CIVIL & ENVIRONMENTAL ENGINEERING

RESEARCH REPORT

Development and Validation of Deterioration Models for Concrete Bridge Decks

Phase 1: Artificial Intelligence Models and Bridge Management System

by

Emily K. Winn

Rigoberto Burgueño

Report No. CEE-RR – 2013/01

June 2013

Research Report for MDOT under Contract No. 2009-0746/Z2
SPR No. 107451

Department of Civil and Environmental Engineering
Michigan State University
East Lansing, Michigan

1. Report No. RC-1587a	2. Government Accession No. N/A	3. MDOT Project Manager Peter Jansson
4. Title and Subtitle DEVELOPMENT AND VALIDATION OF DETERIORATION MODELS FOR CONCRETE BRIDGE DECKS – Phase 1: Artificial Intelligence Models and Bridge Management System		5. Report Date 6/11/2013 6. Performing Organization Code N/A
7. Author(s) Emily K. Winn and Rigoberto Burgueño		8. Performing Org. Report No. N/A
9. Performing Organization Name and Address Michigan State University 3546 Engineering Building East Lansing, MI 48824-1226		10. Work Unit No. (TRAIS) N/A 11. Contract No. 2009-0746 11(a). Authorization No. Z2
12. Sponsoring Agency Name and Address Michigan Department of Transportation Research Administration 8885 Ricks Rd. P.O. Box 30049 Lansing MI 48909		13. Type of Report & Period Covered Final Report 10/20/2009 – 9/30/2012 14. Sponsoring Agency Code N/A
15. Supplementary Notes		

16. Abstract

This research documents the development and evaluation of artificial neural network (ANN) models to predict the condition ratings of concrete highway bridge decks in Michigan. Historical condition assessments chronicled in the national bridge inventory (NBI) database were used to develop the ANN models. Two types of artificial neural networks, multi-layer perceptrons and ensembles of neural networks (ENNs), were developed and their performance was evaluated by comparing them against recorded field inspections and using statistical methods. The MLP and ENN models had an average predictive capability across all ratings of 83% and 85%, respectively, when allowed a variance equal to bridge inspectors. A method to extract the influence of parameters from the ANN models was implemented and the results are consistent with the expectations from engineering judgment. An approach for generalizing the neural networks for a population of bridges was developed and compared with Markov chain methods. Thus, the developed ANN models allow modeling of bridge deck deterioration at the project (i.e., a specific existing or new bridge) and system/network levels. Further, the generalized ANN degradation curves provided a more detailed degradation profile than what can be generated using Markov models. A bridge management system (BMS) that optimizes the allocation of repair and maintenance funds for a network of bridges is proposed. The BMS uses a genetic algorithm and the trained ENN models to predict bridge deck degradation. Employing the proposed BMS leads to the selection of optimal bridge repair strategies to protect valuable infrastructure assets while satisfying budgetary constraints. A program for deck degradation modeling based on trained ENN models was developed as part of this project.

17. Key Words	18. Distribution Statement
Bridge decks; Bridge inspection; Deterioration; Databases; Evaluation; Neural networks, Predictions, Ratings	No restrictions. This document is available to the public through the Michigan Department of Transportation.

19. Security Classification - report	20. Security Classification - page	21. No. of Pages	22. Price
Unclassified	Unclassified	168	N/A

Report No. CEE-RR – 2013/01

DEVELOPMENT AND VALIDATION OF DETERIORATION MODELS FOR CONCRETE BRIDGE DECKS

PHASE 1: ARTIFICIAL INTELLIGENCE MODELS AND BRIDGE MANAGEMENT SYSTEM

by

Emily K. Winn
Graduate Research Assistant

Rigoberto Burgueño, Ph.D.
Associate Professor of Structural Engineering

Research Report to Michigan DOT under Contract No. 2009-0746/Z2
SPR No. 107451

Department of Civil and Environmental Engineering
Michigan State University
East Lansing, MI 48824-1226

June 2013

DISCLAIMER

The opinions, findings, conclusions and recommendations presented in this report are those of the authors alone and do not necessarily represent the views and opinions of Michigan State University or the Michigan Department of Transportation.

This report has also been published as the MS Thesis of Ms. Emily K. Winn under the following reference:

Winn, E.K. Artificial Neural Network Models for the Prediction of Bridge Deck Condition Ratings. MS Thesis, Department of Civil and Environmental Engineering, Michigan State University, 2011.

ABSTRACT

Condition assessment is extremely vital in the decision making process of how millions of taxpayer dollars are spent on repairing or improving aging infrastructure. This research looks at developing artificial neural network (ANN) models to predict the condition ratings of concrete highway bridge decks in Michigan. Historical condition assessments chronicled in the national bridge inventory (NBI) database were used to develop the ANN models. The high complexity of the NBI database due to non-linear variable relationships, subjectivity from manual inspections, and missing data has limited utilization of the database in the development of conventional prediction models. However, ANNs can produce correct responses even in the presence of noise or uncertainty in the training data, and can satisfactorily predict the outcome of complex problems or those with a high degree of nonlinear behavior. Two types of artificial neural networks, multi-layer perceptrons and ensembles of neural networks (ENNs), were developed to predict the condition ratings of concrete bridge decks. The performance of the ANN models was evaluated by comparing them against recorded field inspections and using statistical methods. The MLP and ENN models had an average predictive capability across all ratings of 83% and 85%, respectively, when allowed a variance equal to bridge inspectors. A method to extract the influence of parameters from the ANN models was implemented and the results are consistent with the expectations from engineering judgment. An approach for generalizing the neural networks for a population of bridges was developed and compared with Markov chain methods. Thus, the developed ANN models allow modeling of bridge deck deterioration at the project (i.e., specific existing or new bridge) and system/network levels. Further, the generalized ANN degradation curves provided a more detailed degradation profile than what can be generated using Markov models. A bridge management system (BMS) that optimizes the allocation of repair and maintenance funds for a network of bridges is proposed. The BMS uses a genetic algorithm and the trained ENN models to predict bridge deck degradation. The genetic algorithm aims to minimize the repair costs over a pre-defined planning horizon while maintaining adequate bridge deck conditions. Employing the proposed BMS leads to the selection of optimal bridge repair strategies to protect valuable infrastructure assets while satisfying budgetary constraints.

ACKNOWLEDGMENTS

The research described in this report was carried out under funding from the Michigan Department of Transportation, Contract 2009-0746/Z2, SPR No. 107451, with Mr. Peter Jansson as project manager. The financial support of MDOT and the coordination of Mr. Jansson throughout the execution of the experimental program are gratefully acknowledged. The authors also thank the expert assistance on this research from Dr. Syed Haider on the statistical and probabilistic analysis tasks.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

1 INTRODUCTION AND OBJECTIVES

1.1 Motivation

Bridge deterioration is a serious problem across the United States. A study by the American Society of Civil Engineers (ASCE) found that 27% of the 600,905 bridges in the United States were rated as structurally deficient or functionally obsolete in 2008 (ASCE, 2009). The cost of fixing these deficiencies and maintaining adequate conditions for the next 50 years is estimated at $850 billion. Not only is the economic burden of providing repairs and maintenance large, but the issue of deciding when repairs are to occur adds an additional level of complexity to the problem. The decision of when to intervene is a multiple part problem. First, the future condition of the bridge must be predicted with a high level of certainty. After the future condition is assessed, an optimum repair strategy can be determined. Identifying the optimum repair strategy includes determining which type of repair is the most appropriate and when this repair should occur. The complexity of identifying optimum repair strategies is significantly increased when multiple bridges are considered.

The current method for monitoring the structural health of bridges in the United States is dominated by visual inspections. During a visual inspection, the conditions of structural elements are examined, and a condition rating is assigned based on the presence of damage. The condition rating is measured on an integer scale of 0 to 9, with 9 being excellent condition, and a rating of 4 or below indicating poor condition. The condition records of bridge elements are recorded in the National Bridge Inventory (NBI) database. Although the NBI database contains a wealth of knowledge for the past conditions of bridge elements, no insight is directly available on the future conditions. Thus, repair and rehabilitation options can only be assessed after a completed inspection. This assessment system leads to reactive rather than preventive repair and rehabilitation strategies (Li and Burgueño 2010).

In an effort to establish preventive repair strategies, the Michigan Department of Transportation (MDOT) expands on the national data requirements from the NBI database with more detailed data in the Michigan Bridge Inventory in combination with use of the software Pontis (Robert et al., 2002). Pontis (AASHTO, 2009) is an asset management program that utilizes probabilistic degradation models to predict future bridge element conditions. Although

the system provides beneficial information on planning strategies, the probabilistic degradation models have been shown to contain flawed logic when applied to modeling structural degradation (Huang, 2010). The inadequacy of the probabilistic models to accurately model structural degradation has led to an increased interest among the bridge engineering community to develop alternative methods that provide more accurate results.

The condition ratings from visual inspections provide an overall assessment of the elements condition, but do not provide quantitative and tangible measurements on stresses, strains, cracking, or other physical manifestations of damage. Although mechanistic modeling of deterioration processes in concrete is important, such models are not practical at the state Department of Transportation (DOT) level due to the lack of necessary data, and the diversity in the bridge population. The resources are simply not available to monitor minute details such as chloride concentration to attain an estimate of damage due to rebar corrosion. The NBI database contains a wealth of knowledge when the task of modeling the degradation process is approached in an inverse manner. Although the true process of damage initiation is not known, the condition of bridge elements is known, along with unique parameters that may affect the degradation process. It is the aim of this research to harness these potential relationships to develop concrete bridge deck degradation models based on historical performance data. The condition of the bridge deck is of interest because the deck of a bridge experiences the most severe degradation and is commonly the driving force behind the decision to perform bridge repairs.

The inverse nature of the problem requires the use of nonconventional techniques. Artificial neural networks (ANNs) have been successful in applications of inverse problem mapping, and are proposed for this research. The ability of the ANN models to detect nonlinear relations between parameters due to their unique training method and structure makes them ideal for the current task (Haykin, 1999). ANNs can be conditioned to respond to a stimulus, generally respond correctly even in the presence of noise or uncertainty in the information network, and can satisfactorily predict the outcome of complex problems or those with a high degree of nonlinear behavior (Haykin, 1999).

1.2 Research Objectives

The objective of this research is to use artificial neural networks (ANN) to develop concrete bridge deck degradation models using historic information available in the National Bridge Inventory (NBI) database. The effectiveness of the developed models will be demonstrated through a comparison with current methods employed by MDOT and through implementation in a bridge management system (BMS) that considers repair strategies and life-cycle cost considerations. This main objective was achieved through the following tasks:

- Development of degradation models using artificial neural networks. Two models were developed based on: (i) individual multilayer perceptron (MLP) networks and (ii) ensembles of neural networks (ENN).
- Performance evaluation of the developed ANN models using statistical methods.
- Performance evaluation of the developed ANN models with a Markov probabilistic model.
- Development of network level and project level lifetime degradation curves to illustrate the decline in condition rating of bridge decks over the life of the bridge as predicted by the ANN and Markovian models.
- Perform sensitivity analysis of ANN models to identify the relative influence of input parameters.
- Development of a BMS framework for the identification of optimal maintenance strategies for concrete bridge decks.

1.3 Report Organization

Chapter 2 offers a review of literature concerning concrete degradation and a brief discussion of attempts to model the different processes. A review of the currently utilized Markovian deterioration models, an overview of ANN theory and current uses of ANN models with respect to structural condition assessments, and a review of BMSs is included. An overview to genetic algorithms is also presented. Chapter 3 provides an in depth look at the NBI database. It details the process of tailoring and refining the data for the task at hand and provides a discussion on the inherent complexities and imbalance of the database. Chapter 4 presents the development of ANN and Markovian degradation models. An in depth discussion of the model development,

input parameter selection, and model performance is provided. An investigation into the relative influence of the input parameters for the ANN models is also presented. The development of degradation curves using each of the models is chronicled in Chapter 5. The process for developing both project level and network level degradation curves for the ANN models is given. A comparison of the MLP and ENN model curves is provided along with a comparison of the ANN and Markovian curves. A small investigative study of differences in degradation due to rebar protective coating is also presented. Chapter 6 provides the details on the development of a proposed BMS strategy. The data organization scheme and essential BMS models are discussed. A new approach to model deck degradation after repairs utilizing statistical analyses of the NBI database and information provided by the Michigan Department of Transportation (MDOT) is developed and outlined. The developed BMS is used in a study of a 10-bridge network in the state of Michigan. Results of the optimum solution for the repair strategy for each bridge over a five-year planning horizon are presented.

2 LITERATURE REVIEW

This chapter provides a review of literature pertaining to the methods used in this research. First, concrete degradation modeling methods are reviewed, with an emphasis on the conventional Markov approach. This is followed by a presentation of the artificial neural network (ANN) models that were used for modeling concrete bridge deck degradation in this research. The review includes a theoretical discussion on the ANN models proposed, the multi-layer perceptron (MLP) and the ensemble of neural networks (ENN), along with a review of state-of-the art research with regards to utilizing ANN models for condition assessment. In addition, the concept of a bridge management system (BMS), which aids in identifying repair and rehabilitation strategies, is reviewed. A discussion on optimization tools, with an emphasis on genetic algorithms (GAs), which are used in this research, is provided.

2.1 Concrete Bridge Deck Deterioration Models

Concrete degradation occurs due to multiple physical and chemical processes. Processes include, but are not limited to, alkali-silica reaction, freeze thaw, sulfate attack, and corrosion of embedded reinforcing bars (Metha and Monteiro, 2006). Multiple attempts have been made to accurately model each of these processes (Derucher et al., 1994; Enright and Frangopol, 1998; Bhargava et al., 2006; Isgor and Razaqpur, 2005). These models cover a range of complexities, and in turn accuracy. Accurately modeling a deleterious process is only one piece of the deterioration puzzle. How each process manifests itself as causing damage, and how that damage is propagated through the concrete matrix is another difficult task to model. The effects of just one of these processes are difficult to capture. When multiple types of deterioration processes are to modeled together, the complexity of the problem is compounded, making numerical and mechanistic models difficult to develop. In the context of asset management, not only is the current condition of the concrete necessary, but an extrapolation as to future conditions is also needed. Due to the complexity and need for future condition predictions, stochastic methods are commonly used to model concrete degradation. One common method is the Markov assumption, which as described later, may be inappropriate in the context of modeling concrete degradation. Another emerging method is the reliability-based management

where the condition of the asset is measured in terms of safety (Frangopol et. al., 2001). However, the condition of bridges in the United States is currently monitored using visual inspections that only take into account the observed physical health of the bridge. To utilize the historic data provided by these inspections, ANN models are proposed, which are able to overcome the complexities in the data.

2.2 Markov Modeling for Bridge Condition Prediction

Many current bridge management systems model element degradation as a Markovian process. A Markovian process is stochastic and fulfills the following statement (Benjamin and Cornell, 1970):

$$P\left[X(n+1)=x_{n}+1\middle|\left(X(1)=x_1\right)\cap\left(X(2)=x_2\right)\cap\ldots\cap\left(X(n)=x_n\right)\right]= P[X(n+1)=x_{n}+1|X=x_n] \quad \textbf{2-1}$$

Equation 2-1 states that given the current and all previous values, the probability of $X(n+1)= x_{n+1}$ is equal to the probability of $X(n+1)= x_{n+1}$ when just the current value is given. In other words, Markov process is memoryless, meaning that future behavior depends only on the current state, and not the past history. The behavior of a Markovian process is modeled using transition probabilities, $p_{ij}(n)$, which provide the probability that the process will be in state j at time n given that it was in state i at the previous step. The transition probabilities are commonly displayed in matrix form, known as the transition probability matrix. If a process has r possible states, the transition matrix can be written in the form

$$\Pi=\begin{bmatrix} p_{11} & p_{12} & \cdots & p_{1r} \\ p_{21} & p_{22} & \cdots & p_{2r} \\ \vdots & \vdots & & \vdots \\ p_{r1} & p_{r2} & \cdots & p_{rr} \end{bmatrix} \quad \textbf{2-2}$$

Although Markovian models are commonly used in many capacities including commercial software packages such as Pontis (AASHTO, 2009) and Bridgit (NET, 1994), many researchers have questioned the state independence assumption. Mishalani and Madanat (2002) showed through an empirical study that the history independence assumption is unrealistic if continuous degradation processes, such as chemical processes, are the driving force behind the degradation.

However, the study did show the assumption of history independence may be appropriate for some condition states. Nonetheless, the authors make an argument in favor of time-based models rather than event-base models for improving the accuracy of transition probability matrices. Based on an empirical study of in-service pavements, Ramaswamy (1989) showed that there is an inherent relationship between deterioration and maintenance, which is contrary to the history independence assumption of Markovian processes. Additionally, Madanat et el. (1997) developed a probabilistic deterioration model for concrete bridge decks that provided evidence on the dependence of bridge deck degradation on history, which is inconsistent with the assumption of the Markov method. A study by Huang (2010) used statistical tests aimed at highlighting the weaknesses of the Markov method for estimating the future condition of bridges. Using inspection data from the state of Wisconsin the study generally found that the history independence assumption of the Markov model is invalid in the context of concrete bridge deck deterioration. However, age dependency was found to be less significant once deterioration had started, which is consistent with the Markov assumption. Studies such as these have exposed weaknesses in the Markov assumptions and provide additional support for the development of non-Markovian deterioration models. Although these studies show that the Markov approach is not valid for modeling bridge deck degradation, a Markov model was developed for the deck surface rating in this research to offer a comparison between the novel models developed and the current accepted method.

2.3 Artificial Neural Networks

The alternative to Markovian modeling presented in this research is modeling using ANNs. ANNs process information through a simple structure that attempts to mimic the thought process in the human brain (Liu, 2001). The human brain processes information through a series of interconnected elements, or neurons, which send simple inhibitory or excitation information via connections (Rumelhart et al., 1986). An ANN is a highly interconnected network of many simple linear or nonlinear processors (i.e., functions) connected in a parallel fashion. Each processing unit receives multiple inputs through weighted connections from neurons in the previous layer, performs appropriate computations, and transmits the output to other processing

units. The network performs operations by propagating changes in activation through the weighted connections between the processors (Haykin 1999).

Along with mimicking the human thought process, ANNs also attempt to capitalize on another unique characteristic of the human brain: learning. Artificial neural networks attempt to "learn" the correct input-output mapping through an iterative process. During the learning process, the weighted connections between neurons are modified by using optimization algorithms according to specific properties of the learning scheme. This learning process reduces the error of the ANN model, which provides a more accurate predictor. However, one of the dangers of using ANN models is the potential of overtraining, which is akin to overfitting. This phenomenon occurs when an ANN learns the connections within the training data so well that the ANN is unable to generalize the connections between inputs and outputs and makes the model unfit to use on new or unseen data. Due to their structure, which provides a limited ability to explicitly identify causal relationships, ANNs are also criticized as being 'black box' models (Tu, 1996). However, techniques such as those developed by Olden et al. (2002) can be used to reduce the 'black box' appearance by examining the connection weights between neurons and thus extract variable relationships and perform sensitivity analysis.

2.3.1 Anatomy of a Neuron

In an ANN model, the processing units, or neurons carry out all computations. The job of the neuron is simple: receive input values from neurons on the previous layer, compute an output value by sending the received inputs through a pre-defined function, and send the output to all neurons on the next layer. Figure 2-1 shows a schematic for a neuron, where the subscript n refers to the input for which the weight refers, and k refers to the neuron under examination. The inputs x_1 to x_n are the outputs of neurons from the previous layer. These inputs are multiplied by their respective weights, w_{kn} before entry into the neuron. These weighted inputs are combined via a predetermined propagation rule, commonly summation, to obtain the net input, net_k. The net input is then sent through an activation function, φ, to calculate the neuron output. The output is then passed along to all neurons in the subsequent layer. Equations 2-3 and 2-4 mathematically summarize the process just described (Liu, 2001).

The weights stored in the connections between neurons provide information regarding the level of effect that the first neuron has on the second. A positive weight forms an excitatory connection and a negative forms an inhibitory connection (Rumelhart et al., 1986). Although the weighted inputs are commonly combined using simple summation, more complex propagation rules can be utilized.

$$net_k = \sum_{j=1}^{n} w_{kj} x_j \qquad \textbf{2-3}$$

$$o_k = \varphi(net_k) \qquad \textbf{2-4}$$

Many different types of activation functions, denoted as φ in Figure 2-1, are used in neural network models. In the simplest cases, the activation function is the identity function where the inputs are directly equal to the output of the unit. The activation function can also be a threshold function where the input must exceed a certain value before being sent as an output to other units. Stochastic and continuous activation functions, such as the sigmoid function, are also commonly used.

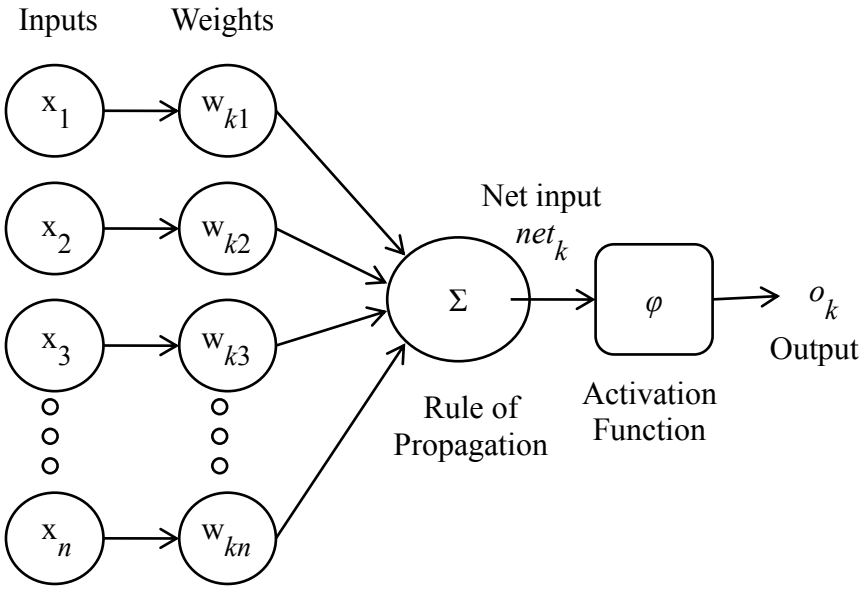


Figure 2-1: ANN neuron schematic

2.3.2 The Multilayer Perceptron

The neuron is the building block of all ANN models. One simple assembly and type of ANN model is the multilayer perceptron (MLP). A MLP consists of one input layer, one output layer, and one or more hidden layers. Input neurons receive input data from an external source. The hidden neurons receive inputs and send outputs within the network, and are not "visible" to outside systems. Output neurons send the signals computed by the system out of the network.

As shown in Figure 2-2, each neuron in a layer is connected to all neurons on the subsequent layer through the weighed connections. As indicated by the arrows, information is sent through the network of neurons in a forward fashion from the input layer to the output layer. Networks with this interconnection and forward-transfer of information have what is called feed-forward architecture.

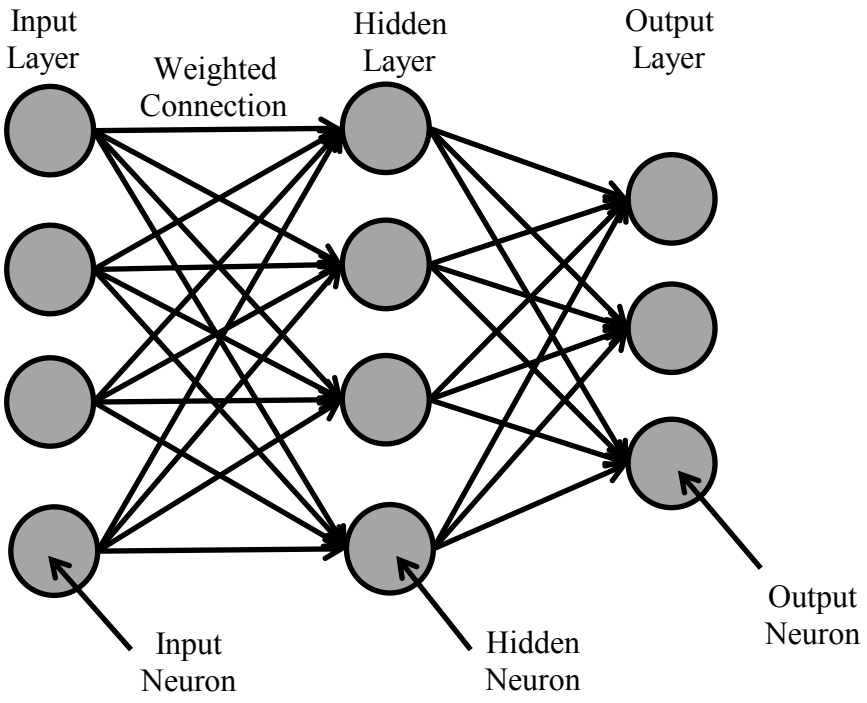


Figure 2-2: MLP schematic

Neural networks learn input output mapping through the presentation of examples. A common learning process is known as error back-propagation. Error back-propagation learning consists of two passes through the network: a forward pass and a backward pass. In the forward pass, an input vector, with a known target output, is applied to the sensory neurons of the network, and its effect propagates through the network layer by layer. During this forward pass, the connection weights of the network are fixed. At the end of the forward pass, the set of outputs produced is the response of the network. The error between the network response and the target response is calculated. During the backward pass, the connection weights are adjusted in accordance with the error back propagation algorithm. The back propagation algorithm is a stochastic gradient descent procedure used to minimize the squared error (Lui, 2001). The algorithm provides a way to calculate the gradient of the error function using the chain rule of differentiation. Bose and Liang, 1996 provide a detailed discussion of the back propagation algorithm and corresponding equations.

2.3.3 Artificial Neural Networks for Structural Condition Assessment

In most applications, neural networks have been used to identify the presence and/or magnitude of damage in structural elements. MLP ANNs were used by Pandey and Barai (1995) to detect and identify damaged structural elements in simple trussed bridge structures based on nodal displacements. The output of the network was the cross-sectional areas of members in an assumed damage zone and from the areas reductions in stiffness were calculated. Neural networks have also been used to classify structural damage in concrete and clay sewer pipes from assessment of photographs of damaged pipes (Moselhi and Shehab-Eldeen 2000). Although the identification of damage through photographs and nodal displacements may reduce the subjectivity associated with manual inspections, the application still requires a reactive maintenance approach. The current problem is one of an inverse nature where the symptoms are known (i.e., extent of damage indicated by the condition rating) but the causes of the symptoms are not. Efforts to develop neural networks used to solve this type of inverse problem discussed in the following. However, because the problem types and data types are vastly different, a direct comparison of model performances is not possible.

An MLP was used by Sobanjo (1997) to predict condition ratings of bridge superstructures using the age of the bridge as the only input. From a testing set of 38 ratings, the network was

able to predict 79% of these ratings correctly. Similarly, Cattan et al. (1997) developed a neural network approach to predicting the condition rating of railway bridges in the Chicago metropolitan area. The overall condition of the bridge on a rating scale of 1 to 5 was the output of the ANN, but the input vector contained physical characteristics of the bridge. Different combinations of input parameters were tested and the best performing network was able to correctly identify 73% of the ratings with unseen testing data. The conditions of underground water main pipes have also predicted using neural networks based on physical, environmental, and operational input parameters (Al-Barqawi and Zayed 2006). The lack of a common rating system between municipalities led to the development of a new, continuous rating system. No information on the size or nature of the data used is provided, but the ANN training software allowed for the contribution of each input variable to be measured. The pipe breakage rate was identified as the most important input parameter. These networks were able to predict 75% of the data correctly within 5% of the true rating, and 92% within 10% of the true rating.

ANNs have also been used to predict the rating of certain bridge elements. Li and Burgueño (2010) compared several ANN methods in their research on predicting bridge abutment condition ratings in the state of Michigan. The ANN models were responsible for predicting the discrete National Bridge Inventory (NBI) abutment rating based on physical and operational bridge parameters. The models that were compared were the MLP, radial basis function (RBF), support vector machine (SVM), supervised self-organizing map (SSOM), fuzzy neural network (FNN), and the ENN. Lifetime deterioration curves were produced based on predictions from the MLP model. The network yielded acceptable performance, correctly identifying the rating 56% of the time, identifying the rating within ±1 of the true rating 87% of the time, and successfully identifying a damaged abutment (rating of 4 or lower) 65% of the time. For the task of predicting concrete bridge deck condition ratings, Morcous (2002) compared an ANN model and a case-based reasoning (CBR) model. Although physical and operational parameters of the bridges were utilized from the Quebec Ministry of Transportation database, the rating scale is continuous and ranges from 1 to 6. When tested, 33% of the predicted ratings fell within a tolerance of ±0.1 of the original rating, and 100% fell within ±1.0 of the original rating. The lack of historical training data motivated the development of ANNs to generate past condition ratings to be used in predictions of future NBI deck ratings (Lee et al., 2008). The networks predicted the condition ratings based on non-bridge parameters pertaining to climate, traffic, and

population changes. Although the NBI rating scale was used as the condition rating, the networks were allowed to make non-discrete rating predictions. The network predicted 79% of the historical data within ±10% of the actual ratings. Using these generated ratings, future condition rating predictions were made based only on past condition ratings and occurred with an average error of 3%. A study by Huang (2010) utilized analysis of variance tests to show that the underlying age and history independence assumptions of Markov deterioration models are invalid in the context of predicting bridge deck deterioration. Eleven parameters in the NBI database that were identified as having a statistical effect in the bridge deck deterioration were used as inputs for an MLP model. The MLP output of the model was the bridge deck condition rating based on the 1 to 5 rating scale utilized by the management program Pontis. The network had a testing prediction performance rate of 75.3%.

2.3.4 Ensembles of Neural Networks

Figure 2-3 shows a schematic of an ensemble of neural networks (ENN), where many individual MLP models are used to determine the ensemble output. The main idea behind the ensemble methodology is that the combination of models, all of which are solving the same original task, provides a more accurate and reliable estimate than any one of the individual models (Rokach, 2005). Ensembles have been shown to effectively reduce the generalization-error (Domingos, 1996; Bauer and Kohavi, 1999) along with reducing variance-error without increasing bias-error (Tumer and Ghosh, 1999; Ali and Pazzani, 1996). Dietterich (2000) suggested that statistical, computational, and representational reasons are behind the improved ability of ensemble methods. The success of ensembles is also attributed to the accurate, but diverse population of the individual models. In a successful ensemble, each model should have a classification accuracy better than chance, and each model may develop its own bias toward a portion of the solution space (Hansen and Salamon, 1990).

The four main characteristics that are used to distinguish ensemble methods are: inter-classifier interaction, voting scheme, diversity generator, and ensemble size (Rokach, 2005). Inter-classifier interaction refers to how each individual classifier affects the other classifiers. The sequential technique incorporates a high level of interaction between models by using knowledge generated in the previous learning iteration to guide the learning in the subsequent iterations. Included in the sequential technique is the boosting method which produces training

sets for subsequent models from data that is not well recognized by previous models (Schapire, 1990). On the other hand, the concurrent methodology encourages much less interaction between models by training each model on their own unique training set. The bagging procedure is commonly used in concurrent ensembles and creates individual training sets by selecting information with replacement from the original data set (Brieman, 1996). This creates individuals that are trained on separate, but not independent data. The data organizational schemes of ensembles enable them to overcome challenges with unbalanced data distribution (Li, 2008). While the boosting and bagging methods differ in their level of inter-classifier connectivity, they both produce classifiers trained on unique data sets, which allows for the individual classifiers to develop a bias toward a portion of the solution space.

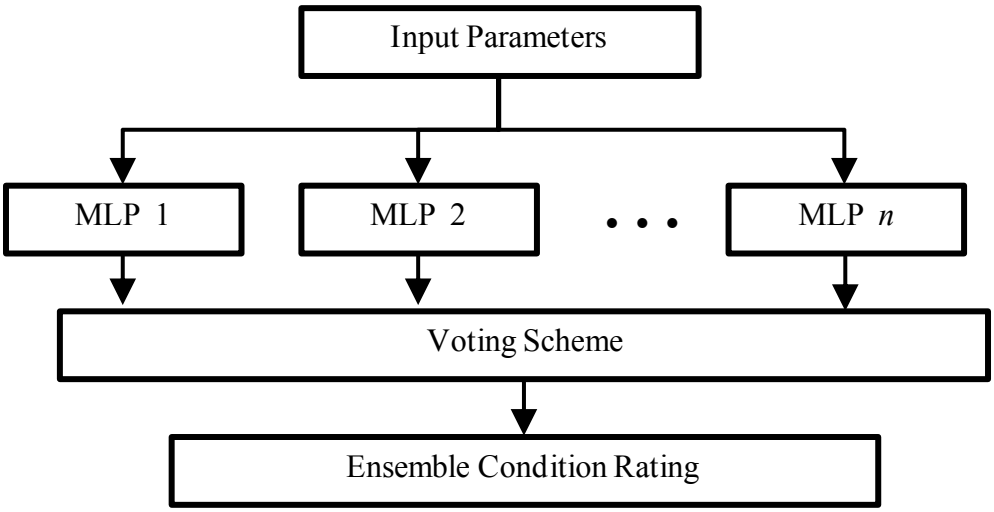


Figure 2-3: ENN diagram

The method in which the individual classifier outputs are combined to achieve the ensemble output is called a voting scheme. While many combination methods have been developed, two of the simplest are uniform voting and Bayesian combination. In uniform voting, all outputs from the individual classifiers are weighted the same. Bayesian voting however, weights the individual classifier outputs based on their posterior probability (Rokach, 2005). More complex combination methods such as stacking, arbiter trees, and grading have also been proposed (Rokach, 2005).

The success of the ensemble is also dependent on the diversity of individual classifiers and the number of classifiers in the ensemble. Diversified classifiers improve classification accuracy due to the development of uncorrelated errors between the classifiers (Hu, 2001). Diversity between classifiers can be introduced by manipulating the training set or by manipulating the classifier creation process (Rokach, 2005). In terms of a neural network, changes in the creation process are represented by different initial connection weights. The classification accuracy of the ensemble improves with an increase in the number of classifiers, but at a certain point the benefit from including more classifiers is negated by computational time and cost. Thus, the number of classifiers should be carefully selected based on the desired accuracy, user preferences, and computational time and computer storage allowed.

Ensembles of neural networks have been utilized in many facets. Mao (1998) used ENNs to improve optical character recognition performance. The investigation considered the effects of bagging, boosting, and basic ENNs. The study showed that ensembles using all three methods provided significant improvements in performance over that of an individual network. The boosting method provided slightly higher accuracy at a zero rejection rate, but this advantage decreased as networks were rejected based on poor performance.

In the financial sector, West et al. (2005) used ENNs to identify the credit health of customers. ENNs were trained to identify if customers had good credit or bad credit based on a description of their credit history including account balances, loan purposes, loan amounts, employment status, and personal information including age, housing, and career. The results from ensembles using three different creation strategies, cross validation, bagging, and boosting, were compared with the results from an individual MLP network. The ensemble methods produced a 3-5% reduction in generalized error that translated to a potential savings of $1.2 billion for the industry.

ENNs are also useful in the medical field as presented in a study by Zhou et al. (2001). The study used ENNs in an early stage lung cancer diagnosis system. The ENNs were utilized to identify cancerous cells based on images of needle biopsies. The images from the needle biopsies were pre-processed to obtain characteristics of the image, such as color component measurements, illumination, saturation, and color proportions. These parameters were used by the ENN to identify if the image contained cancerous cells, and if so, what type of cancer. Because the consequences of a false negative diagnosis, i.e. diagnosing a patient as healthy when

they do in fact have cancer, is so grievous, a two-step network ensemble was composed. The first layer identified with great certainty whether the patient has cancer. If the ENN decided that the patient does have cancer the image information was sent to another ENN trained specifically to identify what type of cancer the image contained. The two-level ENN architecture significantly improved the error measures, especially with regard to the error of false positive diagnoses.

Dackermann et al. (2009) used ENNs to identify damage in timber bridges. The research used the damage index (DI) based on the change in modal strain energy before and after damage to identify the location and severity of defects in timber beam structures. The ENN was composed of five MLPs each of which represents one of the first five shape modes of the beam. The DI values at multiple points on the beam for each respective mode shape were the inputs to the MLPs. The ensemble prediction was decided not by a voting scheme, but by another MLP. The outputs, the location or severity of damage, of the five MLP networks were sent to another MLP, which determined the final location and severity of damage. The ensemble architecture provided a significant improvement in the identification of damage than individual MLP networks.

Li (2008) employed ENNs in the prediction of bridge abutment condition ratings in the state of Michigan. The abutment condition ratings were recorded using the NBI discrete rating scale. As in this research, bridge parameters available in the NBI database were used to predict the abutment rating. The ENNs were developed using a newly proposed modified bagging procedure that drew the same number of training records from each discrete condition rating. This modified bagging procedure helped reduce the data imbalance. Additionally, several voting schemes were compared, with the modified majority voting scheme leading to the highest damage identification ratio. The ENNs provided a 13 to 18% increase in damage identification over the best performing MLP model. The methods used by Li (2008) were adopted in this research due to the success of the developed models and the similarity in the problem type.

2.4 Bridge Management Systems

A bridge management system (BMS) aims to optimize the allocation of maintenance, rehabilitation, and repair funds for one or several bridges or bridge components in a network or

system. A BMS can assess needs at the project level, i.e. an individual bridge, or at the network level. Miyamoto et al., (2000) developed a BMS that assess the needs of one bridge over several years, and is considered a project level BMS. Although it is beneficial to assess the needs of an individual bridge, the health of all bridges in a network is also an important component when it comes to infrastructure management. Network level BMSs considers the needs of multiple bridges and approaches to this end have been developed by Hegazy et al., (2004), Lee and Kim, (2007), and Liu et al., (1997). Lee and Kim (2007) illustrated the usefulness of genetic algorithms to identify optimum solutions for a network of bridges, but the research only considers a planning horizon of one year. The BMSs developed by Hegazy et al., (2004) and Liu et al., (1997) also consider a network of bridges, but aim to identify an optimum solution over a period of several years. BMSs offer support to engineers deciding on funding allocation, and can be tailored to meet the needs of a unique problem.

Although the problem each BMS is solving is similar but different problems, all BMSs share several components. The first of these components is a condition assessment model. Every BMS must have some way of assessing the current and future conditions of the components in question. Liu et al. (1997) used a linear deterioration model developed by Jacobs (1992) that considered the initial condition, yearly deterioration rates (which are a function of bridge age and traffic), and the impact of maintenance activities. The BMS developed by Lee and Kim (2007) only considered one year of planning and thus the direct assessment in the form of a visual inspection was used to identify the condition of the elements in question. Elbehairy et al. (2009) used different deterioration models depending on the bridge element. Joints, bearings, overlays, and finishing are assumed to deteriorate in a linear fashion while the bridge deck, superstructure, and substructure were modeled using a Markovian deterioration method. Transition matrices developed by Jiang (1990) were used for the Markov processes. Huang et al. (2004) utilized a mechanistic deterioration model developed by Babaei et al. (1996) in the assessment of life-cycle strategies for concrete bridge decks. The model used the percentages of spalled and delaminated areas, and the chloride content at the reinforcing bar level to calculate the condition index. Liu and Frangopol (2005a) used a reliability-based method developed by Frangopol et al. (2001) to develop the deterioration profiles of the bridge elements in a multi-objective BMS. Network level BMSs developed by Lui and Frangopol (2005b and 2006) used deterioration profiles developed by Akgül and Frangopol (2004). The lifetime profiles were developed using a

reliability approach using the system reliability and load and resistance models. From these examples, it can be seen that although a degradation model is an integral part of a BMS, multiple different methods have been successfully employed.

Maintenance actions and the effects of these actions on the element deterioration are also an essential part of any BMS, along with the cost of these actions. Liu et al. (1997) divided reconstructive efforts into four categories: routine maintenance, repair, rehabilitation, and replacement. Routine maintenance, repair, and rehabilitation only affect the condition of the bridge deck, which was referred to as the degree of deterioration. The rate of deterioration, or slope of the deterioration curve, remained the same for these repair options. A deck replacement restored the condition of the deck to the initial condition and the initial rate of deterioration. The BMS proposed by Lee and Kim (2007) also adopted the same maintenance scheme. The deterioration degree from Liu et al. (1997) was used along with opinions from bridge engineers to extrapolate the recovering effect in terms of deck condition. As this BMS only examined one year on the planning horizon, no model was proposed, nor necessary, to track the deterioration of the bridge deck after reconstructive efforts. Elbehairy et al. (2009) used a similar improvement model. The condition of the bridge element was improved by a set number of rating points, which was dependent on the repair selected. The deterioration post-repair was calculated using the same transition probability matrices that were used to model deterioration pre-repairs. There was no distinction or change in the deterioration model after reconstruction occurs. Huang et al. (2004) used the repair options of concrete and asphalt overlays, patching, and deck replacements in their life-cycle assessment model. Each repair type affected the spalled and delaminated deck area differently. The deterioration after repairs was modeled as the performance of a deck with black steel with 25-35 years of no deterioration. The reliability-based method developed by Frangopol et al. (2001) was used to model the deterioration after repairs in Liu and Frangopol (2005a). The effect of maintenance was modeled using multiple parameters including the time of first application, time interval of subsequent cyclic application, immediate performance improvement, time delay of deterioration, reduction of deterioration rate, and duration of the maintenance effect. Values for each of these parameters for silane treatment, minor concrete repairs, cathodic protection, and replacement were proposed. The source for the definition of these parameters was indicated as personal communications, and no further discussion on how the values are calculated was provided. Lui and Frangopol (2005b and 2006) adapted the results

of a study by Furuta et al. (2004) to model the effects of reconstructive efforts on the continued bridge deck degradation after repairs. The study included the repair types of resin injection, increasing the slab thickness, attaching a steel plate, and a full deck replacement. Resin injection was modeled as not having an effect in increasing the reliability index, but it was assumed to reduce the deterioration rate and had an effective duration of 15 years. Increasing the slab thickness and attaching a steel plate were assumed to have no effect on the deterioration rate but increased the reliability index by 0.7 and 2.0, respectively. A deck replacement was assumed to return the reliability index to the initial level and did not affect the deterioration rate. As with the condition assessment models, many different methods have been proposed to model the degradation of bridge elements after repairs have occurred. However, as the reviewed methods illustrate, very little knowledge of the true degradation process of concrete elements after repairs is available, and leads to models developed on the basis of assumptions.

Information from the discussed components is utilized in an optimization scheme to identify an optimum repair strategy. Multiple researchers have used genetic algorithms (GAs) as the optimization tool in a BMS (Liu et al., 1997; Lee and Kim, 2007; Elbehairy et al., 2009; Liu and Frangopol, 2005b; Miyamoto et al., 2000; and Liu and Frangopol, 2005a). GAs are well suited for optimization in a BMS framework for several reasons. First, unlike traditional optimization methods, GAs are able to search for multiple points on the solution space at once. This enables the GA to process a large number of potential solutions at one time. This is beneficial in a BMS problem due to the large solution space. GAs also do not rely on sensitivity information to update the solutions. The solutions are updated through probabilistic natural selection rules and not deterministic rules. Additionally, GAs are also able to easily handle multiple objective functions. Although GAs are only one type of optimization tool, as will be discussed in the following section, their ability to offer successful solutions to BMS problems as demonstrated by the reviewed literature, lead to the decision to adopt the GA optimization in this research. A more detailed review of GAs is provided later in the text.

2.5 Optimization Algorithms

In the discussed BMSs, GAs were utilized as the optimization tool to search for the optimal repair strategies. In an optimization problem the maximum or minimum of a function (or

problem) subject to constraints is searched for (Dantzig, 2010). The tools used for optimization depend on the complexity of the problem and range from simple calculus based analysis to complex numerical search algorithms (Snyman, 2005). The discussed GA is one tool that can be utilized in optimization and, for the reasons previously discussed, is well suited for the BMS problem type.

2.6 Genetic Algorithms

In this research, genetic algorithms were selected to search for the optimal solution in the BMS. Genetic algorithms are a type of evolutionary algorithm (EA) that is used in a branch of computing called evolutionary computation. This branch of computing draws inspiration from the natural evolution process (Eiben and Smith, 2003). All EAs utilize natural selection to develop a population that best fits the given environment. In terms of a BMS, this environment is the minimization of the cost of repairs, but maximization of the improvement of the bridge element conditions. Figure 2-4 shows the general scheme of a GA.

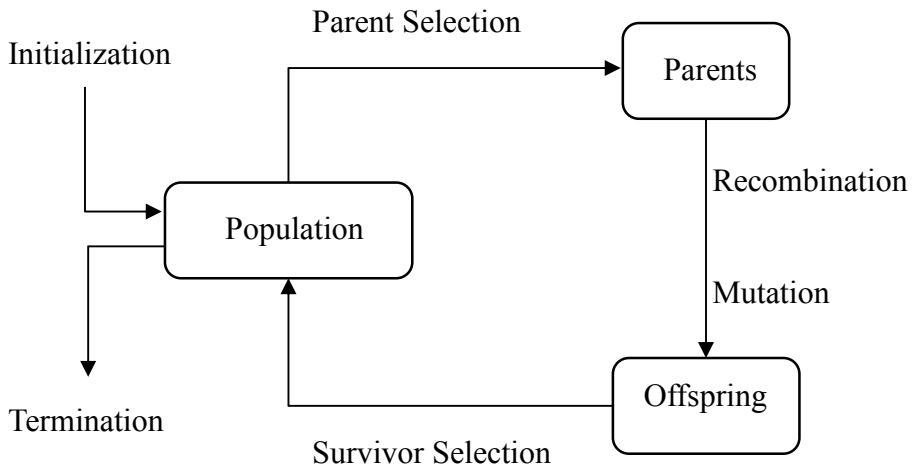


Figure 2-4: GA flowchart

To begin, a population of individuals is randomly created. Each individual represents a potential solution to the problem and is evaluated with regard to a fitness function. This evaluation tells how well the individual performs with regard to the environmental conditions. From the initial population, the individuals with the best fitness, i.e., performance, are chosen as

parents. These parents create offspring, i.e., new solutions, via recombination and mutation processes. The offspring are evaluated, and based on their fitness, may be selected as an individual in the new population. This process is repeated until the algorithm terminates. Common termination criteria include

- Reaching the maximum time limit.
- Reaching the maximum number of fitness evaluations.
- Obtaining a fitness value under a threshold.
- The development of a non-diverse population.

In the real world, an individual, i.e., a solution, is composed of phenotypes. For example, in an integer optimization problem, a sample chromosome may be [1 3 0 2]. Each integer is a phenotype. However, real world data is typically represented in binary form in the GA space. Thus, the chromosome would change to the following representation: [0 1 1 1 0 0 1 0]. Each phenotype is represented as a two-bit genotype.

Figure 2-5 shows a sample of a population, using binary representation, in the GA space. The population has a size *n,* and each chromosome has a length *i.* The size of a genotype in a GA is dependent on the data being represented. In Figure 2-5, each genotype is two bits in length, which represents phenotype values from 0 to 3. The total number of phenotypes in an individual is simply the chromosome length divided by the genotype size.

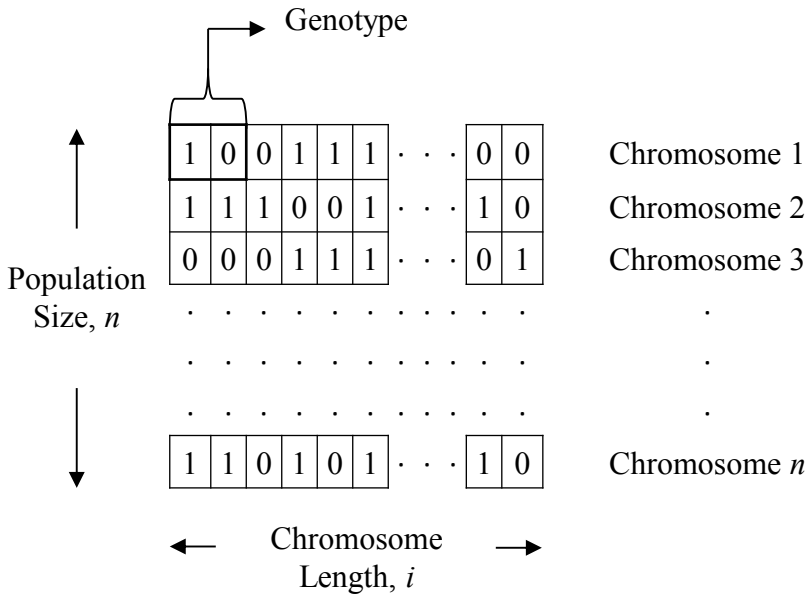


Figure 2-5: GA population

After the initial population is created, the fitness of each individual is determined and several individuals are selected as parents. The selection of parents is a stochastic process in which fitter individuals have a larger chance of being selected. Because the process is stochastic, it is possible that less fit individuals may be selected as parents. However, this prevents the GA from becoming trapped in a local optimum. The selected parents are then subjected to variation operators to create new individuals. The variation operator of mutation takes one parent individual and slightly modifies it to create a child, i.e., a new possible solution. The modification that occurs is selected at random, but the amount of modification can be defined by the user.

Figure 2-6 shows a schematic of mutation. In the parent, the two bits indicated by the gray cells were randomly selected to be modified. As seen in the child, the values of these bits are switched in the child from that of the parent. In genetic algorithms, mutation is used to fill the population with "fresh blood" (Eiben and Smith, 2003). Recombination, or crossover, is another variation operator. Crossover can be thought of as mating, where the traits of both parents are combined to create a child. The parts of each parent that are to be combined, represented by the gray cells in Figure 2-7, are selected at random. The resulting child has traits of both parents but may or may not have an improved fitness.

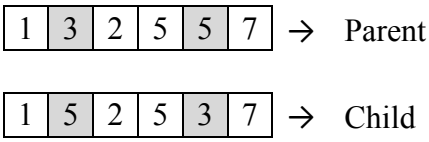


Figure 2-6: GA Mutation

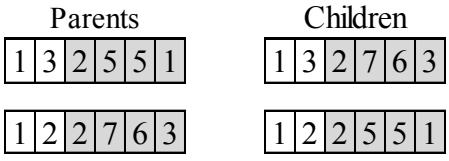


Figure 2-7: GA Crossover

Because the population size of a GA usually remains constant, a decision of whether the children will be included in the new population must be made. This decision is deterministic and

is commonly based the child's fitness value. The commonly accepted rule is that if a child is fitter than the least fit individual, the child replaces the old individual in the new population. Some GAs also utilize the 'elite count' scheme as a reproduction technique. This technique guarantees that a given number of the fittest individuals are members of the new population.

2.7 Summary

This chapter provided a review of the conventional techniques, namely Markov chains, utilized by MDOT to predict concrete bridge element condition ratings. This method of condition assessment was reviewed, and a discussion on the appropriateness of the history independence assumption was provided. Additionally, a review of ANNs, the proposed alternative approach, was presented. Literature for two ANN models, the MLP and the ENN, was provided and illustrates the effectiveness of the neural network approach to provide accurate condition assessments. A review of the selection of an optimal repair strategy using a BMS was also provided along with a detailed discussion of the GA approach to optimization. Each of the reviewed methods will be utilized in the following chapters in the development of ANN models to predict concrete bridge deck condition ratings using information available in the NBI database.

3 NBI DATABASE AND DATABASE REFINEMENT

3.1 Overview

This chapter provides an assessment of the data used in this research. The data used in this research was condition assessments of bridges in Michigan. This data was obtained through the Michigan Department of Transportation (MDOT) which maintains the records as part of the National Bridge Inventory (NBI) database. This chapter provides details on the characteristics and complexities of the NBI database. Additionally, the chapter chronicles the process of preparing the data for use in the degradation models.

3.2 Introduction

The operational conditions of bridges in the United States are archived in the NBI database. The development of the NBI database was mandated by the Federal Highway Administration (FHWA) in the 1970's after the collapse of the Silver Bridge, which connected Ohio and West Virginia across the Ohio River, in 1967. Each state is responsible for maintaining up to date operational, geometric, and condition information on bridges and tunnels in the states' inventory. Geometric parameters include physical bridge measures such as span length, skew angle, deck width, and design load. Operational conditions include average daily traffic (ADT), age, and highway classification. The "health" or structural condition bridges in the database is monitored through bridge inspections. The most common method of determining the health of a bridge during an inspection is through visual assessment. Several different types of bridge inspection types exist. Of these types, routine bridge inspections are the most common. Initial inspections occur at the start of a bridge's life and are used to establish condition baselines for the bridge. Occasionally, in-depth inspections are conducted to identify damage not normally observed during routine inspections. These inspections are used to assist in the maintenance and rehabilitation management of the inventory.

The structural condition (or rating) of bridge components in the NBI rating system is measured on an integer scale of 0 to 9, with 9 being excellent condition, and a rating of 4 or below indicating poor condition. A condition rating of 4 or lower for the bridge superstructure,

substructure, or deck qualifies the bridge as being structurally deficient. Table 3-1 gives the descriptions of the condition ratings for the bridge deck.

Table 3-1: Deck condition rating descriptions

Rating	Description
9	**Excellent -** No noticeable or noteworthy deficiencies.
8	**Very Good -** Minor cracking with no spalling, scaling or delamination.
7	**Good –** Cracking at a spacing of 10 ft or more, with light shallow scaling.
6	**Satisfactory** – Minor deterioration including cracks at a spacing of 5 ft or less, medium scaling, and 2% or less the deck area spalled or delaminated.
5	**Fair** – Minor section loss, between 2% and 10% of the surface area is spalled or delaminated and excessive cracking in the surface.
4	**Poor** – Advanced section loss, large areas of the surface is spalled or delaminated.
3	**Serious** – Deterioration has seriously affected primary structural components and local failures are possible.
2	**Critical** – Advanced deterioration of primary structural elements. Emergency surface repairs required by the crews.
1	**"Imminent" Failure** – Major deterioration present in critical structural components. Bridge is closed to traffic, but corrective action may put the bridge back in service.
0	**Failed** - Bridge closed and is beyond corrective action.

Although the conditions of multiple bridge elements are inspected and monitored, the condition ratings of the bridge decks is the focus in this research. Three separate parameters are used to assess the health of a bridge deck: the top surface condition, the bottom surface condition, and overall condition. This research aims to utilize both the bridge geometric and operational data to predict the condition of a bridge deck in terms of the NBI rating system.

In order utilize the database to its fullest potential, an exploratory investigation of its underlying characteristics was necessary. The investigation examined basic characteristics of the database including the distribution of ratings with time, the distribution of ratings for the entire database, and the relationship between database parameters and the deck condition rating. An

examination of the manual inspection process and the complexities introduced into the database because of the process was also completed.

3.3 Structure of the NBI Database

The NBI database contains valuable information that can be utilized in condition prediction models. However, the database is unbalanced, noisy, contains large amounts of scatter, and due to its relatively recent implementation. Further, each bridge in the database only provides a small window in time for which condition ratings are available. Additionally, the visual inspection process is inherently subjective and subject to human error. The following sections provide an in-depth look at these difficulties.

3.3.1 Availability of Data

Although the creation of the NBI database was chartered in the early 1970s, the earliest relevant inspections in the data available for the state of Michigan are from 1992. The most recent inspections available for this research are from 2010, which means that if a bridge was built before 1992, the absolute longest period of time that one bridge can cover is 18 years. Because inspections are commonly conducted on a biannual basis, the maximum number of inspections that a bridge can contribute to the database is nine. However, the probability that a bridge actually contributes nine inspection records is low because: (i) bridges have been built after 1992, (ii) inspections do not always occur on a two year basis, (iii) some inspection records are missing or miscoded, and (iv) many inspections occur after reconstructive work has been completed on the bridge, and are thus not useful to simulate the main deck degradation process.

Figure 3-1 provides insight as to how many inspections per bridge are available in the database. In the figure, the number of inspections of bridges is on the x-axis and the height of the column represents how many bridges in the database provide the given number of inspections. As seen in the figure, the large majority of bridges contribute less than nine inspection records. In reality, 60% of the bridges contribute five inspection records or less. In the context of a bridge's total lifespan, the inspections cover an extremely small window of time, as shown in Figure 3-2. Thus, for the given bridge, only five inspection records are available in the database.

These five inspection records only provide a glimpse of the bridge's life and subsequent condition rating deterioration.

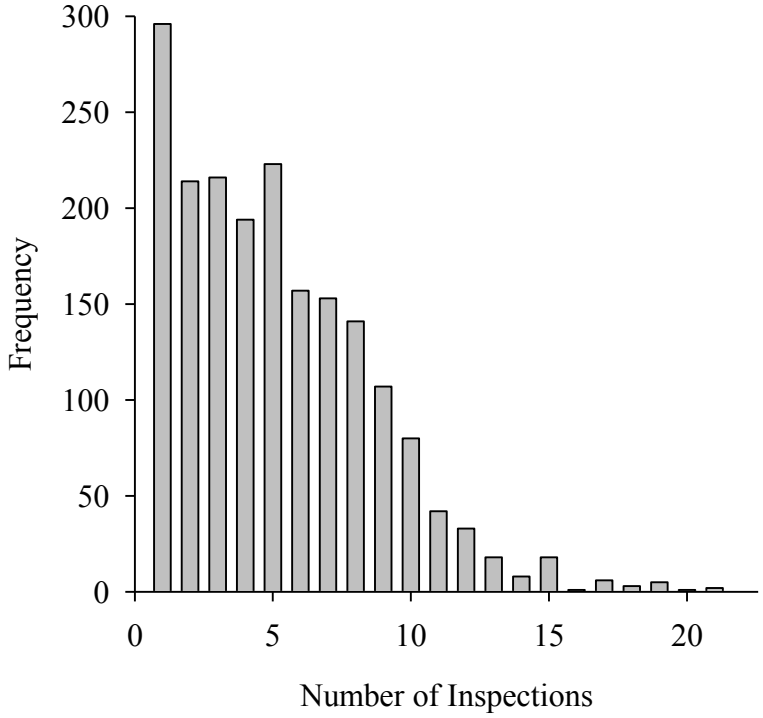


Figure 3-1: Distribution of the number of inspections

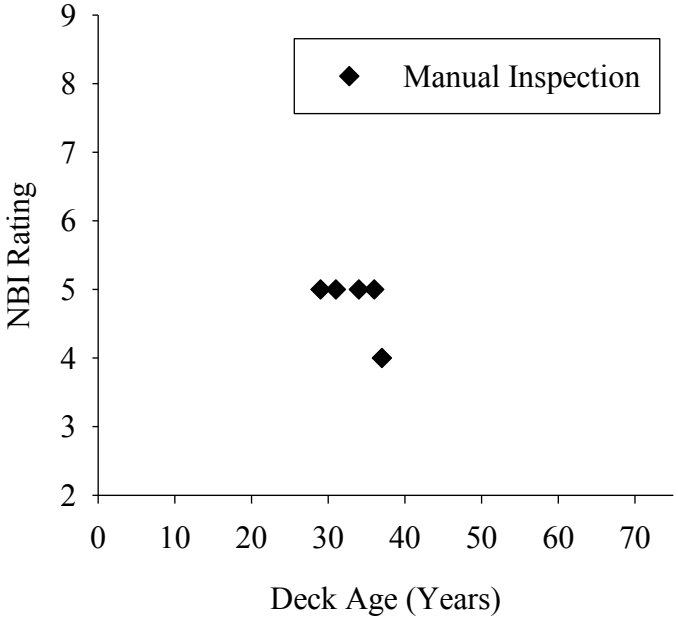


Figure 3-2: Inspection records of bridge 09109035000S130

The small window of time provided by the historical condition data of a single bridge requires the entire database be utilized in order to draw meaningful conclusions about the deck condition in relation to other database parameters. Thus, the entire database must be utilized to draw meaningful conclusions about the condition ratings. However, this is difficult due to the severe scatter in the data.

3.3.2 Database Imbalance

Figure 3-3 provides an example of the imbalance in the NBI database. As seen in the figure, there is a large amount of mid-condition ratings, such as '5', and very few instances of very high and very low condition ratings. Additionally, 84% of ratings can be classified as 'acceptable' by having a condition rating of 5 to 9, and only 16% classify as 'poor' by having a rating of 4 or below. The imbalance in low ratings arises due to reconstructive efforts intended to prevent bridges from reaching poor conditions.

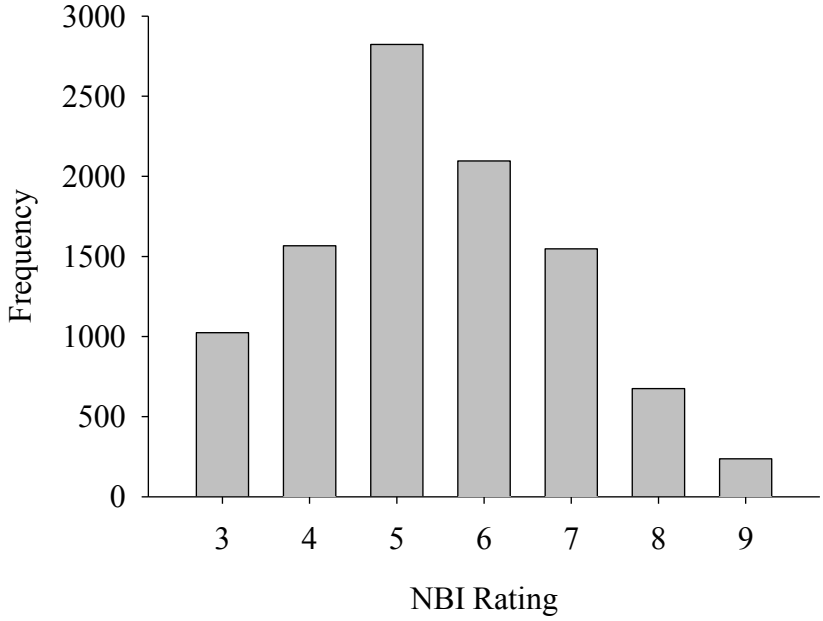


Figure 3-3: Distribution of condition ratings

The disproportionately low number of high condition ratings may be due to: (i) the fact that bridge decks do not remain at such high ratings for extended periods of time, (ii) the fact that few new bridges have been built since the earliest inspection year in the database (1992) and (iii) the

reluctance of a bridge inspector to rate a structural component as in excellent condition. Regardless of the reasons, the imbalance means that very little training data is available for the ANN models to accurately predict ratings with a small number of records.

3.3.3 Database Scatter

As shown in Figure 3-4, if all records are utilized, a large amount of scatter is present between different database parameters and the deck condition rating. Although the relationship between the deck age and the condition rating (as will be shown later) has the strongest correlation with the condition rating, the importance of this relationship unrecognizable in the figure due to the scatter. The scatter is also severe for other parameters, such as the average daily traffic (ADT), skew angle, and percent truck traffic as shown in Figure 3-4. It follows that the success of regression modeling techniques for this type of problem is limited.

3.3.4 Subjectivity in Inspection Process

In addition to the data imbalance and scatter, the subjectivity of the inspection records adds an additional level of complexity. Because visual assessment is the most common tool used in bridge inspections, the process is inherently subjective. Although guidelines and training programs exist for inspectors, human and environmental factors play a large role in the bridge inspection process. The subjectivity in the inspection process was documented in a study of 49 bridge inspectors from 25 state departments inspecting the same bridges (Phares et al., 2001). The subjects were asked to complete seven routine inspections and provide ratings for the deck, superstructure, and substructure for each. The results of the study were normally distributed and it was found that 68% of the ratings fell within an interval of ± 1 rating point of the mean, and 95% fall within an interval of ± 2 rating points of the mean. Multivariate nonlinear analysis determined that the inspector's fear of traffic, near visual acuteness, color vision attributes, formal bridge inspection training, and the inspector's perception of the bridge's maintenance, accessibility, and complexity have a relationship with the inspection results.

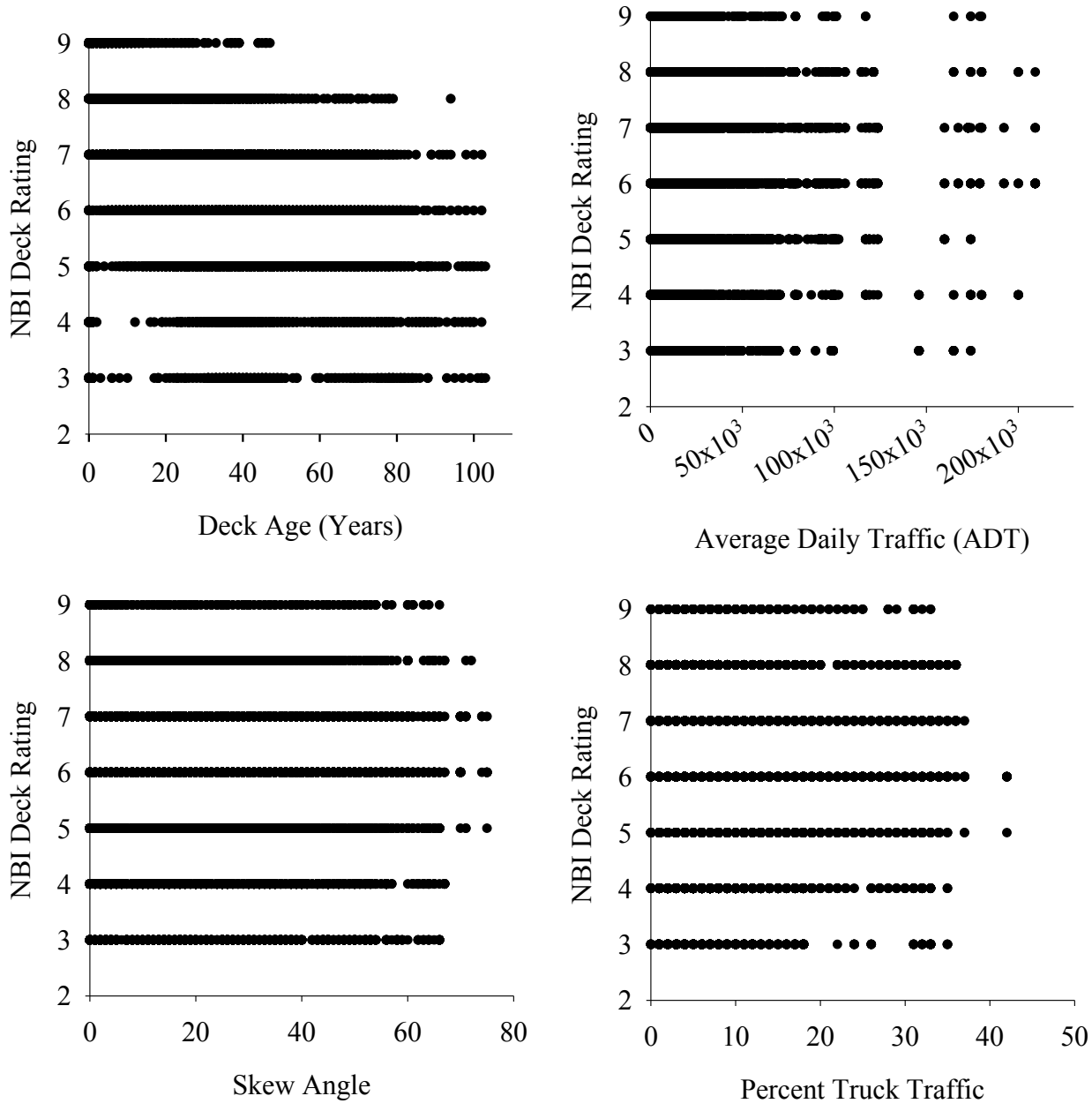


Figure 3-4: NBI database scatter

3.4 Database Preparation

The sever scatter, data imbalance, and limited information provided from individual bridges make the NBI database difficult to use. In addition, the database is littered with erroneous data, and multiple parameters contain outliers. Refinement of the data is necessary in order to streamline the database and minimize any potentially negative effects caused by inherent

difficulties. The database refinement was carried out using multiple different criteria as discussed in the following sections.

3.4.1 Initial Preparation

As the training process of the ANN models requires a bridge condition rating, any records that are missing the overall, deck bottom, or deck surface rating, NBI items 58, 58a, and 58b, respectively, were removed from the database. Initial database screening also included removing records belonging to structure types outside the study's scope (i.e., trunk-line vehicular highway bridges) such as culverts, pedestrian bridges, and railway bridges. Structure types were first identified by using NBI item 8. This item is the structure number, and in this study may not contain a C, X, P, N or T, which indicate culverts, rail road bridges, pedestrian bridges, non-motorized traffic structures, and tunnels, respectively. Not-applicable structure types can also be indicated by NBI items 58, 58a, 59, 60, 107, 108a, 108b, 108c containing an 'N' in place of a rating or measurement. If a record met one of the two criteria discussed above, it was removed on the basis that it was a non-applicable structure type. This initial refinement reduced the size of the database by 47%. The database was further reduced by 42% to reflect the focus of this research on bridges carrying state trunk traffic.

3.4.2 Additional Refinement Based on Database Distributions

After preliminary refinement, histograms showing the distribution of parameters were created to identify any outliers or under populated categories within parameters. Figure 3-5 shows the distribution of the number of records for different structure types in the preliminarily refined database. The predominance of simple or cantilevered steel, continuous steel, and prestressed concrete structure types is obvious. Based on this predominance, the database was refined to contain only these three structure types.

Similarly, the reinforcement protection was investigated and it was found that 99% of the records are for bridges that have either black steel or epoxy coated steel, as show in Figure 3-6. The database was further refined to contain only these two reinforcement protection categories.

Figure 3-7 shows that for the deck structure type, 99% of the records are for concrete cast-in-place structures and the database was further refined to reflect this majority. This refinement is justified due to the focus of the study on the degradation of concrete bridge decks exclusively.

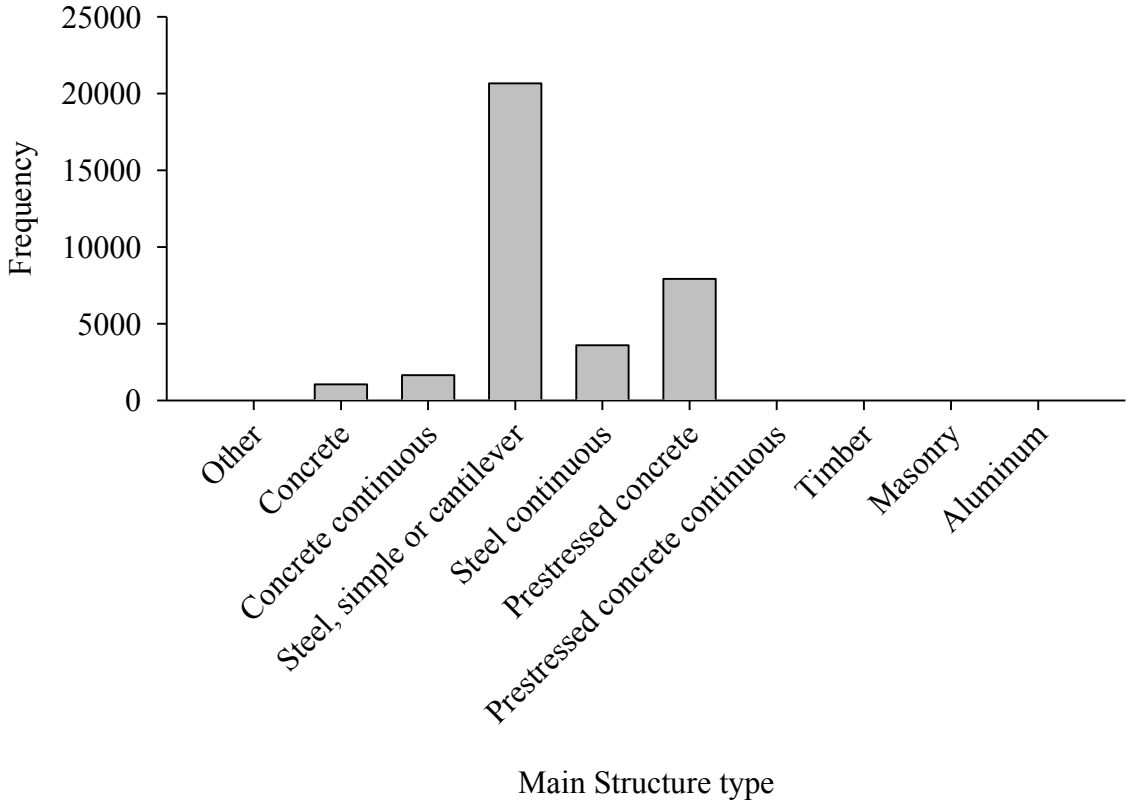


Figure 3-5: Partially refined structure type distribution

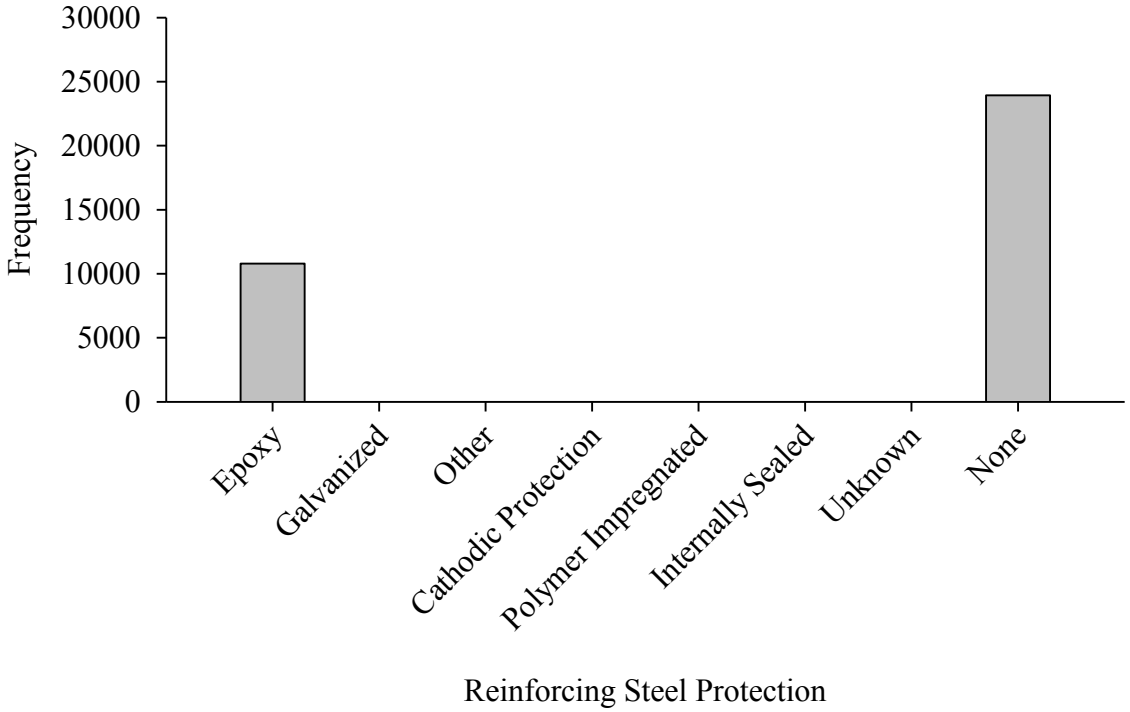


Figure 3-6: Partially refined rebar protection distribution

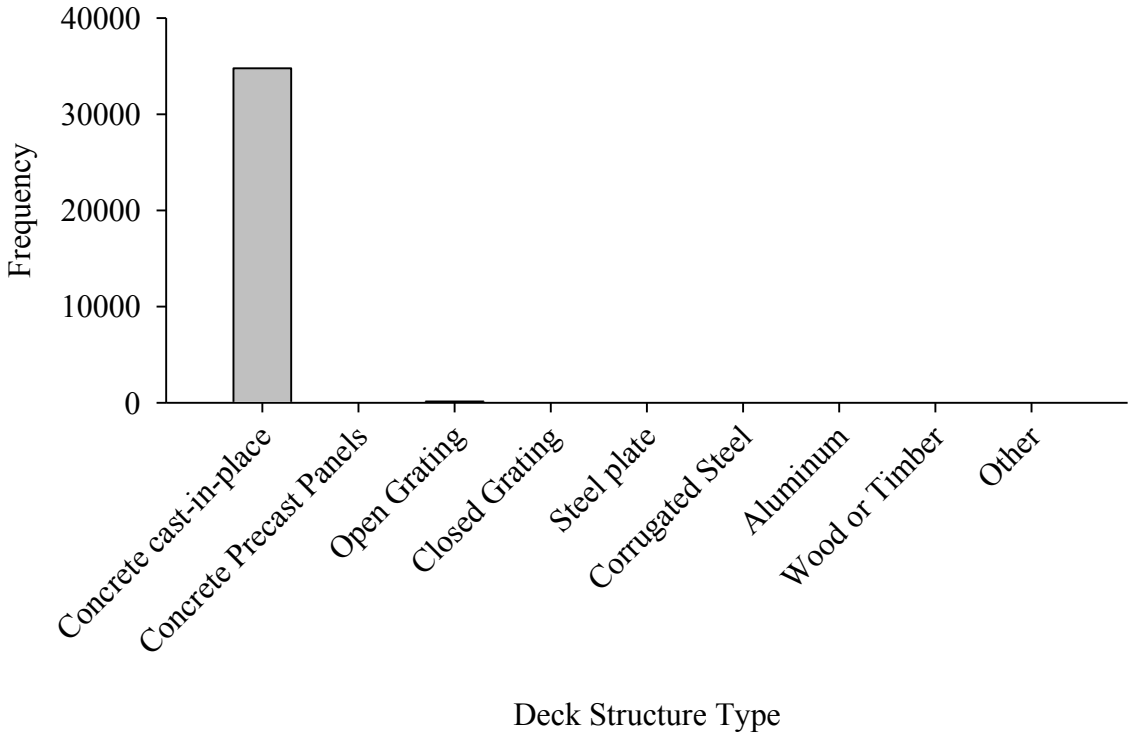


Figure 3-7: Partially refined deck structure type distribution

3.4.3 Missing and Erroneous Data

After this initial refinement, records containing missing data were removed. To be kept in the database, the records had to contain the following parameters: length, width, skew angle, ADT, structure type, year built, and inspection date. The retained information was also checked for errors such as negative ages, zero or negative ADT, no truck traffic percentage, and skews of 90° or 99°. Records containing these errors were removed from the database.

3.4.4 Low Deck Ratings

When a bridge deck reaches a rating of two, restrictions on the loading must be implemented. The damage is so severe that preventative reconstruction practices are no longer applicable. Because these bridges are not functioning at normal operation levels, they were also removed from the database. The records removed accounted for less than 1% of the available data.

3.4.5 Reconstruction or Repair Efforts

The main focus of this research was to develop methods to model the degradation of bridge decks without considering reconstructive or repair measures. Thus, to minimize the effect that repair efforts on the training of the ANNs for rating predictions, inspection records occurring after any recorded reconstruction for the bridge in question were removed. Although the majority of bridge decks that experienced repair were removed using this constraint, it was observed that some bridges still experienced an improvement of condition with an increase in age. It was speculated that this increase in condition was caused by unrecorded repairs. Bridges that experienced an increase in condition rating by two or more inspection points as the age simultaneously increased were identified and examined on an individual basis. An increase of two or more inspection points was selected to account for the potential subjectivity in ratings and was observed to happen both suddenly and gradually with time. Table 3-2 and Table 3-3 provide examples of such records.

Table 3-2: Gradual Increase in deck rating

Deck Rating	Bridge ID	Date of Inspection	Age at Inspection
3	825180801056B01	7/17/1995	48
4	825180801056B01	8/6/1997	50
4	825180801056B01	7/24/1998	51
4	825180801056B01	5/5/1999	52
4	825180801056B01	8/23/2000	53
4	825180801056B01	6/7/2001	54
5	825180801056B01	5/19/2003	56

Shown in Table 3-2 is a gradual increase in the deck rating. The deck goes from a rating of 3 in 1995, to a rating of 5 in 2003. No reconstruction was recorded for this particular bridge. This inconsistency indicates that some form of reconstruction was done on the bridge was done and not recorded, or there may have been a human error in entering the inspection information. Bridges containing this type of gradual increase in rating were removed from the database because without an obvious jump in rating, it is impossible to tell if and/or when reconstruction

or modifications were made to the bridge deck. Therefore, all inspections records for the 18 bridges with this type of increase in rating were removed.

Table 3-3 shows an example for a sharp increase in the deck rating. The deck increased from a rating of 4 in 1995 to a rating of 7 in 1997. As with the bridge in Table 3-2, either reconstruction was performed on the bridge and not recorded, or there was a human error in entering the data.

Table 3-3: Sharp Increase in deck rating

Deck Rating	Bridge ID	Date of Inspection	Age at Inspection
4	03103111000B022	7/28/1995	35
7	03103111000B022	11/11/1997	37

Three hundred and ninety four (394) bridges contained a sudden increase in their respective ratings and were examined on an individual basis to determine which, if any, inspection records should be removed. The bridges experiencing a sudden increase in rating were examined on an individual basis because a sudden increase in rating occurs in several different manners. One type of increase experienced was a jump in records followed by inspection records with similar ratings. For example, a rating may increase from a 4 to a 7 and then stay at ratings of 7 or 6. For these types of records, all inspection records occurring after the sudden increase were removed. Another type of increase is an isolated jump in deck rating. For example, a deck rating may go from a 5 to an 8, but then go back down to 5 at the next inspection. It seems as though the 8 may have been a human error in data entry, and therefore this single record was removed. The third type of jump experienced is actually a dip in the rating. Records may go from a 6 down to a 4, only to jump back up to a 6. As with an isolated increase up, it is possible that this dip is due to human error in data entry also. In this case only the low rating was removed.

Although most data entries affected by reconstruction efforts were identified through the record of reconstruction in the database or by an increase in rating as previously discussed, it has also been noted that minor routine repairs such as patching occur on a regular basis but are not recorded in the NBI database. Because patching may increase the condition rating by only one

point, not all inspection records affected by patching were identified, and thus the database may still contain some records of bridges that have received repair efforts.

3.4.6 Old Bridges with High Ratings

Upon inspection of records in the database, it was also discovered that periodically there were older bridges (30+ years) with very good deck ratings. This poses a problem, as a bridge deck normally lasts only approximately 30 years. Further, there are no reconstruction or overlay years recorded for these decks and the limited years of inspection records provide limited data to detect abnormal behaviors such as those previously discussed. To decrease the confusion that these ratings may cause in the training of the neural network the inspection records were *removed* if they met the following criteria, as the ratings did not accurately reflect the known typical deterioration for a bridge deck:

- If the age ≥ 30 and the deck rating ≥ 6
- If the age ≥ 25 and the deck rating ≥ 7
- If the age ≥ 20 and the deck rating ≥ 8
- If the age ≥ 15 and the deck rating ≥ 9

3.4.7 Young Bridges with Low Ratings

Another type of inconsistency identified was very low inspection ratings for very young bridges, which accounted for 65 inspection records. A frequency analysis was completed on these records, and it was found that 70% of these inspection records belonged to bridges that contained only one span. These one span bridges may be jointless bridges, where expansion joints over the piers are not included in the design, leading to a continuous bridge deck. However, it is acknowledged that these types of bridges can experience cracking very soon after construction (Stringer and Burgueño, 2012). Due to the small number of records with this type of error it was elected to remove inspection records for bridges that had an age of five years and less and had a rating of five or less.

3.5 Summary

The NBI database is a complex compilation of bridge parameters and condition inspection records. The database is a rich source of realistic and historic condition assessments of bridges across the nation. The utilization of these records and the other parameters in the database provide a unique set to develop prediction models. For the current study, the database contains unusable information and some erroneous data. The refinement of the database as discussed in the previous sections results in a comprehensive set to be used for training of neural network models. After data preparation there were 1,956 bridges and 10,034 inspection records included in the database. Nonetheless, complexities inherent to the database add to the level of difficulty in developing robust models using this data.

4 MLP, ENN, AND MARKOV MODEL DEVELOPMENT

4.1 General

This chapter chronicles the development of bridge deck degradation models using artificial neural networks (ANNs) and their performance comparison with a probability-based approach, namely a Markov chain method. The development of the ANN models includes an in depth look at how the input parameters were selected, the selection of the final network architecture, and a brief discussion on the computer codes utilized. The classification performance of the ANN models was examined using confusion matrices, bubble plots, and Press's Q statistic. A study examining the connection weights of the ANN models provides insight on the relative importance of the inputs on deck condition ratings. The Markov model was developed in order to compare the proposed ANN method with a currently accepted form of condition rating prediction used by the Michigan Department of Transportation. The Markov models are produced using the same database as the ANN models, and a discussion of their development is provided.

The single multi-layer perceptron (MLP) model and one based on ensembles of neural networks (ENN) utilized the condition ratings determined via manual inspections conducted on a biannual basis. The condition ratings are the target output for the ANN models. The inputs used to predict the condition ratings were the geometric and operational parameters of the bridge. This allows bridge engineers to utilize known information of the bridge to predict its condition rating. The process by which geometric and operational parameters were selected for input parameters is presented in Section 4.2.1.

4.2 Multi-Layer Perceptron Models

As discussed in depth in Section 2.3.2, the multi-layer perceptron (MLP) is a simple type of neural network model with feed-forward architecture. This chapter chronicles the development of MLP networks to predict the deck condition ratings based on physical and operational characteristics of the bridge. A detailed review of the data selection, architecture selection, and

network performance are presented. MLP models were developed to predict the overall deck condition, deck bottom condition, and deck surface condition rating.

4.2.1 Input Parameter Selection

The bridge parameters in the National Bridge Inventory (NBI) database were used as inputs to the neural network model, while the output was the condition rating (integer value from 3 to 9) of the bridge deck. The selection of input variables is vital to the success of a neural network, and thus only parameters that have an effect on the output should be selected. Therefore, statistical analyzes in the form of correlation analysis and chi-squared hypothesis testing were conducted to identify statistical relationships between database parameters and the deck condition rating. However, because a neural network can model complex nonlinear relationships between variables, simple statistical tests may not identify all pertinent input-output relationships. Thus, additional parameters were selected through engineering judgment and their effect on the ANN models was tested.

4.2.1.1 Correlation Analysis

Correlation analysis draws inferences about the strength of the linear relationship between two or more variables. It provides a quantitative index of the degree to which one or more variables can be used to predict the values of another variable. The correlation coefficient is a value in the range of -1 to 1. The closer the value is to either -1 or 1, the stronger the correlation between the variables being investigated. In this research, the Pearson correlation coefficient was selected to determine correlation between various database parameters and the deck rating. If two variables have a correlation coefficient value less than the absolute value of 0.3, they are considered to be statistically independent, and thus have no linear contribution to each other (Ayyub and McCuen, 2003). Correlation analyses were completed only using the continuous variables in the database, and the results of the tests for the overall, surface, and bottom ratings are shown in Table 4-1.

Table 4-1 shows strong correlations between the deck condition and the deck age and year the bridge was built. These two parameters were the only ones identified to be statistically dependent with the condition rating. The table shows very low correlation coefficient values for physical parameters such as the length, skew angle, and deck width.

Table 4-1: Correlation coefficients of continuous parameters

		Age	Year Built	Maximum Span Length	Number of Spans	Number of Lanes	Length
Pearson Correlation Coefficient	Overall Deck	-0.75	0.72	0.22	-0.11	-0.10	0.09
	Deck Surface	-0.60	0.58	0.17	-0.10	-0.07	0.05
	Deck Bottom	-0.75	0.73	0.22	-0.10	-0.10	0.09
		ADT	Latitude	Truck %	Width	Approach Road Width	Skew Angle
Pearson Correlation Coefficient	Overall Deck	-0.05	0.03	0.12	-0.03	-0.01	-0.08
	Deck Surface	-0.05	-0.01	0.08	0.00	-0.02	-0.09
	Deck Bottom	-0.06	0.02	0.12	-0.02	-0.01	-0.08

4.2.1.2 Chi-Squared Analysis

Chi-squared hypothesis testing was employed to identify associations between discrete parameters and the rating. The null hypothesis of *no association between the deck rating and tested parameter* was tested for all discrete parameters in the database. Valid chi-squared tests with low probabilities of a Type I error (rejecting the null hypothesis when it is in fact true) indicate an association between the parameter in question and the deck rating. Table 4-2 shows the results from the chi-squared tests. The chi-squared value and the probability of a Type I error are listed for each parameter tested. The * symbol indicates that the test may not be valid due to low expected cell counts. For these parameters, the percentage of cells with low expected counts is listed. As seen in Table 4-2, all tests yielded large chi-squared values and low Type I error probabilities. However, only the tests on structure type, rebar protection, and region parameters yielded valid results. Parameters that yielded invalid test results could have been reduced to potentially yield valid test results, but further reduction of the database is undesirable as it is already limited in size and excess data manipulation should be avoided to ensure accurate results.

Table 4-2: Chi-squared results

		χ^2 Value	Type I Error Prob	*Low cell count (%)
Overall Deck Rating	Deck Wearing Surface*	1182	<.0001	20
	Deck Membrane*	252	<.0001	52
	Rebar Protection	1562	<.0001	-
	Region	1125	<.0001	-
	Design Load*	2442	<.0001	29
	Approach Surface Type*	244	<.0001	33
	Structure Material (43a)	2473	<.0001	-
	Design Type*	1555	<.0001	70
Deck Surface Rating	Deck Wearing Surface*	667	<.0001	22
	Deck Membrane*	138	<.0001	64
	Rebar Protection	966	<.0001	-
	Region	517	<.0001	-
	Design Load*	1375	<.0001	29
	Approach Surface Type*	288	<.0001	29
	Structure Material (43a)	1245	<.0001	-
	Design Type*	527	<.0001	70
Deck Bottom Rating	Deck Wearing Surface*	1042	<.0001	22
	Deck Membrane*	206	<.0001	62
	Rebar Protection	1528	<.0001	-
	Region	1065	<.0001	-
	Design Load*	2128	<.0001	29
	Approach Surface Type*	447	<.0001	30
	Structure Material (43a)	2128	<.0001	-
	Design Type*	1291	<.0001	70

4.2.1.3 Knowledge Based and Trial and Error

The parameters of age, year built, rebar protection, region, and structure type were identified to be statistically significant to the overall deck, deck surface, and deck bottom ratings. An initial MLP model was developed using only these five input parameters. The performance of this network served as a baseline when examining the effects of additional parameters on the network performance. The first set of additional parameters to be tested consisted of parameters from the correlation and chi-squared analyses that were just outside the range of being

statistically dependent. These included the maximum span length, number of spans, deck wearing surface, design load, and approach surface type. A second set of additional parameters was defined using engineering judgment and included the parameters of skew angle, ADT, percent truck traffic, average daily truck traffic (ADTT), deck width, and deck surface membrane. The influence of additional parameters was tested by training an MLP model with the statistically significant parameters and the additional parameter. The performance of the newly trained network was then compared against the baseline model and any differences in predictive performance were noted. If an MLP network with an additional parameter exhibited superior predictive ability that parameter was selected to join the statistically dependent parameters as inputs.

4.2.1.4 Final Input Variable Selection

For the overall deck rating and deck bottom surface rating MLP models, eleven parameters were selected as input variables: *age, year built, average daily traffic, percent truck traffic, average daily truck traffic, number of spans, region, steel reinforcement protection, structure type, design load,* and *approach surface type.* For the deck wearing surface rating MLP models, twelve parameters were selected as input variables: *age, year built, average daily traffic (ADT), average daily truck traffic (ADTT), structure type, maximum span length, skew angle, region, steel reinforcement protection, deck surface membrane type, design load,* and *number of spans.*

4.2.2 Network Architecture and Implementation

After selection of the input variables, the size of the input and output layers is known. However, multiple other dimensions of the network architecture remain to be determined. The network architecture can have a serious influence on the predictive performance of the network and requires careful consideration. Additional architectural parameters that are to be defined include the number of hidden layers in the network, the number of nodes on the hidden layer(s), the number of training epochs, the activation function, learning function, and training function. The following sections detail the selection of these parameters and give a brief description of the codes developed using the Matlab Neural Network Toolbox (MathWorks, 2010), which was utilized for all the developed ANN models.

4.2.2.1 Input and Output Layers

The number of inputs dictates the size, i.e., the number of nodes, on the input layer of the MLP model. The number of nodes is also dependent on the structure of the input parameters. For example, quantitative variables, which are continuous in nature such as the age, year built, average daily traffic, etc., can be represented using only one node. However, due to their discrete nature, multiple nodes must be employed to represent a qualitative variable. In this research, quantitative variables were standardized using Equation 4-1 where z is the standardized value, $\overline{x}$ is the non-standardized value, μ is the mean value, and σ is the standard deviation.

$$z = \frac{\overline{x} - \mu}{\sigma} \qquad \textbf{4-1}$$

To represent qualitative variables a binary system was utilized. For example, in the database the structure type is represented by a 3 for simple or cantilevered steel, a 4 for continuous steel, or a 5 for prestressed concrete. For each bridge, this single input was changed to a vector with three elements, one representing a 3, one a 4, and the other a 5. A '1' was recorded in the binary vector for the element corresponding to the qualitative variable in the database. For example, if a bridge was coded as a 4 in the database, the input into the network would read [0 1 0]. Due to the differing data structures, the number of nodes in the input layer is not the same as the number of input parameters. As previously stated, the number of input parameters for the deck overall, bottom, and surface ratings were thirteen, thirteen, and fourteen, respectively. Each discrete input parameter has a unique number of categories, and each category corresponds to one node on the input layer. Because a different number of inputs were selected for the different MLP models, it would be expected that the models have a different number of nodes on the input layer. However, due to the difference in the size of the discrete parameters selected, each MLP model has, by coincidence, 33 nodes on the input layer.

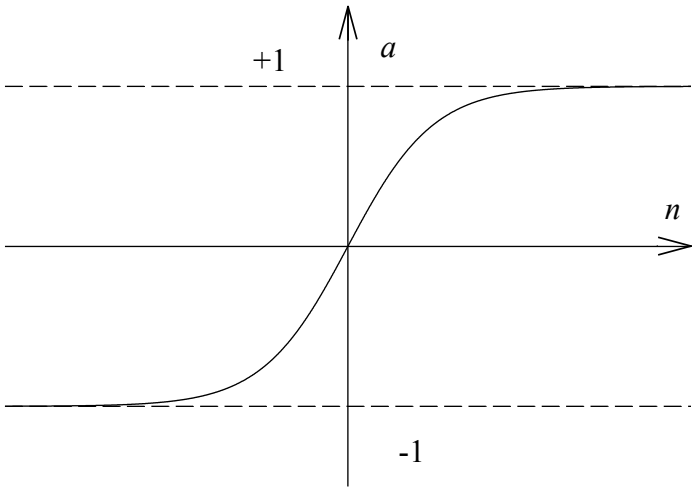
The output for the ANN models was the respective bridge deck condition ratings. Because of the integer scale of the rating system, the output is a discrete variable. The output must be formatted in the same binary manner as the discrete input variables. The output layer of the network thus had seven nodes with each node representing an individual condition rating. The condition ratings were organized in ascending order with the first and last nodes corresponding to a rating of '3' and '9', respectively.

4.2.2.2 Matlab Coding

The Matlab Neural Network Toolbox (MathWorks, 2010) was used to develop the MLP models. The codes written by Li (2008) were modified for the current task. The programs were developed to perform four main tasks: separate the data into two sets (training and testing), create an initial neural network based on architecture specifications, train the network, and test the network. Roughly 2/3 of the data was randomly selected without replacement to train the MLP network and the remaining 1/3 was reserved for network testing. The function 'newff' creates a new feed-forward back propagation MLP neural network with input and output layers that are representative of the data. Using 'newff', additional values for the following architectural parameters were assigned: number of hidden layers, size of the hidden layers, the activation function, backpropagation training function, and the backpropagation learning function.

The number and size of the hidden layers was sequentially modified in order to identify the best performing architecture for the task. A description of this process is offered in Section 4.2.2.3. The hyperbolic tangent sigmoid function, specified as 'tansig' was selected as the activation function for the MLP networks. The hyperbolic tangent sigmoid function is commonly used for continuous activation functions (Rumelhart et al., 1986) . The activation function, as shown in Figure 4-1, transforms the node inputs, n, to node outputs, a, using the 'tansig' function.

The 'tansig' function gives output values between -1 and 1. In Matlab, although 'tansig' is mathematically equivalent to the hyperbolic tangent function 'tanh', 'tansig' is recommended for use in neural networks as it performs faster in the Matlab platform.



a=tansig(n)

Figure 4-1: Tansing activation function

The function 'traingdx' was selected as the training function. This function is responsible for training the neural network based on a series of parameters including the momentum and learning rate. Values for all necessary parameters are specified before network training. The learning function is responsible for calculating the changes in the weight and bias values. The default value of 'learngdm' was selected for the backpropagation weight and bias learning function.

Training of the neural network ensues after the neural network structure, training parameters, and weights and biases were initialized. It is important to note that 'newff' assigns random connection weights and biases for the initialized network. Network training is implemented using the 'train' function, which takes the developed network, training input data, and training output data as arguments. Upon completion of training, the 'train' function returns the trained neural network (optimized connector weights and biases) and the network error. The training process is stopped when one of the following occurs:

- The maximum number of epochs is reached.
- The maximum amount of time is exceeded.
- The performance goal is minimized.
- The performance gradient falls below the minimum specified value.

After training was completed, the performance of the network was measured by examining the ability of the network to correctly predict the outputs for the set of testing data. The testing data is a set of inspection records that was not included in the training set. The trained network is used to predict the condition rating of the inspection rating given the input parameters. Confusion matrices were used to evaluate the performance of the trained network. A detailed discussion on confusion matrices can be found in Section 4.2.3.1.

4.2.2.3 Determination of Final Architecture

While the architecture of a neural network can have a significant impact on its performance, there is no standard method for determining the networks optimal architecture. For this reason, different networks with various architectures were developed in order to identify the best performing neural network for the problem. The performance of the different networks were evaluated after systematic changes to the number of hidden layers, the number of nodes per hidden layer, and the number of training epochs were implemented.

Table 4-3: MLP tested architectures

Epochs	Neurons per Layer	Hidden Layers
2,000	30	1
		2
5,000	30	1
		2
	60	1
		2
10,000	60	1
		2
20,000	60	2

Table 4-3 shows the nine different network architectures tested. To begin, a network with one hidden layer, 30 neurons per layer, and 2,000 epochs was created. The network was then updated to contain two hidden layers, but the same number of epochs and nodes per layer. The next step was to examine the effect of increasing the number of epochs. The number of epochs

was increased to 5,000 for networks with 30 hidden neurons and both one and two hidden layers. The number of neurons per hidden layer was increased from 30 to 60 for the networks with 5,000 epochs and one and two hidden layers. Additional increases in the number of epochs were done. The number of hidden layers for the networks with 10,000 epochs was adopted from the best performing network with 5,000 epochs. Similarly, the number of hidden layers for the 20,000 epoch network was taken from the best performing network with 10,000 epochs. The MLP network that produced the best performance contained two hidden layers with each hidden layer containing 60 neurons. Training with 20,000 epochs produced the network with the best overall performance. The performance of the MLP networks with this architecture is presented in section 4.2.3.

4.2.3 Predictive Performance

The performance of the MLP models was assessed using confusion matrices, bubble plots, and statistical measures. Parameters pertaining to the predictive performance, such as the percentage of correct classifications, were calculated directly from the confusion matrices. Bubble plots fitted with a linear regression to provide another measure of predictive performance. Statistical measures, such as Press's Q-statistic can give a measure of the classification ability of the network. The performance parameters provided in the following sections are for the best performing MLP networks only.

4.2.3.1 Confusion Matrices

Confusion matrices are tabular representations of the prediction ability of the MLP (Kohavi and Provost 1998). The rating predictions of the network are compared against the target outputs and are tabulated. Confusion matrices were developed for both the training and testing data sets. Confusion matrices can be developed using the training data to provide a measure of the prediction ability of the network to examples it has already been exposed to. The confusion matrices developed using the testing set provide a measure of the generalization ability of the network to new, unseen data. Table 4-4 shows the confusion matrix for the overall deck rating MLP developed using training data. The columns represent the true value of the inspection rating and the rows represent the predicted rating by the network.

Table 4-4: Overall deck rating training CM

	Manual Inspection							
Prediction	3	4	5	6	7	8	9	SUM
3	422	115	63	3	0	0	0	603
4	110	467	86	6	0	0	0	669
5	134	355	1426	112	9	0	1	2,037
6	39	91	251	1044	213	10	0	1648
7	1	6	27	214	752	105	9	1114
8	0	0	1	13	61	343	155	573
9	0	0	0	0	2	0	1	3
SUM	706	1034	1854	1392	1037	458	166	6,647
CR (%)	59.8	45.2	76.9	75.0	72.5	74.9	0.6	67.0
AR (%)	75.4	90.6	95.1	98.4	98.9	97.8	94.0	93.8
μ	3.7	4.4	5.1	6.1	6.8	7.7	7.9	-
σ	1.0	0.8	0.6	0.5	0.5	0.5	0.3	-

In the confusion matrix, the value in each cell can thought of as be represented by elements C_{ij} which represent the number of times the network predicted the deck rating as i when the true rating as recorded in the database is j. For instance, $C_{8,7}$ corresponds to row 8, column 7 in Table 4-4. This cell has a value of 61, meaning that there are 61 cases where the rating was predicted by the network as an 8, but the actual rating of the inspection was 7. The dark gray cells along the matrix diagonal represent the number of correct predictions by the network. These cells were used to calculate the correct ratio, or CR. The CR is the percentage of predicted values that correctly matched the actual rating, which are indicated by the dark gray cells in the confusion matrices. The CR is calculated for each individual rating, and for the overall network. As seen in Table 4-4, the network had an overall correct ratio of 67.0% when exposed to the training data.

The subjectivity in the manual inspection process is well recognized and a thus margin of error of ±1 ratings is adopted as representative of this subjectivity. This margin of error is not a representation of the standard deviation, but rather a value chosen based on the work of Phares et al. (2001). The ratings within this margin of error, indicated by the light grey cells, are termed 'acceptable' ratings. The acceptable ratio, or AR, is the percentage of ratings that were predicted within this ±1 rating margin of error. As with the CR, the AR was calculated for the overall

network and for the individual ratings. The AR for the deck overall rating MLP when exposed to the training set was 93.8%, indicating that the network performs very well at predicting ratings within a ±1 rating interval.

The mean (μ) predicted rating for each actual rating was also calculated and presented in the confusion matrix. This value provides a measure of whether the network is over or under predicting specific condition ratings. For example, for ratings that are recorded as a 4 in the database, the network prediction mean is a 4.5. This indicates that for ratings recorded as a 4 in the database, the network is, on average, predicting ratings of a higher value. The standard deviation (σ) for the predicted ratings was also calculated, giving a measure of the spread of the predicted ratings.

Table 4-5: Overall deck rating testing CM

	Manual Inspection							
Prediction	3	4	5	6	7	8	9	SUM
3	172	70	35	4	0	0	0	281
4	59	191	64	2	1	0	2	319
5	72	214	692	70	14	0	1	1,063
6	15	56	159	501	128	4	0	863
7	0	2	19	121	319	48	1	510
8	0	0	0	6	49	163	67	285
9	0	0	0	0	0	2	0	2
SUM	318	533	969	704	511	217	71	3,323
CR (%)	54.1	35.8	71.4	71.2	62.4	75.1	0	61.3
AR (%)	72.6	89.1	94.4	98.3	97.1	98.2	94.4	93
μ	3.8	4.5	5.1	6.1	6.8	7.8	7.8	-
σ	1.0	0.9	0.7	0.6	0.7	0.5	0.8	-

Table 4-6 and Table 4-8 are the confusion matrices for the deck bottom surface and deck top surface ratings, respectively. It is interesting to note that the deck bottom surface has very similar performance, in terms of CR and AR, as the overall deck rating. The mean and standard deviations of the networks were also very similar.

As shown from the results presented in Table 4-8, the network performance for the deck surface rating is lower in terms of the CR and AR than the overall deck rating. Also, larger standard deviations indicate larger spread in the predicted values. Because the same process was

used to determine the optimal input parameters and network architecture, the poorer performance of the deck surface rating network suggests that more variability may be present in the deck surface rating dataset. This increase variability produces a more complex problem, which is in turn more difficult to model.

Table 4-6: Deck bottom rating training CM

	Manual Inspection							
Prediction	3	4	5	6	7	8	9	SUM
3	361	105	44	2	1	0	0	513
4	82	399	111	7	1	1	1	602
5	119	277	1144	85	6	1	0	1,632
6	16	78	222	861	176	7	1	1361
7	0	6	19	199	682	100	4	1010
8	0	0	0	13	55	314	130	512
9	0	0	0	0	0	0	0	0
SUM	578	865	1540	1167	921	423	136	5,630
CR (%)	62.5	46.1	74.3	73.8	74.0	74.2	0.0	66.8
AR (%)	76.6	90.3	95.9	98.1	99.1	97.9	95.6	94.2
μ	3.6	4.4	5.0	6.1	6.8	7.7	7.9	-
σ	0.9	0.8	0.6	0.6	0.5	0.5	0.4	-

Table 4-7: Deck bottom rating testing CM

	Manual Inspection							
Prediction	3	4	5	6	7	8	9	SUM
3	158	51	28	1	1	0	0	239
4	61	182	65	7	0	0	0	315
5	68	156	573	46	12	0	0	855
6	20	30	127	392	100	2	0	671
7	0	0	9	110	290	46	6	461
8	0	0	0	8	46	147	71	272
9	0	0	0	0	0	0	0	0
SUM	307	419	802	564	449	195	77	2813
CR (%)	51.5	43.4	71.5	69.5	64.6	75.4	0	61.9
AR (%)	71.3	92.8	95.4	97.2	97.1	99	92.2	93.2
μ	3.8	4.4	5.0	6.1	6.8	7.7	7.9	-
σ	1.0	0.8	0.7	0.6	0.7	0.5	0.3	-

Table 4-8: Deck surface rating training CM

	Manual Inspection							
Prediction	3	4	5	6	7	8	9	SUM
3	158	45	41	22	11	4	1	282
4	90	434	134	86	33	26	16	819
5	122	136	466	123	44	20	10	921
6	66	169	197	794	190	57	14	1487
7	36	37	65	260	966	198	11	1573
8	1	1	3	22	73	294	120	514
9	0	0	0	0	0	0	0	0
SUM	473	822	906	1307	1317	599	172	5,596
CR (%)	33.4	52.8	51.4	60.7	73.3	49.1	0.0	55.6
AR (%)	52.4	74.8	88.0	90.1	93.3	82.1	69.8	83.6
μ	4.4	4.7	5.1	6.0	6.7	7.2	7.2	-
σ	1.3	1.0	0.9	0.9	0.8	1.1	1.4	-

Table 4-9: Deck surface rating testing CM

	Manual Inspection							
Prediction	3	4	5	6	7	8	9	SUM
3	71	21	17	10	3	4	1	127
4	49	163	94	51	20	17	6	400
5	65	106	200	78	27	12	3	491
6	43	95	131	311	115	24	6	725
7	11	14	38	155	463	110	9	800
8	0	0	1	7	44	141	61	254
9	0	0	0	0	0	0	0	0
SUM	239	399	481	612	672	308	86	2,797
CR (%)	29.7	40.9	41.6	50.8	68.9	45.8	0	48.2
AR (%)	50.2	72.7	88.4	88.9	92.6	81.5	70.9	82.7
μ	4.5	4.8	5.2	5.9	6.7	7.1	7.3	-
σ	1.2	1.0	1.0	0.9	0.8	1.2	1.3	-

The confusion matrices can also be used to provide statistics about the ability of the network to correctly identify damage, and how often the network incorrectly identifies damage. The damage identification ratio (DIR) is defined as the proportion of ratings whose actual value is less than or equal to a 4 and was correctly predicted within the ± 1 confidence band to be less than or equal to a 4. The ability of the network to predict ratings of damaged decks is important

to the success of the network if to be used a damage prediction tool. Although correctly identifying any type of rating is important, the scope of the problem makes correctly identifying damage ratings even more important. Because repairs often occur once the deck has reached, or is close to reaching, a condition rating of 4, the ability to accurately predict ratings in this neighborhood is extremely important to ensure proper allocation of maintenance funds. The false identification ratio (FIR) is the ratio of 'good structures' identified as damaged to all 'good structures' (Burgueño and Li 2008). Essentially, the FIR is the proportion of bridges that are falsely identified as damaged when they are in fact in good condition. Because failing to identify damaged structures has more serious implications than identifying a good structure as damaged, the DIR should be kept as high as possible without significant increases in the FIR. Hypothetically, the FIR should not exceed a set value because a high FIR can lead to misallocated repair funds. However, maintenance and repair decisions would ultimately need field verification from maintenance engineers. The DIR and FIR values for each of the developed networks are listed in Table 4-10. The table includes the performance parameters from both training and testing. The moderately high DIR values and respectively low FIR values of the trained networks are promising and indicate the ability of the networks to be used to successfully identify damage.

Confusion matrices can also be developed for data that has not used during the training of the neural networks. Such data is termed "testing" data and the confusion matrices for the deck overall rating, deck bottom rating, and deck surface rating are shown in Table 4-5, Table 4-7, and Table 4-9, respectively. Comparison of these confusion matrices with those based on the training data shows that the performance of the models reduces only slightly. This is to be expected since the MLP models were not trained on this data. Table 4-10 also provides a direct comparison of the predictive performance of the networks during training and testing. As seen in the table, all of the MLP networks perform better (with respect to all performance measures) when presented with the training data rather than a new and previously unseen data set. The most dramatic reduction in performance is in the DIR of the deck surface MLP. When presented with unseen testing data the DIR of the network drops by nearly 10%. Again, this drop in performance is normal. Overall, the small variation in performance (correct ratio, acceptable ratio, mean value and standard deviation) indicates that both the generalization ability of the neural networks and that the models were not overtrained.

Table 4-10: Performance summary of MLP networks

		DIR (%)	FIR (%)	CR (%)	AR (%)
Overall	Training	64.0	0.3	67.0	93.8
	Testing	57.8	0.6	61.3	93.0
Surface	Training	56.1	5.9	55.6	83.6
	Testing	47.7	6.7	48.2	82.7
Bottom	Training	65.6	0.5	66.8	94.2
	Testing	62.3	0.7	61.9	93.2

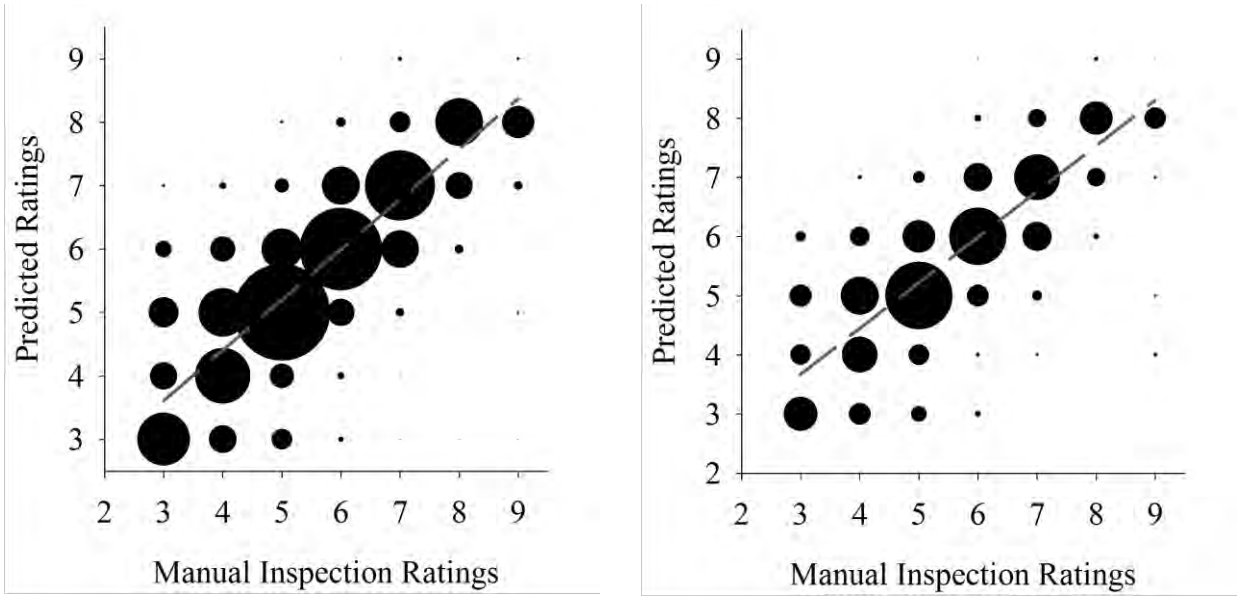
As discussed above, confusion matrices provide valuable information on the classification ability of the networks. Although the confusion matrices show that the MLP networks have satisfactory performance in terms of the AR, CR, FIR, and DIR, a more rigorous assessment of the classification and prediction power was provided through bubble plots and Press's Q statistic, which are discussed in the following sections.

4.2.3.2 Bubble Plots

The bubble plots shown in Figure 4-2 through Figure 4-4 provide a visual representation of the confusion matrices presented in Section 4.2.3.1. Each figure offers a comparison of each of the networks training and testing performance. In the plots, the diameter of the circle is a representation the number of ratings at each point. The same scaling factor was used to create the plots from the training and testing confusion matrices for each network. Because the testing set is half the size of the training set, the bubble sizes of the testing plots are smaller. The larger dots along the diagonal of the plots indicate a higher number of occurrences. The dots off the diagonal are smaller in size, indicating a much smaller number of occurrences.

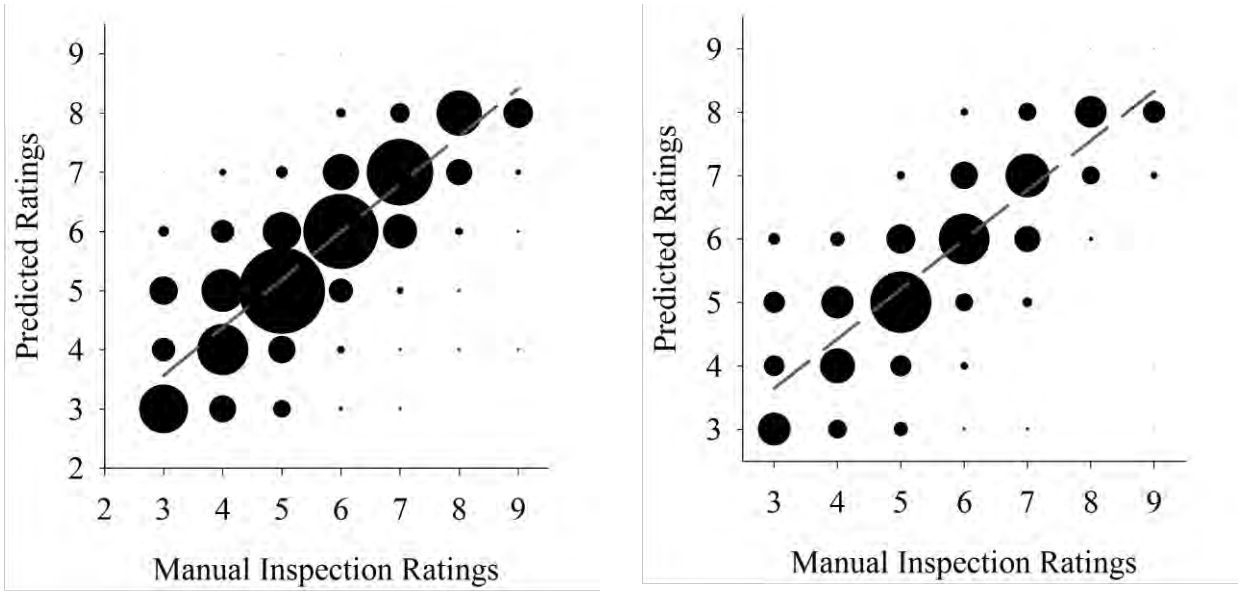
The training and testing data bubble plots show very similar behavior. As a side effect of the different sample sizes, the bubble sizes along the diagonal, representing accurate predictions, on the training plots are larger than those on the testing plots. However, the bubble sizes on the outskirts of the plots are similar in size. Taking into account the population sizes, this shows that when exposed to the training data, the proportion of incorrect predictions away from the true value decreases. This observation is also supported by smaller standard deviation values in the training confusion matrices.

Linear regression conducted on the plots provided additional information on the classification accuracy of the neural networks. The regression lines can be seen in the bubble plots and the equations and respective R^2 values are listed in Table 4-11. The standard error of estimation (SEE) for each linear regression is also listed in Table 4-11.



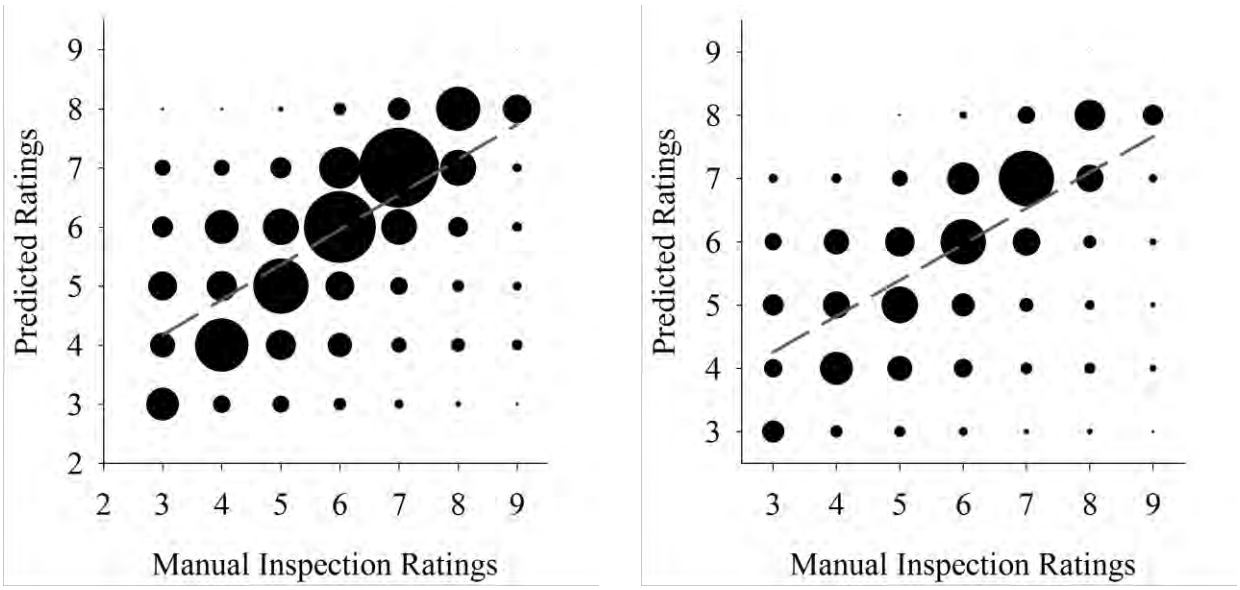
(a) (b)

Figure 4-2: Overall deck rating bubble plots: (a) training, (b) testing



(a) (b)

Figure 4-3: Deck bottom rating bubble plots: (a) training, (b) testing



(a) (b)

Figure 4-4: Deck surface rating bubble plots: (a) training, (b) testing

Table 4-11: Regression equations and parameters

		Equation	R^2	SEE
Overall Rating	Training	$y = 0.7931x + 1.2312$	0.756	0.56
	Testing	$y = 0.7716x + 1.3568$	0.704	0.65
Bottom Rating	Training	$y = 0.8073x + 1.1462$	0.766	0.54
	Testing	$y = 0.7803x + 1.3063$	0.731	0.63
Surface Rating	Training	$y = 0.5943x + 2.3899$	0.467	1.37
	Testing	$y = 0.5677x + 2.555$	0.44	1.46

The linear regression analysis shows that the overall and bottom rating neural networks are very similar in performance. A slope of 1.0 would indicate *perfect* classification by the neural network. The slopes of the overall and bottom rating regression lines demonstrate that these MLP models are *good* at providing correct classifications but they are not *perfect*. The slope of the regression line for the deck surface rating is smaller, and emphasizes the *poorer* classification ability of the network. As expected, the slopes of the regression lines on the training plots are closer to 1 than those of the testing plots for all MLP models.

The R^2 value provides a measure of the fraction of total variance of *y* that is explained by the variation in *x*. The value also represents the ability of a model to predict the output *y* based on the input *x*. Values closer to 1 indicate a strong ability to predict *y* given *x*. In terms of the MLP models the *x* value is not given as the actual manual inspection rating directly but is given as the set of input parameters that correspond to the *x* value. The R^2 for the overall and deck bottom rating models indicate good prediction ability. The deck surface rating MLP has an R^2 value of 0.44 which indicates a poorer ability to predict the condition rating. When exposed to the training data set the MLP models had higher R^2 values indicating better predictive performance, which as described earlier is expected. While the R^2 values produced when the MLP networks were exposed to unseen testing data are less than those of the training set the values do not differ significantly, strengthening the argument that the models are able to generalize and are not overtrained.

The standard error of the estimation (SEE) is the standard deviation of the data about the regression line. The smaller the SEE value is, the stronger the linear relationship between the variables. The SEE values for all regressions are listed in Table 4-11. As seen in the table, the SEE values are smallest for the training plots; and, as expected, the overall deck and deck bottom plots yielded smaller SEE values than the deck surface rating.

4.2.3.3 Press's Q Statistic

Press's Q statistic provides a measure of the discriminatory power of the classification process. Specifically, it tests if the method can classify statistically better than chance. If the Q statistic is larger than a critical value of 6.63 (from a Chi-squared distribution with one degree of freedom), it can be said that the method classifies significantly better than just chance (Hair et al., 1998). However, the Q statistic is affected by the sample size and larger samples are more likely to show significance than small sample sizes with the same correct classification ratio. The following equation is used to calculate the Q statistic

$$\text{Press's Q statistic} = \frac{[N-(nK)]^2}{N(K-1)} \qquad \textbf{4-2}$$

where *N* is the total sample size, *n* is the number of observations correctly classified, and *K* is the number of groups.

Table 4-12 gives the values of Press's Q statistic calculated using the classifications of the MLP networks when presented to the testing set. As seen from the table, all networks show a Q statistic much larger than the critical value of 6.63 which indicates that the networks are able to correctly classify significantly better than chance alone.

Table 4-12: Press's Q statistic

ANN	Press's Q Statistic
Overall MLP	6006
Surface MLP	2632
Bottom MLP	5214

As expected, the Q statistic is smaller for the deck surface rating MLP model. However, although the deck surface MLPs CR is only 48%, the Q statistic indicates that it is able to classify better than chance alone.

4.2.3.4 Discussion on Performance of MLP Models

The confusion matrices, bubble plots with linear regressions, and Press's Q statistic all offer different information regarding the performance of the developed MLP models. As expected, the performance of the models to unseen testing data was worse than the performance of the models when tested with training data. This shows that the models are able to generalize and can be used to successfully predict deck conditions of bridges that were not included in the training set. Although the MLP models provide promising results, they are relatively simple models and there is significant room of improvement. In particular, the CR values for individual ratings are particularly poor. This is particularly evident for ratings with a low number of inspection records, such as 3, 4, and for the highest rating level of 9. Low CR values were obtained for both the training and testing matrices, indicating that the MLP models are not learning these connections well. The reason for the poorer performance is likely due to the severe data imbalance in the database, as presented in Section 3.3.2. This imbalance provides less of an opportunity to learn the connections between the inputs and the rating. The performance of the MLP models is quite remarkable considering the complexity of the problem. In addition to the classification parameters, the ability of the MLP models to predict with a variance comparable to

visual inspectors, documented by Phares et al., (2001), illustrates the effectiveness of the models for predicting the condition ratings. The MLP models provide a method with comparable variance to human inspectors with the added benefit of being able to predict *future* conditions with the same variance.

4.3 Models with Ensembles of Neural Networks

Although the MLP models were able to provide satisfactory predictive performance, as reviewed in the literature, using an ensemble of neural networks has can produce an increase in performance. Figure 2-3 shows a schematic of a neural network ensemble. Ensembles of neural networks are successful due to the 'divide and conquer' approach implemented. A model for the overall condition rating was not developed using ENNs because it is not used in assessments for repair and maintenance by MDOT. Exclusion of an ENN overall condition rating model was also justified through an investigation that revealed that the overall and deck bottom condition ratings at each inspection were almost always identical. The same input parameters used in the MLP models as per Section 4.2.1 were used in the development of the ENN models in order to compare the MLP and ENN models. The following sections outline the development and performance of the proposed ENN models.

4.3.1 Additional Database Refinement

Upon completion of the MLP models, it was discovered that the database contained significant errors with respect to the age of bridge decks with epoxy-coated reinforcement (ECR). An internal MDOT report by Boatman (2010) cites a 1980 mandate requiring the inclusion of ECR in all newly constructed bridge decks as of December 1980. However, as the following figures illustrate, the current database shows considerable deviation from this mandate by containing numerous bridges constructed with ECR before 1980. Rebar protection is an input parameter to the ANN models, and miscoding of this input can have serious implications on the predictions of the models.

Figure 4-5 shows the distribution of unique bridges in the database with respect to the year in which the bridges were built. The number of bridge decks built with black steel is indicated by the black bars and the number of bridge decks built with ECR is indicated by the gray bars. The

figure shows that the majority of bridges were built between 1955 and 1980 and that a considerable number of bridges with ECR decks were constructed prior to 1980.

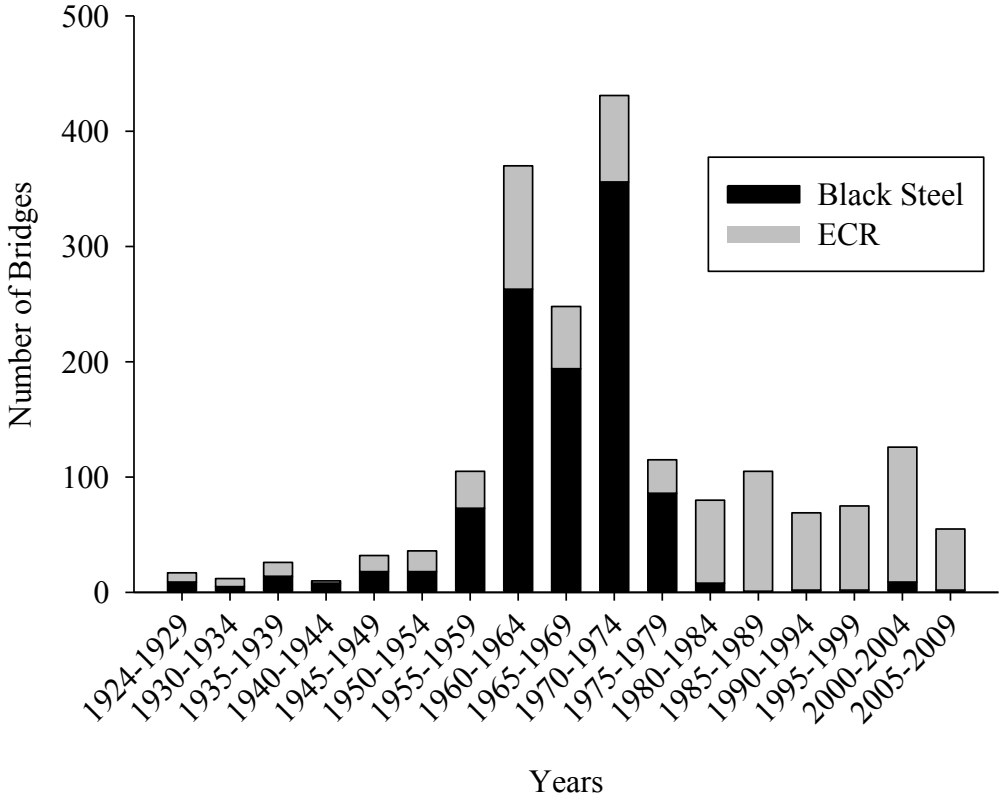


Figure 4-5: Distribution of ECR and black steel reinforcement in bridge decks

Figure 4-6 shows the proportion of black versus ECR bridge decks constructed for the same time periods as in Figure 4-5. As seen in Figure 4-6, the proportion of new bridges with ECR decks significantly increases beginning in 1980. While this reflects the implemented mandate, the figure also shows a surprisingly large proportion of bridges with ECR decks built prior to 1980. While some bridges with ECR decks are to be expected, the large proportions prior to 1980 are cause for concern, and indicate potential coding errors in the database.

Due to the small number of bridges built from 1924-1950, the percentage of ECR bridge decks may appear disproportionately large if there were only a few records coded incorrectly as ECR. Therefore, it is also beneficial to examine the year built distribution of just the ECR deck population, as shown in Figure 4-7. The figure shows that although the database shows some

ECR bridge decks constructed between 1925 and 1951 they make up less that 5% of all ECR bridge decks constructed, and thus are probably a coding error. From the data used to create Figure 4-7 it was calculated that 40% of all ECR bridge decks are coded as being built before 1975.

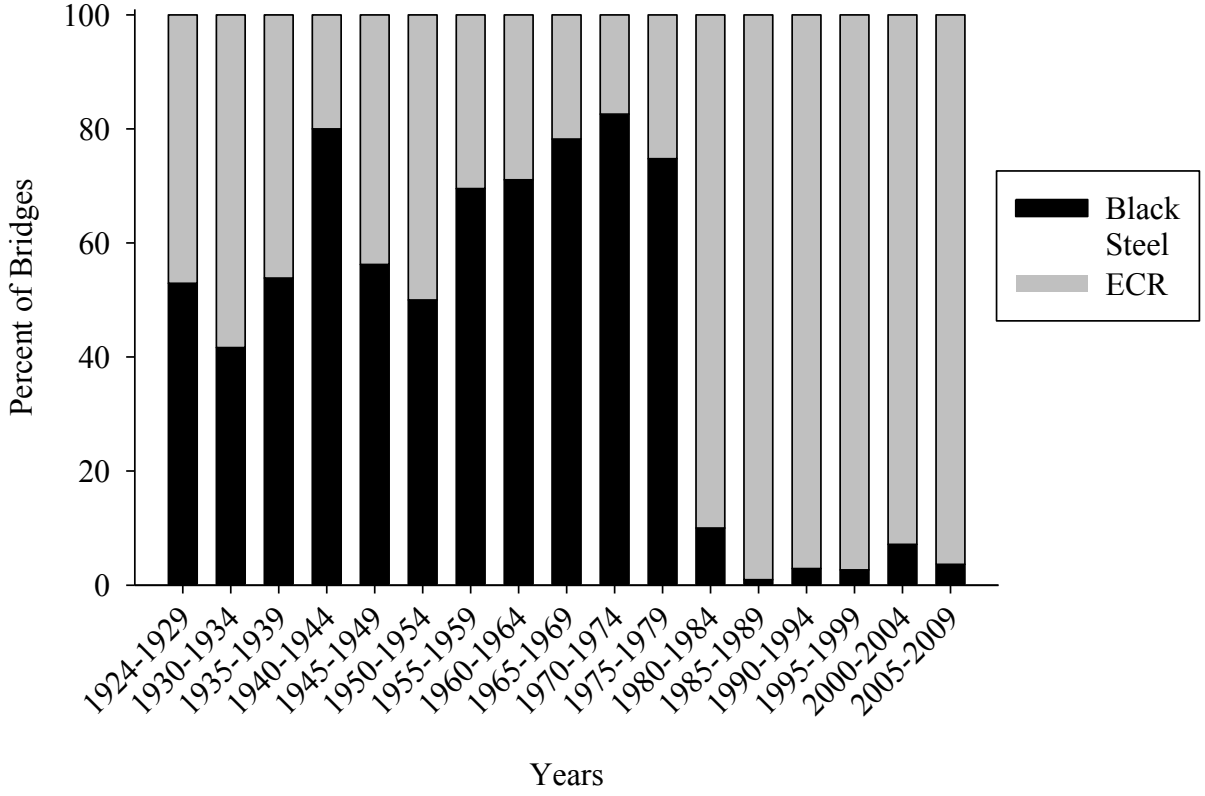


Figure 4-6: ECR and black steel proportion in bridge decks, year built

While it is possible that some bridges were built with ECR decks prior to the 1980 mandate, the statistic that 40% of bridges with ECR decks were built prior to 1975 is almost certainly erroneous. If records with bridge decks built with ECR before 1975 are considered erroneous, this means that 44% of the ECR inspection records in the database are coded incorrectly. This is a very large number of data that is affected and the implications of including these records in the network training data are enormous. Additionally, the rebar protection is an ANN input parameter and the predictive performance of the ANN models may be sensitive to the parameter. Thus, large amounts of incorrect training examples can be deleterious to the predictive performance of the ANN models. Also, it is impossible to discern whether the error in the

records comes from an incorrect year built or a miscoding of the rebar protection coating, and thus no modification to the records to make them usable is possible. For this reason and because the effects of the erroneous data on the ENN performance may prove especially grievous, a decision was made to remove any inspection records that belong to bridges with decks built with ECR prior to 1975. The years between 1975 and 1980 served as a grace period for the implementation of ECR decks in the Michigan bridge inventory. The removal of records with ECR decks built before 1980 reduced the number of ECR training records by 44%.

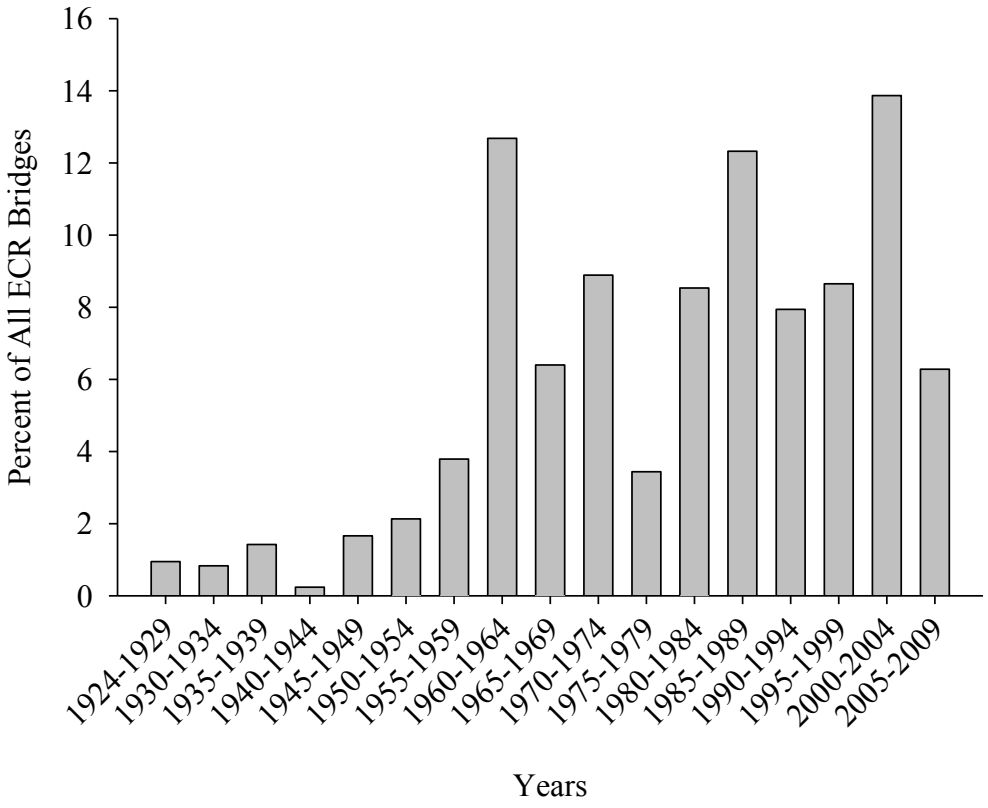


Figure 4-7: ECR bridge year built distribution

4.3.2 Network Architecture and Implementation

As discussed in Chapter 2, the four main network architecture components of an ENN are the inter-classifier interaction, the voting scheme, the diversity generator, and the ensemble size. The following sections offer details on these aspects of the ENN models developed in this study.

4.3.2.1 Individual MLP Architecture and Ensemble Diversity

In this research, the neural network ensembles are composed of individual MLPs using a back propagation algorithm. The MLPs had two hidden layers, 30 neurons per layer, a learning rate of 0.1 and a target mean square error of 0.15. Each MLP was allowed to iterate 2,000 times or until the target error was met. Discussion on the MLP architecture can be found in Section 4.2.2. The networks in the ENN were allowed a smaller number of iterations and a higher error than the MLP models presented in Section 4.2 because as part of an ensemble they need not have very high individual performances. However, the performance of the individual networks does need to meet the criterion of having a mean square error of 0.3 or less. If this criterion is not met the MLP was discarded and not used in the ensemble.

Diversity was introduced into the ensemble in several ways. First, as discussed in the next section, each MLP was trained on a unique set of data. Secondly, each MLP was initialized with random weights and biases. The training data sets and random weights create individual networks that are unique, and thus provide the ensemble with diversity.

4.3.2.2 Data Organization Scheme and Classifier Interaction

In this research, a modified bagging method proposed by Li (2008) was used to prepare the data for entry into the individual networks. Using the bagging method limits the interaction between the classifiers. Each MLP network was individually trained and did not interact with the other MLP networks. Bagging is a procedure used to produce multiple training sets by randomly drawing samples from the original training set with replacement. This procedure was used to create unique training sets for each MLP in the ensemble. The diversity of the ensemble increases because each MLP is trained on a different set of data. Although the training sets are different, because they come from the same large database, they are not independent. Thus, using the modified bagging method for data selection increases the diversity of the ensemble, but maintains some level of interaction between the networks.

As discussed, one area of difficulty for the MLP models is their ability to predict ratings with a low number of training records, such as high and low ratings. This is due to the severe data imbalance which results in a significantly lower number of high and low ratings. To decrease the severity of the data imbalance each MLP training set consisted of a uniform number of inspection records for each individual condition rating. The number of inspection records

selected from each condition rating was determined based on the condition rating containing the smallest number of inspection records. Table 4-13 shows the distribution of inspection records with respect to the condition rating for the deck surface rating.

Table 4-13: Deck surface inspection record distribution

NBI Rating	3	4	5	6	7	8	9
Number of Inspection Records	407	780	1011	1617	1827	846	249

The condition rating with the smallest number of inspection records was a '9' with only 249 records. To decrease the training data imbalance, MLPs developed from this this data will contain 249 inspection records for each condition rating from 3 to 9.

4.3.2.3 Ensemble Voting Scheme

The voting scheme dictates how the outputs of all the MLPs are combined to determine the ensemble output. In this research, plurality voting was used to determine the overall prediction of the ensemble. In this scheme, the prediction of the ensemble is taken as the most common prediction among the individual networks. In other words, the mode of the predicted vales from the MLP networks is the ensemble prediction.

4.3.3 Predictive Performance

Because each MLP was trained on a unique set of data, it is very difficult to track and evaluate the ENN performance based on the training data set. Thus, the performance parameters presented in the following sections are based only the results of the ENN when exposed to the testing set. Due to the method for creating the individual MLP training sets, it is impossible to create a testing set of completely unseen data. Thus, the entire database was used as the testing set.

The performance of an ENN is in part dependent on how many individual MLP networks there are in the ensemble. Therefore, to find the best performing ensemble, the performance of the ENNs were monitored as the number of networks in the ensemble increased. Figure 4-8 and Figure 4-9 graphically show the changes in the CR, AR, FIR, and DIR as the number of

networks was increased for the deck surface and deck bottom rating ensembles. The gray symbols in each figure indicate the performance parameters of the best performing MLPs. As seen in the figures, significant improvements occur in the DIR for both ensembles. Improvements in the CR and AR as the number of MLPs increases are also observed. Small increases in the FIR occur for both models. However, the increases are very small and do not indicate that the ensembles were particularly worse at falsely identifying damage than the MLP networks.

Figure 4-8 and Figure 4-9 were used in conjunction with Table 4-14 to determine the number of MLPs in the ensemble that would yield the best predictive performance. Computation time was also taken into consideration with the selection of the "best" ENN. Table 4-14 gives the exact performance measures and this information was used to differentiate between small differences in ensemble performance. For the deck surface rating, an ensemble with 85 individual MLPs was selected as the best performing individual. Although the DIR of this ensemble was slightly lower than the ensemble with 95 MLPs, the FIR was lower with 85 MLPs and it had a higher CR. For the deck bottom rating an ensemble with 85 MLPs was also selected.

Although the best performing ENN model can be selected using Figure 4-8, Figure 4-9, and Table 4-14, an examination of the ENN confusion matrices can provide additional information on the predictive performance. The following sections provide a detailed examination of the performance of the best performing ensemble using confusion matrices, bubble plots, and Press's Q-statistic.

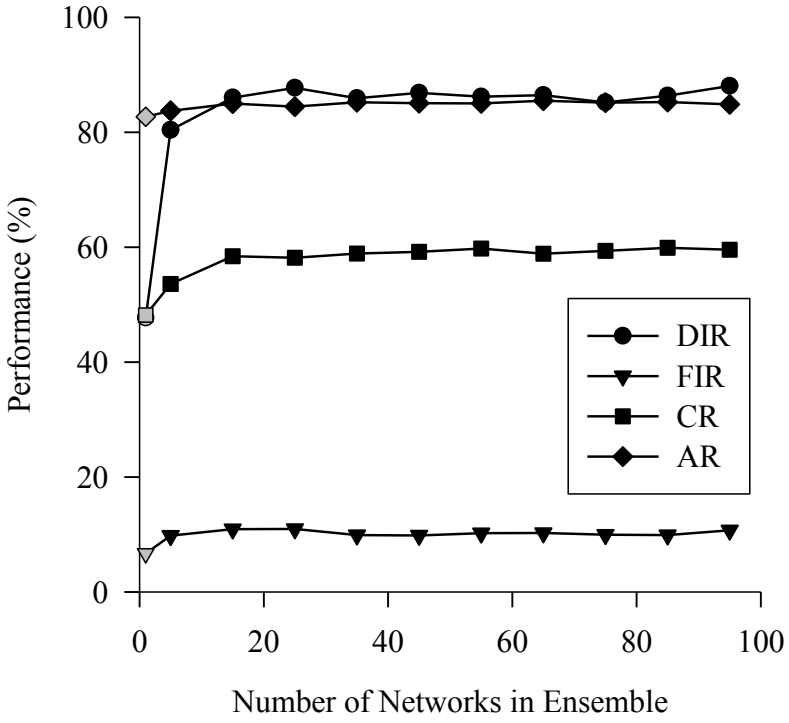


Figure 4-8: Surface ENN performance

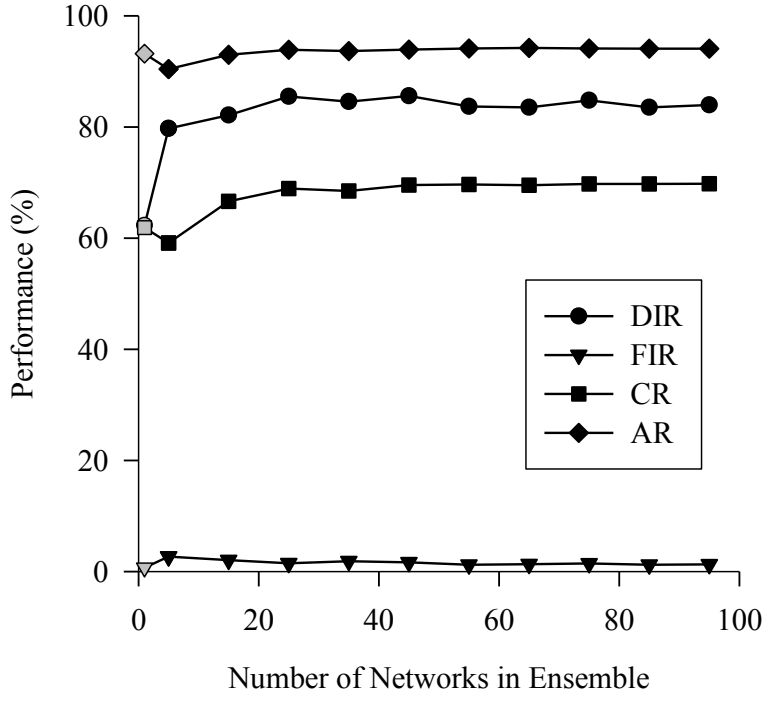


Figure 4-9: Bottom ENN performance

Table 4-14: ENN performance data

	Deck Surface Rating ENN				Deck Bottom Rating ENN			
MLPs in ENN	DIR (%)	FIR (%)	CR (%)	AR (%)	DIR (%)	FIR (%)	CR (%)	AR (%)
5	80.4	9.8	53.6	83.7	79.7	2.7	59.1	90.4
15	86.0	10.9	58.4	85.0	82.1	2.0	66.6	93.0
25	87.7	10.9	58.2	84.5	85.5	1.5	68.9	93.9
35	85.9	9.9	58.9	85.2	84.5	1.9	68.5	93.7
45	86.9	9.8	59.2	85.0	85.6	1.6	69.5	93.9
55	86.2	10.2	59.8	85.0	83.7	1.2	69.7	94.1
65	86.4	10.2	58.9	85.5	83.5	1.3	69.5	94.2
75	85.2	10.0	59.4	85.2	84.8	1.5	69.7	94.1
85	86.4	9.9	59.9	85.2	83.5	1.2	69.8	94.1
95	88.0	10.7	59.5	84.8	83.9	1.3	69.8	94.1

4.3.3.1 Confusion Matrices

The confusion matrices for the deck surface and deck bottom rating are shown in Table 4-15 and Table 4-16, respectively. As for the confusion matrices for the MLP models, see Section 4.2.3.1, the dark gray cells indicate the total number of correct predictions of the ENN. The light gray cells indicate the number of predictions within ±1 rating of the actual rating, which attempts to capture the subjectivity of the manual inspection process (Phares et al., 2001). When compared with the confusion matrices of the MLP models, see Section 4.2.3.1, it can be seen that the overall performance of the ensembles is significantly better.

The overall correct ratio of the deck surface model had an improvement of 10% with the implementation of an ENN. The individual ratings of 3, 4, 5, 8, and 9 showed improvements in the CR ranging from 17% to 77%. These individual ratings also experienced increases in their AR as well. The standard deviation for the individual ratings of 3 and 4 was reduced when the ENN was implemented. Another observation is that the mean predicted value is closer to the true value when predictions were made by the ENN.

Table 4-15: Deck surface ENN confusion matrix

	Manual Inspection							
Prediction	3	4	5	6	7	8	9	SUM
3	340	135	120	90	32	13	7	737
4	29	521	176	192	62	36	17	1033
5	21	85	590	246	87	25	17	1071
6	15	27	88	767	213	34	7	1151
7	1	9	28	247	1111	92	2	1490
8	0	0	4	62	258	512	6	842
9	1	3	5	13	64	134	193	413
SUM	407	780	1011	1617	1827	846	249	6737
CR (%)	83.5	66.8	58.4	47.4	60.8	60.5	77.5	59.9
AR (%)	90.7	95.0	84.5	77.9	86.6	87.2	79.9	85.2
μ	3.3	4.1	4.8	5.7	6.8	7.6	8.1	-
σ	0.8	0.8	1.0	1.2	1.1	1.3	1.8	-

Table 4-16: Deck bottom ENN confusion matrix

	Manual Inspection							
Prediction	3	4	5	6	7	8	9	SUM
3	329	100	121	17	0	0	0	567
4	37	506	271	29	0	0	0	843
5	10	111	1129	75	0	0	0	1325
6	7	37	276	1162	180	2	0	1664
7	0	6	52	281	956	35	0	1330
8	4	3	3	40	156	452	5	663
9	8	6	7	7	43	127	208	406
SUM	395	769	1859	1611	1335	616	213	6798
CR (%)	83.3	65.8	60.7	72.1	71.6	73.4	97.7	69.8
AR (%)	92.7	93.2	90.2	94.2	96.8	99.7	100.0	94.1
μ	3.4	4.2	4.9	6.1	7.0	8.1	9.0	-
σ	1.1	0.9	0.9	0.7	0.6	0.5	0.2	-

Table 4-17 compares the overall performance parameters of the single MLP and ENN models. From the table, it is observed that the deck surface ENN model had a significant improvement of 38.7% in the DIR. For the deck bottom rating, the ENN provided an improvement of 8% for the overall CR. The overall AR also improved by 1%. For the individual ratings, improvements in the CR for the ratings of 3, 4, 6, 7, and 9 ranged from 2.5%

to 97.7%. Improvement in the individual rating AR was also seen for the ratings of 3, 4, and 9. The standard deviation was reduced for the ratings of 7 and 9, and, as for the deck surface model, the ENN produced mean ratings closer to the true value than the single MLP. Error! Not a valid bookmark self-reference. also shows significant improvement in the DIR when the ENN was implemented for the deck bottom rating.

Table 4-17: MLP and ENN performance comparison

		DIR (%)	FIR (%)	CR (%)	AR (%)
Deck Surface Rating	MLP	47.7	6.7	48.2	82.7
	ENN	86.4	9.9	59.9	85.2
Deck Bottom Rating	MLP	62.3	0.7	61.9	93.2
	ENN	83.5	1.2	69.8	94.1

4.3.3.2 Bubble Plots

Figure 4-10 (a) and (b) presents a graphical representation of the ENN confusion matrices in Section 4.3.3.1. The developed bubble plots show larger bubbles on the diagonal, indicating a large number of correct predictions. A linear regression fitted to these plots can be used as another way to quantify the ability of the network to make correct predictions. The linear regression equations and parameters for the ENN bubble plots are listed in Table 4-18.

The slopes of the regression equations for both ENN models are close to 1.0, which indicates that the ENNs are able to predict a large number of ratings correctly. The R^2 values give a measure of the variability of the linear fit. It also provides a measure of how well the model predicts the dependent variable from the explanatory variable. As with the MLP, the explanatory variable in the bubble plots is merely a representation of the network input parameters. The linear regression performed on the bubble plots for the ENN models show an improvement in the R^2 values. The improvement in the R^2 values is analogous to a better predictive performance. When compared to the MLP models, the SEE values for both the ENN models were reduced. The reduction in SEE indicates a stronger linear relationship between the explanatory and dependent variables of the bubble plot.

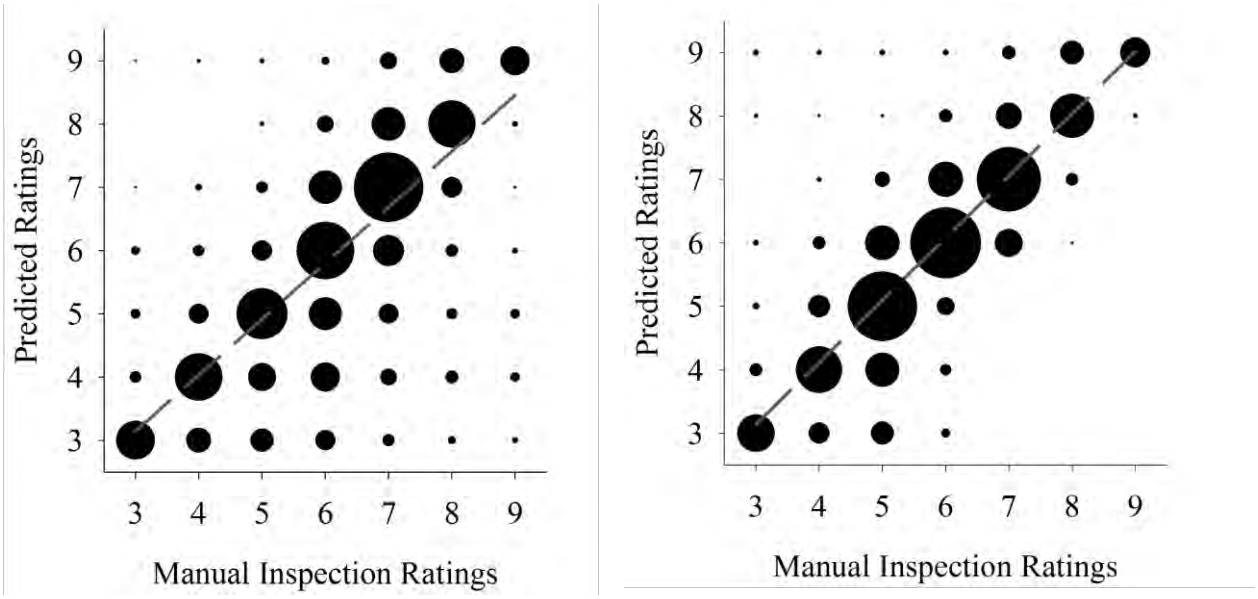


Figure 4-10: ENN bubble plots (a) deck surface rating, (b) deck bottom rating

Table 4-18: ENN bubble plot regression equations and parameters

	Equation	R^2	SEE
Bottom Rating	y=0.9785x+.2041	0.7693	0.60
Surface Rating	y=0.8867x+.4796	0.5907	1.30

4.3.3.3 Press's Q Statistic

As discussed in Section 4.2.3.3, Press's Q statistic provides a measure of the discriminatory power of a classification process. If the statistic is larger than critical value of 6.63 (from a Chi-squared distribution with one degree of freedom) it can be said that the method classifies significantly better than just chance (Hair et al., 1998). As seen in Table 4-19 the Q-statistic was much larger than 6.63 for both the surface and bottom ENN models. This indicates that the ENN models are able to classify better than chance alone. However, this statement must be taken with caution as the Q Statistic is influenced by sample size, and the ENN testing set is almost three times as large as the MLP testing set. The increase in the size of the testing set is due to the fact that the ENNs were tested using all of the available data, which corresponds to both the training and testing sets of the MLP models. Using all of the available data was necessary because the

modified bagging technique used to select the individual MLP models training data severely limits the number of records that remain unused in any of the MLP training sets. Because the number of records unused by the individual MLP models was so limited, it was elected to use the entire database for testing of the network performance. However, it should be stressed that the entire data set was not used in the training of any of the individual MLP networks.

Table 4-19: Press's Q statistic

ANN	Press's Q Statistic
Surface ENN	11437
Bottom ENN	17317

4.3.3.4 Conclusion

When compared to the MLP models, the ENN models show significant improvements in predictive performance. Improvements were observed for both the deck bottom and deck surface rating models in terms of the overall DIR, CR, and AR. The most drastic improvements were observed in the DIR. Nearly a 40% improvement occurred in the ability to identify damage when the ensemble structure was implemented for the deck surface rating. The ability to predict individual ratings for each model also improved with the implementation of the ensemble structure. The bubble plots and linear regression indicated a stronger ability to predict the condition rating given the real rating, or in the terms of the ANN models, a representation of the rating based on the ANN input parameters. Comparison of the performance parameters indicates that the ENN models are better at accurately predicting condition ratings for both the deck surface and deck bottom ratings. The ENN models are also significantly better at identifying damage than the single MLP models.

4.4 Input Parameter Influence

With typical regression models the independent variable coefficient is the measure of the variable influence on the output. However, such an equation with weighted parameter constants is not attainable for an ANN model. Due to the structure of an ANN, a direct measure of an input's influence on the output is not available. This is because an ANN model tracks the

influence of a parameter from neuron to neuron through the entire network. These influences, however, are represented in the connection weights between the neurons. Thus to gain insight on the effect of an input on the output the entire network must be examined. Several methods have been proposed to identify the influence of input parameters in ANN models. The methods proposed include neural interpretation diagrams, Garson's algorithm, and sensitivity analyses (Gevrey et al., 2003). Although multiple methods to assess the influence of the input parameters exist, only one was selected in this research. In a study by Olden and et al. (2004) it was found that a modified connection weight method developed by Olden and Jackson in 2002 provided the best accuracy and precision in determining influential input parameters out of a series of different methods. Thus, this method was adopted in an attempt to determine the influence of the ANN input parameters. The proposed method calculates the product of the input-hidden layer and hidden-output layer connection weights and sums the products across all hidden layers. It is important to note that this method is similar Garson's algorithm, but the sign of the connection weights is utilized. Garson's algorithm uses the absolute value of the weights and can yield considerably different and, as shown in the Olden and et al. (2004) study, inaccurate results.

As previously discussed, the output layers of the ANN models in this study contain one node for each unique condition rating (3 to 9). With this architecture, the influence of the input parameters on the different condition ratings can be examined. Because the condition rating declines as the age of the deck increases, the influence of different inputs on the condition rating can provide insight as to which parameters influence the deck degradation with respect to age.

4.4.1 Input Parameter Study for MLP Models

The results of the study for the MLP models are shown in Table 4-20. The top three influential parameters for each of the seven condition ratings are listed for the overall, bottom, and surface ratings. Because many of the input parameters are categorical in nature, each cell in the table contains the overall input (provided by the abbreviation) and the specific category of the parameter (in parentheses). For example, for the overall deck rating and condition rating of 3, the most influential parameter is the continuous parameter of age. The overall inputs and corresponding abbreviations (MDOT, 2009) are as follows: Continuous (C); Approach Surface (AP); Design Load (DL); Region (R); Structure Type (ST); Surface Protection (SP); and Rebar Protection (RP).

Table 4-20: MLP connection weight results

		Overall Rating		
		1st	2nd	3rd
Deck Rating	3	C (Age)	R (Metro)	ST (Cont Steel)
	4	R (Bay)	AP (Concrete)	R (Southwest Region)
	5	R (Grand)	R (University)	DL (HS20+Mod)
	6	AP (Bituminous surface on gravel)	AP (Brick)	RP (Epoxy Rebar)
	7	AP (Mixed Bituminous surface on concrete)	DL (HS25)	ST (Prestressed Concrete)
	8	AP (Concrete)	R (Metro)	AP (Mixed bituminous surface on concrete)
	9	AP (Concrete)	AP (Mixed bituminous surface on concrete)	R (University)
		Bottom Rating		
		1st	2nd	3rd
Deck Rating	3	AP (Bit Gravel)	C (Age)	AP (Unimproved Earth)
	4	AP (Concrete)	RP (Black Steel)	DL (HS25)
	5	R (Grand)	R (North)	R (University)
	6	AP (Brick)	ST (Simple/Cant Steel)	ST (Prestressed Concrete)
	7	R (Southwest)	C (Age)	AP (Concrete)
	8	R (Metro)	AP (Concrete)	AP (Mix bituminous surface on gravel)
	9	R (North)	DL (HS20)	C (Year Built)
		Surface Rating		
		1st	2nd	3rd
Deck Rating	3	DL (HS20)	ST (Prestressed Concrete)	SP (Built-Up)
	4	SP (Preformed Fabric)	DL (Other/Unknown)	SP (Other)
	5	R (Southwest)	RP (Black Steel)	R (Metro)
	6	DL (HS13.5)	R (Metro)	ST (Simple/Cant Steel)
	7	DL (HS15)	SP (Unknown)	SP (Epoxy)
	8	SP (Other)	R (Bay)	ST (Simple/Cant Steel)
	9	ST (Simple/Cant Steel)	SP (None)	SP (Preformed)

For the overall deck rating the region and the approach surface appear to have the most influence on the condition ratings. For the deck bottom rating these same parameters are also the most influential. For the deck surface rating the region and the deck surface protection type are the most influential parameters. Table 4-21 gives the values from the input parameter study for the parameters listed in Table 4-20. As seen in the table, the connection weight values are larger for lower ratings. The larger the weight value is for a parameter, the more influential that parameter is. For example, for a rating of 3 for the overall deck rating, the most influential parameter was age, and it had a weight value of 2.24. For a rating of 9, the most influential parameter was the approach surface type of concrete, but the weight value was only 0.71. This means that the age is much more influential than the concrete approach surface for their respective ratings even though they are both voted as the most important.

Table 4-21 shows higher weight values for lower ratings, which indicates that the parameters have a larger influence on lower condition ratings than they do on higher condition ratings. This means that age of the deck on a rating of a 3 is more important than the structure type on a rating of a 9, even though both inputs are the most important parameters for their respective rating.

Table 4-21: Parameter influence values, MLP

		Overall Rating			Bottom Rating			Surface Rating		
		1st	2nd	3rd	1st	2nd	3rd	1st	2nd	3rd
Deck Rating	3	2.24	1.36	1.28	2.12	2.03	1.24	0.89	0.56	0.42
	4	1.28	1.15	0.93	1.7	1.46	0.82	1.56	1.49	1.38
	5	2.22	1.74	1.44	2.79	2.14	2.02	1.59	1.31	1.04
	6	2.31	1.32	1.17	1.2	0.89	0.7	1.91	1.82	1.27
	7	0.82	0.57	0.5	0.76	0.68	0.45	1.68	1.61	1.03
	8	0.7	0.42	0.26	0.74	0.58	0.53	0.77	0.55	0.41
	9	0.71	0.57	0.19	0.25	0.24	0.21	0.87	0.34	0.25

Although this information is valuable, many of the parameters indicated as influential are discrete data inputs. Because each discrete variable is represented by multiple nodes, the number of discrete input nodes greatly outnumbers the continuous input nodes. Thus, although a continuous variable may have a large influence on a unique condition rating, it may not appear in

the table because multiple nodes of the same discrete input may be calculated as having more influence. Essentially, because the discrete variables are represented through multiple nodes, the influence of the variables may be double counted in Table 4-20. Table 4-22 was developed to examine the influence of the continuous variables only. Table 4-23 provides the connection weight values for the continuous parameters.

Table 4-22: Connection weight results, continuous variables only

		Overall Rating		
		1st	2nd	3rd
Deck Rating	3	Age	Year Built	ADT
	4	Age	Truck %	ADTT
	5	ADTT	# of Spans	Truck %
	6	Year Built	Truck %	ADT
	7	ADT	ADTT	# of Spans
	8	Year Built	Truck %	# of Spans
	9	Year Built	Truck %	# of Spans
		Bottom Rating		
		1st	2nd	3rd
Deck Rating	3	Age	Year Built	ADTT
	4	Truck %	Year Built	Age
	5	# of Spans	ADTT	ADT
	6	Year Built	# of Spans	Truck %
	7	Age	ADTT	ADT
	8	Year Built	Truck %	# of Spans
	9	Year Built	ADT	ADTT
		Surface Rating		
		1st	2nd	3rd
Deck Rating	3	ADTT	Skew	Year Built
	4	Age	Year Built	Span Length
	5	# of Spans	Span Length	ADTT
	6	Truck %	ADTT	Age
	7	Year Built	ADT	Truck %
	8	ADT	Skew	Span Length
	9	Skew	Year Built	Span Length

From Table 4-22 it is observed that for the overall rating the age and year built have the top influence on the majority of individual deck ratings. This indicates that time is a very important parameter in the prediction of the overall deck rating. The predominance of the age and year built parameters also occurs for the deck bottom rating model. The deck surface model has more diversity in the parameters voted as the most influential for the different individual ratings. This may indicate that at different stages in the deck's life, as indicated by the different condition ratings, different parameters have more of an influence than others. For example, ratings of 8 and 9 occur during the early years of a decks life. For these ratings, the ADT and skew were voted as the most influential continuous input parameters. Thus, it may be inferred that at early stages of the decks life the ADT and skew have the most influence on the deck rating.

Table 4-23: Continuous parameter influence weights

		Overall Rating			Bottom Rating			Surface Rating		
		1st	2nd	3rd	1st	2nd	3rd	1st	2nd	3rd
Deck Rating	3	2.24	0.84	0.46	2.03	0.59	0.24	0.33	0.25	0.16
	4	0.28	0.18	0.06	0.10	0.02	0.01	1.16	0.85	0.31
	5	1.23	0.48	-0.07	0.39	0.14	0.05	0.35	0.15	0.09
	6	0.69	0.18	-0.04	0.63	0.28	0.23	0.12	0.11	0.10
	7	0.15	0.14	0.12	0.68	0.12	-0.06	0.40	0.24	0.17
	8	0.06	0.01	-0.02	0.15	0.08	0.02	0.37	0.11	0.08
	9	0.18	0.03	0.00	0.21	0.13	0.00	0.12	0.02	0.01

The small values of the connection weights listed in Table 4-23 show that some of the top ranked continuous input parameters do not have a large influence on the output ratings. The values also show that some of the 2nd and 3rd most important parameters for some condition ratings are just as important, if not more, than the top ranked parameters for other condition ratings. For example, for the overall deck rating the ADT is ranked as the 3rd most important parameter for a rating of 3 but the value is larger than the most important parameter for a rating of 4, which is age. These values indicate that some condition ratings are influenced by multiple continuous parameters to a greater extent than other ratings are to any continuous parameters.

4.4.2 Input Parameter Study for the ENN Models

The influence of the different input parameters was also studied for the ENN deck surface model. The same method utilized for the MLP models was implemented to identify the top three most influential parameters for each individual MLP in the ensemble. A plurality voting scheme was then used to identify the top three ranked inputs for the ensemble. Using the plurality voting scheme, the input parameter with the most votes from all the individual MLPs (i.e., the mode) was selected as the ensemble vote.

Table 4-24 shows the results of the connection weight study for the ENN model. Each rating, i.e., output, contains three rows. The rows represent the 1st, 2nd, and 3rd ranked variables. The 'Input' column for each row gives the input parameters that received the first, second, and third largest number of votes for the ranking in question. For example, the surface protection of preformed fabric received 20% of the votes as the #1 most important parameter for a rating of nine. The design load 0 (DL0) received 18.8% of votes as the most important parameter for a nine, and the design load 2 (DL2) received 7.5% of votes as the most important parameter for a rating of 9. The representation of each input is the same as in Section 4.4.1 with the overall input provided by an abbreviation and the specific category given in the parentheses. The abbreviations are the same as provided in Section 4.4.1.

An overall examination of the table shows a predominance of surface protective coatings as the most important parameters for ratings of 8 and 9. For ratings of 6 and below, the deck age was voted by the ensemble as having the most influence. The region is also another parameter that populates a large number of cells in the table. Additionally, the rebar coating only appears to have an influence of the ratings of 8 through 5.

Table 4-25 gives the average connection weight values for the top three most influential parameters that received the most votes by the ensemble. The values correspond to the inputs listed in the column "Highest Frequency per Ranking" in Table 4-24. The results presented in Table 4-25 show that the most influential parameter for all condition ratings was age for condition rating 3. Deck surface protection for ratings of 8 and 9 were also among the most influential.

Table 4-24: ENN input rankings

			Top Inputs Within Ranking					
			Highest Frequency per Ranking		2nd Highest Frequency per Ranking		3rd Highest Frequency per Ranking	
			Input	(%)	Input	(%)	Input	(%)
Ranking per Output	9	1st	SP (Preformed Fabric)	20	DL (Other/Unknown)	18.8	DL (HS15)	7.5
		2nd	SP (Preformed Fabric)	15	DL (Other/Unknown)	12.5	C (Year Built)	11.3
		3rd	C (Year Built)	15	R (North)	12.5	DL (Other/Unk.)	10
	8	1st	SP (Other)	23.8	R (Grand)	10	RP (ECR)	6.3
		2nd	SP (Other)	12.5	DL (HS13.5)	8.8	DL (H20)	8.8
		3rd	DL (HS25)	10	C (ADTT)	6.3	R (Grand)	6.3
	7	1st	C (Year Built)	23.8	SP (None)	16.3	DL (DL 3)	7.5
		2nd	C (Year Built)	16.3	SP (None)	13.8	RP (Black Steel)	7.5
		3rd	C (Year Built)	11.3	RP (ECR)	7.5	SP (None)	7.5
	6	1st	C (Age)	17.5	SP (None)	16.3	SP (Unknown)	11.3
		2nd	C (Age)	23.8	R (Metro)	10	SP (Unknown)	7.5
		3rd	C (Age)	11.3	C (Year Built)	8.8	R (University)	7.5
	5	1st	C (Age)	25	RP (Black Steel)	11.3	DL (HS13.5)	7.5
		2nd	C (Age)	15	R (University)	10	RP (ECR)	10
		3rd	DL (HS20+Mod)	11.3	SP (Unknown)	10	R (Bay)	6.3
	4	1st	C (Age)	38.8	SP (Unknown)	7.5	R (North)	6.3
		2nd	C (Age)	20	R (Superior)	12.5	R (Bay)	7.5
		3rd	R (Superior)	13.8	C (Age)	10	R (Bay)	10
	3	1st	C (Age)	53.8	R (North)	7.5	SP (Preformed Fabric)	7.5
		2nd	C (Age)	16.3	SP (Preformed Fabric)	11.3	DL (HS20)	8.8
		3rd	R (North)	12.5	DL (Other/Unk.)	10	R (Bay)	8.8

As with the results from the investigation of the MLP model (see Section 4.4.1) many of the top ranked input parameters were discrete variables. Table 4-26 shows the results of the investigation if only continuous variables are considered. The table is set up in the same fashion

as Table 4-24 with the variables receiving the first, second, and third largest amount of votes for each ranking in the rows. As with Table 4-24, the deck age is selected as an important parameter for ratings of 7 through 3.

Table 4-25: Connection weight study values

		1st	2nd	3rd
Deck Rating	3	10.24	6.66	5.97
	4	6.23	5.25	4.28
	5	6.22	4.90	4.26
	6	7.17	5.06	4.48
	7	7.70	4.67	4.73
	8	7.42	4.65	3.54
	9	7.90	5.76	5.03

As the rating decreases an increasing number of MLPs selected age as the most important parameter. The number of spans appears to have an effect on the lower deck ratings as well. For the rating of 8 both span length and skew angle were selected as the second and third most important parameters. The ADTT was selected a number of times but mainly as the parameter with the third highest frequency of occurrences for the ranking.

Table 4-27 gives the average connection weight values for top three most influential continuous parameters that received the most votes by the ensemble. The values correspond to the inputs listed in the column "Highest Frequency per Ranking" in Table 4-26. The results show that the most influential parameter for all condition ratings is the age for condition rating 3. The table also shows that the continuous parameter of ADTT at a rating of 8 is the least influential of the top parameters. The connection weight values also indicate that the 2nd and 3rd most important parameters were significantly less influential than the top parameter, given their lower weight values.

The connection weight study provides valuable insight as to the influence of the input parameters on the output. The discrete nature of the rating system allows for an evaluation of the inputs on each separate output rating, which provides insight as to how each input influences the condition of the bridge deck at different stages in the deck life. Overall, age was observed to be extremely influential for lower condition ratings. Discrete variables such as region, approach

surface, and deck surface protection were also identified as influential. The connection weights themselves provide a measure of how important each parameter is. Larger connection weights indicate the parameter has more of an influence on the output. This allows for a quantitative comparison of the top influential parameters for each condition rating. Overall, the study shows that different parameters provide different effects at different stages of the degradation process, as represented by the discrete NBI ratings.

Table 4-26: Ranking of top continuous inputs

			Top Inputs Within Ranking					
			Highest Frequency per Ranking		2nd Highest Frequency per Ranking		3rd Highest Frequency per Ranking	
			Input	(%)	Input	(%)	Input	(%)
Ranking per Output	9	1st	Year Built	71.3	ADT	16.3	# of Spans	6
		2nd	ADT	42.5	Year Built	17.5	# of Spans	12
		3rd	# of Spans	30	ADT	22.5	Span Length	18
	8	1st	ADTT	38.8	Span Length	20	Skew	14
		2nd	Span Length	21.3	Skew	20	# of Spans	15
		3rd	Skew	27.5	# of Spans	25	ADT	11
	7	1st	Year Built	76.3	Age	10	ADTT	7
		2nd	Age	22.5	Span Length	21.3	ADTT	16
		3rd	ADTT	25	Span Length	23.8	ADT	11
	6	1st	Age	76.3	Year Built	12.5	ADTT	3
		2nd	Year Built	35	ADT	16.3	ADTT	11
		3rd	Skew	20	Span Length	17.5	ADTT	13
	5	1st	Age	76.3	# of Spans	10	Skew	4
		2nd	# of Spans	25	ADT	20	Skew	16
		3rd	Num of Spans	27.5	Skew	22.5	ADT	16
	4	1st	Age	82.5	Num of Spans	5	Year Built	3
		2nd	Num of Spans	26.3	Span Length	22.5	ADT	15
		3rd	Num of Spans	27.5	Skew	21.3	ADTT	14
	3	1st	Age	92.5	ADT	3.8	Year Built	1
		2nd	ADT	43.8	Skew	26.3	Year Built	10
		3rd	Skew	30	ADT	22.5	Num of Spans	11

Table 4-27: Continuous parameter connection weight values

		1st	2nd	3rd
Deck Rating	3	8.31	2.48	1.23
	4	5.19	1.65	1.31
	5	4.21	1.31	0.66
	6	4.92	3.19	1.07
	7	5.10	2.90	1.05
	8	2.23	1.10	0.34
	9	4.00	1.28	0.29

4.5 Markov Model

As discussed in Section 2.2, a Markovian degradation process is conventionally assumed in bridge degradation models. However, as noted in Section 2.2, several studies have raised concerns about the time-independent assumption in the Markov method for the problem of predicting the degradation of bridge decks. Nonetheless the method is well understood and trusted by many researchers and practitioners. It is thus of interest to compare the performance of the developed ANN models with the results from a Markov approach and to compare the results from both methods. To provide a direct comparison of the predicted degradation from the two types of models, a Markovian model unique to the database used in this study was developed. The Markovian transition probability matrices were developed using the same database from which the ENN training data were selected.

4.5.1 Development

The behavior of a Markovian process is determined through using transition probabilities, $p_{ij}(n)$, which are unique to the given data set. Each value gives the probability that the process will be in state j at time n given that it was in state i at the previous step. The transition probabilities are commonly displayed in a matrix form, known as the transition probability matrix. If a process has r possible states the transition matrix can be written in the form

$$\Pi = \begin{bmatrix} p_{11} & p_{12} & \cdots & p_{1r} \\ p_{21} & p_{22} & \cdots & p_{2r} \\ \vdots & \vdots & & \vdots \\ p_{r1} & p_{r2} & \cdots & p_{rr} \end{bmatrix} \quad \textbf{4-3}$$

To determine the transition probability matrix, the number times a particular change in rating occurs between consecutive inspections for a specific bridge must first be calculated. The assumed time period between consecutive inspections is two years as per the specification of the FHWA. Table 4-28 provides the number of condition transitions for the entire database.

The ratings on the left hand side are the past ratings and the ratings along the top are the future ratings. In other words, the ratings on the left are the ratings that the deck is transitioning *from*, and the ratings on the top are the ratings that the deck is transitioning *to*. For example, the cell highlighted in gray indicates that 91 bridges were identified as transitioning from a rating of a 4 to a rating of a 3.

Table 4-28: Number of rating changes

From/To		Bridge Rating (Future states)							
		3	4	5	6	7	8	9	SUM
Bridge Rating (Past states)	3	239	12	4	7	1	3	7	273
	4	91	425	29	18	8	15	7	593
	5	41	137	510	47	15	9	18	777
	6	12	63	187	881	87	19	12	1261
	7	4	14	43	266	1024	61	6	1418
	8	1	1	3	36	232	354	26	653
	9	0	0	1	4	29	89	50	173

After the number of occurrences was tallied the proportion of each cell with respect to the total number of the respective 'from' ratings was calculated. For example, the 91 occurrences of a ratings transitioning from a 4 to a 3 make up a 0.15 proportion of the 593 ratings starting as a 4. These proportions are the transition probabilities. When compiled, the proportions create the transition probability matrix. Table 4-29 shows the transition probability matrix developed from the results reported in Table 4-28. This transition matrix will be further referred to as TPM1 (Transition Probability Matrix 1).

The dark gray cells on the diagonal of the matrix in Table 4-29 represent the probability that a rating will stay at the indicated rating for the two year inspection period. The results of the Markov model show that there are some occurrences of the rating improving between inspections although the database was checked for these kinds of errors previously. One option for correcting this error is to manually redistribute the probabilities to other cells in the 'from' rows. Table 4-30 shows the updated transition probabilities calculated utilizing this strategy. This transition probability matrix will henceforth be referred to as TPM2. All probabilities of improving the rating were redistributed to the decline in rating side of the matrix. The probabilities for the rating staying the same were not modified.

Table 4-29: Markov Transition Probability Matrix 1 (TPM1)

From/To		Bridge Rating (Future states)						
		3	4	5	6	7	8	9
Bridge Rating (Past states)	3	0.88	0.04	0.01	0.03	0.00	0.01	0.03
	4	0.15	0.72	0.05	0.03	0.01	0.03	0.01
	5	0.05	0.18	0.66	0.06	0.02	0.01	0.02
	6	0.01	0.05	0.15	0.70	0.07	0.02	0.01
	7	0.00	0.01	0.03	0.19	0.72	0.04	0.00
	8	0.00	0.00	0.00	0.06	0.36	0.54	0.04
	9	0.00	0.00	0.01	0.02	0.17	0.51	0.29

Table 4-30: Markov Transition Probability Matrix 2 (TPM2)

From/To		Bridge Rating (Future states)						
		3	4	5	6	7	8	9
Bridge Rating (Past states)	3	1.00	0.00	0.00	0.00	0.00	0.00	0.00
	4	0.28	0.72	0.00	0.00	0.00	0.00	0.00
	5	0.17	0.18	0.66	0.00	0.00	0.00	0.00
	6	0.01	0.14	0.15	0.70	0.00	0.00	0.00
	7	0.00	0.01	0.03	0.23	0.72	0.00	0.00
	8	0.00	0.00	0.00	0.06	0.40	0.54	0.00
	9	0.00	0.00	0.01	0.02	0.17	0.51	0.29

Although improvements in the condition rating indicate some form of reconstructive work occurring on the bridge deck, the variability and subjectivity of the inspection rating process should be reflected in the transition probability matrices. As previously discussed, a subjectivity of ±1 rating point (Phares et al., 2001) is adopted in this research. Table 4-31 shows the transition probability matrix with only the transition probabilities of improvements of more than one rating point redistributed. This transition matrix will be referred to as TPM3.

In this study, the transition probability matrices were developed based on the entire network; therefore, the use of such transitions probability matrices for a specific bridge may not be appropriate. This can be a drawback for the Markov model as the degradation of bridge decks may vary significantly based on different design parameters. It could also make allocating budget for repairs difficult because the needs of an individual bridge cannot be known without an accurate measure of its condition.

Table 4-31: Markov Transition Probability Matrix 3 (TPM3)

From/To		Bridge Rating (Future states)						
		3	4	5	6	7	8	9
Bridge Rating (Past states)	3	0.96	0.04	0.00	0.00	0.00	0.00	0.00
	4	0.23	0.72	0.05	0.00	0.00	0.00	0.00
	5	0.11	0.18	0.66	0.06	0.00	0.00	0.00
	6	0.01	0.07	0.15	0.70	0.07	0.00	0.00
	7	0.00	0.01	0.03	0.19	0.72	0.04	0.00
	8	0.00	0.00	0.00	0.06	0.36	0.54	0.04
	9	0.00	0.00	0.01	0.02	0.17	0.51	0.29

4.6 Summary on Degradation Models

Two types of ANN models, MLP and ENN, were successfully developed to predict the condition rating, or degradation, of bridge decks. Single MLP models were developed to predict the overall deck rating, the deck surface rating, and the deck bottom rating. The MLP models provided satisfactory performance in the identification of damage, predicting the condition within ±1 rating point, and had low false identification ratios. However, their ability to identify the exact condition rating was lacking, especially for high and low ratings. The differences

between the training and testing performances provided evidence that the MLP networks were not overtrained.

Implementing the ensemble method led to significant improvement of the ANNs predictive performance. ENNs were developed to predict the deck surface rating and the deck bottom rating and had a drastic improvement on the overall CR and DIR. Improvements were also made to the individual rating CR and AR. An evaluation of Press's Q statistic showed that both the MLP and ENN models were able to predict the deck condition rating better than chance alone.

A study of the influence parameters of the MLP and ENN models showed that the inputs influence the different condition ratings in different ways. For example, for lower ratings the age of the bridge deck emerged as the most influential parameter. For higher deck ratings the most influential parameters were discrete inputs such as region or approach surface.

The developed Markov model will allow a comparison of the degradation of the bridge deck through its lifespan. Since this approach is currently utilized by MDOT for predictive assessment of its highway inventory, this will provide a good comparison of currently accepted models and the new ANN models developed in this study.

5 DEGRADATION CURVES

The ability of the ANN models to provide satisfactory performance in correctly classifying data that was not used in the network training stage has been established. The developed models can thus be used to model the degradation of a bridge deck during its service life. The developed degradation curves serve an important method to evaluate the life time condition of a bridge deck. Degradation curves were developed using the MLP, ENN, and Markov models presented in Chapter 4. The ANN models can be utilized to predict both the condition of a unique bridge and the condition of a network of bridges. The Markov model can only be utilized in a network setting. The following sections outline the degradation curve development process and highlight the differences between the curves developed by each model. A discussion regarding a parametric study conducted on the effects of black steel and ECR steel decks on deck degradation is also included.

5.1 Development of ANN Model Curves

Two types of degradation curves were developed from the ANN models: bridge specific curves and generalized network curves. Unique bridge curves model the degradation of an individual bridge. Generalized network curves were developed to monitor the degradation of a system of bridges. The following sections outline the process by which the individual and generalized curves were developed for the ANN models.

5.1.1 Unique Bridge Curves

The following steps were used to develop the degradation curve for a specific bridge.

1. Obtain the values of the input parameters for the bridge in question.
2. Format the input parameters for entry into the neural network model (normalization and binary representation).
3. Set the age of the deck to a normalized value of '1'.
4. Present the input parameters to the network for prediction.
5. Increase the age by 1 year, remembering to normalize the value.
6. Repeat steps 4 and 5 for the entire life of the deck.

Following these steps gives a predicted condition rating for each year of the deck's service life. Plotting the condition rating against age provides a visual representation of how the deck rating changes over time. Figure 5-1 shows the overall deck rating predictions made by the MLP model for bridge 09109035000S130.

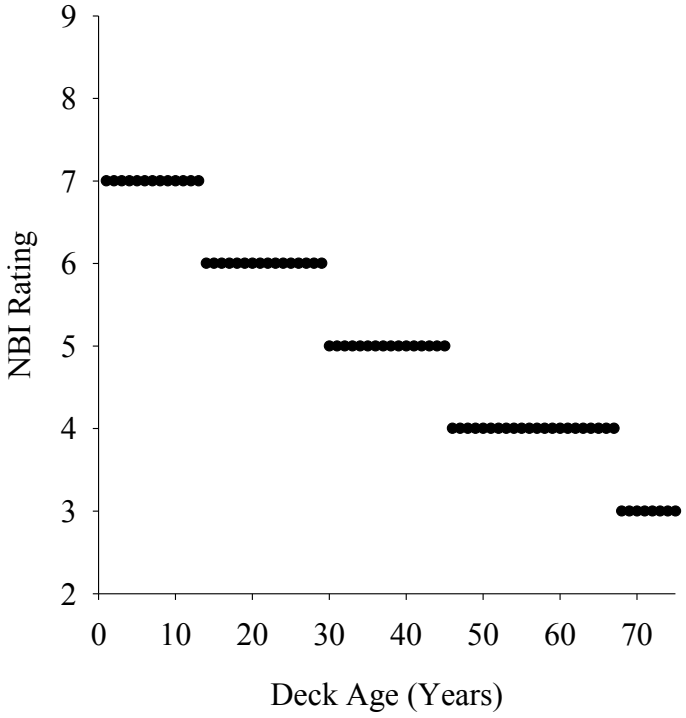


Figure 5-1: Overall deck rating lifetime predictions

This bridge is a continuous steel bridge with two 108 foot spans located in Bay County, Michigan built in 1968. The bridge was built using black steel reinforcement, has zero skew angle, supports two lanes of traffic, and has an approach surface of mixed bituminous on gravel base. Each dot in Figure 5-1 represents the overall deck condition rating for the respective year of the bridge deck's life. As seen in the figure, the condition of the deck decreases over time. It is important to emphasize that the data points presented in Figure 5-1 are neural network predictions. When these predictions are connected with a line, a stepwise trace emerges as shown in Figure 5-2.

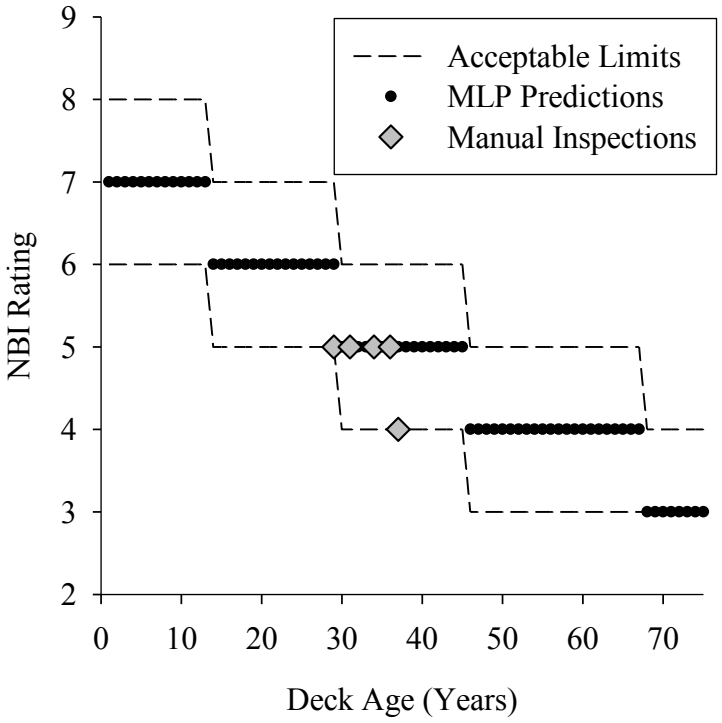


Figure 5-2: Overall deck rating degradation curve

The stepwise nature of the curve is due to the limitation of the network to only predict discrete ratings from 3 to 9, which is dictated by the NBI rating system. In Figure 5-2 the thick solid line is the trace produced by connecting the MLP predictions. The diamonds represent the manual inspections in the database for this particular bridge. Showing the manual inspections alongside the MLP predictions allows for a comparison of how accurately the MLP predictions match the observed conditions for the bridge. The trace defined by the dashed lines defines the bounds to the acceptable predicted value; that is the bandwidth of ±1 to represent inspection subjectivity as discussed for the confusion matrices.

Although a stepwise nature of deck degradation is what is represented in the NBI database, the true degradation of a bridge deck is clearly much more gradual and may be better represented by fitting the predicted ratings with a smooth function. Thus, a logistic function in the form of Equation 5-1 was selected for developing a smoothed degradation curve. Li and Burgueño (2010) proposed Equation 5-1 to fit curves modeling the degradation of bridge abutments.

$$y = a * \frac{1 + m * e^{\left(\frac{b-x}{t}\right)}}{1 + n * e^{\left(\frac{b-x}{t}\right)}} \qquad \textbf{5-1}$$

When the MLP predictions are fitted using Equation 5-1, the resulting curve is a backward S-shape as shown in Figure 5-3. Also shown are the 95% prediction bounds to the logistic fit. It can be seen that all the network predictions fall within these prediction bounds. It is important to emphasize that the logistic fit is made to the MLP predictions. The actual manual inspections represented by the diamonds in Figure 5-2 are not used to calculate the logistic fit.

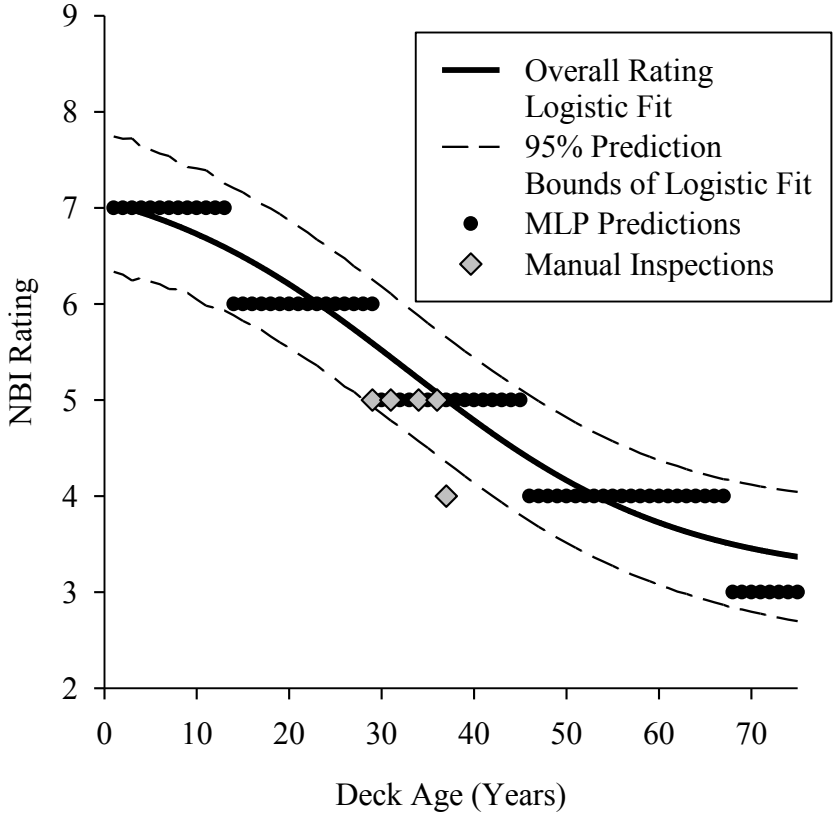


Figure 5-3: Fitted deterioration curve for overall deck rating

The logistic fit provides a continuous representation of the decline in condition of the bridge deck. Both the continuous and discrete degradation curves were developed using both the MLP and ENN models. Upon comparison, it was observed that models produce slight differences in

the degradation curves. The differences in the degradation curves, as discussed in Section 5.3.1, are due to the differences in predictive ability of the models.

5.1.2 Generalized Network Curves

Generalized curves representing the degradation of a bridge network can be developed by statistically combining the predictions for multiple bridges in a network. For illustration of the method, the deck surface rating will be examined and the entire bridge database will be utilized. The first step in the process is to obtain the neural network predictions, as outlined in Section 5.1.1, for all bridges under consideration. In this discussion the MLP model was used for rating predictions. As shown in Figure 5-4, plotting all predicted ratings in a two dimensional plane provides little information as to the nature of which ratings are predicted at a higher frequency at each age. The figure also shows significant scatter.

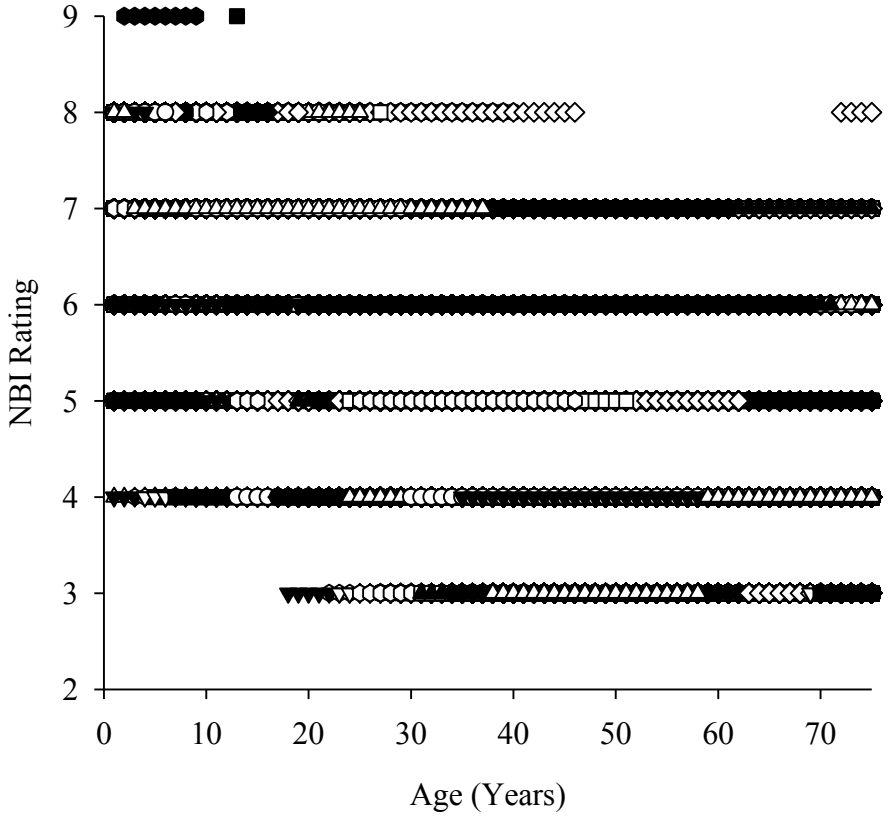


Figure 5-4: All MLP deck surface rating predictions

A 3D representation of the data, as shown in Figure 5-5, offers more insight into the distribution of ratings. Figure 5-5 shows the distribution of predicted ratings per year. On the x-axis is the deck age, and on the y-axis is the NBI rating, just as in Figure 5-4. However, the z-axis shows the number of predictions made for each year for the individual ratings. For example, just over 1,200 bridge decks were predicted to have a rating of an 8 at age 1.

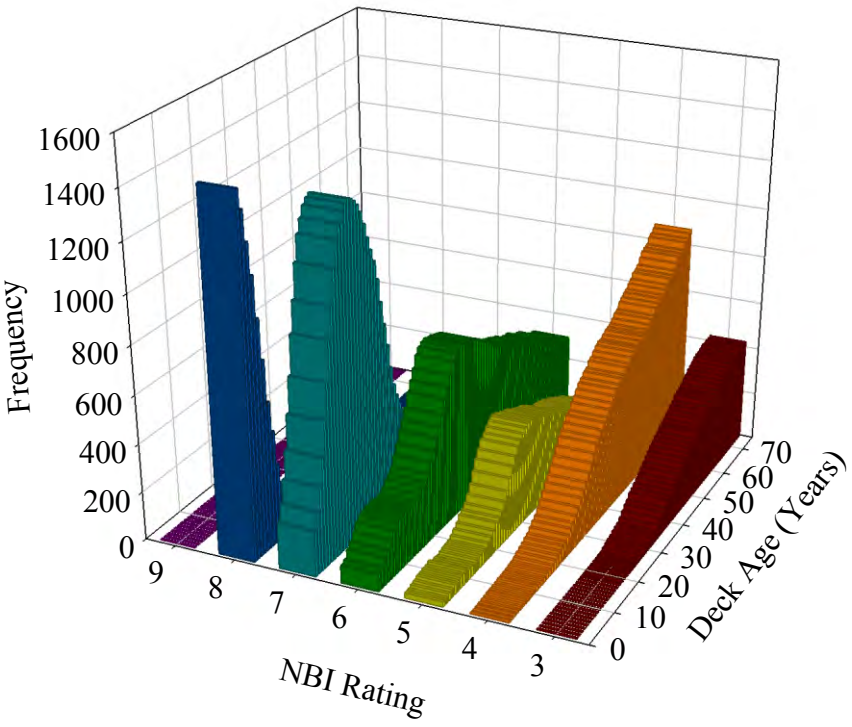


Figure 5-5: MLP predicted deck surface rating distribution. For interpretation of the references to color in this and all other figures, the reader is referred to the electronic version of this report.

The 3D representation allows for a detailed view as to which ratings are predominantly predicted at certain ages. This information was used to calculate the average predicted rating per year. One year can essentially be visualized as a slice of the 3D plot along the yz plane. Figure 5-6 is an example of this visualization at an age of 20 years.

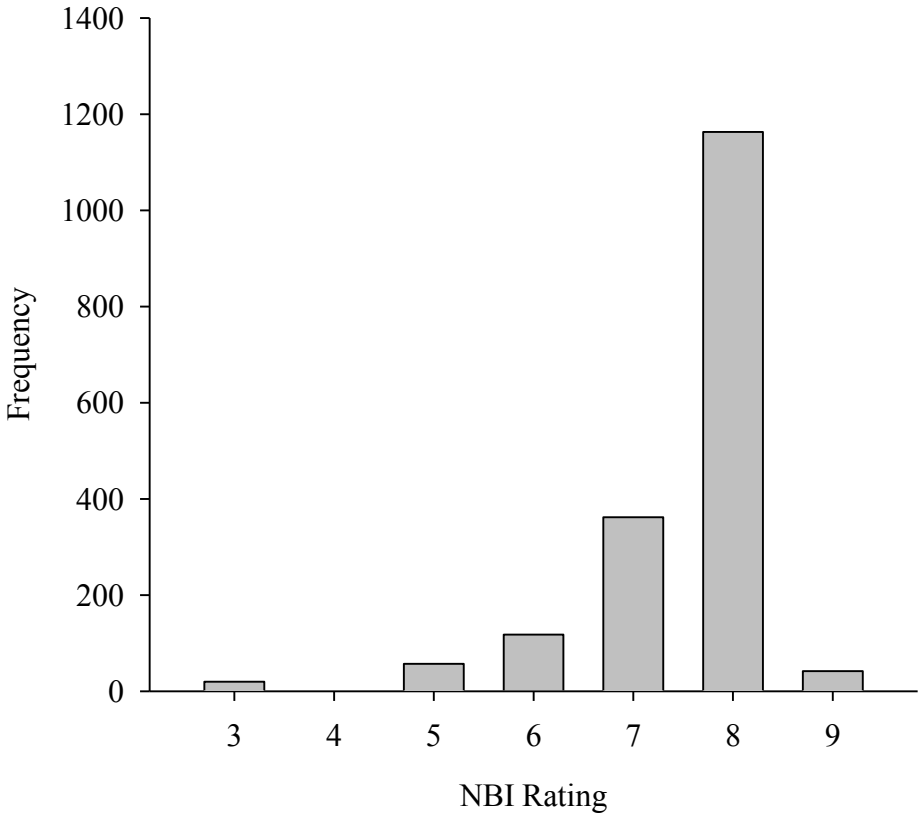


Figure 5-6: Distribution of ratings at age 20

Figure 5-6 shows a predominance of predictions as a rating of 8. However, predictions of other ratings occur frequently. For this particular year, the average predicted rating was 6.6 and the standard deviation was 0.86. The mean and standard deviation for each year was calculated and the resulting plot is shown in Figure 5-7. The standard deviation was used to calculate the confidence intervals for each year.

The generalized curves can be used to monitor the average decline of the condition of the bridge network in question. As the generalized curves model the condition of a network of bridges they can also be used to facilitate a comparison between the ANN and Markov models. The following section discusses the process used to develop network level degradation curves using the Markov model developed in Section 4.5. A comparison of the ANN generalized curves and the Markov model curves is presented in Section 5.3.2.

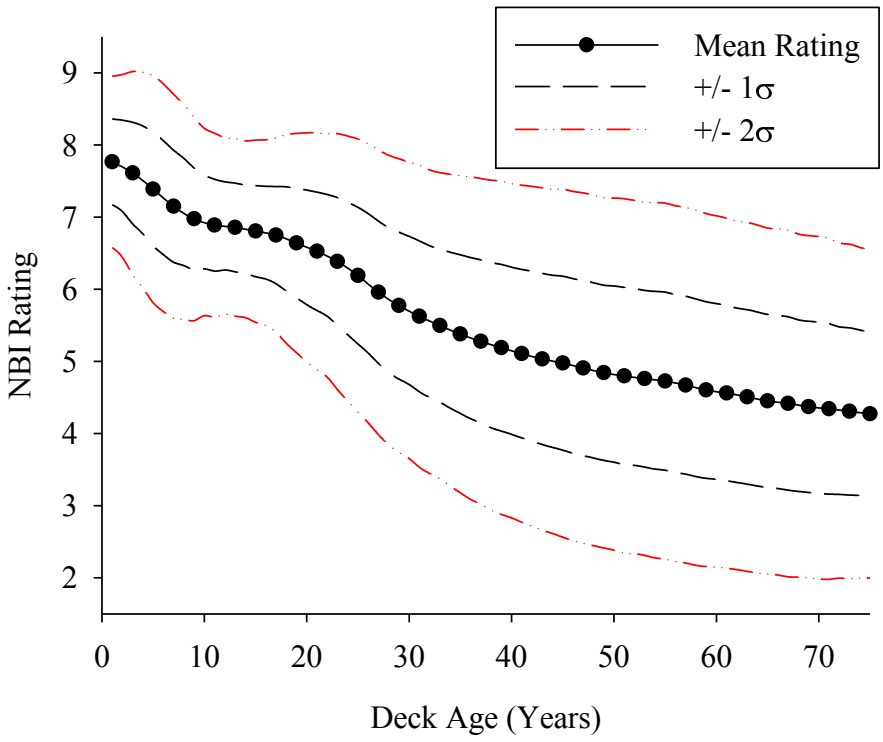


Figure 5-7: Mean and standard deviation curves, deck surface rating

5.2 Development of Markov Model Curves

A Markov degradation model can be used to calculate the average predicted condition of a network of bridges. The average predicted condition calculated using the Markov model is comparable to the network average condition calculated using the generalized degradation curve approach presented in Section 5.1.2. The Markovian transition probability matrices (see Section 4.5) can be used to calculate the average rating of a network of bridges. Figure 5-8 shows the degradation curves developed using the three different probability matrices as presented in Section 4.5.

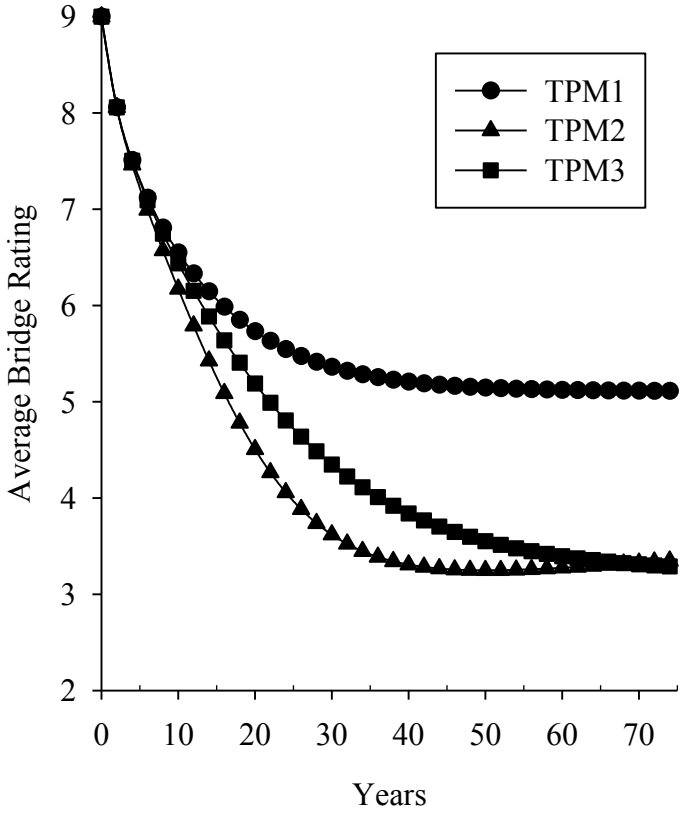


Figure 5-8: Markov degradation curves

To create the degradation curves, the current state of the network must first be determined. The proportion of decks at each condition rating is then calculated. The proportion of bridges at each rating for the first time step (two years) is calculated by performing matrix multiplication using the initial condition vector and the transition probability matrix. The proportion of bridges in each rating at the next time step, or four years of age, is calculated by performing matrix multiplication with the vector of rating proportions for two years and the transition matrix. This iteration is repeated for many subsequent time steps. Once the proportion of ratings is determined, the average rating at each time step is calculated. The average rating of the network can be plotted against age (*time step number* x *time step interval*) to obtain a deterioration curve. The curves plot the average condition of the network as time progresses. The degradation curves predicted by the developed Markov models are similar to other Markov models developed in the reviewed literature (Agrawal et al., 2009).

5.3 Discussion on Degradation Curves

This section provides discussion and comparisons between the different types of degradation curves developed from the three prediction models. Along with comparisons of the curves, the results from a small study investigating the effects of rebar protection on deck degradation are also presented.

5.3.1 MLP versus ENN Curves

Due to the different predictive ability of the MLP and ENN models the degradation curves produced by both vary in some respects. As previously discussed, the ENN models had improved predictive ability, especially in regards to identifying damage. The ENN models were also able to predict ratings of 8 and 9, which as will be seen in the following sections, manifests itself in the developed degradation curves.

5.3.1.1 Unique Bridge Curves

Figure 5-9 shows the predictions from both the ENN and MLP models for bridge 09109035000S130. For illustration purposes, the MLP and ENN predictions are slightly offset from the true predicted NBI rating to allow for easy comparison. As seen in the figure, the ENN and MLP models predict similar, but slightly different conditions throughout the life of the bridge. The ENN model maintains the condition of the bridge deck at the highest predicted rating for longer than the MLP model. The ENN model also predicts lower ratings for older ages of the bridge deck. This trend is likely due to the superior ability of the ENN model to accurately predict higher and lower ratings.

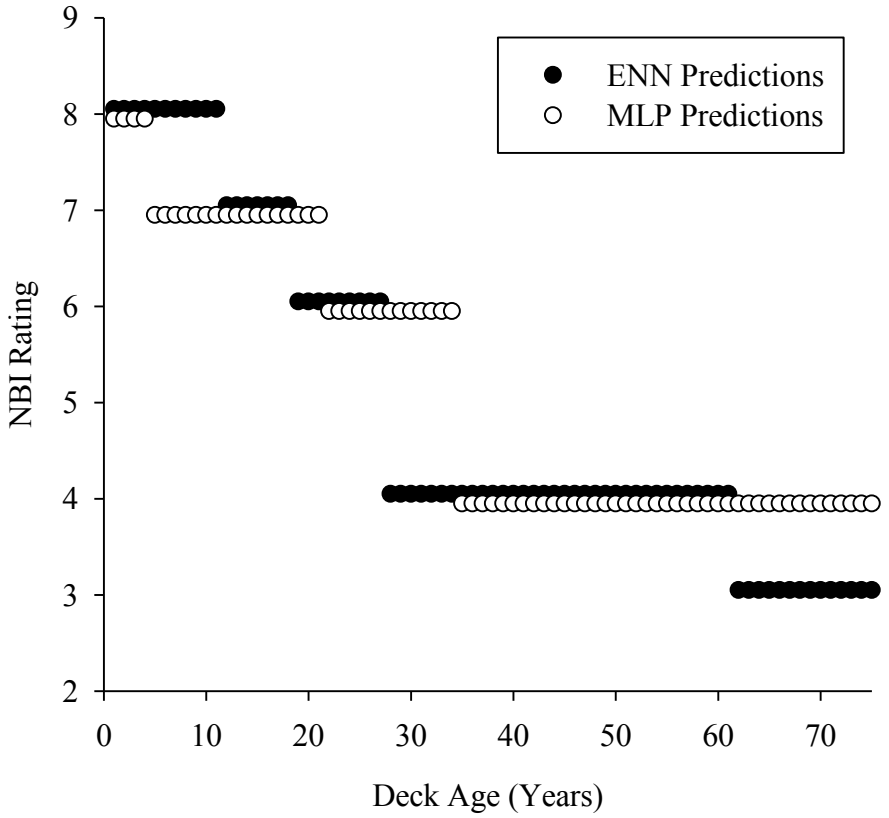


Figure 5-9: ENN and MLP predictions for bridge 09109035000S130

Figure 5-10 offers a comparison of the logistic fits of the predictions made by the MLP and ENN models. As seen in the figure, the fitted logistic curve to the ENN model predictions has a steeper slope in the early ages of the bridge decks life. The curve fitted to the ENN predictions also plateaus at a lower rating than the MLP fit, which is expected due to the ability of the ENN model to make lower predictions by the ENN.

Overall, the ENN model predicts higher early age ratings and lower ratings for older bridge deck ages. This leads to a slightly steeper degradation prediction by the ENN model. These trends were seen in multiple bridge decks that were examined.

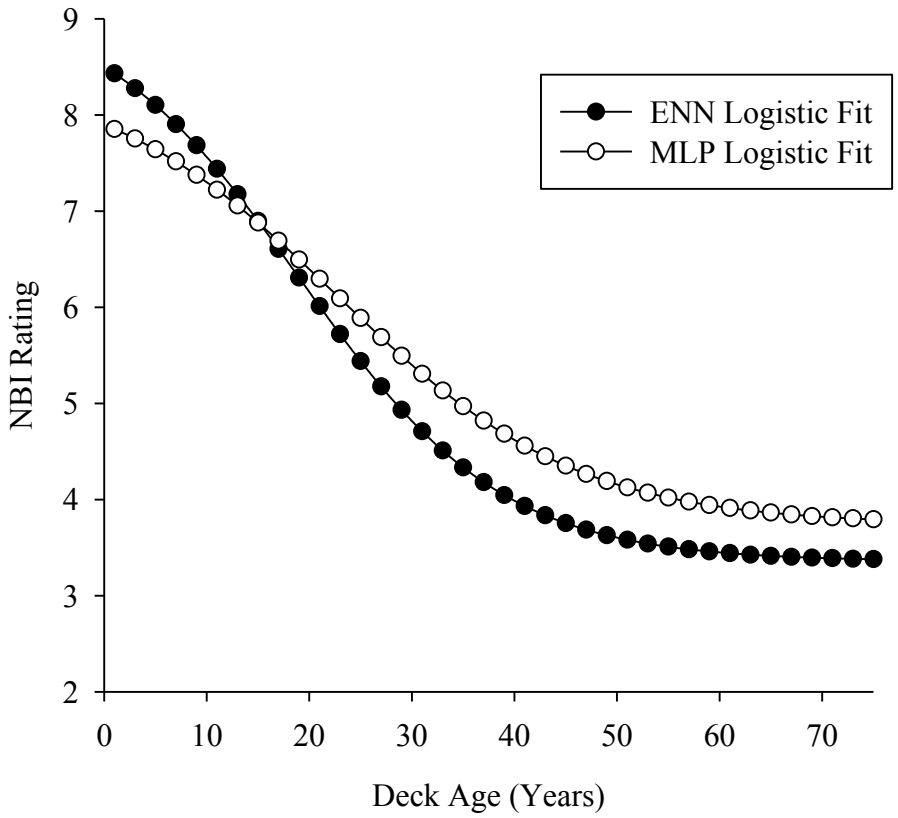


Figure 5-10: MLP and ENN logistic fit curves

5.3.1.2 Generalized Network Curves

The generalized network curves developed using the MLP and ENN models are shown in Figure 5-11 and Figure 5-12, respectively. Both curves share some interesting features. First, the curves both appear to have three distinct phases of deck rating change, indicative of different degradation process over time. The first phase, from years 0 to 10, shows some level of distress is taking place even though the deck is very young. This initial distress is followed by about 10 years of relative latency and then followed by a more rapid degradation after about 20 years of age. Both curves also show a widening of the confidence bounds around 30 years of age. As discussed in more detail in Section 5.3.3, one possibility for this is the increased variance in predictions made for decks with ECR past 30 years of age. The increased variance occurs because inspection records for decks with ECR of ages 30 and above do not exist in the database. As described in Section 4.3.1, these records were removed because ECR was not used in bridge

decks until the late 1970's. The existence of old bridge decks with ECR was determined as erroneous, and thus they were removed. While the removal of these records provided a better training set for the ANN models, it limited the age for which accurate predictions can be made for ECR decks. This lack of accuracy is manifested in the increase in variance in the generalized curves for decks of age 30 and above.

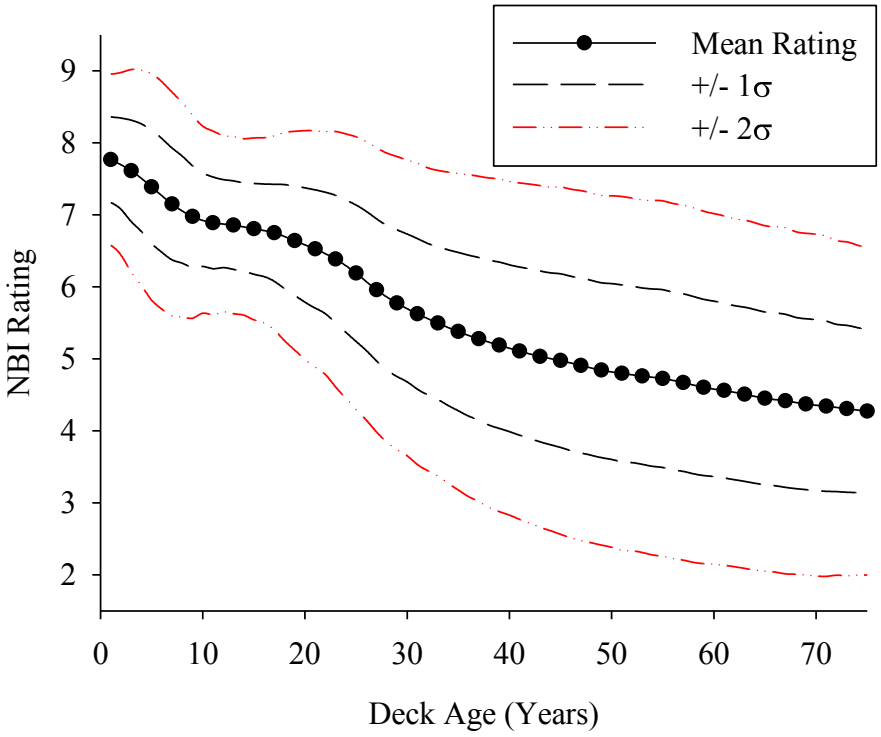


Figure 5-11: Generalized curve from MLP predictions

Another potential explanation for the increase in variance in the 30 to 50-year age range is the possibility that there is more variability in the conditions for bridges in this age range. For very early ages in the bridge's life, it is expected that almost all bridges will have very good condition, thus lowering the variance. However, once degradation processes begin to cause damage in the concrete, every bridge is going to degrade at a different rate. Because the condition of each bridge is degrading at a different rate, there is inherently more variability a network wide condition evaluation. Thus, the increased variance from 30 to 50 years in the ANN generalized curves may be indicative of a real increase in variance in condition ratings for this age range.

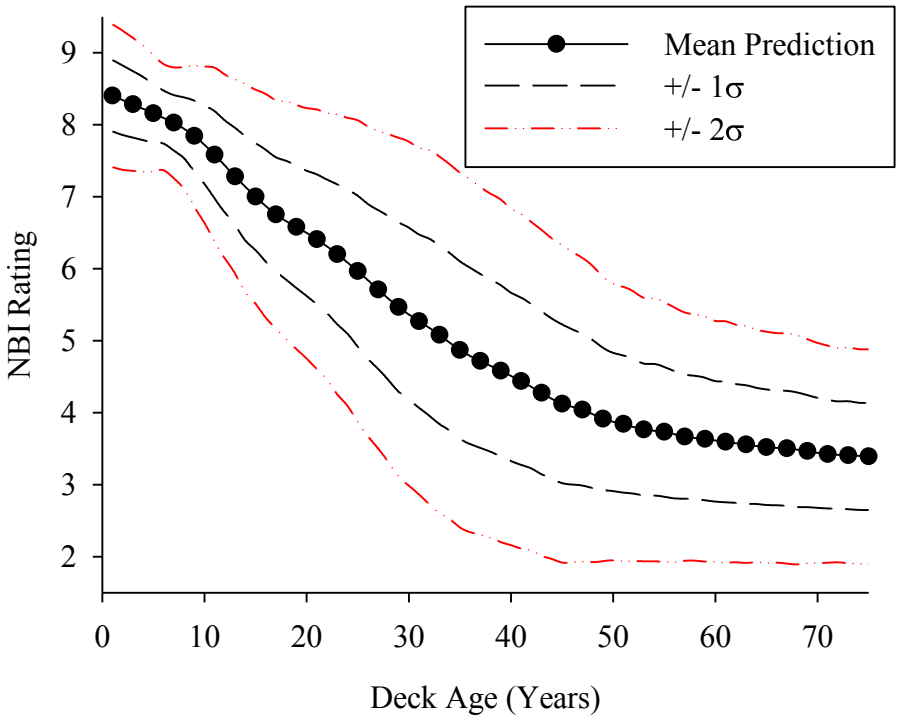


Figure 5-12: Generalized curve from ENN predictions

Upon comparison, the MLP and ENN generalized curves also show some differences. First, the ENN models provide higher mean ratings for early deck ages. This is likely due to the improved ability of the ENN models to predict ratings of 8 and 9. The ENN models also predict lower ratings as the deck age increases. As with the high ratings, it is speculated that this is due to the improved ability of the ENN models to predict lower ratings. Because the ENN generalized curves have both higher ratings for young decks and lower ratings for old decks over the same time period, the rate of deterioration as predicted by the ENN models must be faster at some point over the decks life. From the figures, it can be seen that the generalized ENN curve shows an increase in the rate of deterioration from ages 10 to 50. The confidence band for the MLP generalized curves are also larger than those produced by the ENN. While this observation cannot speak to the predictive ability of the models, it does indicate that the MLP model is predicting more variability in condition than the ENN model.

5.3.2 ANN versus Markov Degradation Curves

The generalized degradation curves allow the average condition of a network of bridges to be assessed, which enables a comparison between the degradation curves developed from the ANN and Markov models. Figure 5-13 compares the degradation curves developed using the MLP model and each of the three Markov transition probability matrices presented in Section 4.5. The curves developed using the Markov transition matrices assume an average initial network condition of a 9. This assumption was made in order to provide the degradation prediction of a newly built network of bridges in pristine condition. As seen in the figure, the Markov models predict a sharp decline in network condition early in the network's life. The Markov models also show a plateau in condition rating around 30 years of age. Compared with the MLP model, the Markov curves appear to have only two phases of degradation: an initially steep decline followed by a plateau around 30 years of age. Conversely, the MLP mean curve seems to capture different degradation phases as indicated by the change in slope.

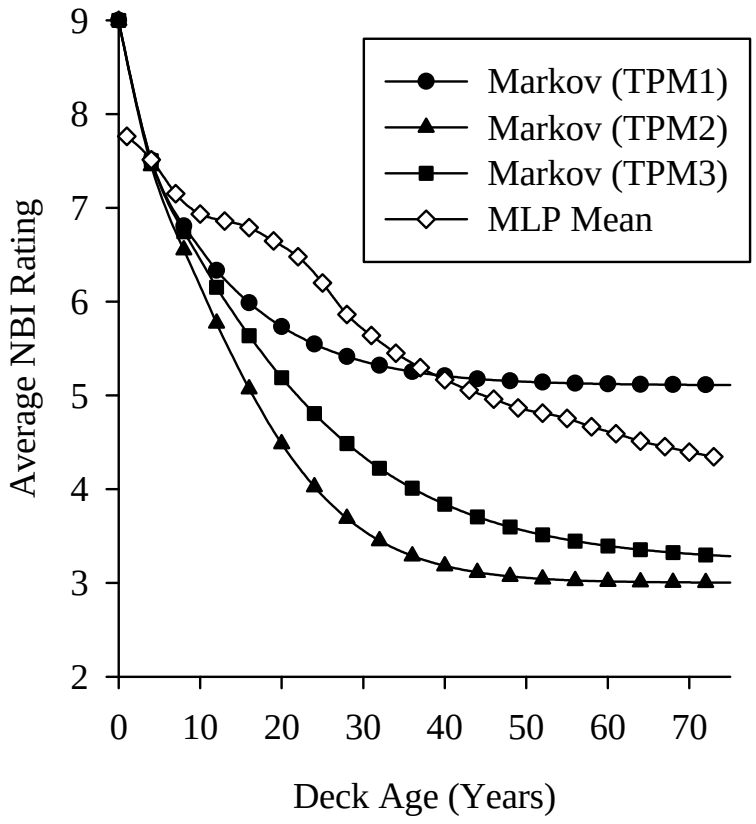


Figure 5-13: Markov versus MLP degradation curves

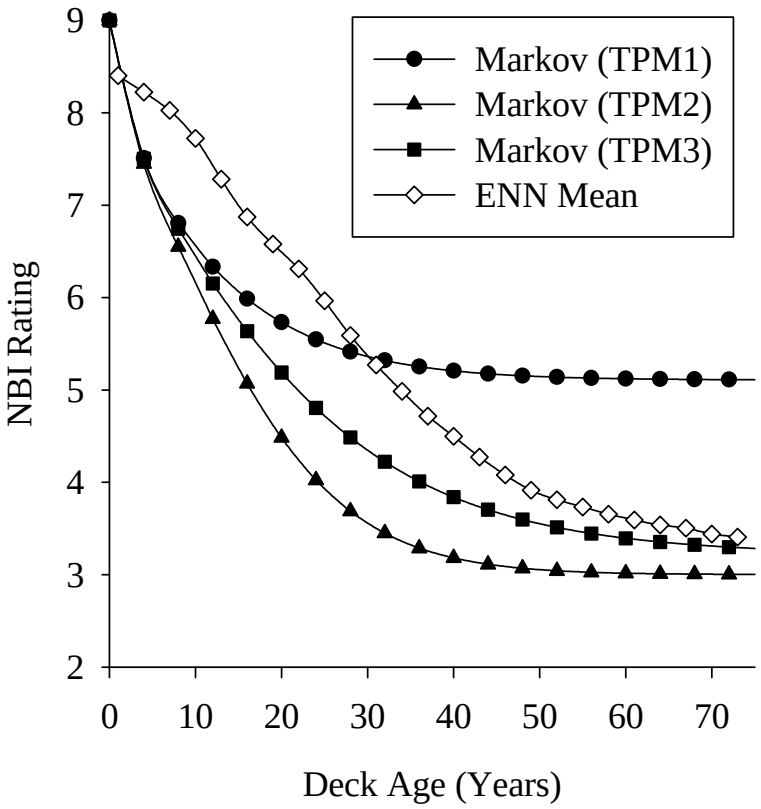


Figure 5-14: Markov versus ENN degradation curves

Figure 5-14 provides a comparison between the ENN and Markov degradation curves. As with the MLP comparison, this figure also shows interesting differences between the Markov and ANN models. The ENN degradation curve shows slower decline in condition in the early ages of the network. In the middle ages of the network, the rate of deterioration of both models is similar. However, the Markov model predicts a plateau in condition several decades before that of the ENN model. It should be noted that the Markov model based on TPM1 is shown here only for discussion purposes since it was shown in Section 4.5.1 that this transition probability matrix had not been removed of all records with jumps in condition due to improvement, which violates proper preparation of a Markov model. Thus, TMP1 is shown only to illustrate the effect of the variability of bridge rating in the study. The Markov model based on TPM1 was developed using the exact same data as the ENN model. Though there were very few instances of these improvement jumps, the effect on the resulting TPM1 Markov degradation curve is significant while the ENN model does not suffer from the same issue. Thus, the ENN is able to generalize and predict lower ratings despite some errors present in the database. The ability of

the ENN to overcome these errors is a known strength of artificial neural networks and can thus be an asset for the use of these methods for condition rating assessment and prediction when dealing with flawed data as that in the NBI database. From the shown comparison it can be seen that the different methods produce greatly different deterioration paths for the same network of bridges. However, it should be noted that Markov models require data preparation in ways slightly different than approached here which can lead to improved performance. When considering the Markov model for TPM3 it can be seen that the Markov and ENN generalized models provide similar performance later in the bridge life. This seems reasonable given that age dependence (and a violation of the assumption of Markovian chains) has been documented to be more important for early age degradation. Performance of the Markov models is also affected by the arbitrarily chosen initial network condition of 9. Nonetheless, it appears that the generalized ANN models are able to represent the degradation for a network of bridges and seem to capture distinct degradation phases, particularly during the eraly ages.

5.3.3 Rebar Protection Investigation

An issue of interest to the Michigan Department of Transportation (MDOT) is examining the benefits of epoxy coated rebar (ECR) in regards to mitigating concrete bridge deck deterioration due to corrosion. Corrosion of embedded reinforcing steel can be separated into two distinct phases: initiation and propagation. During the initiation phase, aggressive chemicals such as CO_2 and chlorides penetrate the concrete and diffuse to the reinforcement level. The initiation phase also includes the depassivisation period in which the passive layer of uncoated reinforcing bars is compromised due to the existence of aggressive chemicals. The duration of the initiation phase is dependent on multiple parameters such as concrete cover, the rate of chemical ingress, which is dictated by concrete quality, the climate in which the concrete element is located, and the critical chloride concentration. The period in which the corrosion of steel occurs and the effects are propagated through the concrete element is known as the propagation period. The model proposed by Weyers (1998) accounts for a free expansion period of corrosion products. In the model, part of the initiation phase accounts for the time necessary to develop of the necessary amount of corrosion products to cause internal stresses large enough to cause cracking. The continuation of cracking due to corrosion products is dictated by the rate of corrosion, which

is a function of the chloride content at the reinforcement steel depth, temperature, concrete cover resistance, and the number of years after corrosion initiation (Liu, 1996).

Several methods, including corrosion inhibitors, surface treatments, and reinforcement treatments can be used to prevent corrosion of reinforcing bars in concrete members. Corrosion-resistant reinforcing members, specifically epoxy-coated rebars, have been adopted in the state of Michigan as the main approach to protect structures against corrosion. Epoxy coating works to prevent corrosion in the following ways:

1. Resistance inhibition
2. Oxygen deprivation
3. Inhibition or aestivation

Multiple studies have been conducted in an attempt to quantify the effects of ECR with mixed results. Many studies provided evidence pointing to improved corrosion resistance using ECR (Samples and Ramirez, 1999; Fanous et al., 2000; and Lee and Krauss, 2003). However, the positive effects of ECR is highly dependent on the extent of cracking of the concrete member, concrete cover, the number of holidays (imperfections in the epoxy coating), and the thickness of the epoxy coating (Samples and Ramirez, 1999). An investigation of bridge decks in Virginia by Brown (2002) concluded that the service life of ECR decks was extended by about 5 years. The small improvements in service life lead to the conclusion that ECR is not a cost effective option for corrosion prevention in Virginia. The Long Key Bridge investigation in Florida provided a drastically different view of the benefits of ECR. The study revealed significant corrosion of ECR in marine structures after only 6 to 10 years (Sagues et al., 2001). The study concluded that the main cause of the excessive corrosion was likely due to disbondment and extensive damage of the epoxy coating.

The mixed results of field studies and the lack of service life data available makes it difficult to accurately assess the benefits of ECR. Because the subject is of very high interest to MDOT, the developed artificial neural network degradation models were employed in an attempt to draw conclusions regarding the presumed positive effects that epoxy coating has on concrete deck deterioration. Specifically, generalized degradation curves were developed for both the population of black steel decks and ECR decks in order to compare the average degradation of decks with the two types of rebar protection. Figure 5-15 shows the generalized curves for the

black and ECR populations as predicted by the ENN models. The ENN models were selected due to their superior predictive ability.

Because the database was modified to only contain ECR decks built after 1975, the oldest possible inspection records of ECR decks are 35 years old. Because the neural networks have no ECR training records beyond this age, predictions by the neural network for ECR decks beyond 35 years are guesses. Although the neural network may be able to draw inferences from other data available, there is no guarantee that the decks behave in the predicted manner because this data is simply unavailable at this point. Thus, predictions made past the age of available data are distinguished in Figure 5-15 by the light-colored diamonds.

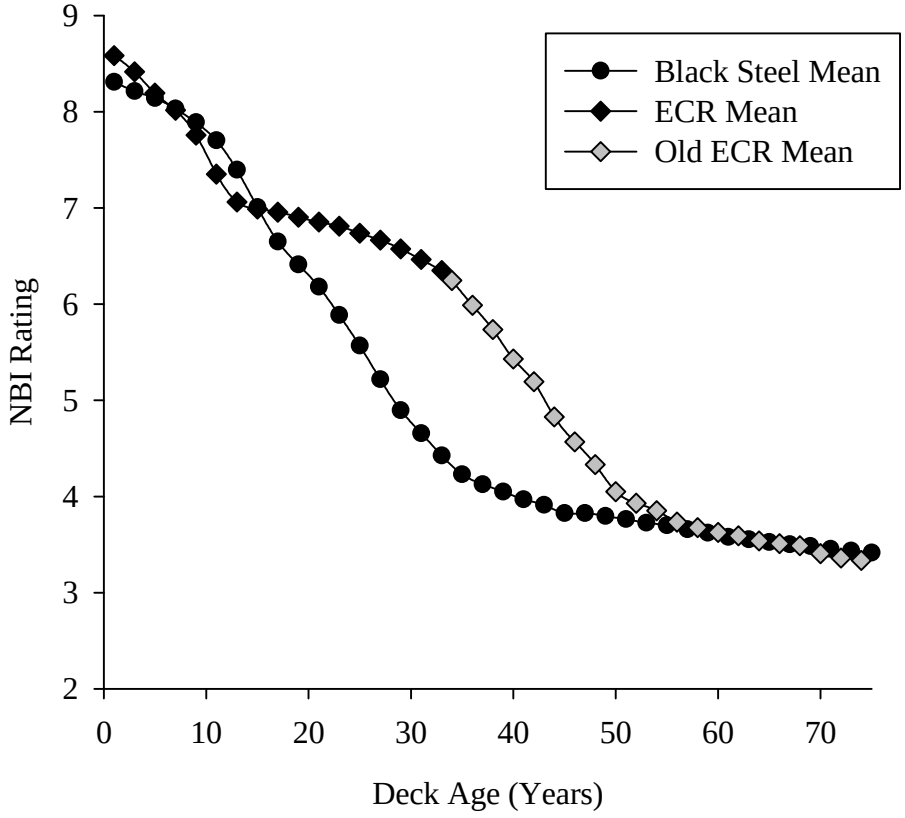


Figure 5-15: Black and ECR generalized ENN curves

As seen in Figure 5-15, the average rating of both the ECR and black steel decks are predicted to be almost identical for the first 15 years. This indicates that mechanisms other than corrosion are the driving forces of early age deck deterioration. The degradation of the ECR

decks deviates from that of the black steel decks around 15 years of age. The ECR decks are predicted to experience a much slower decline in rating for ages 15 through 35. It is speculated that this behavior is due to the protection against corrosion that is offered by the epoxy coating.

After 35 years of age, it can be seen that the *rate* of deterioration of the ECR decks is almost identical to that of the black steel decks. Because inspections for epoxy coated decks for ages beyond 35 do not exist, it cannot be stated with a high degree of confidence that this is the true degradation process for ECR decks. However, if this is indeed the true representation of the degradation process, it indicates that after the bridge deck accrues a certain amount of damage, the remaining degradation process is independent of rebar coating. This observation is consistent with the sentiment that ECR only delays the onset of corrosion, and does not inhibit damage once the process has begun.

To provide a comparison to the currently accepted method for degradation prediction, separate Markov models were developed to predict the condition ratings for black steel decks and ECR decks using the exact same data sets. The degradation curves developed using these matrices are shown in Figure 5-16 along with the generalized ENN model predictions.

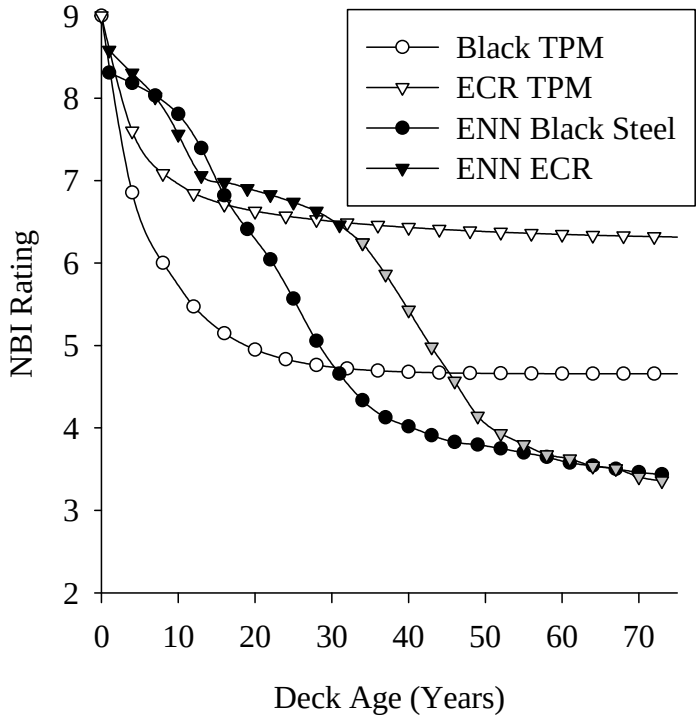


Figure 5-16: Markov and ENN curve comparison for black and ECR bridge decks

As seen in Figure 5-16, the Markov models produce very different degradation profiles. The transition matrix developed using only ECR bridge records plateaus at a rating of 7. This is probably due to an error in the transition probability matrices, which, for purpose of comparison was the same as used for the ENN model and contained records with bridge condition improvement. Proper definition of the probability transition matrices for the Markov model are expected to lead to degradation curves that predict ratings below 4 and thus more consistent with the ENN curve results. However, the Markov model curve will most likely still not be able to capture different rates of degradation that may be produced by the effects of ECR.

5.4 Summary on Degradation Curves

The developed ANN models were used to model both project level and network level degradation. At the project level, the unique degradation of a bridge deck was predicted, and was fitted with a logistic curve to provide a continuous representation of the degradation process. The condition prediction of a network of bridges was statistically combined to provide a network level assessment of condition. Additionally, the statistically combined network level assessment was compared with the degradation predicted by the developed Markov model (see Section 4.5). The developed curves show that the ANN models indicate several different phases of degradation, which may be representative of true differences in the concrete degradation process. The Markov model results in a smoother degradation curve and thus cannot capture different degradation phases. Additionally, as shown in the study of the effects of ECR, the ENN model is able to identify changes in degradation due to different design parameters on the network level.

6 BRIDGE MANAGEMENT SYSTEM

6.1 General

As discussed in Section 2.4, a Bridge Management System (BMS) aims to optimize the allocation of maintenance, rehabilitation, and repair funds for one or several bridges or bridge components. Figure 6-1 provides a flowchart for the BMS process. The entire BMS consists of major components that include the models, constraints, and decision support. The model component includes a degradation model, a cost model, and an improvement model. The support model component includes the ANN models used to develop the degradation curves, and the information from the NBI database to support the ANN models. Also included in the BMS support models is the optimization platform. The constraints are supplied to the optimization platform and include budgetary, condition, user-defined, and DOT constraints.

The Michigan Department of Transportation currently uses the software Pontis (AASHTO, 2009) and their own unique Bridge Condition Forecasting System (BCFS) (MDOT, 2009) to manage bridge repairs. Both Pontis and BCFS use Markovian assumptions for element deterioration, which, as previously discussed, contain invalid assumptions in the context of structural degradation. Moreover, Markovian models are limited to providing predictions only for a network of bridges and are unable to provide accurate predictions at the project level. The ANN models developed in this research serve as alternatives to Markov degradation models and can be utilized as the integral deterioration model in a BMS. The proposed BMS uses genetic algorithms to search for the optimum rehabilitation strategy for the proposed network. Ten bridges from the Lansing, Michigan area serve as the trial network in this study. Due to the lack of post-repair degradation models available, a novel method was developed based on behavior observed in MDOT's National Bridge Inventory (NBI) database.

6.2 Overview of Bridge Management System

Figure 6-1 provides a schematic of a BMS. The main components in the system are the BMS models, the constraints, and the decision support. The BMS models include a degradation model, a cost model, and an improvement model. As discussed in Section 2.4, a bridge

management system can be developed unique to the owner's interests. Multiple different models (predictive, repair and cost) can be implemented, and different optimization schemes can be utilized to identify an optimum solution. Also, the constraints on the decision support system (i.e., optimization problem) can be customized to fit needs of the agency, or owner. The BMS developed in this research follows the structure and framework presented by Hegazy et al. (2004). The proposed model is similar in the optimization scheme, namely the genetic algorithm (GA) set up, as is discussed later in Section 6.8. The current BMS differs from the work by Hegazy et al. (2004) in the models utilized. The developed ensemble of neural networks (ENN) models was utilized as the degradation model, which differs from the Markovian approach used by Hegazy et al. (2004). Additionally, the improvement model in this research was developed based on an investigation of bridge repairs provided by the state of Michigan, as presented in Section 6.5. Further, a post-reconstruction degradation model was developed for the current research, as detailed in Section 6.6. Hegazy et al. (2004) simply utilized the same Markovian transition probability matrices to model degradation after a bridge deck was repaired. In summary, the BMS developed in this research used unique models for degradation, improvement, cost, and post-repair models based on information provided by MDOT.

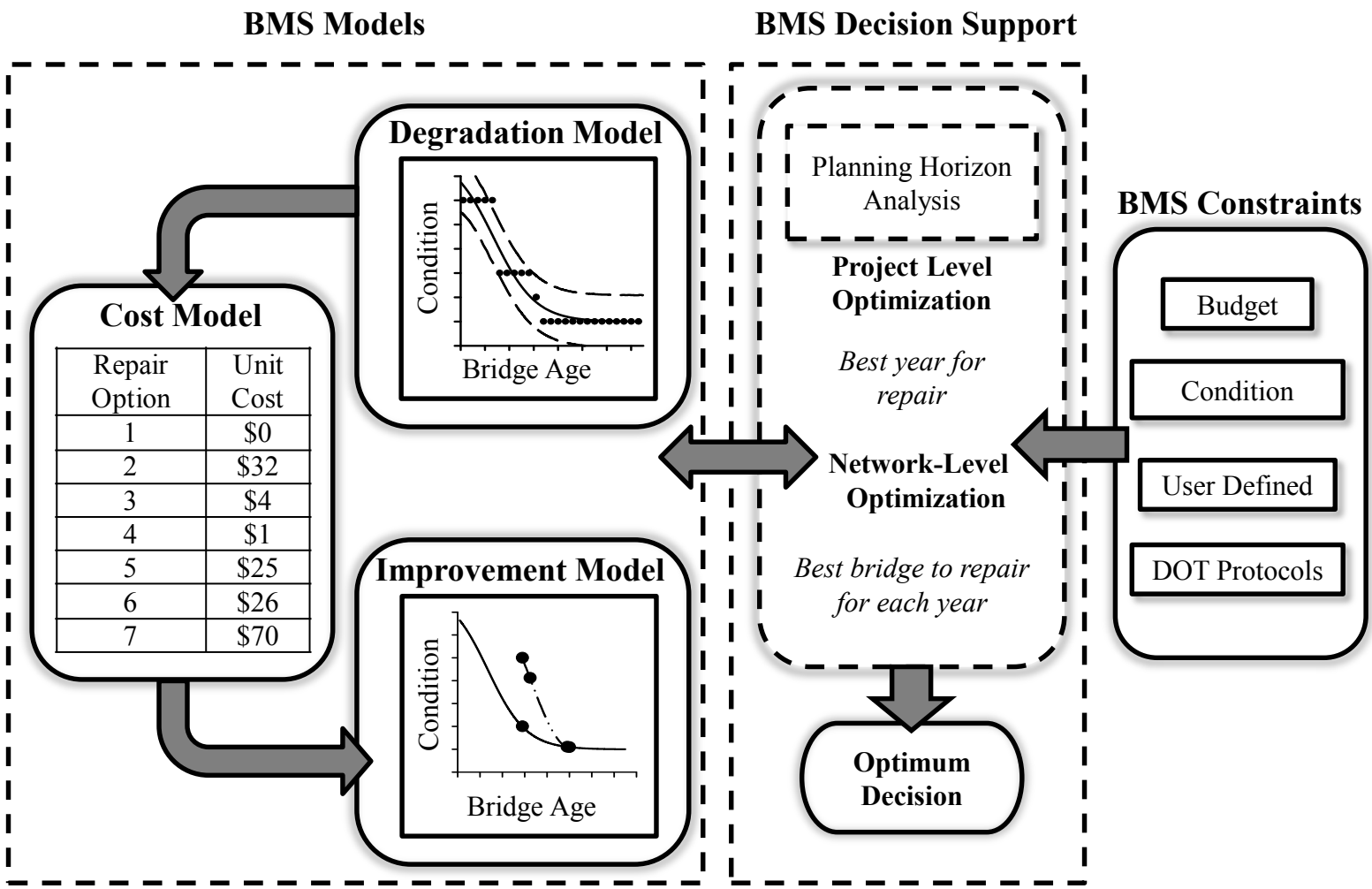


Figure 6-1: BMS flowchart

6.3 Selection of Bridges for Study

To the author's knowledge, reconstruction funding for bridges and roads in the state of Michigan is allocated by Region (groupings of state counties). In an attempt to realistically model the task at hand, the University Region was selected as the area of interest for the BMS Plan. The University Region includes the city of Lansing along with Clinton, Shiawassee, Eaton, Ingham, Livingston, Jackson, Washtenaw, Hillsdale, Lenawee, and Monroe counties. In the current database, there are 311 bridges in the University Region.

The current study is of a much smaller scale than that ordinarily considered by MDOT. One method of developing a smaller population is to randomly select bridges from the original population. However, using this method may lead to a population that is lacking in diversity with regard to parameters that are influential to deck degradation. Using a non-diverse population limits the applicability of the developed model. In order to use the model successfully on a larger scale, diversity in the sample population must be ensured. To select a diverse population of bridges, a method utilizing the parameters indicated by the study of the ENN connection weights (see Section 4.4) was developed. First, the continuous parameters that were identified as having the most influence on each NBI rating, as per Section 0 were recorded. The year built, ADTT, and age were identified as the unique parameters with the highest influence. Of these parameters, age is not important for the development of the degradation curves. Because this leaves only two unique parameters, the second most influential inputs were examined. In the examination of the second most influential parameters, three additional inputs were identified: ADT, span length, and the number of spans. In all, five parameters were used to select the bridge population. For each parameter, one bridge with the mean value of the parameter was selected from the University Region. The bridges were selected without replacement, meaning that if a bridge satisfied more than one parameter, it could only be used once in the population. To ensure equal representation of decks with black steel rebars and decks with ECR, this process of selecting three values for each of the five parameters was completed twice; with one group representing black steel decks and the other ECR. The population of ten bridges selected for the BMS study is presented in Table 6-1.

Table 6-1: List of bridges selected for BMS study

	Black Steel Decks	ECR Decks
Year Built	58158033000S020	30130071000B050
ADT	58158033000B052	19119042000S140
ADTT	33133035000S070	47147082000R020
Number of Spans	81181105000S090	76176024000S060
Max Span	23123061000S030	58158152000B041

6.4 Deterioration Model

As shown in the BMS flowchart (Figure 6-1), the ANN models, specifically the ENN model, are a decision support component. Although the ENN models were used to predict the condition of the bridge decks over their life, they were not used as the immediate degradation model. The actual degradation model is the fitted degradation curves to the ENN predictions (see Section 5.1.1). Only the deck surface rating model was used in the BMS and the fitted degradation curves of all ten bridges in the network were developed, with Figure 6-2 serving as an example. The degradation curves for all bridges in the network are provided in Appendix A.

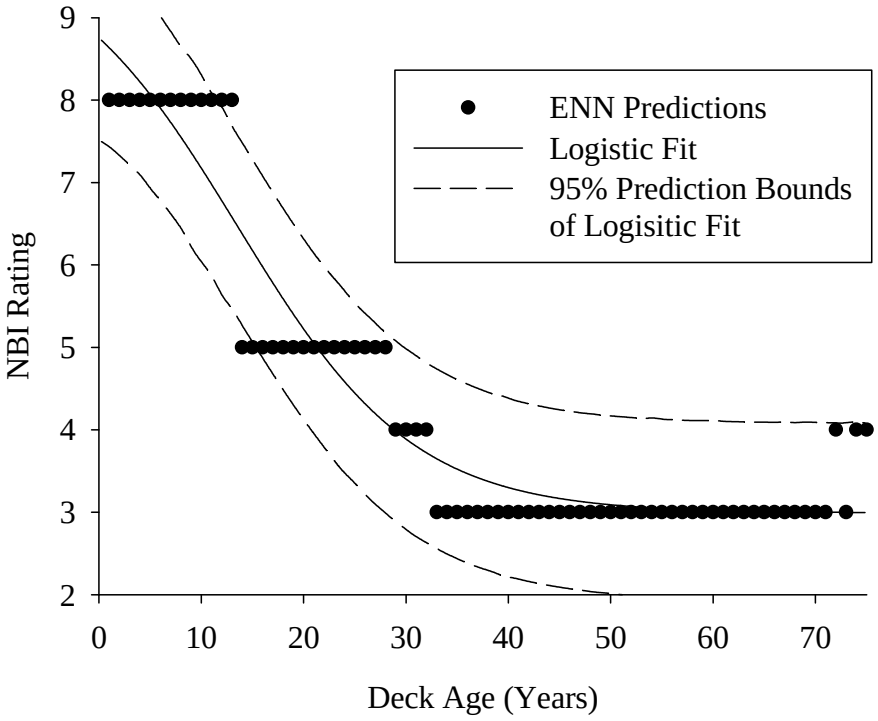


Figure 6-2: Bridge 58158033000S020 surface degradation curve

6.5 Repair Model

To make decisions regarding the selection of repairs, MDOT currently uses the bridge deck preservation matrix (BDPM) shown in Table 6-2 (Kelley, 2010). In this matrix, the type of reconstruction is selected based on the condition rating of both the deck surface and deck bottom. The BDPM also provides information on the *expected* results for each repair option. An estimate on when the bridge deck will require additional maintenance is also given. The BDPM provides information regarding crack sealing, epoxy overlays, deck patching, concrete overlays, HMA overlays, and deck replacements.

Table 6-2: Original MDOT bridge deck preservation matrix (BDPM)

DECK CONDITION STATE				REPAIR OPTIONS	POTENTIAL RESULT TO DECK BSIR		NEXT ANTICIPATED EVALUATION
Top Surface		Bottom Surface					
BSIR #58a	Deficiencies % (a)	BSIR #58b	Deficiencies % (b)		BSIR #58a	BSIR #58b	
≥ 5	N/A	N/A	N/A	Hold	No Change	No Change	1 to 8 years
				Seal Cracks/Healer Sealer			
	≤ 5%	> 5	≤ 2%	Epoxy Overlay	8, 9	No Change	10 to 15 years
	≤ 10%	≥ 4	≤ 25%	Deck Patch	Up by 1 pt.	No Change	3 to 10 years
4 or 5	10% to 25%	5 or 6	≤ 10%	Deep Concrete Overlay	8, 9	No Change	25 to 30 years
		4	10% to 25%	Shallow Concrete Overlay	8, 9	No Change	10 to 15 years
				HMA Overlay with water-proofing membrane	8, 9	No Change	8 to 10 years
		2 or 3	> 25%	HMA Cap	8, 9	No Change	2 to 4 years
≤ 3	>25%	> 5	< 2%	Deep Concrete Overlay	8, 9	No Change	20 to 25 years
		4 or 5	2% to 25%	Shallow Concrete Overlay	8, 9	No Change	10 years
				HMA Overlay with water-proofing membrane	8, 9	No Change	5 to 7 years
		2 or 3	>25%	HMA Cap	8, 9	No Change	1 to 3 years
				Replace Deck	9	9	40+ years

Although the improvement level and longevity of the repairs are estimates based on the experience of MDOT, personal communication with MDOT engineers has indicated that some are skeptical about the reliability of the BDPM. However, MDOT has unfortunately not been able to obtain a better estimate of the longevity and improvement level of the repairs. Because of the noted skepticism, a study was conducted to examine the accuracy of the anticipated improvements to the deck surface rating for each repair option given in the BDPM utilizing reconstruction records provided by MDOT. The repair options of deck patching, deep and shallow concrete overlays, HMA overlays, epoxy overlays, and deck replacements were investigated. Bridges receiving these repair options were identified and the inspections prior to and after the repair date were collected. The change in ratings from the inspections before and after the repair was calculated. Because the repairs may not have taken place on the exact date specified, three inspections after the repair date were considered in the calculation of the change in rating. The maximum difference in rating was taken as the overall change in rating and the improvement for that bridge from that repair option.

If just the deck surface rating is taken into account, the BPM provides a very large range of rating improvement for all repair options under investigation. For concrete overlays and HMA overlays, the MDOT BDPM states that the deck rating can be at a rating of 3, 4, or 5, and be restored to a rating of 8 or 9. This means that any one of these repairs could provide an improvement of anywhere between three rating points (5 to 8) and six rating points (three to 9). This range is very large, especially considering the discrete nature of the NBI rating scale. Epoxy overlays are expected to provide an improvement of two or three rating points, and deck patching is expected to provide a one point improvement. Histograms were developed to identify which improvement in rating occurred the most often in the provided dataset. Figure 6-3 shows the distributions of maximum deck surface rating changes calculated for deck patching, epoxy overlays, shallow concrete overlays, deep concrete overlays, and HMA overlays. The black bars indicate the number of rating points that the repair option may provide as per the BDPM.

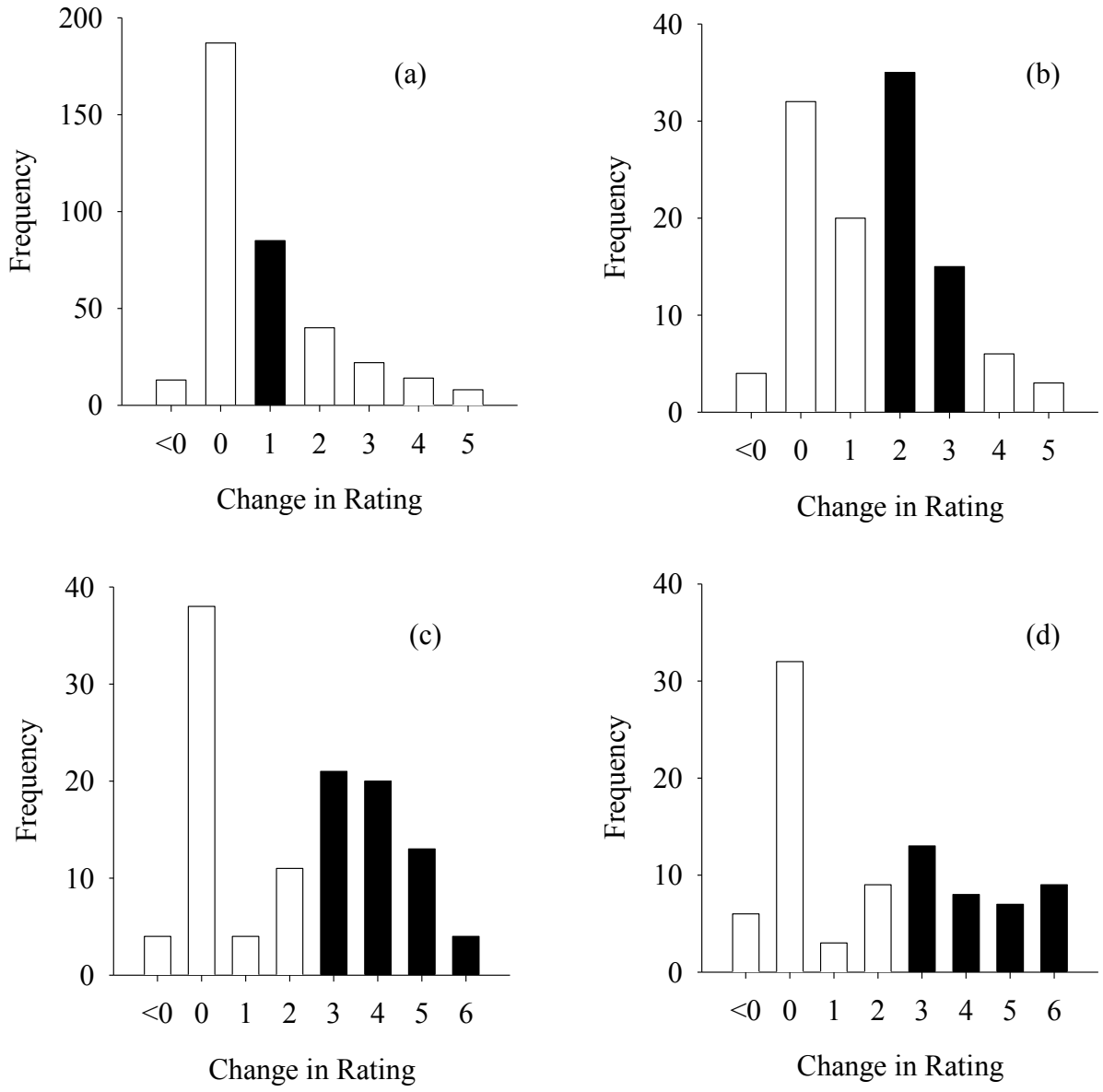
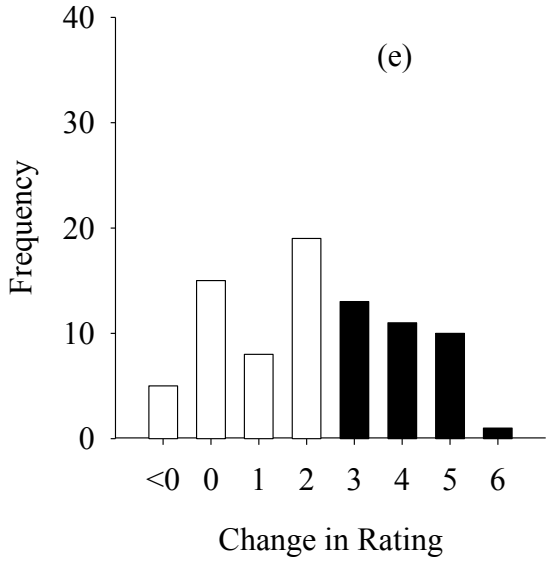


Figure 6-3: Repair of effect distributions on surface rating; (a) patching, (b) epoxy overlay, (c) shallow concrete overlay, (d) deep concrete overlay, (e) HMA overlay



As seen in Figure 6-3, the data indicates that the most common outcome from deck patching, shallow concrete overlays, and deep concrete overlays was no change of rating. For HMA and epoxy overlays, an improvement of two rating points was observed most commonly. The histograms show that the most commonly occurring increase in rating almost never matches with the MDOT BDPM predicted improvement (see Table 6-2). Of all the repair options, only epoxy overlays commonly produced rating improvements that match what is predicted by the BPM. For all of the repair options the data also show that there are instances where the condition of the bridge deck actually declines after the repair occurs. As this is very unlikely to occur, these instances are probably due to errors in recording the repair date in the database. It should be recognized, however, that if the deck bottom surface is severely cracked the rehabilitation method might not result in the expected rating increase.

Table 6-3 through Table 6-5 show the results of the study in terms of the statistical mean, median, and mode for each repair option. Because a decline in rating after a repair is unlikely, statistical measures without taking into account declines in rating were also calculated. Repair efforts causing no change in the rating were also disregarded in one set of calculations. Table 6-3 shows the statistics for the distributions using all the changes in ratings per reconstruction type. Table 6-4 shows the statistics for the rating improvement distributions without negative jumps in ratings. Table 6-5 contains the statistics for the distributions only containing positive changes in

rating. In all three tables, the expected improvement as given in the MDOT BDPM (see Table 6-2) is also provided.

Table 6-3: All rating change statistics

	HMA	Epoxy	Patch	Deep Concrete	Shallow Concrete
Mean	0.9	1.1	0.5	1.9	2.2
Median	0.0	1.0	0.0	2.0	3.0
Mode	2	2	0	0	0
MDOT BDPM	3 to 6	2 to3	1	3 to 6	3 to 6

Table 6-4: Non-negative rating change improvement statistics

	HMA	Epoxy	Patch	Deep Concrete	Shallow Concrete
Mean	1.4	1.2	0.6	2.2	2.3
Median	1.0	1.0	0.0	2.0	3.0
Mode	2	2	0	0	0
MDOT BDPM	3 to 6	2 to 3	1	3 to 6	3 to 6

Table 6-5: Positive rating change improvement statistics

	HMA	Epoxy	Patch	Deep Concrete	Shallow Concrete
Mean	2.6	2.2	1.9	3.7	3.5
Median	2	2	1	3	4
Mode	2	2	1	3	3
MDOT BDPM	3 to 6	2 to 3	1	3 to 6	3 to 6

Not included in the calculations above is the improvements observed by deck replacements. A deck replacement is expected to return the deck condition to a rating of a 9. Thus, the success of deck replacements was determined by calculating the percentage of bridge decks that achieved this rating after receiving a deck replacement. Due to the subjectivity in manual inspections, ratings of 8 were also included in this calculation. It was determined that 85% of decks improved to the condition of 8 or 9 after a deck replacement occurred.

Overall, this investigation shows that the expected increases in rating due to repair options outlined in MDOT's BDPM (see Table 6-2) are not being achieved. From this observation, it is

proposed that the BDPM be modified to reflect the results of this study. It is proposed to adopt the median rating change using non-negative results, as presented in Table 6-4. The anticipated improvements to the bridge deck in the BDPM were modified to reflect the results of this study as shown in Table 6-6. The BPM was also modified to only contain the repair options examined in this study. The time to the next anticipated maintenance was also simplified by assuming the minimum value provided in the MDOT BDPM.

Table 6-6: Proposed bridge deck preservation matrix (BDPM) for BMS Improvement Model

DECK CONDITION STATE		REPAIR OPTIONS	POTENTIAL RESULT TO DECK BSIR		NEXT ANTICIPATED EVALUATION
Top Surface	Bottom Surface		Top Surface	Bottom Surface	
≥ 5	> 5	Epoxy Overlay	1	No Change	10 years
	≥ 4	Deck Patch	0	No Change	3 years
4 or 5	5 or 6	Deep Concrete Overlay	2	No Change	25 years
	4	Shallow Concrete Overlay	3	No Change	10 years
		HMA Overlay with water- proofing membrane	1	No Change	8 years
≤ 3	> 5	Deep Concrete Overlay	2	No Change	20 years
	4 or 5	Shallow Concrete Overlay	3	No Change	10 years
		HMA Overlay with water- proofing membrane	1	No Change	5 years
	2 or 3	Replace Deck	9	9	40 years

6.6 Post Reconstruction Degradation

An essential part of any bridge management system is modeling how the deterioration of the element occurs after repair work has taken place. Modeling the deterioration of a bridge deck after reconstruction work is a complex task and the literature presented in Section 2.4 provides a review of current methods. From the review, it is noted that no universal method is accepted, and typically simplistic models are used to model this deterioration. A new method for modeling deck degradation after reconstruction is proposed for the BMS developed in this research. The post-reconstruction modeling is similar to work by Frangopol et al. (2001). This work could not be used directly as there is no discussion on how the values of the post-repair degradation slopes were attained. In this work, the post-repair degradation was modeled as a cubic function that is dependent on the repair option selected. The cubic function model was selected after several additional modeling options were identified and considered, as discussed in the following section.

6.6.1 Post-reconstruction Degradation Modeling Options

Because no post-repair degradation modeling technique is readily available, several options were considered before arriving at the proposed method. The first option involves simply increasing all condition ratings after the repair effort by the amount specified by the repair. Figure 6-4 provides an example of what the post-repair degradation would look like using this approach. Notice that the degradation curve experiences a shift upwards at a given point in time (repair rating increase) but the degradation curve is unchanged after the repair intervention.

The drawback of this method is that if the repair is selected later in life, the degradation curve declines in condition very little, and produces a plateau in condition, as seen in the figure. If the condition is simply increased by a set number of rating points, this plateau occurs at a higher condition rating, and, as seen in the figure, degradation past this condition is unlikely.

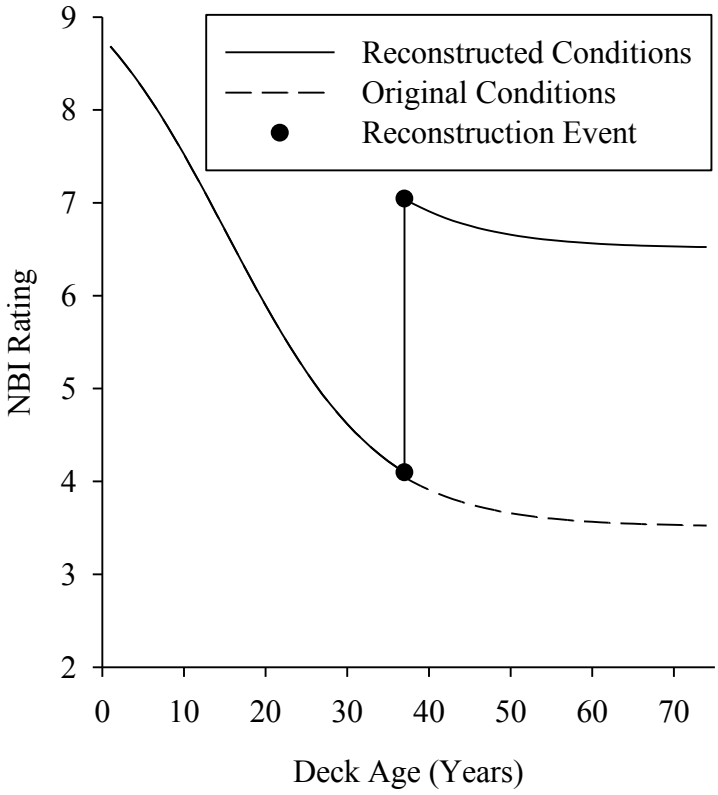


Figure 6-4: Post-repair degradation model option 1

Another option, and one that is has been used by Elbehairy et al., (2009) and Hegazy et al., (2004), is to restore the rate of degradation back to the original degradation rate at age 0. Figure 6-5 provides an example of what the post-repair degradation would look like if this option is chosen. Although there is no longer a plateau in condition at a high rating, the underlying assumption that the deck returns to its original degradation rate seems flawed. Repair efforts are intended to provide a patch to the physical manifestation of deleterious effects. A repair may slow the degradation rate, but it does not restore the bridge deck to its original condition. The effect of the repair depends on the underlying degradation mechanisms that cause the damage. Thus, providing the repairs does not restore the bridge deck to its original, pristine condition; and although the repair efforts appear to fix problems, they are not mitigating the actual causes of deterioration. Thus, the deterioration processes, such as corrosion, are still allowed to continue at their present rate. Because the repair efforts, excluding deck replacements, do not restore the bridge deck to its original condition, modeling the deterioration rate after a repair as such is not accurate.

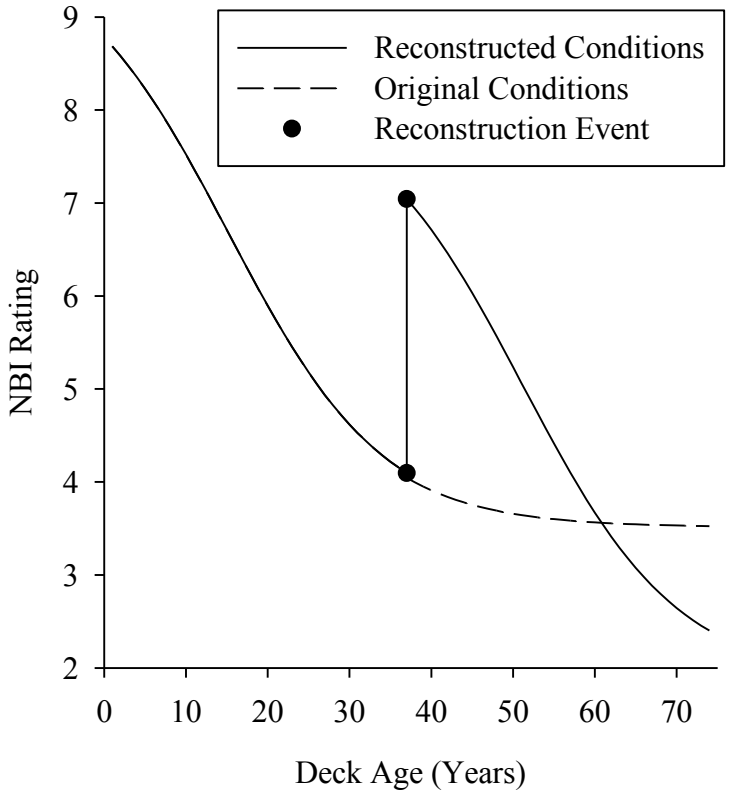


Figure 6-5: Post-repair degradation model option 2

6.6.2 Novel Post-reconstruction Degradation Model

Due to the shortcomings of the previous methods, an alternative post-repair model is proposed for this research. Degradation was modeled using cubic regression based on the immediate post-repair degradation rate of a repair effort and the estimated time until the next maintenance action as per the MDOT BDPM (see Table 6-2). The first step is to identify the immediate response in terms of degradation rate that a repair effort has on a bridge deck. This was achieved by utilizing records from MDOT bridges that had received repairs, or maintenance, and had received inspections immediately following the repair. Bridges with three inspection records after reconstruction were identified and the average change in rating between the first and second, and second and third ratings was calculated using Equation 6-1.

$$\alpha = \frac{1}{2}(\theta_1 + \theta_2)(\Delta_1 + \Delta_2) \qquad \textbf{6-1}$$

where α is the average change in rating, θ is the slope, and Δ is the time period. The subscripts indicate the change in rating between inspections. Change between inspections 1 and 2 is indicated by the subscript 1, and that between inspections 2 and 3 is indicated by the subscript 2. Table 6-7 shows the results from using Equation 6-1 for each reconstruction type. The time period was calculated as the total of $\Delta_1 + \Delta_2$ for a set of inspections. The average time period for all sets of inspections was calculated and is recorded in Table 6-7. As with the calculation of the improvement in rating from the reconstruction effort (see Section 6.5), difficulties ensued due to some inspection records showing the opposite expected behavior. In this case, some of the inspection records continued to show an *increase* in the condition rating after the initial improvement occurred immediately after reconstruction date. Thus, in a similar manner, the change in condition was calculated using all data, data without increases in ratings, and data with only decreases in ratings.

Table 6-7: Immediate rating changes using Equation 6-1

	All		Without Increase		Decrease Only	
	Average Rating Change	Average Time Period	Average Rating Change	Average Time Period	Average Rating Change	Average Time Period
HMA	0.45	3.39	-0.93	3.49	-1.43	3.42
Epoxy	0.00	3.36	-0.86	3.52	-1.32	3.53
Patch	0.08	3.33	-0.49	3.35	-1.31	3.47
Replace	0.44	3.38	-0.71	3.39	-1.29	3.38
Deep Concrete	1.11	2.92	-1.17	3.06	-1.34	3.07
Shallow Concrete	0.25	3.43	-0.88	3.43	-1.24	3.45

As seen in Table 6-7, all the ratings cannot be used to model the immediate degradation after reconstruction efforts as the average change in rating considering all inspections is positive for all reconstruction types. Due to this and the fact that a decline in rating is the natural progression with time, it is proposed to use the average change in ratings calculated without considering increases in rating. In conjunction with the immediate decline in rating, the method for modeling

the reconstruction uses a cubic regression to match the slope and rating of the original degradation curve at twice the time to the next anticipated evaluation as specified on the modified BDPM (see Table 6-2).

Figure 6-6 shows an example of the degradation of a bridge deck condition both before and after repair work was performed. The function *f(x)* is the degradation of the deck developed from fitting the ENN predictions for the bridge. The function *g(x)* is the degradation of the deck after the reconstruction work. *g(x)* is determined by performing a cubic regression using the points B, B', C and C', as shown in Figure 6-6.

Figure 6-6 schematically shows that reconstruction work occurs at point *A*, which has the coordinates (r,t_1). From the modified BDPM matrix (Table 6-6), the expected change in condition rating (Δr), the immediate decline in condition after the reconstruction work (θ_b), and the time to the next anticipated evaluation of the deck for work (Δt) can be obtained for a particular repair option. From this information, points B and B' can be determined. Point B represents the immediate change in condition rating due to the repair option selected, and has the coordinates (t_1, $r+\Delta r$). The location of point B' is determined using the post-reconstruction slope from the intervention matrix.

In the development of the post-reconstruction curve, the following conditions must be satisfied.

$$g(t_2) = f(t_2) \qquad \textbf{6-2}$$

$$g'(t_2) = f'(t_2) \qquad \textbf{6-3}$$

where for *g(x)*: t_2 is equal to $2\Delta t$, and for *f(x)*: t_2 is equal to $(t_1+2\Delta t)$. The conditions state that after a period of $2\Delta t$, the condition and the rate of deterioration of the deck post-reconstruction will match that of the original degradation curve without intervention. The period of $2\Delta t$ is twice the period suggested by MDOT for additional rehabilitation options to be considered (see Table 6-2 and Table 6-6); and it is assumed that at this time any benefits resulting by reconstruction or repair will cease to have an impact on the degradation of the bridge deck. Using the conditions outlined above, points *C* and *C'* can be calculated. A cubic regression is then completed using points *B*, *B'*, *C*, and *C'* to determine the equation for *g(x)*.

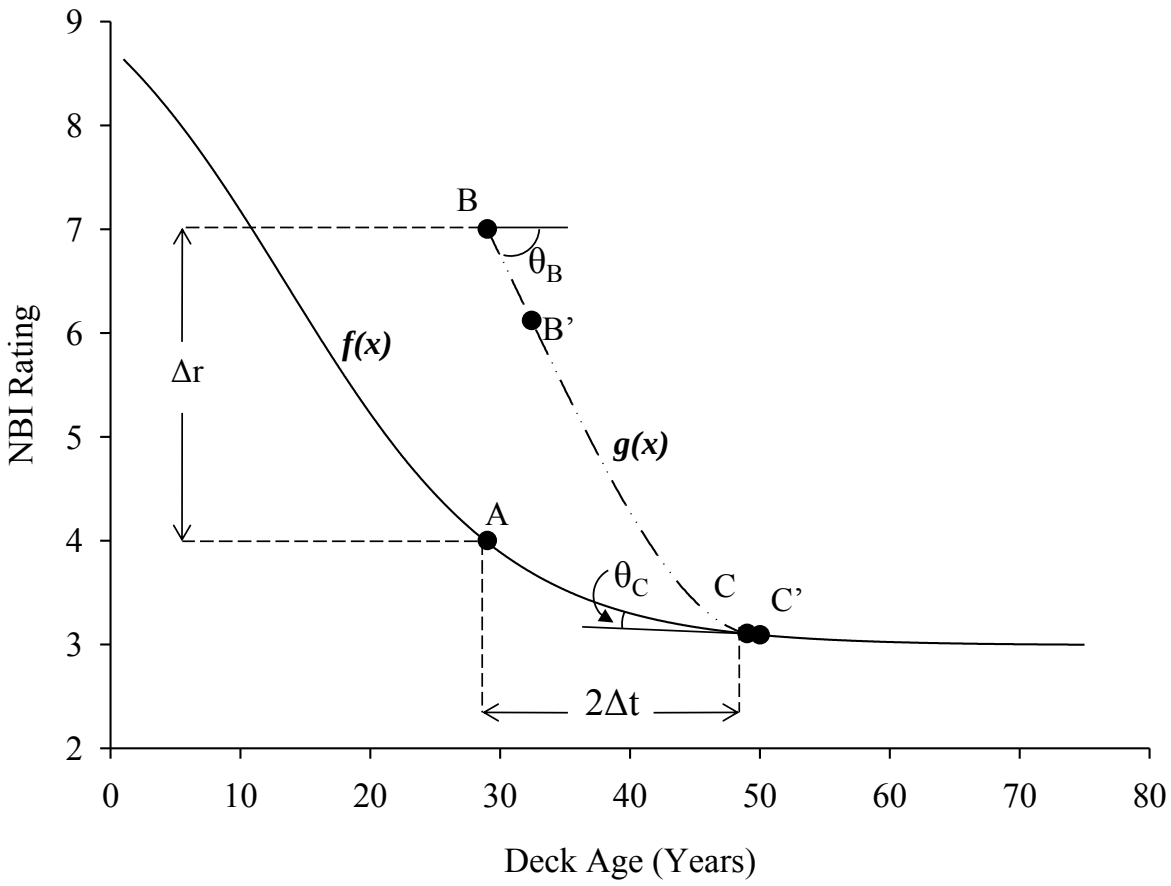


Figure 6-6: Degradation curve with post-rehabilitation deterioration, for bridge 58158033000S020

The process for creating the degradation after reconstruction work, *g(x)*, is summarized in the flowchart shown in Figure 6-7.

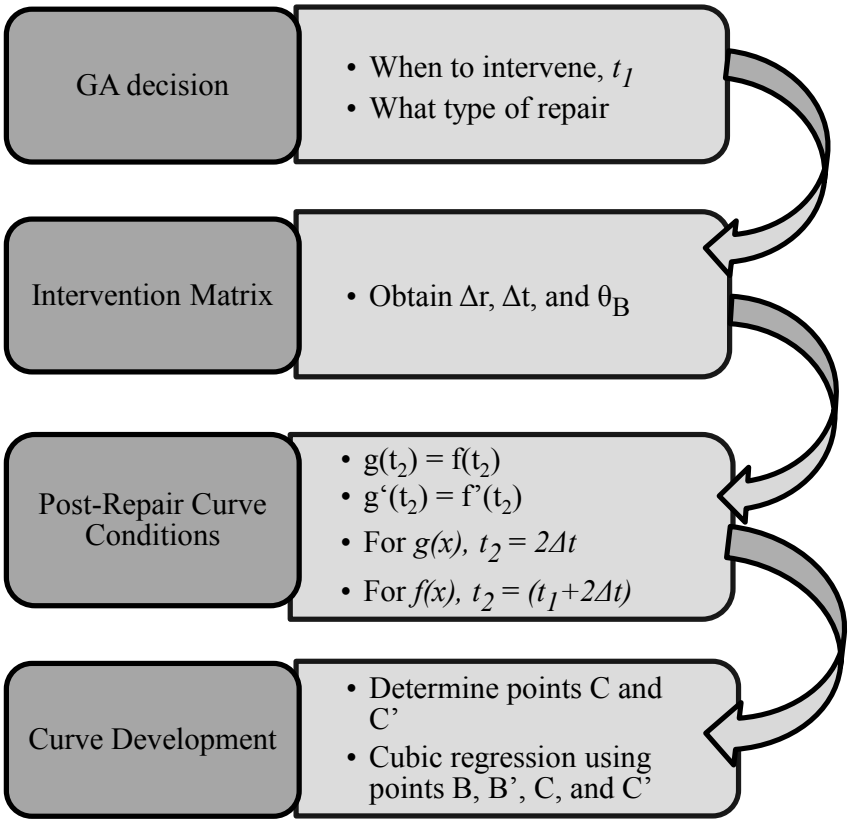


Figure 6-7: Post-reconstruction curve development flowchart

6.7 Cost Model

The initial cost of the network was calculated in terms of the value of the bridge decks only. The cost of a new bridge deck is listed as $70 per square foot by MDOT (Kelley, 2010). The initial cost of the 10-bridge network in question was calculated as $5,933,830. This cost is solely based on the cost of the bridge decks. Because degradation models were only developed for the deck element of the bridge, the depreciation of the network can only be determined in terms of the bridge deck. Thus, the initial cost of the bridge decks is the more appropriate parameter to use in this current study.

Table 6-8 shows the unit costs for each reconstruction option (Kelley, 2010). Costs are listed in unit prices. The prices listed are general costs and do not include additional costs incurred due to necessary traffic control during road projects. The costs also do not allow for any distinction to be made if a bridge is located in a more difficult access area, such as over water.

Table 6-8: Cost matrix (Kelley, 2010)

	Unit Price	Notes
Patching	$32/ft^2	Includes hand chipping
Epoxy Overlay	$34/yd^2 = $3.80/ft^2	
HMA/Bit Overlay	$1.25/ft^2	$5/ft^2 with water-proofing membrane
Shallow Concrete Overlay	$25/ft^2	Includes joint replacement and hydro
Deep Concrete Overlay	$26/ft^2	Includes joint replacement and hydro
Deck Replacement	$70/ft^2	Includes removal of old deck and new railings

Because the available repair costs are per square foot, it is necessary to know the width and length dimensions of each bridge. Because deck patching is a localized repair effort it is not applied to the entire deck. Thus, if deck patching is selected as the reconstruction option, it is necessary to calculate the area of the deck that requires patching. The area that requires patching is measured by the deck area that is spalled or delaminated, which is dependent on the deck condition rating. The NBI rating notes (MDOT, 2009) offer specific percentages for deck spalling and delamination for each rating. The percentages listed in Table 6-9 are used for calculating the damaged area for which patching is required, and in turn the cost associated with the patching.

Table 6-9: Spalled or delaminated percentages per surface rating (MDOT, 2009)

Rating	Spalled or Delaminated Area (%)	Adopted Value
9	-	-
8	-	-
7	-	-
6	< 2%	1%
5	2% - 10%	10%
4	10% - 25%	25%
3	> 25%	50%

6.8 Genetic Algorithms for BMS

The goal of the BMS is to identify the optimum repair strategies for a network of bridges over the planning horizon, which was taken as 5 years in this research. The optimum solution should minimize the cost of repairs while maintaining deck conditions at an acceptable level. The fitness function aims to minimize the total cost of the repairs for all bridges over the entire planning horizon, and is defined is as:

$$min \sum_{j=1}^{T} \sum_{i=1}^{N} A_i \, C_{R_{ij}} \qquad \textbf{6-4}$$

where N is the number of bridges, T is the number of years on the planning horizon, A is the area of the bridge in question, and C_R is the unit cost of repair R. While minimizing the fitness function the solution must also satisfy specified constraints. Two different constraints were applied to the GA. The first requires that all deck conditions be maintained above a '4' for every year on the planning horizon, as represented by Equation 6-5; where CR is the condition rating, and i and j are the same as above. A second constraint restricts the number of times a repair option can be selected on the planning horizon to one or zero. The effects of these two constraints on the repair strategy selected are presented in Section 6.9.1.

$$CR_{ij} > 4 \qquad \textbf{6-5}$$

While conventional GAs use a binary data representation scheme, an integer data representation scheme was selected in this research. An integer data representation removes the need for transforming the phenotypes into genotypes (see Section 2.6) and allows for easy translation between the real world space and the GA space. Integer data representation requires the use of special creation, crossover, mutation, and selection functions (see Section 2.6), all of which are predetermined if using the integer representation option in the Matlab (MathWorks, 2011) *ga* function. Using an integer data representation also allows for easy addition of the time variable, as represented by the planning horizon. Figure 6-8 shows how an individual is represented using the integer representation and five year planning horizon.

The total length of the chromosome is T x N years where T is the planning horizon and N is the number of bridges. Each bridge is represented by a 5-bit string. As shown in the figure, the first five numbers belong to bridge 1, the second belong to bridge 2, and the last five numbers belong to bridge *N*. Each bit represents one year on the planning horizon, and the value of each bit is representative of the repair option selected for that year. For example, the schematic in Figure 6-8 shows that for Bridge 1, the repair option of '1' was selected for years one and two, option '2' was selected in year three, option '5' in year four, and option '6' in year five. This data representation style allows for easy modification if a longer planning horizon is desired.

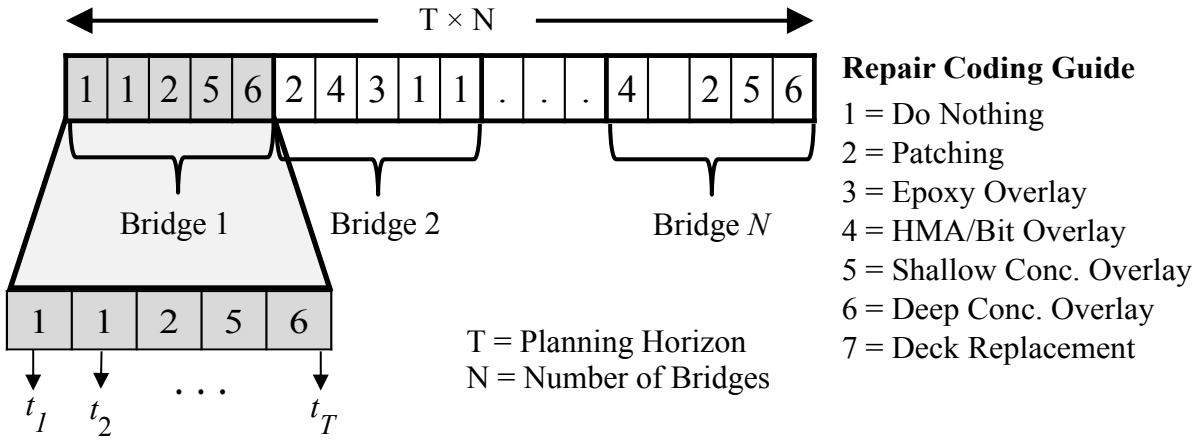


Figure 6-8: BMS chromosome example

6.9 Implementation of the Bridge Management System

The results of two separate BMS tests are provided. The first BMS was subject only to the constraint requiring the condition rating of every bridge in the network be maintained at or above a 4. The second BMS was subject to the rating constraint and also to the constraint limiting the number of repairs efforts for a bridge in a five year period to two of fewer. Both BMS tests were optimized using the same GA structure, population size, options, and fitness function.

6.9.1 Results

Table 6-10 presents the repair options selected by the GA when subject only to the rating constraint. The coding guide to the repairs is given in Table 6-11. The GA was terminated because the penalty value from the nonlinear constraint was below the tolerance value. As seen in the table, the optimum solution consists of multiple repair options selected for each bridge over the planning horizon. It is interesting to note that none of the repair options selected are a deck replacement.

Table 6-12 provides the costs associated with each of the repair options selected. As seen in the table, some repairs do not have an associated cost. This is because the repair option selected was patching, and the condition rating at the selection time was high enough that according to Table 6-9, no damaged area was present. The total cost of all the repairs over the five year planning horizon was $ 980,000.

Table 6-10: BMS selected repair selections; rating constraint only

	Year on Planning Horizon				
Bridge	1	2	3	4	5
1	6	4	2	3	2
2	3	5	4	4	1
3	3	3	4	1	4
4	4	3	3	2	2
5	3	4	2	2	2
6	4	5	3	2	2
7	4	2	3	2	2
8	3	2	4	3	3
9	4	3	2	3	1
10	3	2	4	2	4

Table 6-11: Repair Coding Guide

1 = Do Nothing
2 = Patching
3 = Epoxy Overlay
4 = HMA Overlay
5 = Shallow Concrete Overlay
6 = Deep Concrete Overlay
7 = Deck Replacement

Table 6-12: BMS repair costs; rating constraint only

	Year on Planning Horizon				
Bridge	1	2	3	4	5
1	$ 207,000	$ 10,000	$ 3,000	$ 30,000	-
2	$ 22,000	$ 143,000	$ 7,000	$ 7,000	-
3	$ 22,000	$ 22,000	$ 7,000	-	$ 7,000
4	$ 6,000	$ 19,000	$ 19,000	-	-
5	$ 91,000	$ 30,000	-	-	-
6	$ 2,000	$ 39,000	$ 6,000	-	-
7	$ 9,000	-	$ 26,000	-	-
8	$ 30,000	-	$ 10,000	$ 30,000	$ 30,000
9	$ 10,000	$ 30,000	-	$ 30,000	-
10	$ 47,000	-	$ 15,000	-	$ 15,000

Figure 6-9 shows the degradation curve for Bridge 1 when the selected repairs are modeled. The solid line represents the degradation curve with repairs, and the dashed line is the degradation curve if not repairs had occurred. The post-repair degradation curves for all other bridges in the study are shown in Figure B-1, located in Appendix B. As seen in the figure, the repairs provide a significant improvement in rating during the planning horizon. However, the figure also shows steep decline in condition shortly after the end of the planning horizon, indicating that the repairs provide are very limited in longevity. Upon closer examination, it was determined that the algorithm used to calculate the post-reconstruction degradation curve should be modified. Currently, the algorithm uses the last selected repair in the planning horizon to model the degradation in the years following. However, this assumption is flawed if the last

repair selected is patching, as is the case for Bridge 1, shown in Figure 6-9, because patching is a localized repair. Patching is only completed on damaged areas are of the bridge deck, and thus it is not appropriate to model the degradation of the entire deck based of the degradation characteristics of the patching repair. Thus, the algorithm should be modified to provide the post-repair degradation using the last repair on the planning horizon that was applied to the entire deck. This modification in the algorithm has been left for a future study.

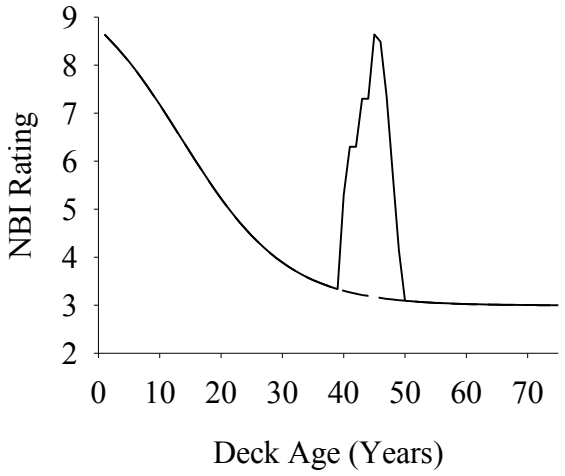


Figure 6-9: BMS GA results for bridge 58158033000S020; rating constraint only

Table 6-13 provides the condition ratings of the repaired bridges over the planning horizon. As seen in the table, all bridge decks are maintained at a condition above a 4 for all years on the planning horizon. Additionally, the table shows a large number of bridges at very high conditions. This is likely due to the GA selecting multiple repairs for the bridge decks.

Table 6-14 shows the optimum repair strategy selected by the GA when both the rating and the number of repairs selected is constrained. The GA was constrained to selecting two or fewer repair options on the planning horizon. As seen in Table 6-14, this constraint was satisfied (a '1' refers to no repair selected). Additionally, it was observed that the repair strategy consists almost entirely of full deck replacements. The selection of the full deck replacements likely occurs because a full deck replacement provides the largest improvement to the condition rating. However, a full deck replacement is the most costly repair available, and results in much higher costs per bridge, as seen in Table 6-15. The total cost of this repair strategy is $7,330,000.

Table 6-13: Condition ratings over the planning horizon; rating constraint only

	Year on PH				
Bridge	1	2	3	4	5
1	5	6	6	7	7
2	5	8	9	9	9
3	5	6	7	7	8
4	6	7	8	8	8
5	6	7	7	7	7
6	9	9	9	9	9
7	7	7	8	8	8
8	9	9	9	9	9
9	8	9	9	9	9
10	9	9	9	9	9

Table 6-14: BMS selected repair selections; rating and selection constraint

	Year on PH				
Bridge	1	2	3	4	5
1	7	1	1	7	1
2	7	1	1	7	1
3	7	1	1	1	3
4	1	7	7	1	1
5	1	1	1	1	7
6	1	7	1	3	1
7	1	7	1	1	1
8	1	1	4	1	1
9	7	7	1	1	1
10	3	1	1	7	1

Table 6-15: BMS repair costs; rating and selection constraint

	Year on PH				
Bridge	1	2	3	4	5
1	$ 557,000	-	-	$ 557,000	-
2	$ 399,000	-	-	$ 399,000	-
3	$ 411,000	-	-	-	$ 22,000
4	-	$ 342,000	$ 342,000	-	-
5	-	-	-	-	$ 1,673,000
6	-	$ 110,000	-	$ 6,000	-
7	-	$ 479,000	-	-	-
8	-	-	$ 10,000	-	-
9	$ 558,000	$ 558,000	-	-	-
10	$ 47,000	-	-	$ 858,000	-

Although this repair strategy is much more costly, because full deck replacements were selected, the longevity of the repair is considerably longer. As seen in Figure 6-10, the condition of the bridge deck remains above a rating of a 4 for the remainder of its service life. This behavior is consistent with all other bridges included in the study, as seen in Figure B-2. Although the conditions of the bridge decks remains high for many years after the selected repairs, the large cost of the repair strategy makes it unlikely candidate for an actual repair strategy for MDOT.

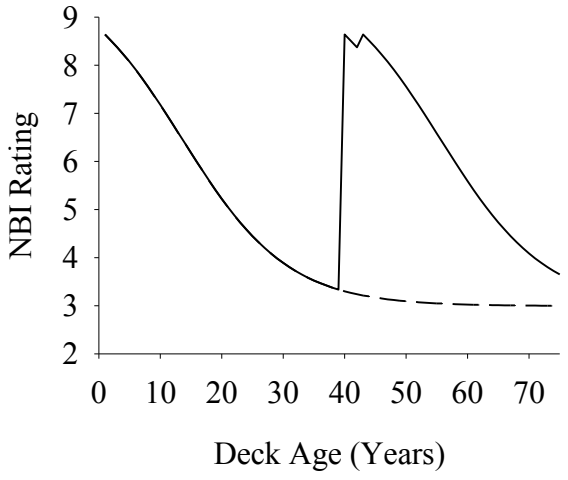


Figure 6-10: BMS GA results for bridge 58158033000S020; rating and selection constraint

6.10 Alternative Approaches, Discussion, and Future Work

The results of the 10-bridge study presented in Section 6.9.1 illustrate the ability of the GA optimization to be used to select repair strategies for a network of bridges. The results from two tests using different constraints also illustrate the ability of the proposed BMS to be tailored to fit the user preferences. However, genetic algorithms are only one approach to the optimization of repairs on infrastructure. Closed form solutions to optimizing when repairs should occur have been proposed by Haider and Dwaikat (2010). Their model determines when a repair effort should occur on a pavement to maximize the area under the degradation curve. The maximization of the area is analogous to the best repair effort over the lifetime of the pavement.

Although the GA model proposed in this research has produced promising results, refinement of the optimization problem and more sophisticated methodologies are possible and needed. Specifically, while the GA solution for the BMS framework is considered very promising, improvement on the definition of constraints is needed. For example, the solutions presented (see Table 6-10, Table 6-13, and Table 6-14) did not consider the constraint of repair sequence. That is, certain repairs are clearly not done after others. For example, an HMA overlay (repair 4) will not be done after a deep concrete overlay repair (repair 6). Also missing from the optimization formulation is the longevity of the repair. This means that it is expected that the bridge will not need intervention for a certain time after a repair has taken place, and clearly that time depends on the repair type. Redefinition of these constraints was not possible within the time constraints of the project but their definition is feasible and does not limit the applicability of the GA optimization strategy or the framework presented for the BMS.

Extensions to the BMS presented in this study could also include more constraints, a larger bridge network, and a more in depth study of the GA structure to identify the best performing parameters. Additionally, as discussed in Section 6.9.1, modifications to the post-repair degradation algorithm should be made to account for the localized effect of the repair option of patching.

7 CONCLUSIONS

While the National Bridge Inventory (NBI) database contains very valuable historical data concerning the condition of bridges across the country, the database itself is challenging to work with. The database contains large amounts of scatter, is subject to human error both in data entry and in the inspection process, and many relationships between the variables are not easily identifiable with simple statistical tools. However, the NBI database is publicly available and because the upkeep of data is mandatory for a state's Department of Transportation (DOT), the potential of the database is only growing. Additionally, utilizing the database in condition prediction models limits the need for costly sensors and equipment that are required in other condition assessment models.

Artificial neural network (ANN) models were selected to develop prediction models due to their ability to handle complex and nonlinear data. The ANN models were developed to predict the bridge deck condition rating from physical, geometric, and operation parameters of a bridge. A simple type of neural network, the multi-layer perceptron (MLP), was able to correctly classify condition ratings better than chance alone, and was able to identify 48% to 62% of decks in poor condition. Although the MLP models are elementary, they were able to overcome the database complexities and error better than a developed Markov model. Ensembles of neural network (ENN) models significantly improved the ANN predictive performance. A drastic increase in the ability to correctly classify conditions and to identify decks in poor condition was observed with the implementation of the ENN models. The ENN models were correctly able to identify decks with condition ratings indicating damage 84% to 86% of the time. Additionally, both the MLP and ENN models were able to predict the condition ratings with a variance comparable to visual inspectors in the field. This is a remarkable level of performance considering the added advantage of the ANN models to provide *future* predictions of condition. It is noted however, that the ANN models are general, and cannot provide the detailed information that a mechanistic model can. However, due to the large diversity in design parameters of a state's bridge inventory and the information maintained about each bridge in the inventory, the implementation mechanistic-based degradation models are currently not feasible for implementation in the degradation modeling and bridge management preservation strategy of a highway network by a state DOT.

The ANN models can be easily utilized to track the degradation of a bridge deck over its life. A unique feature of the ANN models is their ability to predict the degradation process at the project level, or for a single bridge deck. Conversely, commonly used techniques such as Markov models are limited to predicting the condition of a network. The degradation predictions by the ANN models for multiple bridges can however be statistically combined to come up with a generalized network degradation curves. When compared to the Markov degradation curves, the generalized ANN models show distinct trends in the decline of the bridge deck condition as the age of the bridge increases.

In addition to the degradation models developed, an investigation of the improvements provided by reconstruction repairs used by the Michigan DOT was conducted. The study found that the anticipated improvements in condition rating were not being achieved, and modifications were suggested and utilized in the developed improvement model.

The developed degradation, repair, cost, and improvement models were successfully utilized in a genetic algorithm (GA) based bridge management system (BMS). The models provided the degradation curves for a network of ten bridges in the Lansing, MI area, and the GA was charged with identifying the most cost effective solution for maintaining the network. The GA aimed to minimize the cost of all repairs over a five-year period while maintaining the condition of the bridge decks above a rating of a four for the planning horizon.

Overall, ANN models provide a new approach to model bridge deck conditions utilizing data that is readily available in the NBI database of state DOTs. The developed models provide better results than the currently used Markov models, and their real-world application potential was successfully demonstrated through implementation in a GA based BMS. The potential of ANNs as tools for inventory assessment and condition evaluation has been established, and continued development of ANN models for condition evaluation is recommended.

APPENDICES

APPENDIX A: BMS DEGRADATION CURVES

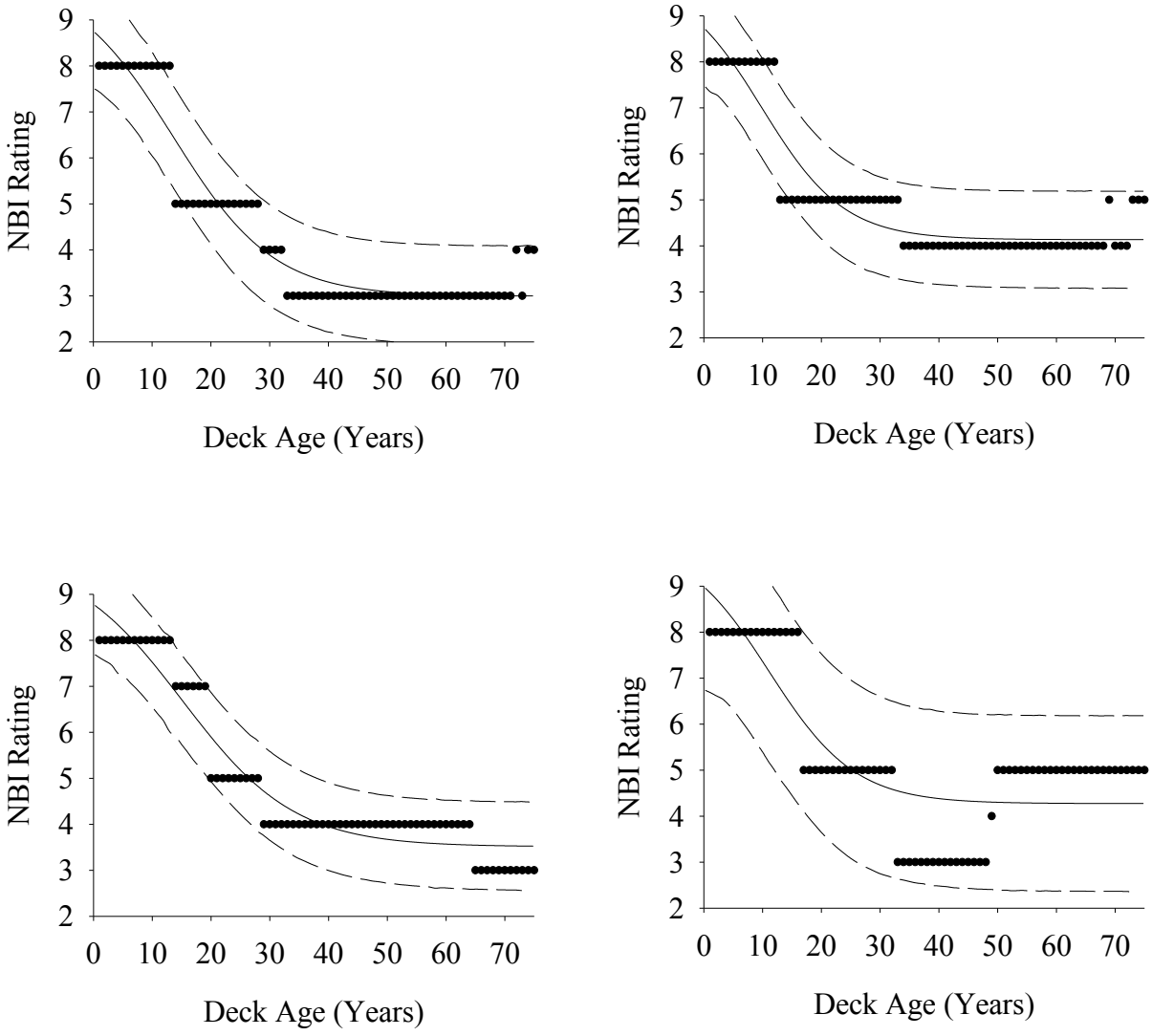


Figure A-1: BMS degradation curves

Figure A-1 (cont'd)

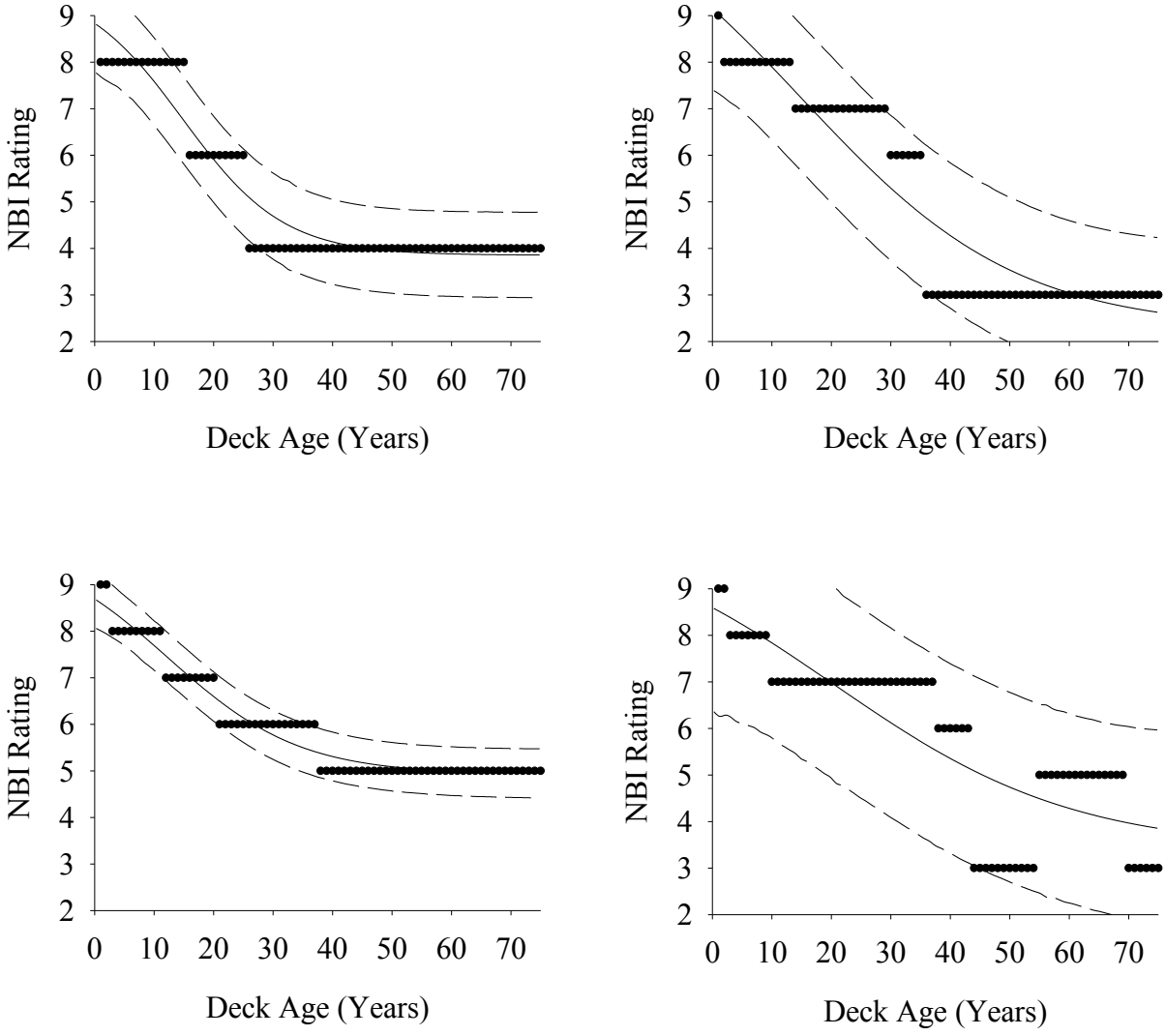
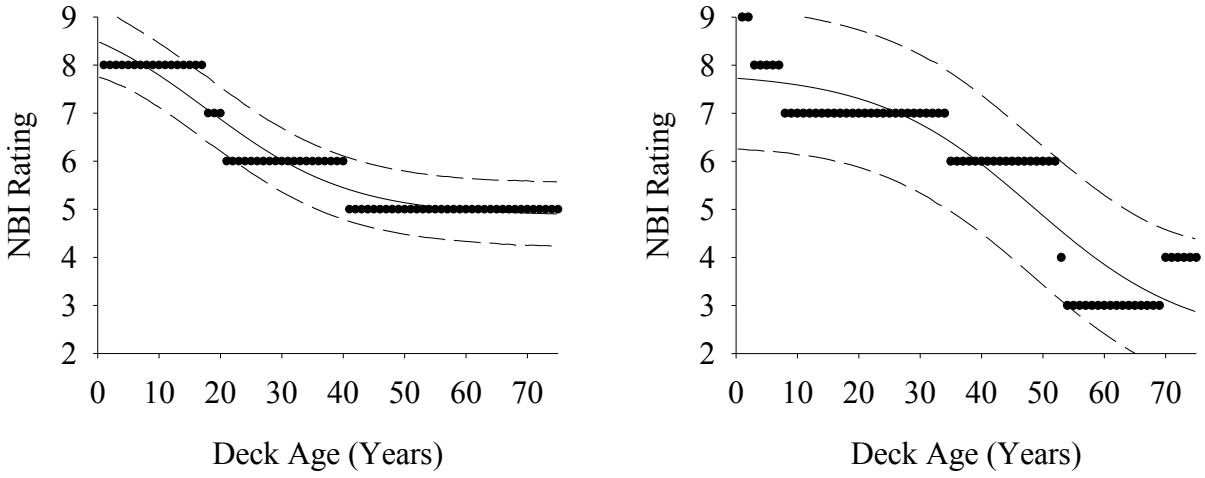


Figure A-1 (cont'd)



APPENDIX B: BMS RESULTING DEGRADATION CURVES

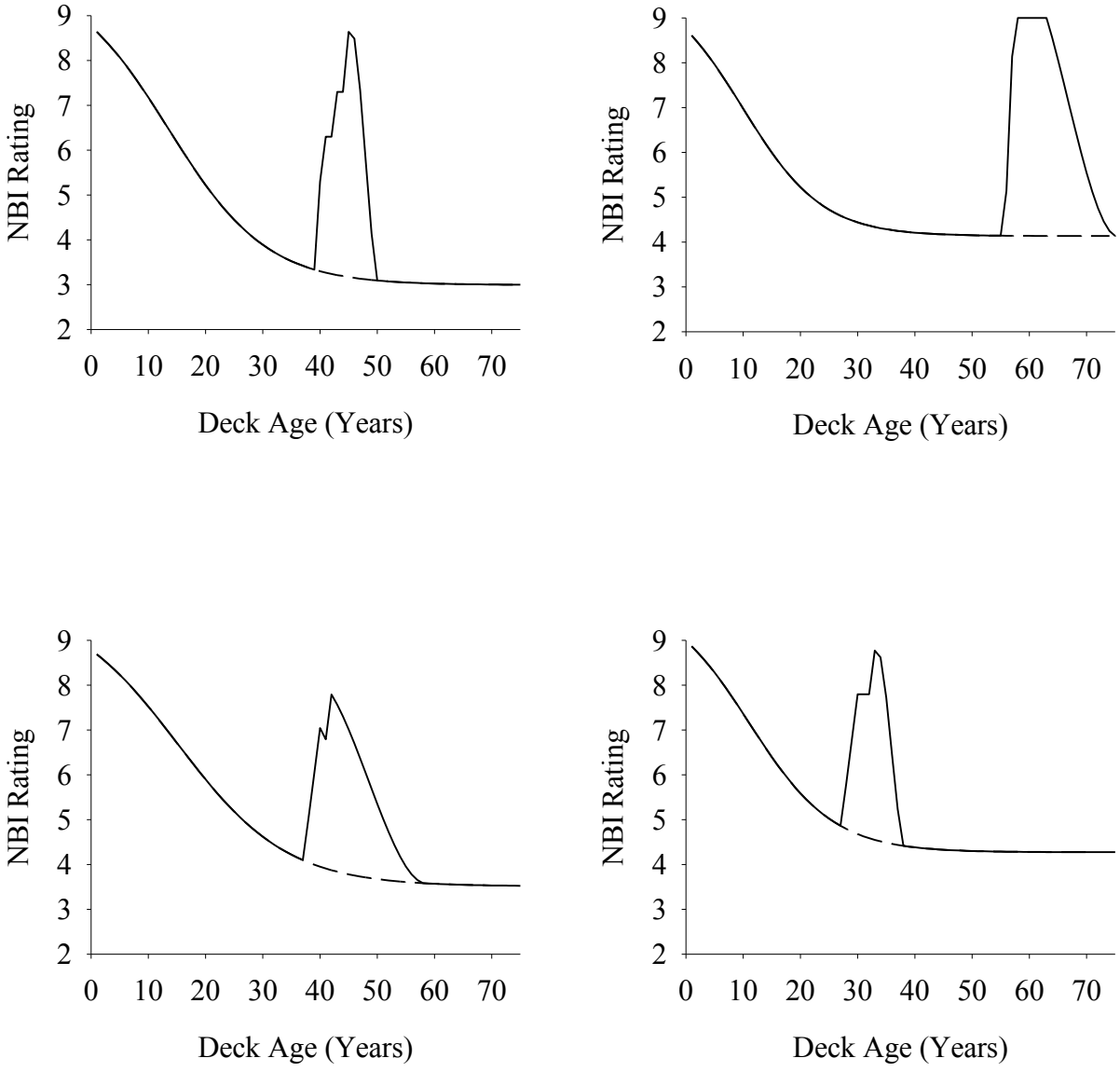


Figure B-1: BMS results; rating constraint only

Figure B-1 (cont’d)

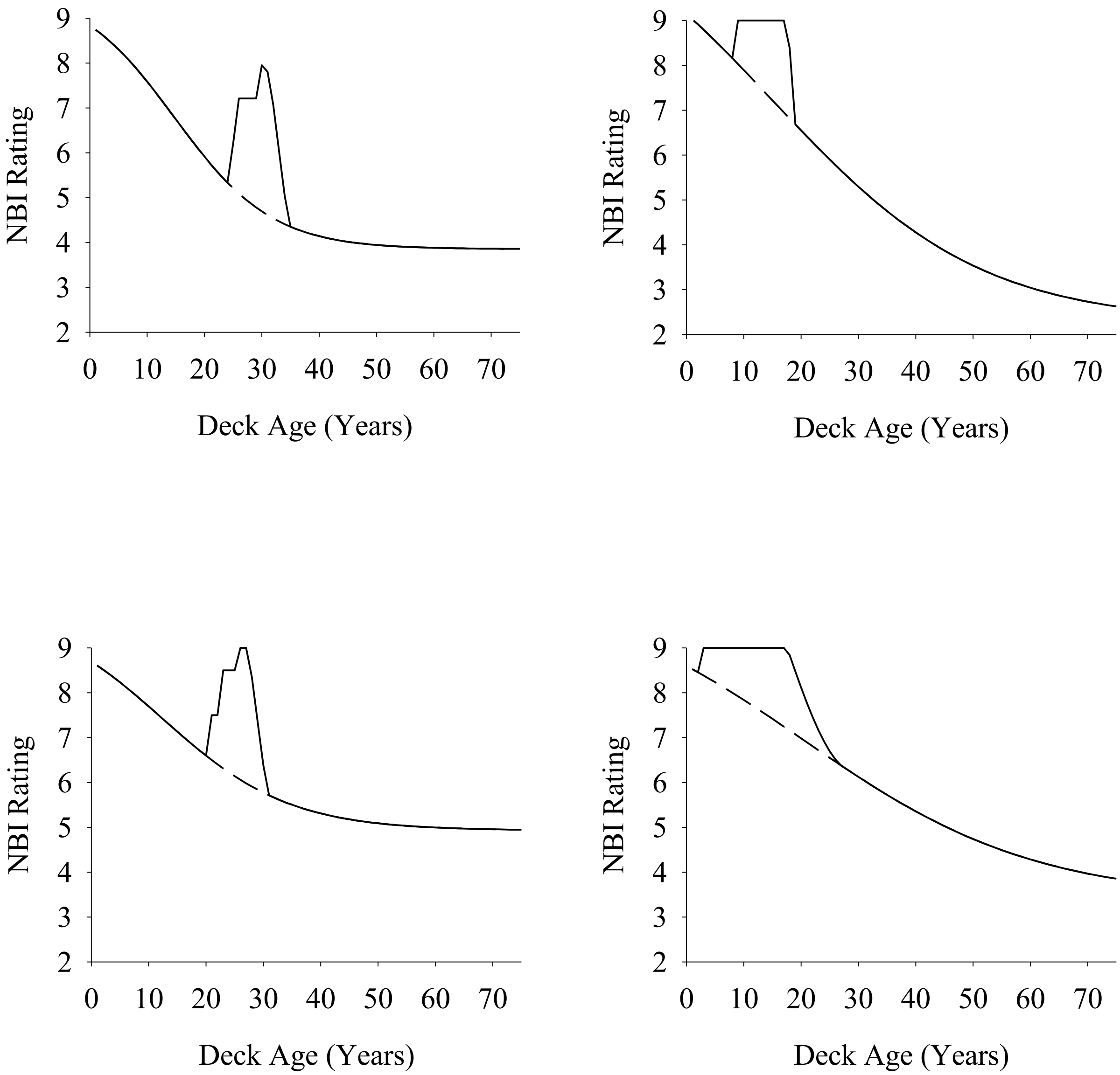
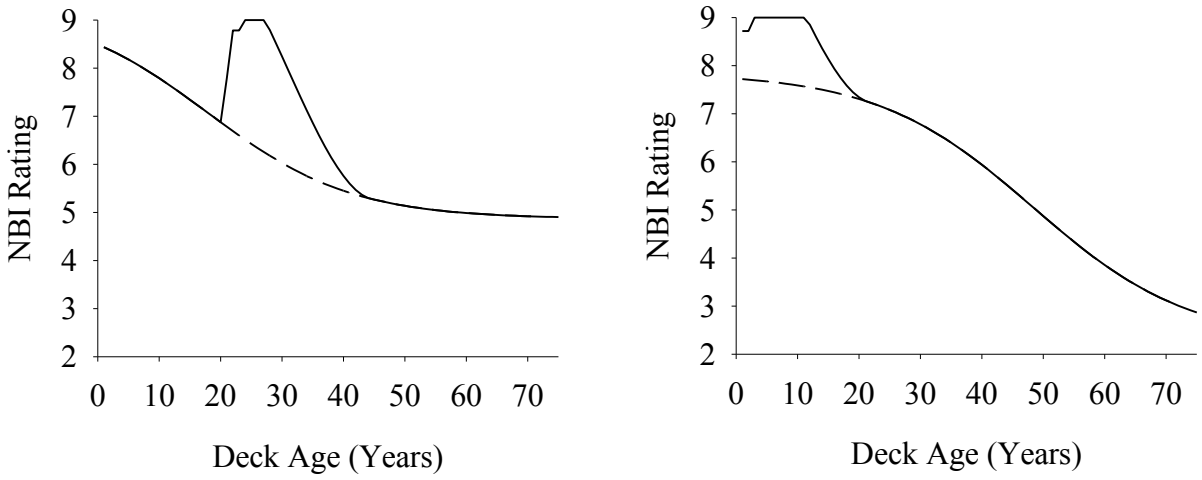


Figure B-1 (cont'd)



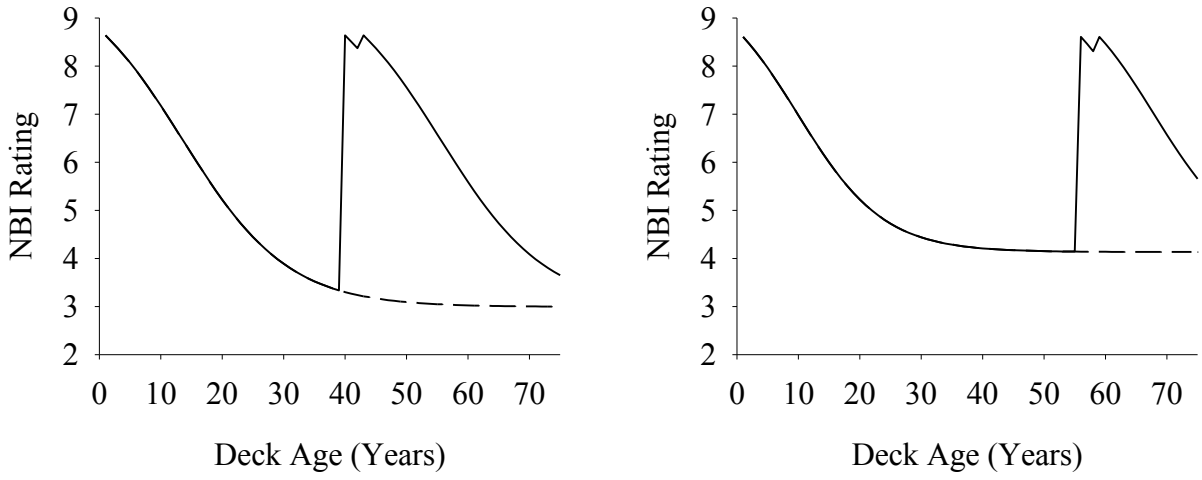


Figure B-2: BMS results; rating and repair limit constraints

Figure B-2 (cont'd)

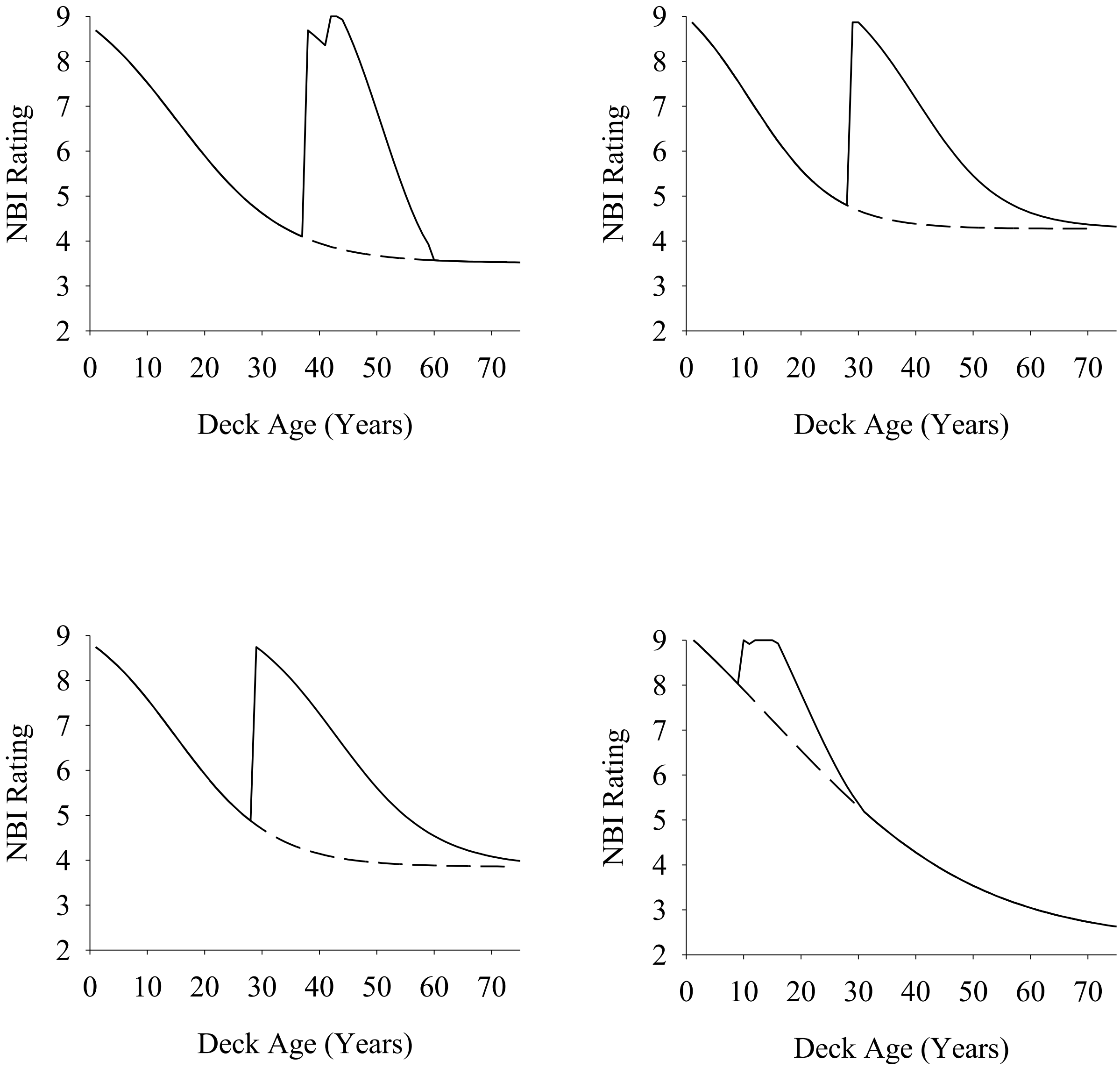
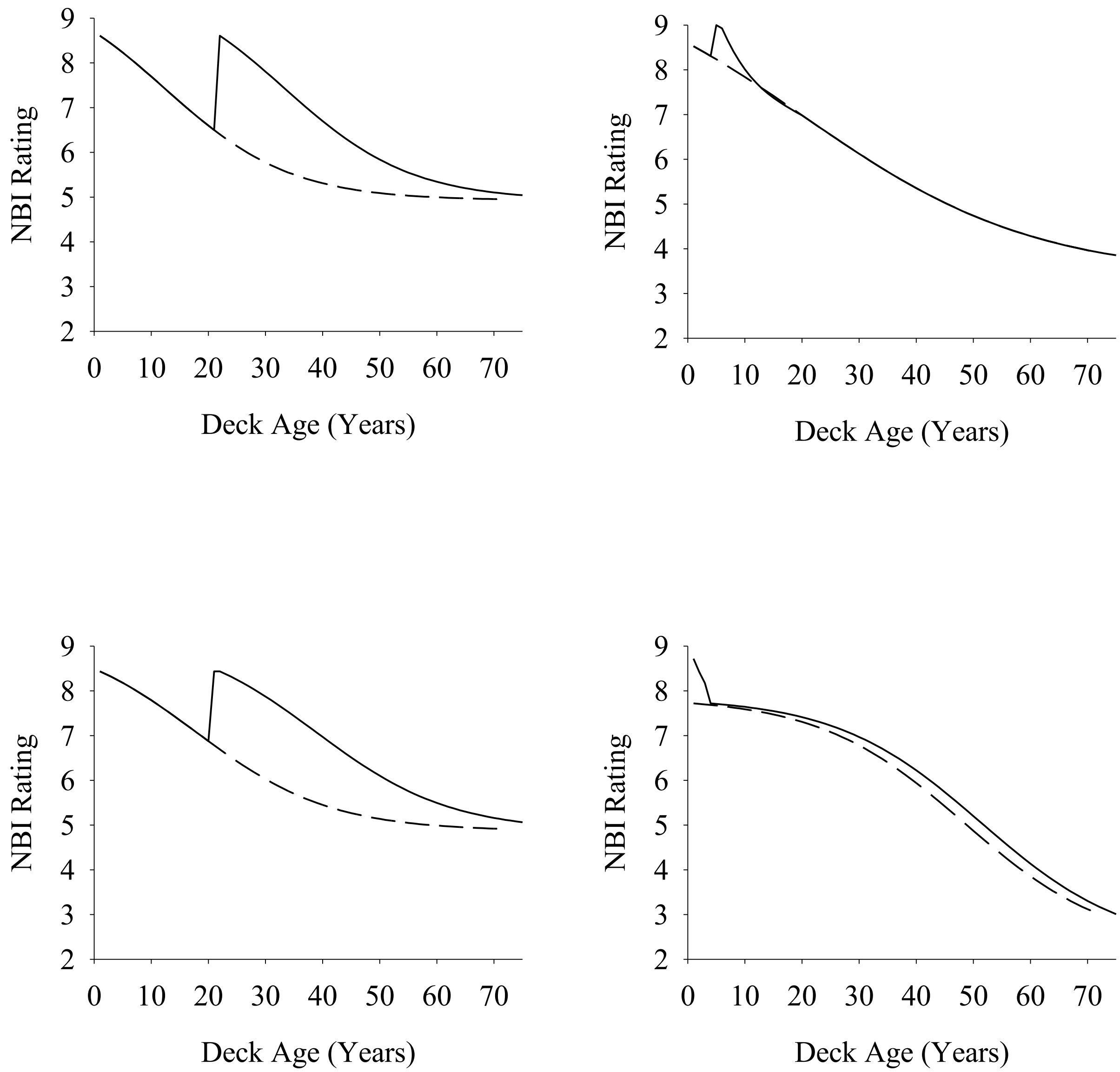


Figure B-2 (cont'd)



APPENDIX C: ENSEMBLE OF NEURAL NETWORKS (ENN 1.1) USER'S MANUAL

Rigoberto Burgueño, Emily K. Winn and Nan Hu
Michigan State University

C. 1 Introduction

ENN-Deck (V 1.1) is a program developed by Michigan State University for the Michigan Department of Transportation (MDOT) as part of a research project to model the degradation of highway bridge decks. Deck degradation is modeled through an ensemble of artificial neural networks (ENN). ENN-Deck is a stand-alone executable compiled from Matlab (Mathworks 2012) codes using a graphic user interface (GUI).

ENN-Deck contains a degradation model and a post-repair model. For the prediction of degradation, the program uses an ensemble of neural networks trained with historic information from the National Bridge Inventory (NBI) database. The program can make predictions for new bridges using design parameters or for existing bridges using the MDOT Bridge ID. Existing and new inspection data can be plotted on the degradation curve. In addition to the degradation models developed, an investigation of the improvements provided by reconstruction repairs used by the MDOT was conducted. The program also allows determining the improvement on deck condition ratings from repair methods and its subsequent post-repair degradation.

Overall, this program provides a new approach to model bridge deck conditions utilizing data from the NBI database of state DOTs. It can assist the MDOT on the prediction of the condition ratings and the decision making of the maintenance actions on highway bridge decks in Michigan. It is noted however, that the neural network predictions are based on learned performance from empirical inspection data. Thus, its accuracy is affected by the subjectivity of the visual inspection process. For that, while there are ways to infer variable significance in the neural network model, such capability was not included in this program.

C.2 Installation

C.2.1 Copy CD files to destination folder

Copy the files in the CD to the location where ENN is to be executed.

C.2.2 Set up MCRInstaller.

Install the Matlab Component Runtime (MCR) through the MCRInstaller. MCRInstaller is required before running all the Matlab stand-alone application, because it contains the toolboxes that the application needs. You need to have administrative rights in the computer/account in order to do this installation. Double click to open MCRInstaller and follow the InstallShield Wizard instructions, as shown through Figure C-1 to Figure C-5. After finishing the installation, the MCRInstaller file may be deleted.

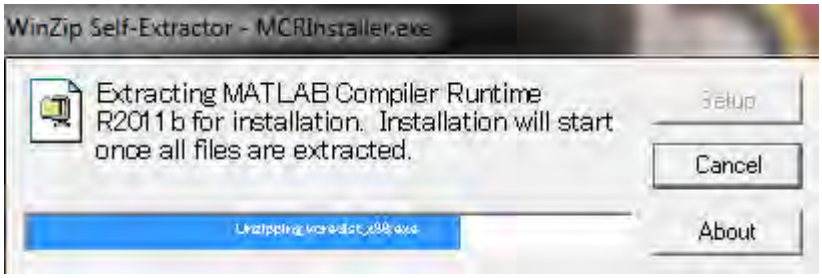


Figure C-1 Run the MCRInstaller.exe

The Matlab Complier Runtime Installer will start. Click 'Next'.

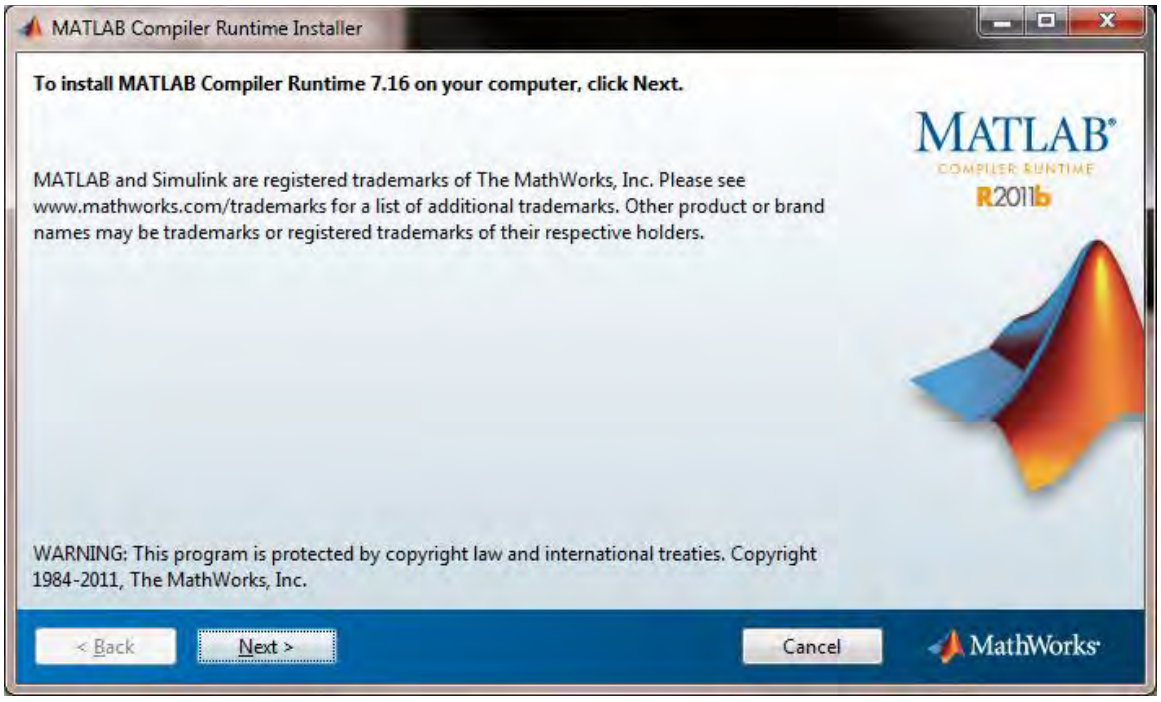


Figure C-2 Matlab Complier Runtime Installer

Click ‘Install’.

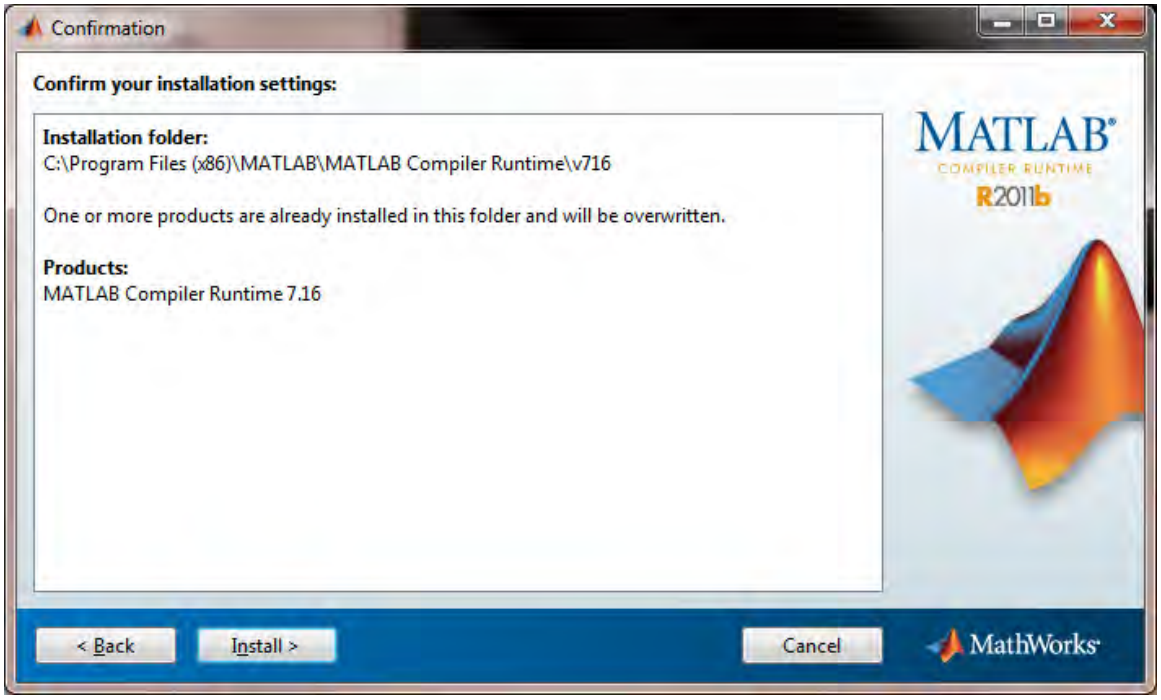


Figure C-3 Installation setting of Matlab Complier Runtime

The program should now be installing.

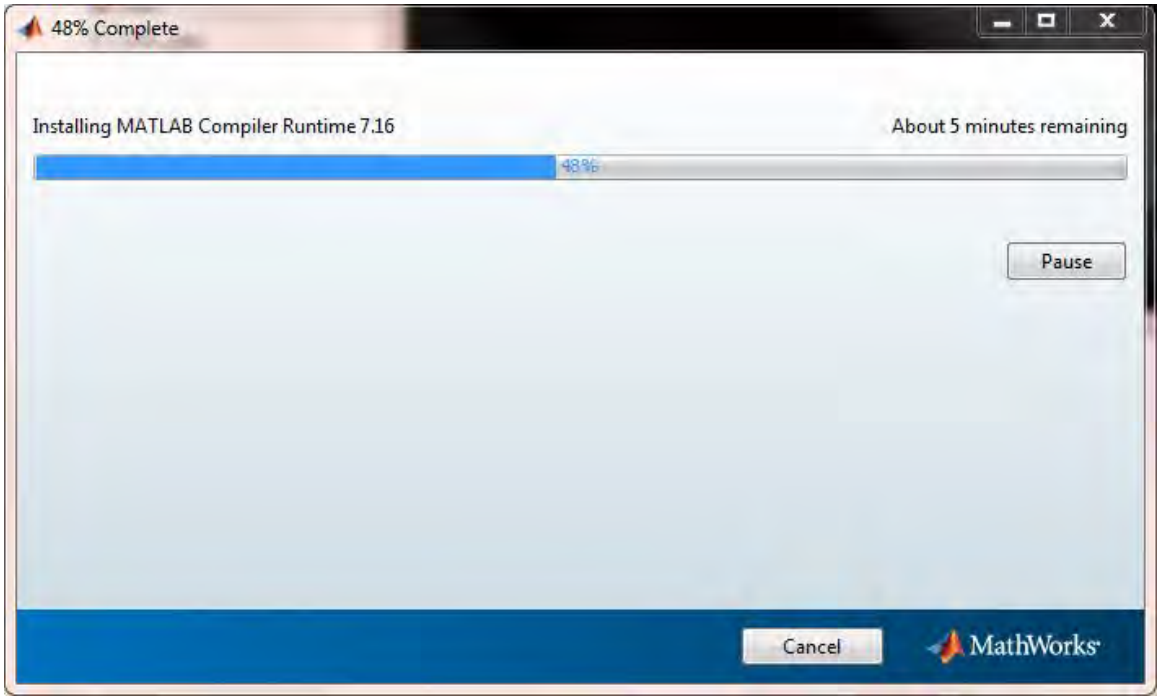


Figure C-4 Installation process of Matlab Complier Runtime

After installation, click ‘Finish’.

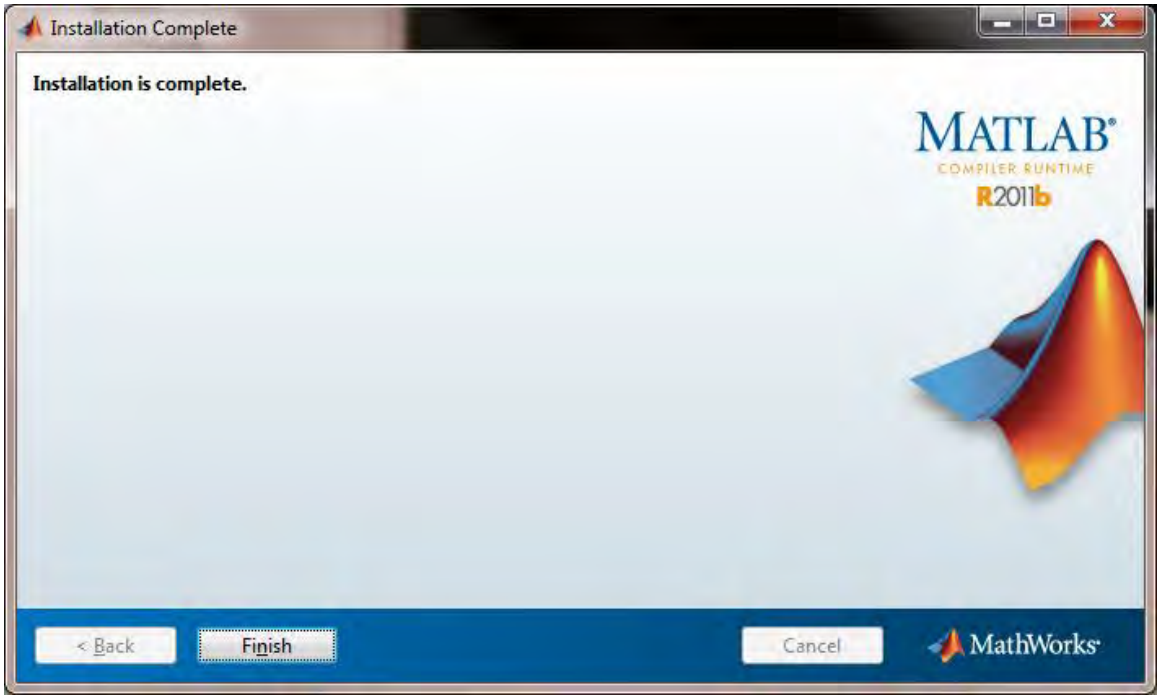


Figure C-5 Completion of installation process

C.2.3 Install ENN_pkg

Figure C-6 Extract of the ENN-Deck package

ENN-Deck 1.1 can be installed by double-clicking the ‘ENN_Deck_pkg’, as shown in Figure C-6. The program file and a readme file will be shown in the same folder.

C.2.4 Execute ENN-Deck 1.1

ENN-Deck 1.1 can be run by double-clicking the ‘ENN’ icon.

C.3 User Interface

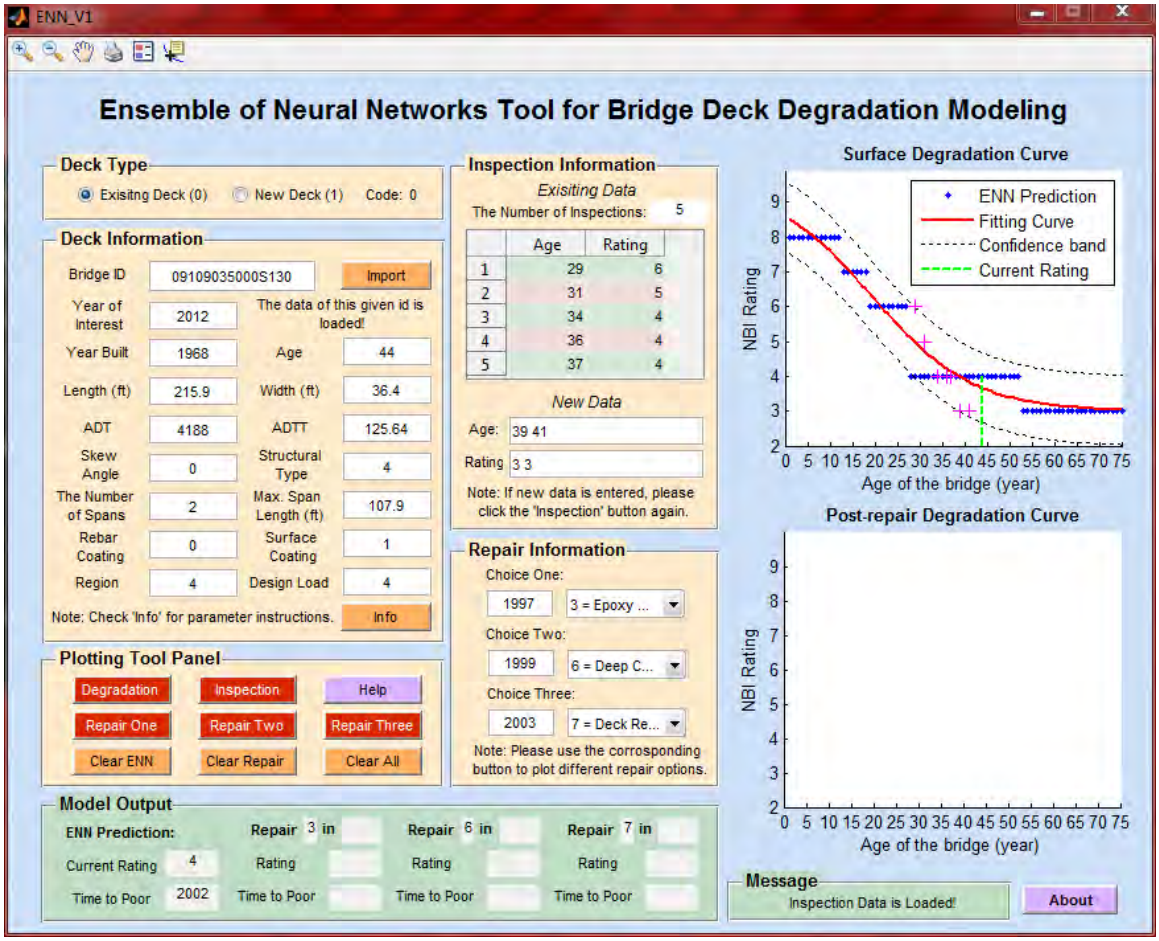


Figure C-7 Interface of ENN 1.1 (predict using bridge ID)

The user interface of ENN program is shown in Figure C-7. It contains three main zones: model input, plotting and model output. Detailed explanation of key parameters is given next.

C.3.1 Predictions Using MDOT Bridge ID (Existing Deck)

The first example is using an existing MDOT bridge to show the entire modeling process of a degradation curve by the ENN model and the determination of a post repair curve.

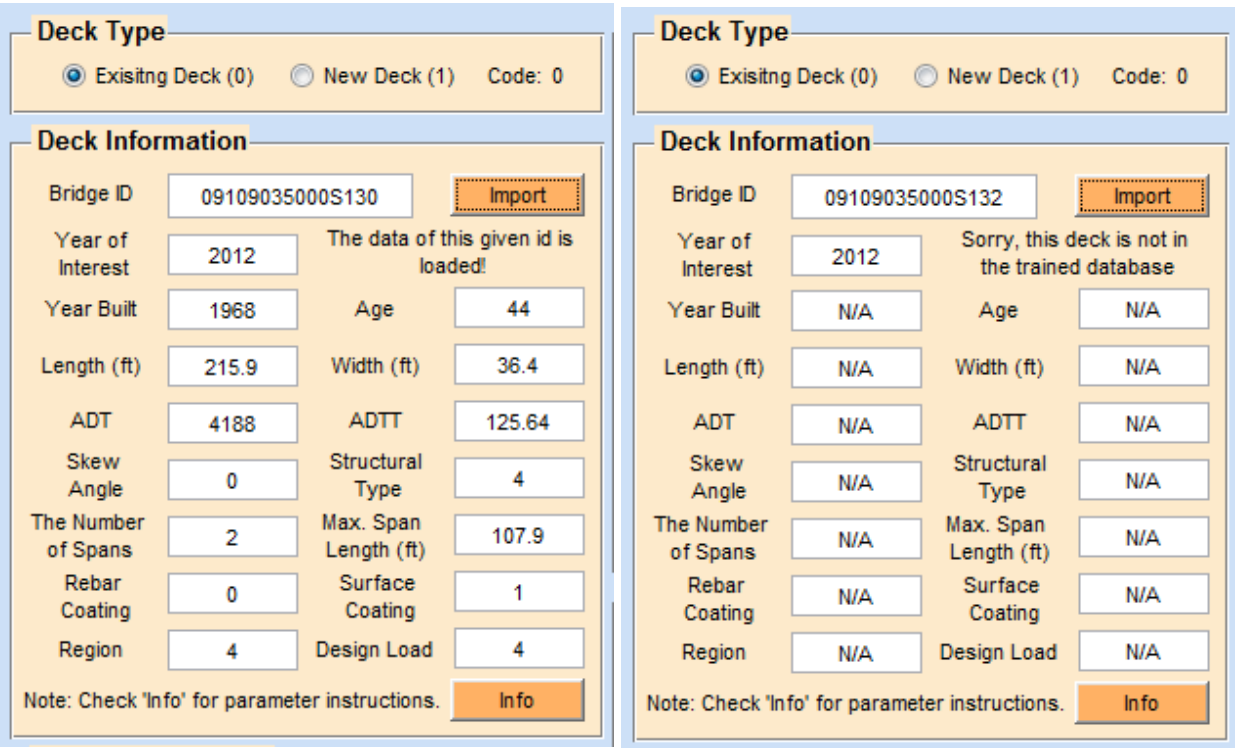
Step 1: Choose the Deck Type

The first step is choosing the deck type: select 'Existing Deck' radio button if you would like to predict the deterioration curve for an existing MDOT bridge, refer to **Figure C-**8. Once

the radio button is clicked, a ‘0’ in the code zone indicates what deck type was selected by the user.

Step 2 Input Bridge ID and Year of Interest, Import the Deck Information

The second step is to import the deck information with a valid bridge ID. The format needs to conform to the Michigan Structure Inventory and Appraisal Coding Guide and a capital "S" must be used. An example input for this step is: ‘09109035000S130’, as shown in **Figure C-8** (a). The deck information cannot be loaded, if the user makes an input error or the deck is not included in the training database (a total number of 1911), as shown in **Figure C-8** (b).



(a) **(b)**

Figure C-8 Deck information of a given existing deck

Below is guidance on some key parameters for defining the deck information:

- ADT: Average daily traffic (Min: 70, Max: 209200);

- ADTT: average daily truck traffic (Min: 0, Max: 15761);
- Skew Angle: the angle between the centerline of a pier and a line normal to the roadway centerline (Min: 0, Max: 72);
- Number of Spans: (Min: 1, Max: 15);
- Maximum Span Length: (Min: 21 ft, Max: 929 ft)
- Structural Type (3, 4 and 5): 3, for prestressed concrete bridge with I-girders; 4, for prestressed concrete bridge with adjacent box girders; and 5, for prestressed concrete bridge with spread box girder.
- Rebar Type(0 and 1) : 0 black steel reinforcement, 1 epoxy coated reinforcement;
- Surface Coating (1, 2, 3, 8, 9 and 0): 1 monolithic concrete (concurrently placed, with structural deck), 2 integral concrete (separate non-modified layer of concrete added to structural deck), 3 latex concrete or similar additive, 8 gravel, 9 other, 0 none (no additional concrete thickness or wearing surface is included in the bridge deck);
- Design Load (2, 3, 4, 5, 6, 9 and 0):: Use the code below to indicate the live load for which the structure was designed: 2 M 13.5 H 15, 3 MS 13.5 HS 15, 4 M 18 H 20, 5 MS 18 HS 20, 6 MS 18+Mod HS 20+Mod, 9 MS 22.5 HS 25, 0 Other or Unknown (describe on inspection reporting form).
- Region: 1-Superior, 2-North, 3-Grand, 4-Bay, 5-Southwest, 6-University, 7-Metro;

Step 3 Plotting Deterioration Curves

Click the 'Degradation' pushbutton in the 'Plotting Tool Panel'. A warning dialog will show up, saying the "Please wait for 10 seconds". After that, the ENN prediction result will show in the model output box. The current rating is calculated based on the year of interest, while time to poor is the year on the fitting curve when the rating reaches to a value of '4'.

Figure C-9 ENN prediction for bridge id '09109035000S130'

Results from the ENN model for a valid bridge ID include the ENN prediction, the fitting curve, the confidence bands and the current rating, as shown in Figure C-10. The user may open

a file (namely‘ENN.txt’) in the same folder of the application, which contains a 75×5 matrix, including the age of deck, the ENN prediction, the fitting curve, upper bound of confidence band and lower bound of confidence band, as shown in Figure C-11.

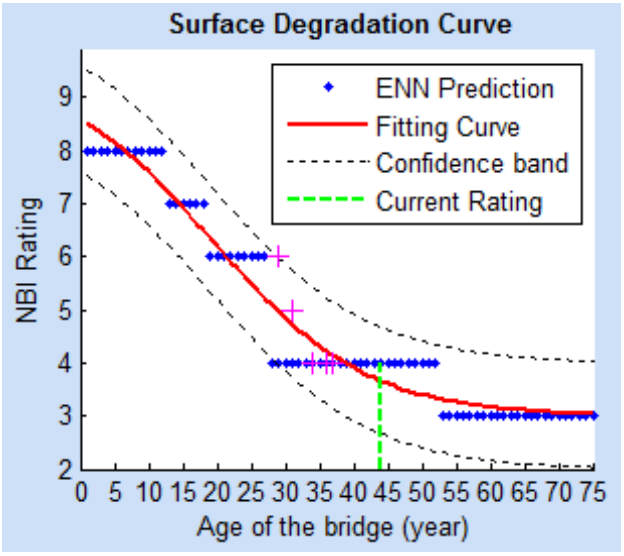


Figure C-10 surface deterioration curve of deck ‘09109035000S130’

1	8	8.51	9.51	7.51
2	8	8.43	9.43	7.43
3	8	8.34	9.34	7.34
4	8	8.25	9.25	7.25
5	8	8.15	9.15	7.15
6	8	8.05	9.05	7.05
7	8	7.94	8.94	6.94
8	8	7.83	8.83	6.83
9	8	7.71	8.71	6.71
10	8	7.59	8.59	6.59
11	8	7.46	8.46	6.46
12	8	7.33	8.33	6.33
13	7	7.20	8.20	6.20
14	7	7.06	8.06	6.06

Figure C-11 the output file of prediction of deck ‘09109035000S130’

If the bridge ID is not included in the ENN trained database, an error dialog box will appear as shown in Figure C-12.

Figure C-12 Error dialog box for a bridge id not in the trained database

Step 4 Loading the Inspection Data

Click the 'Inspection' pushbutton. Existing data will be loaded and the data will be shown on the top figure, see the Figure C-13. Bridge '09109035000S130' has five inspection records in the database. Note that current trained database only includes the inspection data up to March 2010. Therefore, the user can enter new inspection data (age and rating). The numbers for the age and rating should be entered separated by a space and the vectors should be of the same length. Otherwise, there will be an error as shown in Figure C-14. In this example, three new inspection data are entered in the 'New Data' field. Click the 'Inspection' button again or enter the new data at the beginning. It can be seen that the inspection data are shown in the figure. The user can insert or close the legend with the button of the toolbox at the top of the interface.

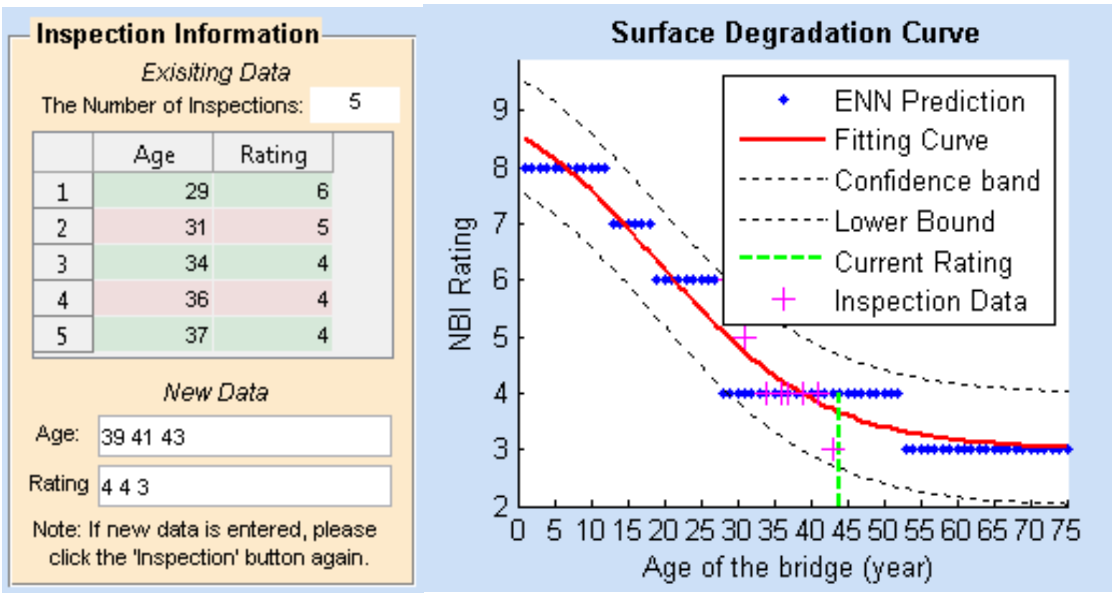


Figure C-13 Inspection data and degradation curve for deck '09109035000S130'

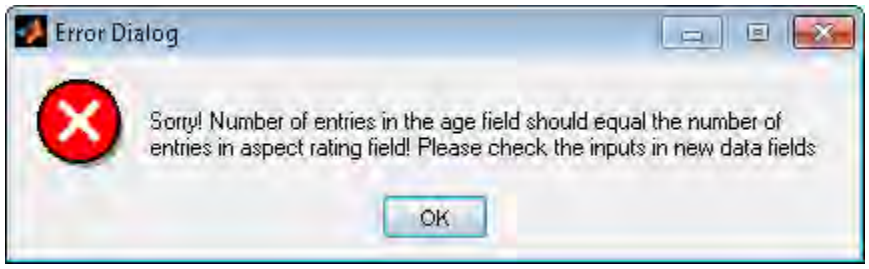


Figure C-14 Error dialog box for entering new inspection data

Step 5 Post-repair Curves

To make decisions regarding the selection of repairs, the MDOT currently uses the bridge deck preservation matrix (BDPM). Based on a statistical analysis of the NBI database, the expected rating improvement for the different repair options in the BDPM was modified. Degradation was modeled using a cubic regression based on the immediate post-repair degradation rate of a repair effort and the estimated time until the next maintenance action as per the original MDOT BDPM. A detail explanation of the approach can be found in [1]. In this part, the user will be asked which repair option is to be implemented and when the repair is to be done as shown in Figure C-15. There are six repair options in the pop-up menu.

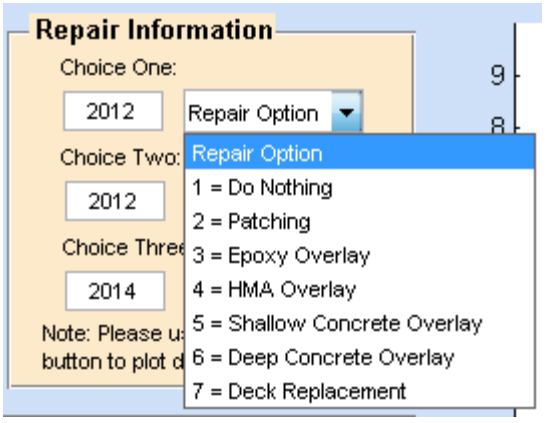


Figure C-15 Repair options in pop-up menu

Enter the repair year and select one out of six repair options. After clicking the 'Repair One' pushbutton, the post repair curve one will be given, as shown in Figure C-16.

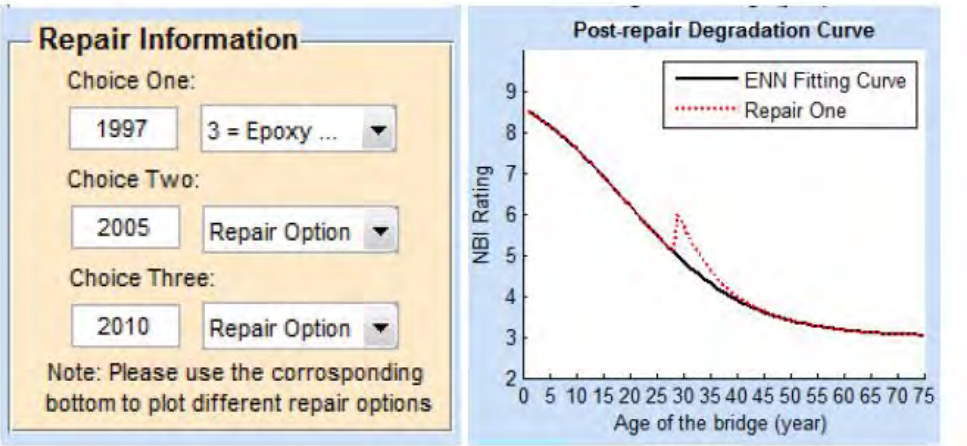


Figure C-16 Repair option one and corresponding post repair curve for deck '09109035000S130'

The user can try three different repair scenarios and plot the post-repair curves by clicking different buttons in the plotting tool panel. As shown in Figure C-17, three post-repair curves will be given in the figure. The user can also close the legend or zoom in the figure with the available tools. In the model output zone the rating immediately after repair and time to poor for three repair options will be listed so that the user can compare the repair effect. Again, the user may open a file (namely'Repair.txt') in the same folder, which contains the age of deck, the fitting curve, post-repair curve one, post-repair curve two and post-repair curve three.

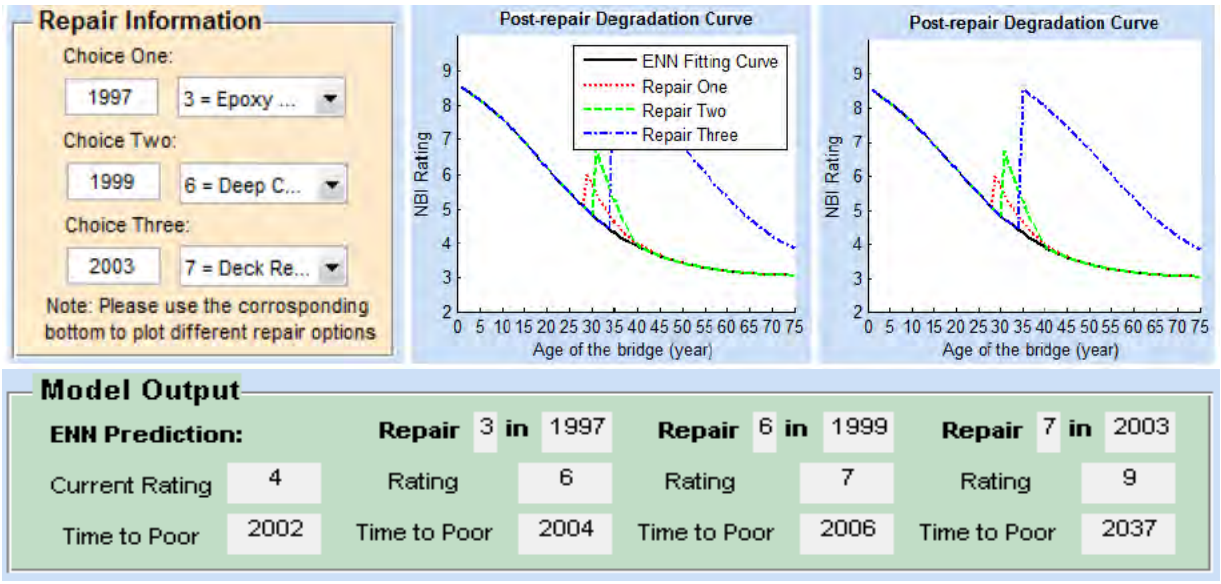


Figure C-17 Three repair scenarios for deck '09109035000S130'

Step 6 Reset and Try a New Bridge ID

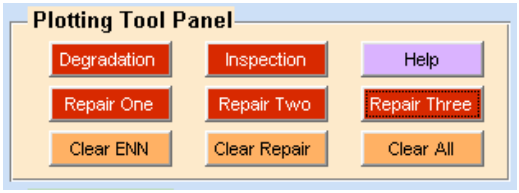


Figure C-17 Plotting Tool Panel

In the bottom row of the plotting tool panel the user can click the 'Clear ENN' button to reset the surface degradation curve; click the 'Clear Repair' button to reset the post-repair degradation curve; and/or click the 'Reset All' button to reset all the figures and data, as shown in Figure C-17. There is a toolbar at the top of the interface. As shown in Figure C-18, the first three buttons (zoom in, zoom out and pan) can help the user to inspect the figures in detail. The next one allows printing of the figures. The one following the print button is to remove or insert the legend, because the legend in some cases will overlap with the degradation or repair curves. The final button on the toolbar is the data cursor, which allows checking all the data on the figures.

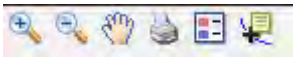


Figure C-18 Toolbar

The user may click the 'Help' button to open a quick direction text. For details, the user may click the 'User Manual' button at the bottom as shown in Figure C-19.

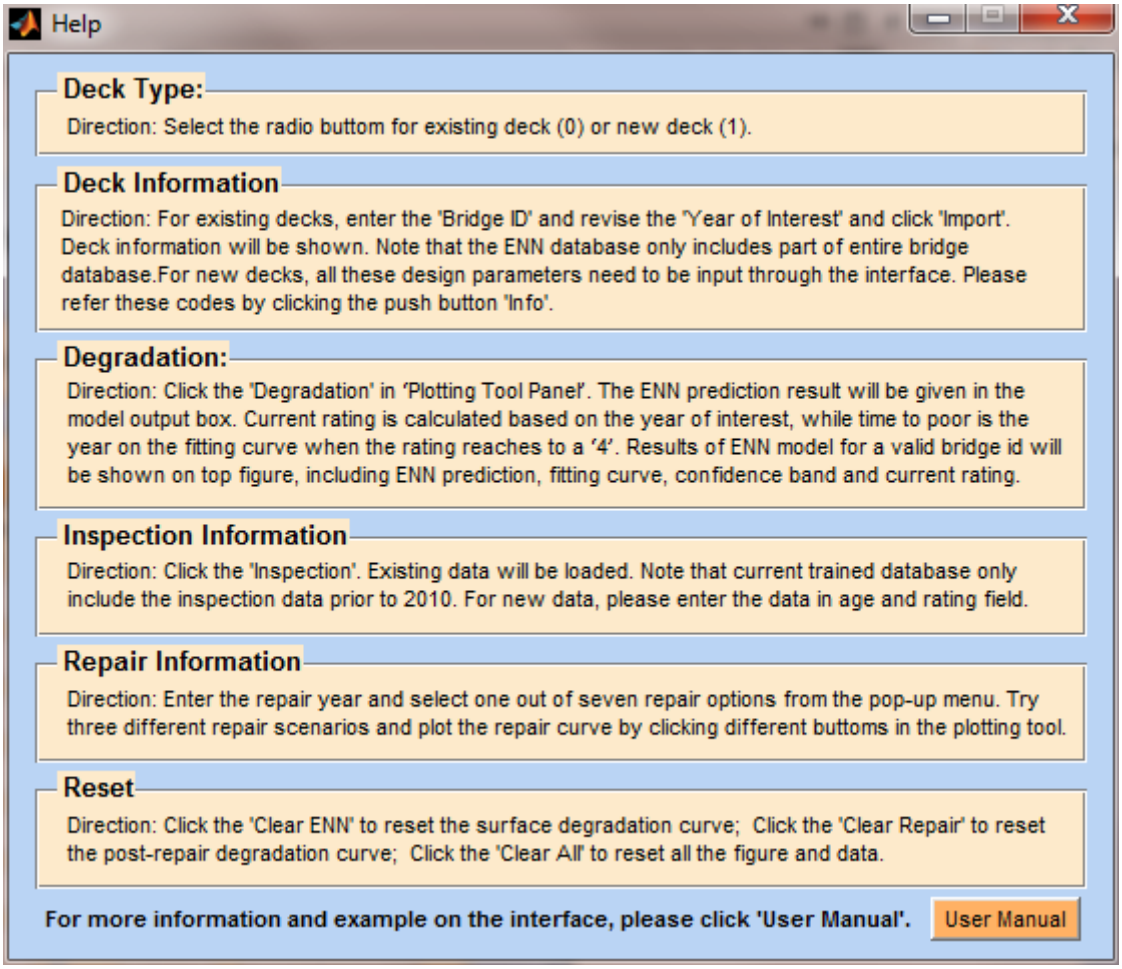


Figure C-19 Help window for a quick instruction

Every time the user clicks the pushbuttons in the plotting tool panel and calculations are done, there should be a message in the box that indicates the action has been completed, as shown in Figure C-20. For both the degradation curve and post-repair curve, it only takes about 10 seconds to obtain the figure.

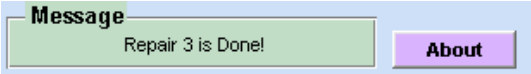


Figure C-20 Message text box

Finally, the 'About' button may be clicked to find the information about the program, as shown in Figure C-21.

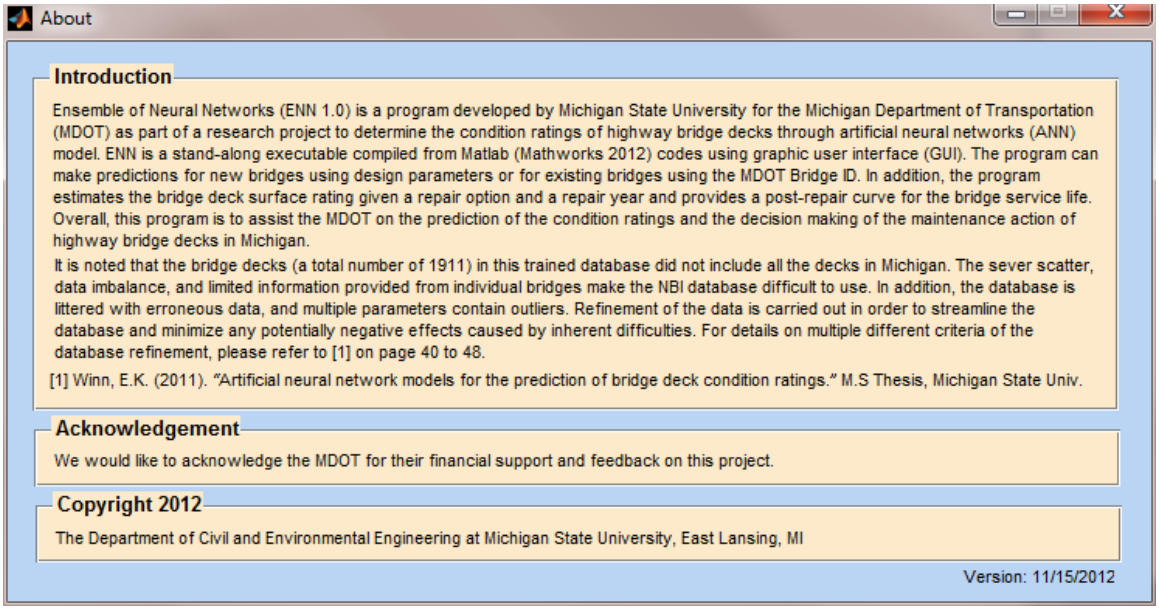


Figure C-21 'About' pushbutton for interface information

C.3.2 Predict Using MDOT Bridge Design Parameters (New Deck)

The second example is using MDOT bridge design parameters as input. The only difference between evaluating an existing deck or a new deck is that the user must enter the deck information according to the MDOT design parameters.

Step 1: Choose the Deck Type

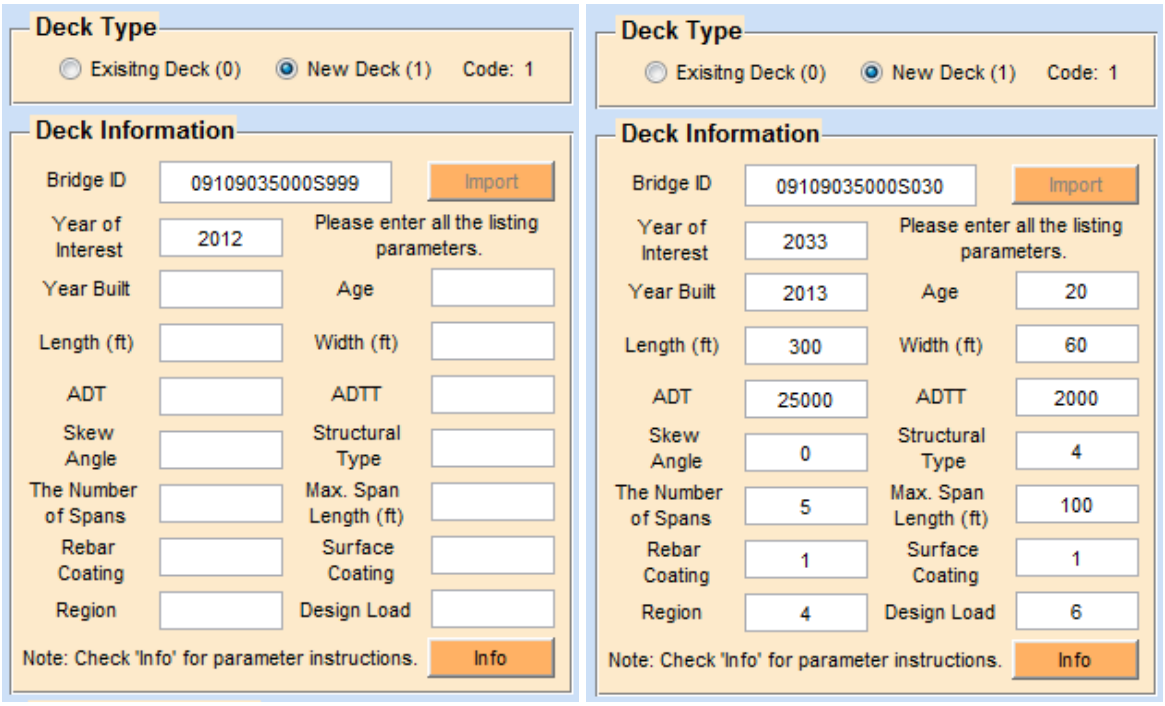
The first step is still choosing the deck type: in this case select the 'New Deck' radio button, see to Figure C-22 (a). Once the radio button is clicked, a '1' in the code zone indicates what deck type was selected.

Step 2 Input All the Listed Deck Information

The second step is to enter the deck information. Here is an example for a new deck: '09109035000S999', as shown in Figure C-22 (b). Assuming that this deck will be built in 2013 and the year of interest is 2033. All these design parameters need to be input through the interface. Please refer to the code information provided in previous section for an existing deck. It should be recognized that several parameters from the bridge appraisal inventory are not included. For example, in the full database, there are 10 different structural types. However, in the trained database, only types 3, 4 and 5 are available. It is noted that the ENN model is not perfect. Since the accuracy of the ENN model is dependent on the trained database, the prediction results could be wrong if the parameters used deviate considerably from the training database. Therefore, the program only provides a number of options for the structural types, rebar type, surface coating and design load. For other parameters, the program allows a range based on the existing data in order to improve the accuracy of the ENN model.

Step 3 Plotting Deterioration Curves

After entering the deck information, there is no difference between an existing deck and new deck for the following four steps. Here is an example for a new deck as shown in Figure C-23. Click the 'Degradation' pushbutton in the 'Plotting Tool Panel'. The ENN prediction results will be given in the model output box. The current rating is '7' at the year of interest (2033), while the time to poor is '2058' on the fitted curve when the rating reaches to a rating of '4'.



(a) **(b)**

Figure C-22 Deck information of a new deck '09109035000S999'

Step 4 Entering the Inspection Data

Since a new deck does not have existing data, new inspection data will be entered in this step.

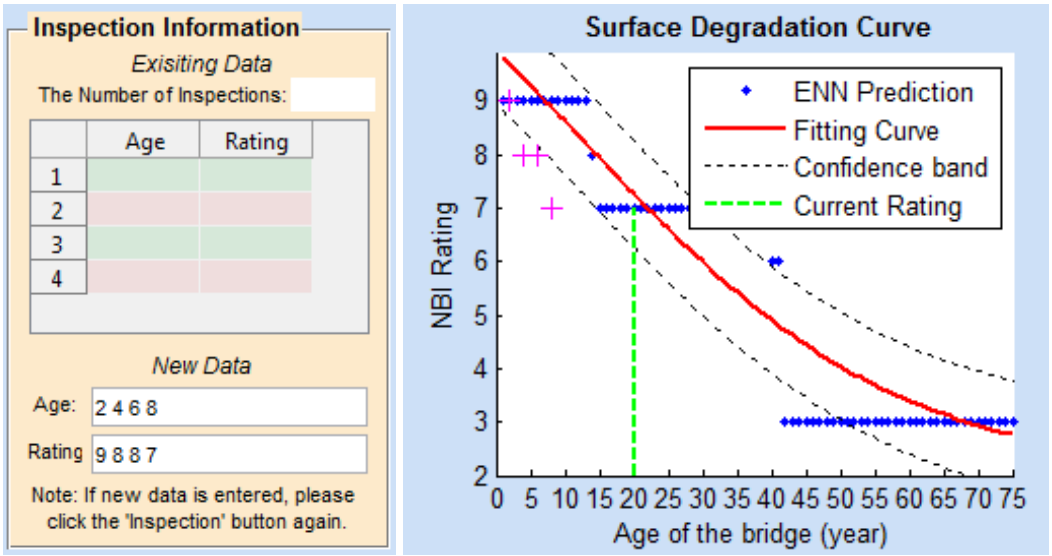


Figure C-23 inspection data and degradation curve for new deck '09109035000S999'

Step 5 Post-repair Curves

Results for the example are shown in Figure C-24. Three post-repair curves are given in the figure. Note that repair 3, a deck replacement in 2055, the time to poor is not available in the given range of 0 to 75 years.

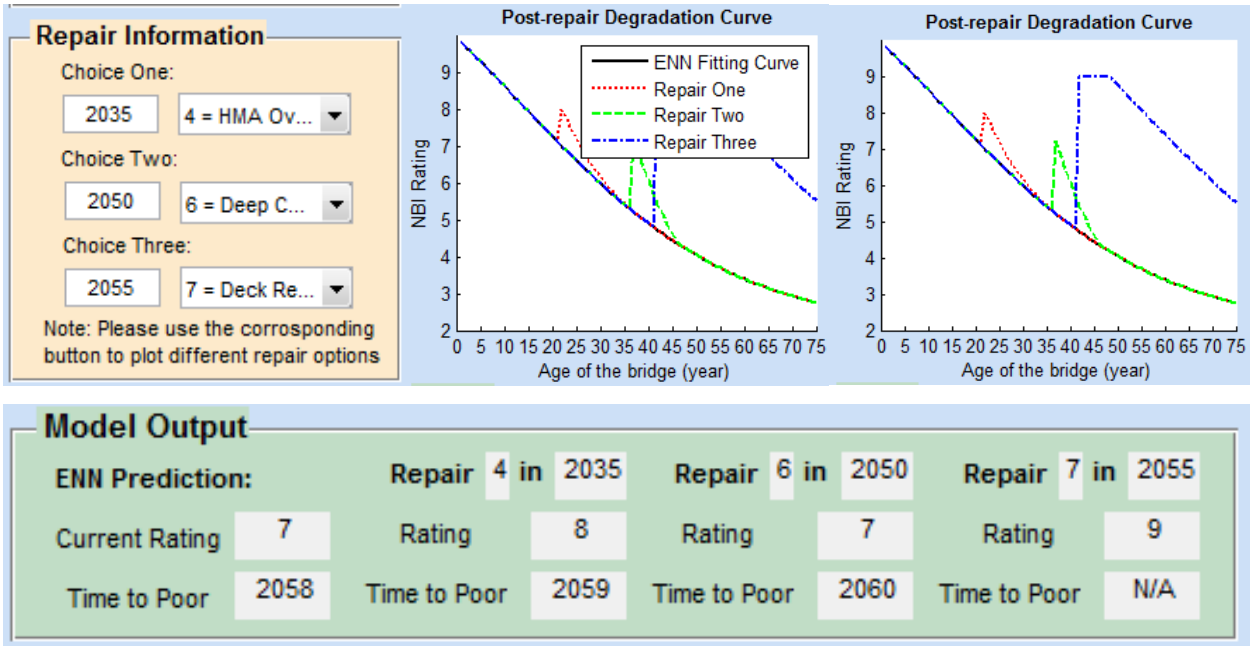


Figure C-24 Three repair scenarios for deck '09109035000S130'

Step 6 Reset Data

Reset all the data and try different design parameter.

REFERENCES

REFERENCES

AASHTO (2009). Pontis Bridge Management. American Association of State Highway and Transportation Officials, Washington, D.C

Agrawal, A.K., Kawaguchi, A, and Chen, Z.. (2009). "Bridge element deterioration rates." Final Research Report, Project No. C-01-51. New York State Department of Transportation.

Akgül, F. and Frangopol, D. (2004). "Computational platform for predicting lifetime system reliability profiles for different structure types in a network." *Journal of Computing in Civil Engineering.* 18(2), 92-104.

Al-Barqawi, H., and Zayed, T. (2006). "Condition rating model for underground infrastructure sustainable water mains." *Journal of Performance of Constructed Facilities,* 20(2), 129-135.

Ali, K. and Pazzani, M. (1996). "Error reduction through learning multiple description." *Machine Learning*. 24(3), 173 – 202.

ASCE. (2009). "Report card for America's infrastructure."<http://www.infrastructurereportcard.org/> (September 27, 2010).

Ayyub, B. McCuen, R. *Probability, Statistics, and Reliability for Engineers and Scientists.* 2nd ed. Boca Raton, FL: Chapman & Hall: 2003.

Babaei, K., Purvis, R., Clear, K., and Weyers, R. (1996). "Methodology for concrete removal, protection, and rehabilitation." *Report Prepared for the U.S. Department of Transportation, Federal Highway Administration,* Wilbur Smith Associates, Falls Church, VA.

Bauer, E. and Kohavi, R. (1999). "An empirical comparison of voting classification algorithms: bagging, boosting, and variants." *Machine Learning*. 35, 1-38.

Benjamin, J. and Cornell, C. (1970). "Common probabilistic models." *Probability, Statistics, and Decision for Civil Engineers.* McGraw-Hill, New York, NY.

Bertolini, L., Elsener, B., Pedeferri, P., and Polder, R. (2004). *Corrosion of Steel in Concrete*. Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim.

Bhargava, K., Ghosh, A., Mori, Y., and Ramanujam, S. (2006). "Analytical Model for Time to Cover Cracking in RC Structures Due to Rebar Corrosion." *Nuclear Engineering and Design.* 236, 1123-139.

Boatman, B. (2010) "Epoxy coated rebar bridge decks: expected service life." MDOT Bridge Operations Unit Report

Bose, N.K., Liang, P. (1996). "Neural network fundamentals with graphs, algorithms, and applications." McGraw-Hill.

Breiman, L. (1996). "Bagging Predictors." *Machine Learning*. 24(2), 123-140.

Cattan, J., and Mohammadi, J. (1997). "Analysis of bridge condition rating data using neural networks." *Microcomputers in Civil Engineering,* 12(6), 419-429.

Dackerman, U., Li, J., and Samali, B. "Damage identification in timber bridges utilizing the damage index method and neural network ensembles." *Australian Journal of Structural Engineering*. 9(3), 181-194.

Dantzig, G.B. (2010). "The nature of mathematical programming." *Mathematical Programming Glossary.* INFORMS Computing Society.

Derucher, K., Korfiatis, G., and Ezeldin, S. (1994). *Materials for Civil and Highway Engineers.* 3rd edition. Prentice-Hall., Englewood Cliffs, N.J.

Dietterich, T. (2000). "Ensemble methods in machine learning." Multiple Classifier Systems. Lecture Notes in Computer Science. Springer Berlin / Heidelberg. Volume 1857. Pgs 1-15.

Domingos, P. (1996). "Using partitioning to speed up specific-to-general rule induction." *Proceedings of the AAAI-96 Workshop on Integrating Multiple Learned Models.* 29-34.

Eiben, A., and Smith, J. (2003). "Genetic algorithms." *Introduction to Evolutionary Computing.* Springer, Verlag Berlin Heidelberg, Germany.

Elbehairy, H., Hegazy, T., and Soudki, K. (2009). "Integrated multiple-element bridge management system." *Journal of Bridge Engineering.* 14(3), 179-187.

Enright, M., and Frangopol, D. (1998). "Probabilistic Analysis of Resistance Degradation of Reinforced Concrete Bridge Beams under Corrosion." *Engineering Structures* 20(11), 960-71.

Fanous, F., Wu, H., and Pape, J. (2000). "Impact of deck cracking on durability." Center for Transportation Research and Education Project 97-5. Iowa DOT Report TR-405

Frangopol, D., Kong, J., and Gharaibeh, E. (2001). "Reliability-based life-cycle management of highway bridges." *Journal of Computing in Civil Engineering.* 15(1), 27-34.

Furuta, H., Kameda, T., Fukuda, Y., and Frangopol, D. (2004). "Life-cycle cost analysis for infrastructure systems: Life cycle cost vs. safety level vs. service life." Keynote paper in *Life-cycle performance of deteriorating structures: Assessment, design and management*, D. Frangopol, E. Brühwiler, M. Faber, and B. Adey, eds., ASCE, Reston, VA., 19-25.

Gevrey, M., Dimopoulos, I., and Lek, S. (2003). "Review and comparison of methods to study the contribution of variables in artificial neural network models." *Ecological Modelling*. 160(2003). 249-264.

Haider, S. and Dwaikat, M. (2010). "Estimating optimum timings for preventive maintenance treatments to mitigate pavement roughness." *Journal of Transportation Research Board*.

Hair, J., Anderson, R., Tatham, R., and Black, W. (1998). "Multiple discriminant analysis and logistic regression." *Multivariate Data Analysis.* 5th edition. Upper Saddle River, New Jersey.

Hansen, L., and Salamon, P. (1990). "Neural network ensembles." IEEE Transactions on Pattern Analysis and Machine Intelligence. 12(10), 993-1001.

Haykin, S. (1999). "Introduction." *Neural networks: A comprehensive foundation*, 2nd Ed., Prentice-Hall, Upper Saddle river, N.J, 842.

Hegazy, T., Elbeltagi, E., and Elbehairy, H. (2004). "Bridge deck management system with integrated life cycle cost optimization." *Transportation Research Record: Journal of Transportation Research Board*, No. 1866, TRB, National Research Council, Washington D.C. 44-50.

Hu, X. (2001). "Using rough sets theory and database operations to construct a good ensemble of classifiers for data mining applications." Proceedings of the 2001 IEEE International Conference on Data Mining. 233-240.

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

Huang, Y.H., Adams, T., and Pincheira, J. (2004). "Analysis of life-cycle maintenance strategies for concrete bridge decks." *Journal of Bridge Engineering.* 9(3), 250-258.

Isgor, O., and Razaqpur, A. (2006). "Modelling Steel Corrosion in Concrete Structures." *Materials and Structures* 39, 291-302.

Jacobs, T. (1992). "Optimal long-term scheduling of bridge deck replacement and rehabilitation." *Journal of Transportation Engineering.* 118(2), 312-322.

Jiang, Y. (1990). "The development of performance prediction and optimization models for bridge management systems." Ph.D. thesis, Purdue University, West Lafayette, IN.

Kelley, R. (2010). Personal Communication.

Kohavi, R., and Provost, F. (1998). "Glossary of terms." *Machine Learning,* 30(2-3), 271-274.

Lee, C.K, and Kim, S.K. (2007). "GA-based algorithm for selecting optimal repair and rehabilitation methods for reinforced concrete (RC) bridge decks." *Automation in Construction*. 16(2007). 153-164.

Lee, J., Sanmugarasa, K., Blumenstein, M., and Loo, Y. C. (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.

Lee, S.K. and Krauss, P. (2003). "Service life extension of northern bridge decks containing epoxy-coated reinforcing bars." Concrete Reinforcing Steel Institute: Schaumburg, IL.

Li, Z. (2008). "Soft computing." *Soft computing for damage prediction and cause identification in civil infrastructure systems*. Ph.D. Dissertation, Dept. of Civil and Environmental Engineering, Michigan State Univ., East Lansing, Mich, 225.

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

Liu, G.P. (2001). "Neural networks." Nonlinear Identification and Control: A neural network approach. Springer, London. 210.

Liu, M. and Frangopol, D. (2005a). "Multiobjective maintenance planning optimization for deteriorating bridges considering condition, safety, and life-cycle cost." *Journal of Structural Engineering.* 131(5), 833-842.

Lui, M. and Frangopol, D. (2005b). "Balancing connectivity of deteriorating bridge networks and long-term maintenance cost through optimization." *Journal of Bridge Engineering.* 10(4), 468-481.

Liu, M. and Frangopol, D. (2006). "Optimizing bridge network maintenance management under uncertainty with conflicting criteria: Life-cycle maintenance, failure, and user costs." *Journal of Structural Engineering.* 132(11), 1835-1845.

Liu, Y. (1996). "Modeling the time-to-corrosion cracking of the cover concrete in chloride contaminated reinforcement concrete structures." Ph.D. Dissertation, Virginia Polytechnic Institute, Blacksburg, VA, 171.

Lui, C., Hammad, A., and Itoh, Y. (1997). "Maintenance strategy optimization of bridge decks using genetic algorithm." *Journal of Transportation Engineering.* 123(2). 91-100.

Madanat, S., Karlaftis, M., and McCarthy, P. (1997). "Probabilistic infrastructure deterioration models with panel data." *Journal of Infrastructure Systems.* 3(1), 4-9.

Mao, J. (1998). "A case study on bagging, boosting, and basic ensembles of neural networks for OCR." *Neural Networks Proceedings, 1998. IEEE World Congress on Computational Intelligence.* Vol. 3, 1828-1833.

MathWorks Inc., The (2010). "Neural network toolbox." MATLAB 2010a. Natick, Mass.

MDOT, (2009). "Michigan structure inventory and appraisal coding guide." Michigan Department of Transportation, Lansing, MI.

Mehta, P. K., and Monteiro, P. (2006). "Durability." *Concrete: Structure, Properties, and Materials*. Second ed. Englewood Cliffs, N.J.: Prentice Hall.

Mishalani, R., and Madanat, S. (2002). "Computation of infrastructure transition probabilities using stochastic duration models." *Journal of Infrastructure Systems.* 8(4), 139-148.

Miyamoto, A., Kawamura, K., and Nakamura, H. (2000). "Bridge management system and maintenance optimization for existing bridges." *Computer-Aided Civil and Infrastructure Engineering.* 15(2000). 45-55.

Morcous, G. (2002). "Comparing the use of artificial neural networks and case-based reasoning in modeling bridge deterioration." *Proc., 30th Annual Conference of Canadian Society for Civil Engineering,* Curran Associates, Inc, 2471-2479.

Moselhi, O. and Shehab-Eldeen, T. (2000). "Classification of defects in sewer pipes using neural networks." *Journal of Infrastructure Systems,* 6(3), 97-104.

National Engineering Technology (NET) Corporation, (1994). "Bridgit Bridge Management System: Technical Manual—Version 1.00." NCHRP Project 12-28(2)A, National Cooperative Highway Research Program, Transportation Research Board, Washington, DC.

Olden, J. and Jackson, D. (2002). "Illuminating the "black box": a randomization approach for understanding variable contributions in artificial neural networks." *Ecological Modeling*. 154(2002). 135-150.

Olden, J., Joy, M., and Death, R. (2004). "An accurate comparison of the methods for quantifying variable importance in artificial neural networks using simulated data."

Pandey, P. C., and Barai, S. V. (1995). "Multilayer perceptron in damage detection of bridge structures." *Computers and Structures*, 54(4), 597-608.

Phares, B. M., Rolander, D. D., Graybeal, B. A., and Washer, G. A. (2001). "Reliability of visual bridge inspection." *Public Roads*, 64(5), 22-29.

Ramaswamy, R. (1989). "Estimation of latent pavement performance from damage measurements." PhD dissertation, Dept. of Civil Engineering, MIT, Cambridge, Mass.

Robert, W. E., Marshall, A. R., Shepard, R. W., and Aldayuz, J. (2002). "The Pontis bridge management system: state-of-the-practice in implementation and development." *Cambridge Systematics Publications*.

Rokach, L. and Maimon, O. (2005). "Ensemble methods for classifiers." *Data Mining and Knowledge Discovery Handbook.* Springer, US. 957-980.

Rumelhart, McClelland, and the PDP Research Group. (1986). "A general framework for parallel distributed processing." *Parallel Distributed Processing, Vol. 1*. The MIT Press, Cambridge, MA

Sagues, A., Powers, R., and Kessler, R. (2001). "Corrosion performance of epoxy-coated rebar in Florida Keys bridges." Final Paper 01642, NACE International, Houston, TX.

Samples, L. and Ramirez, J. (1999). "Methods of corrosion protection and durability of concrete bridge decks reinforced with epoxy-coated bars – Phase I." Publication FHWA/IN/JTRP-98/15. , Indiana Department of Transportation and Purdue University, West Lafayette, Indiana.

Schapire, R. (1990). "The strength of weak learnability." *Machine Learning*. 5(2), 197-227.

Snyman, J. A. (2005). "Practical mathematical optimization: an introduction to basic optimization theory and classical and new gradient-based algorithms." Springer. 257. New York, NY.

Sobanjo, J. (1997). "*A neural network approach to modeling bridge deterioration."* Proc, 4th Congress on Computing in Civil Engineering. ASCE, New York, 623-626.

Stringer, D. and Burgueño, R. (2010). "Identification of Causes and Solution Strategies for Deck Cracking in Jointless Bridges Progress Report One." *Progress Report for the Michigan Department of Transportation*. Dept. of Civil and Environmental Engineering, Michigan State University., East Lansing, Michigan

Tu, J. (1996). "Advantages and disadvantages of using artificial neural networks versus logistic regression for predicting medical outcomes." *Journal of Clinical Epidemiology.* 49(11). 1225-1231.

Tumer, K. and Ghosh, J. (1999). "Linear and order statistics combiners for pattern classification." *Combining Artificial Nets*. Springer – Verlag. 127-162.

West, D., Dellana, S., and Qian, J. (2005). "Neural network ensemble strategies for financial decision applications." *Computers and Operations Research.* 32, 2543-2559.

Weyers, R. (1998). "Service life model for concrete structures in chloride laden environments." *ACI Materials Journal*. 95(4), 445-451.

Zhou, Z., Jiang, Y., Yang, Y., and Chen, S. (2002). "Lung cancer cell identification based on artificial neural network ensembles." *Artificial Intelligence in Medicine*. 24, 25-36.