12-15-2020

Condition Rating Prediction Using an Interactive Deterioration Model Development Package

Minwoo Chang

Mitchell S. Maguire

applied sciences

Article

Condition Rating Prediction Using an Interactive Deterioration Model Development Package

Minwoo Chang [1,*] and Marc Maguire [2]

[1] Department of Civil and Environmental Engineering, Myongji University, Yongin-si, Gyeonggi-do 17058, Korea
[2] Durham School of Architectural Engineering and Construction, University of Nebraska-Lincoln, Lincoln, NE 68588, USA; marc.maguire@unl.edu
* Correspondence: cmw321@mju.ac.kr

Received: 12 November 2020; Accepted: 13 December 2020; Published: 15 December 2020

Abstract: This paper presents an advanced method to determine explanatory variables required for developing deterioration models without the interference of human bias. Although a stationary set of explanatory variables is ideal for long-term monitoring and asset management, the penalty regression results vary annually due to the innate bias in the inspection data. In this study, weighting factors were introduced to consider the inspection data collected for several years, and the most stationary set was identified. To manage the substantial amount of inspection data effectively, we proposed a software package referred to as the Deterioration Model Development Package (DMDP). The objective of the DMDP is to provide a convenient platform for users to process and investigate bridge inspection data. Using the standardized data interpretation, the user can update an initial dataset for the deterioration model development when new inspection data are archived. The deterministic method and several stochastic approaches were included for the development of the deterioration models. The performances of the investigated methods were evaluated by estimating the error between the predicted and inspected condition ratings; further, this error was used for estimating the most effective number of explanatory variables for a given number of bridges.

Keywords: bridge condition rating; deterioration model; bridge monitoring system; explanatory variables; structural health monitoring

1. Introduction

The increasing demand for structural health monitoring (SHM) has demonstrated the need for effective civil infrastructure management systems [1–3]. The objective of such management systems is to inform an appropriate maintenance actions depending on the condition state of the infrastructure in question. Deterioration models have been widely adopted by infrastructure management agencies to evaluate the changes in the condition ratings during the service life of civil infrastructures, such as bridges, sewers, and asphalt pavements [4–6]. Inspection data monitored from all assets over several decades can be analyzed to extract the features of the deterioration process. Among the various types of civil infrastructure, deterioration models mainly focus on bridges, owing to the importance of public safety and the substantial maintenance expenses involved [3,7].

The deterioration of civil infrastructure is defined as a continuous decline in the condition ratings from the initial operational condition due to gradual changes in the performance of the structural elements [8]. The health status of an individual bridge can be expressed by the condition rating, which provides a comprehensive measure on the basis of field inspections and ranges from “0 (failed)” to “9 (excellent)” [9]. Using the Pontis bridge management system, the element levels of inspection data are converted into the National Bridge Inventory (NBI) condition ratings for the decks, superstructures, substructures, channels and their protection, and culverts [10]. Numerous studies have focused on the decks, superstructures, and substructures in the development of deterioration models [4,11–14].

Many studies have been conducted classify of bridge assets and demonstrated the benefit of using a specific deterioration model type for effective bridge monitoring systems (BMS); however, these tasks require significant computation time to manage bridge inspection data and should interpret new inspection data for model update [15]. For further investigation, a comprehensive toolsuite for the development of deterioration model would aid this process for engineers, managers, and researchers.

This study presents a method for determining the explanatory variables of condition ratings from inspection data. Although several variables are commonly identified as explanatory variables, the final variable sets may vary from year to year. In the proposed method, nearly stationary sets of explanatory variables, regardless of the inspection year, are identified for bridge assets. To provide a convenient platform for developing deterioration models, we propose a software package referred to as the Deterioration Model Development Package (DMDP). The essence of DMDP is to allow the user to update the inspection data if the previously archived data exist and to provide multiple options for the developing deterioration models. A deterministic model and several stochastic deterioration models were encoded for the deck, superstructure, substructure, and culvert elements. Wyoming bridges were then investigated to validate the proposed method and DMDP by predicting the condition ratings. The ability to conduct performance comparisons in the DMDP allows the optimal number of explanatory variables that minimizes the normalized prediction error for each condition state to be determined.

2. Background

The technical literature on the bridge deterioration model and BMS is reviewed in this section. The deterioration models provide a mathematical evidence for decision making process including life cycle cost (LCC) estimation [16,17] and establishment for maintenance strategies [18–20]. The deterioration models are classified as deterministic or stochastic on the basis of the contribution of uncertainty to the transition between condition ratings. Deterministic models are developed by estimating the average condition ratings of the bridge inventory according to age (or year built) [21]. Although deterministic deterioration models are convenient for analyzing the inspection data, they do not reflect the transitions between condition ratings or other historic information.

Stochastic deterioration models are designed to overcome these issues and have been adopted in many U.S. states [4,19,21]. The Markov chain process has been implemented extensively to estimate the transition probabilities between condition ratings. The cumulative damage is assessed as a decrease in condition states over transition periods [22]. The elements of transition probability matrix are estimated using the condition history of various civil infrastructures, i.e., asphalt pavement [6], sewers [23], and bridges [4,24–26]. Although Markovian models are difficult in terms of implementing current condition and expressing condition states for all elements in a group of bridges with a single number, Markov chain process is still effective in considering historical inspection data, being widely used to reflect uncertainties in the deterioration process [27,28]. Alternatively, the sojourn time in each condition rating is modeled statistically. In reliability, the hazard rate function is modeled with Weibull distribution and the sojourn time is then estimated on the basis of the exponential of cumulated hazard function [29,30]. The sojourn time is used to estimate elements of transition probability matrix for the Markov chain process, being named as the semi-Markov method [31]. Bayesian techniques have been implemented to quantify the uncertainties from the prior information such as the assumed structural conditions and inspection data, and are used to update the posterior models [32–34]. Recently, artificial intelligence techniques including artificial neural network, data fusion, and machine learning has been used to develop deterioration models and to establish maintenance strategies [35–38].

NBI inspection data is often used for the development of deterioration models. Owing to the increase in inspection data, the more sophisticated deterioration model can be developed by separating bridges into multiple groups. Butt et al. (1987) proposed a zoning technique to estimate transition probability matrices for every 6-year period [39]. Explanatory variables are used to group bridges and the deterioration models are developed for each [25,40–43]. Instead of leaning on the expert judgement for the selection of explanatory variables, Chang et al. (2017) proposed a framework based on a penalty regression quantifying the significance target variables [20]. Zhang and Marsh (2020) identified significant features among NBI data using random forest technique and developed Bayesian network models for deterioration prediction for each group [35].

Several studies focused on the performance comparison among existing and newly proposed methods. Agrawal et al. (2010) compared the deterioration models developed by Markov chain and Weibull distribution approaches and concluded that the Weibull shows better performance due to the duration independence assumption in the inspected data [4]. Wellalage et al. (2015) developed the Markov chain Monte Carlo simulation technique, with the superior performance being verified by the prediction model when it was compared with two other existing methods [44]. Chang et al. (2019) compared several Markov chain approaches and demonstrated the effectiveness of using logistic regression to model bridge deterioration [24]. Many studies used the comparison of condition ratings between inspected and predicted from deterioration models for validation [44–46].

3. DMDP

The NBI requires regular inspections to be conducted biannually [47]. The resulting increase in the inspection data pool has enabled deterioration models to be developed by means of computational processes to access and analyze the inspection data from the numerous bridge assets. Maintenance agencies are required to monitor bridge conditions continuously and use such models to aid in the decision-making processes. Therefore, a comprehensive platform for conducting these tasks is necessary [12,48].

The DMDP is designed to conveniently interpret NBI records and develop deterioration models [49]. All bridge deterioration and inspection data used can be found at the Federal Highway Administration (FHWA) site [50]. One significant feature of the DMDP is that it allows the user to recall the deterioration model analysis results and update these when new inspection data are archived. In addition to an efficient protocol for loading inspection data, the DMDP provides multiple methods for the development of deterioration models, and thus the users can compare their performances in a straightforward manner.

Figure 1 presents a flowchart of the DMDP, which is composed of three subtasks: the initiation of the DMDP, selection of explanatory variables, and development of deterioration models. Each subtask is matched to a MATLAB-based window developed using a graphical user interface platform. Once the user decides to initiate the analysis, a second window appears for the selection of explanatory variables. If a previously archived inspection dataset is accessible for a specific bridge element, the user can select the update option. Otherwise, the DMDP prompts the user to load the inspection data from the beginning and conducts an analysis to identify the significant variables. An explanatory variable list is obtained and used to divide the bridge data into multiple groups depending on the classification levels. Thereafter, a deterioration model is developed for each group. The window for the development of the deterioration models displays the deterioration curves for the selected group and the condition rating prediction results depending on the classification tree, if available.

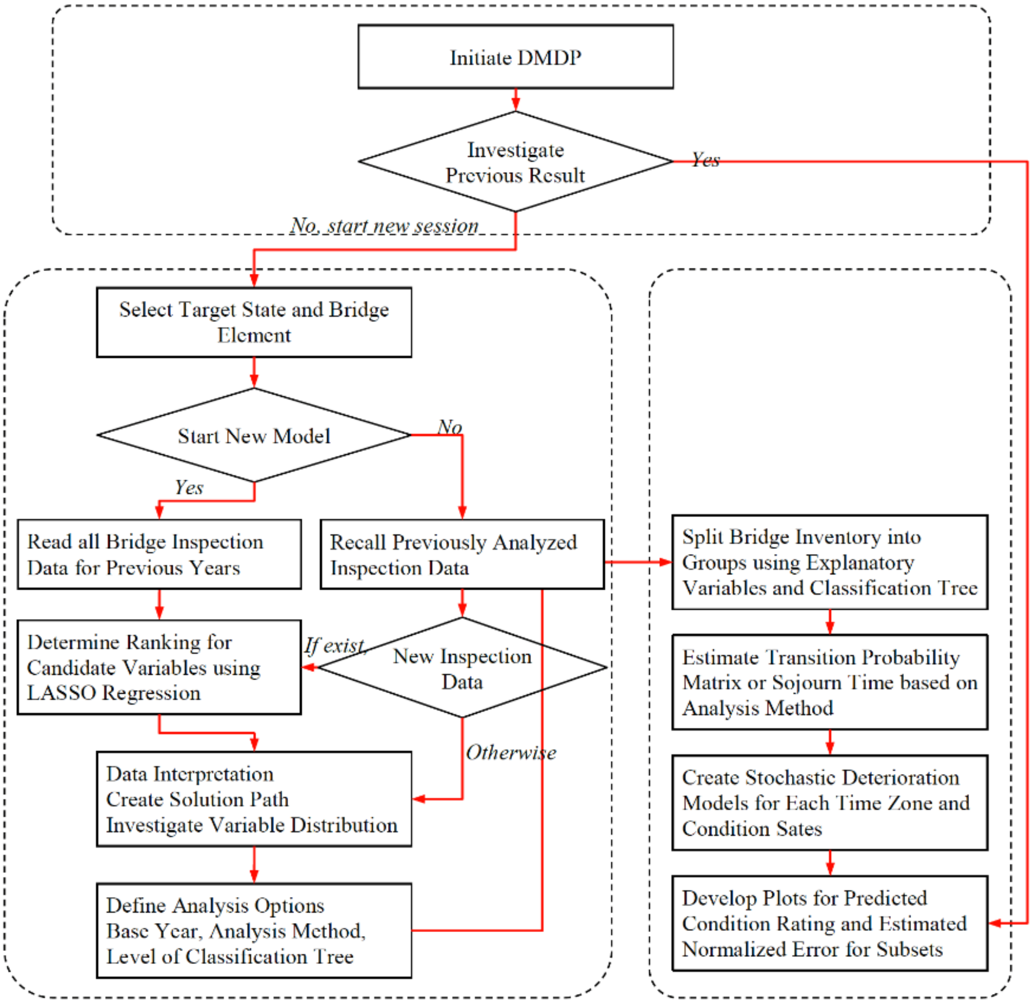


Figure 1. Flowchart for Deterioration Model Development Package (DMDP).

In the following section, NBI regulations, the weighted least absolute selection and shrinkage operator (LASSO), and an algorithm for the comparison of the condition rating prediction error are briefly described.

3.1. NBI Regulation

For effective bridge monitoring, the Federal Highway Administration (FHWA) requires each state to compile a bridge inspection database, known as the NBI database [47]. Over 100 inspection records for each bridge have been archived biannually since 1992 and opened to the public [50]. Several studies have focused on the selection of explanatory variables from these data, which are defined as parameters to which the condition status level is most sensitive from among the candidate variables following the NBI standards [21,31]. The list of candidate variables, presented in Table 1, is conservatively defined from among all inspection records. Since not all states record weather/climate accurately, this information is vanished from the candidates. Each variable includes more than two indices and is discretized manually with then indices if they are almost continuous variables, such as the length of the main span and average daily traffic.

Table 1. List of candidate variables.

Index No.	Candidate Variables
1	Route signing prefix
2	Highway agency district
3	Base highway network
4	Maintenance responsibility
5	Functional classification of inventory route
6	Year built (age)
7	Lanes on the structure
8	Lanes under the structure
9	Average daily traffic
10	Design load
11	Skew
12	Type of service on bridge
13	Type of service under bridge
14	Kind of material and/or design
15	Type of design and/or construction
16	Number of spans in main unit
17	Inventory route, total horizontal clearance
18	Length of maximum span
19	Structure length
20	Bridge roadway width
21	Deck width
22	Deck structure type
23	Type of wearing surface
24	Type of membrane
25	Deck protection
26	Average daily truck traffic
27	Designated national network

3.2. Weighted LASSO

Chang et al. (2017) [21] proposed a framework for the selection of explanatory variables using penalized regression, known as LASSO [51]. By controlling tuning parameter in LASSO, the particular contribution for candidate variables is determined. Under the given condition ratings, which possibly contain innate human error, the explanatory variables can be determined without bias. To establish an efficient bridge management system with a nearly stationary set of explanatory variables, we adopted the weighted LASSO in the DMDP.

Suppose that the condition ratings and inspection data collected from m bridges in the ith year are $\mathbf{Y}$ and $\mathbf{X}$, respectively, for which the linear regression model can be defined as follows:

$$\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon} \tag{1}$$

In Equation (1), $\mathbf{X} \in \Re^{m \times n}$ is a matrix form of the inspection data for n explanatory variables, and $\boldsymbol{\beta}$ is the regression coefficient vector that minimizes the sum of the squared error, $\boldsymbol{\varepsilon}$. The traditional LASSO estimates the regression coefficient $\widehat{\boldsymbol{\beta}}$ using the following formula:

$$\widehat{\boldsymbol{\beta}} = \arg\min \left\{ \|\mathbf{Y} - \mathbf{X}\boldsymbol{\beta}\|_2^2 + \lambda \sum_{j=1}^{n} \left|\beta_j\right| \right\} \tag{2}$$

In Equation (2), λ is the tuning parameter, and the penalty is defined as the absolute sum of β. The amount of shrinkage is generally defined as the dot product of these two terms. The priority of the candidate variables is determined by controlling this shrinkage parameter. Although LASSO successfully identifies the explanatory variables using rigorous evidence, the result varies depending on the inspection year.

Equation (2) is modified to use the inspection data from multiple years, such that $\mathbf{X} = [w_1\mathbf{x}_1^{\mathrm{T}} \quad w_2\mathbf{x}_2^{\mathrm{T}} \quad \cdots \quad w_s\mathbf{x}_s^{\mathrm{T}}]^{\mathrm{T}}$ and $\mathbf{Y} = [w_1\mathbf{y}_1^{\mathrm{T}} \quad w_2\mathbf{y}_2^{\mathrm{T}} \quad \cdots \quad w_s\mathbf{y}_s^{\mathrm{T}}]^{\mathrm{T}}$, where s denotes the total number of previous inspection years. The weighting parameter w_i is introduced to estimate the contribution of the inspections of previous years for LASSO regression and is assumed to be half for each inspection cycle.

The elements of Wyoming bridges were analyzed in this study to investigate the consistency of explanatory variable selection using the weighting factor in LASSO regression. The binary function used to quantify the difference in the ranking system for adjacent years is given by

$$f_{ij} = \begin{cases} 1 \text{ if } v_i = v_{i-1} \\ 0 \text{ if } v_i \neq v_{i-1} \end{cases} \tag{3}$$

In Equation (3), v_i is the ith element of a list of n explanatory variables in the jth year. The normalized consistency, c_j, is defined as follows:

$$c_j = \frac{1}{n}\sum_i f_{ij} \tag{4}$$

Figure 2 presents the normalized consistency for the deck, superstructure, and substructure elements in the Wyoming bridges considered in this study. Regardless of the element type, the normalized consistency generally increases when using inspection data from multiple years. When eight years of inspection data were used, the set of explanatory variables was almost identical to the set obtained when using all inspection data. Accordingly, the DMDP was developed to use eight years of inspection data with weighting factors to determine the explanatory variables.

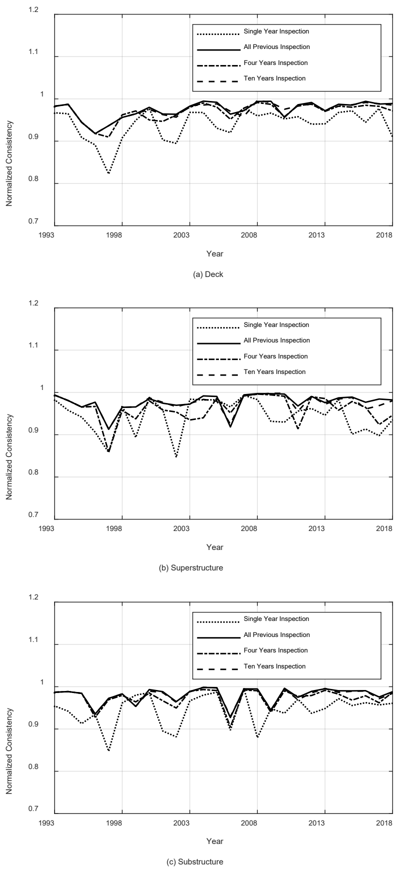


Figure 2. Normalized consistency comparison for (**a**) deck, (**b**) superstructure, and (**c**) substructure elements.

3.3. Prediction of Condition Ratings

The DMDP provides an option to investigate the previous inspection data, as illustrated in Figure 1. Supposing that the year r is used for the selection of the explanatory variables, the inspection data prior to the rth year are employed to develop the deterioration models. For stochastic deterioration models, the remaining inspection data can be used to estimate the prediction error. The normalized average error for a future year, denoted by e_r, is estimated as follows:

$$e_r = \frac{\sum_1^m |1-\rho|}{m} \tag{5}$$

In Equation (5), ρ is the ratio between the predicted and inspected condition ratings. The prediction of condition rating is determined either from the dot product of the transition probability matrix and state vector when the Markov chain is utilized, or from the curve-fitting of sojourn times when the Weibull distribution is used. The normalized average error, ρ, denotes the closeness of the prediction to the inspection results, and its ideal value is equal to unity.

4. Development of Deterioration Models

4.1. Deterministic Deterioration Models

The NBI condition ratings for the bridge elements are determined by converting of the inspection data using the mapping criterion [10]. The deterministic deterioration model uses the current status of the NBI condition ratings by taking their mean according to each age and applying a curve-fitting algorithm to connect them. In the DMDP, a power function is employed to develop the deterioration curve, for which the number of bridges corresponding to each condition rating is used as a weighting factor.

4.2. Stochastic Deterioration Models using Markov Chain

The uncertainty of the condition state transition over time is modeled using a Markov chain process [22,25]. A typical Markov chain problem involves developing a transition state matrix between any two inspection periods, which is defined as

$$\mathbf{P} = \begin{bmatrix} p_{11} & 1-p_{11} & 0 & \cdots & 0 \\ 0 & p_{22} & 1-p_{22} & \cdots & 0 \\ 0 & 0 & p_{33} & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \end{bmatrix} \tag{6}$$

In Equation (6), p_{ii} is the probability that the condition rating of state i remains the same. Accordingly, the probability representing a decrease in the condition state is determined as $1-p_{ii}$. The elements of $\mathbf{P}$ are equal to zero when $i<j$ because the states cannot be improved without intervention. Assume that the initial distribution probability of the bridges in each condition state vector is q_0. Using total probability theory, the transition probability matrix for next inspection state can be obtained by the powers of $\mathbf{P}$, and the probability condition state after the nth inspection periods is defined as

$$q_n = q_0 \mathbf{P}^n \tag{7}$$

For most inspection data, the year built (age) is ranked highly in the set of explanatory variables. The zoning technique, in which the inspection data are grouped for identical age periods, is therefore used to consider the effects of the age and to improve the accuracy of the deterioration models [36]. The transition state matrices are estimated for each group and the initial vectors are updated accordingly. Several methods, including percentage prediction, logistic regression, and optimization-based approaches, have been developed to estimate the elements of the transition state matrix [11,24,25].

4.3. Stochastic Deterioration Models using Weibull Distribution

A stochastic process has been developed to model the uncertainty in the estimation of the sojourn time, which represents the duration of a bridge element remaining at a particular condition rating [4,30,52]. The probability of survival $S_i(t)$ indicates that a bridge at condition rating i remains in the same condition state when the sojourn time T_i exceeds t years, and is modeled using the cumulative distribution function of Weibull distribution, as follows:

$$S_i(t) = \exp\{-\Lambda_i(t)\} \tag{8}$$

In Equation (8), $\Lambda_i(t)$ is defined as the integrated value of the hazard rate, $\lambda_i(t)$, which can be modeled using Weibull distribution as follows:

$$\lambda_i(t) = \frac{\eta_i}{\theta_i}\left(\frac{t}{\theta_i}\right)^{\eta_i - 1} \exp\left\{-\left(\frac{t}{\theta_i}\right)^{\eta_i}\right\},\ t > 0,\ \theta_i > 0,\ \eta_i > 0,\ i = 2, \ldots, 9 \tag{9}$$

In Equation (9), η_i and θ_i are the shape and scale parameters, respectively, which can be estimated using a curve-fitting mechanism. As an example, the distribution of Wyoming bridges regarding age belonging to each condition rating was modeled using Weibull distribution, as shown in Figure 3. The higher condition rating showed that the shape of the Weibull function generally skewed right more, and vice-versa. The sojourn time, T_i, is defined as a mean of the estimated Weibull distribution and can be calculated by

$$\mathrm{E}(T_i) = \eta_i \Gamma\left(1 + \frac{1}{\theta_i}\right) \tag{10}$$

In Equation (10), $\Gamma(.)$ denotes the gamma function. For each subset of the bridge inventory, a maximum of nine sojourn times can be estimated for all condition states, except the "0 (failure)" condition, indicating the average duration for which the bridges remain in a particular condition state. Third-order polynomial functions are implemented in the DMDP to develop deterioration models from these discretized data.

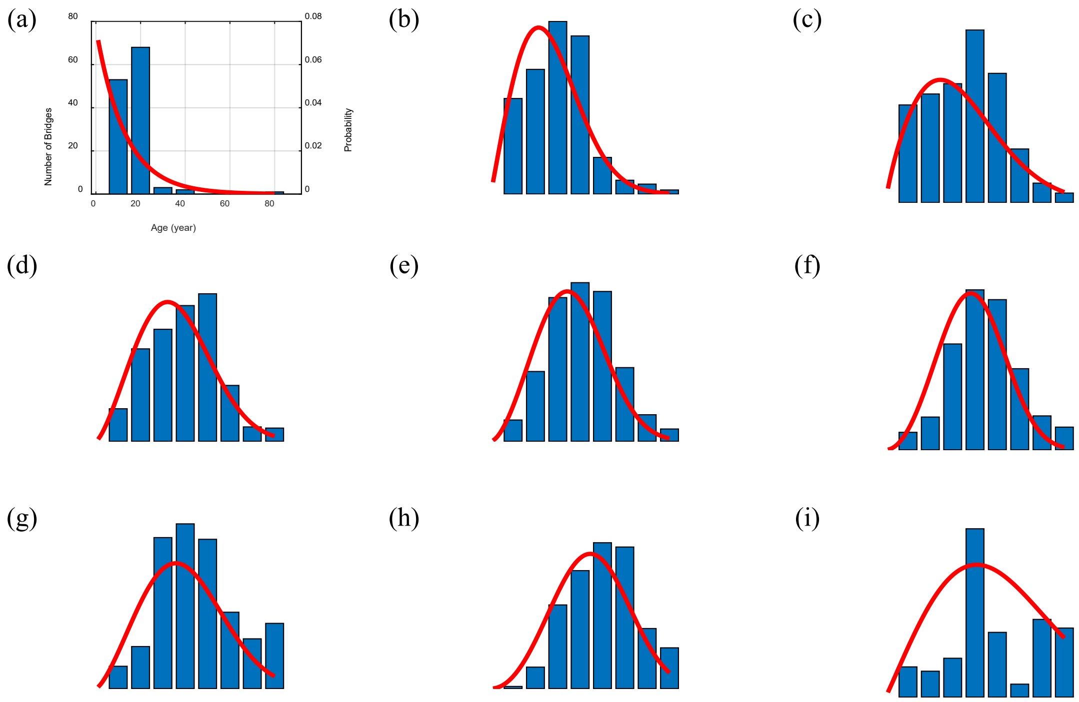


Figure 3. Modeling of the bridge age distribution using a Weibull function for condition ratings of (**a**) "9 (excellent)", (**b**) "8 (very good)", (**c**) "7 (good)", (**d**) "6 (satisfactory)", (**e**) "5 (fair)", (**f**) "4 (poor)", (**g**) "3 (serious)", (**h**) "2 (critical)", and (**i**) "1 (imminent failure)".

5. Effect of using Weighted LASSO

Wyoming bridges were investigated to verify the DMDP and to provide performance comparison of the deterioration models. Regular bridge inspections have been conducted in Wyoming since 1992. More than 3100 bridges are currently operating in Wyoming, and 257 bridges (8.21%) were classified as deficient as of the end of 2018, which is greater than the national average (7.64%), and thus will require substantial repair and rehabilitation efforts to improve public safety [53]. To establish an effective BMS, the Wyoming Department of Transportation has focused on the development of deterioration models [12]. As an extension of the previous study into deterioration mode development, this work validates the use of weighted LASSO for the selection of explanatory variables and compares the performance of several algorithms including deterministic and stochastic deterioration models.

5.1. Selection of Explanatory Variables

To investigate the efficacy of the weighted LASSO, we identified the lists of explanatory variables for the deck, superstructure, and substructures using the inspection data from a single year (existing method) and the previous eight years (proposed method), as indicated in Tables 2–4, which show the index numbers of the candidate variables defined in Table 1. In general, more consistent results were observed when the proposed method was applied to the selection of explanatory variables. For the deck elements identified using the proposed method, the type of wearing surface, which had an index number of 23, was generally considered to be the most significant, followed by the structure length, which had an index number of 19, whereas this ranking was inconclusive when only the single year inspection data were used. The inspection records for a single year can possibly be affected by specific conditions including human error, natural disasters, and maintenance actions across the entire state whether they are recorded. The expansion of inspection data to include previous years helps to diminish such biased effects. Similar results were observed for the other elements; therefore, the proposed method identifies the top-ranked explanatory variables with superior consistency compared to the existing method.

These explanatory variable sets determined by LASSO are used to divide the bridge data into multiple groups using a classification tree and to develop deterioration models, the performances of which were evaluated by the normalized prediction error defined in Equation (5). The number of classification tree levels was set to three and logistic regression was applied to estimate the transition probability matrices. For example, Figures 4 and 5 illustrate the specific classification trees for the deck element based on the inspection data from a single year (2013) and from multiple years (2006 to 2013), respectively. The minimum number of bridge assets was set to 50 for each group.

Table 2. List of explanatory variables for deck since 2009 using least absolute selection and shrinkage operator (LASSO) with and without weights.

Year	Single-Year Inspection					Eight-Year Inspection				
2009	23	9	4	10	22	23	10	4	5	19
2010	23	22	19	10	9	23	10	4	19	9
2011	23	22	2	19	4	23	4	19	10	22
2012	23	19	2	4	22	23	19	2	4	22
2013	19	23	4	9	2	23	19	2	4	22
2014	23	19	4	9	2	23	19	4	2	9
2015	23	4	19	5	9	23	19	4	2	9
2016	23	4	19	9	10	23	19	4	9	2
2017	23	19	4	10	9	23	19	4	10	9
2018	23	10	19	4	26	23	19	4	10	2

Table 3. List of explanatory variables for superstructure since 2009 using LASSO with and without weights.

Year	Single-Year Inspection					Eight-Year Inspection				
2009	22	10	21	18	5	22	10	21	17	4
2010	22	20	10	21	5	22	10	21	17	4
2011	22	20	10	21	18	22	10	20	21	17
2012	22	20	10	4	18	22	10	20	21	18
2013	22	20	18	21	5	22	20	10	21	5
2014	22	20	15	5	18	22	20	10	5	18
2015	22	10	5	4	19	22	20	10	5	18
2016	22	23	4	20	14	22	20	4	5	19
2017	22	10	4	19	8	22	20	4	10	19
2018	10	22	4	19	8	22	10	4	19	20

Table 4. List of explanatory variables for substructure since 2009 using LASSO with and without weights.

Year	Single-Year Inspection					Eight-Year Inspection				
2009	10	20	23	22	21	10	20	21	22	23
2010	20	10	23	22	1	10	20	22	23	21
2011	20	23	10	22	1	10	20	23	22	1
2012	23	10	20	5	4	10	20	23	22	1
2013	23	20	10	5	22	10	23	22	20	1
2014	23	10	20	5	26	23	20	10	22	5
2015	23	10	20	5	22	23	10	20	5	22
2016	23	10	20	22	5	23	10	20	5	22
2017	10	23	22	20	5	23	10	20	22	5
2018	10	22	26	20	5	10	23	22	20	5

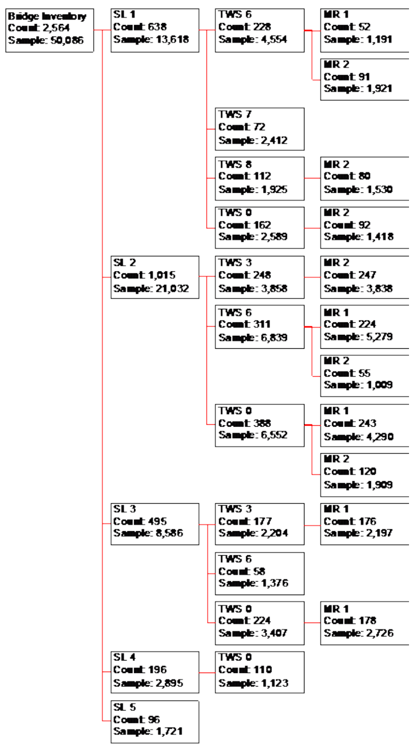


Figure 4. Classification tree for Wyoming bridges using explanatory variables from a single year (2013) of inspection data for deck elements.

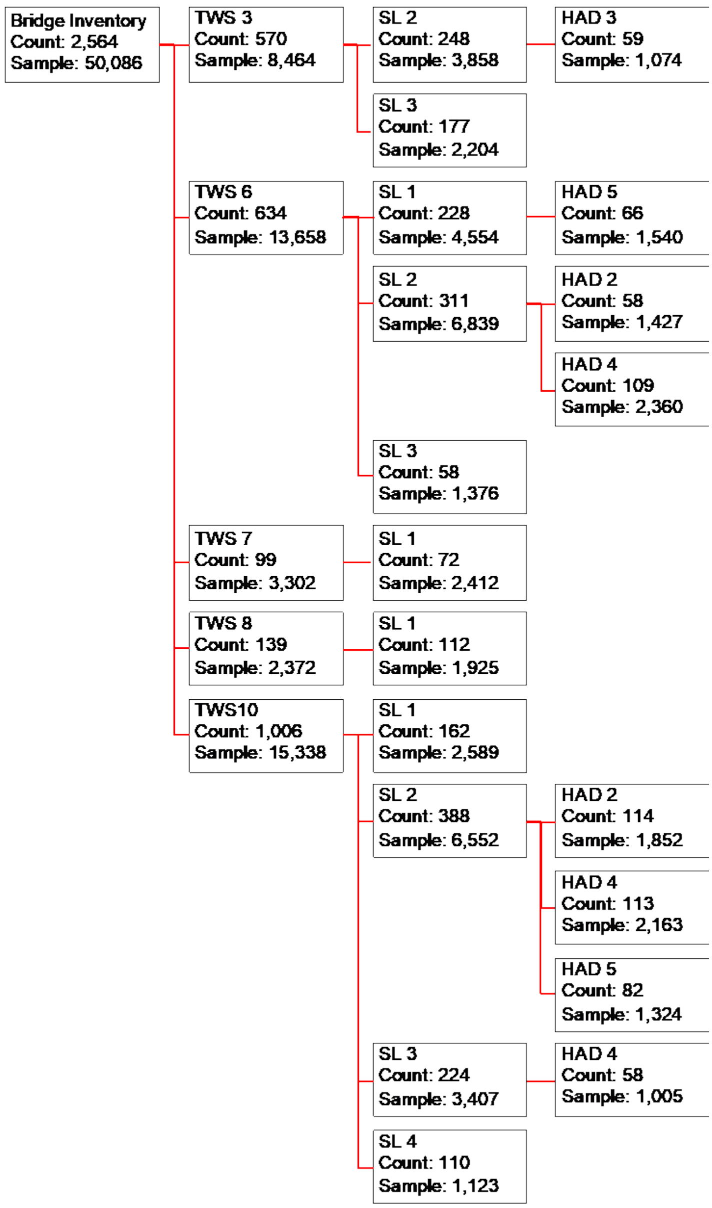


Figure 5. Classification tree for Wyoming bridges using explanatory variables from eight years (2006-2013) of inspection data for deck elements.

Logistic regression was utilized to develop the deterioration models. Then, the normalized prediction errors for the 2014 to 2018 inspection data were compared, as shown in Figure 6, in which it can be observed that the normalized error generally increased with increasing prediction year. Although the normalized error decreased by an average of only 1.6% over all prediction periods, this was still a remarkable improvement considering the large number of bridges and inspection data points. Using the inspection data from multiple years reduced the chance of selecting explanatory variables that would only exhibit high correlation with a particular year. This provided additional benefits to bridge owners for more effectively managing their bridge assets.

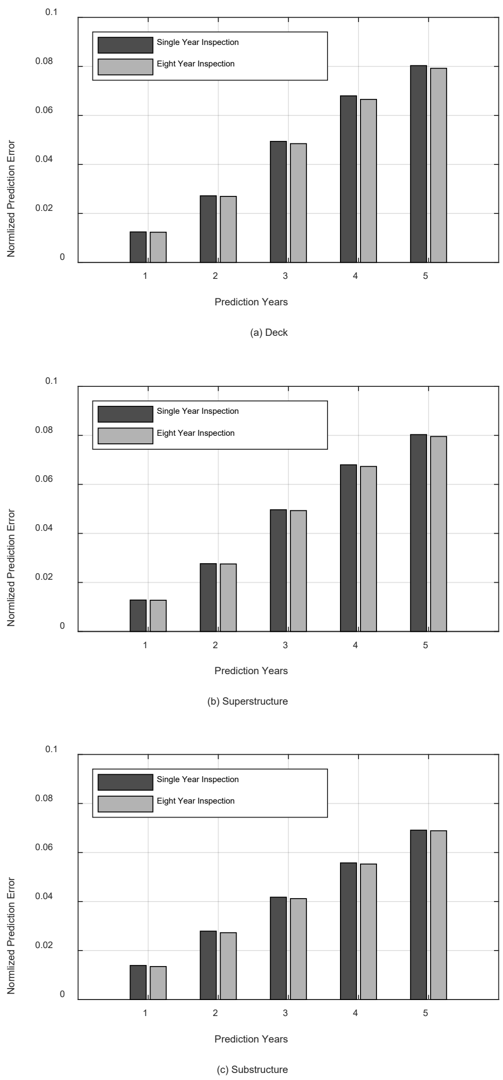


Figure 6. Normalized prediction error comparison for (**a**) deck, (**b**) superstructure, and (**c**) substructure elements when explanatory variables were determined by using single- and eight-year inspection data.

5.2. Effective Number of Explanatory Variables

The number of bridges varies according to the states, affecting the level of classification scheme used to improve the prediction accuracy. Although LASSO generally suggests a large number of explanatory variables to optimize the penalty regression model, this is unnecessary and impossible since the number of bridges and quantity of inspection data are often insufficient to effectively group. Indeed, a high level of classification does not guarantee an improvement in the prediction result and occasionally shows negative effects. Figure 7 illustrates the decrease in normalized prediction error for the deck elements from randomly selected states. The classification scheme was applied until the bridges could be grouped using the required number of bridges and inspection data. For example, the sixth explanatory variable for Alaska could not create a bridge group and thus the classification ended at the fifth level. In many cases, the minimum error was observed when the number of explanatory variables was less than the maximum, and thus excessive classification should accordingly be avoided.

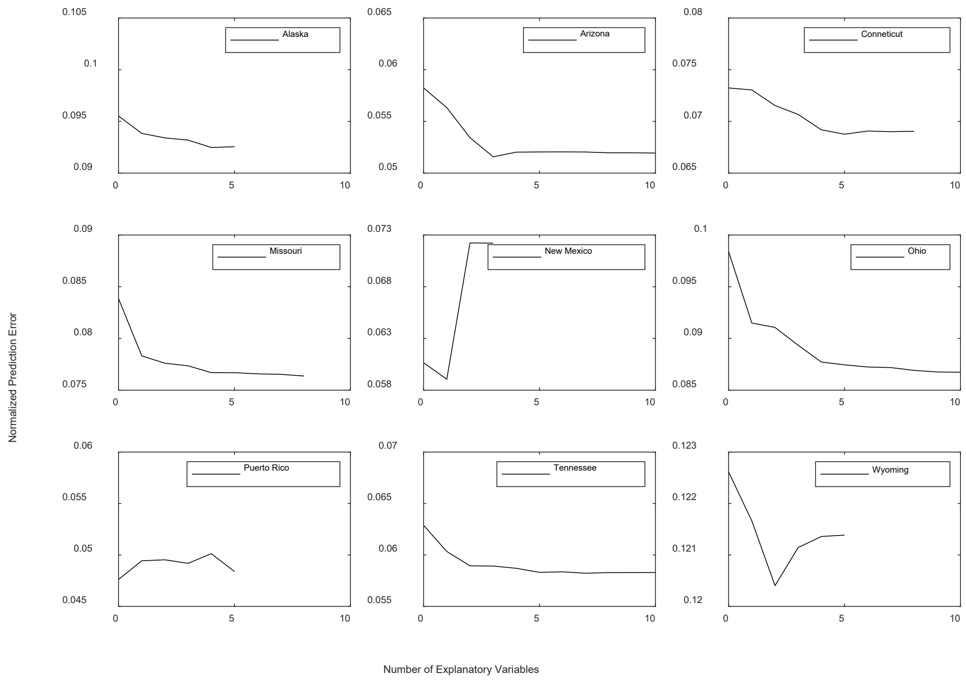


Figure 7. Normalized prediction error depending on the number of explanatory variables for the deck elements in randomly selected states.

To suggest a reasonable classification tree level, we compared the prediction error for five years using inspection data from various states. The optimal number of explanatory variables was determined when the improvement in the normalized prediction error was less than 1E-4 with further increase in the level of classification. Figure 8 shows the optimal number of explanatory variables depending on the number of bridges using all inspection data from all states. The trendline was developed using a linear function and is shown in blue, and the 95% confidence interval is plotted in red. In general, the optimal number of explanatory variables increased as the number of considered bridges increased. Regardless of the number of bridges, four explanatory variables should be considered as a minimum, and no state would require more than eight explanatory variables.

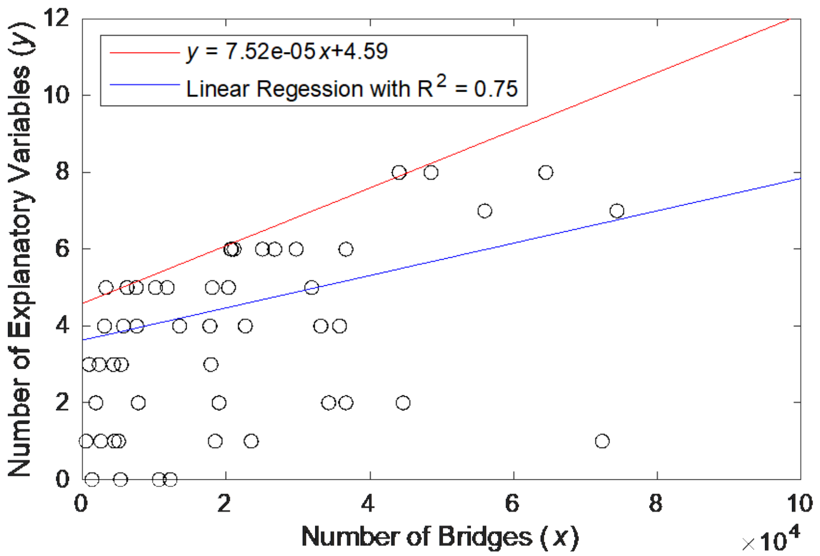


Figure 8. Number of optimal explanatory variables for all U.S. states to minimize the normalized prediction error.

6. Comparison of Deterioration Models for Wyoming Bridges

6.1. Deterministic Deterioration Models

Figure 9 presents all of the deterministic deterioration models for the available sets when three classification tree levels were used to investigate Wyoming bridges. The bridges that did not belong to these subsets were considered to follow the previous classification level. This rule was applied to the stochastic deterioration models in the same manner. The deterioration was modeled using a power function of the mean bridge age for each condition rating, which are indicated by the bullet symbols (○) in the figure.

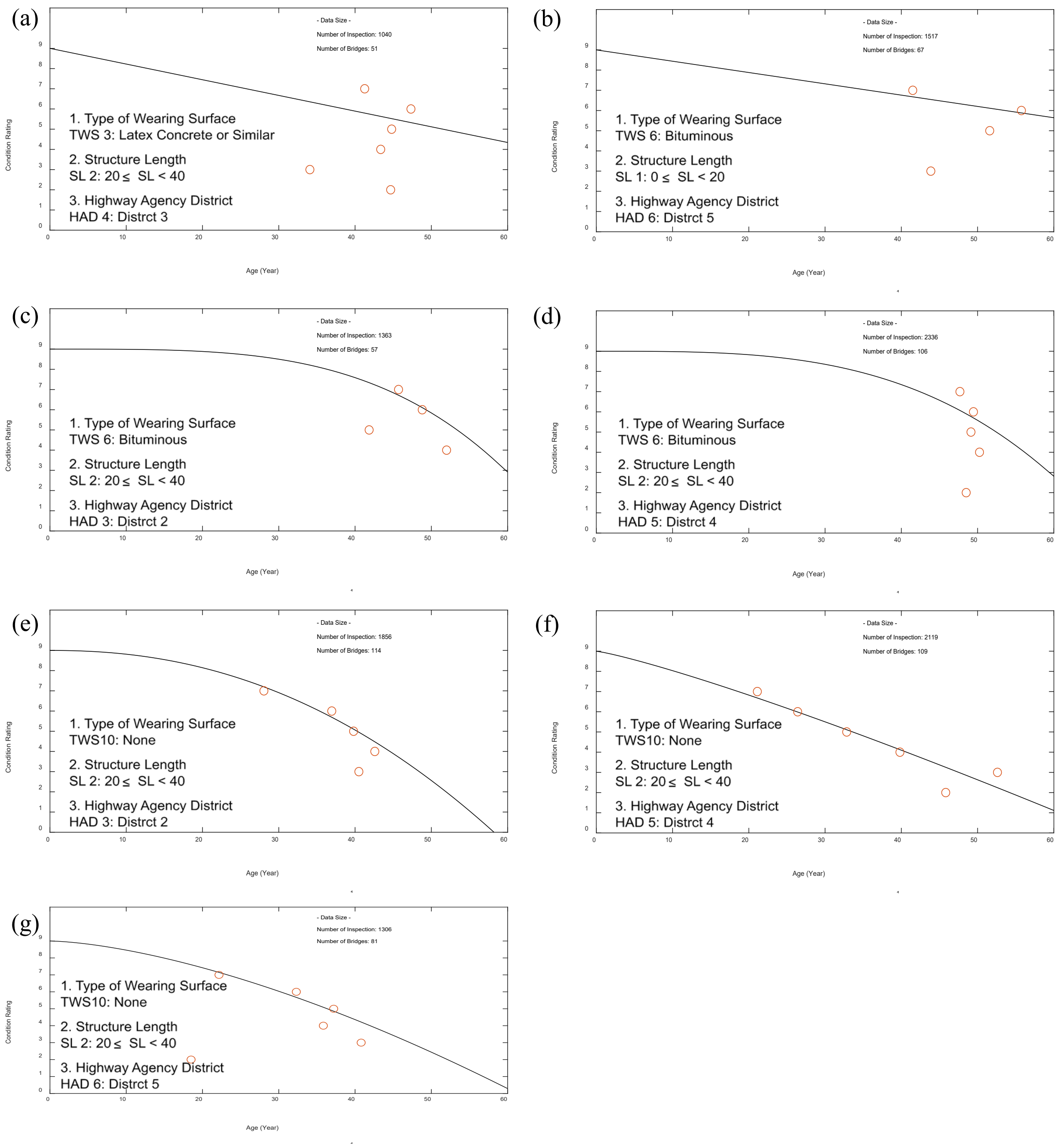


Figure 9. Deterministic deterioration models corresponding to the third level of the classification tree, for which the explanatory variables were identified using 2006-2013 National Bridge Inventory (NBI) data; (**a**) subset 1, (**b**) subset 2, (**c**) subset 3, (**d**) subset 4, (**e**) subset 5, (**f**) subset 6, and (**g**) subset 7.

Overall, it can be observed that the bridges in Figure 9a,d are relatively old on the basis of the mean age distribution. The bridge assets illustrated in Figure 9b,c included only four condition ratings, therefore, the deterioration curves were easily distorted. None of the figures include highquality bridges corresponding to the condition ratings of "8 (very good)" and "9 (excellent)". Thus, it is difficult to develop a representative model using a single curve and to predict future condition ratings. However, the deterministic models remain effective for identifying the distribution of currently operating bridges depending on their condition ratings. On the basis of the age information, the user can simply identify the types of bridges built in a specific era. The bullets in Figure 9f,g corresponding to condition ratings greater than "7 (good)" are plotted up to an age of nearly 20, which indicates that the bridges in these subsets have been constructed relatively recently.

6.2. Stochastic Deterioration Models

Four stochastic deterioration algorithms were coded in the DMDP: three were based on the Markov chain process (percentage prediction, logistic regression, and optimization-based) and one used the Weibull distribution to estimate the sojourn time. The deterioration models for the deck elements based on all of the inspection data are plotted in Figure 10. The deterioration curves for the percentage prediction and logistic regression algorithms were found to be similar. The optimization-based algorithm exhibited a rapid decrease during the first years, following which the deterioration process was almost flat, regardless of the initial condition ratings. The proportion of condition ratings between "4 (poor)" and "7 (good)" from the historical inspection data resulted in a small decrease in the entire age range for the optimization-based algorithm. The sojourn times for the condition ratings, indicated by the bullets in the deterioration models, were estimated using the Weibull distribution, and polynomial curve-fitting was applied to plot the deterioration model illustrated in Figure 10d. The deterioration curves starting from "7 (good)" and "5 (fair)" were shifted compared to that starting from "9 (excellent)". The curve-fitting algorithm thus significantly affects the deterioration shape. In fact, a rapid decrease is generally observed in the other methods, but differences exist between the deterioration curve and bullets.

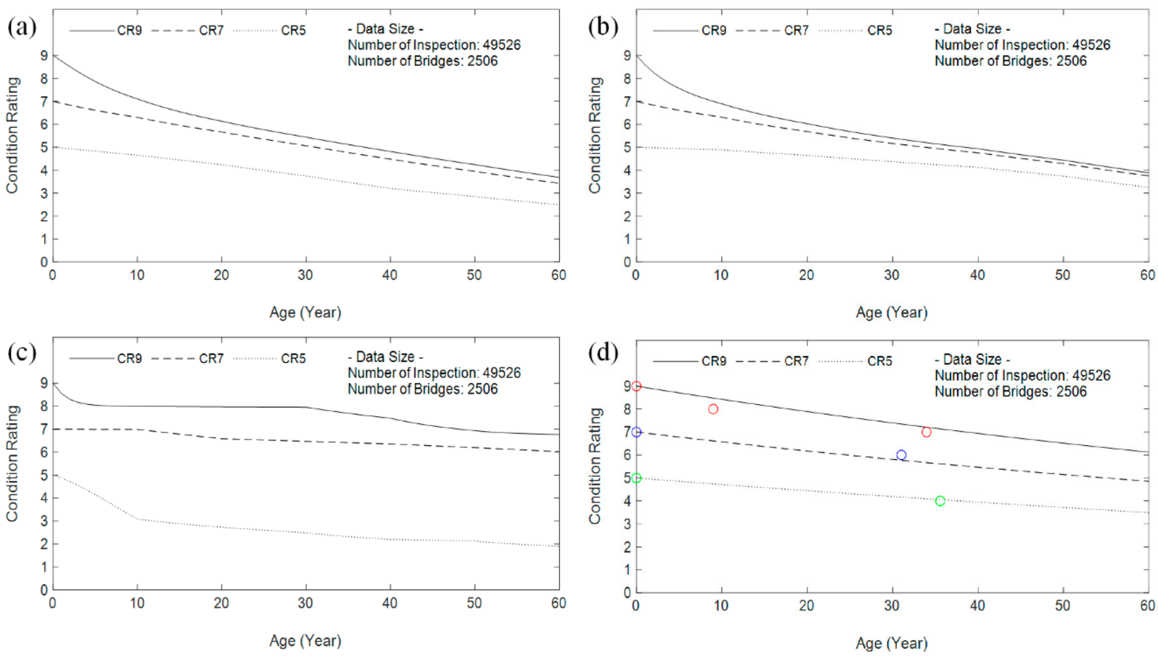


Figure 10. Deck deterioration models for all Wyoming bridges using (**a**) percentage prediction, (**b**) logistic regression, (**c**) optimization algorithm, and (**d**) Weibull distribution.

Further analysis was carried out to investigate the effect of the classification tree level applied in the stochastic methods. Figure 11 presents an example of the superstructure deterioration models for a subset associated with the three explanatory variables "DST 1" (cast-in-place concrete), "BRW 4" (range between 4 and 6 m), and "DL 5" (MS 18). The quantity of inspection data and bridges were slightly higher than required for these variables. Similar to the deterioration models for all of the inspection data, the models for the percentage prediction and logistic regression algorithms were almost identical. The condition ratings decreased rapidly in early age and became almost flat for the remaining service life. The decrease in the condition ratings under the optimization-based algorithm was mainly concentrated in the first two decades. The number of components corresponding to the high-level of condition ratings used for the optimization process nearly flattened this curve. The deterioration model for the Weibull distribution was similar to those for the percentage prediction and logistic regression algorithms, but differences were evident between the bullets representing the estimated sojourn time and the deterioration curve resulting from the curve-fitting process.

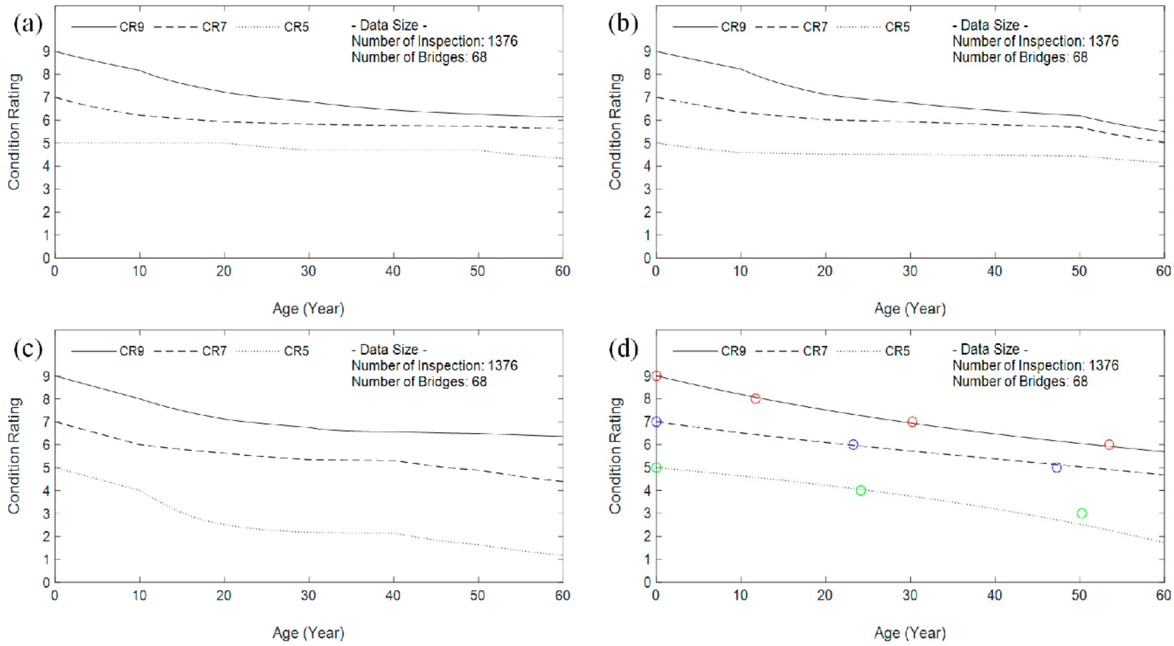


Figure 11. Superstructure deterioration models for a subset of Wyoming bridges associated with "DST 1" (cast-in-place concrete), "BRW 4" (range between 4 and 6 m), and "DL 5" (MS 18), using (**a**) percentage prediction, (**b**) logistic regression, (**c**) optimization, (**d**) Weibull distribution.

6.3. Assessment of Various Modeling Strategies using DMDP

To compare the performances of the stochastic models, we evaluated the normalized prediction errors for the deck, superstructure, and substructure elements. Figure 12 presents the normalized prediction errors for the condition ratings five years later (i.e., using data from 2013 to predict the conditions in 2018) depending on the classification tree level. In general, the prediction error slightly decreased, particularly for the percentage prediction and logistic regression algorithms. As the quantity of inspection data was insufficient for optimization when a high classification tree level was considered, the result was inferior in terms of the prediction error. A similar effect was observed for the Weibull distribution. Although the deterioration shapes for the percentage prediction and logistic regression algorithms were almost identical across numerous subsets, the logistic regression algorithm generally exhibited superior performance to the percentage prediction algorithm. Overall, the classification tree scheme, while common, does not seem to provide dramatically improved predictions.

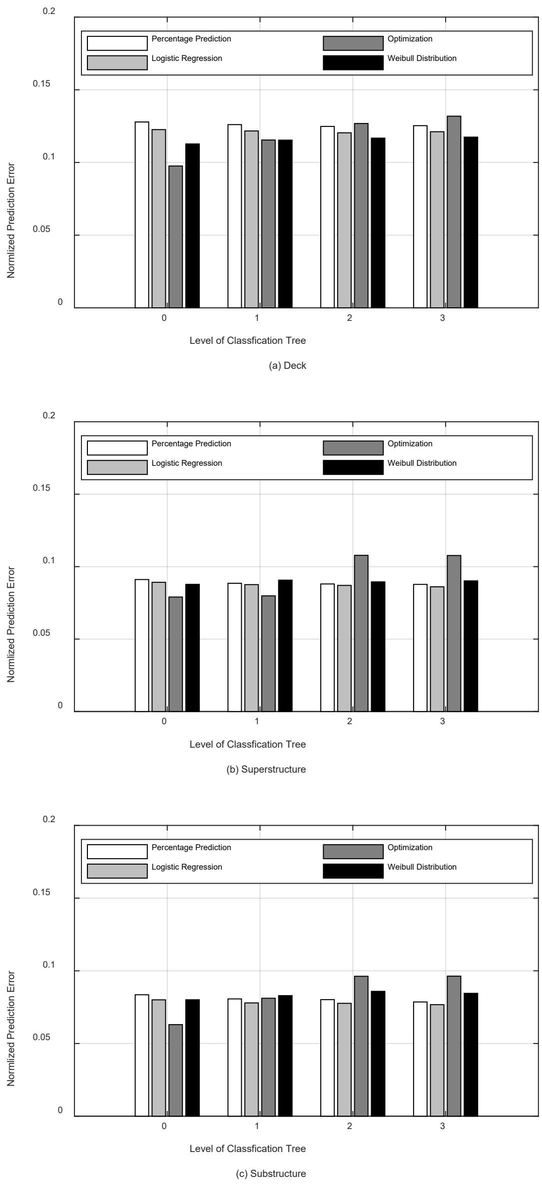


Figure 12. Normalized prediction error comparison of stochastic methods depending on the level of classification tree for (**a**) deck, (**b**) superstructure, and (**c**) substructure.

Figure 13 presents the normalized prediction errors for the investigated stochastic methods depending on the prediction years, for which a three-level of classification tree was used. The Weibull distribution exhibited the best result for the deck, whereas the logistic regression algorithm was superior for the other elements.

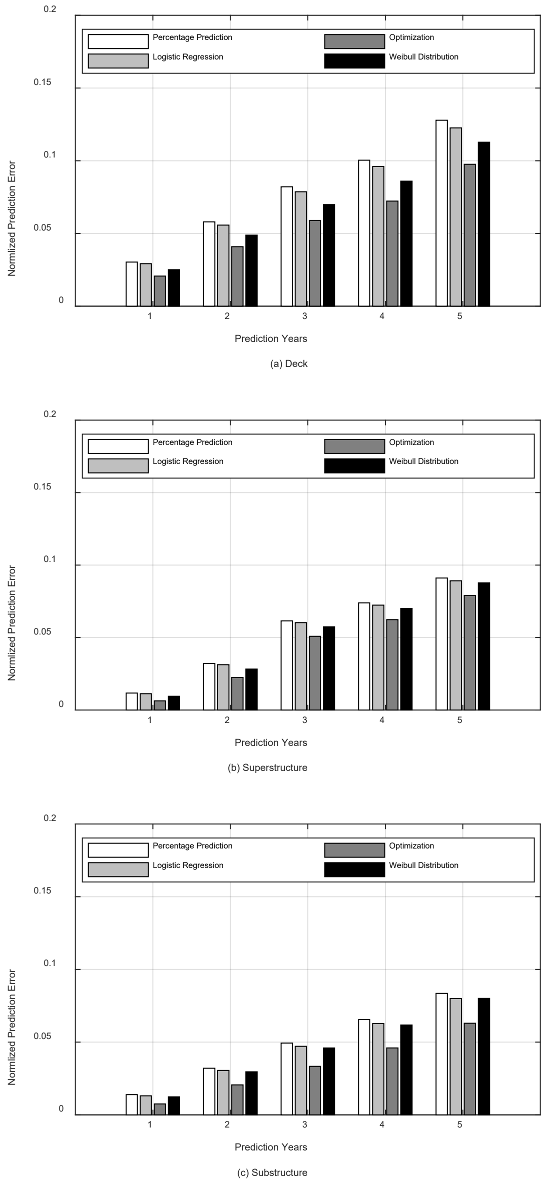
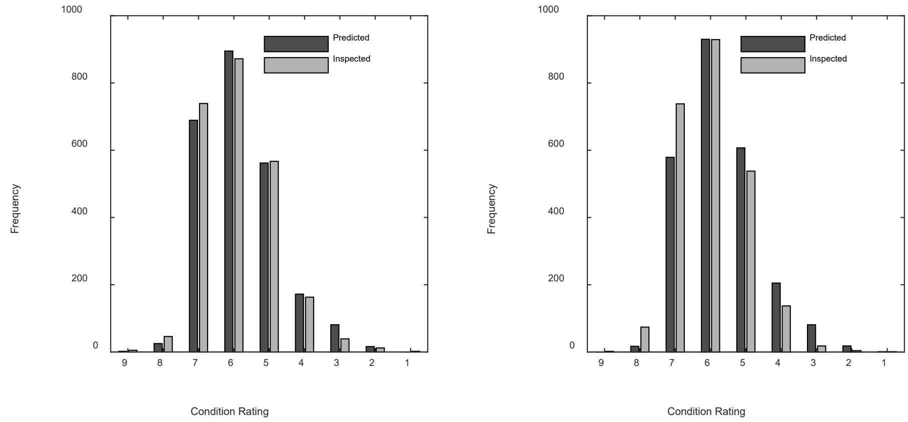


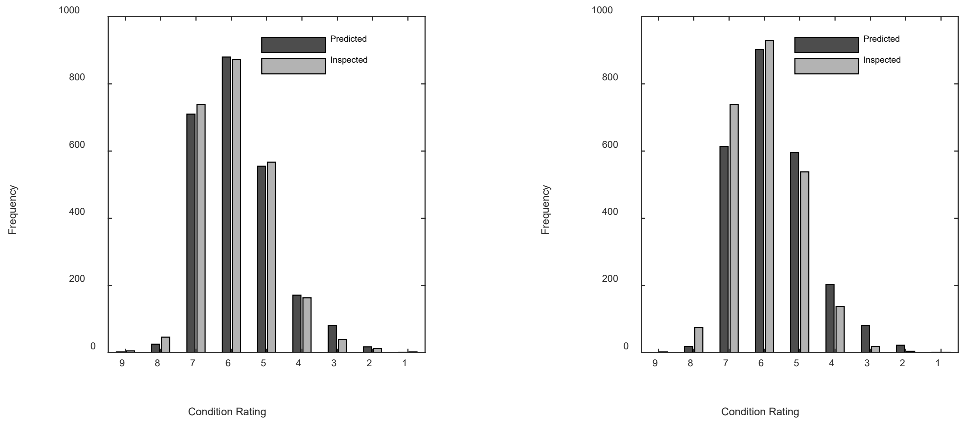
Figure 13. Normalized prediction error comparison of stochastic methods depending on prediction years for (**a**) deck, (**b**) superstructure, and (**c**) substructure.

The DMDP provides plots for the distribution of predicted condition ratings for the bridges inspected a particular year, which can be selected by the user. The predicted condition ratings were rounded to count the number of bridges for each condition. The predicted distribution was compared with the actual inspection data to examine the performance of the investigated algorithm. Figure 14 illustrates the distribution of two- and four-year predictions (i.e., using data from 2013) for the deck of all Wyoming bridges when the logistic regression algorithm was used without considering classification. Likely, due to the lack of repair information, the increase in the condition ratings was unable to be captured. In same manner, the distribution comparison was conducted using the second level of classification, as shown in Figure 15. Although minor improvement was observed when compared to Figure 14, the overall forecasting pattern only resulted in minor changes.



(**a**) (**b**)

Figure 14. Bridge distribution comparison between inspection and predicted condition ratings without considering classification for (**a**) two-year and (**b**) four-year prediction.



(**a**) (**b**)

Figure 15. Bridge distribution comparison between inspection and predicted condition ratings using second level of classification tree for (**a**) two-year and (**b**) four-year prediction.

7. Conclusions

Deterioration models should provide reasonable deterioration estimates to support the maintenance decision-making process and to enable governments to allocate an appropriate budget for SHM, including repair, rehabilitation, and reconstruction. Numerous local governments have developed deterioration models for the effective monitoring of their bridge assets. To enable easy access to deterioration models, this study presented a toolbox, known as the DMDP, and compared the performance of its embedded algorithms.

The DMDP was successfully used to develop deterioration models, and its convenient usability as a potential tool for BMS was demonstrated. The weighted LASSO was newly implemented for the consistent selection of explanatory variables with improved performance. Additional investigation was conducted to suggest an appropriate number of explanatory variables, which were used to classify the bridge inventory into groups with common factors. On the basis of the comparison of normalized prediction errors according to the classification tree level, we were able to suggest the optimum number of explanatory variables in terms of the number of bridges in the target state's inventory.

Deterioration models were then developed for Wyoming bridges using various methods embedded in the DMDP, and the normalized prediction errors were analyzed to compare their performance. Although the deterministic deterioration models were unsuitable for predicting future conditions, they were still effective for observing the distribution of current bridge assets and construction trends. The performance of the stochastic methods varied on the basis of the purpose of the developed deterioration models. The optimization-based and Weibull distribution-based approaches exhibited superior performance for near-future prediction without considering classification. Although the logistic regression and percentage prediction algorithms were both preferable for long-term monitoring and high classification tree levels, the logistic regression algorithm generally exhibited better performance.

The error comparison also demonstrated the importance of considering the bridge performance improvement to develop deterioration models. The current guide did not request to record the maintenance action such as repair and rehabilitation, and thus the deterioration curve only decreased. Indeed, the historical inspection record included an increase in condition ratings without proper explanation. To improve future predictive models, the condition rating changes followed by such maintenance should be modeled as well.

Author Contributions: Conceptualization and methodology: M.C. and M.M.; software, validation, formal analysis, investigation, data curation, writing—original draft preparation, visualization, supervision, project administration, and funding acquisition: M.C.; resources, writing—review and editing: M.M. All authors have read and agreed to the published version of the manuscript.

Funding: This research is supported by a grant from R&D Program of Myongji University (2019 year) and the Korea Agency for Infrastructure Technology Advancement (KAIA) funded by the Ministry of Land, Infrastructure, and Transport (no. 20SCIP-B128492–04).

Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results.

References

1. Thompson, P.D.; Small, E.P.; Johnson, M.; Marshall, A.R. The Pontis bridge management system. *Struct. Eng. Int.* **1998**, *8*, 303–308.
2. Miyamoto, A.; Kawamura, K.; Nakamura, H. Bridge management system and maintenance optimization for existing bridges. *Comput. Aided Civ. Infrastruct. Eng.* **2000**, *15*, 45–55.
3. Frangopol, D.M.; Saydam, D.; Kim, S. Maintenance, management, life-cycle design and performance of structures and infrastructures: A brief review. *Struct. Infrastruct. Eng*. **2012**, *8*, 1–25.
4. Agrawal, A.K.; Kawaguchi, A.; Chen, Z. Deterioration rates of typical bridge elements in New York. *J. Bridge Eng.* **2010**, *15*, 419–429.
5. Egger, C.; Scheidegger, A.; Reichert, P.; Maurer, M. Sewer deterioration modeling with condition data lacking historical records. *Water Res.* **2013**, *47*, 6762–6779.
6. Anyala, M.; Odoki, J.B.; Baker, C.J. Hierarchical asphalt pavement deterioration model for climate impact studies. *Int. J. Pavement Eng.* **2014**, *15*, 251–266.
7. Sobanjo, J.; Mtenga, P.; Rambo-Roddenberry, M. Reliability-based modeling of bridge deterioration hazards. *J. Bridge Eng.* **2010**, *15*, 671–683.
8. Abed-Al-Rahim, I.J.; Johnston, D.W. Bridge element deterioration rates. *Transp. Res. Rec.* **1995**, *1490*, 9–18.
9. Federal Highway Administration (FHWA). *Recording and Coding Guide for the Structure Inventory and Appraisal of the Nation's Bridges*; Report No. FHWA-PD-96-001; FHWA: Washington, DC, USA, 1995.
10. Hearn, G.; Frangopol, D.; Chakravorty, M.; Myers, S.; Pinkerton, B.; Siccardi, A.J. Automated generation of NBI reporting fields from Pontis BMS database. In *Infrastructure: Planning and Management*, Gifford, J.L., Uzarski, D.R., McNeil, S., Eds.; ASCE: Reston, VA, USA, 1993; pp. 226–230.
11. Morcous, G. Performance prediction of bridge deck systems using Markov chains. *J. Perform. Constr. Facil.* **2006**, *20*, 146–155.
12. Chang, M.; Maguire, M. *Developing Deterioration Models for Wyoming Bridges*; Report No. FHWA-WY-16/09F; WY. Dept. Transp.: Cheyenne, WY, USA, 2016.
13. Thomas, O.; Sobanjo, J. Comparison of Markov chain and semi-Markov models for crack deterioration on flexible pavements. *J. Infrastruct. Syst*. **2013**, *19*, 186–195.
14. Cavalline, T.L.; Whelan, M.J.; Tempest, B.Q.; Goyal, R.; Ramsey, J.D. *Determination of Bridge Deterioration Models and Bridge User Costs for the NCDOT Bridge Management System*; Report No. FHWA/NC/2014-07; NC. Dept. Transp.: Carthage, NC, USA, 2015.

15. Assaad, R.; El-adaway, I.H. Bridge infrastructure asset management system: Comparative computational machine learning approach for evaluating and predicting deck deterioration conditions. *J. Infrastruct. Syst.* **2020**, *26*, 04020032.
16. Frangopol, D.M.; Lin, K.Y.; Estes, A.C. Life-cycle cost design of deteriorating structures. *J. Struct. Eng.* **1997**, *123*, 1390–1401.
17. Kong, J.S.; Frangopol, D.M. Cost–reliability interaction in life-cycle cost optimization of deteriorating structures. *J. Struct. Eng.* **2004**, *130*, 1704–1712.
18. Liu, C.; Hammad, A.; Itoh, Y. Maintenance strategy optimization of bridge decks using genetic algorithm. *J. Transp. Eng.* **1997**, *123*, 91–100.
19. Grall, A.; Bérenguer, C.; Dieulle, L. A condition-based maintenance policy for stochastically deteriorating systems. *Reliab. Eng. Syst. Saf.* **2002**, *76*, 167–180.
20. Dromey, L.; Ruane, K.; Murphy, J.J.; O'Rourke, B.; Lacey, S. A bridge-rehabilitation strategy based on the analysis of a bridge-inspection data set. *Infrastruct. Asset Manag.* **2020**, *7*, 25–35.
21. Chang, M.; Maguire, M.; Sun, Y. Framework for mitigating human bias in selection of explanatory variables for bridge deterioration modeling. *J. Infrastruct. Syst.* **2017**, *23*, 04017002.
22. Brown, M.; Chaganty, N.R. On the first passage time distribution for a class of Markov chains. *Ann. Probab.* **1983**, *11*, 1000–1008.
23. Scheidegger, A.; Hug, T.; Rieckermann, J.; Maurer, M. Network condition simulator for benchmarking sewer deterioration models. *Water Res.* **2011**, *45*, 4983–4994.
24. Chang, M.; Maguire, M.; Sun, Y. Stochastic modeling of bridge deterioration using classification tree and logistic regression. *J. Infrastruct. Syst.* **2019**, *25*, 04018041.
25. Hatami, A.; Morcous, G. Deterioration models for life-cycle cost analysis of bridge decks in Nebraska. *Transp. Res. Rec.* **2012**, *2313*, 3–11.
26. Jiang, Y.; Saito, M.; Sinha, K.C. Bridge performance prediction model using the Markov chains. *Transp. Res. Rec.* **1988**, *1180*, 25–32.
27. Madanat, S.; Kalaftis, M.G.; McCarthy, P.S. Probabilistic infrastructure deterioration models with panel data. *J. Infrastruct. Syst.* **1997**, *3*, 4–9.
28. Sianipar, P.M.M.; Adams, T.M. Fault-tree model of bridge element deterioration due to interaction. *J. Infrastruct. Syst.* **1997**, *3*, 103–110.
29. Black, M.; Brint, A.T.; Brailsford, J.R. A semi-Markov approach for modelling asset deterioration. *J. Oper. Res. Soc.* **2005**, *56*, 1241–1249.
30. Mishalani, R.G.; Madanat, S.M. Computation of infrastructure transition probabilities using stochastic duration models. *J. Infrastruct. Syst*. **2002**, *8*, 139–148.
31. Sobanjo, J.O. State transition probabilities in bridge deterioration based on Weibull sojourn times. *Struct. Infrastruct. Eng.* **2011**, *7*, 747–764.
32. Yan, W.J.; Papadimitriou, C.; Katafygiotis, L.S.; Chronopoulos, D. An analytical perspective on Bayesian uncertainty quantification and propagation in mode shape assembly. *Mech. Syst. Signal Process.* **2020**, *135*, 106376.
33. Enright, M.P.; Frangopol, D.M. Condition prediction of deteriorating concrete bridges using Bayesian updating. *J. Struct. Eng.* **1999**, *125*, 1118–1125.
34. Straub, D. Stochastic modeling of deterioration processes through dynamic Bayesian networks. *J. Eng. Mech.* **2009**, *135*, 1089–1099.
35. Zhang, H.; Marsh, D.W.R. Multi-state deterioration prediction for infrastructure asset: Learning from uncertain data, knowledge and similar groups. *Inf. Sci.* **2020**, *529*, 197–213.
36. Huang, Y.H. Artificial neural network model of bridge deterioration. *J. Perform. Constr. Facil.* **2010**, *24*, 597–602.
37. Lee, J.S.; Hwang, S.H.; Choi, I.Y.; Kim, I.K. Prediction of track deterioration using maintenance data and machine learning schemes. *J. Transp. Eng. Part A Syst.* **2018**, *144*, 04018045.
38. Liu, K.; El-Gohary, N. Ontology-based semi-supervised conditional random fields for automated information extraction from bridge inspection reports. *Autom. Constr.* **2017**, *81*, 313–327.
39. Butt, A.A.; Shahin, M.Y.; Feighan, K.J.; Carpenter, S.H. Pavement performance prediction model using the Markov process. *Transp. Res. Rec.* **1987**, *1123*, 12–19.
40. Mauch, M.; Madanat, S. Semiparametric hazard rate models of reinforced concrete bridge deck deterioration. *J. Infrastruct. Syst.* **2001**, *7*, 49–57.

41. Manafpour, A.; Guler, I.; Radlińska, A.; Rajabipour, F.; Warn, G. Stochastic analysis and time-based modeling of concrete bridge deck deterioration. *J. Bridge Eng.* **2018**, *23*, 04018066.
42. Baik, H.S.; Jeong, H.S.; Abraham, D.M. Estimating transition probabilities in Markov chain-based deterioration models for management of wastewater systems. *J. Water Resour. Plan. Manag.* **2006**, *132*, 15–24.
43. Veshosky, D.; Beidleman, C.R.; Buetow, G.W.; Demir, M. Comparative analysis of bridge superstructure deterioration. *J. Struct. Eng.* **1994**, *120*, 2123–2136.
44. Wellalage, N.K.W.; Zhang, T.; Dwight, R. Calibrating Markov chain–based deterioration models for predicting future conditions of railway bridge elements. *J. Bridge Eng.* **2015**, *20*, 04014060.
45. Le Son, K.; Fouladirad, M.; Barros, A.; Levrat, E.; Iung, B. Remaining useful life estimation based on stochastic deterioration models: A comparative study. *Reliab. Eng. Syst. Saf.* **2013**, *112*, 165–175.
46. Bu, G.; Lee, J.; Guan, H.; Blumenstein, M.; Loo, Y.C. Development of an integrated method for probabilistic bridge-deterioration modeling. *J. Perform. Constr. Facil.* **2014**, *28*, 330–340.
47. Federal Highway Administration (FHWA). *Bridge Preservation Guide: Maintaining a State of Good Repair Using Cost Effective Investment Strategies*; Report No. FHWA-HIF-11-042; FHWA: Washington, DC, USA, 2011.
48. Morcous, G.; Hatami, A. *Developing Deterioration Models for Nebraska Bridges*; Report No. SPR-P1 M302; NE Dept. Roads: Lincoln, NE, USA, 2011.
49. Chang, M. Deterioration Model Development Package User's Guide. Available online: http://home.mju.ac.kr/home/index.action?siteId=cmw321 (accessed on 1 December 2020).
50. Federal Highway Administration (FHWA). Download NBI ASCII Files. Available online: https://www.fhwa.dot.gov/bridge/nbi/ascii.cfm (accessed on 1 October 2020).
51. Themelis, K.; Rontogiannis, A.A.; Koutroumbas, K. Semi-Supervised Hyperspectral Unmixing via the Weighted Lasso. In Proceedings of the 2010 IEEE International Conference on Acoustics, Speech and Signal Processing, Dallas, TX, 14–19 March 2010; Institute of Electrical and Electronics Engineers; IEEE: Piscataway, NJ, USA, 2010.
52. Yeh, R.H. State-age-dependent maintenance policies for deteriorating systems with Erlang sojourn time distributions. *Reliab. Eng. Syst. Saf.* **1997**, *58*, 55–60.
53. American Road & Transportation Builders Association (ARTBA). 2018 Annual Report. Available online: https://www.artba.org/wp-content/uploads/2019/01/ARTBA-Annual-Report_1.17.19.FINAL_.pdf (accessed on 31 December 2019).