section 5.2.

When developing the preliminary model for the inventory and operating ratings, a thorough analysis was conducted to determine the best-performing and ideal number of predictors to minimize model error predictions. Analysis results from backward selection ANOVA test and forward selection showed preliminary models for inventory (Equation 4.1) and operating (Equation 4.2) ratings. The preliminary models satisfactorily captured the variance, achieving a 40% capture rate for inventory rating pristine training bridges and a 35% capture rate for the pristine training operating bridges. Both models (Equations 4.1 and 4.2) met the assumption of normal distribution with a mean of residuals centered around zero. However, the inventory rating preliminary model performed better than the operating rating when applied to the full training dataset, demonstrating an R^2 of 0.50 for the inventory rating compared to 0.49 for the operating rating. The inventory rating also had less error overall in the model, with an MSE of 128 for the testing dataset compared to an MSE of 368 for the operating rating.

One significant uncertainty is that bridges without design plans were used to construct the preliminary model since the predictor DL (Design_Load_31) used bridges with design load 0, which denotes that the modeled design load is unknown. A total of 715 and 755 training bridges with $DL = 0$ were present in the training datasets for the inventory and operating ratings, respectively. Overall, however, the preliminary model was useful for estimating baseline load ratings for bridges in pristine condition without prior design plans. In addition, the model is based on readily available bridge inventory data that can be easily collected and used to estimate baseline load ratings. However, further validation and refinement of the model may be necessary to improve its accuracy and applicability for different bridge types and conditions.

5.2 Adjustment Factors to Account for Bridge Condition Degradation

This study developed a systematic sliding scale adjustment factor to account for changes in load rating due to bridge deterioration and distress. The adjustment factor (φ) is a normalized value between 0 and 1, which is intended to scale a bridge from its predicted pristine rating (condition rating of “8”) to its likely deteriorated load rating (condition rating “2” to “7”) given information from a superstructure evaluation. This approach is reasonable and valid because both the training and testing datasets displayed similar diminished reported load ratings with decreasing condition rating (Figures 3.7 and 3.8).

However, uncertainty was present regarding the sliding scale adjustment factors (φ) obtained from the reported load ratings with decreasing condition ratings. Therefore, a bootstrapping simulation was performed, which included random sampling from load ratings with various condition rating subsets (10,000 times) to obtain an uncertainty distribution for the sliding scale adjustment factor with decreasing condition ratings. The bootstrapping simulation results verified a downward trend in the IQR uncertainty with a reduction in condition rating (Tables 4.8 and 4.9). Because the sampling of distributions is random, occasional bootstraps of a lower condition rating (e.g., 7) would have a higher load rating than a pristine condition bridge (i.e., 8). This can be observed in the 75th percentile of an 8-to-7 adjustment factor (Table 4.7 and Figure 4.9), which indicates as adjustment factor of 1.07, that is the load rating would increase despite the drop in condition rating. However, this is not typically the case; and the median (50th percentile) 8-to-7 adjustment factor was 0.80, suggesting a 20% drop in load rating. Overall, the adjustment factors accurately capture the overall downward trend, with decreased reported load ratings with reduced condition rating, as expected given bridge deterioration.

5.3 Developed Model Application and Limitations

The fully developed models for the inventory and operating ratings with the applied adjustment factors are displayed in Equations 5.1–5.4 (reported in metric tons):

$$\hat{y}_{Inv} = -591.345 + 0.313Age + 1.347DL + -1.08Struc + 0.413W - 0.907Type$$

Equation 5.1

$$\hat{y}_{Inv-95\%} = \varphi_i(\hat{y}_{Inv}) \pm 1.96 \times 11.32$$

Equation 5.2

$$\hat{y}_{Opr} = -924.08 + 0.49Age + 2.06DL - 2.109Struc + 0.739W$$

Equation 5.3

$$\hat{y}_{Opr-95\%} = \varphi_i(\hat{y}_{Opr}) \pm 1.96 \times 18.87$$

Equation 5.4

Where:
$\hat{y}_{Inv}$ = developed baseline load rating for inventory rating,
$\hat{y}_{Inv-95\%}$ = fully developed inventory rating model with applied adjustment factor corresponding to the 95% prediction intervals,
$\hat{y}_{Opr}$ = developed baseline load rating model for operating rating,
$\hat{y}_{Opr-95\%}$ = fully developed operating rating model with applied adjustment factors,
11.32 (metric tons) = standard error from the fully developed inventory rating model,
18.87 (metric tons) = standard error from the fully developed operating rating model (section 4.2.3), and
1.96 = z-score corresponding to the two-tailed central limits of the central 95% probability of a normal distribution (p-values of 2.5% and 97.5%) (Aityan, 2022).

The predictors used in the models are Year Built (*Age*), Design Load (*DL*), Structure Kind (*Struc*), Deck Width (*W*), and Structure Type (*Type*). The descriptions of each predictor, their units, and the applied adjustment factors are described in Tables 5.1–5.7. Tables 5.5–5.7 use the same

NBI code for the model input because this model uses readily available data from the NBI database, meaning the model input for each predictor uses the same input as the database.

Table 5.1: Load Rating Regression Model for Pristine Inventory Rating Bridges

Predictor	NBI Code	Predictor Name	Units
X_1	YEAR_BUILT_027	Age	(-)
X_2	DESIGN_LOAD_031	DL	Refer to table 5.5
X_3	STRUCTURE_KIND_043A	Struc	Refer to table 5.6
X_4	DECK_WIDTH_MT_052	W	Meters
X_5	STRUCTURE_TYPE_043B	Type	Refer to Table 5.7

Table 5.2: Load Rating Regression Model for Pristine Operating Rating Bridges

Predictor	NBI Code	Predictor Name	Units
X_1	YEAR_BUILT_027	Age	(-)
X_2	DESIGN_LOAD_031	DL	Refer to Table 5.5
X_3	STRUCTURE_KIND_043A	Struc	Refer to Table 5.6
X_4	DECK_WIDTH_MT_052	W	Meters

Table 5.3: Median Inventory Rating at a Reported Condition Rating and Associated Adjustment Factor (φ) Based on Training Dataset

Superstructure Condition	Inventory Rating Median	Adjustment Factor (φ_i)
2	5.85	0.15
3	7.65	0.19
4	16.2	0.40
5	17.5	0.44
6	20.85	0.52
7	31.5	0.79
8	40.1	1

Table 5.4: Median Operating Rating at a Reported Condition Rating and Associated Adjustment Factor (φ) Based on Training Dataset

Superstructure Condition	Operating Rating Median	Adjustment Factor (φ_i)
2	9.05	0.13
3	12.6	0.19
4	24.5	0.36
5	27.0	0.40
6	33.8	0.50
7	52.2	0.77
8	67.65	1

Table 5.5: Design Load (DL)/DESIGN_LOAD_031 Units and Description

Design Load NBI Code	Model Code	Metric Design Description	English Design Description
1	1	M9	H 10
2	2	M 13.5	H 15
3	3	MS 13.5	HS 15
4	4	M 18	H 20
5	5	MS	HS 20
6	6	MS 18+Mod	HS 10+Mod
9	9	MS 22.5	HS 25
0	0	Other or Unknown	-

Table 5.6: Structure Kind (Struc)/STRUCTURE_KIND_043A Units and Construction Material Description for Model Development

Structure Type NBI Code	Model Code	Construction Material Description
1	1	Concrete
2	2	Concrete Continuous
5	5	Prestressed Concrete*
6	6	Prestressed Concrete Continuous *

Note: *Post-tensioned concrete uses the same code as prestressed concrete.

Table 5.7: Structure Type (Type)/STRUCTURE_TYPE_043B Units and Description for Model Development

Structure Type NBI Code	Model Code	Structure Description
1	1	Slab
2	2	Stringer/Multi-beam or Girder
3	3	Girder and Floorbeam System
4	4	Tea Beam
5	5	Box Beams or Girders – Multiple
7	7	Frame (excluding culverts)
11	11	Arch – Deck
12	12	Arch – Thru
20	20	Mixed Types
22	22	Channel Beam
0	0	Other

The fully developed models (Equations 5.2 and 5.4) explained 50% and 49% of the variance in the observed inventory and operating rating. Equations 4.1 and 4.2 were used for training and validation using predictors from section 4.1.1. The developed models (Equations 5.2 and 5.4) are intended to estimate inventory and operating ratings for all Kansas concrete bridge structures without design plans and that follow the selection criteria outlined in the methods. Extrapolation of the model to bridge conditions and types that are outside of the ranges used in training could result in erroneous prediction.

The models developed in this study compared reasonably well to the model by Ruiz (2020), which had an R^2 of 0.51 and standard error of 6.51 metric tons. The fully developed inventory rating model had an R^2 of 0.50 with a standard error of 11.31 metric tons, while the operating rating model had an R^2 of 0.49 and standard error of 18.87 metric tons when applied to the training datasets. When the developed models with the applied adjustment factors were validated against the testing dataset, approximately 94.9% and 94.0% of the testing bridges' reported ratings were within the 95% prediction intervals for inventory and operating ratings, respectively. The remaining 5.1% of the bridges for the inventory rating model had design rankings of HS20, and the remaining 6.0% of the bridges for the operating rating model were HS20 designs that were not captured by the model.

The inventory and operating rating models developed in this study were shown to beneficially estimate baseline load ratings for bridges in pristine condition with no original design plans. These models use accessible bridge inventory data to estimate baseline load ratings by applying a reduction factor to accommodate changes in bridge condition rating. However, because the models reflect trends among Kansas bridges within the NBI and not engineering judgment, further validation and refinement of the models is needed to enhance their accuracy and applicability to various bridge types and conditions.

One crucial aspect needing further consideration is the uncertainty associated with omitting outliers in the developed inventory rating model, as discussed in section 3.1.4, since the developed model is only applicable to inventory load ratings up to 70.2 metric tons. This limitation introduces a potential source of bias that may affect the accuracy of the model predictions for inventory ratings exceeding 70.2 tons. Therefore, the developed models should be used to augment engineering judgment rather than replace it. In addition, the NBI data used in this study had inherent limitations, such as the presence of incorrect data entries or a lack of predictors (e.g., slab thickness) that could enhance the model's performance (Ruiz, 2020). Machine learning models could also be employed as an alternative approach to approximate concrete bridge load ratings.

Chapter 6: Conclusions

The primary objectives of this study were to develop a linear regression model to estimate load ratings for concrete bridges in Kansas with no original design plans using data from the NBI database and to establish adjustment factors to account for decreases in load rating correlated with bridge condition rating. These models reflect trends among Kansas bridges within the NBI, not engineering judgment. This approach sought to answer the following question for a given bridge: Knowing nothing more about the structure than what is available within the NBI, what is the *expected* rating based on similar bridges in similar condition in Kansas?

The developed linear regression models considered predictor variables of bridge age, design load, structure kind, deck width, and structure type. The preliminary models exhibited satisfactory performance, effectively capturing a substantial portion of variance of the observed data for the reported inventory ($R^2 = 0.4$) and operating ($R^2 = 0.35$) ratings. To account for changes in load rating due to bridge condition, a sliding scale adjustment factor was applied to normalize the median reported load rating for bridges with a condition rating of 8 or 9 (i.e., very good condition). The adjustment factor was validated with a bootstrapping simulation that demonstrated a downward trend in adjustment factor uncertainty with low condition ratings. However, the sliding scale approach does not apply universally to all bridges in the datasets since some bridges with low condition ratings have higher reported load ratings than pristine bridges.

The final developed model with the applied reduction factors showed satisfactory performance, capturing half the variance in the observed data for inventory ($R^2 = 0.51$) and operating ($R^2 = 0.45$) ratings within the 95% prediction limits when applied to the testing datasets. The models are best used to enhance engineering judgment, help identify outliers and potential errors, and establish expected load ratings for bridges by capturing approximately half the variance in the data. These models also provide a comprehensive approach to estimating load ratings by incorporating adjustment factors to account for decreases bridge loading rating and bridge condition degradation. The model prediction intervals account for the probabilistic uncertainty with predicted load ratings to establish a range engineers can utilize based on familiarity with the respective bridge structure. Further validation and refinement of the models are recommended to improve accuracy and applicability for various bridge types and conditions throughout Kansas.

References

American Association of State Highway and Transportation Officials (AASHTO). (2018). *The manual for bridge evaluation* (3rd ed.). American Association of State Highway and Transportation Officials.

Aityan, S. K. (2022). Confidence intervals. In: *Business Research Methodology*: *Classroom Companion*. Springer. https://doi.org/10.1007/978-3-030-76857-7_13

Awang, Z., Afthanorhan, A., & Asri, M. A. M. (2015). Parametric and non parametric approach in structural equation modeling (SEM): The application of bootstrapping. *Modern Applied Science*, *9*(9), 58.

Federal Highway Association (FHWA). (1995). *Recording and coding guide for the structure inventory and appraisal of the nation's bridges*.

Federal Highway Association (FHWA). (2011). *Revisions to the recording and coding guide for the structure, inventory and appraisal of the nation's bridges (coding guide) – item 31, design load, and items 63 and 65, method used to determine operating and inventory ratings* [Memorandum].

Gao, L. (2013). *Load rating highway bridges in accordance with load and resistance factor rating method*. Outskirts Press, Inc.

Khainge, Y. M., Yee, M. M., & Aung, E. E. (2019). Forecasting stock market using multiple linear regression. *International Journal of Trend in Scientific Research and Development*, *3*(5), 2174–2176.

Kim, T. K. (2017). Understanding one-way ANOVA using conceptual figures. *Korean Journal of Anesthesiology, 70(1)*, 22. https://doi.org/10.4097/kjae.2017.70.1.22

Lequesne, R. D., & Collins, W. (2019). *Synthesis of rating methodologies for concrete bridges without plans* (SL Report 19-2).

Liu, Z., Jiang, M., & Luo, T. (2022). Leveraging low-fidelity data to improve machine learning of sparse high-fidelity thermal conductivity data via transfer learning. *Materials Today Physics*, *28*, 100868. https://doi.org/10.1016/j.mtphys.2022.100868

Perez, H., & Tah, J. H. (2020). Improving the accuracy of Convolutional neural networks by identifying and removing outlier images in datasets using t-SNE. *Mathematics*, *8*(5), 662. https://doi.org/10.3390/math8050662

Rouder, J., Schnuerch, M., Haaf, J. M., & Morey, R. D. (2021). *Principles of model specification in ANOVA designs*. https://doi.org/10.31234/osf.io/e56ab

Ruiz, E. (2020). *Estimating the load rating of reinforced concrete bridges without plans* [Doctoral dissertation, Mississippi State University]. Theses and Dissertations. https://scholarsjunction.msstate.edu/cgi/viewcontent.cgi?article=2931&context=td

Schneider, A., Hommel, G., & Blettner, M. (2010). Linear regression analysis. *Deutsches Ärzteblatt international*. https://doi.org/10.3238/arztebl.2010.0776

Thomopoulos, N. T. (2013). *Essentials of monte carlo simulation, statistical methods for building simulation models*.

Uyanık, G. K., & Güler, N. (2013). A study on multiple linear regression analysis. *Procedia - Social and Behavioral Sciences*, *106*, 234–240. https://doi.org/10.1016/j.sbspro.2013.12.027

van der Goot, R. 2021. We need to talk about train-dev-test splits. In *Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing*. Online and Punta Cana, Dominican Republic. https://aclanthology.org/2021.emnlp-main.368/

Wilks, D. S. (2020). 7: Statistical forecasting. In *Statistical Methods in the Atmospheric Sciences* (4th ed.,).

Yu, Z., Guindani, M., Grieco, S. F., Chen, L., Holmes, T. C., & Xu, X. (2022). Beyond T test and ANOVA: Applications of mixed-effects models for more rigorous statistical analysis in neuroscience research. *Neuron*, *110*(1), 21–35. https://doi.org/10.1016/j.neuron.2021.10.030

K-TRAN

KANSAS TRANSPORTATION RESEARCH AND NEW-DEVELOPMENT PROGRAM

Kansas
Department of Transportation