2023-05-01

A Multiple Model Approach To Bridge Deterioration Modeling

Jin A. Collins
University of Texas at El Paso

A MULTIPLE MODEL APPROACH TO BRIDGE DETERIORATION MODELING

Jin Collins

Doctoral Program in Civil Engineering

APPROVED:

Jeffrey Weidner, Ph.D., Chair

Carlos M. Ferregut, Ph.D.

Cesar Tirado, Ph.D.

Amit J. Lopes, Ph.D.

Stephen L. Crites, Jr., Ph.D.
Dean of the Graduate School

A MULTIPLE MODEL APPROACH TO BRIDGE DETERIORATION MODELING

by

JIN COLLINS, M.S.

DISSERTATION

Presented to the Graduate Faculty of

The University of Texas at El Paso

in Partial Fulfillment

of the Requirements

for the Degree of

DOCTOR OF PHILOSOPHY

Department of Civil Engineering

THE UNIVERSITY OF TEXAS AT EL PASO

May 2023

Acknowledgements

I would like to express my most appreciation and sincere gratitude to my advisor, Dr. Jeffrey Weidner, for his support, guidance, advice, and encouragement throughout the Ph.D. program and research.

I would like to thank my committee, Dr. Carlos Ferregut, Dr. Cesar Tirado, and Dr. Amit Lopes, for dedicating their time by serving on my committee.

I would like to thank all the University of Texas El Paso staff for assisting with my research.

I would also like to thank all the members of the PASS lab for their support and friendships.

Finally, I thank my family for their support and love.

Abstract

Bridge Management Systems (BMSs) have been used to support decision-making in bridge projects, including maintenance, rehabilitation, and replacement under financial constraints. These decisions are predicated on the ability to estimate bridges' future condition states to optimize bridge network systems' performance. Deterioration models are commonly used tools to help transportation agencies make predictions of future condition states of infrastructure facilities, schedule capital investments, and make comparative decisions. Numerous deterioration models have been developed over the last few decades. In particular, Markov chain models utilize bridge data collected from the National Bridge Inventory (NBI) database to make predictions. The accuracy of predictions is measured by comparing the estimated future condition states to the ensemble mean of observed condition states. To improve the accuracy, researchers have applied explanatory variables' effect on the deterioration process in modeling. The selection of explanatory variables and application methods in the modeling are different by researchers, creating uncertainty and bias.

This research presents the development of a framework for a novel deterioration approach utilizing multiple models through a case study. The Texas concrete decks were collected from the NBI database. Markov chain models, time-based Weibull, and corrosion-induced mechanistic models were collected from the literature review. In the development of the model, the transition probability matrices of each model were estimated. A transition probability matrix (TPM) is a product of the Markov chain models. The same explanatory variables were applied in the modeling process except for the mechanistic-based model. The proposed model was developed by integrating these TPMs. The products of the model were a single deterioration curve which presents the future condition rating of a component over time and a range of possible condition ratings at a given time. The results provided reasonable accuracy.

The deterioration model can be utilized to estimate the future condition states of infrastructure facilities more confidently. The estimations can be used to plan more effective and efficient maintenance, rehabilitation, and replacement (MR&R) activities under a limited budget, to prevent potential failure and to maintain an acceptable service level.

Table of Contents

List of Tables

List of Figures

Chapter 1: Introduction

1.1 Motivation

The Intermodal Surface Transportation Efficiency Act of 1991 (ISTEA) marked the start of a new era in transportation in the United States. ISTEA required transportation agencies to take a more proactive approach to planning and asset management. It included a requirement for management systems for pavement, bridge, safety, congestion, public transportation, and intermodal systems. Bridge Management Systems (BMSs) have been used to support decision-making in bridge projects, including maintenance, rehabilitation, and replacement under financial constraints (Agrawal et al., 2010). These decisions are predicated on the ability to estimate bridges' future condition states to optimize bridge network systems' performance. Deterioration models are commonly used tools to help transportation agencies make predictions of future condition states of infrastructure facilities, schedule capital investments, and make comparative decisions. Numerous deterioration models have been developed over the last few decades. There is no perfect model to predict the future conditions of the facilities. Each model has advantages and disadvantages. Deterministic, stochastic, and artificial intelligence models are based on historical condition rating records. Mechanistic models are based on actual data related to corrosion, and reliability-based models are based on a computation of the probability of failure. Stochastic models can capture the inherently probabilistic nature of the deterioration process, which the deterministic models cannot (Agrawal et al., 2010). Mechanistic models describe the deterioration behavior of bridge components induced by corrosion which stochastic models cannot capture. Reliability-based models estimate the probability of failure of bridge structures, and they are not based on historical records. Artificial intelligence models can simulate complex and nonlinear phenomena that classic statistical techniques cannot, but big data are required (Yosri et al., 2021).

This research presents the development of a framework for a new deterioration modeling approach that integrates multiple existing, standalone deterioration modeling approaches. The deterioration models can estimate the future condition states of infrastructure facilities. The estimations can be used to plan more effective and efficient maintenance, rehabilitation,

and replacement (MR&R) activities under a limited budget, to prevent potential failure and to maintain an acceptable service level. This research can provide a better understanding of the deterioration of bridge components for developing and implementing an efficient and effective bridge management system.

1.2 Research Aim and Objectives

This thesis aims to develop a framework of a more robust deterioration model to predict the future condition distribution of bridge components and condition rating of a bridge component by utilizing viable deterioration models, including state-based Markov chain, time-based Weibull and Mechanistic-based models through a case study. The objectives of this research are:

- Understanding the methods of existing deterioration models using the NBI database.
- Analyzing the variability of the models.
- Developing a method to estimate transition probability matrices from the time-based Weibull distribution and mechanistic-based models.
- Developing methods for integrating models.
- Evaluating the result.

1.3 Thesis Organization

This thesis is organized as follows:

Chapter 1 includes the motivation and research objectives.

Chapter 2 reviews the overviews of the National Bridge Inventory database and relevant literature, including stochastic and mechanistic models.

Chapter 3 describes the methodology of developing a multiple-model approach, including single deterioration models, state-based Markov chain, time-based Weibull deterioration, and corrosion-induced mechanistic deterioration models.

Chapter 4 contains a case study. Texas bridge data are collected and used in deterioration modeling. The evaluation method is explained, and the results are presented. In addition, the effect of the explanatory variable, weather condition, is explored.

Chapter 5 includes another case study. Pennsylvania bridge data are collected and used in deterioration modeling.

Chapter 6 contains a discussion, conclusion, and future work.

Chapter 2: Literature Review

2.1 Bridge Deterioration Models

Bridge deterioration models describe the physical condition of bridge structures. The state continuously degrades if no intervention exists, such as MR&R activities. External factors, also known as explanatory variables, such as age, traffic, climate, etc., can affect the deterioration process (Agrawal et al., 2010). The effects of explanatory variables that the external factors influence deterioration are applied implicitly or explicitly in stochastic deterioration modeling. The condition of bridge components is represented using condition rating in the NBI database or the reliability index, which is the probability of failure in reliability-based modeling. These models can be generally classified into deterministic, stochastic, mechanistic, reliability-based, and artificial intelligence (neural network) models. Deterministic models illustrate the relationship between explanatory variables influencing bridge component deterioration and its condition state using statistical methods such as average and linear regression (Yanev and Chen, 1993). These models do not include the intrinsic probabilistic nature of the bridge deterioration process. Stochastic models can capture the uncertainty and randomness of deterioration by having one or more random variables in the process (Agrawal et al., 2010). Stochastic models can be further classified into two groups: state-based and time-based models (Mauch and Madanat, 2001). The Markov chain is commonly used in state-based stochastic deterioration modeling of infrastructure (Butt et al., 1987; Camahan et al., 1987; Jiang et al., 1988; Mauch and Madanat, 2001; Tran, 2007; Baik et al., 2006; Wellalage et al., 2015; Cavalline et al., 2015; Goyal et al., 2020). Jiang et al. (1988) showed the statistical suitability of the Markovian process for bridge deterioration modeling. The Weibull-based probability density function is commonly used in time-based stochastic deterioration modeling of infrastructure (Agrawal et al., 2010). The mechanistic models describe the corrosion-induced deterioration process (Morcous and Lounis, 2007; Nickless and Atadero, 2018). In reliability-based models, the condition states are categorized into five groups: excellent, very good, good, fair, and acceptable conditions corresponding to the reliability index estimated based on a state function which is the difference between the resistance of a structure and stress (Frangopol et al., 2004;

Barone and Frangopol, 2014). The artificial intelligence approach makes use of techniques such as artificial neural networks (Morcous and Lounis, 2005; Tran et al., 2009), the ensemble of neural networks (Bu et al., 2015; Lee et al., 2014; Li and Burgueño, 2010), backward prediction (Pandey and Barai, 1995; Lee et al., 2008), and multi-layer perception (Li and Burgueño, 2010).

2.1.1 The National Bridge Inventory

The safety of bridges in the United States, one of the crucial elements in the transportation system, became a national focus as a result of the collapse of the Silver Bridge in December 1967 (Witcher, 2017). In response, the US Department of Transportation developed and implemented the National Bridge Inspection Standards (NBIS) to assess existing bridges' condition state. Every two years, trained inspectors assess three broad components of a bridge (i.e., deck, superstructure, and substructure) and rate on a 0-9 scale. The collected data are reported to the Federal Highway Administration (FHWA) for inclusion in the National Bridge Inventory database. The conditions of bridge components have been stored since 1992. In addition, the conditions of bridge elements have been reported to the FHWA for inclusion in the National Bridge Element Inventory since 2015. Table 2.1 summarizes the rating codes and descriptions for bridge components in the NBI database (Administration and Transportation, 2012).

The NBI database is the primary source of data for most deterioration modeling in the United States because of the relatively long history of data available. However, it is not without it's drawbacks. NBI data is subjective. Component condition is assigned by individual inspectors, and while inspectors are trained consistently, there is inherent subjectivity in their assessments. The data also lacks the fidelity to differentiate between actions that could improve component condition (MR&R) and simple variation in assessment over subsequent cycles (e.g., inspector B believes condition state is actually higher than inspector A from two years prior). Even data input errors are indistinguishable. Despite these issues, NBI data is the best available data source to explore deterioration modeling.

Table 2.1: The Condition Rating Codes and Descriptions

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

2.2 Stochastic Approaches to Deterioration Modeling

In stochastic modeling, the deterioration process is assumed to contain one or more random variables that capture the uncertainty and randomness of the probabilistic nature of the process (Agrawal et al., 2010). Stochastic models can be classified into state-based and time-based models (Mauch and Madanat, 2001). In the state-based model, such as the Markov chain, transition probabilities define the likelihood of either staying in a given state or moving to another state in a given period. The calculation of these probabilities is presented in the following sections.

2.2.1 State-based Markov chain Deterioration Models

Markov chain models are widely used for state-based deterioration modeling. The BMSs need two predictions to support bridge management; the future distribution of bridge components with specific condition ratings at any given time and the future condition ratings of a bridge component at any time (Jiang et al., 1988). These two elements can be obtained using state-based Markov chain deterioration models. The main goal of these models is to estimate a

transition probability matrix.

2.2.1.1 Markov chain

A Markov chain is a discrete-time stochastic model created by Andrey Markov, a Russian mathematician (Kouemou and Dymarski, 2011). It states that the probability of each event in a sequence of events depends only on the state of the previous event. The model is used in various domains that deal with sequential data such as traffic, weather, finance, business, medical, internet, etc. It also has been used in infrastructure deterioration modeling; pavements (Butt et al., 1987; Camahan et al., 1987), stormwater pipe (Micevski et al., 2002), sewer pipe (Baik et al., 2006), and bridges (Jiang et al., 1988; Ranjith et al., 2013). In bridge deterioration modeling, condition ratings of bridge components are the states, and the probability is the probability of transitioning from a condition rating to a lower condition rating. A matrix form of these probabilities is called the transition probability matrix (TPM) or transition matrix (Jiang et al., 1988). Numerous Markov chain models can be used to determine a TPM.

2.2.1.2 Transition Probability Matrix

A TPM is an n by n matrix that contains transition probabilities of condition ratings, where n is the number of states. The sum of probabilities in each row of the matrix will be one. In bridge deterioration modeling, a common TPM is a seven-by-seven matrix including condition rating 9 to 3. According to the data, the lowest rating is most likely 3, indicating that bridge components are typically subjected to MR&R activities at a rating not less than 3 (Jiang et al., 1988). Equation 2.1 shows an example of a typical TPM used in modeling. The P_{99}, P_{88}, ..., P_{44} are transition probabilities of staying at current condition ratings 9, 8, ..., 4, respectively. The $1-P_{99}$, $1-P_{88}$, ..., $1-P_{44}$ are the probabilities of transitioning to the next lower condition ratings 8, 7, ..., 3 respectively.

$$\boldsymbol{P} = \begin{bmatrix} p_{9,9} & 1-p_{9,9} & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & p_{8,8} & 1-p_{8,8} & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & p_{7,7} & 1-p_{7,7} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & p_{6,6} & 1-p_{6,6} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & p_{5,5} & 1-p_{5,5} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & p_{4,4} & 1-p_{4,4} & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \tag{2.1}$$

The condition distribution of bridge components at a network level can be estimated using Equation 2.2, where $P(t)$ is a state vector of probabilities of distribution of the components at any given time t, $P(0)$ is an initial state vector, and P is a TPM. Expected condition ratings of a bridge component $E(t, P)$ at any time t are estimated using Equation 2.3, where R is a condition rating vector. The expected condition ratings are needed to develop a deterioration curve of a bridge component.

$$\boldsymbol{P}(t) = \boldsymbol{P}(0)\boldsymbol{P}^t \tag{2.2}$$

$$E(t, \boldsymbol{P}) = \boldsymbol{P}^t\boldsymbol{R} \tag{2.3}$$

There are numerous methods to estimate TPMs. In this research, the regression nonlinear optimization model proposed by Jiang et al. (1988), the Bayesian maximum likelihood model proposed by Tran (2007) and Wellalage et al. (2015), the ordered probit-based model proposed by Madanat et al. (1995), Poisson-based and Negative binomial-based models proposed by Madanat and Ibrahim (1995), and Proportional hazards-based model proposed by Cavalline et al. (2015) and Goyal et al. (2020) are presented.

2.2.1.3 Regression Nonlinear Optimization

Jiang et al. (1988) proposed two different methods to estimate TPMs. One is for predicting the condition distribution of bridge components using Equation 2.4. It is based on actual

condition rating data in a sample. Transition probabilities $P_{i,j}$ from condition rating i to j in a given time can be calculated by dividing the number of component transitions $n_{i,j}$ by the total number of components before transitions n_i where i = 9, 8, ..., 4 and $j = i, (i-1)$.

$$P_{i,j} = \frac{n_{i,j}}{n_i} \tag{2.4}$$

The other method proposed by Jiang et al. (1988) is developing a deterioration rate curve that provides a relationship between condition rating and age. Multiple TPMs are estimated because the authors adopt the zone technique developed by Butt et al. (1987), which was used to estimate pavement condition. The sample data are divided into age groups consisting of six years. A deterioration rate in each group is assumed to be homogeneous. The TPMs of each group are estimated by optimizing an objective function, Equation 2.5, where the coefficients (*A*, *B*, *C* and *D* in Equation 2.5) of a nonlinear regression curve *S*(*t*) are obtained by evaluating goodness-of-fit. Once the TPMs of individual groups are obtained, expected condition ratings at any time *t*, *E*(*t*,***P***) can be estimated using Equation 2.2 with an initial state vector. The initial state vector of the first group is $\begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$. The initial state vector of the second group is the last state vector of the first group. Therefore, each group's last state vector becomes the next group's initial state vector.

$$\begin{aligned} & min \sum_{t=1}^{T} |S(t) - E(t, \boldsymbol{P})| \\ & \text{where} \quad S(t) = A + Bt + Ct^2 + Dt^3 \end{aligned} \tag{2.5}$$

2.2.1.4 Bayesian Maximum Likelihood

Like the regression nonlinear optimization method, transition probabilities are obtained by optimizing an objective function. The objective function is derived from Bayes' theorem (Equation 2.6), which states that the conditional probability of event θ given event Y is proportional to the product of the probability of event θ and the conditional probability of event Y given event θ.

$$P(\theta|Y) = \frac{P(\theta)P(Y|\theta)}{P(Y)} \tag{2.6}$$

The $P(\theta|Y)$ can be described as the likelihood of θ given Y. It can be simplified using the joint probability theory and converted into a logarithmic form. Transition probabilities of staying at the current condition ratings can be estimated by maximizing the logarithmic likelihood function shown in Equation 2.7, where T is the largest age in the data set, N_i^t is the number of bridge components in condition rating i at time t, and $\boldsymbol{C}_{it}$ is a vector of probabilities of a condition rating i at time t as a function of TPM by Equation 2.3 (Wellalage et al., 2015).

$$log[L(\theta|Y)] = max \sum_{t=1}^{T} \sum_{i=9}^{4} N_i^t log(\boldsymbol{C}_{it}) \tag{2.7}$$

2.2.1.5 Ordered probit-based

Madanat et al. (1995) introduced an ordered probit in bridge component deterioration modeling. The ordered probit is often used for predicting ordinal variables in social science. Ordinal variables are variables that are categorical and ordered. For example, poor, fair, good, and excellent indicate states. Equation 2.8 is an ordered probit model for bridge component deterioration proposed by Madanat et al. (1995). It describes a relationship between unobserved deterioration U and a set of explanatory variables X where β is a regression coefficient, and ϵ is a random error term assumed to take the form of a normal cumulative distribution function.

$$log(U) = \beta X + \epsilon \tag{2.8}$$

The unobserved deterioration, U, can be mapped into the discrete value of the change in condition rating Z, which is bounded by thresholds θ_j, and θ_{j+1} shown in Equation 2.9 for condition state j. The probability of Z of each bridge component in a normal cumulative distribution function can be calculated using Equation (2.10).

$$Z = j \quad \text{if and only if} \quad \theta_j < U \leq \theta_{j+1} \tag{2.9}$$

$$P(Z = j) = F(\theta_{j+1} - \beta X) - F(\theta_j - \beta X) \tag{2.10}$$

The unknown variables β, θ_j, and θ_{j+1} can be estimated by maximizing the logarithm likelihood function. In bridge component deterioration modeling, bridge components are grouped according to previous condition ratings. The authors assume that the change in condition rating is a realization of different mechanistic deterioration processes. The unknown variables are estimated for each group using Equation 2.11. The d_{nj} is when the *Z* of a bridge component *n* equals *j*; otherwise, it is zero. The variable *N* is the total number of bridge components in the group.

$$\log[L(\beta, \theta|X_n, d_{nj})] = \sum_{n=1}^{N} \sum_{j=0}^{i-1} d_{nj} \log[F(\theta_{j+1} - \beta X_n) - F(\theta_j - \beta X_n)] \tag{2.11}$$

The estimated parameters are substituted in Equation 2.10, and the transition probabilities of individual components in the groups are calculated. The probabilities are grouped again according to the change in condition rating to estimate mean transition probabilities $P_{i,(i-j)}$ of transition from condition rating *i* to $(i - j)$ of condition rating group *i* using Equation 2.12.

$$\begin{aligned} P_{i,(i-j)} &= \frac{1}{N} \sum_{n=1}^{N} p(j|X, i) \\ \text{for} \quad & j = 0, 1, 2, ..., i \end{aligned} \tag{2.12}$$

2.2.1.6 Poisson and Negative Binomial Regression Models

Poisson regression is often used to model discrete and non-negative data. It can predict a dependent variable that is assumed to have a Poisson distribution given one or more independent variables. The Poisson distribution is the probability of a given number of events occurring in a discrete interval. A property of the Poisson distribution is that its mean equals its variance. If a set of count data is assumed to have the Poisson distribution, but its mean is not equal to its variance, the data is considered over-dispersed. Negative binomial regression is used to model over-dispersed count data. It is based on a mixture of the Poisson and gamma distributions. Madanat and Ibrahim (1995) applied Poisson regression and Negative binomial regression to bridge component deterioration modeling. The procedure of estimating transition probabilities using proposed models are the same as the ordered probit-based model previously discussed.

Bridge component data are grouped according to condition rating, and parameters are estimated by maximizing a logarithm likelihood function. Transition probabilities of individual component are calculated by substituting the estimated parameters. Transition probabilities of the bridge components are obtained by averaging the sum of transition probabilities of each component which are grouped again by the change in condition rating. The following objective and probability functions are used to estimate parameters and transition probability of a component, n. Equations 2.12 and 2.13 are the logarithm likelihood function (objective function) and the probability function in Poisson regression, respectively. Equations 2.14 and 2.15 are the logarithm likelihood function and probability function in Negative binomial regression, respectively. These Poisson-based and Negative binomial-based approaches can explicitly model the structure of the deterioration process within a discrete time-period as a function of explanatory variables.

$$\log(L) = \sum_{n=1}^{N} Z_n(\beta X_n) - e^(\beta X_n) \tag{2.13}$$

$$p(Z_n = j) = \frac{e^{\lambda_n}\lambda_n^j}{j!} \tag{2.14}$$

$$\begin{aligned}\log(L) = \sum_{n=1}^{N} \log\left(\Gamma\left(\frac{1}{\alpha} + Z_n\right)\right) - \log\left(\Gamma\left(\frac{1}{\alpha}\right)\right) - \log(Z_n!) \\ + \frac{1}{\alpha}\log\left(\frac{1}{1+\alpha\lambda_n^*}\right) + Z_n \log\left(1 - \frac{1}{1+\alpha\lambda_n^*}\right)\end{aligned} \tag{2.15}$$

$$P(Z_n = j) = \frac{\Gamma\left(\frac{1}{\alpha} + j\right)}{\Gamma\left(\frac{1}{\alpha}\right) j!}\left(\frac{1}{1+\alpha\lambda_n^*}\right)^{\frac{1}{\alpha}}\left(1 - \frac{1}{1+\alpha\lambda_n^*}\right)^j \tag{2.16}$$

where Z_n = the change in condition rating; j = 0, 1, 2, . . . , i that i is a condition rating of a group; β = the regression coefficient; X_n = the explanatory variable of a component n; $\lambda_n = e^{\beta X_n}$; $\lambda_n^* = e^{\beta X_n + \epsilon_n}$ where ϵ_n = random error of a component n; $\Gamma()$ = the gamma function; and α = the rate of over-dispersion.

2.2.1.7 Proportional Hazard Model

Cavalline et al. (2015) and Goyal et al. (2020) proposed a bridge component deterioration model from the Cox proportional hazards model, which is commonly used in medical research to study patients' survival time and explanatory variables. The model consists of two parts: baseline survival function and hazard ratio. The baseline survival function estimates transition probabilities of staying at a particular condition rating. Equation 2.17 describes a TPM where $p_{9,9}$, $p_{8,8}$, ..., $p_{4,4}$ are baseline transition probabilities, which average transition probabilities over condition rating duration and HR_9, HR_8, ..., HR_4 are hazard ratios corresponding to condition rating 9, 8, ..., 4.

$$\boldsymbol{P} = \begin{bmatrix} p_{9,9}^{HR_9} & 1-p_{9,9}^{HR_9} & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & p_{8,8}^{HR_8} & 1-p_{8,8}^{HR_8} & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & p_{7,7}^{HR_7} & 1-p_{7,7}^{HR_7} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & p_{6,6}^{HR_6} & 1-p_{6,6}^{HR_6} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & p_{5,5}^{HR_5} & 1-p_{5,5}^{HR_5} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & p_{4,4}^{HR_4} & 1-p_{4,4}^{HR_4} & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \tag{2.17}$$

The baseline transition probability $p_i(t)$ is a survival probability $S_i(t)$ at time *t* in a group of condition rating *i* (Equation 2.18), where $F_i(t)$ is the cumulative probability function of failure in which a condition rating transitions to lower condition ratings.

$$\begin{aligned} & p_i(t) = S_i(t) \\ & \text{where} \quad S_i(t) = 1 - F_i(t) \end{aligned} \tag{2.18}$$

The hazard ratio (*HR*) is the proportionality constant when the relative risk of transitioning to a lower condition rating is evaluated to baseline explanatory variables shown in Equation 2.19, where z^1 is a vector of explanatory variables and z^0 is a vector of the baseline values. The HR describes the effect of explanatory variables. An HR value of less than 1 indicates that an

explanatory variable decreases the deterioration rate, and a value greater than one means that an explanatory variable increases the deterioration rate.

$$HR = e^{\beta(z^1 - z^0)} \tag{2.19}$$

2.3 Summary of State-based Markov Chain models

State-based Markov chain models mentioned above are summarized in terms of random variables in optimization, functions in computational process, assumption in deterioration rate, moreover, ways explanatory variables applied in modeling. Table 2.2 summarizes regression nonlinear optimization (RNO), Bayesian maximum likelihood (BML), ordered probit model (OPM), Poisson regression (PR), negative binomial regression (NBR), and proportional hazard model (PHM). The RNO and BML estimate transition probabilities by optimizing objective functions. The PHM is an empirical data-based model that does not use an optimization method. Explanatory variables are expressed explicitly in PR, NBR, and PHM. It means that the values of transition probabilities can be changed by applying the number and effect of explanatory variables. The BML is based on the assumption of a constant deterioration rate for the entire service life of the component. The RNO, OPM, PR, NBR, and PHM utilize variable deterioration rates according to groups, age groups and condition rating groups.

Table 2.2: Summary of state-based Markov chain deterioration models

<table>
<tr><th>Model</th><th>Optimization
(Random Variables)</th><th>Function</th><th>Explanatory
Variables</th><th>Assumption
(Deterioration Rate)</th></tr>
<tr><td>RNO</td><td rowspan="2">Transition probability</td><td>Third order polynomial</td><td rowspan="3">Implicit</td><td>Constant in age group</td></tr>
<tr><td>BML</td><td>Conditional probability</td><td>Constant in the entire service life</td></tr>
<tr><td>OPM</td><td>Threshold,
Regression coefficient</td><td>Normal</td><td rowspan="4">Constant in condition rating group</td></tr>
<tr><td>PR</td><td rowspan="2">Regression coefficient</td><td>Poisson</td><td rowspan="3">Explicit</td></tr>
<tr><td>NBR</td><td>Binomial</td></tr>
<tr><td>PHM</td><td>NA</td><td>Survival, Hazard ratio</td></tr>
</table>

2.4 Time-based Weibull Deterioration models

The Weibull distribution is commonly used in reliability and survival analyses. Agrawal et al. (2010) proposed a Weibull-based deterioration model to estimate probabilities of a duration of staying at a given condition ratings of bridge elements. The Weibull distribution can be described with two random variables: the shape and scale parameters. A shape parameter, β, affects the distribution's general shape and represents the failure rate behavior. If β is less than 1, then the rate decreases; if it is greater than 1, then the rate increases; and if it is equal to 1, then the rate is constant. A scale parameter, η, affects the spread of the distribution on the horizontal axis. The larger η, the more stretched the distribution. These parameters are obtained by fitting data for different condition ratings. The mean duration, T_i, of staying at the condition rating i can be estimated using Equation 2.20 where Γ represents the gamma function which is an extension of factorial function that $\Gamma(n+1) = n!$ for all non-negative numbers n including complex numbers.

$$E(T_i) = \eta_i \Gamma\left(1 + \frac{1}{\beta_i}\right) \tag{2.20}$$

2.5 Mechanistic-based Deterioration models

2.5.1 Current Mechanistic deterioration models

Researchers state current mechanistic-based deterioration models are not perfect because of the lack of parameter information and the basis of assumptions in modeling. State-based or time-based deterioration models were based on condition rating, determined the overall condition of components by visual inspection, and were helpful in network-level management (Morcous and Lounis, 2007). Mechanistic-based deterioration models were developed based on physical parameters, identifying the extent and severity of the deterioration mechanism of components by condition survey. The mechanistic-based models were valuable for estimating a more accurate accumulation of damage. However, its computational complexity and empirical data requirement suppress use in deterioration modeling (Stewart and Rosowsky, 1998; Morcous and Lounis, 2007; Hu et al., 2013; Nickless and Atadero, 2018). An ideal mecha-

nistic deterioration model included the effects of multiple deterioration processes, interactions between elements, protective systems, and MR&R activities. There were gaps between the models proposed and the actual bridge condition because of the lack of experimental and field data for modeling (Nickless and Atadero, 2018). Moreover, the authors stated that currently available bridge condition data were insufficient to evaluate mechanistic deterioration models because the condition data were crude and subjective. Hu et al. (2013) proposed a mechanistic model that included three mechanistic models to estimate corrosion initiation time, crack initiation time, and time of crack propagation to the surface to predict the surface cracking time of reinforced concrete decks. It took about 20 to 50 years to reach CR 3, depending on the surface chloride ($5.4kg/m^3$ to $1.8kg/m^3$) content.

2.5.2 Chloride-induced corrosion mechanistic deterioration model

Corrosion occurs due to chemical reactions between metals and their surrounding environment. Carbonation and chloride are the main reasons for inducing corrosion in reinforced concrete. Carbonation in the concrete causes a reduction of the pH. If the pH level may drop below 8, then reinforced steel cannot retain its passive film, which is subsequently broken by chloride ions. If the chloride accumulates at certain level, then the corrosion is initiated. The most common resource of chloride is deicing salts used due to the policy which started in the 1960s (Berver et al., 2001). In winter, anti-icing chemical solutions and deicing salts are applied to keep roadways and bridge decks clear of ice. The anti-icing solutions weaken the formation of bonded snow and ice for easy removal and deicing salts (i.e., sodium chloride $(NaCl)$, magnesium chloride $(MgCl_2)$, and calcium chloride $(CaCl_2)$) break the bond between the ice and the surface (Lawson et al., 2017).

Unlike stochastic deterioration models that use a qualitative performance measurement such as condition ratings, mechanistic-based models use quantitative measurements to assess condition states of bridge components, such as the amount of spall, delamination, or contaminated area due to chloride. Mechanistic models predict damage, damage growth, and damage impact on the safety and serviceability of a facility due to applied loads and environment (Morcous and Lounis, 2007). The deterioration of reinforced concrete decks in the U.S. have been ex-

periencing deterioration associated with corrosion due to applying deicing salts to keep them clear of ice in winter (Berver et al., 2001). The corrosion-induced deterioration process of reinforced concrete decks is described in the initiation and propagation stages (Tuutti, 1982). The initiation stage corresponds to corrosion initiation. It can be described as the time interval from the exposure to chlorides until the start of corrosion. It is assumed that diffusion governs the chloride ingress process in concrete decks to estimate the corrosion initiation time. The propagation stage corresponds to damage initiation, such as cracking, delamination, or spalling. Equation 2.21 is a mechanical model called Crank's closed-form solution of Fick's second law of diffusion. It is used to estimate the diffusion of chloride concentration $C(d_c, t)$ as a function of surface chloride concentration C_s and Gaussian error function erf which is in terms of cover depth d_c, diffusion coefficient D, and time t. Probabilistic modeling is needed to include uncertainty in prediction (Morcous and Lounis, 2007).

$$C(d_c, t) = C_s \left[1 - erf\left(\frac{d_c}{2\sqrt{Dt}}\right)\right] \tag{2.21}$$

Chapter 3: Methodology

Numerous deterioration models have been developed to estimate prediction of future conditions of bridge components or elements. However, there is only limited work on deterioration modeling approaches that leverage more than one model. Morcous and Lounis (2007) proposed a combination of probabilistic and mechanistic models that could support network and project decision-making. However, they did not introduce a method to integrate models. A multiple model approach proposed for deterioration modeling will be developed by integrating models. This approach can (1) reduce uncertainty causing by engineering judgment in modeling process and (2) be more realistic by including deterioration behavior information. The methodology of the framework includes data analysis, the development of stochastic (state-based Markov chain and time-based Weibull) and Mechanistic-based deterioration models, the development of integration methods of these models, and the analysis of prediction reliability of the proposed model. Figure 3.1 describes a workflow of the proposed methodology.

This research will propose integration methods of single models based on Markov transition probability matrix (TPM). There are more state-based Markov chain deterioration models than other models, time-based, and mechanistic-based deterioration models proposed in the literature. The state-based Markov chain and time-based Weibull models will be developed to be suitable for integrating with concrete decks collected from the NBI database. Researchers have implemented different ways to improve the prediction accuracy measured by comparing the estimated conditions to the means of the observed condition in deterioration modeling. The methods are based on various assumptions, which are different by the researchers. Stochastic deterioration models collected from the literature are developed by applying the same assumptions. Then the results are evaluated and compared. Unlike the state-based Markov chain models, TPMs cannot be obtained directly from time-based Weibull and Mechanistic-based models. It is necessary to develop a method to estimate transition probabilities from the results of these models. The two methods will be applied to integrate the models. The proposed models will estimate the future condition distribution of bridge components and condition ratings of a component. Finally, the reliability of the proposed models will be measured by comparing

it to the collected sample data.

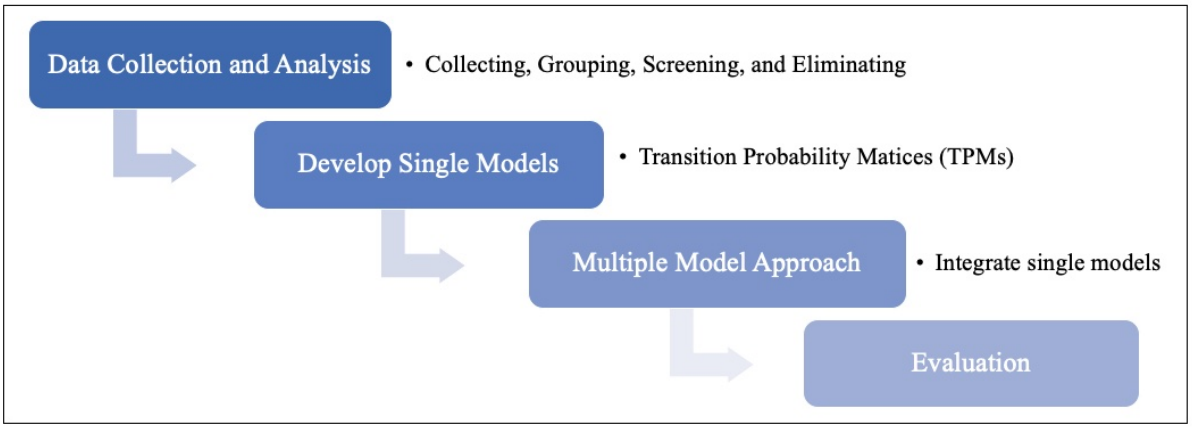


Figure 3.1: Workflow of methodology

3.1 Data Collection and Preparation

3.1.1 Collecting/Grouping

The NBI database contains information about bridges that can influence bridge deterioration, such as structural material type, design type, design load, and so on. In deterioration modeling, these characteristics are called ”explanatory variables” because they are considered to influence deterioration. In this research, bridges are grouped according to six explanatory variables, described below. The number in the parentheses is the item number in the NBI database. The details of each item are in Appendix A. This research grouped concrete decks not subjected to maintenance, rehabilitation, and replacement in deterioration modeling.

- Design_Load (31): Item 31 is the information of the live load which was used for structural design. The numerical codes are used to classify the live load. Bridges that are code 7 (pedestrian) and 8 (railroad) are excluded.

- Open_Closed_Posted (41): Item 41 provides the information on the operational status. The status is indicated with an alphabetic code. All bridges, except bridges that are code A (Open, no restriction), are excluded.

- Deck_COND (58): Item 58 contains the deck condition rating of a bridge. The numerical codes are used to indicate deck conditions. Bridges rated the code N (Not Applicable)

are excluded (Table 2.1).

- Work_Done_By (75B): Item 75 is the information on proposed work to improve bridge condition. Item 75B contains whether the proposed work will be done by contract or force account using numerical codes. All bridges which contain numerical values are excluded.

- Year_Reconstructed (106): Item 106 contains the year of most recent reconstruction. All reconstructed bridges are excluded.

- Deck_Structure_Type (107): Item 107 is the information on the type of deck system using numerical codes. Bridges with code 1 (concrete cast-in-place), 2 (concrete precast panels), and 9 (other) are included in this research.

3.1.2 Screening

Two successive inspection period data are required for ordered probit, Poisson regression, negative binomial regression, and proportional hazard modeling. Data are eliminated if there are discrepancies between the two data sets, if bridges records are not in both data sets, and if there are increases in condition ratings through visual inspection.

3.1.3 Eliminating Outliers

The extreme data points can considerably influence deterioration modeling (Butt et al., 1987). This research uses a boxplot method that graphically shows a statistical summary, such as the mean, dispersion, and skewness of a data set through its quartiles, to identify outliers. First, the sample data are divided into groups according to condition rating. Then the data of each group are analyzed by the boxplot method.

3.2 Development of Single Deterioration Models

The models are developed to estimate a TPM for the entire service life of a deck having the same format. The same sample data are used in state-based Markov chain and time-based Weibull deterioration modeling. Matlab (R2022b) and Excel (2016) are used for modeling.

Excel is used to adjust the data for application in each model.

3.2.1 State-based Markov chain Model

A state-based Markov chain model can estimate the probability of staying at a particular condition rating or transitioning to a lower condition rating.

3.2.1.1 Regression Nonlinear Optimization (RNO)

The RNO approach leveraged here is developed based on the model proposed by Jiang et al. (1988); multiple TPMs were used to predict the future condition distribution of bridge components and condition ratings of a component. The authors assumed that the deterioration rate was not consistent for the entire service life of a bridge component. They adopted the zone technique developed by Butt et al. (1987) divided the sample data into age groups of six years. A deterioration rate in each group was assumed to be constant. In this research, the deterioration rate is assumed to be consistent for the entire service life of a deck, and a TPM is estimated.

3.2.1.2 Bayesian Maximum Likelihood (BML)

The logarithm likelihood function proposed by Wellalage et al. (2015) estimates transition probabilities of staying at current condition ratings. The sample data are organized by rows and columns. The first column is age from 1 to the oldest in each group. The next to the last column are condition ratings. Each cell contains the number of components in each condition rating at a given age.

3.2.1.3 Ordered Probit Model (OPM), Poisson Regression (PR), and Negative Binomial Regression (NBR)

These models are developed based on models proposed by Madanat et al. (1995) and Madanat and Ibrahim (1995). The sample data are divided into groups according to condition ratings. The decks of each group had the same condition rating as the previous year. Instead of computing probabilities of staying at the current condition rating and transitioning to mul-

tiple lower condition ratings, which Madanat et al. (1995) and Madanat and Ibrahim (1995) estimated, only a transition probability of staying at the current condition rating is estimated in this research.

3.2.1.4 Proportional Hazard Model (PHM)

In this model, a transition probability consists of two parts, baseline probability and hazard ratio (Cavalline et al. 2015; Goyal et al. 2020). Like OPM, PR, and NBR, the sample data are grouped according to the previous condition rating. A baseline probability of each group can be estimated using the Matlab function, empirical cumulative distribution function (ecdf). The hazard ratio of each group is estimated using the Matlab function, Cox proportional hazard regression (coxphfit).

3.2.1.5 Summary of the developed state-based models

Table 3.1 summarizes the developed models. The RNO, BML, OPM, PR, and NBR estimate random variables by optimizing objective or logarithm likelihood functions. The random variables in RNO and BML are transition probabilities, and in OPM, PR, and NBR are parameters that are input to estimate transition probabilities such as thresholds (OPM) and regression coefficients (OPM, PR, and NBR).The PHM is not used as an optimization method to obtain parameters or probabilities. It is an empirical data-based model.

Table 3.1: Summary of the key characteristics of the models used in the developed approach

Model	**Optimization**	**Function**	**Assumption**
RNO	✓	Polynomial	Constant deterioration rate for the entire service life
BML	✓	Conditional	
OPM	✓	Normal	Different deterioration rate according to condition rating group
PR	✓	Poisson	
NBR	✓	Negative Binomial	
PHM	NA	Survival	

3.2.2 Time-based Weibull Model

The probability of duration of staying at a particular condition rating can be estimated using the time-based Weibull model. Figure 3.2 shows the procedure to estimate a TPM.

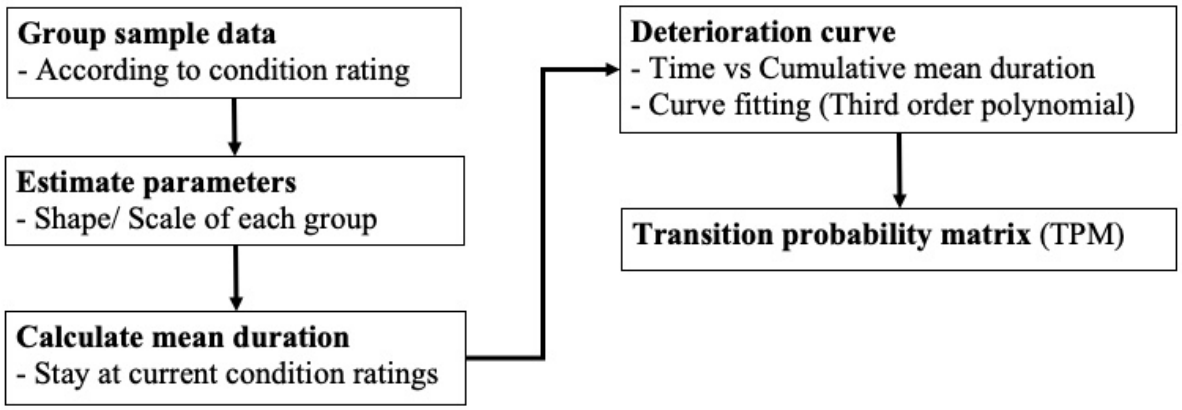


Figure 3.2: Workflow to estimate a TPM for a Weibull model

First, the sample data are divided into groups of condition ratings. Second, each group's scale (η) and shape (β) parameters are obtained using the Matlab function (wblfit). Third, with inputs of the parameters, the expected mean duration at given condition ratings are estimated. Fourth, a third-order polynomial equation of the meantime, $S(t)$, is obtained using the curve fitting function of the Matlab. Finally, transition probabilities of staying at given condition ratings can be estimated by optimizing the objective function, which is used in the RNO modeling (Equation 2.5). Figure 3.2 shows the work process to estimate a transition probability matrix (TPM).

3.2.3 Mechanistic-based Deterioration Model

A transition probability matrix (TPM) is developed using a corrosion-induced mechanistic model. Figure 3.3 shows a procedure of mechanistic based deterioration modeling.

There are different mechanistic models to determine different physical phenomena, such as corrosion initiation, cracking initiation, and failure, so it is necessary to select a mechanistic model. This research uses Crank's closed-form solution of Fick's second law of diffusion model for corrosion-induced modeling. Equation 3.1 is reorganized from Equation 2.21 to calculate the initiation time of corrosion (T_i). Unlike the stochastic deterioration model, mechanistic

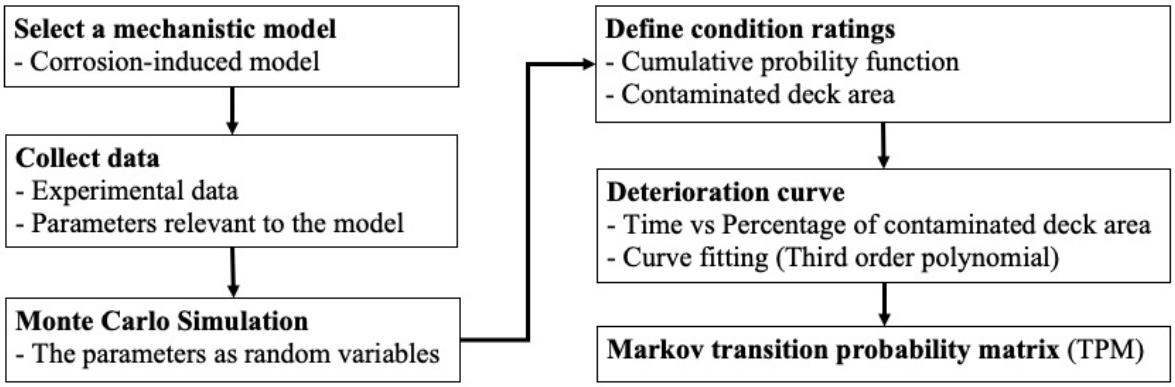


Figure 3.3: Workflow to estimate a TPM for a mechanistic deterioration model

models are commonly based on experimental data. The data are collected through a literature review. Four parameters are required in Equation 3.1.

- Surface chloride concentration (C_s): It is the maximum chloride concentration at a certain depth (typically 0.5 inches) below the surface (Hu et al., 2013). The concentration values depend on the type of deicing salts and the frequency of applying the salts. The higher frequency increases concentration (Xi et al., 2018).

- Chloride diffusion coefficient (D): It is defined by multiple factors such as temperature, concrete age, moisture content, and so on (Hu et al., 2013).

- Chloride threshold level (C_{th}) represents the minimum chloride concentration required to initiate corrosion (Nickless and Atadero, 2018). It is influenced by the water-cement ratio, temperature, materials and so on (Stewart and Rosowsky, 1998). It is also varied depending on the reinforcing rebar. The epoxy-coated reinforcing rebar is higher than black rebar (Hu et al., 2013). The chloride threshold level will vary, so it is treated as a uniformly distributed random variable within the 0.6 to 1.2 kg/m^3 (Stewart and Rosowsky, 1998).

- Concrete cover depth (d_c): A typical cover depth of concrete bridge decks in Texas is about 2 inches (Transportation, 2012).

$$T_i = \frac{d_c^2}{4D\left[erf^{-1}\left(1-\frac{C_{th}}{C_s}\right)\right]^2} \tag{3.1}$$

where T_i = corrosion initiation time and erf^{-1} = inverse Gaussian error function. The Monte Carlo simulation technique is utilized to create a probabilistic density function and cumulative distribution function of the time to corrosion initiation due to the uncertainty in the prediction of deterioration. Surface chloride concentration and chloride diffusion coefficient of reinforced concrete in the U.S. are log-normally distributed random variables with means of 3.5 kg/m^3 and 63.11 $mm^2/year$, respectively (Stewart and Rosowsky, 1998; Hu et al., 2013; Nickless and Atadero, 2018). The chloride threshold level of the black bar is a uniformly distributed random variable with a range of 0.6 to 1.2 kg/m^3 Stewart and Rosowsky (1998), a range of 0.4 to 2.4 kg/m^3 (Nickless and Atadero, 2018) in U.S. and log-normally distributed with a mean of 1.35 kg/m^3 and standard deviation of 0.14 in Canada (Morcous and Lounis, 2007). In this research, parameters proposed by Stewart and Rosowsky (1998) are used because the parameters are an overall mean in the U.S. nationwide. Moreover, the other researchers used the parameters in their modeling. Table 3.2 summarizes parameters, values, and distributions applied in the Monte Carlo simulation.

Table 3.2: Parameters for Monte Carlo simulation of corrosion initiation time

Parameter	**Value**	**Unit**	**Distribution**
Surface chloride concentration, C_s	μ = 3.5 σ = 1.75	kg/m^3	Log-normal
Chloride diffusion coefficient (D)	μ = 63.11 σ = 47.33	$mm^2/year$	Log-normal
Chloride threshold level (C_{th})	min = 0.6 max = 1.2	kg/m^3	Uniform
Concrete cover depth (d_c)	μ = 50.8 σ = 12.7	mm	Normal

A probability of corrosion initiation at a given time is obtained. Figure 3.4 shows the cumulative distribution and probability density function at corrosion initiation time. The mean initiation time is about 25 years which is greater than 21 years, as estimated by Morcous and Lounis (2007) and Hu et al. (2013), and between 12 and 22 years estimated, estimated by Nickless and Atadero (2018). The result can vary depending on the concrete cover depth (Lethanh et al., 2017).

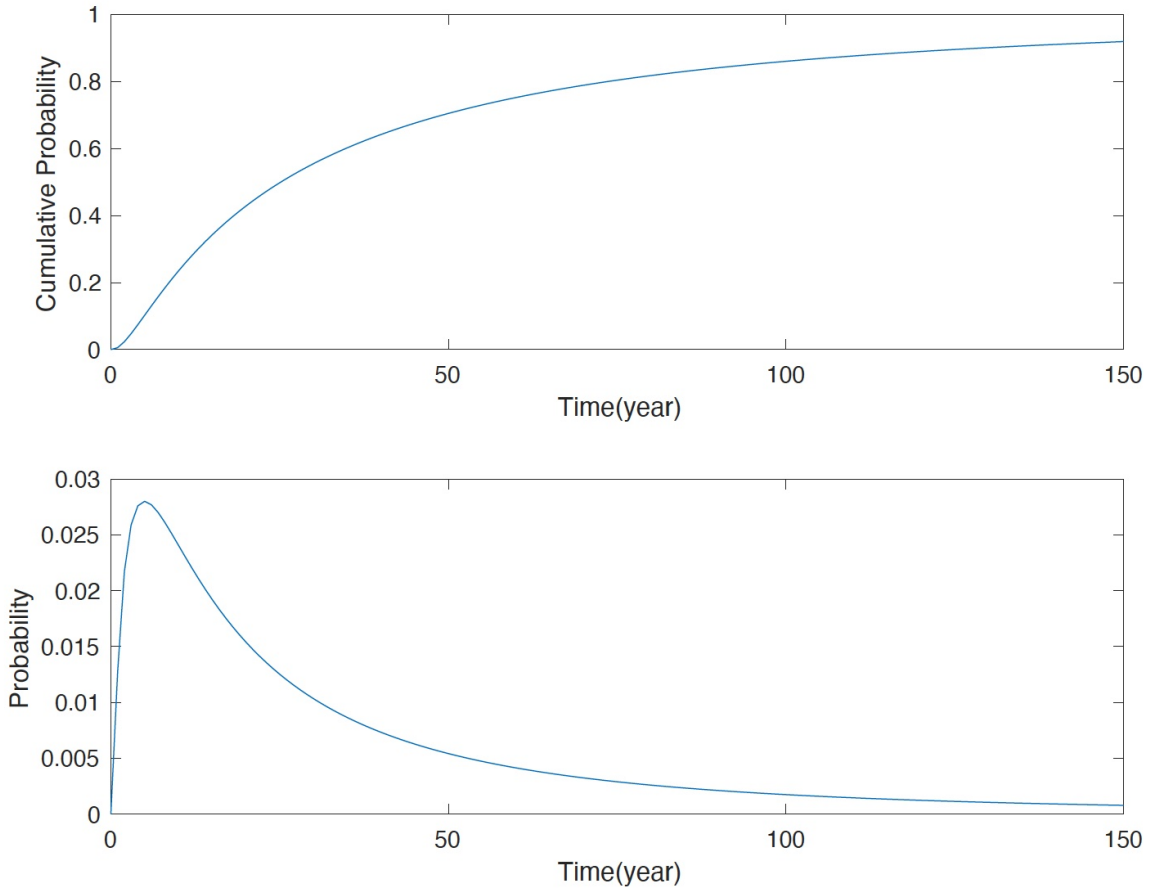


Figure 3.4: Cumulative distribution (top) and probability density (bottom) functions of corrosion initiation time developed via Monte Carlo simulation

Condition ratings can be defined by the cumulative probability. Morcous and Lounis (2007) stated that a cumulative probability is equal to the percentage of a concrete deck's chloride-contaminated or corrosion-induced damaged area at a given time. Table 3.3 shows the concrete bridge deck condition rating by condition indicators (Administration and Transportation, 2012). This research uses chloride content (column 5 in Table 3.3) to define condition ratings.

Table 3.4 contains the cumulative probabilities, corresponding times, and defined condition ratings. The probabilities equal the contaminated concrete deck area percentage (Morcous and Lounis, 2007). A concrete deck transitions to the next lower condition rating at the corresponding times. For example, the deck will transition to CR 6 at three years of age and stay at CR 6 for five years before transitioning to CR 5. At 35 years, the deck will be rated at CR 3 because the contaminated area of the deck will reach 60 percent.

After determining condition ratings corresponding to the cumulative probabilities, a re-

Table 3.3: Mapping concrete bridge deck condition ratings to corrosion indicators as a function of deck area

Rating (1)	Condition Indicators (% Deck Area)			
	Spall (2)	Delamination (3)	Electrical potentials (4)	Chloride content (lb/yd^3) (5)
9	None	None	0	0
8	None	None	None >0.35	None >1.0
7	None	<2%	45% <0.35	None >2.0
6	<2% spall or sum of all deteriorated or contaminated deck concrete <20%			
5	<5% spall or sum of all deteriorated or contaminated deck concrete <20 - 40%			
4	>5% spall or sum of all deteriorated or contaminated deck concrete <40 - 60%			
3	>5% spall or sum of all deteriorated or contaminated deck concrete >60%			
2	Deck structural capacity grossly inadequate			
1	Deck repaired by replacement only			
0	Holes in deck – danger of other sections of deck falling			

Table 3.4: Define condition rating (CR) from cumulative probability

Time (year)	**Cumulative Probability (%)**	**Condition Rating (CR)**
0	0	CR 9
1	1	CR 8
2	2	CR 7
3	5	CR 6
9	21	CR 5
18	40	CR 4
35	60	CR 3

gression curve is generated using the curve fitting function of the Matlab. Finally, transition probabilities of staying at given condition ratings can be estimated by optimizing an objective function, which is the same one used in the RNO modeling (Equation 2.5).

3.3 Development of Multiple Model Approach

In this research, a framework of a novel deterioration modeling approach is proposed by integrating multiple models. Two integration methods are explored; method 1 and method 2 shown in Figure 3.5. Method 1 uses the expected condition ratings, and method 2 uses the TPMs obtained from the single models. The details of developing processes are explained in the following sections.

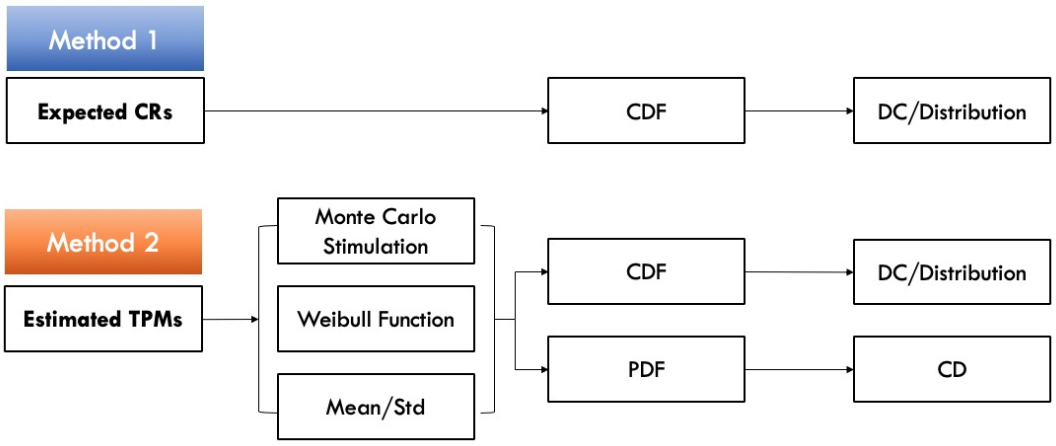


Figure 3.5: Two proposed multiple model approaches. Method 1 combines models at the CR level while Method 2 combines models at TPM level.

3.3.1 Method 1: Using Expected Condition Rating

The expected condition ratings at a given time from the single models are used to construct a cumulative distribution function (CDF) and a probability density function (PDF) at a given time. For an example, Table 3.5 shows the expected condition ratings obtained from the single models of zone 1 in Texas. A CDF and a PDF are obtained from 1 to 80 years using eight expected condition ratings at a given age. A deterioration curve is a curve that connects each mean of the probability density function at a given age. The CDF is used to obtain the condition rating distribution (Distribution in Figure 3.5) at a given age.

Table 3.5: Expected Condition Ratings obtained from all single models of zone 1 in Texas

Age	RNO	BML	OPM	PR	NBR	PHM	WDM	MDM
1	8.83	8.33	8.28	8.27	8.31	8.21	8.96	8.13
2	8.64	8.05	7.90	7.85	7.83	7.85	8.93	7.36
3	8.46	7.90	7.64	7.55	7.51	7.62	7.89	6.69
4	8.27	7.80	7.45	7.33	7.32	7.44	8.86	6.12
5	8.10	7.72	7.30	7.15	7.19	7.28	8.82	5.60
6	7.95	7.64	7.17	7.00	7.12	7.15	8.78	5.14
7	7.81	7.58	7.07	6.87	7.07	7.02	8.75	4.77
8	7.68	7.52	6.98	6.76	7.04	6.91	8.71	4.50
9	7.57	7.46	6.90	6.66	7.03	6.81	8.68	4.31
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
77	6.35	6.18	5.30	5.15	6.98	5.02	6.58	4.00
78	6.34	6.17	5.29	5.14	6.98	5.02	6.55	4.00
79	6.34	6.16	5.29	5.14	6.98	5.02	6.53	4.00
80	6.33	6.15	5.28	5.14	6.98	5.02	6.50	4.00

3.3.2 Method 2: Using Transition Probability Matrix (TPM)

In this method, TPMs estimated from the models above are used. The steps are following.

1. A cumulative probability vector at a given time ***C*(t)** is constructed from a state vector (using Equation 2.2) where p_9^t, p_8^t, ... and p_3^t are transition probabilities of CR 9, 8, ... and 3 and C_9^t, C_8^t, ... and C_3^t are cumulative transition probabilities of CR 9, 8, ... and 3 at given time t.

$$\textbf{\textit{state vector}(t)} = \begin{bmatrix} p_9^t & p_8^t & p_7^t & p_6^t & p_5^t & p_4^t & p_3^t \end{bmatrix}$$

$$C_9^t = p_9^t$$

$$C_8^t = p_9^t + p_8^t$$

$$C_7^t = p_9^t + p_8^t + p_7^t$$

$$C_6^t = p_9^t + p_8^t + p_7^t + p_6^t$$

$$C_5^t = p_9^t + p_8^t + p_7^t + p_6^t + p_5^t$$

$$C_4^t = p_9^t + p_8^t + p_7^t + p_6^t + p_5^t + p_4^t$$

$$C_3^t = p_9^t + p_8^t + p_7^t + p_6^t + p_5^t + p_4^t + p_3^t$$

$$\boldsymbol{C}(\mathbf{t}) = \begin{bmatrix} C_9^t & C_8^t & C_7^t & C_6^t & C_5^t & C_4^t & C_3^t \end{bmatrix}$$

2. The Monte Carlo simulation generates random one hundred thousand (100,000) samples between 0 and 1. For example, if a sample is between 0 and C_9^t, then the sample is a condition rating of 9; if a sample is between C_9^t and C_8^t, then the sample is a condition rating of 8, and so on.

3. Figure 3.6 is an example of a histogram of the distribution of generated samples from 1 to 6 years with a TPM obtained from the RNO in zone 1. More histograms (from 7 to 78 year) are in Appendix C. The figure is an example of a histogram of the distribution of generated samples from 1 to 6 years with a TPM obtained from the RNO in zone 1. As shown in the histograms, the samples for each year have a different distribution. Therefore, the Weibull distribution is considered suitable to obtain a mean and standard

deviation of the distribution of the samples of each year.

4. Mean and standard deviation are computed by averaging means and standard deviations of probability density functions of single models at a given time.

5. Normal probability density function and cumulative distribution function at a given time are generated using the mean and standard deviation from step 4. The central limit theorem states that the distribution of sample means is approximately normally distributed as the sample size gets larger and larger.

6. A deterioration curve containing a range of condition ratings at a given time is constructed and the expected condition distribution is estimated.

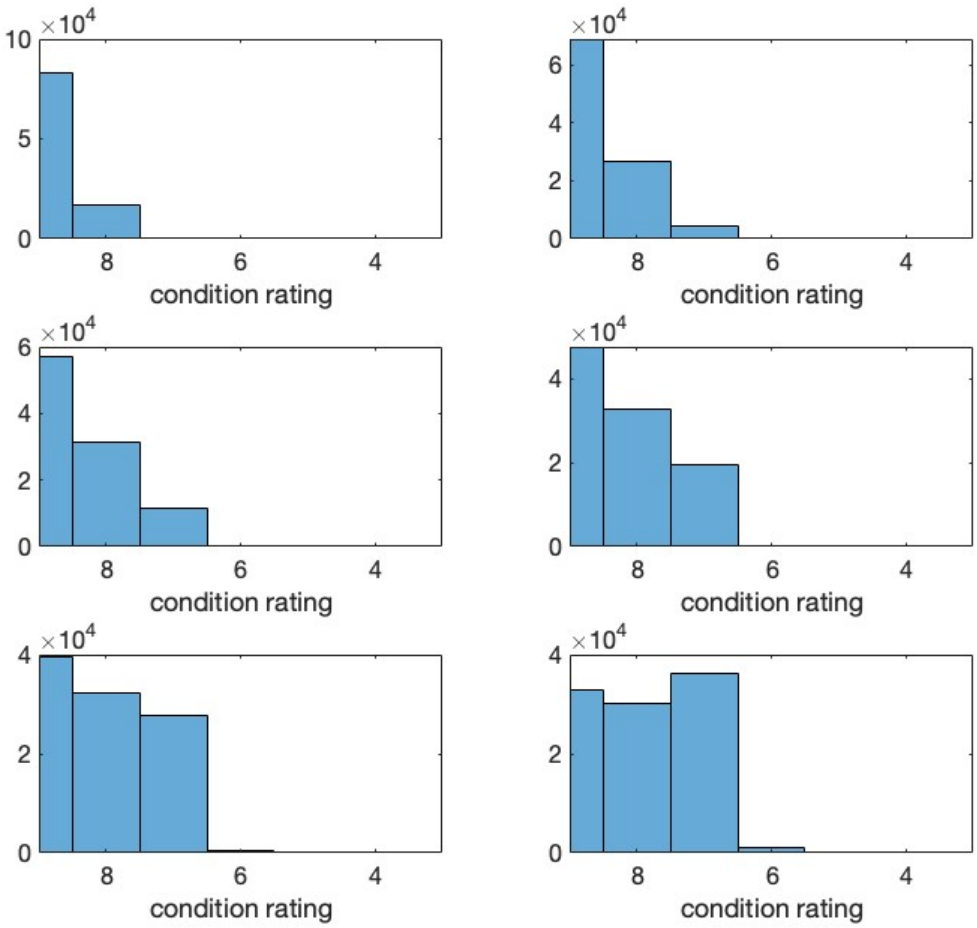


Figure 3.6: Histogram of RNO sample data

Chapter 4: Case Study 1

4.1 Data Selection and Preparation

This study focused on bridges in Texas. Texas bridge concrete deck data from 2008 and 2010 were collected from the National Bridge Inventory (NBI) database to demonstrate the development of a framework for proposed deterioration modeling.The total numbers of bridges are 58,709 and 51,454 respectively. The bridges are inspected every two years. It assumes the condition ratings of bridge components are the same for two years.

4.1.1 Collection

Jackson et al. (2017) classified the winter weather across Texas into four zones based on the frequency of snow and ice. Figure 4.1 shows the zones with the number of freezing days. In Zone 1, 23 or more freezing days annually with frequent snow and rare ice. Zone 2 has 15 to 22 freezing days with occasional snow and ice. Zone 3 has 6 to 14 freezing days with infrequent snow and occasional ice. In Zone 4, 5 or fewer freezing days with very occasional snow and ice. Deterioration models are developed for each zone, and Texas bridge deck data are collected corresponding to the zone from the NBI database using item 3 (county code). Texas consists of 254 counties, of which 48, 105, 81, and 20 are in Zone 1, 2, 3, and 4, respectively.

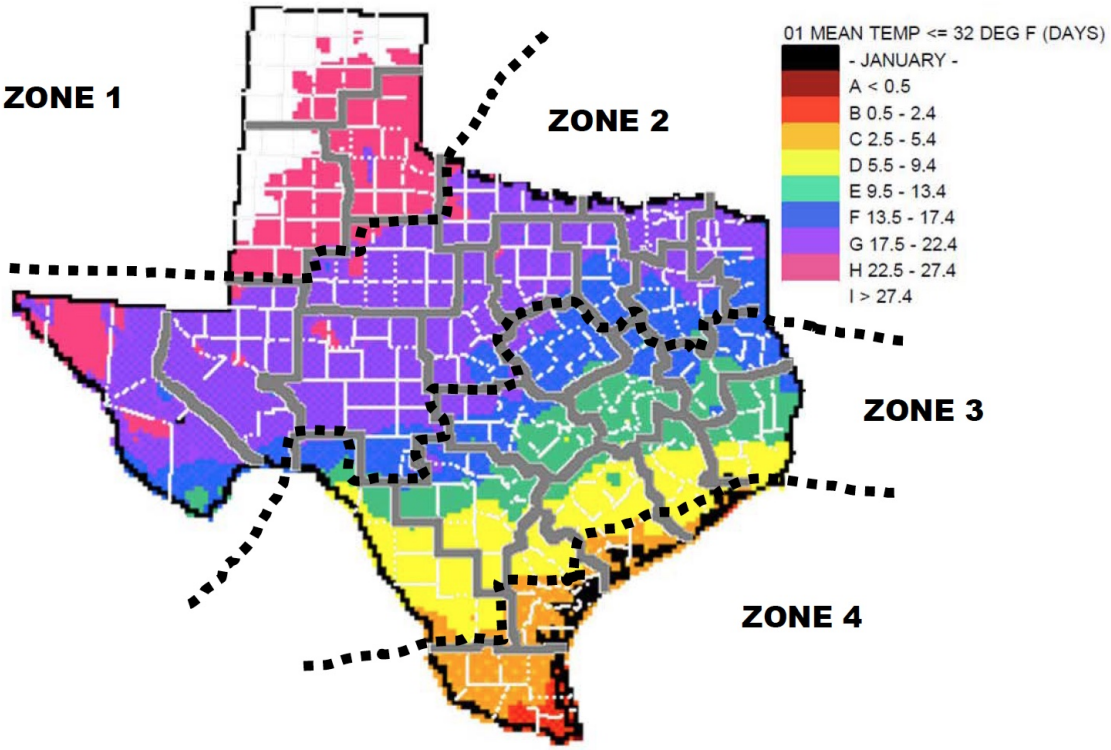
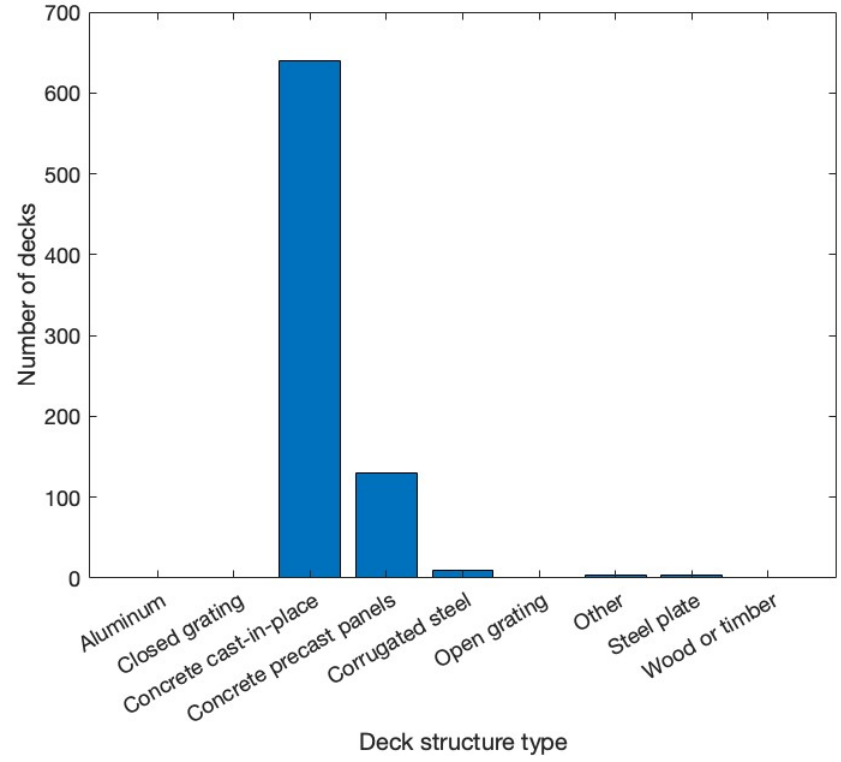


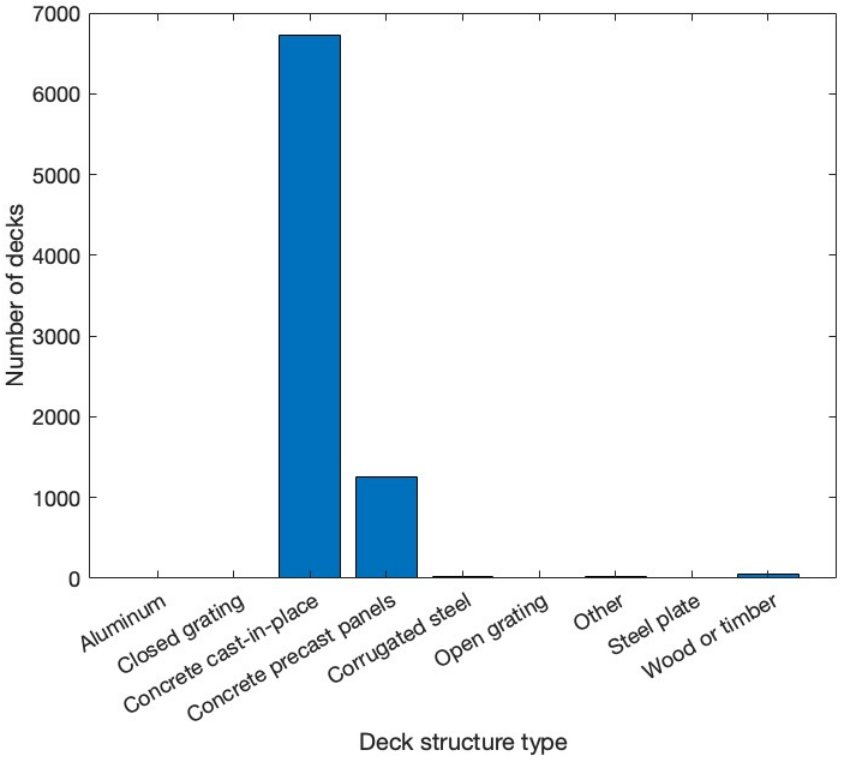
Figure 4.1: Texas winter weather classification by zone (Jackson et al., 2017)

4.1.2 Grouping

The collected bridge decks from the NBI database are grouped according to the explanatory variables. Figure 4.2 contains the number of decks in 2010 according to the deck structure type in zone 1 (773), zone 2 (8086), zone 3 (7689), and zone 4 (1630). This research uses ”concrete cast-in-place”, ”concrete precast panels”, and ”other” as sample data. The sample data were divided into four zones according to winter snowfall.

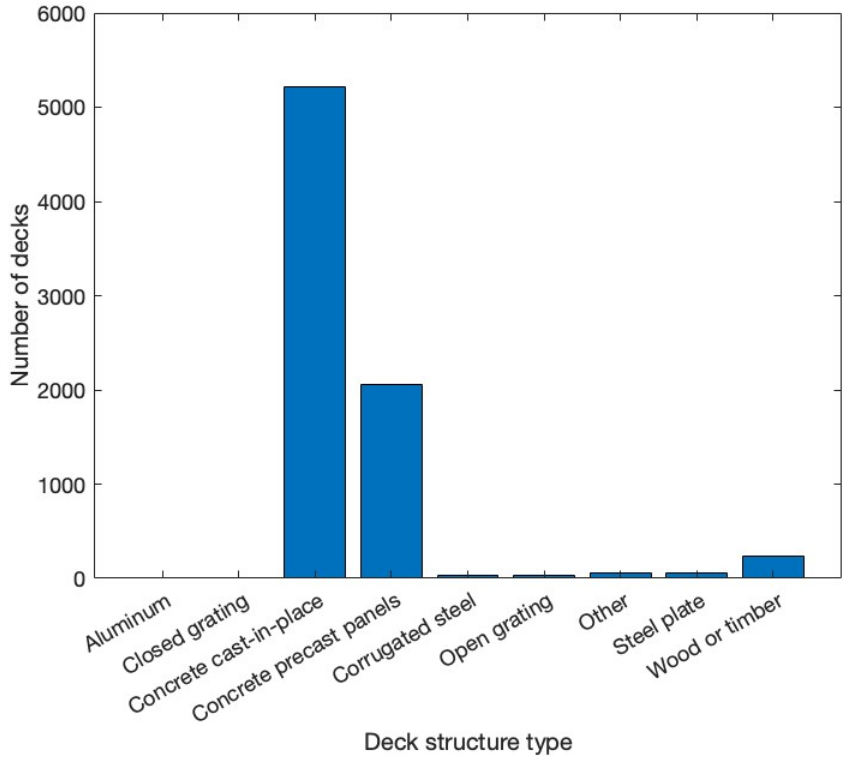


(a) Zone 1

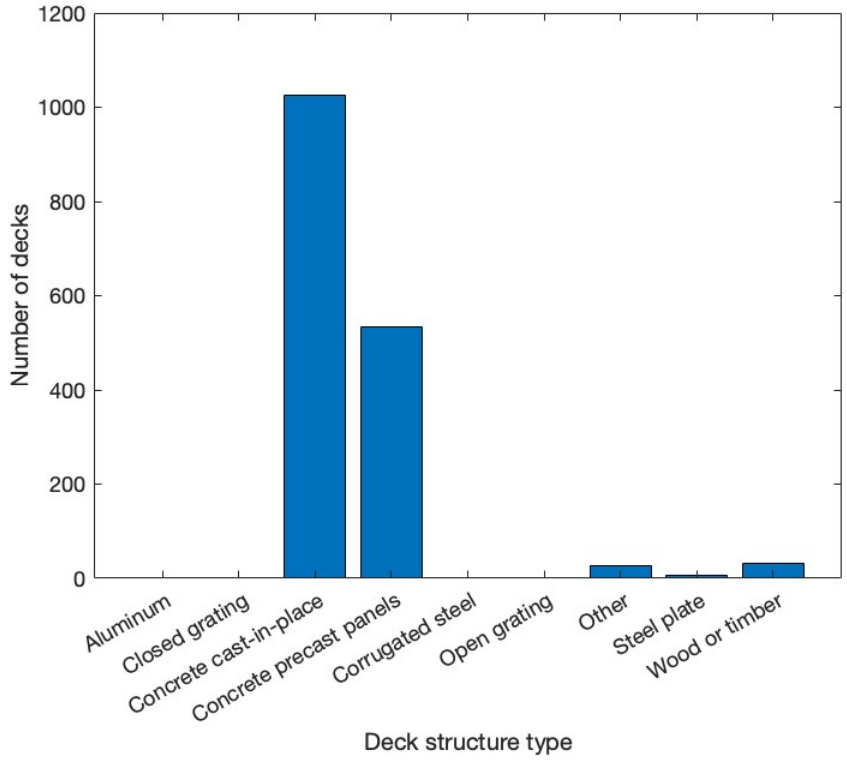


(b) Zone 2

Figure 4.2: Number of Decks according to deck structure type and zone



(c) Zone 3



(d) Zone 4

Figure 4.2: Number of Decks according to deck structure type and zone (cont.)

4.1.3 Screening

The sample data set includes bridges only recorded in both 2008 and 2010. In addition, bridges are also included in the data set if bridges were recorded only in 2010 but built in 2008, 2009, and 2010 and rated CR 9. The collected data are checked for discrepancies by comparing 2008 and 2010. Table 4.1 summarizes the number of decks corresponding to the type of discrepancy. The location discrepancy defines that the structure identity number is identical, but latitude or/and longitude are not matched. For example, the locations of 7, 487, 118, and 27 bridges recorded the same structure ID number of zone 1, 2, 3, and 4, respectively, in 2008 and 2010 are different.

Table 4.1: Number of discrepancies in the sample data by zone in Texas

Type of Discrepancy	**Zone 1**	**Zone 2**	**Zone 3**	**Zone 4**
Location	7	487	118	27
Deck structure type	1	134	245	83
Year_built	0	21	31	4

4.1.4 Eliminating Outliers

The collected deck data are divided into five groups according to condition rating from 9 to 5 because there are no bridges rated below 5 in the data set. The data of each group are analyzed. Figure 4.3 shows the boxplot and number of bridges of zone 1 grouped by condition rating (CR). In Figure 4.3(a), a box is the age range between the 25th and 75th quartiles. A red line in the box is the median of ages. Figure 4.3(b) shows the distribution of the decks of condition ratings 9, 8, 7, 6, and 5. The boxplot and number of bridges of each CR in the other zones are in Appendix B.

Table 4.2 contains the total number of decks of samples in each zone. Zone 1 has the smallest sample size, and Zone 2 has the largest sample size.

Table 4.2: Number of sample bridge decks by zone in Texas

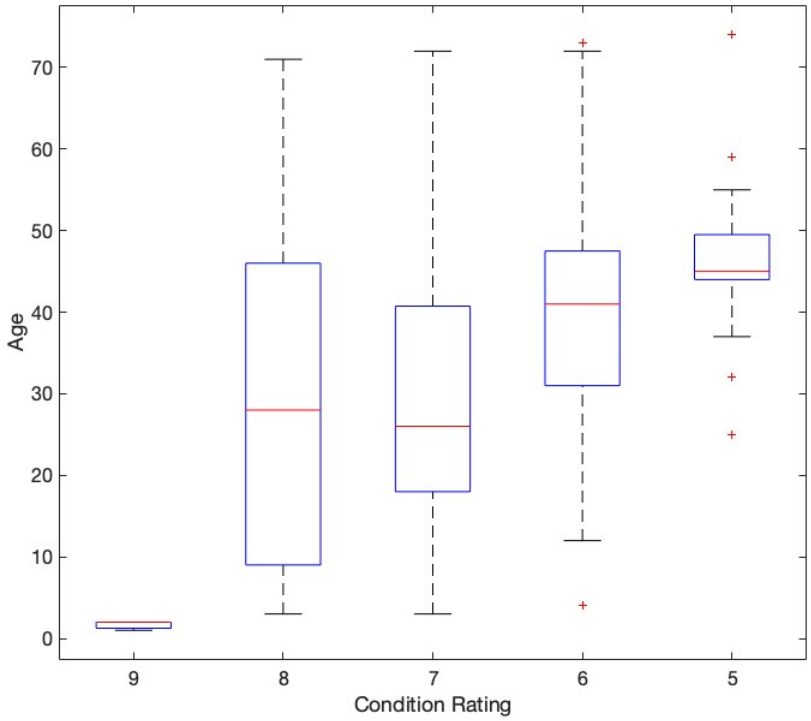
Total number of decks	**Zone 1**	**Zone 2**	**Zone 3**	**Zone 4**
After grouping	773	8,004	7,327	1,589
After screening	699	6,546	6,050	1,296
After eliminating outliers	692	6,446	5,989	1,292

Table 4.3 summarizes the statistical information of the boxplot of Zone 1. The data points away from an interquartile range which is the difference between 25th and 75th more than 1.5 times are considered outliers. The outliers at CR 5 are ages greater than 59 and less than 32. The tables for the other zones are in Appendix B.

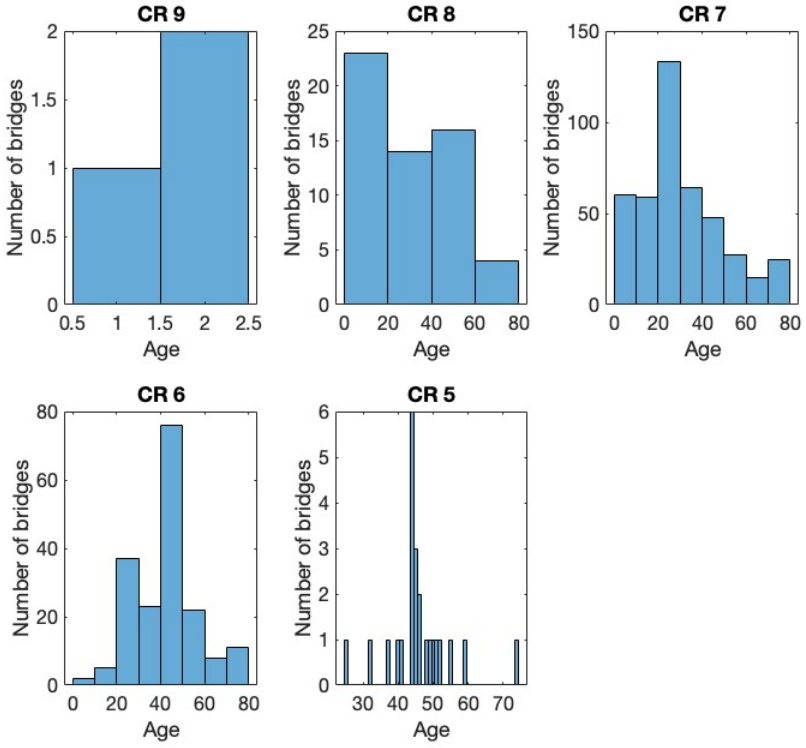
Table 4.3: Summary of boxplot results for outlier identification

	CR 9	**CR 8**	**CR 7**	**CR 6**	**CR 5**
Maximum age	2	71	72	72	74
Minimum age	1	3	3	4	25
Number of outliers	0	0	0	3	4
Outliers (age)				73, 4	>=59, <=32
Upper adjacent (age)		71	72	72	55
75th percentile (age)		46	40.75	47.5	49.5
Median (age)	2	28	26	41	45
25th percentile (age)	1.25	9	18	31	44
Lower adjacent (age)	1	3	3	12	37
Interquartile range	0.75	37	22.75	16.5	5.5

Figure 4.4 is the histogram of deck samples corresponding to age after eliminating outliers. The x-axis is the age from 0 to 80 years, and the y-axis is the number of decks from 0 to 600. Figure 4.5 is the histogram of samples corresponding to condition rating. The y-axis is the number of decks from 0 to 4000.

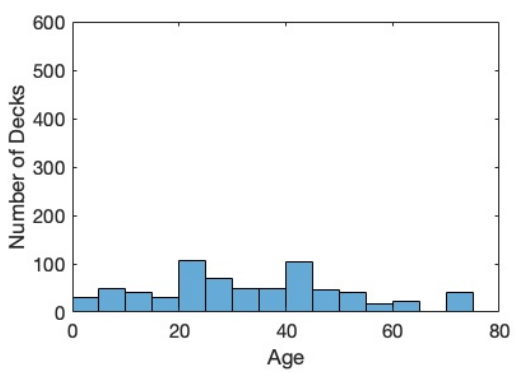


(a) Boxplot showing outliers

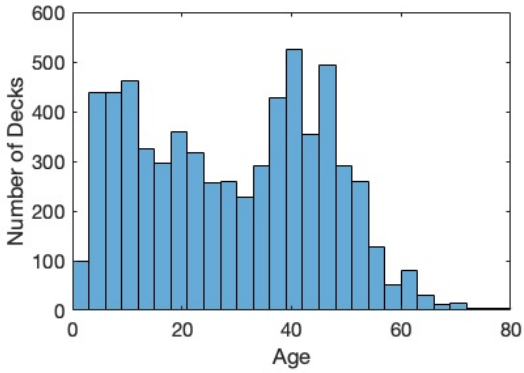


(b) Number of bridges versus age according to condition rating group

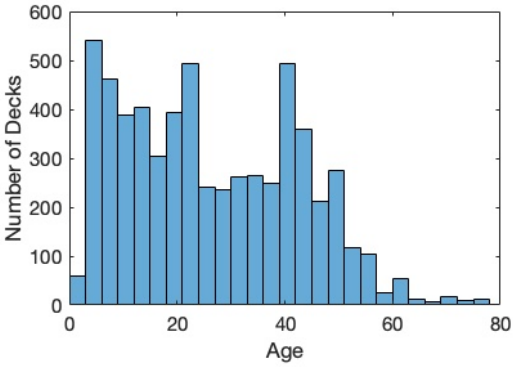
Figure 4.3: Statistical summary of collected deck data of zone 1 before eliminating



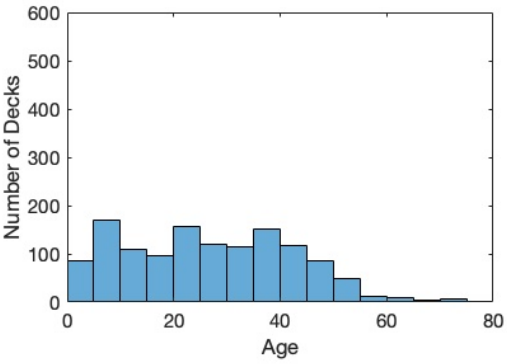
(a) Zone 1



(b) Zone 2

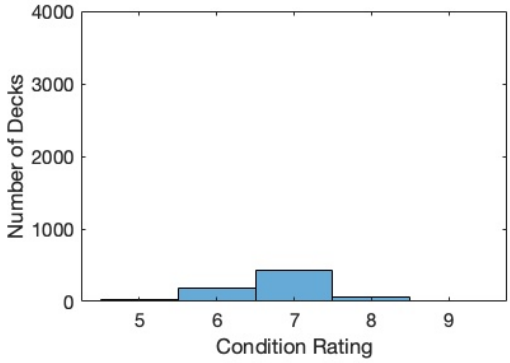


(c) Zone 3

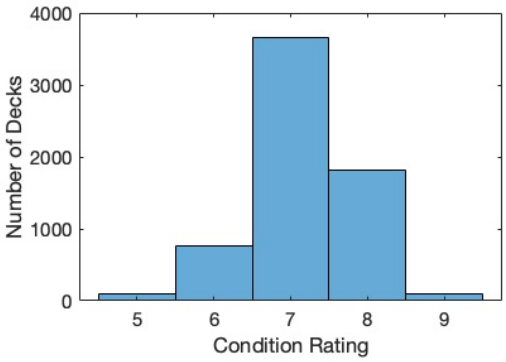


(d) Zone 4

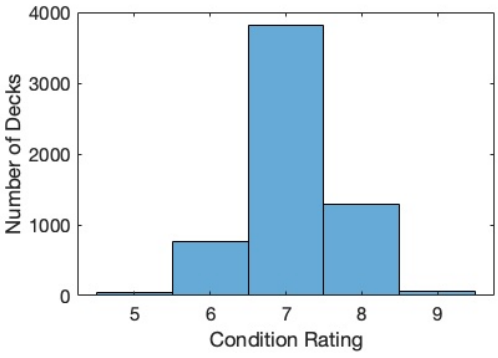
Figure 4.4: Distribution of bridge decks by age and zone



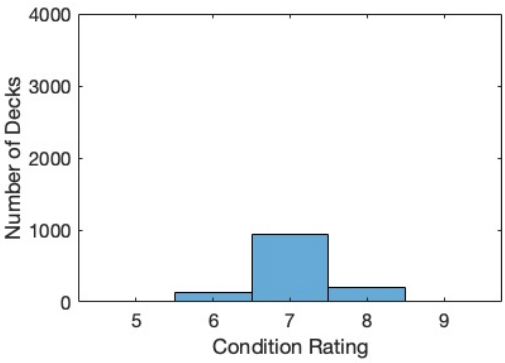
(a) Zone 1



(b) Zone 2



(c) Zone 3



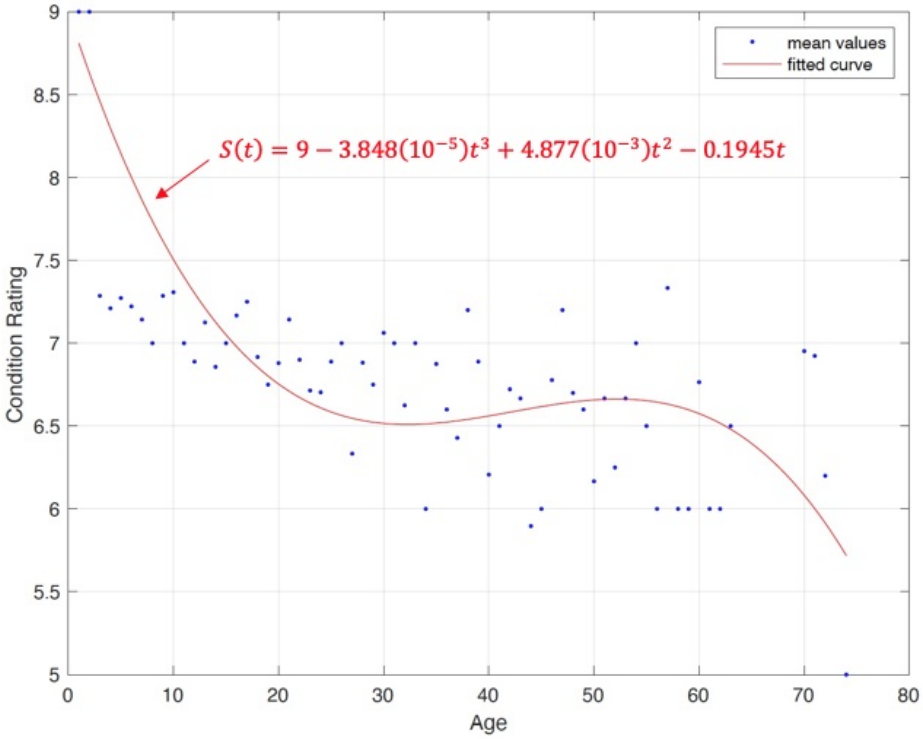
(d) Zone 4

Figure 4.5: Distribution of bridge decks by condition rating and zone

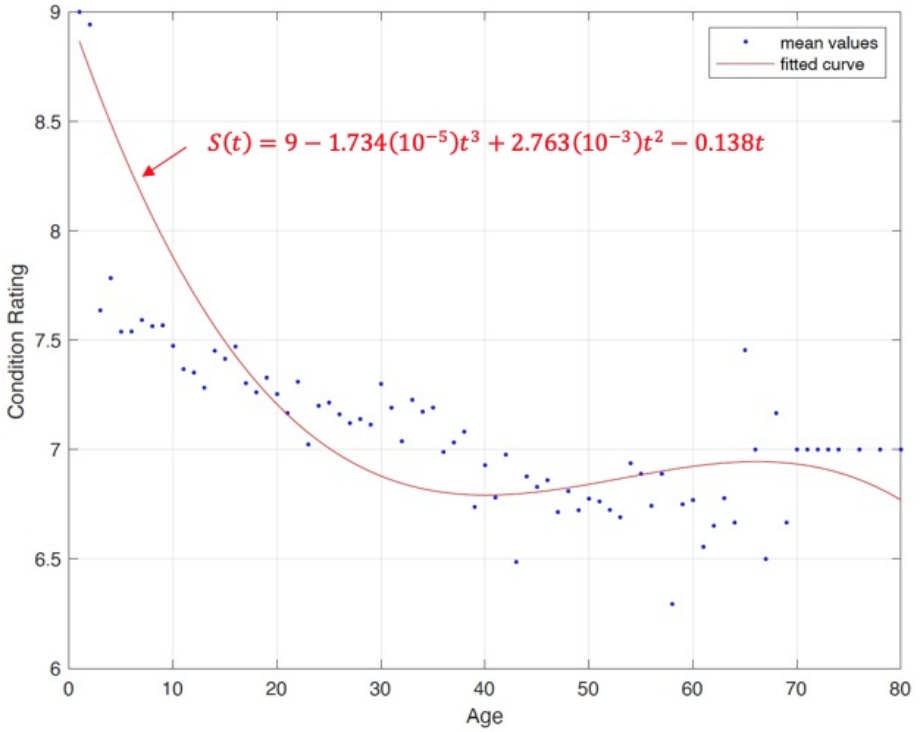
4.2 Single Model Approach

4.2.1 Regression Nonlinear Optimization (RNO)

Using the curve fitting tool in Matlab (2022b), coefficients of a third polynomial equation, $S(t)$, for the sample data are estimated. Figure 4.6 shows the mean values of the sample data (blue dots) and a curve (red line) obtained by the curve fitting method.

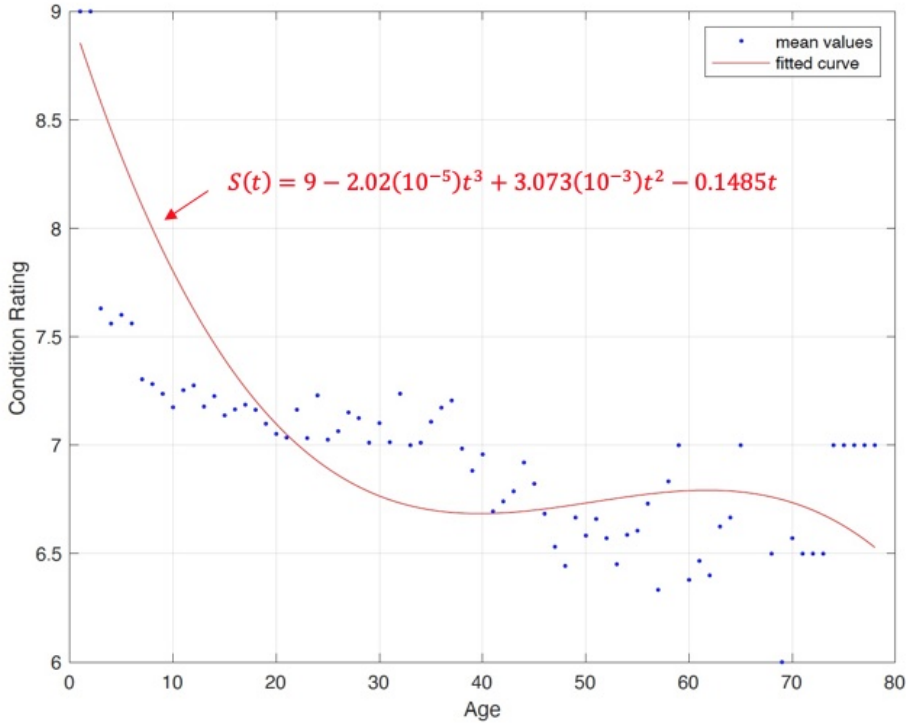


(a) Zone 1

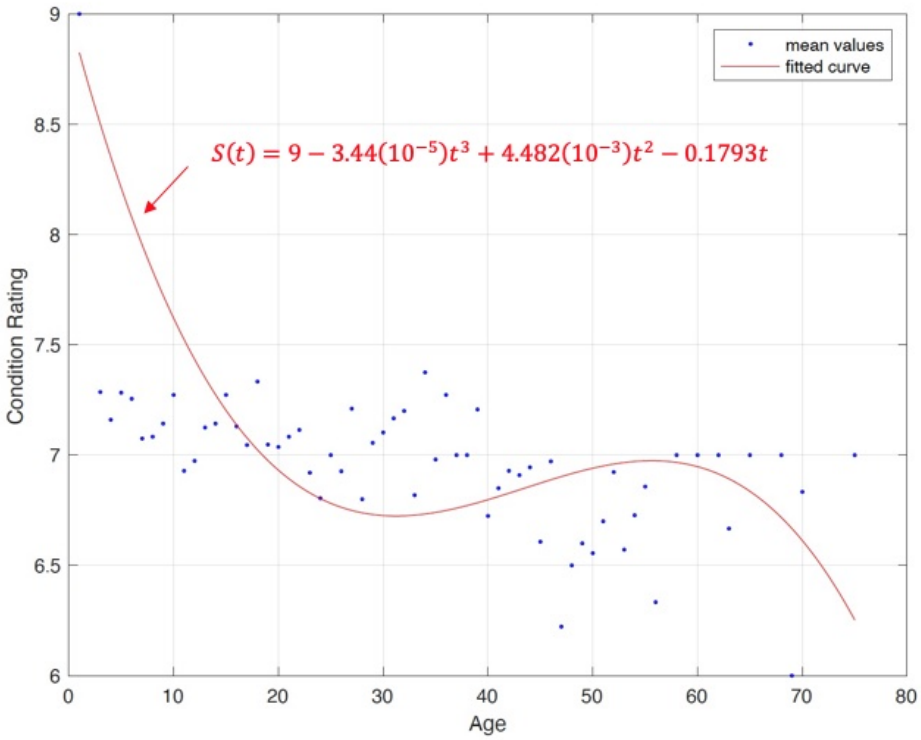


(b) Zone 2

Figure 4.6: Regression curves for RNO approach



(c) Zone 3



(d) Zone 4

Figure 4.6: Regression curves for RNO approach (cont.)

4.2.2 Bayesian Maximum Likelihood (BML)

The sample data are organized by rows and columns. For example, In Table 4.4, the first column (Age) is age from 1 to the oldest in the sample data. The columns, CR 9, 8, 7, 6, 5, and 4, contain the number of decks rated from condition rating 9 to 4 at a given age. The TPM presents in the result section.

Table 4.4: Sample data structured for implementation in BML approach

Age	CR 9	CR 8	CR 7	CR 6	CR 5	CR 4
0	0	0	0	0	0	0
1	1	0	0	0	0	0
2	2	0	0	0	0	0
3	0	2	5	0	0	0
4	0	4	15	0	0	0
⋮	⋮	⋮	⋮	⋮	⋮	⋮
40	0	0	7	21	1	0
⋮	⋮	⋮	⋮	⋮	⋮	⋮
74	0	0	0	0	1	0

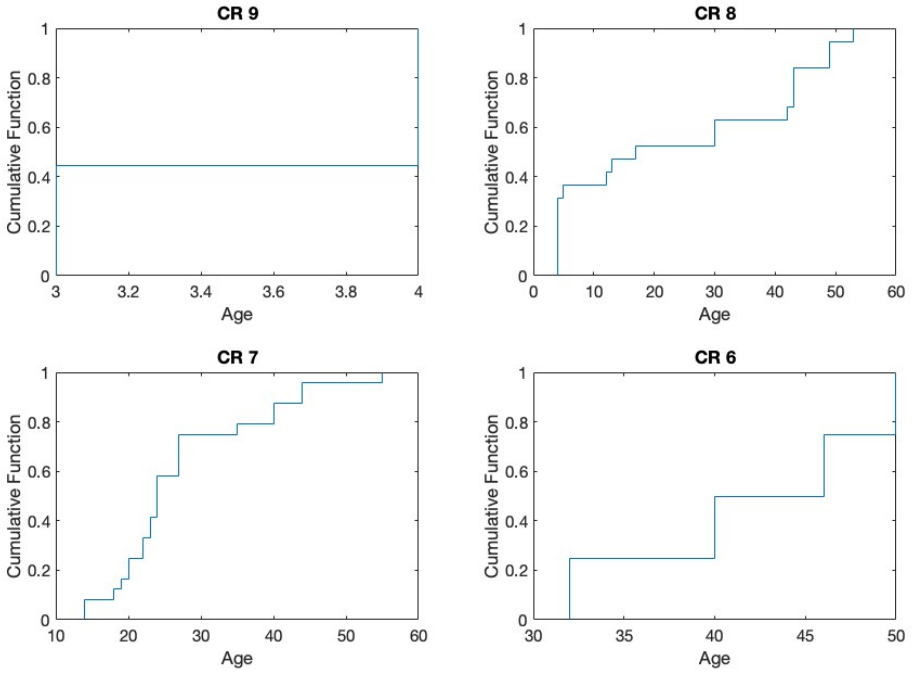
4.2.3 Ordered Probit Model (OPM), Poisson Regression (PR), and Negative Binomial Regression (NBR)

The sample data are divided into five groups from CR 9 to 5. The decks of each group had the same condition rating as the 2008. The TPMs are presented in the result section.

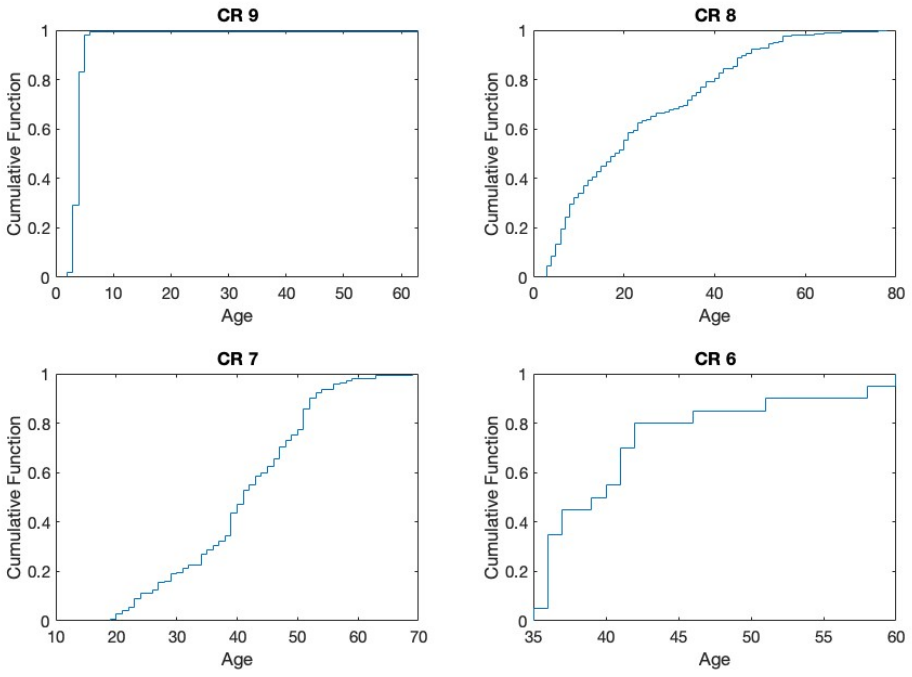
4.2.4 Proportional Hazard Model (PHM)

A baseline probability of each group can be estimated in Figure 4.7 from CR 9 to CR 6. In CR 5, all components are rated at CR 5.

Table 4.5 summarizes the baseline probabilities (Baseline) and hazard ratios (HR) of zone 1, 2, 3, and 4 corresponding to the group. The HR values are greater than 1 in CR 9 group of all zones.

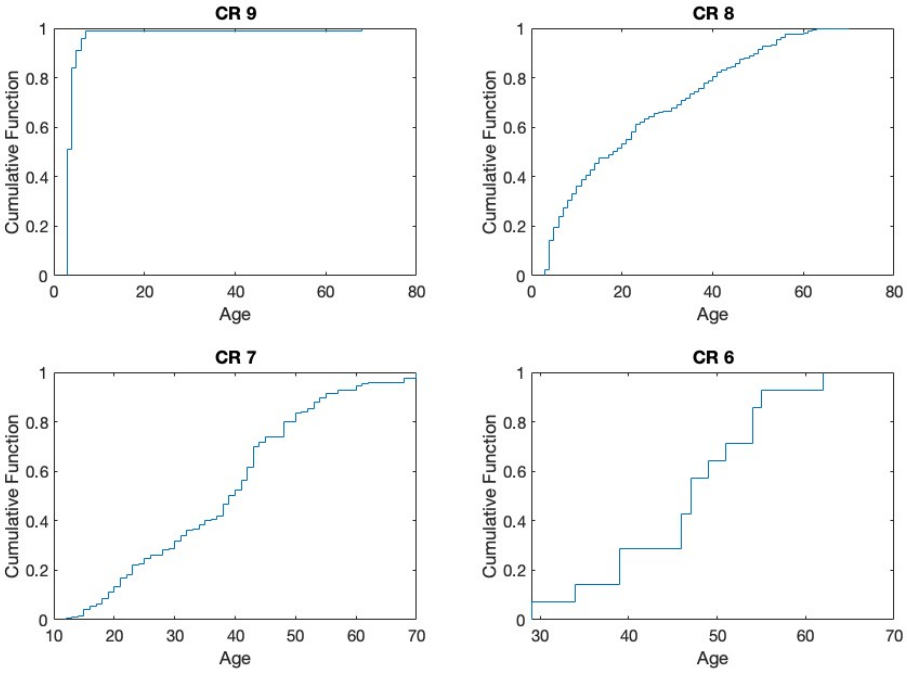


(a) Zone 1

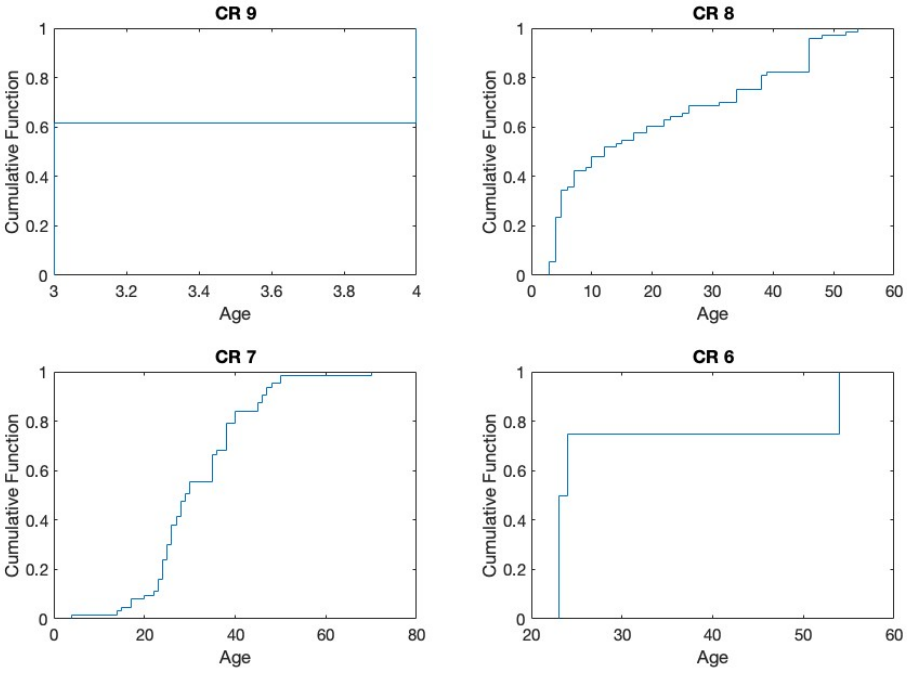


(b) Zone 2

Figure 4.7: Empirical cumulative distribution functions of condition rating groups of each zone for PHM approach



(c) Zone 3



(d) Zone 4

Figure 4.7: Empirical cumulative distribution functions of condition rating groups of each zone for PHM approach (cont.)

Table 4.5: Baseline probabilities and hazard ratios (HRs)

Group	Zone 1		Zone 2		Zone 3		Zone 4	
	HR	Baseline	HR	Baseline	HR	Baseline	HR	Baseline
CR 9	2.6800	0.5556	5.4911	0.9787	3.5226	0.4900	1.6459	0.3846
CR 8	0.7302	0.6842	1.1763	0.9530	1.5410	0.9780	1.0828	0.9452
CR 7	0.7577	0.9167	1.4526	0.9930	1.7074	0.9944	1.1435	0.9841
CR 6	0.8212	0.7500	0.6804	0.9500	1.3278	0.9286	0.7412	0.5000
CR 5	1	1	1	1	1	1	1	1

4.2.5 Time-based Weibull Model

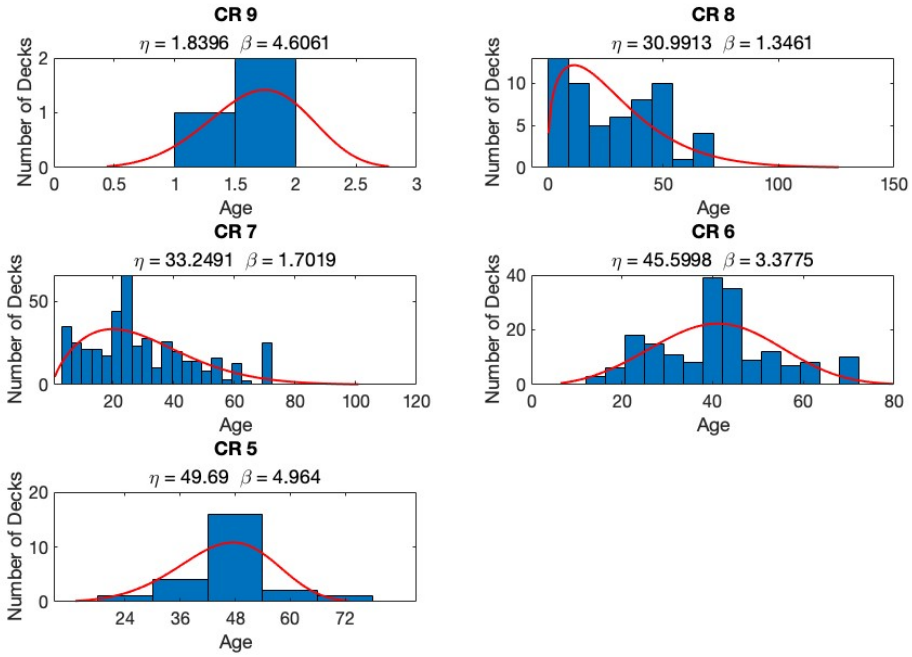
The mean duration and cumulative time of all zones summarized in Table 4.6 are obtained using Equation 2.20 with the parameters of condition rating groups shown in Figure 4.8. The histograms are the numbers of components according to age and the red lines are the Weibull functions. The blue dots in Figure 4.9 are cumulative times.

Table 4.6: Mean duration and cumulative time

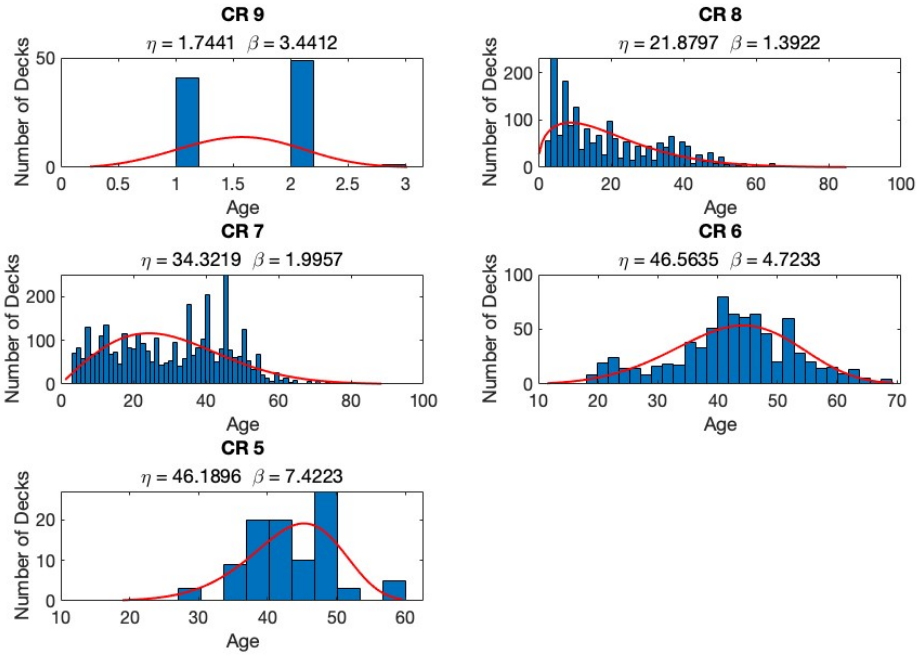
Zone	Measurement	CR 9	CR 8	CR 7	CR 6	CR 5
1	Mean Duration (year)	2	28	30	41	46
	Cumulative Time (year)	2	30	60	101	146
2	Mean Duration (year)	2	20	30	43	43
	Cumulative Time (year)	2	22	52	95	138
3	Mean Duration (year)	2	17	26	41	46
	Cumulative Time (year)	2	19	45	86	131
4	Mean Duration (year)	1	21	26	35	35
	Cumulative Time (year)	1	22	48	83	118

Third-order polynomial equations of the mean time of zone 1, 2, 3, and 4, $S(t)$, a red line in Figure 4.9, is obtained using the curve fitting function of the Matlab.

- Zone 1: $S(t) = 9 - 9.124(10^{-7})t^3 + 2.521(10^{-4})t^2 - 0.05097t$
- Zone 2: $S(t) = 9 - 1.131(10^{-6})t^3 + 3.256(10^{-4})t^2 - 0.05241t$
- Zone 3: $S(t) = 9 - 1.465(10^{-6})t^3 + 4.182(10^{-4})t^2 - 0.06016t$
- Zone 4: $S(t) = 9 - 9038(10^{-7})t^3 + 2.503(10^{-4})t^2 - 0.05086t$

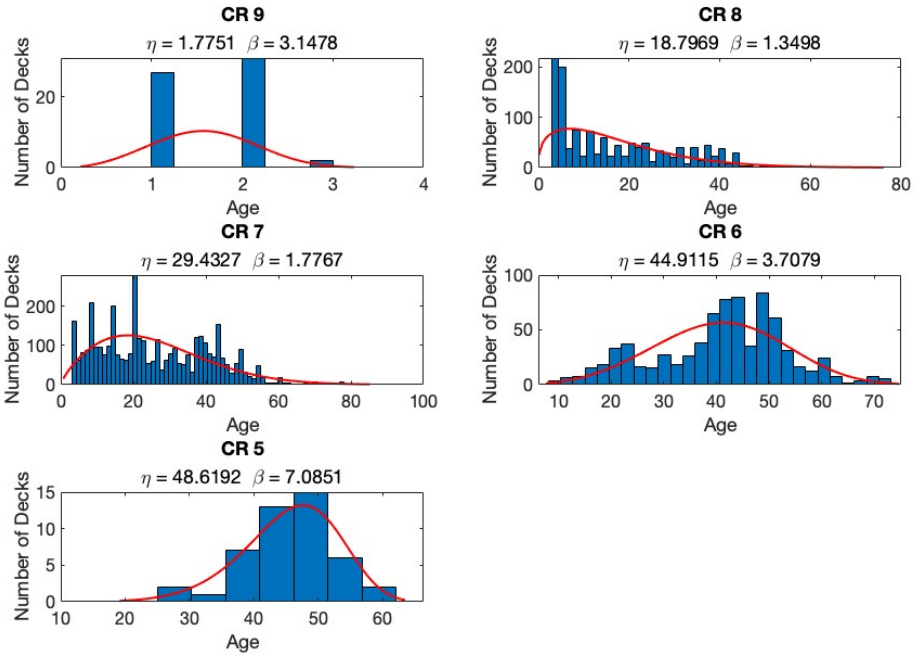


(a) Zone 1

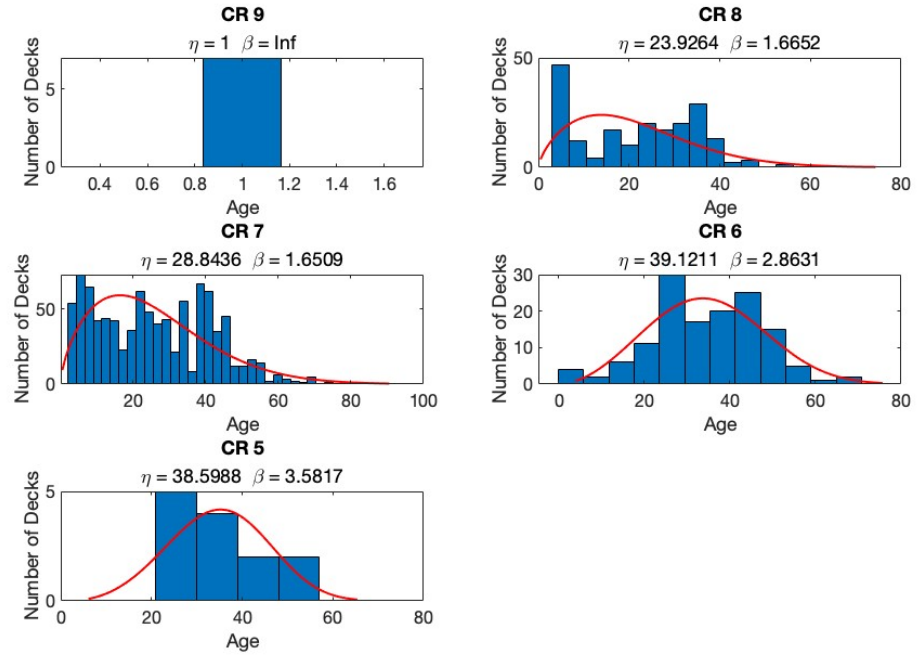


(b) Zone 2

Figure 4.8: Scale (η) and shape (β) parameters of condition rating groups of each zone for Weibull approach

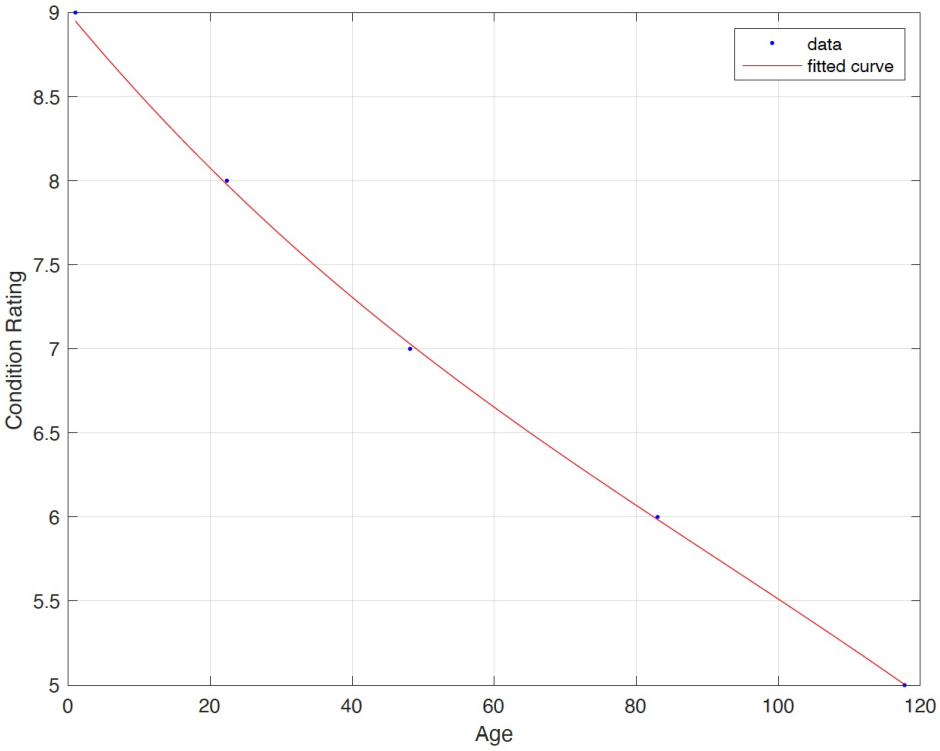


(c) Zone 3

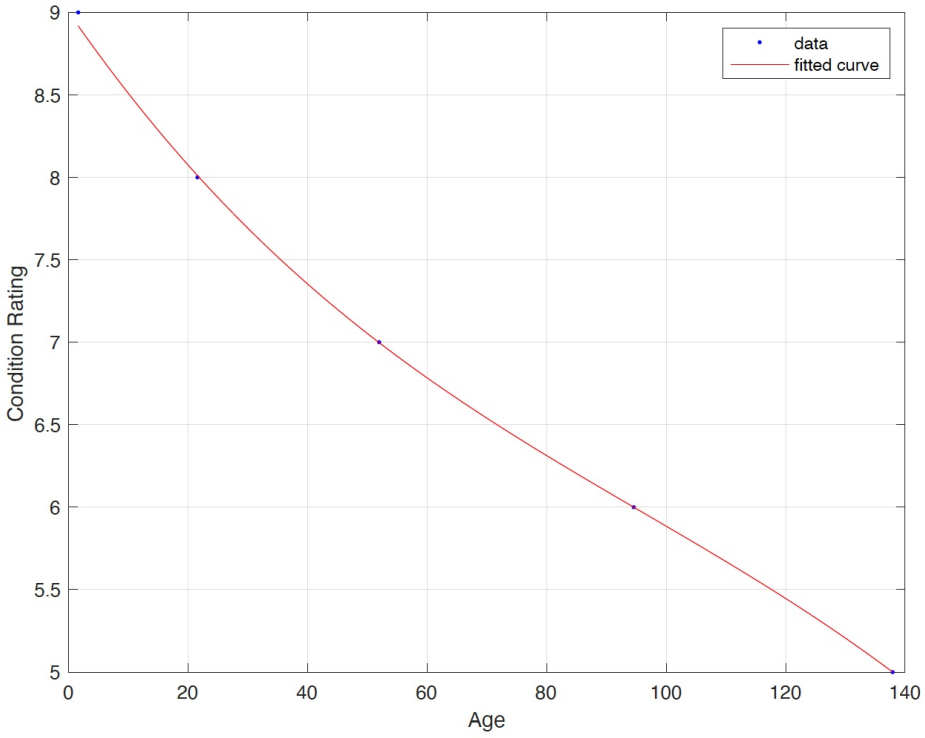


(d) Zone 4

Figure 4.8: Scale (η) and shape (β) parameters of condition rating groups of each zone for Weibull approach (cont.)

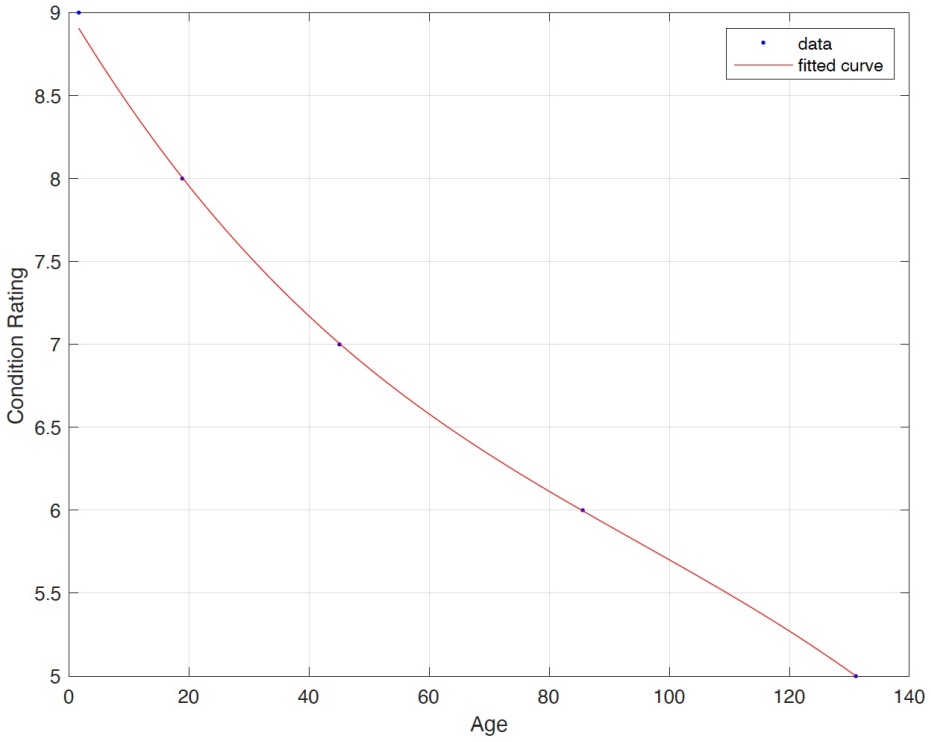


(a) Zone 1

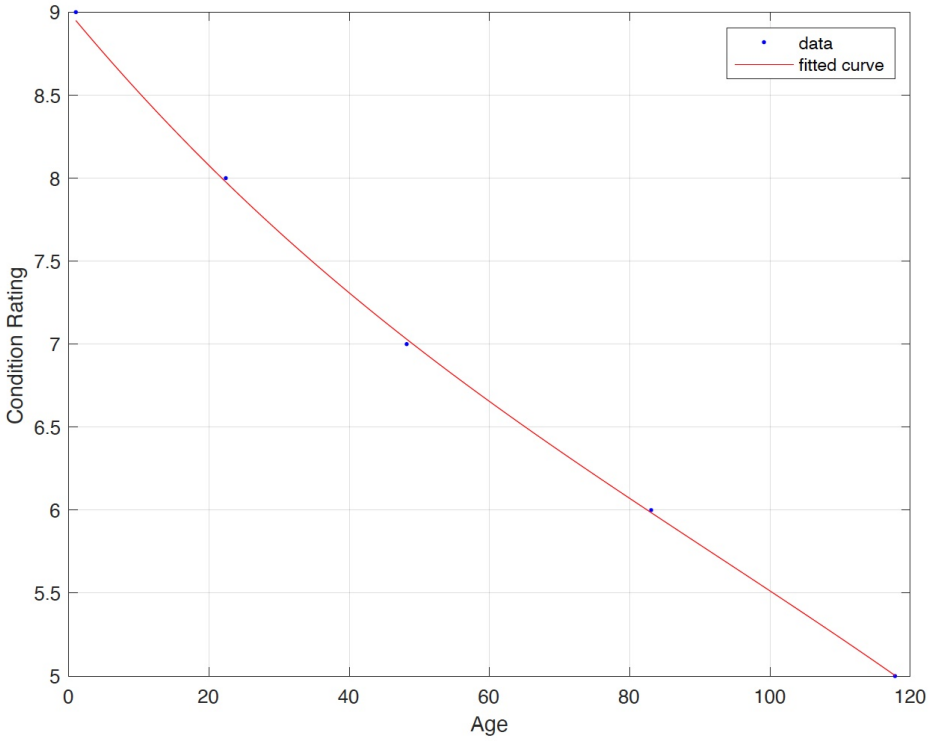


(b) Zone 2

Figure 4.9: Regression curves of mean duration for Weibull approach



(c) Zone 3



(d) Zone 4

Figure 4.9: Regression curves of mean duration for Weibull approach (cont.)

4.2.6 Mechanistic-based Deterioration Model

From Table 3.4, using the data points from CR 9 to CR 4, a regression curve, $S(t) = 9 - 0.0006002t^3 + 0.03589t^2 - 0.6642t$, is generated (the red line in Figure 4.10) using the curve fitting function of the Matlab. Finally, transition probabilities of staying at given condition ratings can be estimated by optimizing the objective function, which is used in the RNO modeling (Equation 2.5).

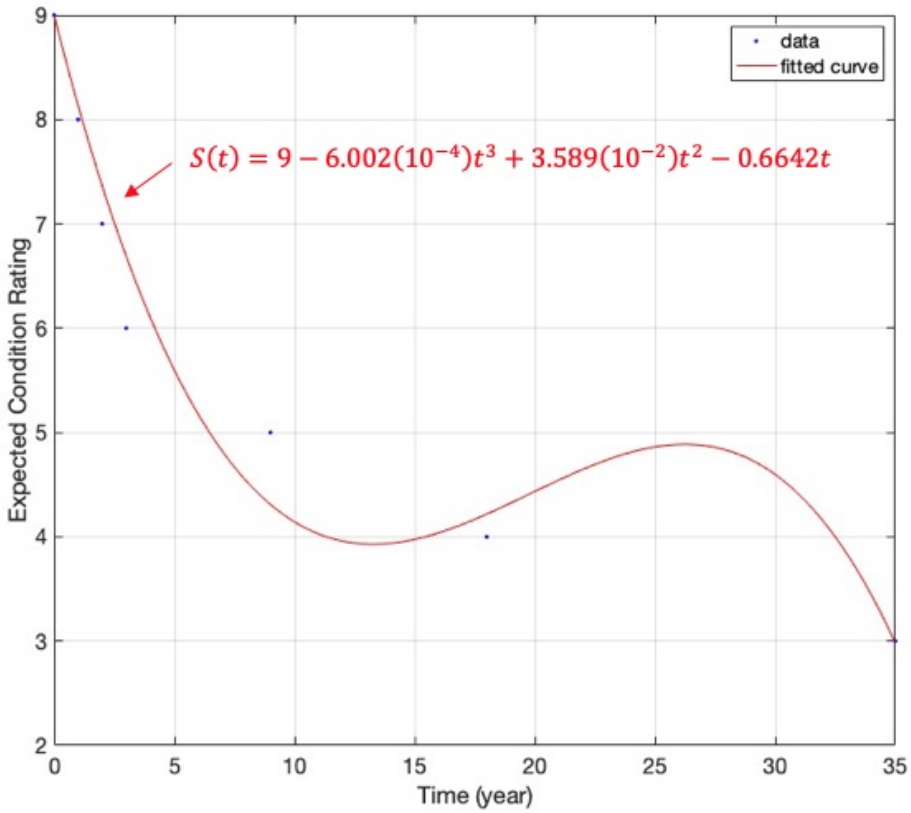


Figure 4.10: Regression curve of defined condition rating given time for Mechanistic-based deterioration modeling

4.2.7 Result

Table 4.7 shows transition probabilities of staying at a particular condition rating by the models and zones. $p_{9,9}$, $p_{8,8}$,..., and $p_{5,5}$ represent the transition probability of staying at condition rating 9, 8, ..., and 5, respectively. The regression nonlinear optimization (RNO), Bayesian maximum likelihood (BML), Weibull deterioration model (WDM), and mechanistic deterioration model (MDM) do not contain a transition probability of one from $p_{9,9}$ to $p_{5,5}$. However, the ordered probit model (OPM), Poisson regression (PR), negative binomial regression (NBR), and proportional hazard model (PHM) include the transition probability of one at $p_{6,6}$ or/and $p_{5,5}$. It means that the condition ratings of decks will stay the same over time when they reach these condition ratings.

Figure 4.11 shows deterioration curves (DCs) of zone 1, 2, 3, and 4. The y-axis is the expected condition rating from 4 to 9, and the x-axis is the age from 0 to 80. The gray circles indicate the distribution of decks of the sample data. The size of circles is different according to the number of decks. Zone 1 contains the curves obtained from all models, including the MDM. The MDM has a higher deterioration rate than the others. The DC of the WDM is linear. Zone 2, 3, and 4 contain the curves from all models except the MDM. The PR has a higher deterioration rate in the three zones.

Table 4.8 summarizes expected condition distributions (unit: percent) of the 2010 year in zone 1, 2, 3, and 4. P_{zone1},P_{zone2},P_{zone3}, and P_{zone4} are the initial state vectors that contain the distribution of condition ratings of 2008 of zone 1, 2, 3, and 4, respectively. The expected condition ratings of 2010 are estimated using Equation 2.2. In zone 1, CR 7 is the highest percentage in all models except the MDM. The CR 6 is the highest, about 73 percent in the MDM. CR 7 is also the highest percentage in all models of zone 2. In zone 3, CR 7 is the highest in all models except the WDM. The CR 6 is the highest, with about 57 percent in the WDM. In zone 4, CR 7 is the highest in all models except the WDM. The CR 5 is the highest, with about 73 percent in the WDM.

$$P_{zone1} = \begin{bmatrix} 0.01724 & 0.10345 & 0.62069 & 0.22989 & 0.02874 & 0 \end{bmatrix}$$

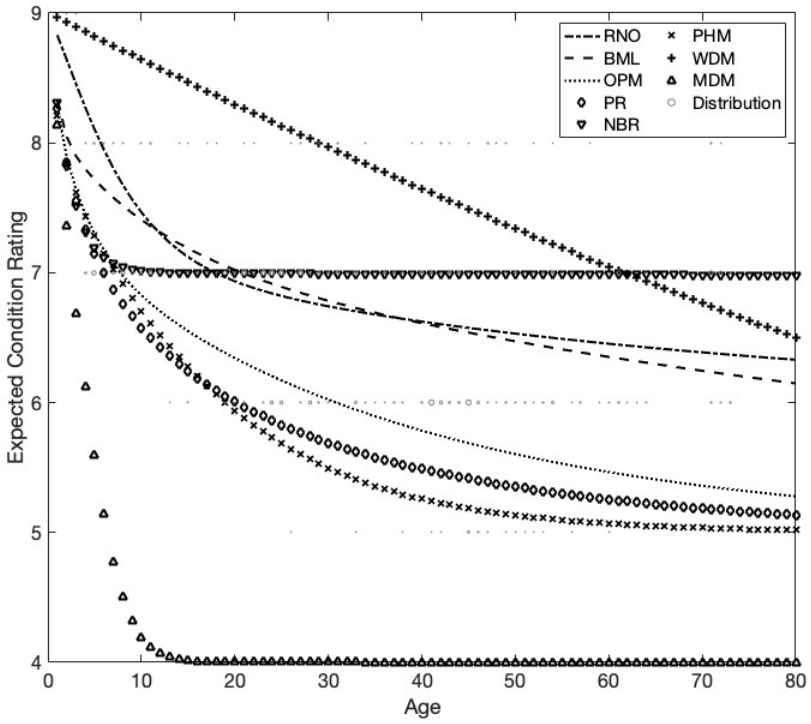
$$P_{zone2} = \begin{bmatrix} 0.03692 & 0.32454 & 0.52823 & 0.09898 & 0.01132 & 0 \end{bmatrix}$$

$$P_{zone3} = \begin{bmatrix} 0.02672 & 0.29621 & 0.57305 & 0.09901 & 0.00501 & 0 \end{bmatrix}$$

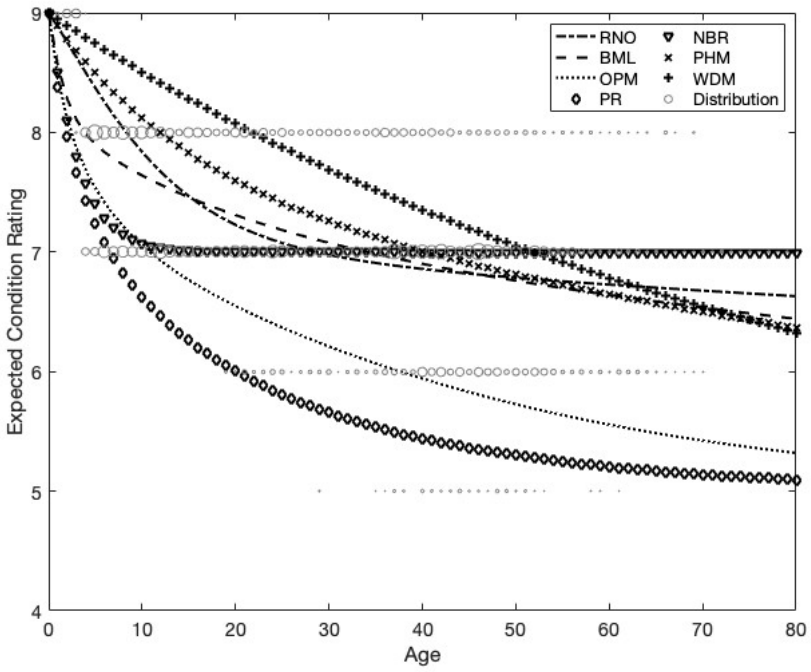
$$P_{zone4} = \begin{bmatrix} 0.02558 & 0.19767 & 0.71163 & 0.06357 & 0.00155 & 0 \end{bmatrix}$$

Table 4.7: Transition probability of staying at current condition rating

	RNO	BML	OPM	PR	NBR	PHM	WDM	MDM
ZONE 1								
$p_{9,9}$	0.8303	0.3321	0.2807	0.2692	0.3075	0.2070	0.9637	0.1323
$p_{8,8}$	0.7336	0.9096	0.7465	0.6905	0.6119	0.7580	0.9667	0.2464
$p_{7,7}$	0.9843	0.9827	0.9444	0.8990	0.9998	0.9362	0.9730	0.3697
$p_{6,6}$	**1**	0.9929	0.9750	0.9690	**1**	0.7896	0.9840	0.5503
$p_{5,5}$	0.5856	**1**	**1**	**1**	**1**	**1**	0.5041	0.4512
ZONE 2								
$p_{9,9}$	0.8793	0.5475	0.4499	0.3761	0.4967	0.8886	0.9480	
$p_{8,8}$	0.8287	0.9461	0.8204	0.7164	0.7026	0.9449	0.9598	
$p_{7,7}$	0.9930	0.9891	0.9594	0.8991	0.9999	0.9898	0.9744	
$p_{6,6}$	**1**	0.9925	0.9687	0.9632	**1**	0.9657	0.9913	
$p_{5,5}$	0.4609	**1**	**1**	**1**	**1**	**1**	0.5044	
ZONE 3								
$p_{9,9}$	0.8707	0.4593	0.4040	0.3541	0.4388	0.0810	0.9382	
$p_{8,8}$	0.8073	0.9247	0.6966	0.5920	0.6174	0.9663	0.9764	
$p_{7,7}$	0.9885	0.9892	0.9490	0.8759	0.7381	0.9904	0.4135	
$p_{6,6}$	0.9998	0.9964	0.9764	0.9614	**1**	0.9049	0.9929	
$p_{5,5}$	0.5048	**1**	**1**	**1**	**1**	**1**	0.5104	
ZONE 4								
$p_{9,9}$	0.8413	0.2546	0.2345	0.2494	0.3007	0.2075	0.9470	
$p_{8,8}$	0.7651	0.9145	0.7222	0.6722	0.6304	0.9408	0.9867	
$p_{7,7}$	0.9960	0.9918	0.9384	0.8426	0.8886	0.9819	0.0466	
$p_{6,6}$	0.9999	0.9934	0.9512	0.9378	0.9990	0.5982	0.0080	
$p_{5,5}$	0.5028	**1**	**1**	**1**	**1**	**1**	0.9780	

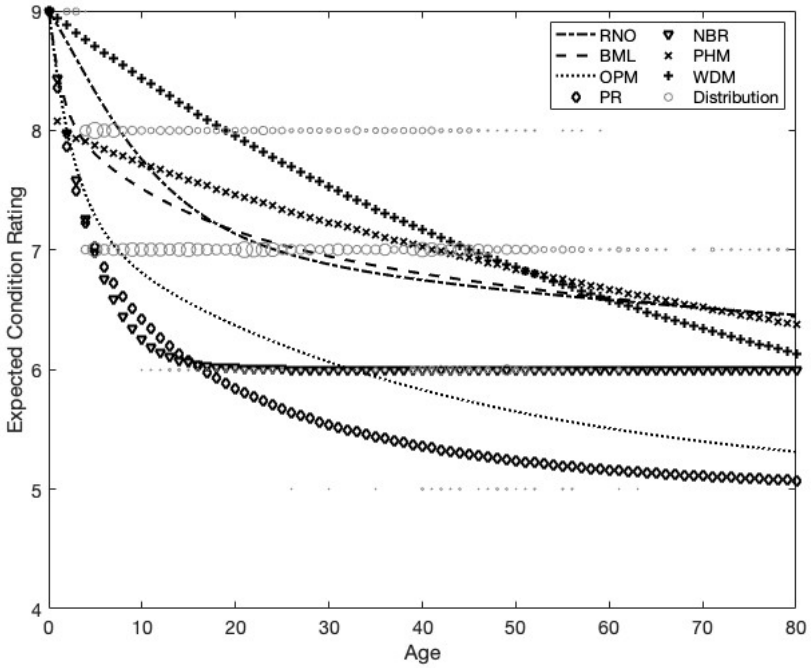


(a) Zone 1

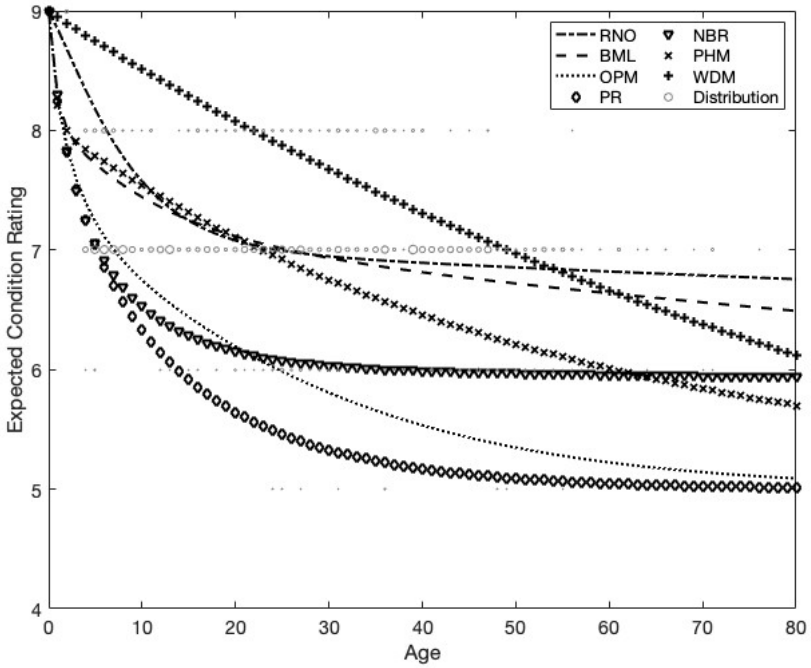


(b) Zone 2

Figure 4.11: Expected condition rating from all models by zones



(c) Zone 3



(d) Zone 4

Figure 4.11: Expected condition rating from all models by zones (cont.)

Table 4.8: Expected condition distribution of 2010 bridge decks by model and zone

Model	CR 9	CR 8	CR 7	CR 6	CR 5	CR 4	Model	CR 9	CR 8	CR 7	CR 6	CR 5	CR 4
Zone 1							**Zone 3**						
RNO	1.2	6.0	65.0	25.0	1.0	1.9	RNO	2.0	19.9	66.3	11.3	0.1	0.4
BML	0.2	10.0	61.8	24.8	3.2	0.0	BML	0.6	27.3	60.5	11.1	0.6	0.0
OPM	0.1	7.0	60.1	28.6	4.1	0.0	OPM	0.4	16.1	66.9	15.5	1.0	0.0
PR	0.1	6.1	55.6	33.6	4.5	0.0	PR	0.3	12.0	62.4	23.7	1.5	0.0
NBR	0.2	5.0	69.0	23.0	2.9	0.0	NBR	0.5	12.9	47.2	39.0	0.5	0.0
PHM	0.1	7.3	59.0	21.3	12.4	0.0	PHM	0.0	30.2	58.2	9.2	2.3	0.0
WDM	1.6	9.8	59.4	25.5	1.3	2.3	WDM	2.4	28.6	10.8	57.4	0.5	0.4
MDM	0.0	1.2	14.4	47.9	28.5	8.0							
Zone 2							**Zone 4**						
RNO	2.9	23.1	62.3	10.7	0.2	0.9	RNO	1.8	12.2	78.9	6.9	0.0	0.1
BML	1.1	31.5	55.2	10.9	1.3	0.0	BML	0.2	18.8	73.4	7.4	0.2	0.0
OPM	0.7	24.4	59.4	13.7	1.8	0.0	OPM	0.1	12.2	72.3	14.4	1.0	0.0
PR	0.5	19.2	58.2	20.0	2.0	0.0	PR	0.2	10.7	61.0	26.6	1.6	0.0
NBR	0.9	18.3	69.8	9.9	1.1	0.0	NBR	0.2	9.5	67.9	22.1	0.2	0.0
PHM	2.9	29.7	55.2	10.3	1.8	0.0	PHM	0.1	19.8	71.0	4.3	4.8	0.0
WDM	3.3	30.3	52.7	12.4	0.4	0.9	WDM	2.3	19.5	0.4	4.0	73.7	0.1

4.2.8 Evaluation

4.3 Evaluation

The mean absolute error (MAE) shown in Equation 4.1 is used to evaluate the results obtained from the models where n is the total number of data points, E_i is an estimated value, and O_i is an observation value.

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|E_i - O_i| \tag{4.1}$$

Table 4.9 contains the MAE values of estimated condition rating and distribution according to the zone. The MAE values of expected condition ratings are calculated to the minimum, mean, and maximum of the sample data. The model which shows the highest accuracy of prediction of future condition rating to ensemble means in sample data may not show the highest accuracy to a minimum and maximum condition rating. For example, in zone 1, the NBR is the smallest value to the mean, the OPM is the smallest respect to the minimum, and the RNO is the smallest respect to the maximum, as indicated by bold text in Table 4.9. The model that shows the highest accuracy of prediction of future condition rating of a component can be different from the model that shows the highest accuracy of prediction of future condition distribution of

components. For instance, the BML is the smallest value to the expected condition distribution. However, it is not the smallest of the minimum, mean, and maximum in the expected condition rating.

This analysis demonstrates the variability in the deterioration model process and inconsistencies in critical deterioration model outputs. A single model cannot represent, quantify, or even acknowledge the variability in CR for a population of bridges at a given age that is observed in the data. As adding explanatory variables reduces the population size, this limitation appears more pronounced.

Table 4.9: Mean Absolute Error (MAE) of all single models by zones

Model	Expected Condition Rating			Expected Condition Distribution	Expected Condition Rating			Expected Condition Distribution
	Minimum	Mean	Maximum		Minimum	Mean	Maximum	
Zone 1					**Zone 3**			
RNO	0.85	0.33	0.70	0.02	1.10	0.24	0.80	**0.01**
BML	0.82	0.31	0.73	**0.01**	1.08	**0.19**	0.81	0.02
OPM	**0.50**	0.89	1.32	**0.01**	0.60	0.93	1.70	0.02
PR	0.53	1.15	1.58	0.03	0.58	1.33	2.10	0.04
NBR	0.98	**0.19**	**0.58**	0.02	**0.41**	0.83	1.60	0.09
PHM	0.56	1.28	1.69	0.03	1.25	0.29	0.66	0.03
WDM	1.88	1.05	0.78	0.02	1.47	0.53	**0.58**	0.18
MDM	1.86	2.77	3.16	0.19				
Zone 2					**Zone 4**			
RNO	1.10	0.21	0.83	0.03	0.99	0.27	0.58	0.02
BML	1.08	**0.19**	0.86	**0.01**	0.93	**0.25**	0.66	**0.01**
OPM	0.63	0.95	1.74	0.02	0.64	1.09	1.70	**0.01**
PR	**0.60**	1.36	2.15	0.03	0.78	1.44	2.05	0.05
NBR	1.00	0.23	0.83	0.04	**0.38**	0.80	1.42	0.04
PHM	1.25	0.31	0.75	**0.01**	0.82	0.45	0.91	0.03
WDM	1.52	0.54	**0.60**	0.02	1.50	0.70	**0.42**	0.26

4.4 Multiple Model Approach

In zone 1, four models are proposed, the first model using method 1, including all the single models; the second model using method 2, including all the models; the third model using method 1, including all the single models except the mechanistic deterioration model (MDM), and fourth model using method 2, including all the models except the MDM. In zone 2, 3, and 4, two models are proposed, the first model using method 1, including all the single models except the MDM, and the second model using method 2, including all the models except the MDM.

4.4.1 Results

The results include the prediction of the future condition ratings of a bridge deck and the prediction of the future condition distribution of bridge decks of a network system in section 4.2.7. Section 4.4.1.1 is divided into two subsections, methods 1 and 2. Method 1 shows the results of the first model of zone 1, 2, 3, and 4. Method 2 shows the results of the second model of all the zones.

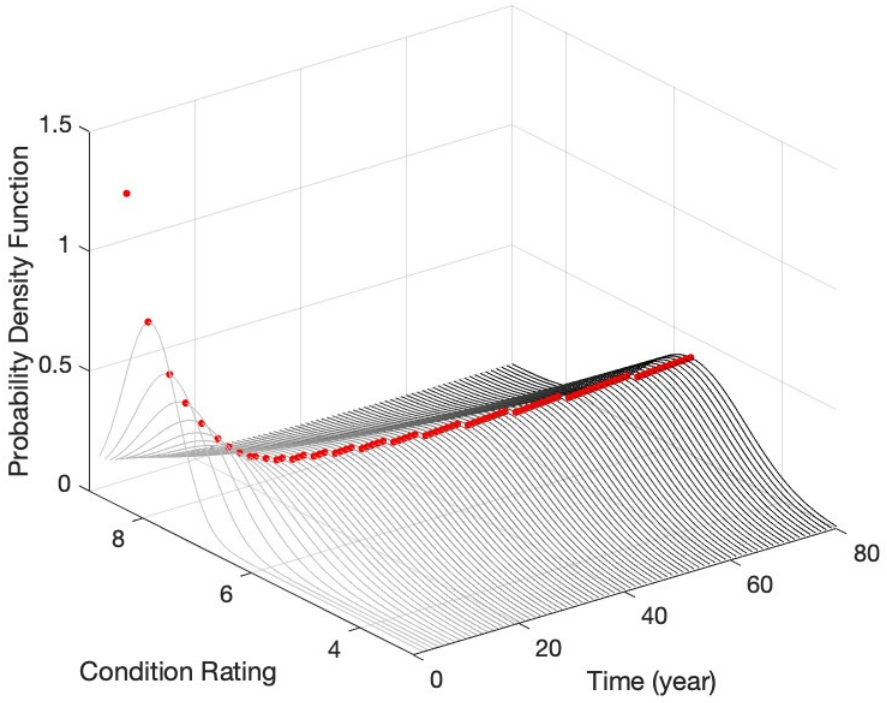
4.4.1.1 Expected Condition Rating

It is a prediction of the future condition rating of a bridge concrete deck. The initial condition rating of the deck is assumed at CR 9 (excellent condition) for a new deck.

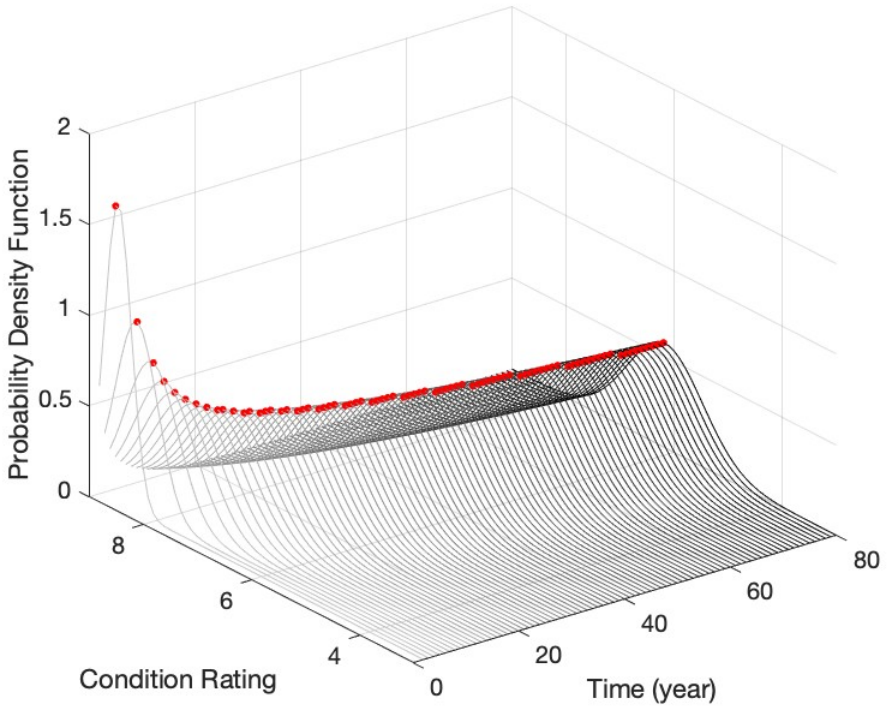
Method 1

Method 1 uses expected condition ratings obtained from the single models to produce a deterioration curve. Figure 4.12 shows 3D deterioration curve plots. The vertical-axis is the time from 0 (front) to 80 year (back). The horizontal-axis is the condition rating (CR) from 3 (right) to 9 (left). The z-axis is the probability density function from 0 (bottom) to 1.4 (top) in zone 1, 3, and 4 and 1.8 (top) in zone 2. The red dots indicate means of the density functions at a given time t. Figure 4.13 presents continuously expected condition ratings (black dots), including all possible CRs (gray dots) at a given time t which is a range of condition ratings. The horizontal-axis is the expected CRs from 3 (bottom) to 9 (top). The vertical-axis is the time

from 0 to 80 years. The deterioration rate is higher in order of zone 1, 4, 3, and 2. However, the difference in the rate between zones is small. The range of possible CRs at a given time is wider in zone 1 than in others. Figure 4.14 contains deterioration curves of all zones. The deterioration rate is higher in order of zone 1, 4, 3, and 2. Figure 4.15 includes the probability of expecting a particular CR at a give time (gray circles). A bigger circle is a higher probability of expecting at the CR. A deck in all zones can be rated CR 8 at 80 years old. A deck in zone 1 can be rated CR 4 from about 10 years old; in zone 3 and 4 can be from 20 years old, and in zone 2 can be from 60 years old.

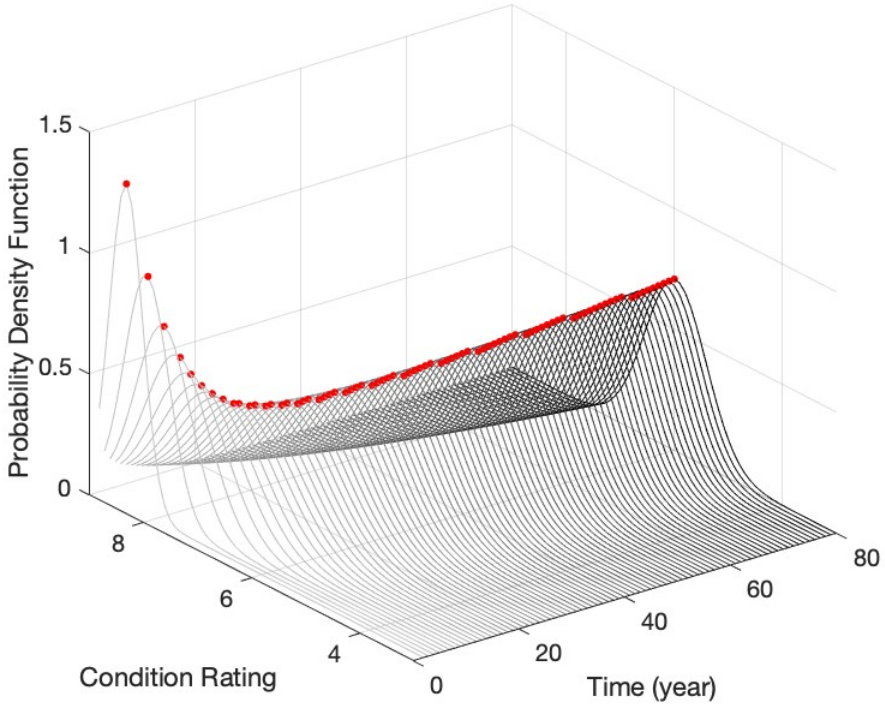


(a) Zone 1

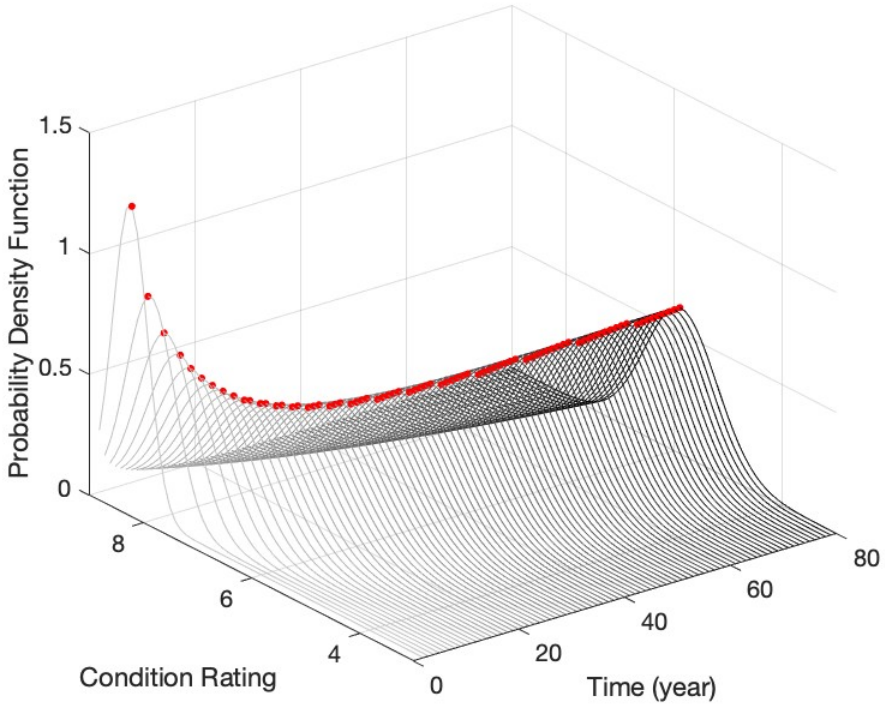


(b) Zone 2

Figure 4.12: 3D deterioration curve plots of the proposed models using Method 1 by zone

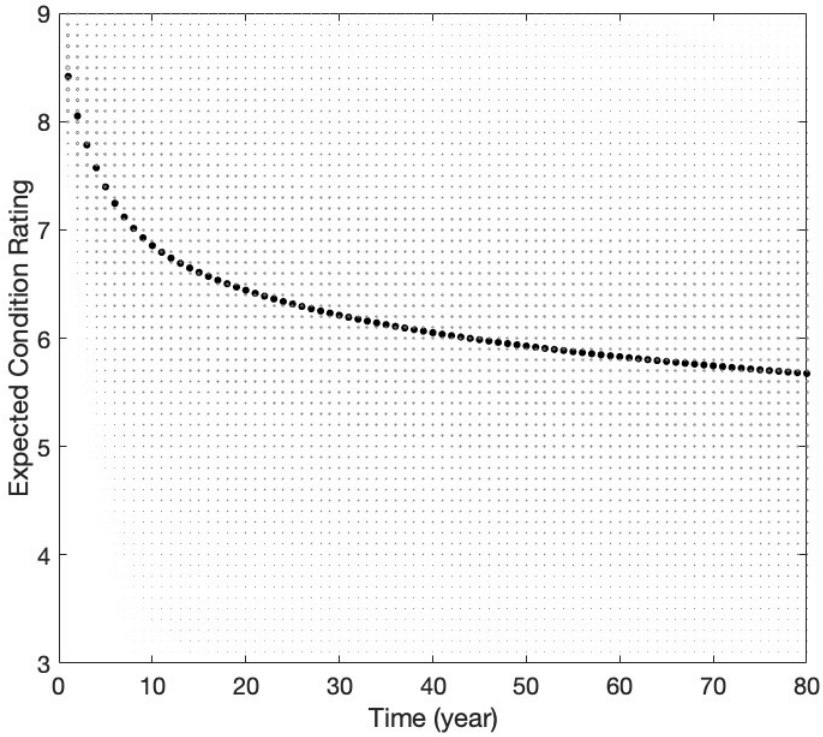


(c) Zone 3

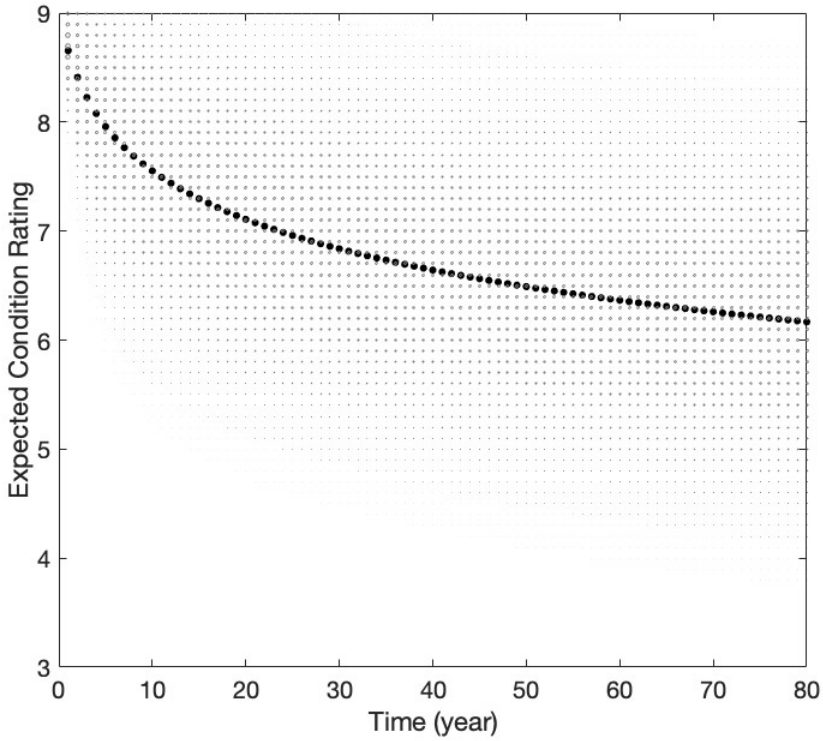


(d) Zone 4

Figure 4.12: 3D deterioration curve plots of the proposed models using Method 1 by zone (cont.)

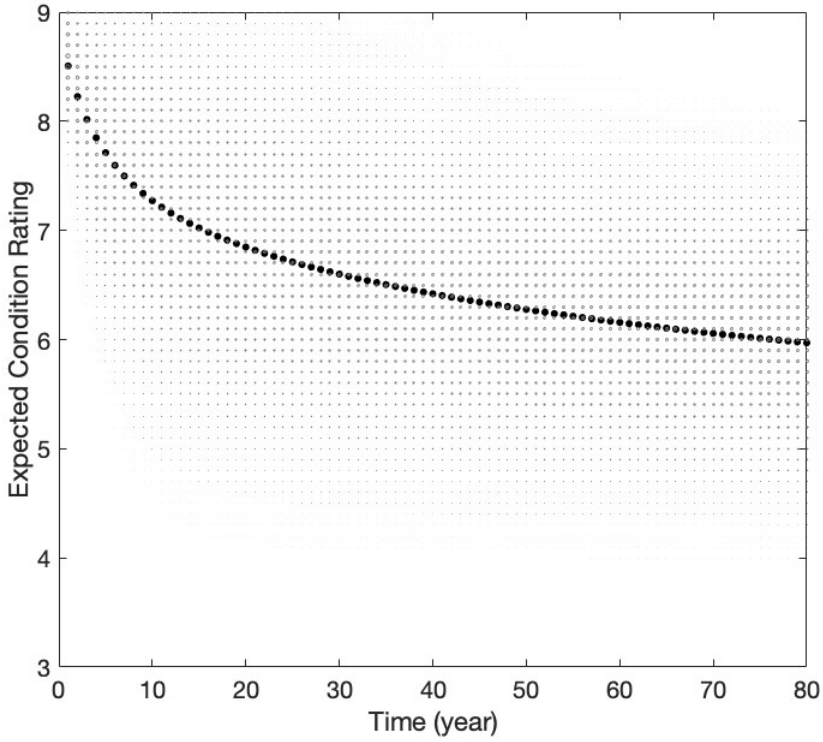


(a) Zone 1

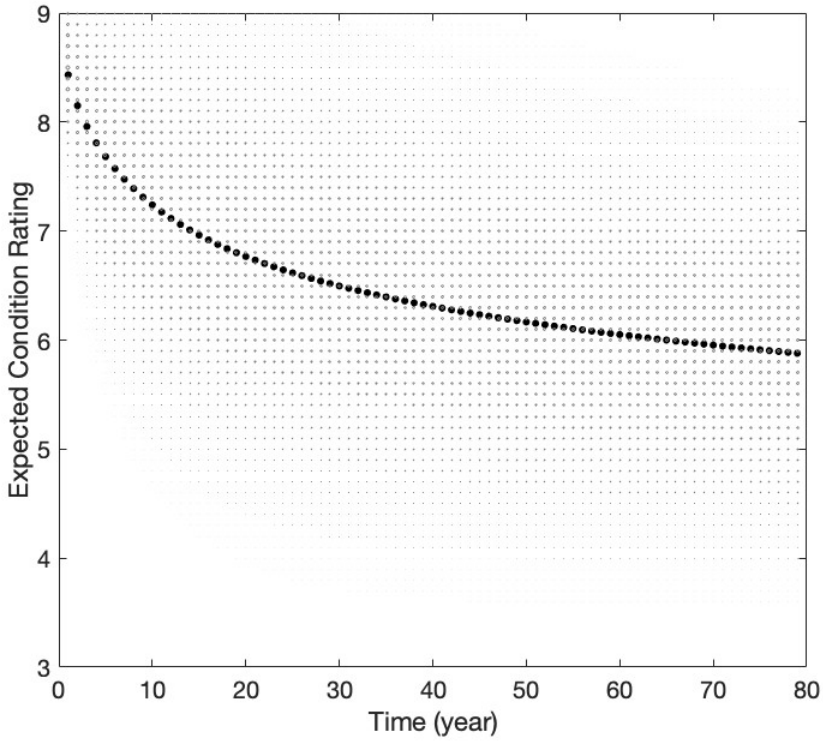


(b) Zone 2

Figure 4.13: 2D Deterioration curves including all possible CRs using Method 1 by zone



(c) Zone 3



(d) Zone 4

Figure 4.13: 2D Deterioration curves including all possible CRs using Method 1 by zone (cont.)

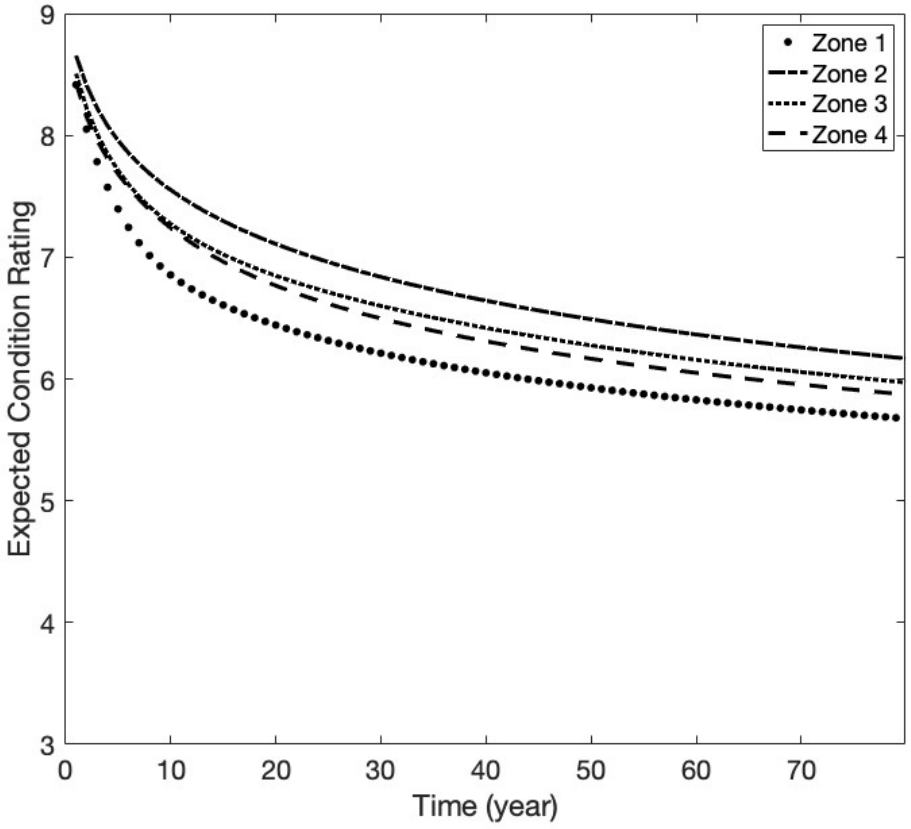
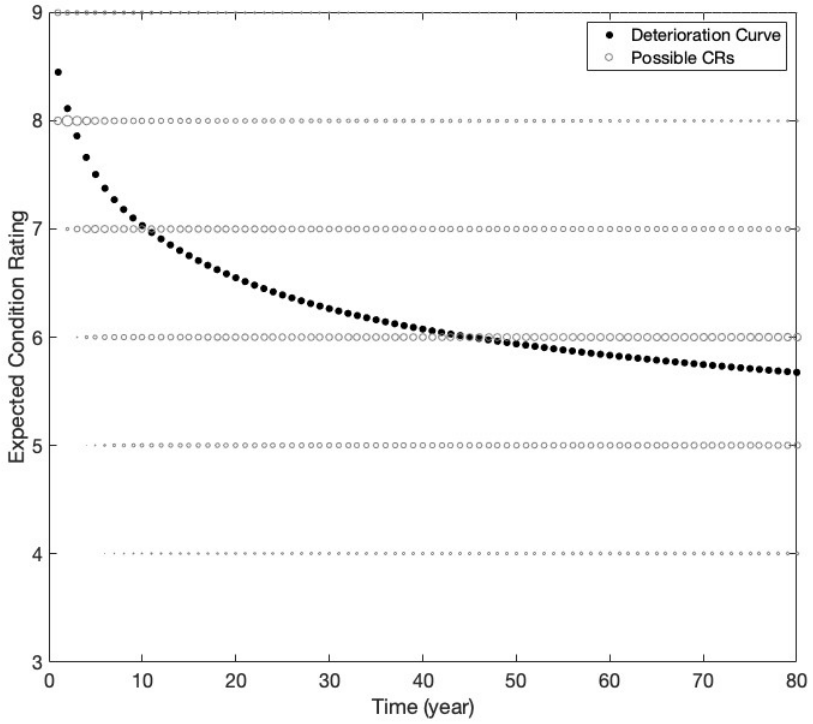
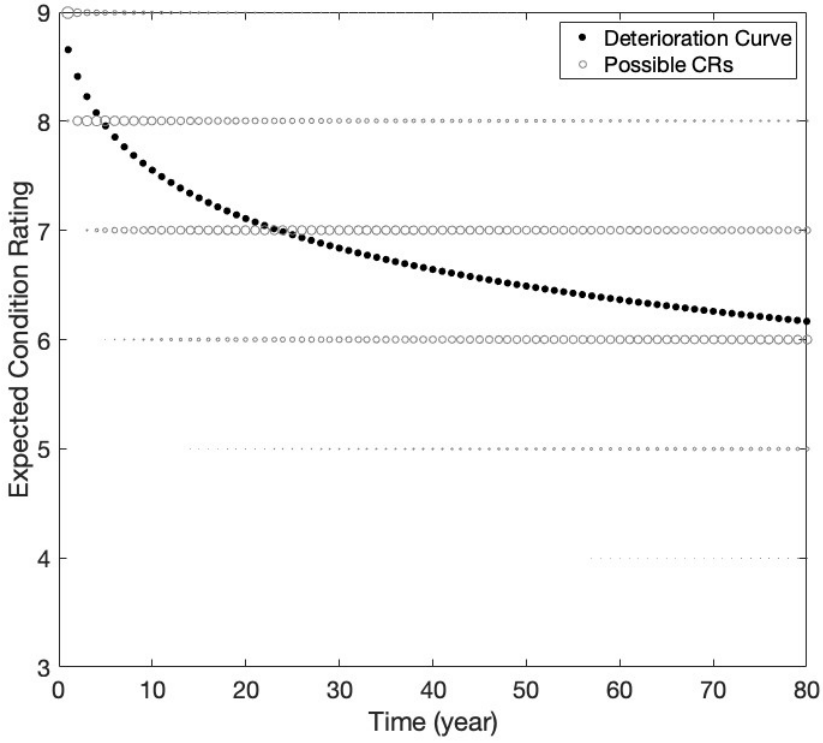


Figure 4.14: Comparison of deterioration curves by zone using Method 1

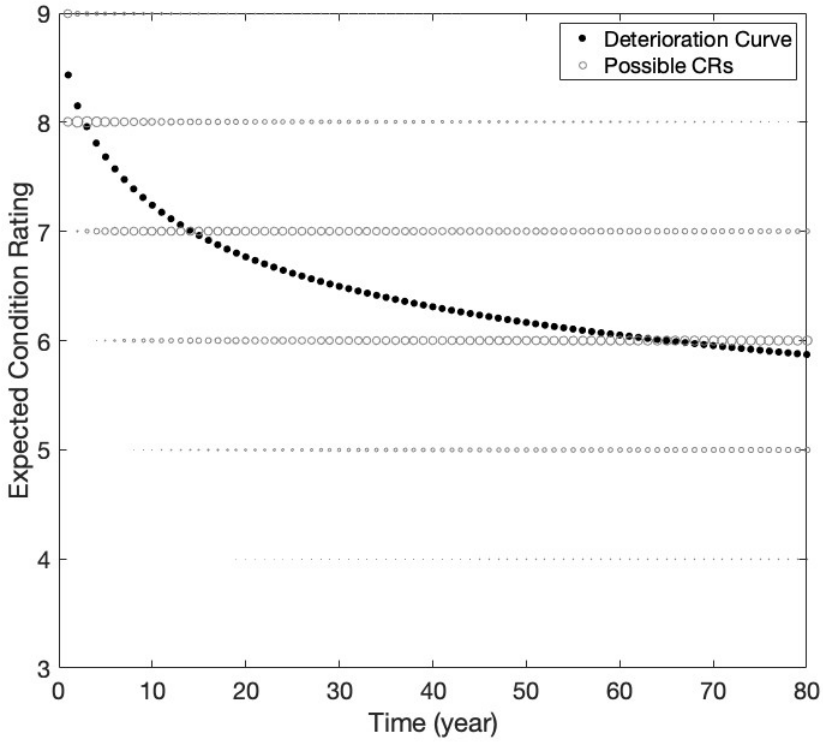


(a) Zone 1

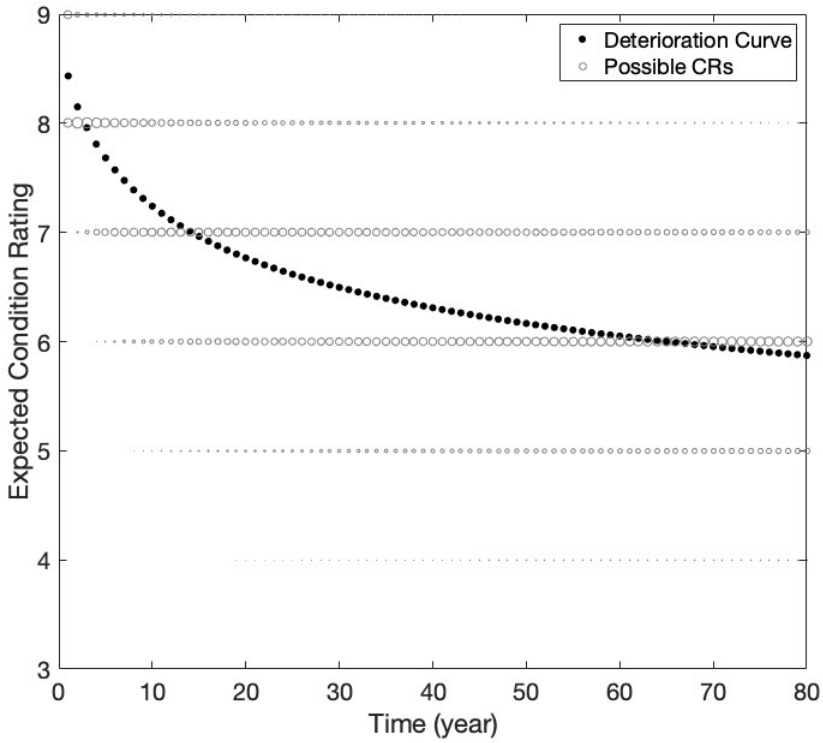


(b) Zone 2

Figure 4.15: 2D Deterioration curves including possible CRs at given times using Method 1 by zone



(c) Zone 3

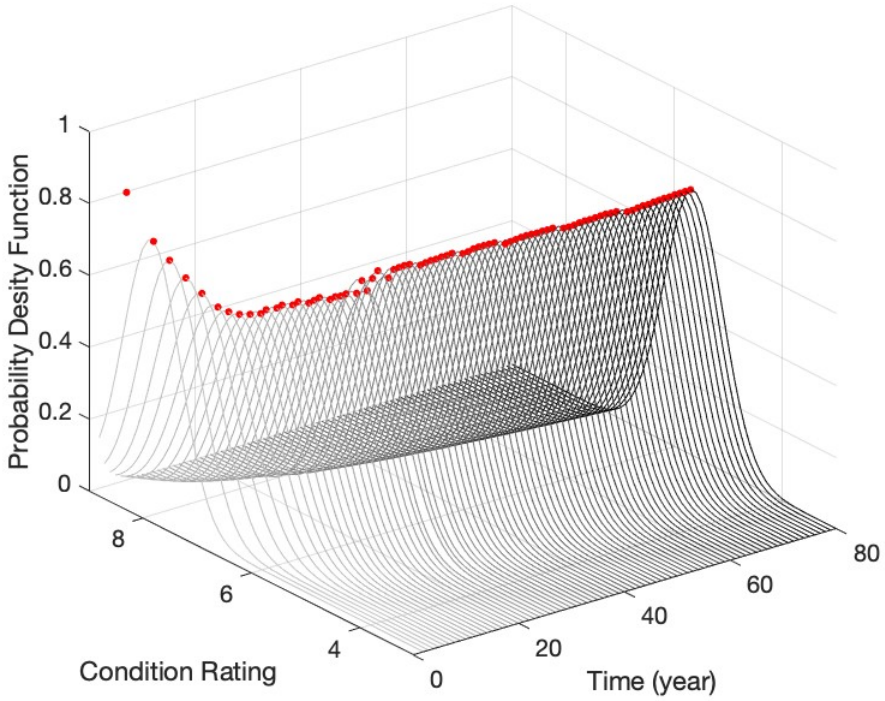


(d) Zone 4

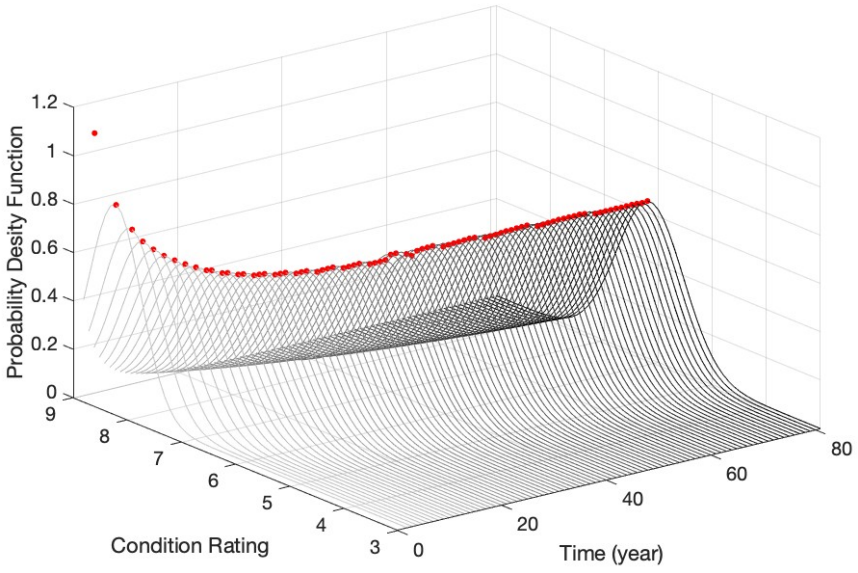
Figure 4.15: 2D Deterioration curves including possible CRs at given times using Method 1 by zone (cont.)

Method 2

Method 2 uses transition probability matrices (TPM) obtained from the single models. Figure 4.16 shows 3D deterioration curve plots. The x-axis is the time from 0 (front) to 80 years (back). The y-axis is the condition rating (CR) from 3 (right) to 9 (left). The z-axis is the probability density function from 0 (bottom) to 1 (top) in zone 1 and 3, to 1.2 (top) in zone 2, and 0.9 (top) in zone 4. Figure 4.17 presents continuously expected condition ratings (black dots), including all possible condition ratings (gray dots) at a given time t. The size of the dots indicates the degree of probability of expecting at particular CRs. The higher-order of deterioration rate is the same as in method 1. The range of possible CRs at a given time is similar in all zones. Figure 4.18 contains deterioration curves of all zones. The deterioration rate is higher in order of zone 1, 4, 3, and 2. This is the same as the one in method 1. In Figure 4.19, a deck in zone 2, 3, and 4 can be rated CR 8 at 80 years old. In zone 1, a deck can be rated CR 8 up to about 40 years old. A deck in zones 1 and 2 may not be rated CR at 80. In zone 3 and 4, a deck can be rated CR 4 from about 50 years and about 30 years, respectively.

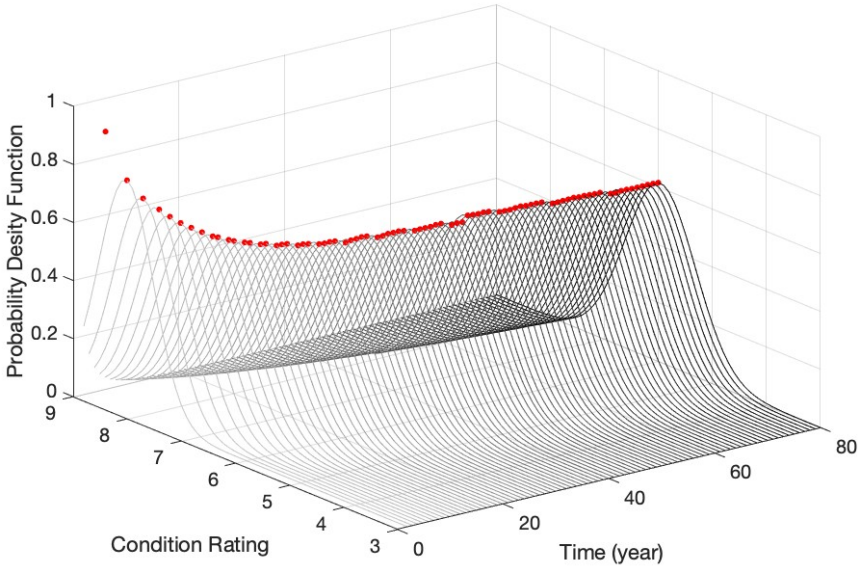


(a) Zone 1

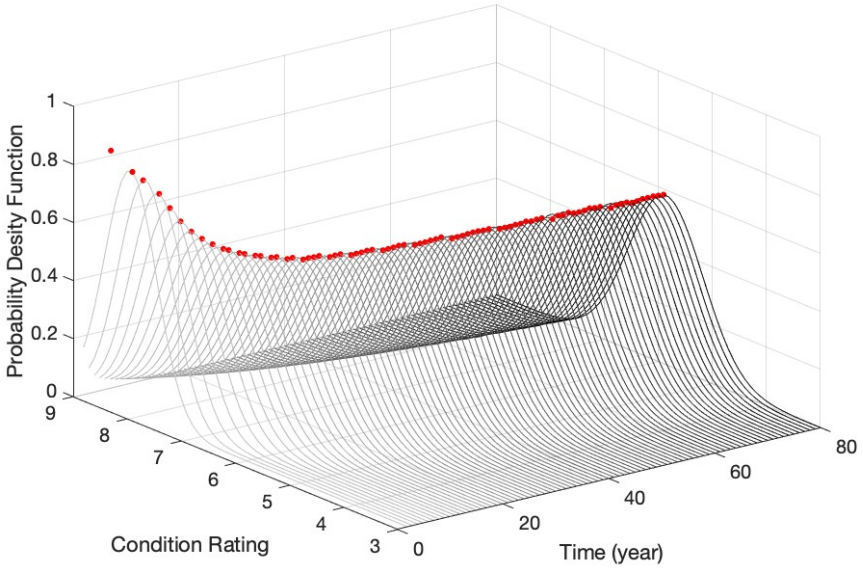


(b) Zone 2

Figure 4.16: 3D deterioration curve plots of the proposed models using Method 2 by zone

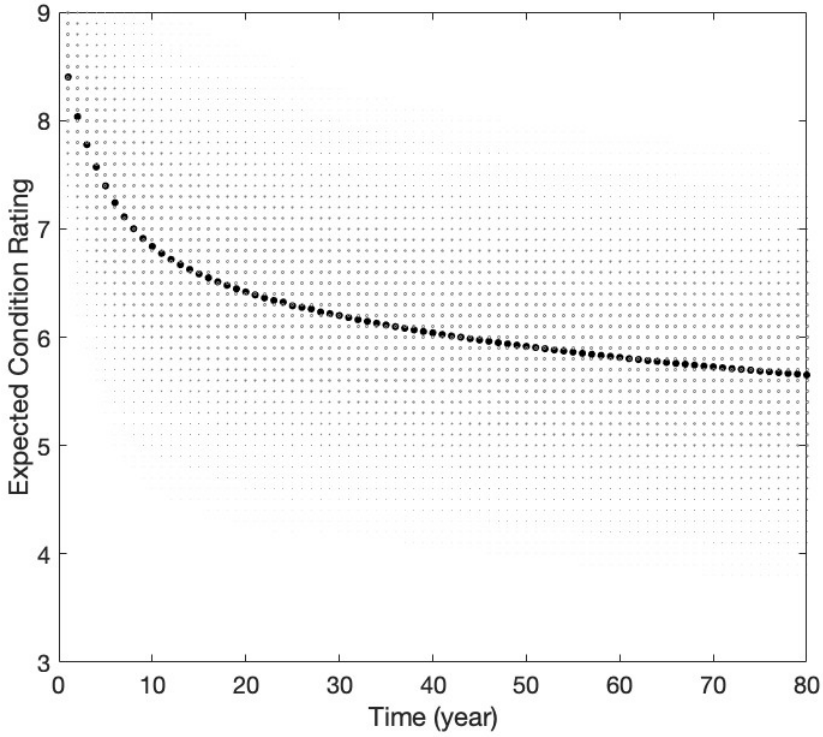


(c) Zone 3

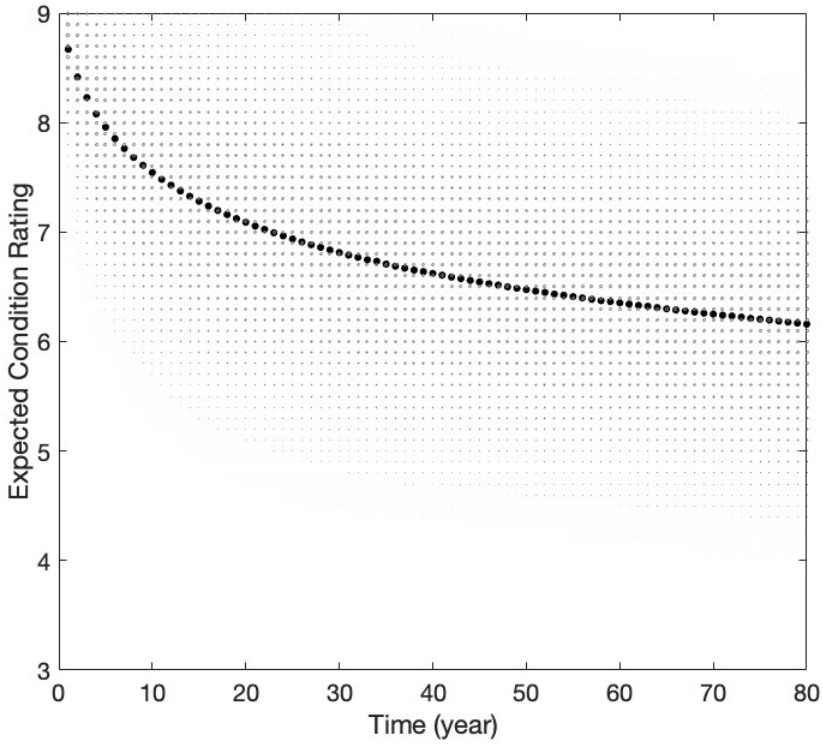


(d) Zone 4

Figure 4.16: 3D deterioration curve plots of the proposed models using Method 2 by zone (cont.)

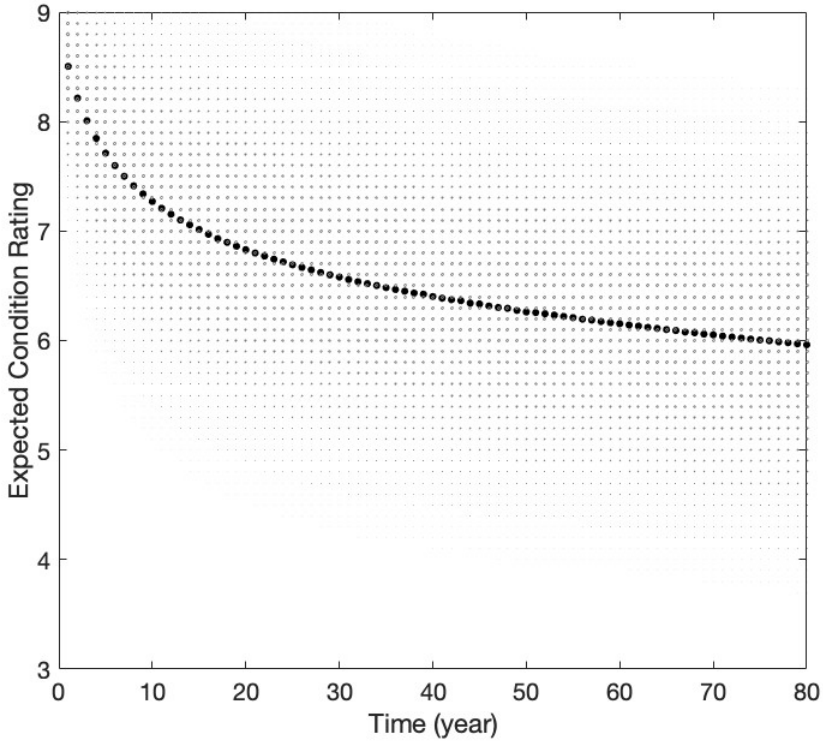


(a) Zone 1

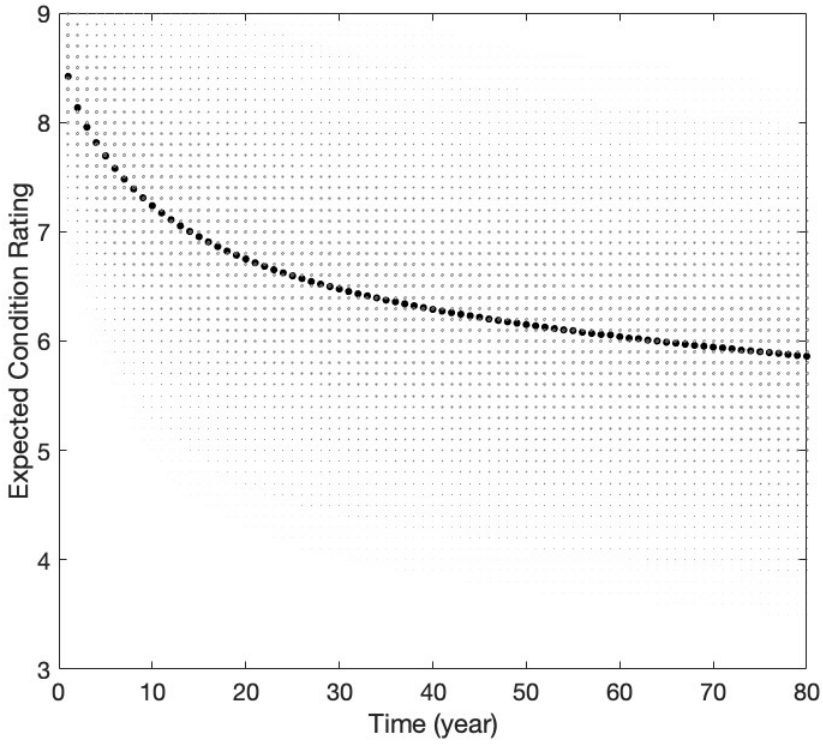


(b) Zone 2

Figure 4.17: 2D Deterioration curves including all possible CRs using Method 2 by zone



(c) Zone 3



(d) Zone 4

Figure 4.17: 2D Deterioration curves including all possible CRs using Method 2 by zone (cont.)

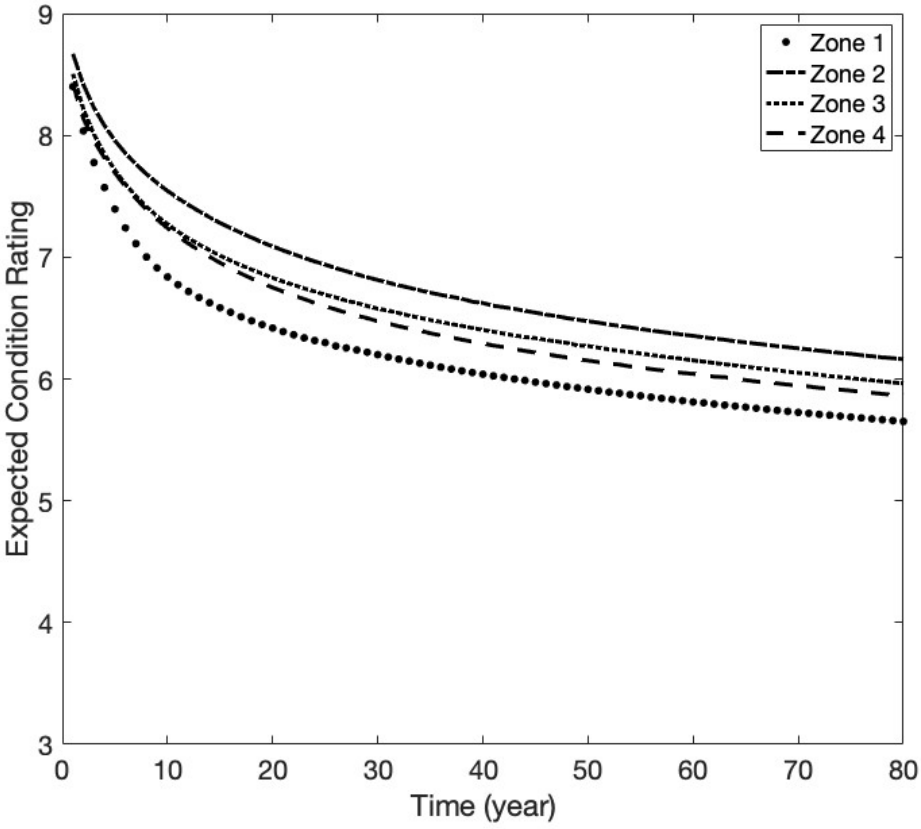
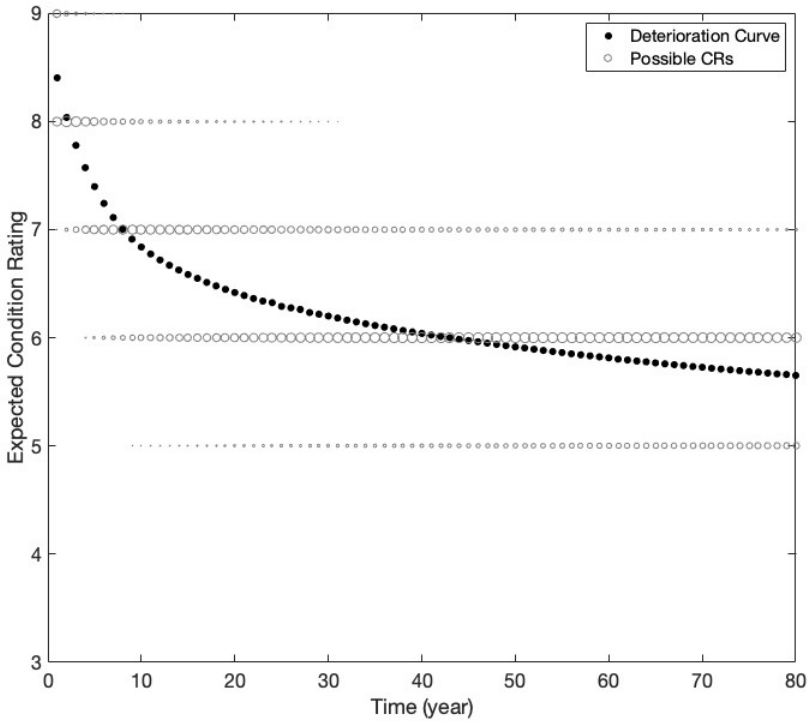
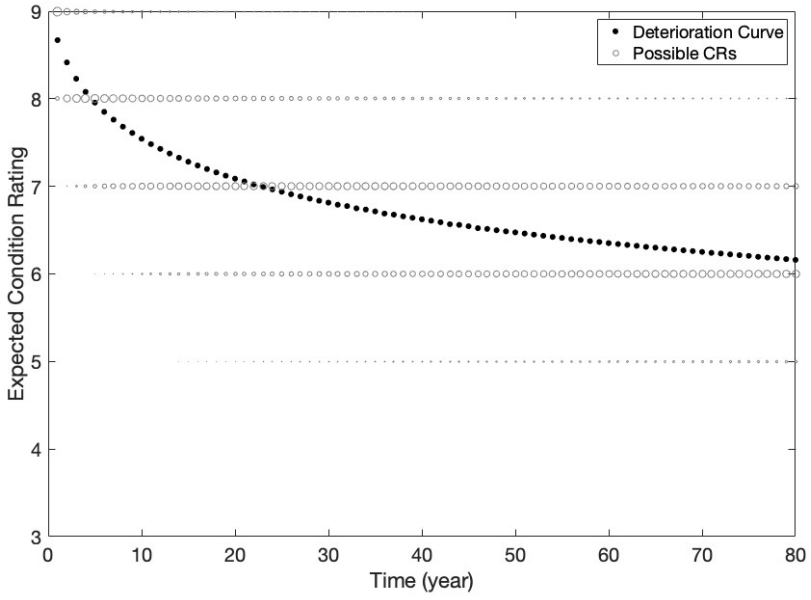


Figure 4.18: Comparison of deterioration curves by zone using Method 2

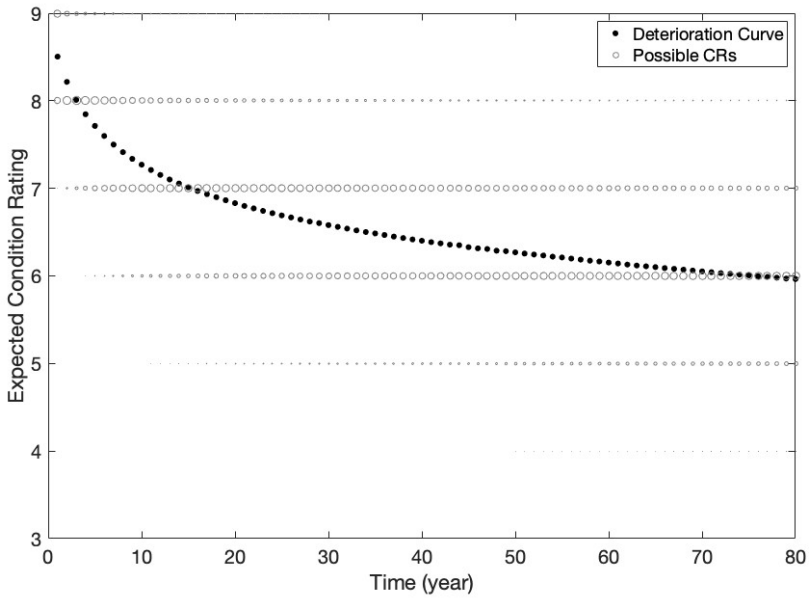


(a) Zone 1

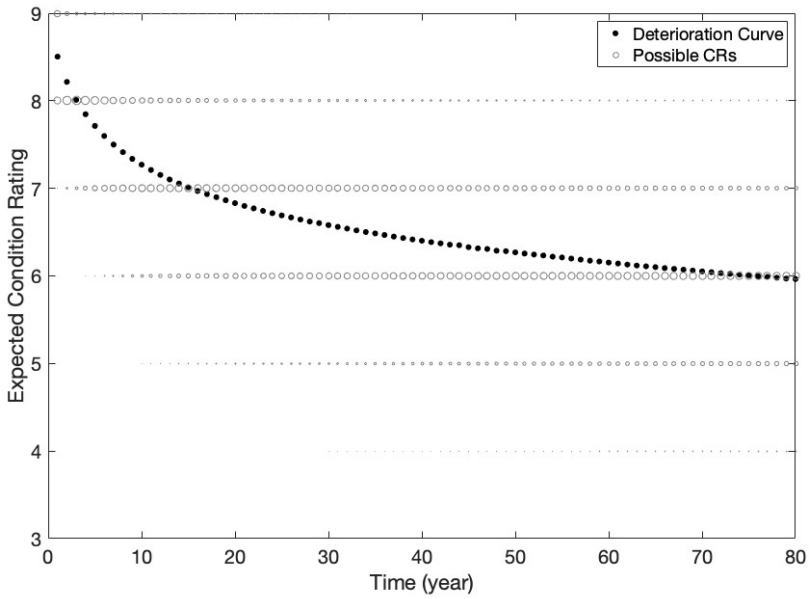


(b) Zone 2

Figure 4.19: 2D Deterioration curves including possible CRs at given times using Method 2 by zone



(c) Zone 3



(d) Zone 4

Figure 4.19: 2D Deterioration curves including possible CRs at given times using Method 2 by zone (cont.)

4.4.1.2 Expected Condition Distribution

Table 4.10 shows the expected condition rating distributions of bridge concrete decks of the network system of 2010, 2020, and 2030. The future condition rating distribution of the decks is estimated using method 2 with the initial condition rating distribution of 2008. Method 1 cannot be used to estimate the expected condition distribution directly because it uses the predicted condition ratings calculated from the models, not the TPMs. The percent of decks at CR 9, 8, and 7 decreases, and the percent of decks at CR 6, 5, 4, and 3 increases over time. For examples, decks at condition rating 9 are 0.41, 0.03, and 0.01 percent in 2010, 2020, and 2030, respectively in zone 1, and decks at condition rating 7 are about 50, 43, and 34 percent in 2010, 2020, and 2030, respectively in zone 3. In zone 2, decks at CR 6 are about 19, 33, and 43 percent in 2010, 2020, and 2030, respectively. In zone 4, decks at condition rating 4 are about 0.11, 0.60, and 1.28 percent in 2010, 2020, and 2030, respectively.

Table 4.10: Expected condition rating distribution of concrete decks (unit: percentage)

Year	**CR9**	**CR8**	**CR7**	CR6	**CR5**	CR4	**CR3**
Zone 1							
2010	0.38	9.45	44.47	39.08	6.42	0.20	0.00
2020	0.01	1.03	22.55	58.18	17.62	0.63	0.00
2030	0.00	0.28	13.50	60.11	25.15	0.99	0.02
Zone 2							
2010	2.75	26.62	50.45	18.74	1.37	0.02	0.00
2020	0.30	11.19	53.43	32.53	2.52	0.02	0.00
2030	0.07	5.60	47.08	42.93	4.25	0.05	0.00
Zone 3							
2010	1.75	20.81	50.25	24.65	2.46	0.05	0.00
2020	0.12	6.06	43.45	43.97	6.28	0.13	0.00
2030	0.03	2.71	34.32	52.82	9.89	0.23	0.00
Zone 4							
2010	0.96	15.17	48.47	31.23	4.06	0.11	0.00
2020	0.14	5.01	35.08	47.07	12.10	0.60	0.01
2030	0.04	2.23	25.36	51.91	19.16	1.28	0.02

4.4.2 Reliability of Prediction

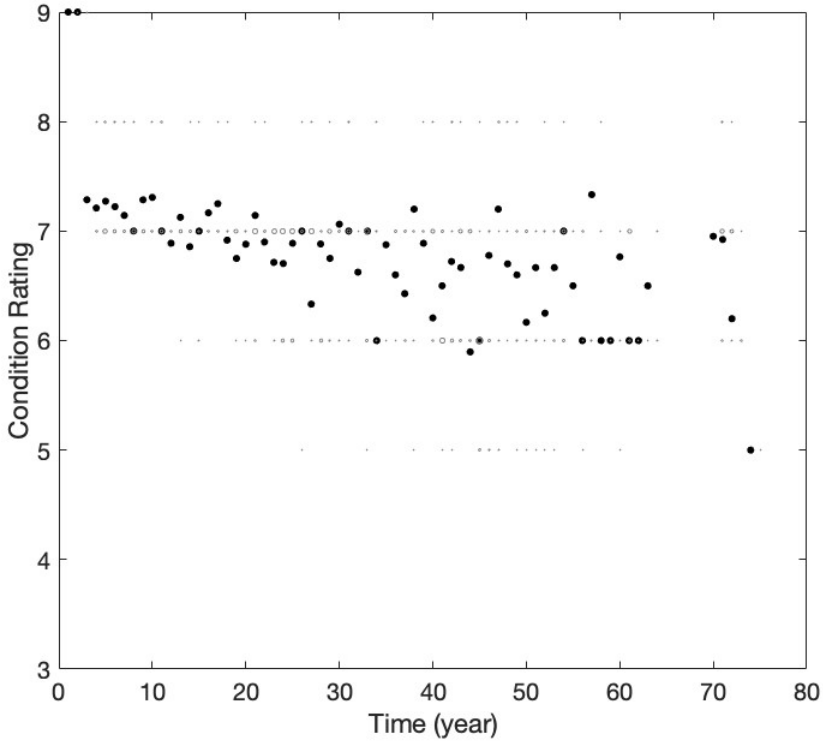
Figure 4.20 shows the sample data in zone 1, 2, 3, and 4. The x-axis is the time from 0 to 80 years, and the y-axis is the condition rating from 3 to 9. The black dots in the graphs indicate the means of the sample at a given age. The mean values are noisy year to year, but exhibit a general decreasing trend. Comparing expected CRs and the means is not a suitable method to evaluate the reliability of the proposed models.

In this research, two validation approaches are used to evaluate the reliability of the prediction of proposed models. The first approach is to compare probabilities of particular condition ratings at a given time. The sample data is organized according to particular CRs at a given time, and then the number of the CRs is divided by the sum of the total number of the CRs at a given time. They are compared to the probabilities of possible CRs at a given time, shown in Figure 4.15 and Figure 4.19 for Method 1 and Method 2, respectively. The second approach is a comparison of the most likely CRs. The CRs, the largest number and highest probability at a given time from the sample data and the proposed models, respectivel, are extracted and compared. Figure 4.21 includes the most likely expected CRs obtained using methods 1 (blue circles) and 2 (red circles) and the most likely observed CRs from the sample data (black dots). The patterns of the most likely expected CRs from methods 1 and 2 in all zones are similar.

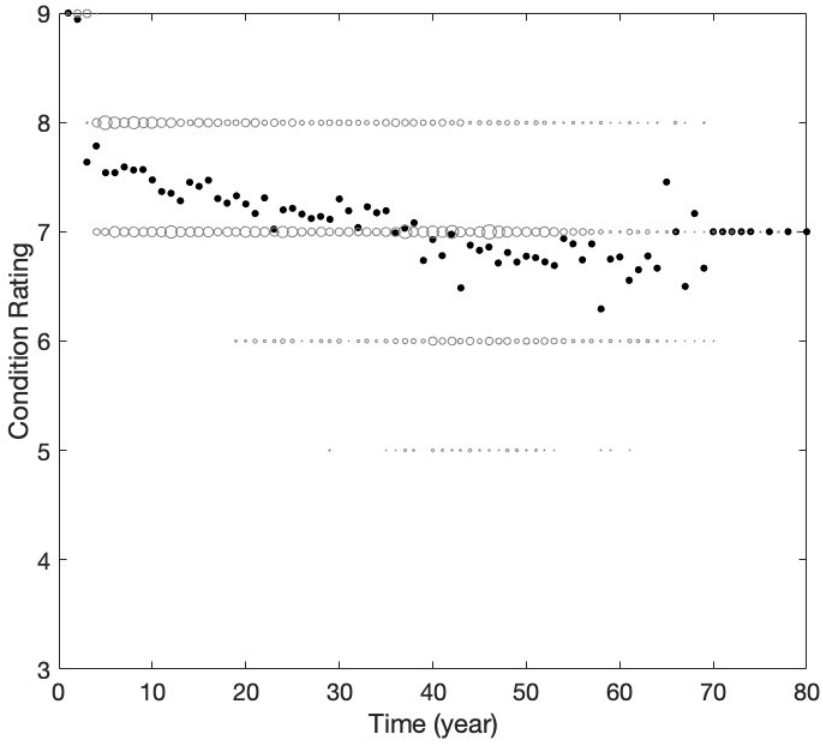
Table 4.11 summarizes the MAE values calculated using Equation 4.1. The values according to CRs from 9 to 3 are similar between method 1 and 2. The overall CR is the averaged MAE values over all CRs and ages. The most likely CR is the average of the sum of the differences between the most possible expected CR and observed CR at a given time. The mean is the MAE value between the continuous expected CRs and means from the sample data. The MAE values of the overall CR, most likely CR, and mean are smaller using method 1 than method 2.

Table 4.12 contains the MAE value of the expected condition distribution, that is, the error between the estimated condition distribution of the 2010 year and the sample data of the 2010 year. The value of zone 1 is 0.05, and it is smaller than the ones (0.14) of the other zones.

Figure 4.22 shows the prediction error is the difference between the most likely estimated and observed CRs at a given time; if the estimated and observed CR is equal, it is zero; If an

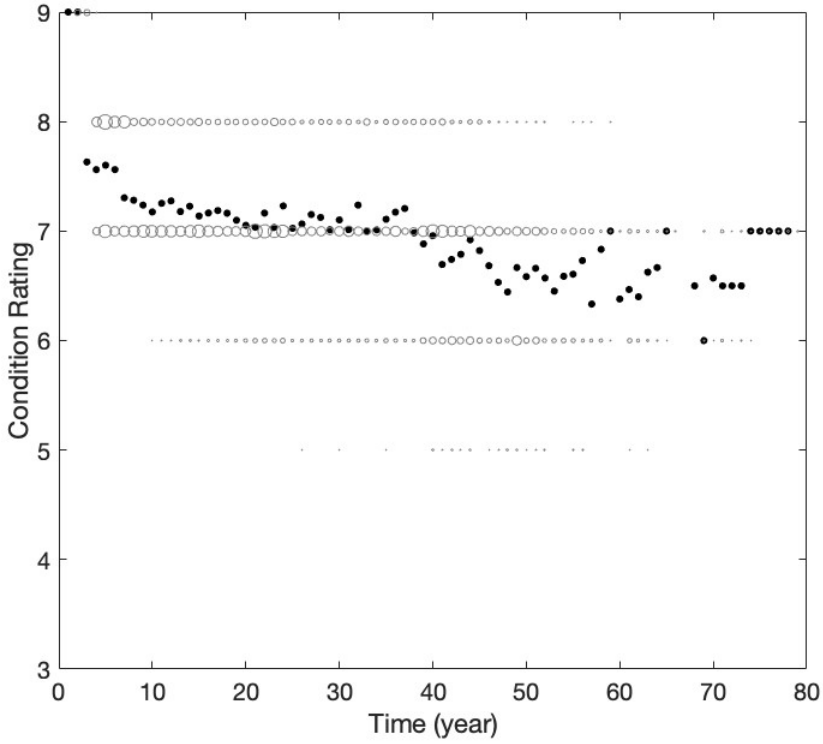


(a) Zone 1

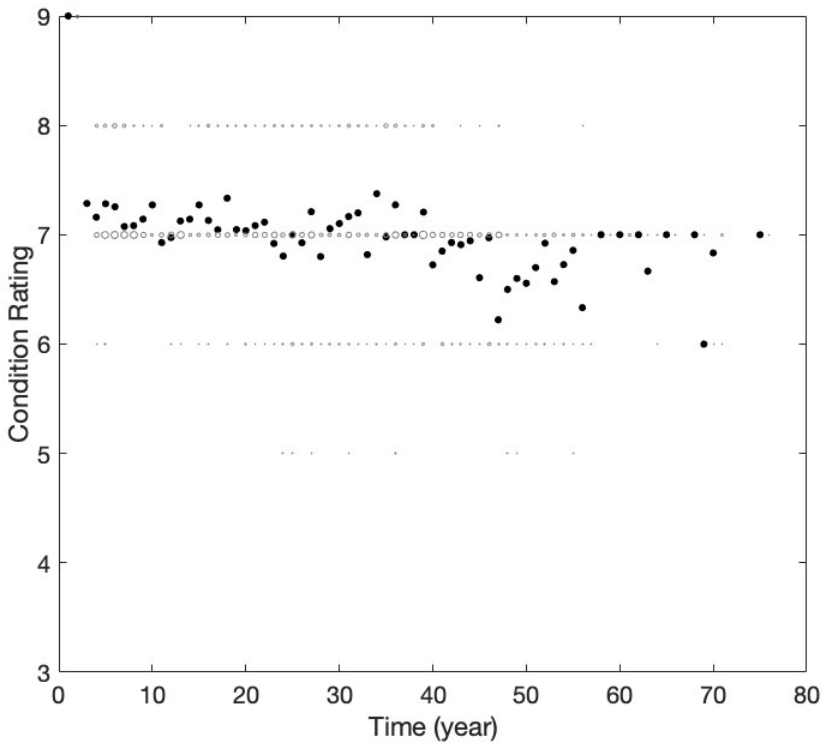


(b) Zone 2

Figure 4.20: Mean and condition rating distribution of sample data across zones



(c) Zone 3



(d) Zone 4

Figure 4.20: Mean and condition rating distribution of sample data across zones (cont.)

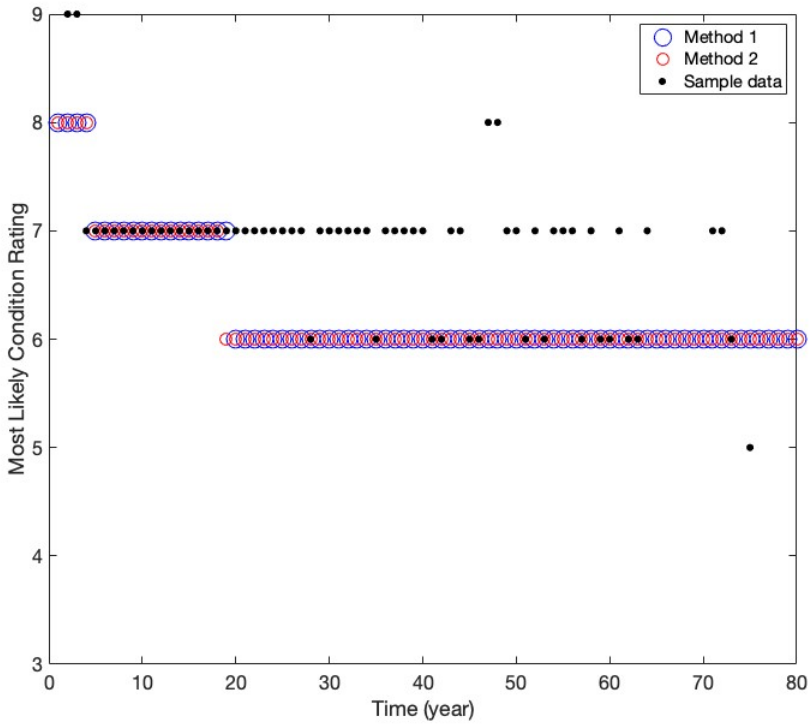
Table 4.11: Mean Absolute Error (MAE) of expected condition ratings by zones

	Zone 1		**Zone 2**		**Zone 3**		**Zone 4**	
	Method 1	Method 2	Method 1	Method 2	Method 1	Method 2	Method 1	Method 2
CR 9	0.05	0.03	0.03	0.03	0.03	0.03	0.03	0.02
CR 8	0.11	0.09	0.09	0.10	0.07	0.08	0.11	0.11
CR 7	0.34	0.27	0.17	0.17	0.27	0.24	0.32	0.31
CR 6	0.24	0.33	0.17	0.20	0.25	0.24	0.31	0.32
CR 5	0.18	0.13	0.04	0.03	0.05	0.06	0.10	0.10
CR 4	0.06	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Overall CR	0.16	0.14	0.08	0.09	0.11	0.11	0.14	0.14
Most Likely CR	0.52	0.54	0.35	0.38	0.47	0.47	0.60	0.62
Mean	0.53	0.54	0.35	0.36	0.39	0.40	0.50	0.52

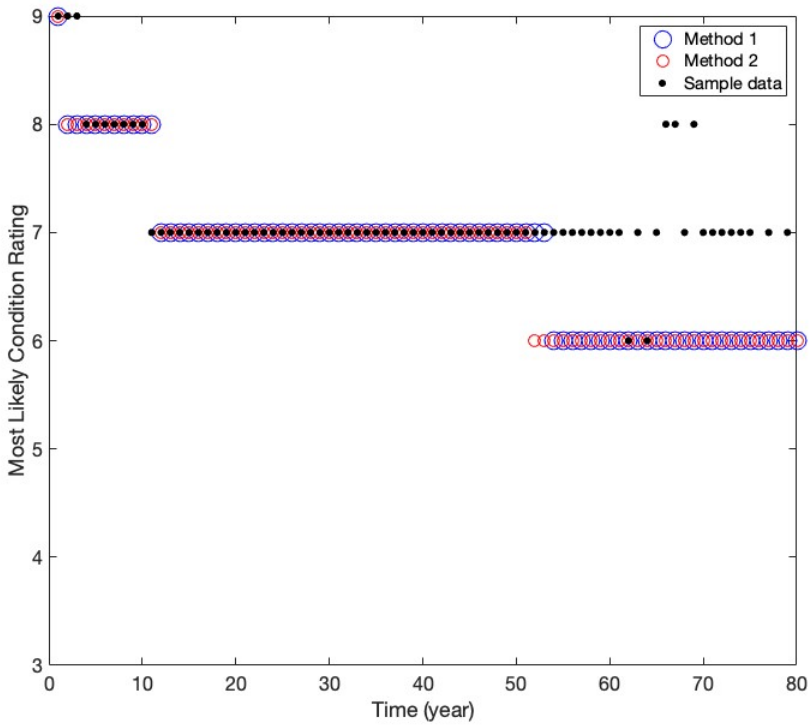
Table 4.12: MAE of expected condition distributions of the zones

Year	**Zone 1**	**Zone 2**	**Zone 3**	**Zone 4**
2010	0.05	0.14	0.14	0.14

observed CR is lower (higher) than the estimated CR, it is negative (positive). Overall, the prediction error of a proposed model using method 2 is slightly higher than that of method 1.

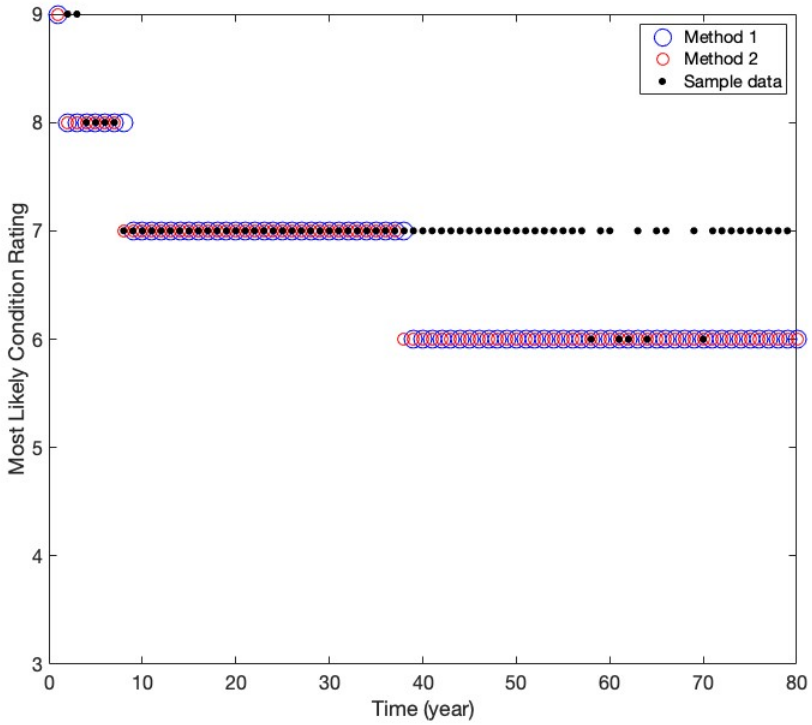


(a) Zone 1

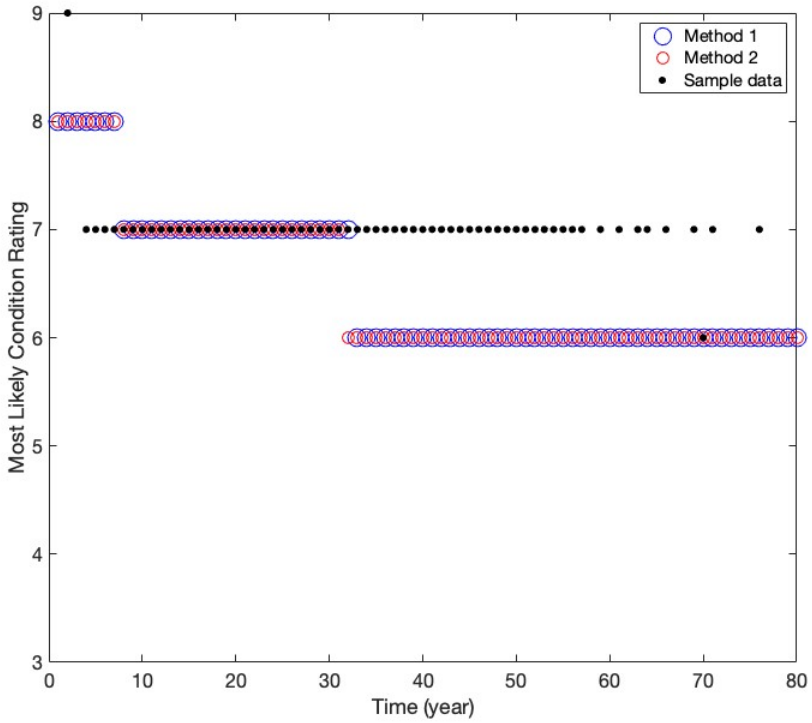


(b) Zone 2

Figure 4.21: Comparison of expected and observed most likely condition ratings by zone

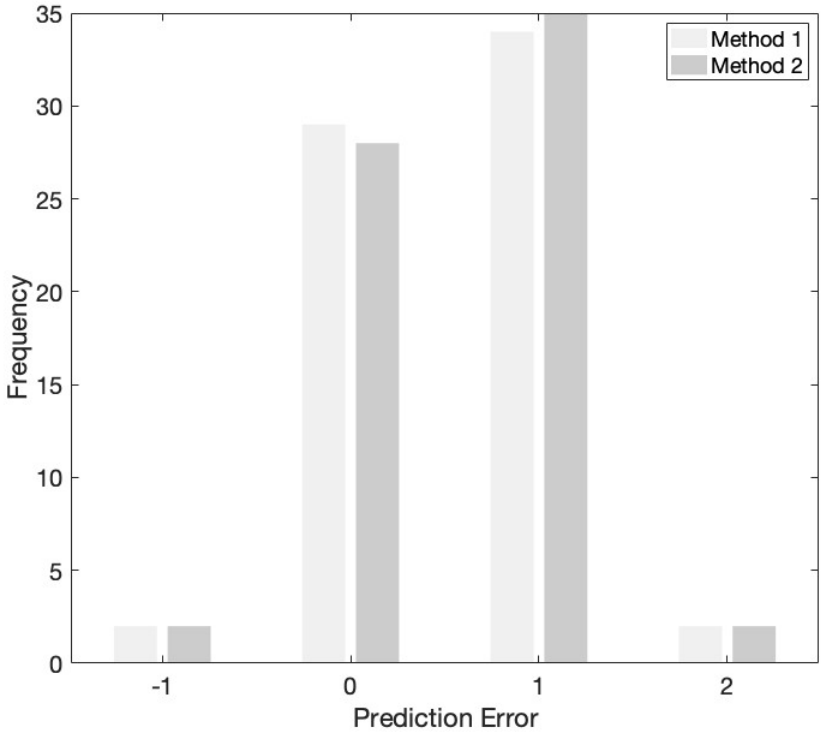


(c) Zone 3

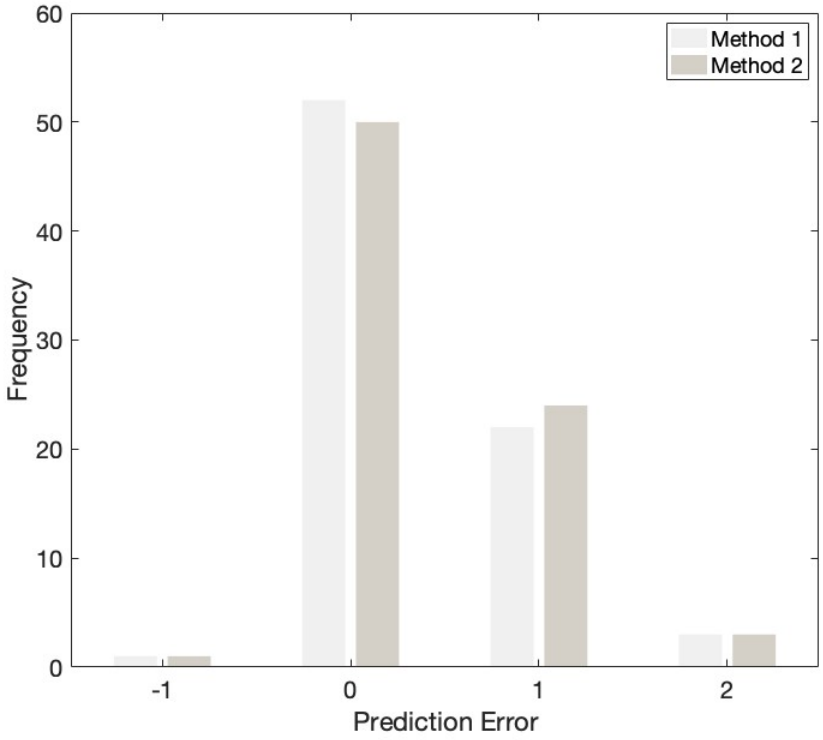


(d) Zone 4

Figure 4.21: Comparison of expected and observed most likely condition ratings by zone (cont.)

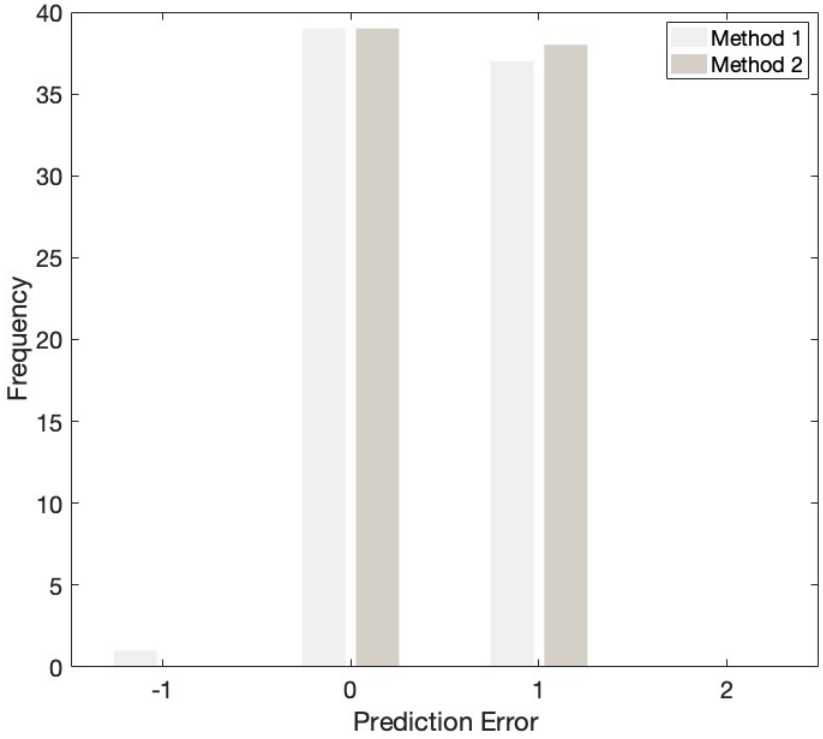


(a) Zone 1

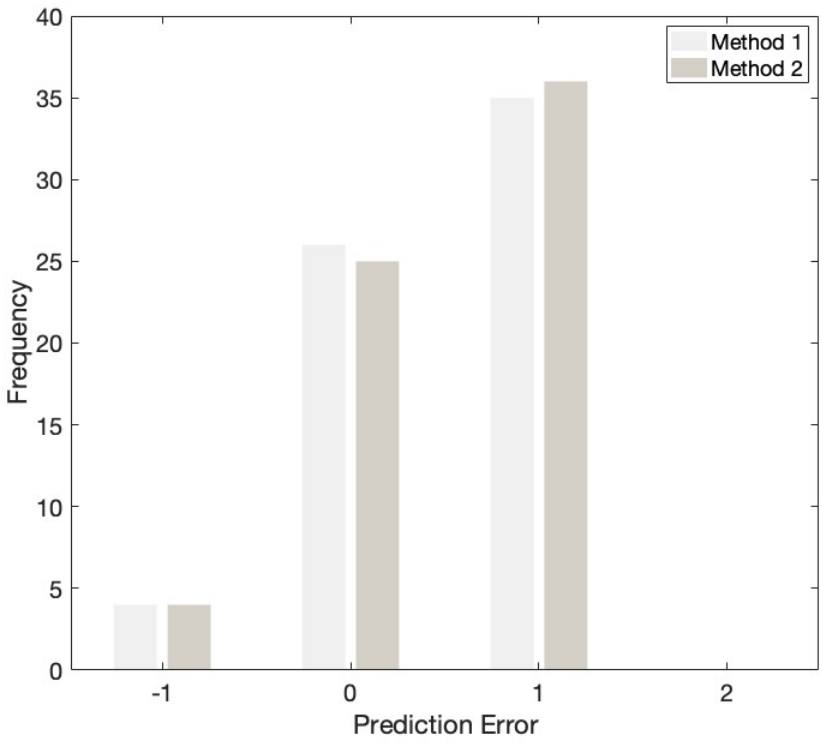


(b) Zone 2

Figure 4.22: Prediction Error in most likely condition rating by zone



(c) Zone 3



(d) Zone 4

Figure 4.22: Prediction Error in most likely condition rating by zone (cont.)

4.4.3 The effect of explanatory variable

In this section, the effect of the explanatory variable (i.e., environment) is explored by comparing two models of zone 1: One is that all models are included in the modeling, and the other is that all models, except the mechanistic deterioration model (MDM), are included.

Method 1

Figure 4.23 shows two 2D deterioration curve plots: 4.23(a) shows results with all expected condition ratings (CRs) obtained from all the models included in the modeling, and 4.23(b) shows results with all expected CRs except CRs from the MDM included. A deck deteriorates faster in 4.23(a) than in 4.23(b). The time to a possible CR 4 ranges from about 10 years and about 20 years in 4.23(a) and 4.23(b), respectively.

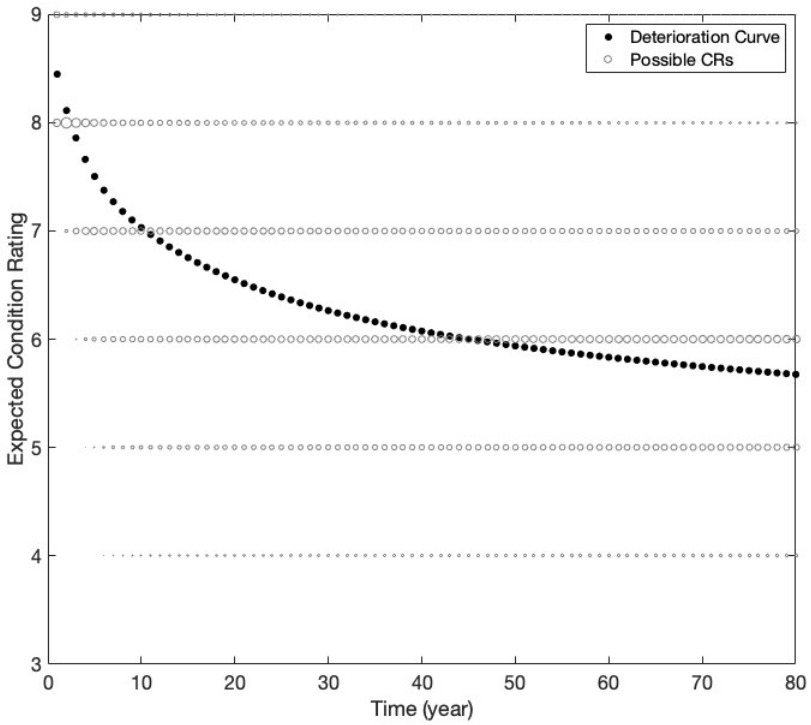
Figure 4.24 includes the most likely estimated and observed CRs. The bule circles indicate the most likely CRs of a proposed model, using expected CRs obtained from all models. The red circles are the most likely CRs of a proposed model, using expected CRs obtained from all models except the MDM. The black dots are the observed most likely CRs of the sample data. The CRs (red circles) of the proposed model, excluding the MDM, seem closer to the CRs (black dots) of the sample data. Like the black dots, the red circles stay up to 34 years at CR 7.

Table 4.13 summarizes the mean absolute error (MAE) values of the two models. The MAE values of the most likely CR and mean of the model, which exclude the MDM, are smaller than the other model.

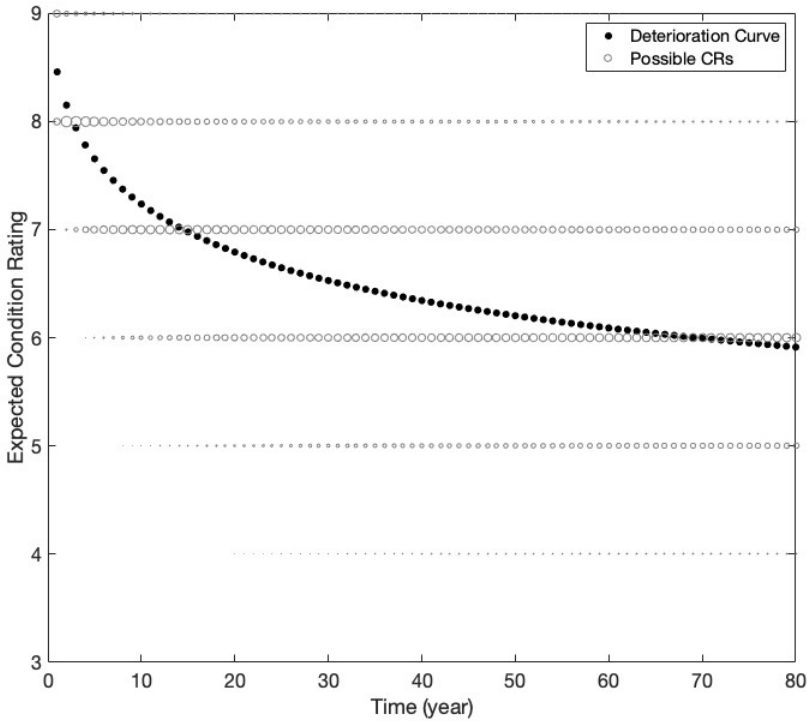
Table 4.13: MAE values of the proposed models of Zone 1 using Method 1

	Overall CR	**Most Likely CR**	**Mean**
Include all expected CRs	0.16	0.52	0.53
except CRs from MDM	0.13	0.22	0.47

Figure 4.25 shows the most likely CR prediction error at a given time. The lighter gray bars is a model that uses all expected CRs in modeling. The darker gray bars are a model that uses all expected CRs except MDM CRs. The overall prediction errors of the lighter gray bars are less than the others.



(a) All expected CRs



(b) Except MDM CRs

Figure 4.23: 2D Deterioration curve plots of Zone 1 using Method 1

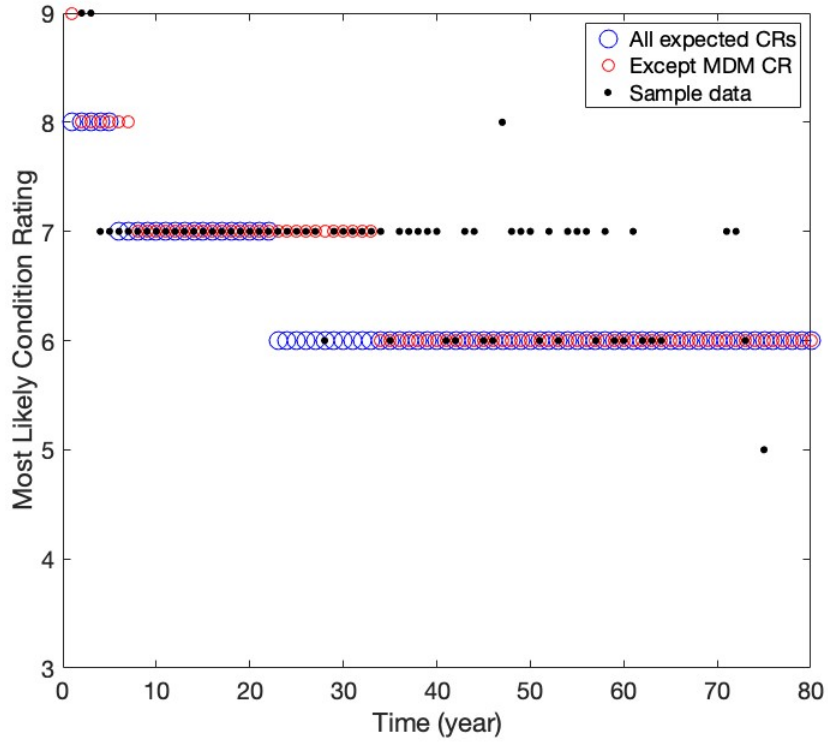


Figure 4.24: Comparison of expected and observed most likely CRs of Zone 1 using Method 1

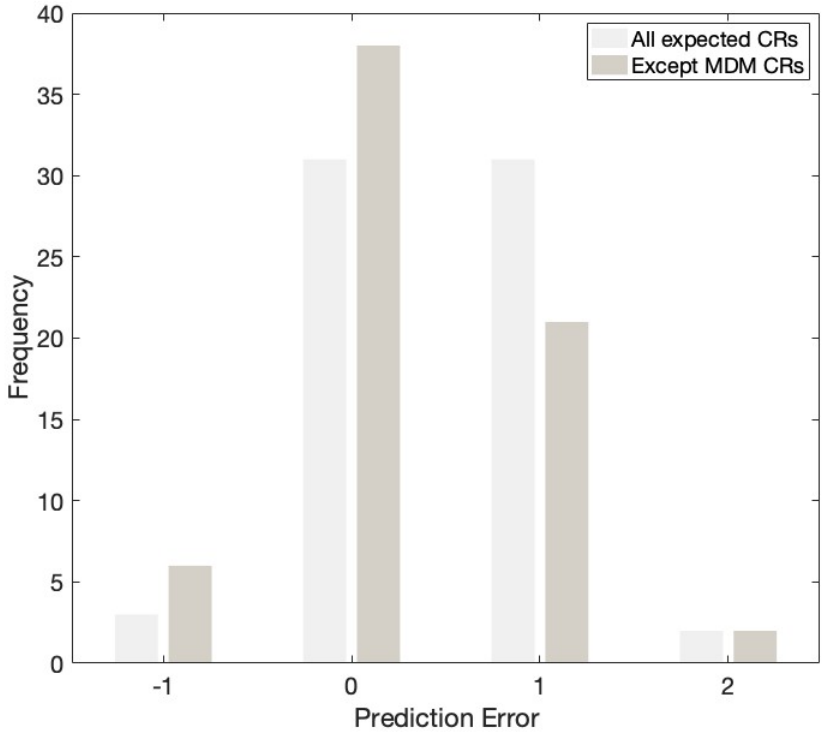


Figure 4.25: Comparison of prediction error in most likely CRs of Zone 1 using Method 1

Method 2

Figure 4.26 contains deterioration curves: 4.26(a) uses all the models in modeling, and (b) uses all the models except the MDM. Figure 4.27 includes the most possible estimated and observed CRs. The blue circles indicate the most likely CRs of a proposed model, using TPMs obtained from all models. The red circles are the most likely CRs of a proposed model, using TPMs obtained from all models except the MDM. The black dots are the observed most likely CRs of the sample data. The CRs (red circles) of the proposed model, excluding the MDM, seem closer to the CRs (black dots) of the sample data. Like the black dots, the red circles stay up to 34 years at CR 7. It is the same result as the one in method 1. Figure 4.28 shows prediction error.

Table 4.14 summarizes the MAE values of the two models; a model includes all models in the modeling. The other model uses all models except the MDM. The model, including all models, has a smaller MAE value of the overall CRs than the other model. However, excluding the MDM, the model has smaller MAE values of the most likely CR and mean than the other model.

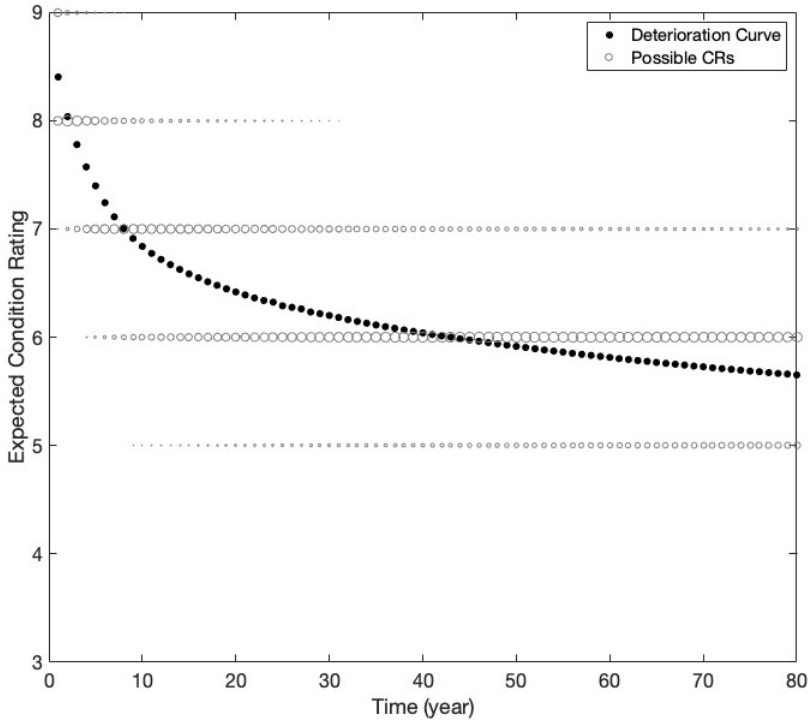
Table 4.14: MAE values of the proposed models of Zone 1 using Method 2

	Overall CR	**Most Likely CR**	**Mean**
All models	0.15	0.54	0.54
Except the MDM	0.15	0.54	0.49

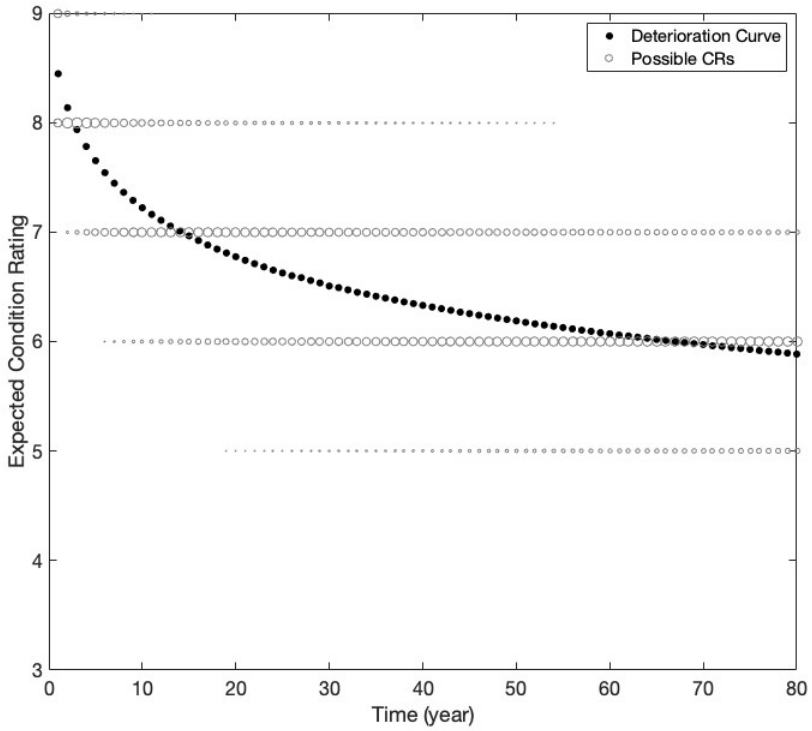
Table 4.15 summarizes the prediction errors of the proposed models. One model includes all expected CRs or all TPMs, and the other excludes the MDM's expected CRs or TPM. The prediction errors using methods 1 and 2 are similar to each other.

Table 4.15: Prediction error of Method 1 and 2 of Zone 1

Prediction Error	**Method 1**		**Method 2**	
	All	Except MDM	All	Except MDM
-1	3	6	2	5
0	31	38	28	40
1	31	21	35	21
2	2	2	2	3



(a) All models



(b) Except the MDM

Figure 4.26: 2D Deterioration curve plots of Zone 1 using Method 2

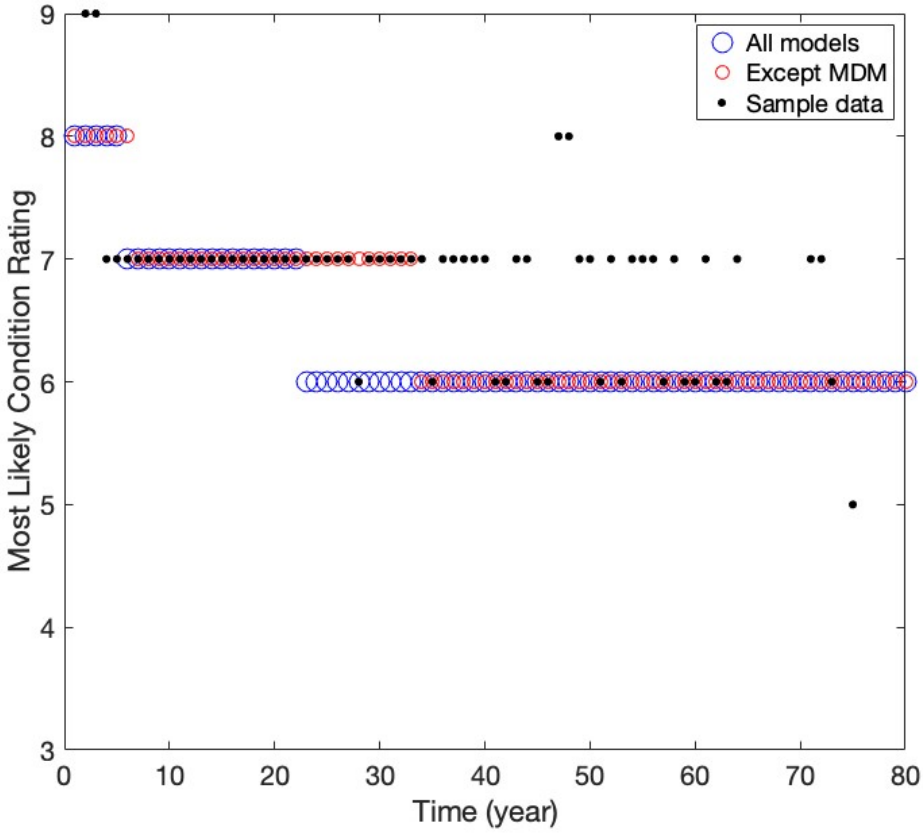


Figure 4.27: Comparison of expected and observed most likely CRs of Zone 1 using Method 2

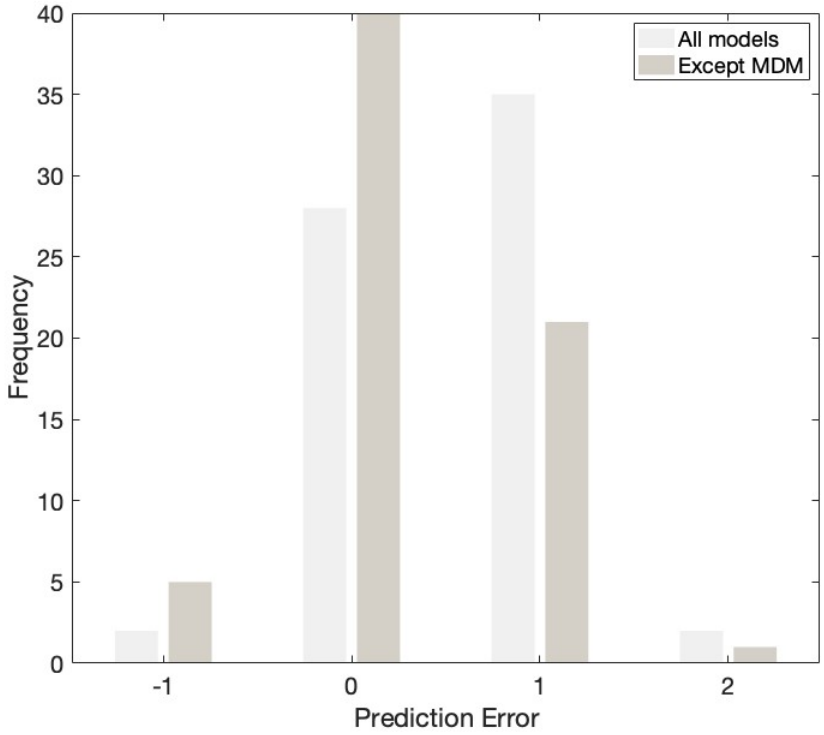


Figure 4.28: Comparison of prediction error in most likely CRs of Zone 1 using Method 2

Chapter 5: Case Study 2

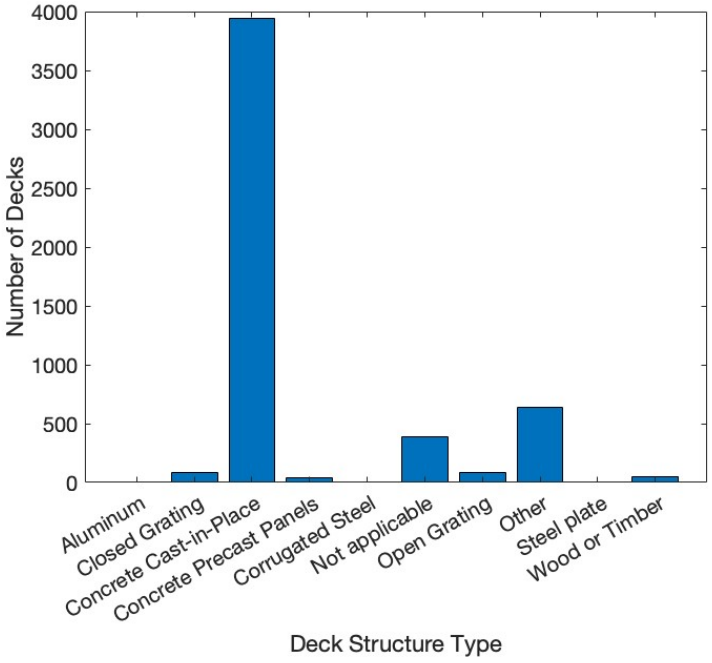
5.1 Data Selection and Preparation

This study focused on bridges in Pennsylvania. The bridge data from 2008 and 2010 were collected from the NBI database, and the total numbers of bridges are 29,414 and 22,318, respectively. The development processes of single deterioration models and the proposed models are the same as the ones in case study 1. The detail of the processes are not presented in this chapter, and only the results are presented.

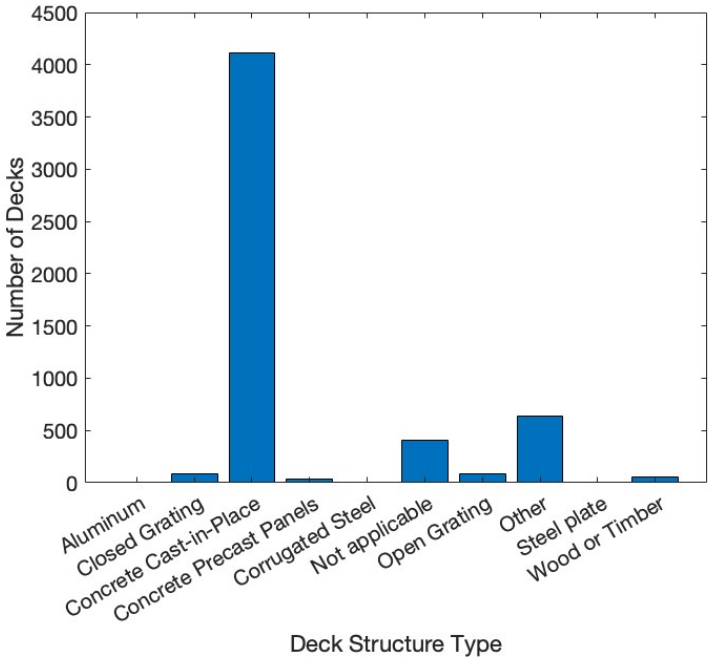
5.1.1 Grouping

The collected decks were grouped according to the same explanatory variables as the ones in case study 1. Figure 5.1 contains the number of decks of 2010 according to the deck structure type in Pennsylvania. After grouping, the total number of decks was 5,232 and 5,415 in 2008 and 2010, respectively. This research uses ”concrete cast-in-place”, ”concrete precast panels”, and ”other” as sample data.

Unlike Texas, the sample data were not divided into zones according to winter weather conditions. The winter weather condition in Pennsylvania is shown in Figure 5.2 from the National Oceanic and Atmospheric Administration (NOAA). The snowfall is from 12 to 110 inches.



(a) 2008 year



(b) 2010 year

Figure 5.1: Number of Decks according to deck structure type

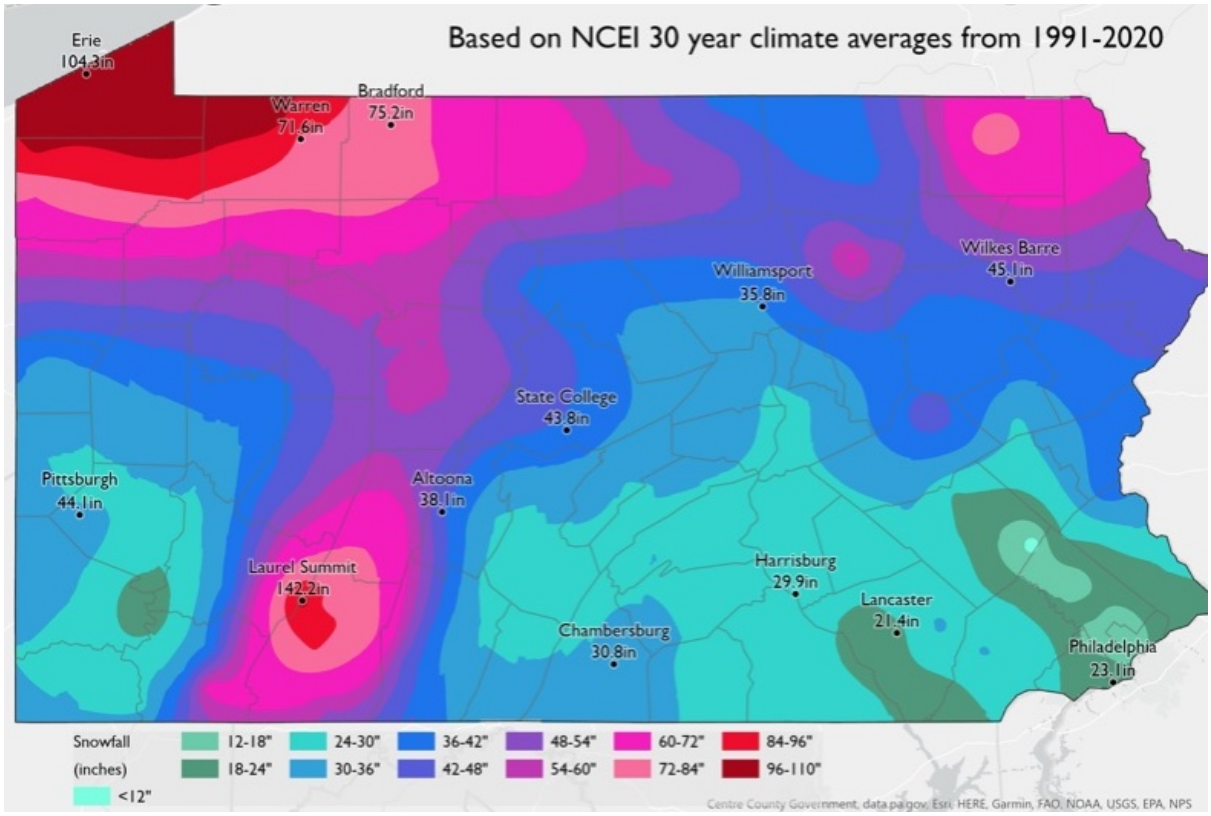


Figure 5.2: Average annual snowfall map in Pennsylvania

5.1.2 Screening

The sample data includes bridges recorded only in both the 2008 and 2010 years. The 54 bridges built in 2008 and 2009 rated CR 9 were included in the sample data, and The 15 bridges rated CR 3 were excluded in sample data. Table 5.1 shows the type and number of discrepancies between the 2008 and 2010 year data. The 1,324 decks were eliminated by screening.

Table 5.1: Summary of discrepancy in the NBI database

Type of Discrepancy	**Number of Discrepancy**
Structure ID	109
Location	457
Year_built	2
Deck structure type	754

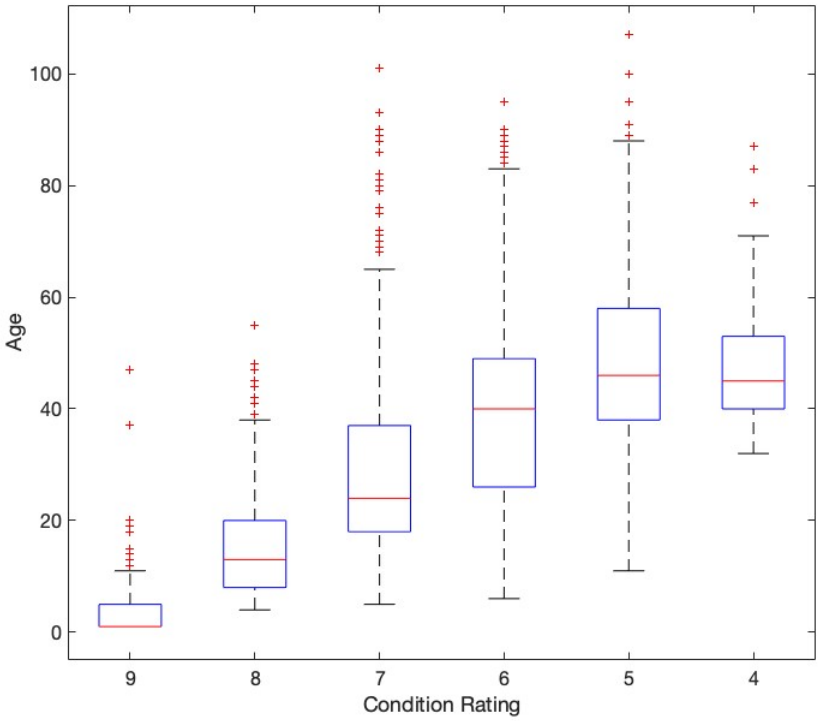
5.1.3 Eliminating Outliers

Figure 5.3 contains the boxplot and the sample data histogram. And Table 5.2 summarizes the values of the boxplot. The 87 decks were eliminated, and the total number of decks in the

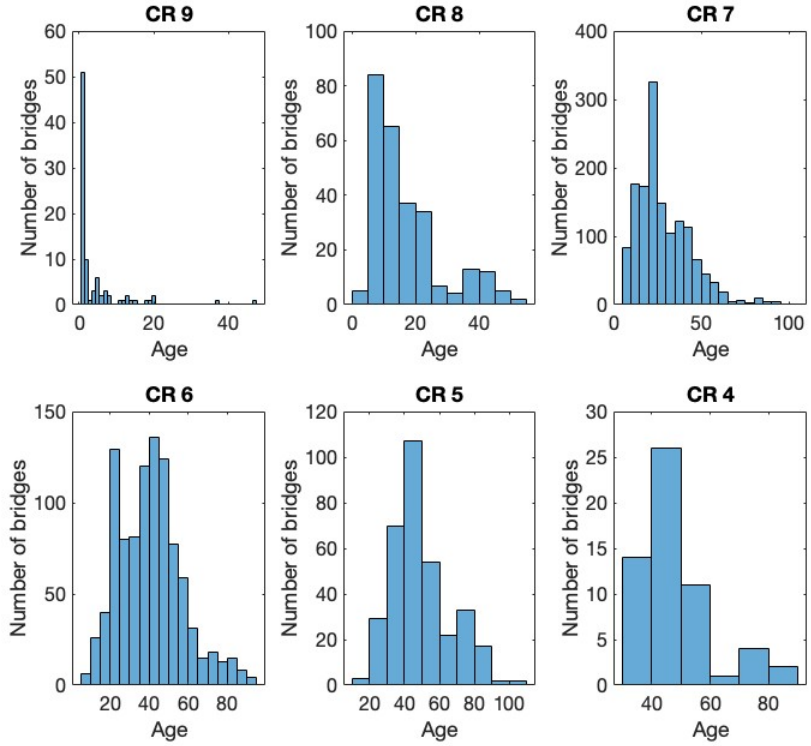
sample data is 3,091.

Table 5.2: Summary of boxplot of sample data in Pennsylvania

	CR 9	**CR 8**	**CR 7**	**CR 6**	**CR 5**	**CR 4**
Maximum age	47	55	101	95	107	87
Minmum age	1	4	5	6	11	32
Number of outliers	11	22	30	15	6	3
Outliers (age)	>= 12	>= 39	>= 68	>= 84	>= 89	>= 77
Upper adjacent (age)	11	38	65	83	88	71
75th percentile (age)	5	20	37	49	58	53
Median (age)		13	24	40	46	45
25th percentile (age)		8	18	26	38	40
Lower adjacent (age)	1	4	5	6	11	32
Interquartile	4	12	19	23	20	13



(a) Boxplot



(b) Number of bridges according to condition rating group

Figure 5.3: Statistical summary of sample data of Pennsylvania

5.2 Single Model Approach

Unlike the TPMs of Texas, a 6 by 6 matrix, TPMs of Pennsylvania use a 7 by 7 matrix, including decks rated at CR4 because the population of CR 4 is not small. The processes of developing the single and proposed models are the same as in case study 1. Therefore, case study 2 presents only the results.

5.2.1 Regression Nonlinear Optimization

The regression curve $S(t)$ in a red line and the means at given CRs (blue dots) of the sample data are shown in Figure 5.4. The equation of third polynomial is $S(t) = 9 - 1.885(10^{-5})t^3 + 0.00297t^2 - 0.1617t$. The transition probability matrix (TPM) is

$$\begin{bmatrix} 0.83902 & 0.16098 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0.88883 & 0.11117 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0.88984 & 0.11016 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0.99783 & 0.00217 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0.94951 & 0.05049 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0.45912 & 0.54088 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

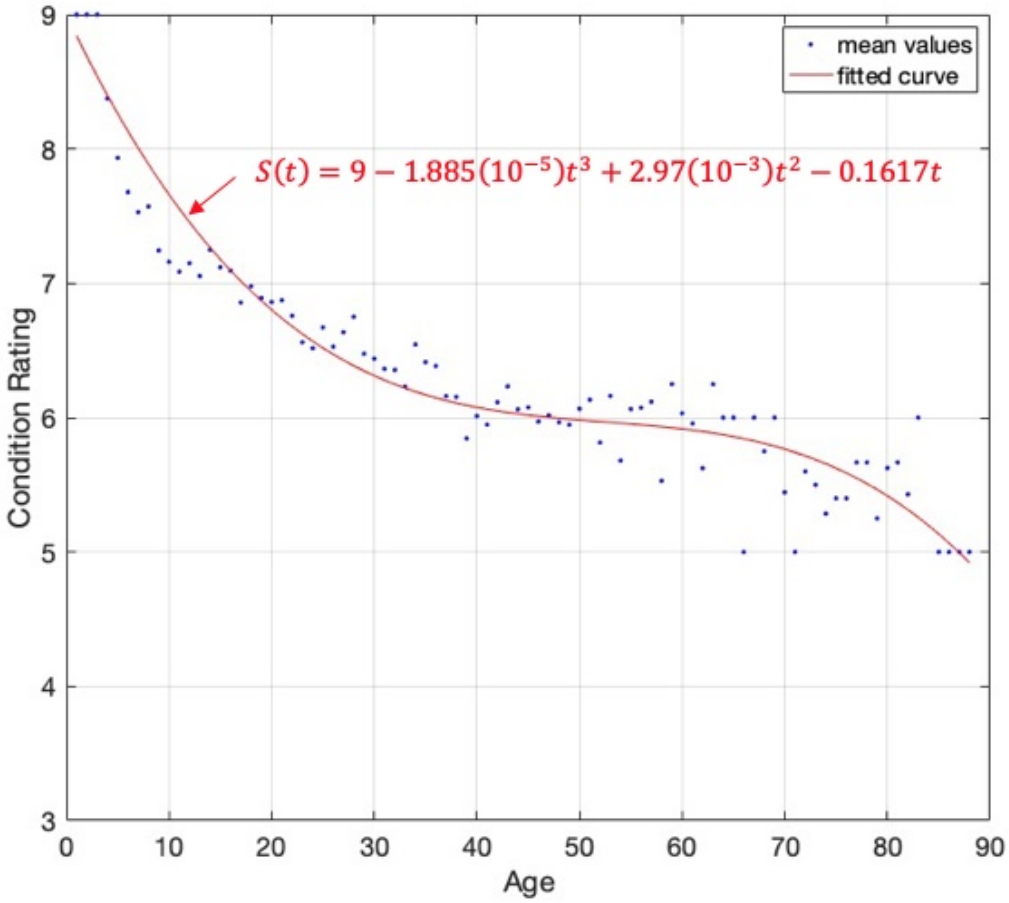


Figure 5.4: Regression curve of Pennsylvania

5.2.2 Bayesian Maximum Likelihood

The transition probability matrix (TPM) is estimated by optimizing the logarithm likelihood function, Equation 2.7.

$$\begin{bmatrix} 0.69842 & 0.30158 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0.85942 & 0.14058 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0.96773 & 0.03227 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0.98101 & 0.01899 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0.98943 & 0.01057 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0.99999 & 0.00001 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

5.2.3 Ordered Probit Model, Poisson, and Negative Binomial Regression

The transition probability of CR 5 of OPM and PR is less than 1, and NBR is 1. It means a concrete deck stays at CR 5, not transitioning to the lower condition rating even over time. The transition probability matrix (TPM) of OPM is

$$\begin{bmatrix} 0.84579 & 0.15421 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0.77801 & 0.22199 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0.92987 & 0.07013 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0.94219 & 0.05781 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0.96000 & 0.04000 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

The transition probability matrix (TPM) of PR is

$$\begin{bmatrix} 0.38364 & 0.61636 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0.66279 & 0.33721 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0.84228 & 0.15772 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0.90189 & 0.09811 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0.95349 & 0.04651 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

The transition probability matrix (TPM) of NBR is

$$\begin{bmatrix} 0.71458 & 0.28542 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0.68375 & 0.31625 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0.98327 & 0.01673 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0.99993 & 0.00007 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

5.2.4 Proportional Hazard Model

The transition probability of PHM is estimated with two components, baseline probability, and hazard ratio. Empirical cumulative distribution functions of condition rating groups are shown in the Figure 5.5 used to obtain baseline probabilities of each condition rating. Table 5.3 includes the baseline transition probabilities (survival probabilities) and hazard ratios (HRs) at condition ratings.

Table 5.3: Baseline probability and Hazard ratio (HR)

	CR 9	**CR 8**	**CR 7**	**CR 6**	**CR 5**	**CR 4**
Survival probability	0.97368	0.98361	0.98734	0.98810	0.98876	0.98611
Hazard Ratio (HR)	3.02559	2.14609	1.39565	1.05736	0.70798	1

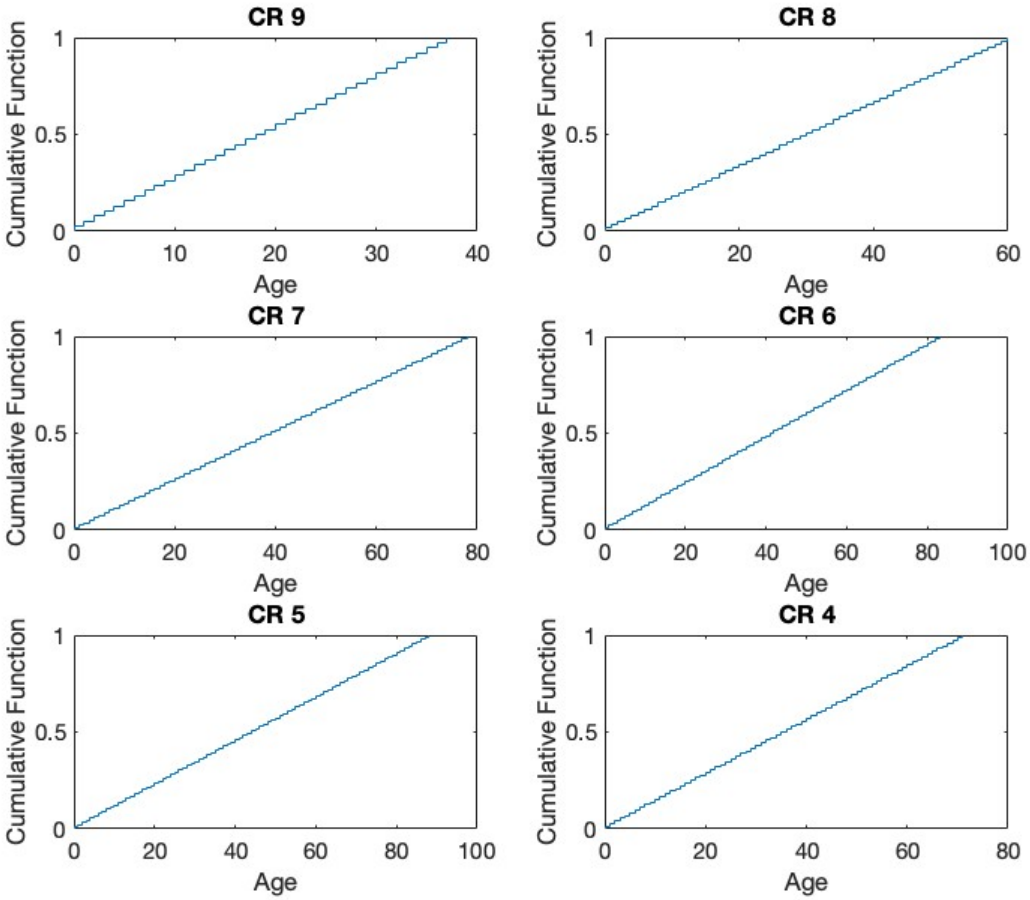


Figure 5.5: Empirical cumulative distribution functions

$$\begin{bmatrix} 0.92248 & 0.07752 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0.96515 & 0.03485 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0.98238 & 0.01762 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0.98742 & 0.01258 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0.99203 & 0.00797 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0.98611 & 0.01389 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

5.2.5 Time-based Weibull deterioration Model (WDM)

Figure 5.6 shows the histograms of condition rating groups including the scale (η) and shape (β) parameters. These parameters are used to calculate the mean duration of staying at a particular condition rating.

Figure 5.7 contains the regression curve $S(t)$ in red line. The blue dots are the cumulative mean times at condition ratings. The equation of regression curve is $S(t) = 9 - 1.102(10^{-}6)t^3 + 0.0003786t^2 - 0.061t$. The transition probability matrix (TPM) of WDM is

$$\begin{bmatrix} 0.93927 & 0.06075 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0.95254 & 0.04746 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0.96182 & 0.03818 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0.99417 & 0.00583 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0.50746 & 0.49254 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0.50288 & 0.49712 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

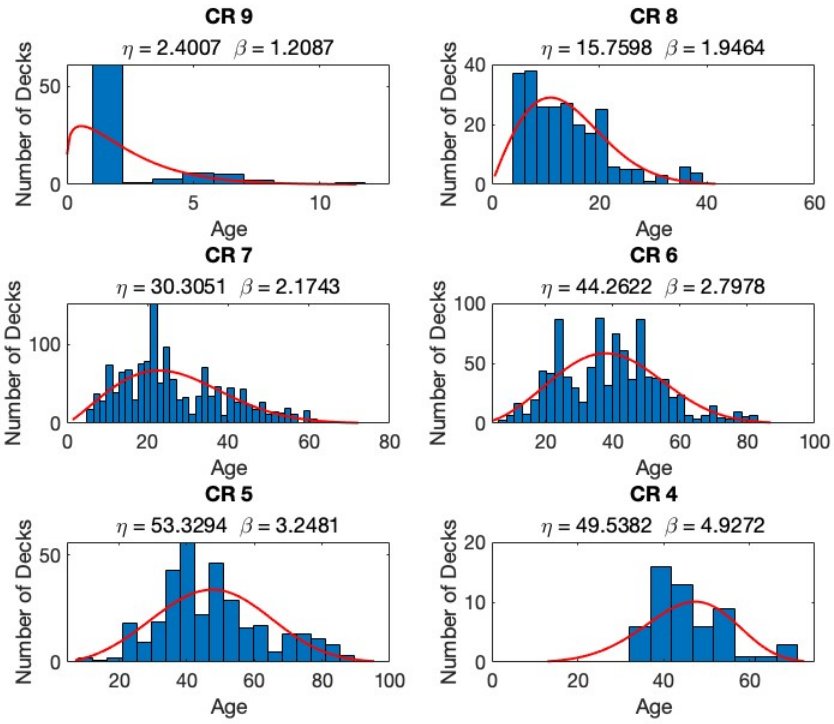


Figure 5.6: Histogram including shape and scale parameters

5.2.6 Mechanistic-based deterioration Model (MDM)

The same parameters and regression curve as the ones in case study 1 are used to estimate transition probabilities. The transition probability matrix (TPM) of MDM is

$$\begin{bmatrix} 0.13234 & 0.86766 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0.24640 & 0.75360 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0.38863 & 0.61137 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0.27332 & 0.72668 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0.89936 & 0.10064 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0.89548 & 0.10452 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

5.2.7 Deterioration curves of single models

Figure 5.8 shows the deterioration curves of all single models. The gray circles are the condition rating distribution of sample data.

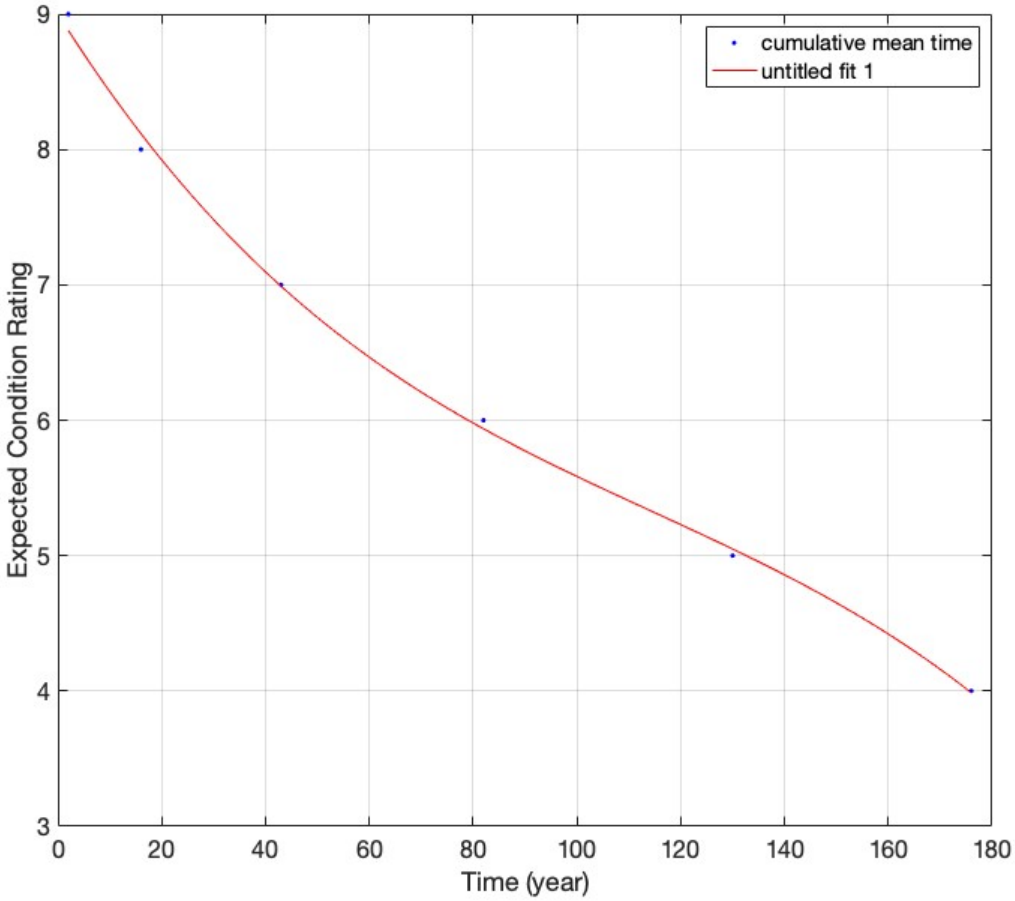


Figure 5.7: Regression curve

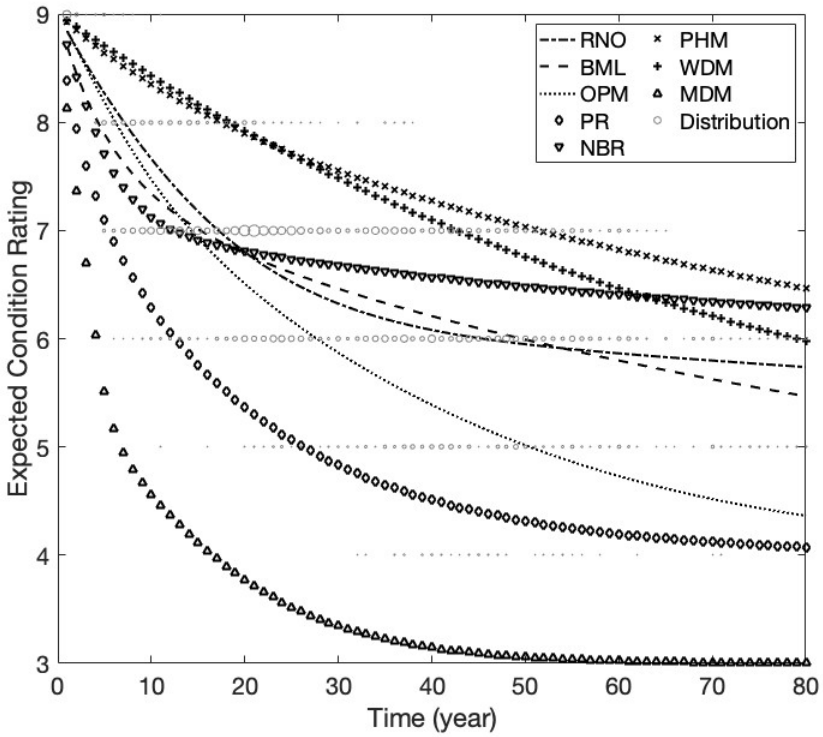


Figure 5.8: Deterioration curves estimated from single models

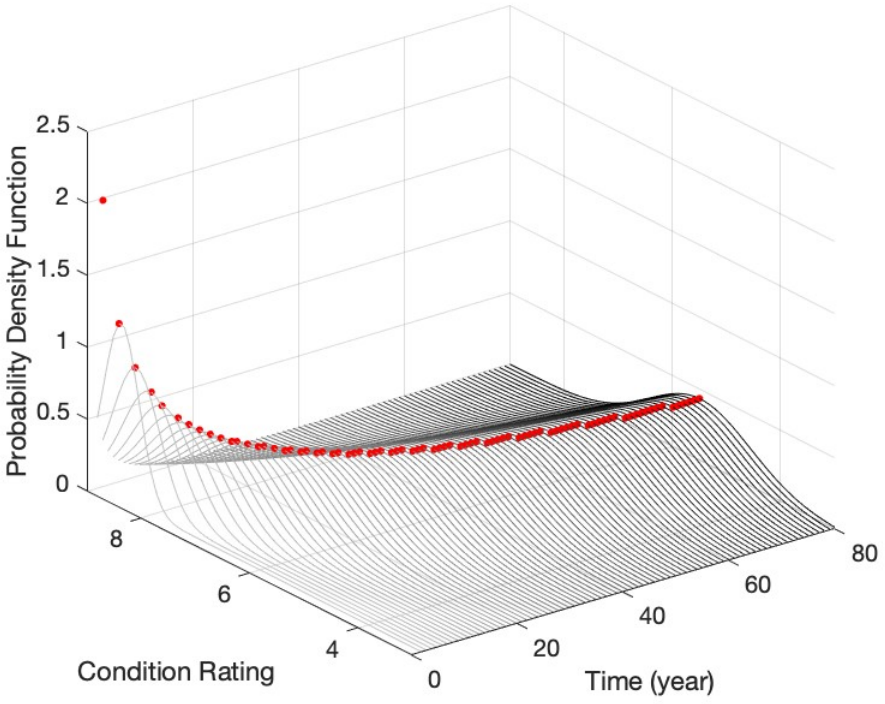
5.3 Multiple Model Approach

Four deterioration models are proposed; the first model includes all the single models and uses method 1. The second model includes the same models as the first model but uses method 2. The third model uses method 1, including all the single models except the mechanistic deterioration model (MDM). The fourth model uses method 2, including the same models as the third model.

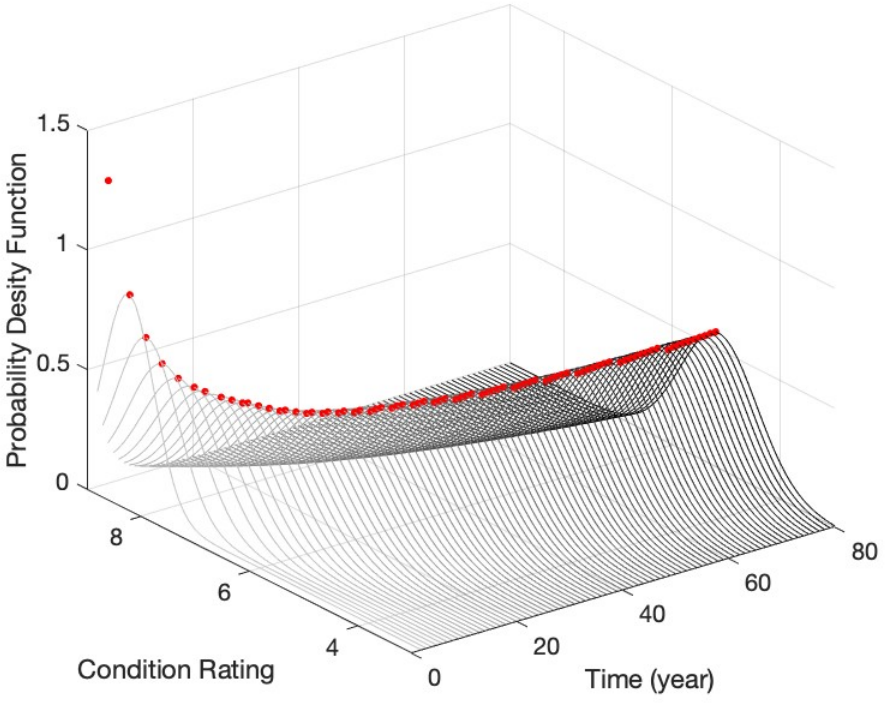
5.3.1 Result/Comparison

Figure 5.9 shows the 3D deterioration curve plots the four proposed models. The red dots are the means of probability density functions at given times and are the expected condition ratings at given times. The shape of the first and third models and the second and fourth models are similar. Figure 5.10 contains the 2D deterioration curves of the models, including possible condition rating distributions (circles). The possible condition ratings at a given time differ between the first and second models. For example, in the first model, condition rating 8 presents

continuously for up to 80 years. However, in the second model, condition rating 8 presents up to 50 years. Figure 5.11 includes only deterioration curves of the models to compare. The first and second models have almost the same deterioration rates. Also, the deterioration rates of the third and fourth models are almost identical.

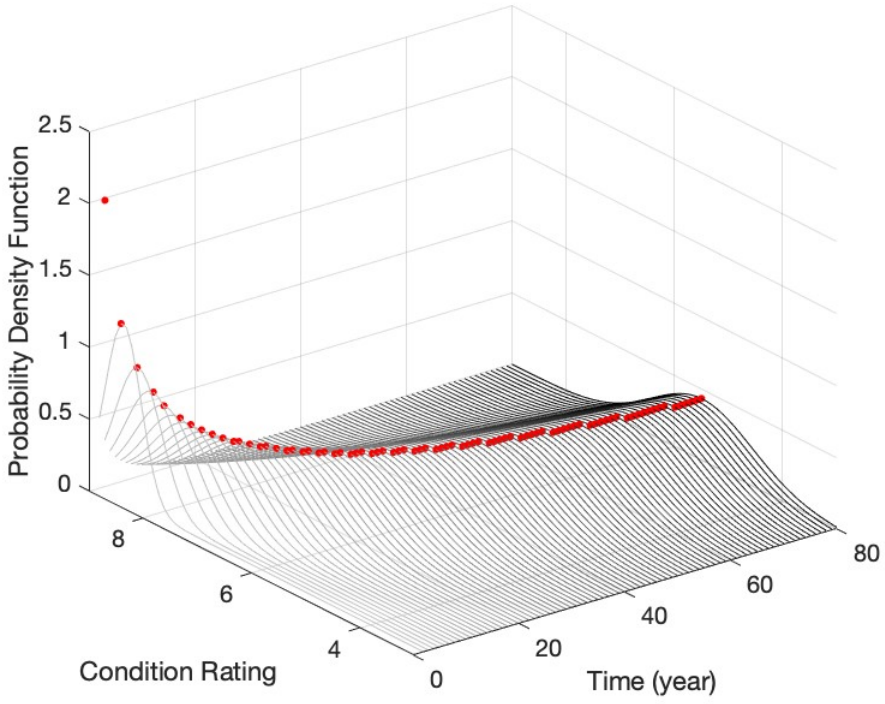


(a) First proposed model

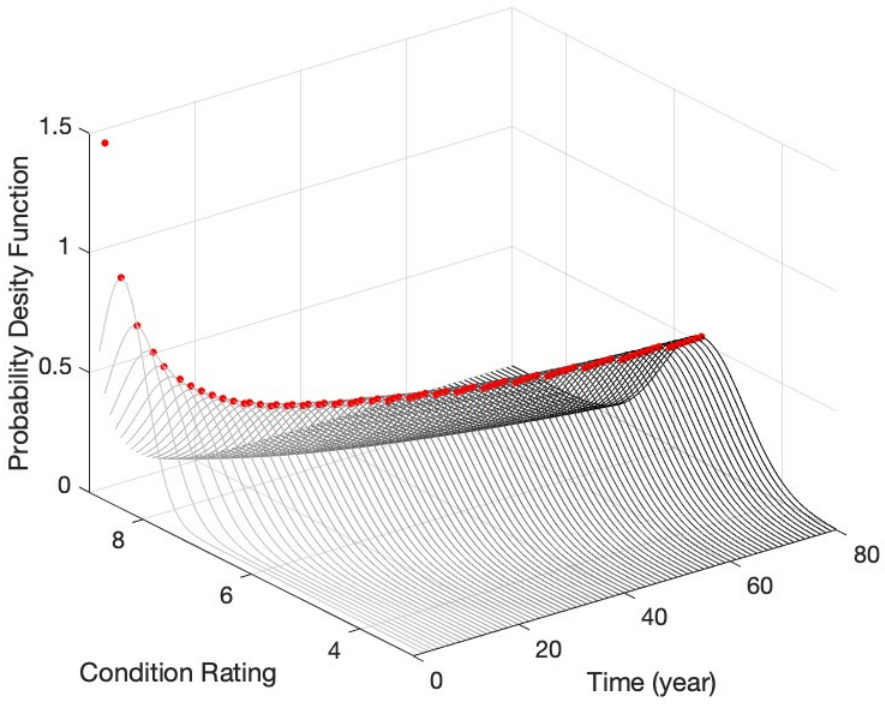


(b) Second proposed model

Figure 5.9: Probability density function of proposed multiple models

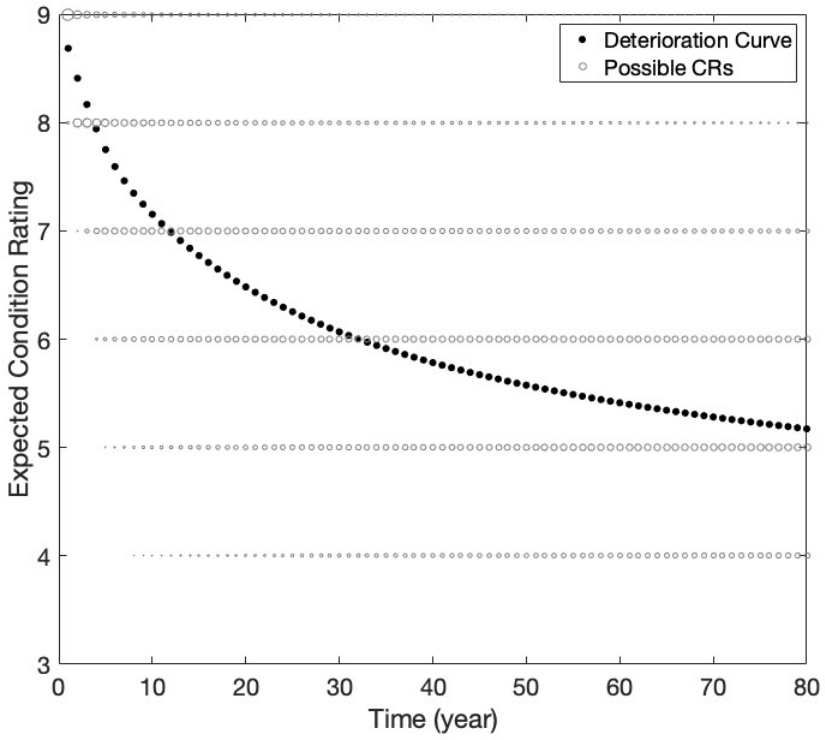


(c) Third proposed model

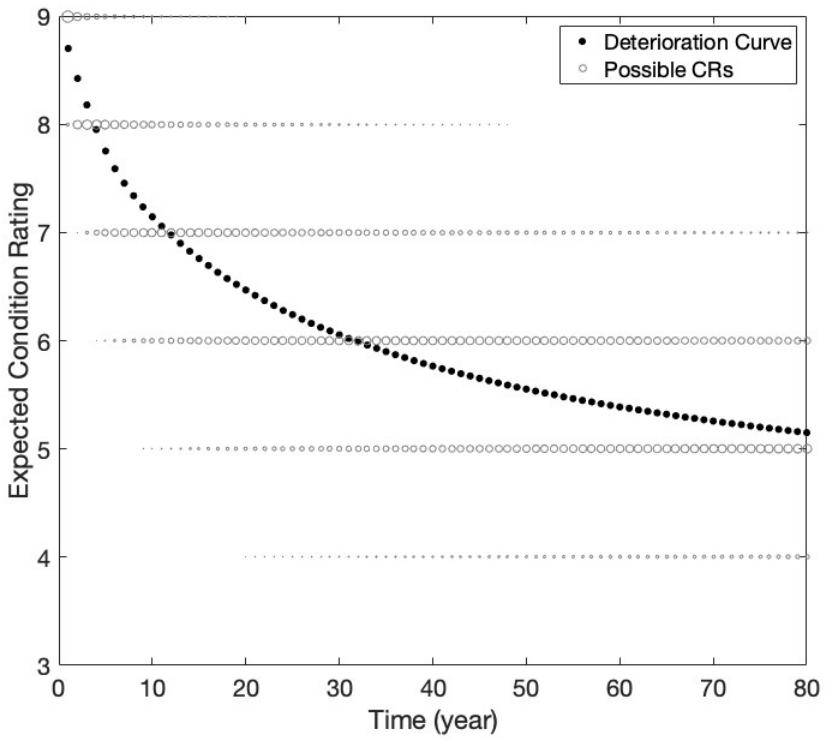


(d) Fourth proposed model

Figure 5.9: Expected condition ratings of proposed models (cont.)

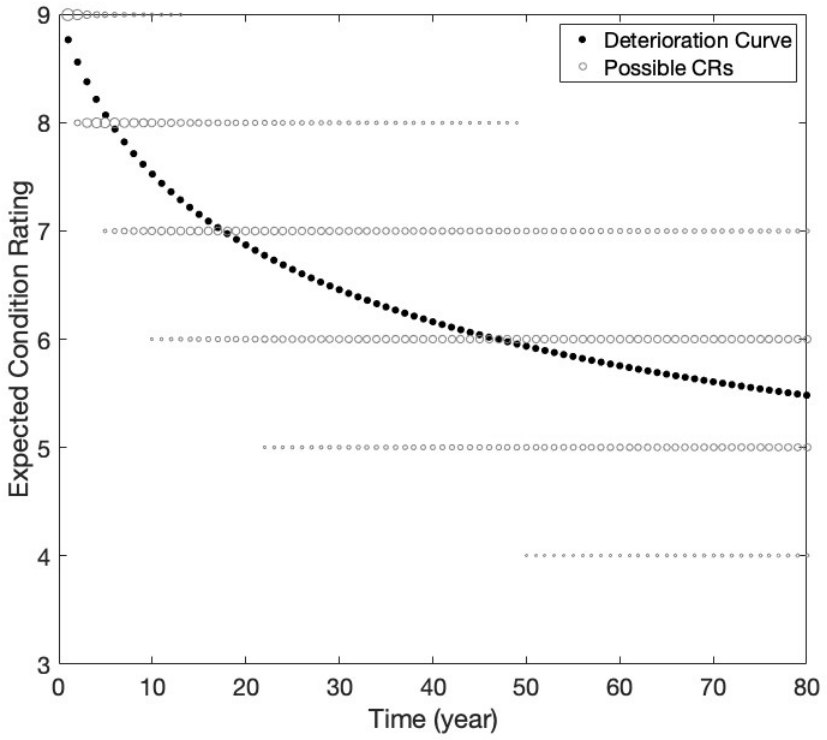


(a) First proposed model

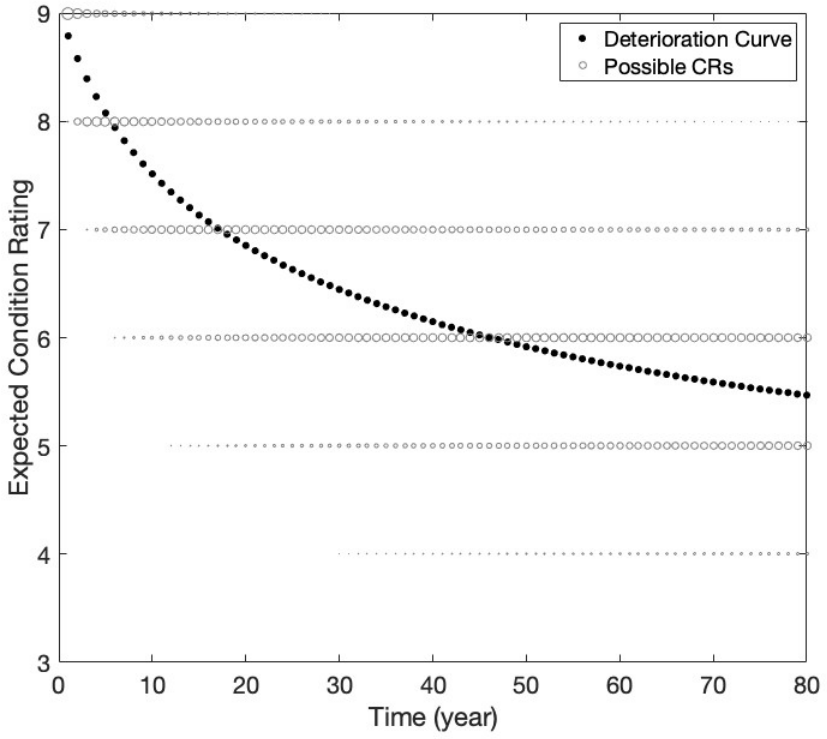


(b) Second proposed model

Figure 5.10: Expected condition ratings of proposed models



(c) Third proposed model



(d) Fourth proposed model

Figure 5.10: Expected condition ratings of proposed models (cont.)

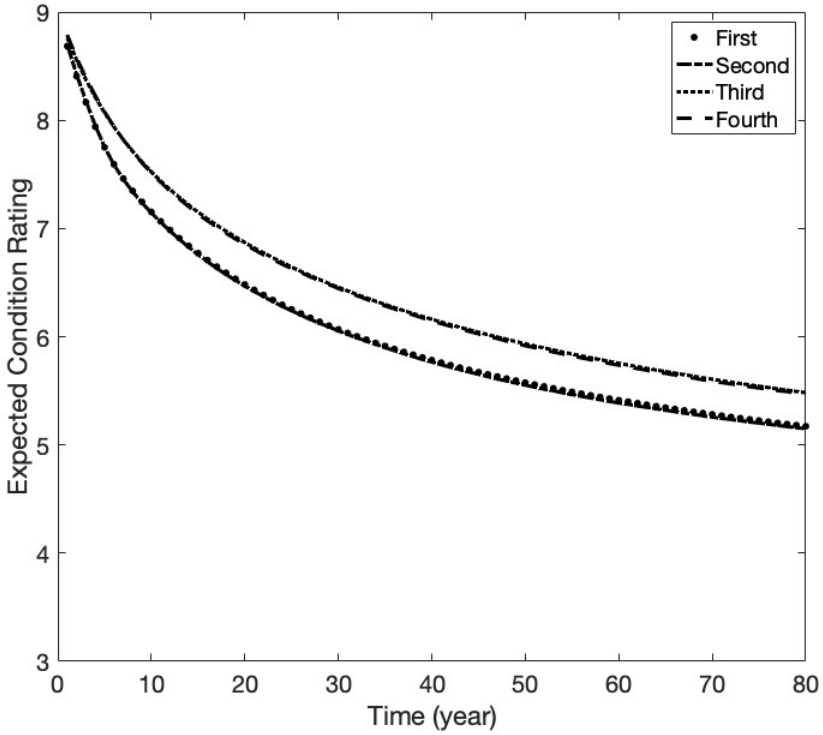


Figure 5.11: Deterioration curves of proposed models

5.3.2 Evaluation

The prediction accuracy of the proposed models are evaluated by comparing them to the sample data. Figure 5.12 contains the means of condition ratings and condition rating distribution of the sample data (2010) at given times. Unlike Texas, the means at given times degrade over time despite some fluctuations. Figure 5.13 shows the most likely expected and observed condition ratings. The blue circles are the first model, the red circles are the second model, the green circles are the third model, the magenta circles are the fourth model, and the black dots are the sample data. The observed condition ratings in the range of over 60 years old are 7, 6, or 5. In the range, the expected condition rating from the first and second models is 5, and from the third and fourth models is 6. The prediction error in most likely condition rating at a given time is calculated as the difference between the most likely expected and observed condition ratings shown in Figure 5.14. The zero of the prediction error means that the observed and expected condition ratings are identical at given times. The negative (positive) value means that the expected condition rating is higher (lower) than the observed one. The third (orange

bars) and fourth (purple bars) models have less prediction errors than the first (blue bars) and second (red bars) models overall.

Table 5.4 summarizes the mean absolute error (MAE) values of the models, including expected condition ratings according to condition ratings from 9 to 4, most likely condition ratings, means, and expected condition rating distribution of 2010. The expected condition rating distribution of decks of the network system can be estimated from state vectors at given times. The second and fourth models can be used to estimate the future condition rating distribution because the models use TPMs that calculate state vectors, unlike the first and third models uses expected condition ratings that estimate future condition ratings without the step of obtaining state vectors. In the expected condition rating, the MAE values of the third and fourth models are less than the ones of the other two models. The MAE value of the fourth model is less than that of the second model in the expected condition rating distribution. However, the results are reasonable based on the researcher's findings that there were differences between the expected conditions estimated from the mechanistic deterioration models proposed and the observed bridge conditions (Nickless and Atadero, 2018).

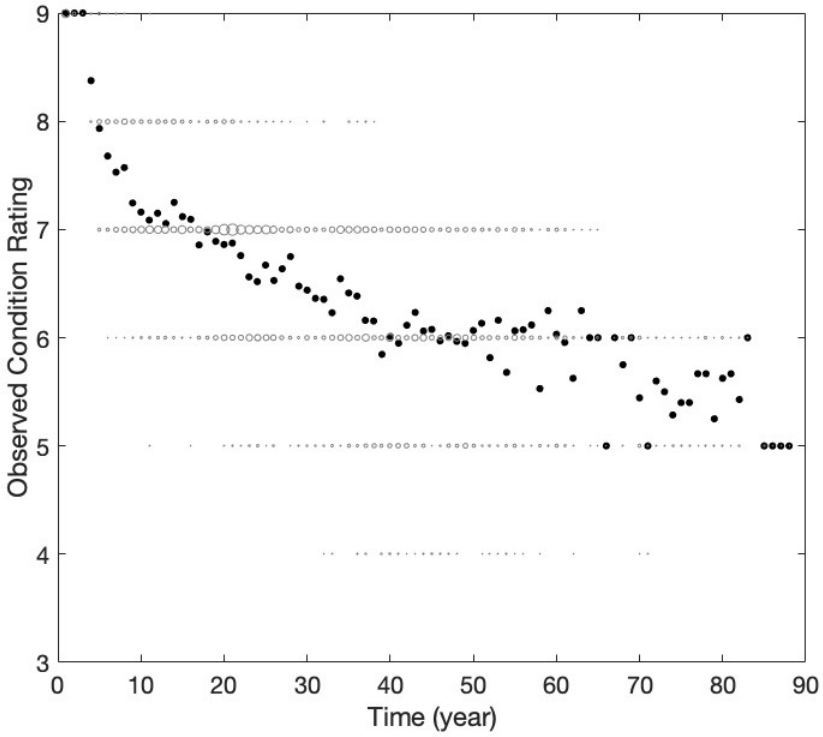


Figure 5.12: Sample data, 2010, Pennsylvania

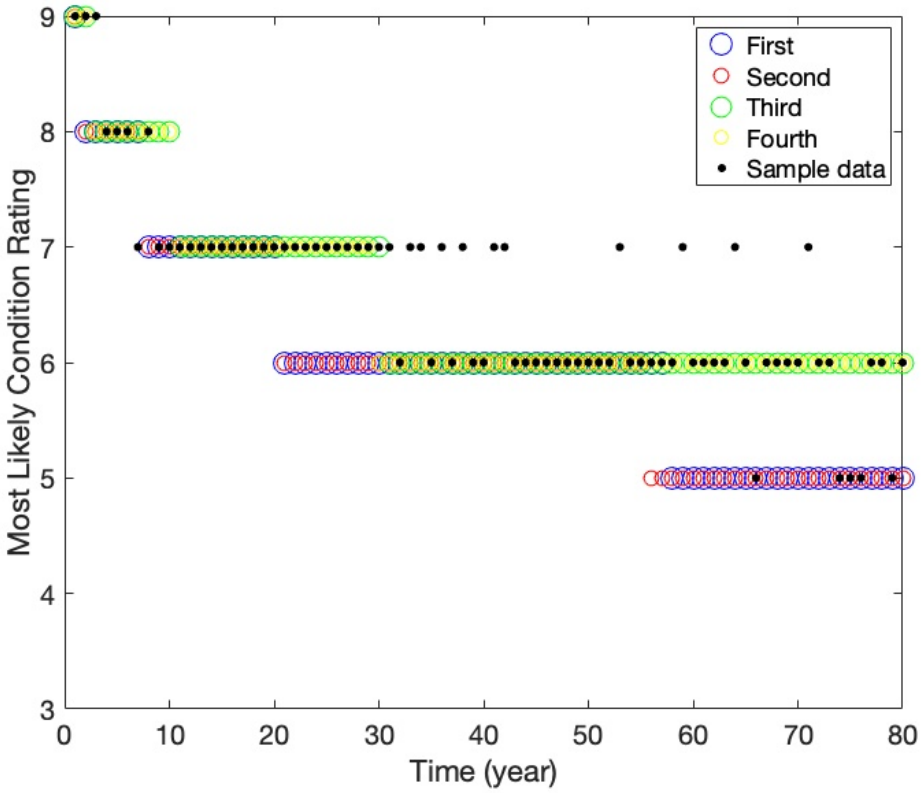


Figure 5.13: Compare most likely condition ratings

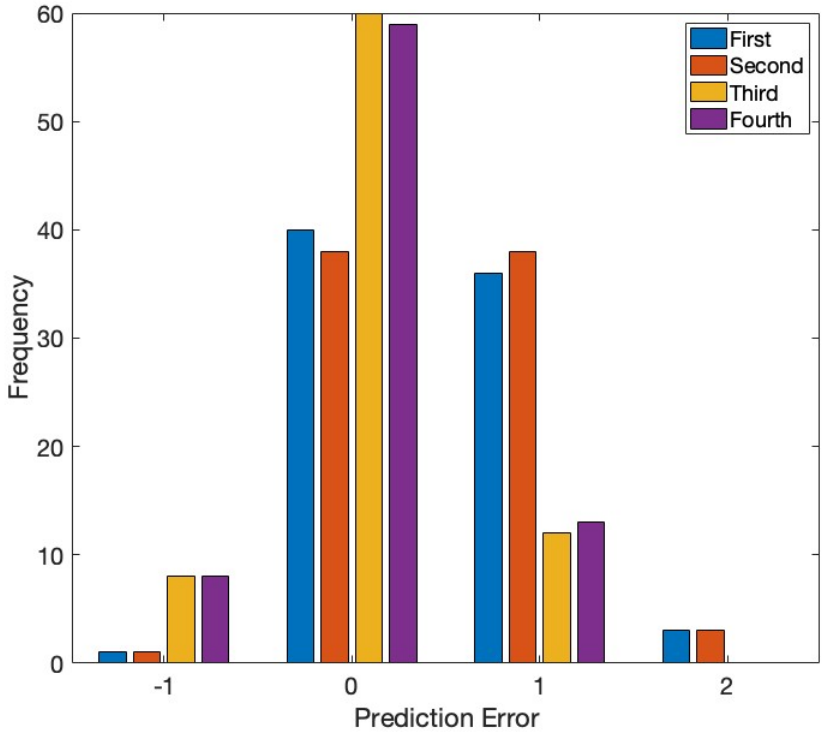


Figure 5.14: Prediction errors of proposed models

Table 5.4: MAE values of the proposed models

	First	**Second**	**Third**	**Fourth**
Expected Condition Rating				
CR 9	0.05	0.02	0.03	0.03
CR 8	0.08	0.03	0.08	0.06
CR 7	0.19	0.16	0.13	0.12
CR 6	0.16	0.14	0.10	0.10
CR 5	0.12	0.13	0.08	0.08
CR 4	0.10	0.04	0.04	0.03
Overall CR	0.12	0.09	0.07	0.06
Most likely CR	0.50	0.53	0.03	0.04
Mean	0.37	0.39	0.18	0.18
Expected Condition Rating Distribution				
2010 year		0.04		0.03

Chapter 6: Discussion, Conclusion, Future work

This research includes an analysis of existing and common approaches to deterioration modeling of bridges that rely on stochastic approaches. In particular, state-based approaches leverage a Transition Probability Matrix or TPM. There was variability between the models that calls into question our ability to use these models confidently to make predictions. If all models are reasonable, but all models produce different results, which model do we use?
This research presented a simple, novel approach to combining deterioration models, which inherently reduces the uncertainty of using a single model. The multiple-model approach combines various models, including state-based Markov, time-based Weibull, and mechanistic-based models. The results show good prediction accuracy and, more importantly, reflect the uncertainty associated with all the viable models that could be used. This work is a preliminary step towards a more robust multiple-model deterioration approach that considers more than stochastic models.

The practical application of current deterioration models, including state-based Markovian deterioration models, time-based Weibull deterioration model, and mechanistic-based deterioration models, does not include analytical assessment of different model approaches relative to one another (i.e., comparison of the results of each model obtained under the same inputs to the observed data), the influence of explanatory variable selection or the relative applicability of the model to the desired outputs. This research has demonstrated that these inputs (i.e., the effect of explanatory variables) influence the ”accuracy” of the results, which is the end user's concern.

A flexible framework that utilizes multiple deterioration model approaches is proposed to address the deleterious effects of an arbitrarily-selected population of bridges (i.e., explanatory variables) in single deterioration modeling.

This novel approach is built on the following key characteristics:

- The basic actionable information from this multiple-model framework remains a single deterioration curve. However, this curve is constructed considering the TPMs of two or more deterioration modeling approaches. The difference from current models is that it

contains a range of possible condition ratings besides the expected condition rating at a given time.

- The time-based Weibull and mechanistic deterioration models in the literature estimate only the deterioration curve, not transition probability matrix. This research devises a method to generate a TPM from the deterioration curves obtained from these two models for integration with other models.

- The framework combines models by producing expected condition ratings or transition probabilities from all the component TPMs. This research presents two integration methods to estimate the future condition rating of a component. One of the methods (method 2) is to estimate the future condition rating distribution of components in a network system.

- The framework retains the information from each component model. This results in a combined deterioration prediction that inherently reflects the variability across models and the flexibility to explore that variation directly.

- The framework explores model performance versus explanatory variable selection to achieve more ”accurate” results. It does not guarantee more accurate results by applying the effect of the explanatory variables. For example, in the case studies, proposed models, which contain the effect of the environmental condition (including a mechanistic-based deterioration model), offers less ”accurate” results than a proposed models that excluded it. In case study 2, the sample data has a condition rating of 7, 6, or 5 over 60 years old. The proposed model, excluding the mechanistic-based model, predicts a condition rating of 6, and the proposed model, including the mechanistic-based model, estimates a condition rating of 5. The bridge component conditions are rated by visual inspection biannually. Since the visual inspection can be subjective, the recorded condition data can also be subjective.

- The framework demonstrates the reduction of uncertainty and bias in modeling. Researchers agreed that there were differences between the predicted future conditions and

the observed conditions due to the need for more accurate field data and the complexity of mechanistic-based deterioration modeling. Assumptions are necessary to create a mechanistic-based deterioration model, and the assumptions differ depending on the researchers. There is uncertainty and bias in modeling, creating differences between predicted and observed condition states. The expected condition ratings obtained from the proposed model are closer to observed condition ratings than those obtained from only the mechanistic-based model.

- The framework can be more robust by integrating deterministic and different mechanistic models. The mechanistic deterioration models can be more reliable by obtaining better field data using nondestructive engineering and LiDAR techniques. The framework can be used in bridge element deterioration modeling.

Appendix A: The NBI items

Table A.1: Item 31 - Design Load

Code	Metric Description
1	M 9
2	M 13.5
3	MS 13.5
4	M 18
5	MS 18
6	MS 18 + Mod
7	Pedestrian
8	Railroad
9	MS 22.5
0	Other or Unknown

Table A.2: Item 41 - Structure Open, Posted, or Closed to Traffic

Code	Description
A	Open, no restriction
B	Open, posting recommended but not legally implemented
D	Open, would be posted or closed except for temporary shoring
E	Open, temporary structure in place to carry legal loads
G	New structure not yet open to traffic
K	Bridge closed to all traffic
P	Posted for load
R	Posted for other load-capacity restriction

Table A.3: Item 75B - Work Done by

Code	Description
1	Work to be done by contract
2	Work to be done by owner's forces

Table A.4: Item 107 - Deck Structure Type

Code	Description
1	Concrete Cast-in-Place
2	Concrete Precast Panels
3	Open Grating
4	Closed Grating
5	Steel plate
6	Corrugated Steel
7	Aluminum
8	Wood or Timber
9	Other
N	Not applicable

Appendix B: Data Analysis

The tables summarize the boxplots of zone 2, 3, and 4. The figures are a boxplot and histogram of sample data of zone 2, 3, and 4. In the boxplot, the x-axis is condition rating from 9 to 5, and the y-axis is age. The red plus is an outlier. This plot shows a median age at a particular condition rating. The median age at CR 6 and 5 are identical in all zones. Considering the median, a component in zone 4 reaches CR 6 and 5 earlier than other zones. The histogram shows the number of components at a given age according to condition rating.

Table B.1: Summary of boxplot, Zone 2

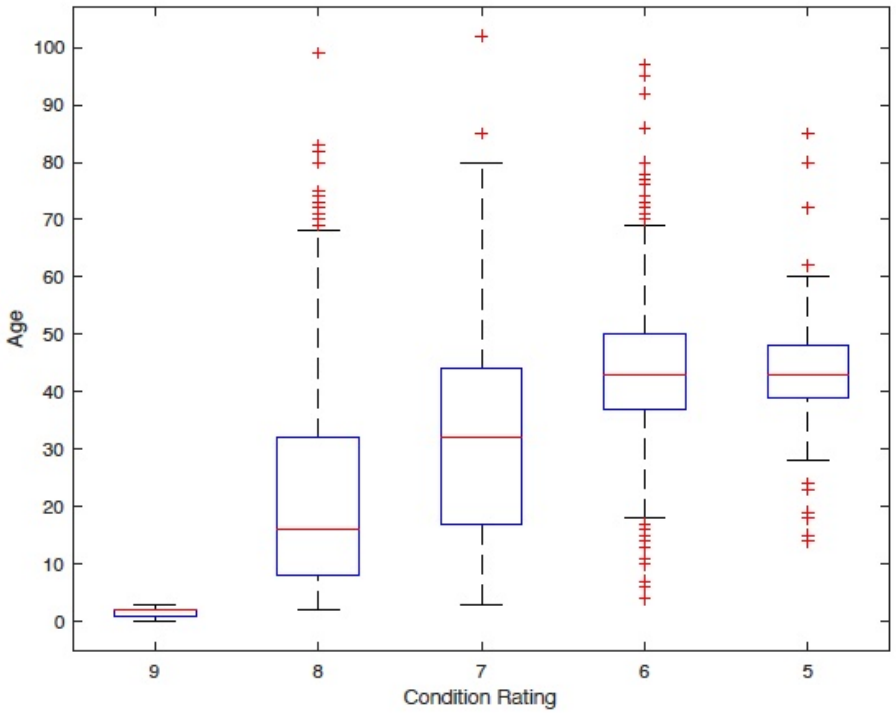
	CR 9	**CR 8**	**CR 7**	**CR 6**	**CR 5**
Maximum age	3	99	102	97	85
Minimum age	0	2	3	4	14
Number of outliers	0	29	4	54	13
Outliers (age)		>=69	>=85	>=70, <=17	>=62, <=24
Upper adjacent (age)	3	68	80	69	60
75th percentile (age)		32	44	50	48
Median (age)	2	16	32	43	43
25th percentile (age)	1	8	17	37	39
Lower adjacent (age)	0	2	3	18	28
Interquartile	1	24	27	13	9

Table B.2: Summary of boxplot, Zone 3

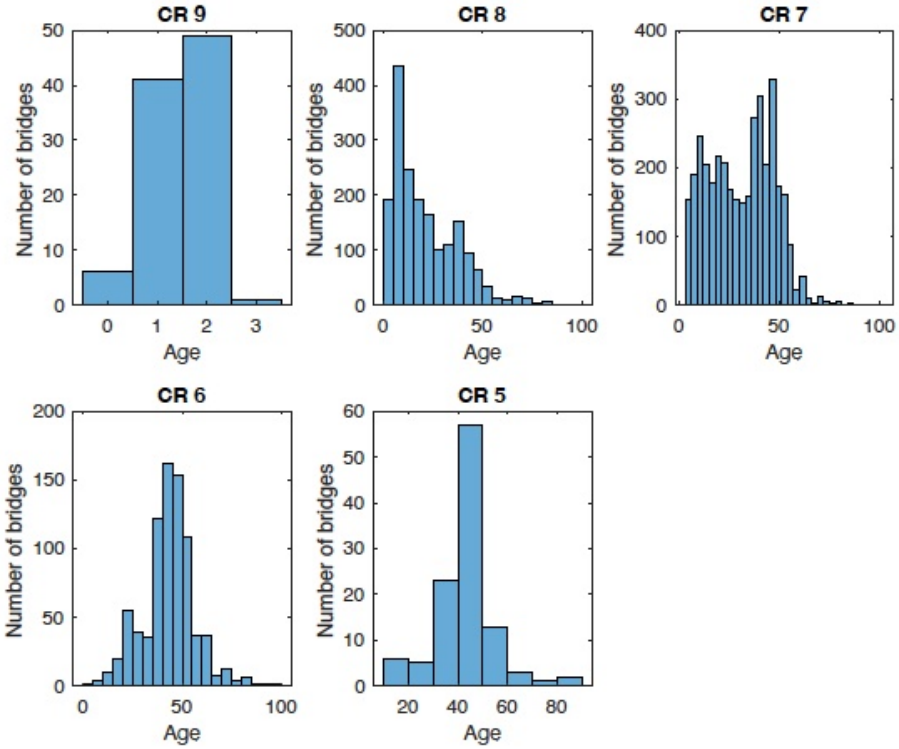
	CR 9	**CR 8**	**CR 7**	**CR 6**	**CR 5**
Maximum age	3	86	90	88	74
Minimum age	1	3	3	7	10
Number of outliers	0	14	18	17	12
Outliers (age)		>=60	>=79	>=74, <=8	>=70, <=22
Upper adjacent (age)	3	58	78	73	62
75th percentile (age)		27	39	49	50
Median (age)	2	14	23	42	43
25th percentile (age)		6	13	33	39
Lower adjacent (age)	1	3	3	9	25
Interquartile	1	21	26	16	11

Table B.3: Summary of boxplot, Zone 4

	CR 9	**CR 8**	**CR 7**	**CR 6**	**CR 5**
Maximum age	2	55	81	70	54
Minimum age	0	3	3	3	23
Number of outliers	3	0	1	0	0
Outliers (age)	0, 2		81		
Upper adjacent (age)		55	75	70	54
75th percentile (age)		32.75	39	45	47
Median (age)		23	24.5	35	35
25th percentile (age)		7	12	25	24
Lower adjacent (age)	1	3	3	3	23
Interquartile	0	25.75	27	20	23

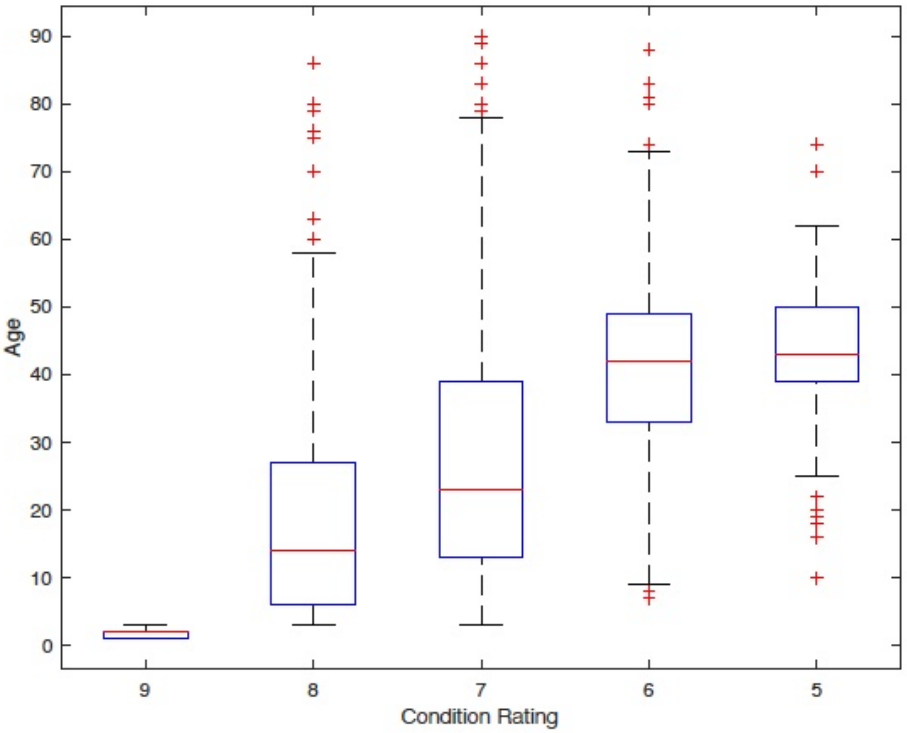


(a) Boxplot

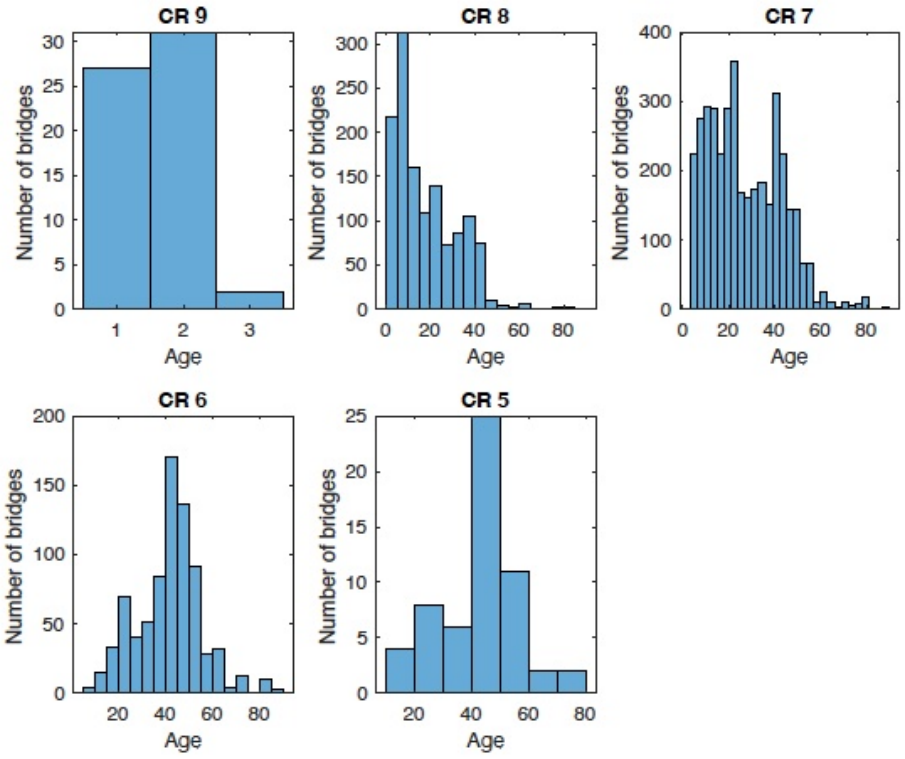


(b) Histogram

Figure B.1: Zone 2 Sample data analysis

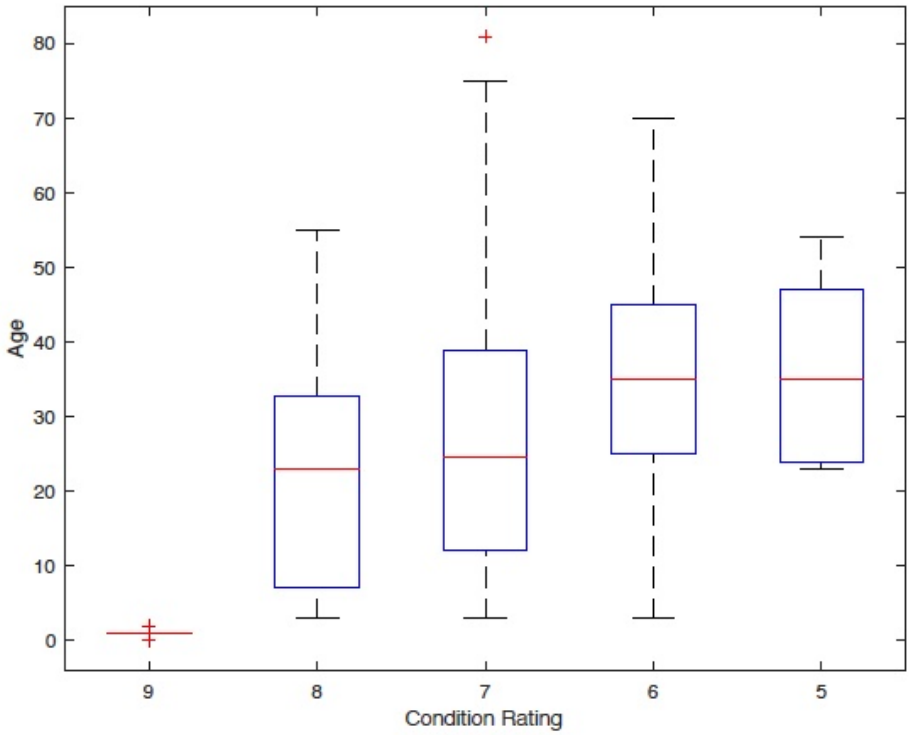


(a) Boxplot

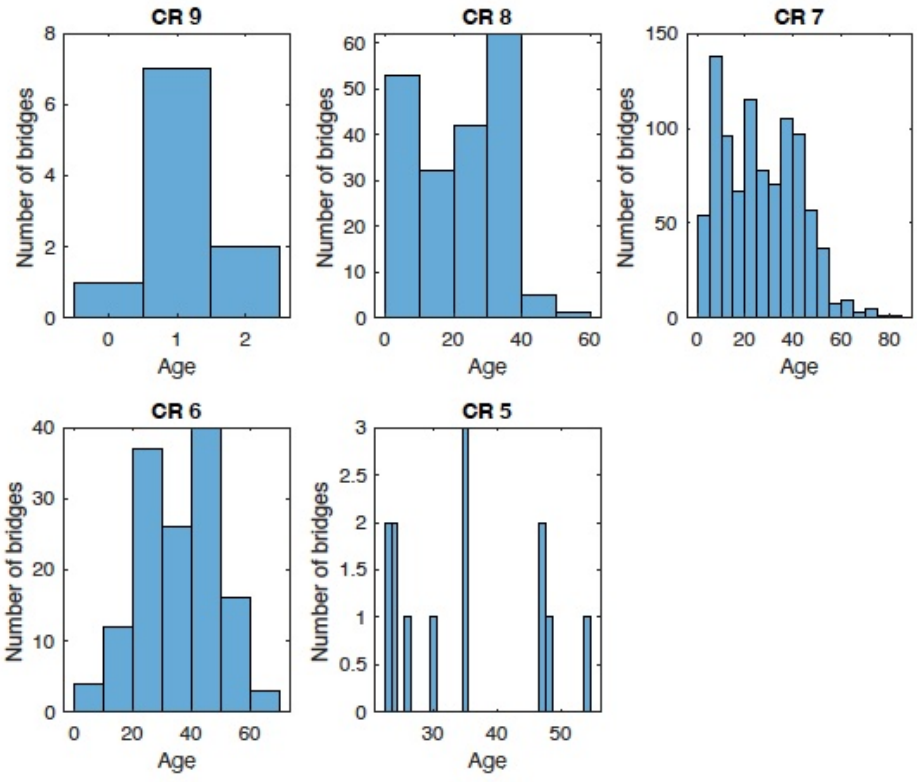


(b) Histogram

Figure B.2: Zone 3 Sample data analysis



(a) Boxplot

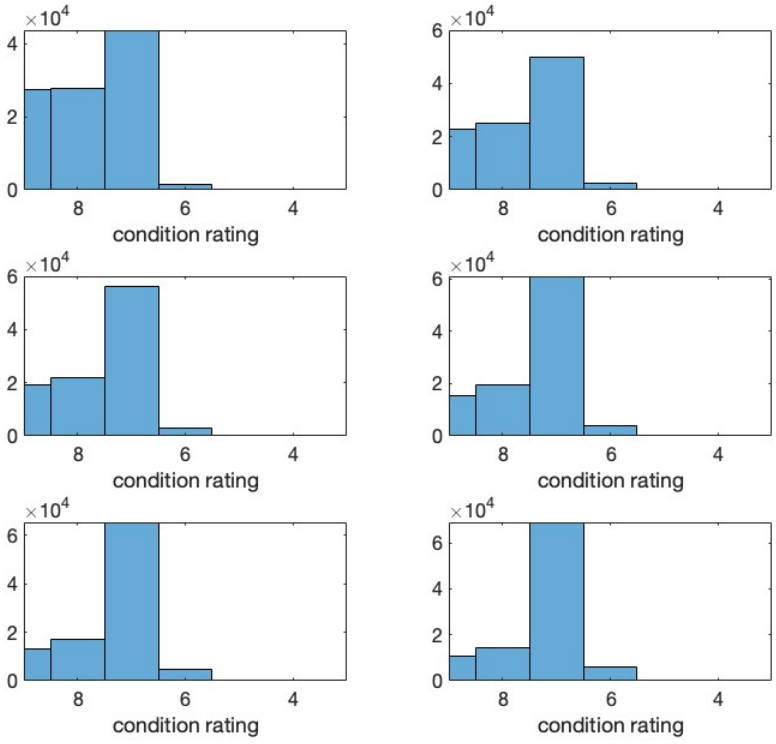


(b) Histogram

Figure B.3: Zone 4 Sample data analysis

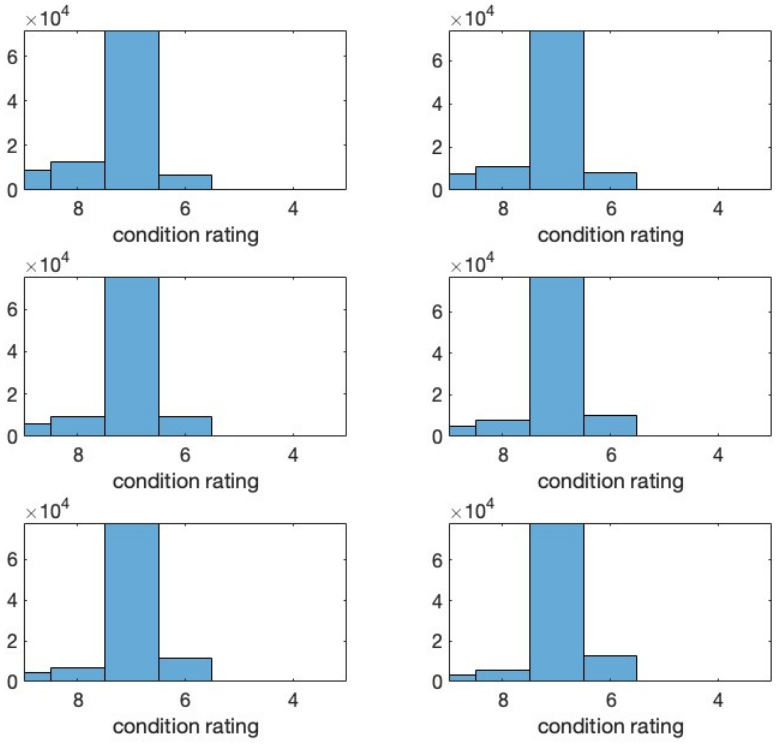
Appendix C: Sample data histogram

Figure C.1 shows a sample of histograms of the sample data of RNO generated using the Monte Carlo simulation with a uniformly distributed random numer (n = 100,000).

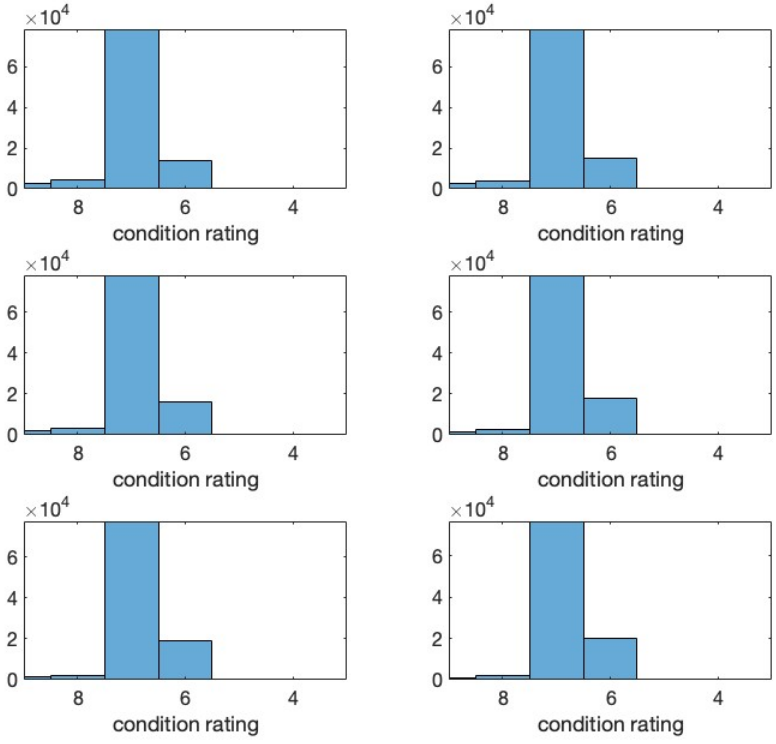


(a) from 7 to 12

Figure C.1: Sample data histograms of RNO of Zone 1

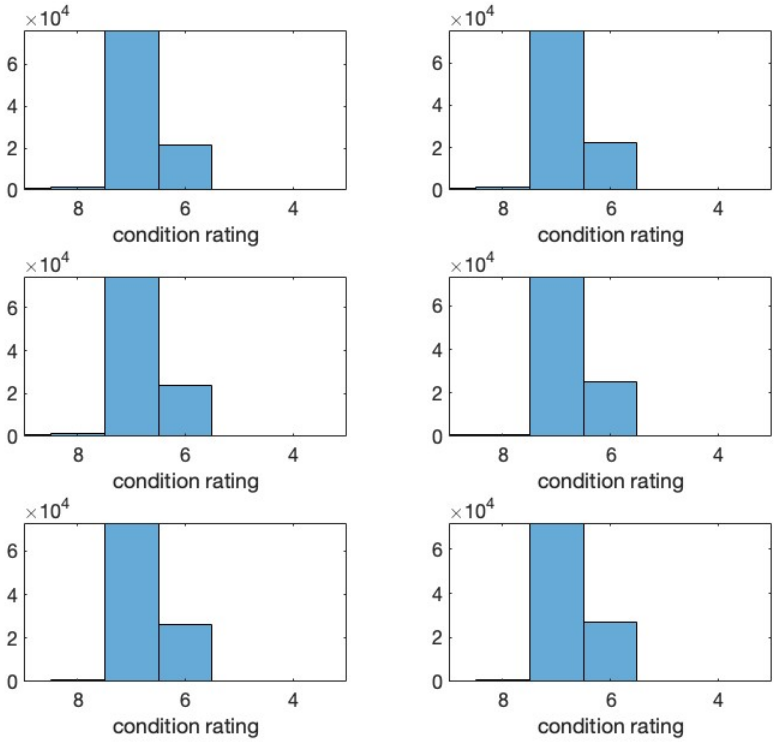


(b) from 13 to 18

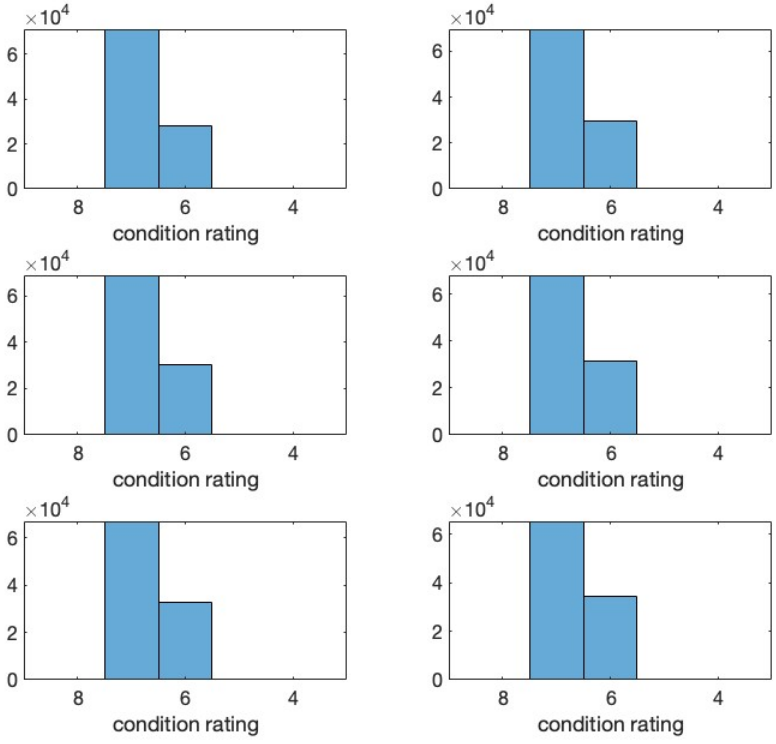


(c) from 19 to 24

Figure C.1: Sample data histograms of RNO of Zone 1 (cont.)

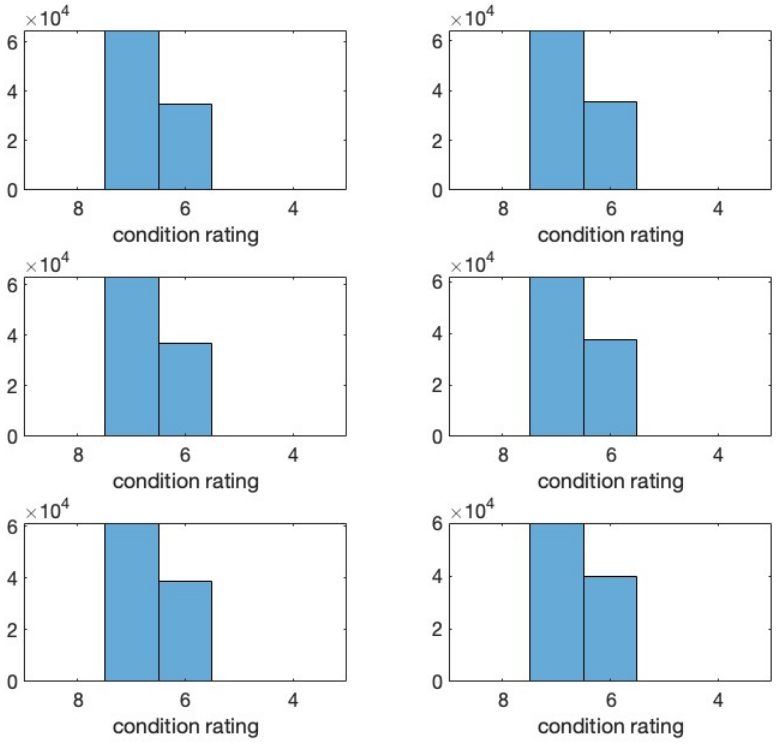


(d) from 25 to 30

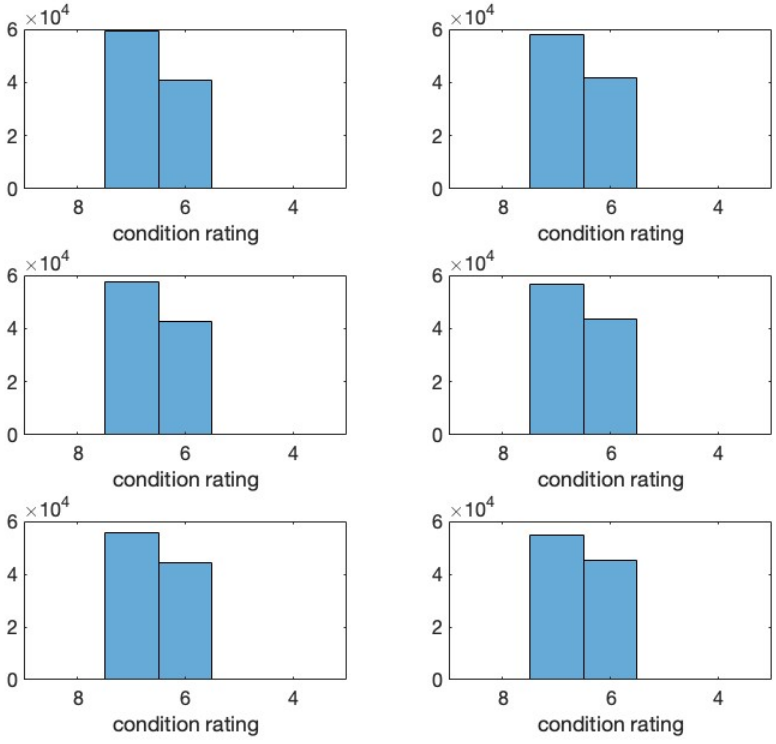


(e) from 31 to 36

Figure C.1: Sample data histograms of RNO of Zone 1 (cont.)

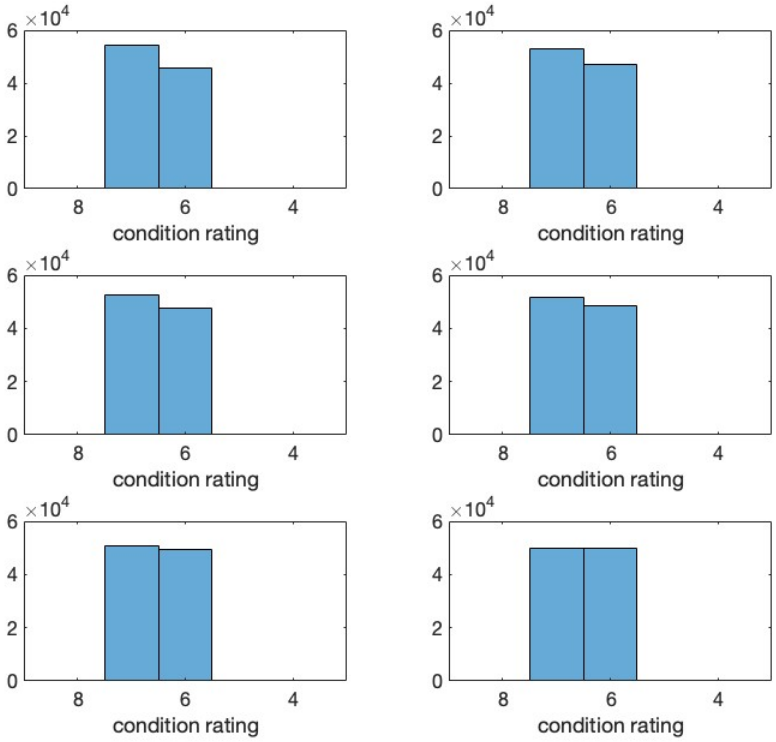


(f) from 37 to 42

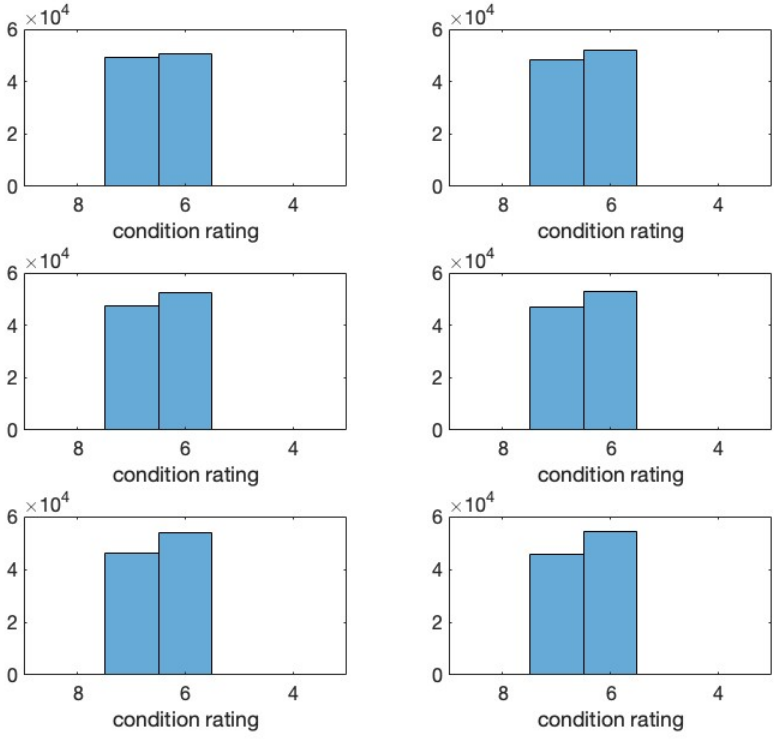


(g) from 43 to 48

Figure C.1: Sample data histograms of RNO of Zone 1 (cont.)

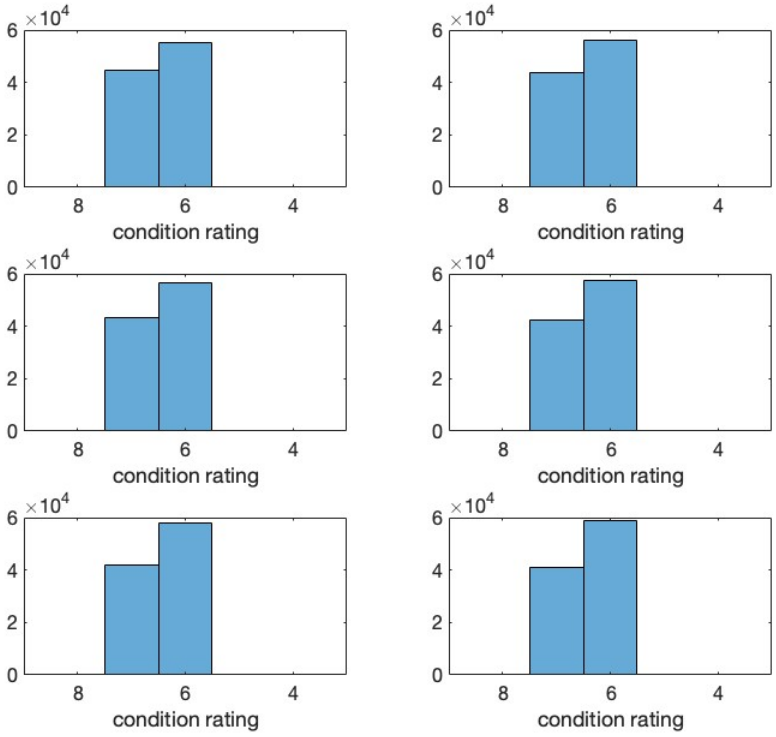


(h) from 49 to 54

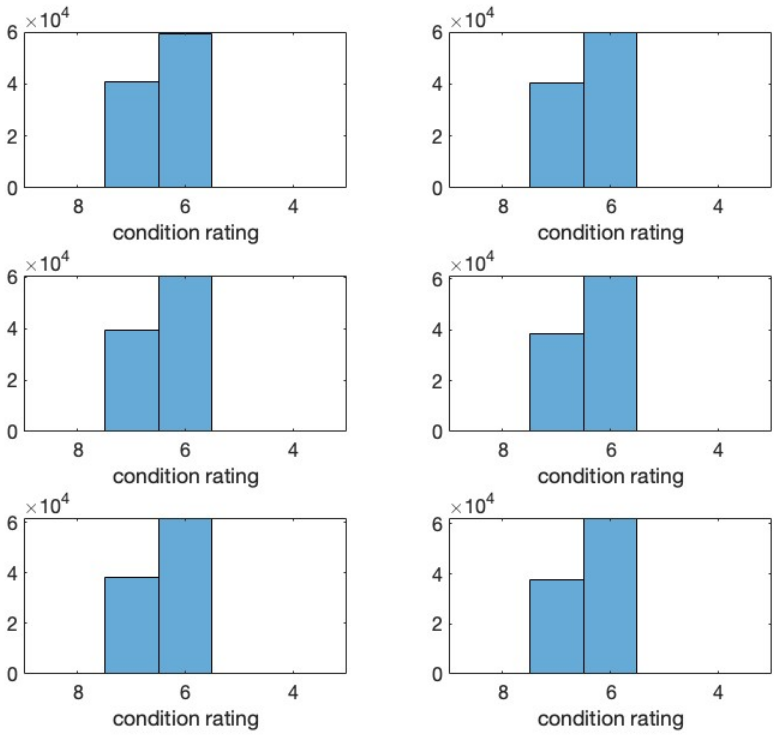


(i) from 55 to 60

Figure C.1: Sample data histograms of RNO of Zone 1 (cont.)

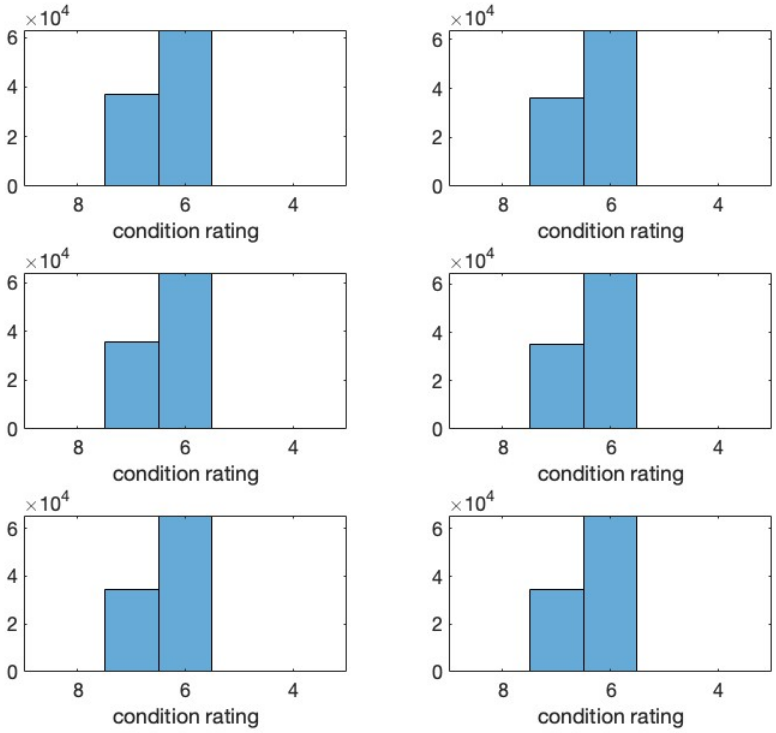


(j) from 61 to 66



(k) from 67 to 72

Figure C.1: Sample data histograms of RNO of Zone 1 (cont.)



(l) from 73 to 78

Figure C.1: Sample data histograms of RNO of Zone 1 (cont.)

References

Administration, F. H. and U. Transportation (2012). *Recording and Coding Guide for the Structure Inventory and Appraisal of the Nation's Bridges*. CreateSpace Independent Publishing Platform.

Agrawal, A. K., A. Kawaguchi, and Z. Chen (2010). Deterioration rates of typical bridge elements in new york. *Journal of Bridge Engineering 15*(4), 419–429.

Baik, H.-S., H. S. Jeong, and D. M. Abraham (2006). Estimating transition probabilities in markov chain-based deterioration models for management of wastewater systems. *Journal of water resources planning and management 132*(1), 15–24.

Barone, G. and D. M. Frangopol (2014). Reliability, risk and lifetime distributions as performance indicators for life-cycle maintenance of deteriorating structures. *Reliability Engineering & System Safety 123*, 21–37.

Berver, E. W., J. O. Jirsa, D. Fowler, H. Wheat, and T. Moon (2001). Effects of wrapping chloride contaminated concrete with fiber reinforced plastics. Technical report.

Bu, G., J. Lee, H. Guan, Y. Loo, and M. Blumenstein (2015). Prediction of long-term bridge performance: Integrated deterioration approach with case studies. *J. Perform. Constr. Facil 29*(3), 04014089.

Butt, A. A., M. Y. Shahin, K. J. Feighan, and S. H. Carpenter (1987). *Pavement performance prediction model using the Markov process*. Number 1123.

Camahan, J., W. Davis, M. Shahin, P. Keane, and M. Wu (1987). Optimal maintenance decisions for pavement management. *Journal of Transportation Engineering 113*(5), 554–572.

Cavalline, T. L., M. J. Whelan, B. Q. Tempest, R. Goyal, and J. D. Ramsey (2015). Determination of bridge deterioration models and bridge user costs for the ncdot bridge management system. Technical report.

Frangopol, D. M., M.-J. Kallen, and J. M. v. Noortwijk (2004). Probabilistic models for life-cycle performance of deteriorating structures: review and future directions. *Progress in structural engineering and materials 6*(4), 197–212.

Goyal, R., M. J. Whelan, and T. L. Cavalline (2020). Multivariable proportional hazards-based probabilistic model for bridge deterioration forecasting. *Journal of Infrastructure Systems 26*(2), 04020007.

Hu, N., S. W. Haider, R. Burgueño, et al. (2013). Development and validation of deterioration models for concrete bridge decks-phase 2: mechanics-based degradation models. Technical report, Michigan. Dept. of Transportation.

Jackson, A., K. Rainwater, W. D. Lawson, S. Senadheera, D. Liang, J. Surles, A. Morse, W. Yan, et al. (2017). Snow and ice control materials for texas roads; volume 1: Literature and best practices review; volume 2: Field trials and laboratory study.

Jiang, Y., M. Saito, and K. C. Sinha (1988). *Bridge performance prediction model using the Markov chain*. Number 1180.

Kouemou, G. L. and D. P. Dymarski (2011). History and theoretical basics of hidden markov models. *Hidden Markov models, theory and applications 1*.

Lawson, W. D., W. A. Jackson, K. W. Rainwater, S. Senadheera, and D. Liang (2017). Guidelines on selection and use of snow and ice control materials. Technical report, Texas Tech University. Center for Multidisciplinary Research in Transportation.

Lee, J., H. Guan, Y.-C. Loo, and M. Blumenstein (2014). Development of a long-term bridge element performance model using elman neural networks. *Journal of infrastructure systems 20*(3), 04014013.

Lee, J., K. Sanmugarasa, M. Blumenstein, and Y.-C. Loo (2008). Improving the reliability of a bridge management system (bms) using an ann-based backward prediction model (bpm). *Automation in Construction 17*(6), 758–772.

Lethanh, N., J. Hackl, and B. T. Adey (2017). Determination of markov transition probabilities to be used in bridge management from mechanistic-empirical models. *J. Bridge Eng 22*(10), 04017063.

Li, Z. and R. Burgueño (2010). Using soft computing to analyze inspection results for bridge evaluation and management. *Journal of Bridge Engineering 15*(4), 430–438.

Madanat, S. and W. H. W. Ibrahim (1995). Poisson regression models of infrastructure transition probabilities. *Journal of Transportation Engineering 121*(3), 267–272.

Madanat, S., R. Mishalani, and W. H. W. Ibrahim (1995). Estimation of infrastructure transition probabilities from condition rating data. *Journal of infrastructure systems 1*(2), 120–125.

Mauch, M. and S. Madanat (2001). Semiparametric hazard rate models of reinforced concrete bridge deck deterioration. *Journal of Infrastructure Systems 7*(2), 49–57.

Micevski, T., G. Kuczera, and P. Coombes (2002). Markov model for storm water pipe deterioration. *Journal of infrastructure systems 8*(2), 49–56.

Morcous, G. and Z. Lounis (2005). Prediction of onset of corrosion in concrete bridge decks using neural networks and case-based reasoning. *Computer-Aided Civil and Infrastructure Engineering 20*(2), 108–117.

Morcous, G. and Z. Lounis (2007). Probabilistic and mechanistic deterioration models for bridge management. *Computing in Civil Engineering*, 364–373.

Nickless, K. and R. A. Atadero (2018). Mechanistic deterioration modeling for bridge design and management. *Journal of Bridge Engineering 23*(5), 04018018.

Pandey, P. and S. Barai (1995). Multilayer perceptron in damage detection of bridge structures. *Computers & structures 54*(4), 597–608.

Ranjith, S., S. Setunge, R. Gravina, and S. Venkatesan (2013). Deterioration prediction of timber bridge elements using the markov chain. *Journal of Performance of Constructed Facilities 27*(3), 319–325.

Stewart, M. G. and D. V. Rosowsky (1998). Time-dependent reliability of deteriorating reinforced concrete bridge decks. *Structural safety 20*(1), 91–109.

Tran, D., B. C. Perera, and A. Ng (2009). Comparison of structural deterioration models for stormwater drainage pipes. *Computer-Aided Civil and Infrastructure Engineering 24*(2), 145–156.

Tran, H. D. (2007). *Investigation of deterioration models for stormwater pipe systems*. Ph. D. thesis, Victoria University.

Transportation, T. D. (2012). *Snow and Ice Control Operations Manual*. Texas Department of Transportation.

Tuutti, K. (1982). *Corrosion of steel in concrete*. Cement-och betonginst.

Wellalage, N., T. Zhang, and R. Dwight (2015). Calibrating markov chain-based deterioration models for predicting future conditions of railway bridge elements.

Witcher, T. (2017). From disaster to prevention: The silver bridge. *Civil Engineering Magazine Archive 87*(11), 44–47.

Xi, Y., Y. Jing, R. Railsback, et al. (2018). Surface chloride levels in colorado structural concrete. Technical report, Colorado. Dept. of Transportation. Research Branch.

Yanev, B. and X. Chen (1993). Life-cycle performance of new york city bridges. *Transportation Research Record 1389*, 17.

Yosri, A., Y. Elleathy, S. Hassini, and W. El-Dakhakhni (2021). Genetic algorithm-markovian model for predictive bridge asset management. *Journal of Bridge Engineering 26*(8), 04021052.

Vita

Jin Collins was born in Seoul, Korea, and immigrated to America in 2001. She attended the Long Beach community college in California and transferred to the California State University Long Beach. She graduated with the university's outstanding student of the year with a Bachelor of Science Degree in Civil Engineering in the spring of 2007. After graduation, she worked for the Public Work, the County of Los Angeles in California as a civil engineering assistant until 2011. She attended the University of Texas El Paso and studied Physics. She graduated with the Academic and Research Excellence Undergraduate Student of the year from the university in the spring of 2017. In the fall, she attended the Master's program in Civil Engineering. She completed her Master's research, focused on bridge deterioration modeling, in the summer of 2019. She then began a doctorate to continue her research in the fall, funded as a teaching assistant. She defended her research in the spring of 2023, and after graduation, she plans to work for a government agency.