

University of Mississippi

eGrove

---

Electronic Theses and Dissertations

Graduate School

---

8-1-2022

## Development and Evaluation of Performance Models for Asphalt, Concrete, and Composite Pavements using Machine Learning

Rulian Ferreira de Almeida Barros

Follow this and additional works at: <https://egrove.olemiss.edu/etd>

---

### Recommended Citation

Ferreira de Almeida Barros, Rulian, "Development and Evaluation of Performance Models for Asphalt, Concrete, and Composite Pavements using Machine Learning" (2022). *Electronic Theses and Dissertations*. 2369.

<https://egrove.olemiss.edu/etd/2369>

This Dissertation is brought to you for free and open access by the Graduate School at eGrove. It has been accepted for inclusion in Electronic Theses and Dissertations by an authorized administrator of eGrove. For more information, please contact [egrove@olemiss.edu](mailto:egrove@olemiss.edu).

DEVELOPMENT AND EVALUATION OF PERFORMANCE MODELS FOR ASPHALT,  
CONCRETE, AND COMPOSITE PAVEMENTS USING MACHINE LEARNING

A Dissertation

Presented for the Degree of

Doctor of Philosophy

Department of Civil Engineering

The University of Mississippi

Rulian Ferreira de Almeida Barros

August 2022

Copyright © 2022 by Rulian Ferreira de Almeida Barros  
ALL RIGHTS RESERVED

## ABSTRACT

Transportation infrastructures account for a considerable portion of public investments, which serve as the backbone of a country's economy by providing essential services to businesses and people. In the United States, public investments in transportation infrastructure assets represent trillions of dollars. The U.S road network consists of about 4 million miles, being the world's largest, longest, and biggest transportation system. Paved roads account for 2.6 million miles, and 93% of them are surfaced with asphalt. However, a portion of the paved roads consists of asphalt overlaid concrete pavements, also known as composite pavements. When concrete pavements start to fail, they are overlaid with Hot Mix Asphalt (HMA). Compared to flexible or rigid pavements, this offers better performance measures both structurally and functionally, and accordingly, it can be considered a cost-effective alternative.

Several performance indicators have been used to assess pavement surface conditions, but the Pavement Condition Rating (PCR) and the International Roughness Index (IRI) are the most widely used and well-recognized pavement performance indicators. Transportation agencies use these indexes to evaluate and classify the conditions for the road networks in the long term. If maintenance and rehabilitation (M&R) interventions are not performed timely, the pavement damage caused by environmental impacts and traffic repetitions can lead the roads to early deterioration. Billions of dollars are spent every year on M&R. However, a shortage in federal and state funds led roads and bridges to poor conditions since M&R interventions were not carried out timely. Therefore, there is a need to develop pavement performance prediction models that can

support and allow decision-makers to prioritize M&R actions due to the limited budget allocation and estimate the rate of pavement deterioration.

Traditionally, linear, non-linear, multiple linear regression analysis, Markov chains, mechanistic-empirical relations, survivor curves, semi-Markov, and Bayesian models have been used for predicting pavement performance. However, simple statistical approaches do not account for the complex relations among input variables and pavement performance. A growing body of literature is exploring the use of more advanced modeling techniques for pavement performance prediction. Among these techniques, the Artificial Neural Networks (ANNs) approach has shown the most significant improvements with consistent and reliable results. However, most performance models did not consider M&R history in the model development.

This doctoral research presents new pavement performance models incorporating the M&R history and activities for composite pavements of the LTPP database. Additionally, a more comprehensive approach was developed for flexible, rigid, and composite pavements of the Mississippi Department of Transportation (MDOT) database, accounting for the influence of M&R history. This dissertation successfully utilized the ANNs modeling technique to obtain accurate and promising prediction results for pavement performance. Furthermore, the development of a simple, low-cost, and easy-access graphical user interface (GUI) tool brings a significant contribution to the enhancement of agencies' pavement management system (PMS) by predicting future pavement conditions, identifying rehabilitation needs, and allowing a better budget allocation for critical pavement sections without the need of distress data.

## DEDICATION

I would like to dedicate this dissertation to my family and my parents who gave me the best conditions and education to prepare me for real life. In particular, my mother, Ana Lucia, who has always been by my side throughout my personal, academic, and professional journey. Thank y'all.

This dissertation is also dedicated to my wife who has also been by my side throughout this time and has always supported me. Thank you, Thamiris Vitoria. This dissertation would not have been possible without your love and support.

## LIST OF ACRONYMS

AASHTO	American Association of State Highway and Transportation Officials
ANNs	Artificial Neural Networks
ASE	Average Square Error
BPNN	Back-Propagation Neural Networks
CESAL	Cumulative Equivalent Single Axle Load
CN	Construction Number
CRCP	Continuously Reinforced Concrete Pavement
CS&O	Crack, Seat, and Overlay
ESAL	Equivalent Single Axle Load
FHWA	U.S. Department of Transportation's Federal Highway Administration
GEP	Gene Expression Programming
GPS	General Pavement Studies
GUI	Graphical User Interface
HMA	Hot-Mix Asphalt
IMS	Information Management System
IRI	International Roughness Index
IRI <sub>0</sub>	Initial International Roughness Index
IRI <sub>Left</sub>	Left (inside) Wheel Path International Roughness Index
IRI <sub>Mean</sub>	Mean International Roughness Index

IRI <sub>Right</sub>	Right (outside) Wheel Path International Roughness Index
IRRE	International Road Roughness Experiment
JCP	Jointed Concrete Pavement
JPCP	Jointed Plain Concrete Pavement
JRCP	Jointed Reinforced Concrete Pavement
LOGSIG	Logarithmic-Sigmoidal
LTPP	Long-Term Pavement Performance
M&R	Maintenance and Rehabilitation
MAPE	Mean Absolute Percentage Error
MARE	Mean Absolute Relative Error
MDOT	Mississippi Department of Transportation
MEPDG	Mechanistic-Empirical Pavement Design Guide
ML	Machine Learning
MLR	Multiple Linear Regression
NCHRP	National Cooperative Highway Research Program
NMDOT	New Mexico Department of Transportation
PCC	Portland Cement Concrete
PCI	Pavement Condition Index
PCR	Pavement Condition Rating
PMS	Pavement Management Systems



PSI	Present Serviceability Index
PSR	Present Serviceability Rating
R	Coefficient of Correlation
r	Pearson Correlation Coefficient
R <sup>2</sup>	Coefficient of Determination
RF	Random Forest
RMSE	Root Mean Squared Error
SD	Standard Deviation
SF	Site Factor
SHRP	Strategic Highway Research Program
SN	Structural Number
SPS	Specific Pavement Studies
SVM	Support Vector Machine
USDOT	United States Department of Transportation

## ACKNOWLEDGMENTS

I would like to express my deepest gratitude to Dr. Waheed Uddin, Professor and Director of the Center for Advanced Infrastructure Technology (CAIT), my first Ph.D. advisor, and the person responsible for my doctoral studies in the United States. Without your unstoppable desire to work, study, and research, nothing would be possible. Your incredible knowledge, guidance, support, and help made a difference in my life. I hope you are proud of all your student's accomplishments. Thank you very much, Dr. Waheed Uddin.

I am extremely grateful to my advisor, Dr. Hakan Yasarer, who accepted me as his Ph.D. student and guided me through the last years of my doctoral program. Thanks for all the support, knowledge, time, and friendship. It is always good to work with you and learn new abilities that can be applied not just to the academic field but also my personal and professional life. Special to Dr. Najjar who was always there to help me in my doctoral program. I am also extremely grateful to my friend and colleague, Dr. Salma Sultana, for helping me all the time. It would be extremely difficult to graduate if we did not share all of our questions, concerns, difficulties, and friendship.

I am also thankful to the Department of Civil Engineering for the financial support during all these years; Thanks, Dr. Cristiane Surbeck. I would like to thank my committee members Dr. Hakan Yasarer as the committee chair, Dr. Yacoub Najjar, Dr. Hunain Alkhateb, and Dr. Tyrus Mc Carty. Thank you for your contribution to my dissertation.

Words cannot express my gratitude to my family, especially my mother Ana Lucia, and my wife Thamiris Vitoria. Thanks for all your love and support throughout my Ph.D. program.

## TABLE OF CONTENTS

ABSTRACT.....	ii
DEDICATION.....	iv
LIST OF ACRONYMS .....	v
ACKNOWLEDGMENTS .....	viii
TABLE OF CONTENTS.....	ix
LIST OF TABLES .....	xiv
LIST OF FIGURES .....	xvi
CHAPTER I: INTRODUCTION.....	20
1.1. Overview .....	20
1.2. Problem Statement .....	23
1.3. Research Objectives .....	24
1.4. Dissertation Organization.....	25
CHAPTER II: BACKGROUND .....	29
2.1. Machine Learning .....	29
2.2. Overview of Pavement Performance Indicators.....	40
2.3. Long-Term Pavement Performance .....	43
2.4. Mississippi Department of Transportation Performance Models .....	46
CHAPTER III: LITERATURE REVIEW .....	49
3.1. Pavement Performance Models.....	49
CHAPTER IV: ROUGHNESS MODELING FOR COMPOSITE PAVEMENTS USING MACHINE LEARNING .....	62
4.1. Abstract .....	62

4.2.	Introduction .....	63
4.3.	Literature Review.....	65
4.4.	Model Development.....	68
4.5.	ANN Model.....	71
4.6.	Results and Discussions .....	74
4.7.	Sensitivity Analysis.....	76
4.8.	Conclusions .....	78
<b>CHAPTER V: THE EFFECT OF CLIMATE AND TRAFFIC ON COMPOSITE PAVEMENT</b>		
<b>ROUGHNESS MODELING USING MACHINE LEARNING .....</b>		
		<b>80</b>
5.1.	Introduction .....	80
5.2.	Literature Review of IRI Prediction Models.....	82
5.3.	Model Development.....	88
5.4.	ANN Development.....	99
5.5.	ANN Model Results.....	105
5.6.	In-depth Analysis .....	109
5.7.	Conclusion.....	112
<b>CHAPTER VI: INTERNATIONAL ROUGHNESS INDEX MODEL FOR COMPOSITE</b>		
<b>PAVEMENTS IN THE LTPP WET NON-FREEZE CLIMATE REGION: MACHINE</b>		
<b>LEARNING AND REGRESSION APPROACHES .....</b>		
		<b>115</b>
6.1.	Abstract .....	115
6.2.	Introduction .....	116
6.3.	Literature Review.....	119
6.4.	Model Development.....	122
6.5.	ANN Development.....	127
6.6.	Regression Model.....	131
6.7.	Discussion .....	132
6.8.	Comparison Between ANN and MLR Model Results.....	134

6.9. Conclusions .....	136
CHAPTER VII: COMPOSITE PAVEMENT ROUGHNESS MODELING FOR LTPP WET FREEZE CLIMATE REGION USING MACHINE LEARNING.....	139
7.1. Abstract .....	139
7.2. Introduction .....	140
7.3. Literature Review .....	142
7.4. Model Development.....	143
7.5. Ann Development .....	146
7.6. Discussion .....	149
7.7. Model comparison.....	150
7.8. Conclusions .....	152
CHAPTER VIII: PAVEMENT PERFORMANCE MODELING FOR COMPOSITE PAVEMENTS IN THE LTPP WET FREEZE CLIMATE REGION INCORPORATING MAINTENANCE AND REHABILITATION .....	154
8.1. Introduction .....	154
8.2. Model Development.....	155
8.3. Discussion .....	159
8.4. Conclusions .....	160
CHAPTER IX: PAVEMENT PERFORMANCE MODELING CONSIDERING MAINTENANCE AND REHABILITATION FOR COMPOSITE PAVEMENTS IN THE LTPP WET NON-FREEZE REGION USING NEURAL NETWORKS .....	161
9.1. Abstract .....	161
9.2. Introduction .....	162
9.4. Model Development.....	164
9.5. Discussion .....	173

9.6.	Implementation of Model 1 for Random Sections .....	175
9.7.	Conclusion.....	176
CHAPTER X: DEVELOPMENT OF PERFORMANCE MODELS FOR FLEXIBLE PAVEMENTS USING THE MDOT DATABASE .....		178
10.1.	Model Development .....	178
10.2.	MDOT Flexible Pavement Performance Model.....	181
10.3.	Implementation of the Developed Models via GUI for the ANN MDOT Asphalt Model 190	
10.4.	Recommendations from the ANN MDOT Model.....	192
CHAPTER XI: PERFORMANCE PREDICTION MODEL FOR JOINTED CONCRETE PAVEMENTS IN MISSISSIPPI USING MACHINE LEARNING.....		194
11.1.	Introduction .....	194
11.2.	Literature Review .....	196
11.3.	Model Development .....	198
11.4.	Discussion.....	204
11.5.	Implementation of Model 1 via GUI .....	206
11.6.	Conclusion .....	208
CHAPTER XII: PERFORMANCE EVALUATION OF COMPOSITE PAVEMENTS ON MISSISSIPPI HIGHWAYS VIA MACHINE LEARNING .....		210
12.1.	Abstract.....	210
12.2.	Introduction .....	211
12.3.	Literature Review .....	213
12.4.	Model Development .....	214
12.5.	Results And Discussion .....	217
12.6.	Conclusions .....	221
CHAPTER XIII: GRAPHICAL USER INTERFACE (GUI) .....		223

13.1. Graphical User Interface Development .....	223
CHAPTER XIV: SUMMARY, CONCLUSIONS, AND RECOMMENDATIONS .....	227
14.1. Summary.....	227
14.2. Conclusions .....	231
14.3. Recommendation for Future Research .....	237
LIST OF REFERENCES .....	239
APPENDIX.....	252
Appendix A Pavement Condition Rating Forms and Key Forms [57] .....	253
VITA.....	259

## LIST OF TABLES

Table 1. Independent and Dependent Variable Configuration for Five ANN Models.....	72
Table 2. ANN Model Results.....	73
Table 3. Summary of Climate Zones .....	90
Table 4. Descriptive Statistics of Output Variables.....	91
Table 5. Descriptive Statistics of Input Variables .....	97
Table 6. Categorical Variables Summary .....	98
Table 7. Input and Output Variables Configuration .....	102
Table 8. Summary of Model 1 Network Statistics.....	103
Table 9. Summary of Model 2 Network Statistics.....	104
Table 10. Summary of Model 3 Network Statistics.....	105
Table 11. Summary of Best Network Performance for each ANN Model.....	106
Table 12. Descriptive Statistics of Input Variables .....	127
Table 13. Summary of ANN Model Network Statistics.....	129
Table 14. Descriptive Statistics of Input Variables .....	146
Table 15. Input Variables Configuration for Five ANN Models.....	147
Table 16. ANN Model Results.....	149
Table 17. Independent and Dependent Variable Configuration for Five ANN Models.....	156
Table 18. ANN Model Results.....	158
Table 19. States Located in the Wet Non-freeze LTPP Climate Zone .....	166



Table 20. Input and Output Variables for ANN Models .....	171
Table 21. ANN Model Results.....	172
Table 22. Independent and Dependent Variables for MDOT Asphalt Pavements .....	184
Table 23. ANN Model Results for MDOT Flexible Pavements (Complementary PCR Output).....	185
Table 24. ANN Model Results for MDOT Flexible Pavements (IRI Output).....	186
Table 25. Independent and Dependent Variables for ANN Models .....	202
Table 26. ANN Model Results for JCP Pavements (Complementary PCR and IRI Outputs) ...	203
Table 27. Independent and Dependent Variables Configuration.....	216
Table 28. ANN Model Results for Complementary PCR Output .....	218
Table 29. ANN Model Results for IRI Output .....	218

## LIST OF FIGURES

Figure 1. Example of Typical ANNs Architecture .....	31
Figure 2. Sigmoidal Function .....	34
Figure 3. Dynamic-Sequential Network Structure [46] .....	40
Figure 4. Pavement Condition Rating (PCR) Scale [57] .....	42
Figure 5. LTPP Climate Zone Map [59] .....	45
Figure 6. $IRI_{Mean}$ Measurement (m/km) .....	69
Figure 7. $IRI_{Mean}$ Measurement, Section 06-7455, California, 1989-2015 .....	70
Figure 8. Network Architecture of the Best Performing ANN Model (Model 5) .....	74
Figure 9. Graphical Comparison of Accuracy Measures between Developed ANN Models .....	74
Figure 10. Observed and Predicted $IRI_{Mean}$ .....	75
Figure 11. Observed vs. Predicted $IRI_{Mean}$ for ANN Model 5 .....	76
Figure 12. Observed and Predicted Plot of $IRI_{Mean}$ for Section 01-0604 .....	76
Figure 13. ANN Future Prediction $IRI_{Mean}$ for Asphalt Overlay on Concrete Section 01-0604 ..	77
Figure 14. LTPP Climate Zones .....	90
Figure 15. $IRI_{Left}$ and $IRI_{Right}$ Measurements .....	92
Figure 16. Two-tailed t-test Probability Distribution Graph .....	93
Figure 17. Independent Variables Versus $IRI_{Right}$ .....	95
Figure 18. ANN Network Plot .....	99
Figure 19. Sigmoidal Function .....	100

Figure 20. Graphical Comparison of Accuracy Measures .....	107
Figure 21. Network Architecture of the Best Performing ANN Model (model 3).....	107
Figure 22. Model 3 - Observed vs Predicted Values for $IRI_{Right}$ .....	108
Figure 23. Observed vs. Predicted $IRI_{Right}$ for ANN Model 3 .....	109
Figure 24. $IRI_{Right}$ Observed vs Predicted for Section 04-0607 - Arizona.....	109
Figure 25. $IRI_{Right}$ Observed vs Predicted for Section 05-A607 - Arkansas.....	110
Figure 26. $IRI_{Right}$ Observed vs Predicted for Section 18-0607 – Indiana.....	111
Figure 27. $IRI_{Right}$ Observed vs Predicted for Section 48-7165 - Texas .....	112
Figure 28. Pavement Performance Modeling Methodology.....	122
Figure 29. LTPP Climate Zones [2,59].....	124
Figure 30. $IRI_{Left}$ and $IRI_{Right}$ Measurements .....	125
Figure 31. Independent Variables Versus $IRI_{Right}$ .....	126
Figure 32. Architecture of the Best-Performing Network for the ANN Model.....	130
Figure 33. Observed vs. predicted $IRI_{Right}$ for ANN model.....	132
Figure 34. Observed vs. predicted $IRI_{Right}$ for MLR model.....	133
Figure 35. Observed vs. Predicted $IRI_{Right}$ Values for Verification Sections .....	134
Figure 36. $IRI_{Right}$ Section 05-A608 - ANN and MLR Comparison.....	135
Figure 37. $IRI_{Right}$ Section 41-7018 - ANN and MLR Comparison.....	136
Figure 38. Independent Variables Versus $IRI_{Right}$ .....	145
Figure 39. Graphical Comparison of Accuracy Measures .....	149

Figure 40. Observed Versus Predicted $IRI_{Right}$ for ANN Model 5 .....	150
Figure 41. Observed and Predicted Plots of $IRI_{Right}$ for Sections 29-0661 and 46-7049.....	151
Figure 42. Observed vs. Predicted $IRI_{Right}$ for ANN Model 1 and ANN Model 2, respectively	159
Figure 43. Pavement Performance Modeling Methodology.....	165
Figure 44. $IRI_{Right}$ measurements, Section 06-0661, California, 1993-2015 .....	169
Figure 45. Graphical comparison of accuracy measures for all developed models.....	173
Figure 46. Model 1 Predicted vs. Observed $IRI_{Right}$ .....	174
Figure 47. Model 1 Observed and Predicted Plots of IRI for Sections 40-0603 and 48-5154...	175
Figure 48. Pavement Performance Modeling Methodology for the MDOT Database.....	179
Figure 49. Spatial Map of MDOT Flexible Pavement Sections .....	182
Figure 50. ANN Architecture for the Best Performing Model for MDOT Asphalt Pavement ..	187
Figure 51. Graphical Comparison of Accuracy Measures for MDOT ANN Models.....	188
Figure 52. Observed vs. Predicted Complementary PCR for MDOT Flexible ANN Model 5 ..	189
Figure 53. Observed vs. Predicted $IRI_{Mean}$ for MDOT Flexible ANN Model 5 .....	189
Figure 54. Preliminary GUI Application for ANN Model for MDOT Asphalt Pavement.....	190
Figure 55. Observed and Predicted Plot of PCR for MDOT Asphalt Pavement Section #14....	191
Figure 56. Observed and Predicted Plot of $IRI_{Mean}$ for MDOT Asphalt Pavement Section #14	192
Figure 57. MDOT PCR Scale [93] .....	197
Figure 58. Graphical Comparison of Accuracy Measures for All Developed Models.....	205
Figure 59. Observed vs. Predicted PCR and IRI for Model 1 .....	205

Figure 60. GUI Application for Model 1 .....	207
Figure 61. Observed and Predicted Plot of PCR and IRI for JCP Pavement Section #903 .....	208
Figure 62. Observed Versus Predicted PCR for All Developed Models .....	219
Figure 63. Observed Versus Predicted IRI for All Developed Models .....	220
Figure 64. GUI for Flexible Pavement .....	223
Figure 65. User Form for Approximate ESAL Increase .....	224
Figure 66. Pavement Performance Predictions .....	225
Figure 67. Predictions for PCR and IRI.....	225

## CHAPTER I: INTRODUCTION

### 1.1. Overview

Transportation infrastructures (roads, railways, airports, transit), account for a considerable portion of public investments, which serve as the backbone of a country's economy and society by providing essential services to businesses and people [1]. A country's road network is considered one of the largest public infrastructure assets worldwide [2]. Transportation infrastructure assets represent public investments of trillions of dollars in the United States (U.S.) [1]. The U.S. road network consists of about 4 million miles, comprising 1.4 million miles of unpaved roads and 2.6 million miles of paved roads. About 93 percent of the 2.6 million miles of paved roads and highways are surfaced with asphalt [2]. However, a portion of the paved highway surfaces comprises composite pavements, which are made of asphalt overlaid concrete pavements. Most composite pavements are a result of concrete pavement rehabilitation. When concrete pavements start to fail, they are overlaid with Hot Mix Asphalt (HMA). Compared to flexible or rigid pavements, composite pavements can offer better performance levels both structurally and functionally and accordingly can be a more cost-effective alternative [3].

Over the last decades, state and federal transportation agencies have established several performance indicators such as Present Serviceability Rating (PSR), Present Serviceability Index (PSI), Pavement Condition Rating (PCR), Pavement Condition Index (PCI), and International Roughness Index (IRI) [1] to evaluate the effectiveness and efficiency of their service provision

[4]. Among all pavement condition indices used to assess pavement surface conditions, the PCR and IRI are the most widely used and well-recognized pavement performance indicators. The Mechanistic-Empirical Pavement Design Guide (MEPDG) [5] designed to update the 1993 American Association of State Highway and Transportation Officials (AASHTO) [6] uses the IRI measurements of longitudinal roughness to indicate pavement smoothness. The IRI measurements are based on the quarter-car analysis method, a mathematical model of a vehicle that represents a body and a single wheel [7], with standardized parameter values and a reference simulation speed of 80 km/h [8]. It can be expressed in two types of units, in/mile or m/km. A higher IRI value indicates a rough pavement profile, which affects the ride quality experienced by road users. A lower IRI value indicates a smooth pavement profile, causing a better ride quality for the road users. If pavement smoothness decreases induced by climatic and traffic attributes, the IRI value increases [1]. The PCR is a rating method based upon pavement distresses. Although the relationship between pavement distress and performance is hard to understand, the ability of pavement to sustain traffic loads safely and smoothly is adversely affected by the incidence of observable distresses. Hence, the PCR offers a weighted measure for uniformly identifying and describing pavement distresses in terms of severity and extent.

Efficient and well-maintained road networks are essential to ensure acceptable conditions for the road networks in the long term. If maintenance and rehabilitation actions are not performed timely, the pavement damages inflicted by environmental impacts and traffic repetitions may lead to poor conditions rapidly that can cause life-threatening for road users [2]. A substantial amount of financial support is required for the maintenance and rehabilitation of a road network. However, the United States Department of Transportation (USDOT) encountered a shortage in its funds as

of 2015 since most of the funds were transferred to state agencies for new constructions and maintenance and rehabilitation (M&R) treatments of roads and bridges in poor condition. Thus, a slow reimbursement rate can delay M&R actions leading to a higher degree of deterioration for the infrastructure assets [9]. Without proper funds, local and state agencies cannot maintain all highway pavements at acceptable levels. For this reason, pavement performance prediction models are an essential part of agencies' Pavement Management Systems (PMS), allowing decision-makers to prioritize M&R actions due to budget allocation and reduce the rate of pavement deterioration. Performance models provide an estimation of pavement conditions, rehabilitation needs, and enable agencies to prioritize the worst sections. Furthermore, performance models are easy to understand and can deliver deeper insights by converting performance indices into operational measures to inform how long and how well the road will continue to serve the users [10]. However, some agencies utilize probability models that do not include essential parameters such as pavement design parameters, rehabilitation interventions, and traffic, leading to inaccurate and ineffective M&R decisions.

For this reason, there is a need to develop more inclusive, reliable, and accurate pavement performance models that can estimate future pavement conditions, identify rehabilitation needs and analyze rehabilitation impacts. A growing body of literature explores different modeling techniques for pavement performance prediction, and several studies explore the use of regression models to predict pavement roughness. However, due to the complex relations between pavement structure and its behaviour under climatological and traffic variables and the interaction among all these elements [11], more advanced modeling techniques using machine learning proved to be promising. The Artificial Neural Networks (ANNs) approach was used in several studies [2,10,12–



20], offering significant improvements over traditional techniques, such as regression, for processing large volumes of data with a higher degree of accuracy.

This dissertation presents the use of ANNs approach for developing performance prediction models, new approaches incorporating M&R, exploring key inputs that are not considered in most models of the literature, and develops a graphical user interface (GUI) to implement the best-developed models for real-life use.

## 1.2. Problem Statement

One of the main concerns of transportation agencies is to ensure appropriate conditions for the road networks in the long term. If frequent maintenance and rehabilitation are not performed, the pavement damage caused by environmental impacts and traffic repetitions may cause an early deterioration decreasing the pavement life. Performance models are imperative to assess road's future conditions, estimate pavement deterioration, identify M&R needs, and prioritize budget allocation for a better management system.

Traditionally, linear, non-linear, multiple linear regression analysis, Markov chains, mechanistic-empirical relations, survivor curves, semi-Markov, and Bayesian models have been used for predicting pavement performance [20]. However, simple statistical approaches such as linear regression do not result in significant accuracy measures for performance predictions due to the complexity/non-linearity of the relationships among variables [21]. Numerous research investigations [2,9,15,19,22–27] have been conducted to discover different methods that can generate efficient, rational, and practical prediction models. Among these approaches, the ANN technique has been showing significant improvements with efficient and reliable results. However, the literature indicates that most performance models did not consider M&R history in the model

development. Other key features such as climatological factors and geographical location were also not included in many models.

Recently, some studies [2,9,28–36] included M&R history as input variables for flexible and rigid pavements using the Long-Term Pavement Performance (LTPP) database. However, there were no studies including M&R interventions for composite pavements using LTPP. For other databases different than the LTPP, there is also a lack of research models that account for the influence of M&R history and activities during the service life of the pavement, which affects the accuracy and reliability of the prediction [37].

Therefore, this dissertation presents research papers, in which performance models incorporating the effects of M&R history and activities for composite pavements and the development of prediction models for specific climate zones of the LTPP database. Additionally, a more complete and inclusive approach considering maintenance and rehabilitation for asphalt, concrete, and composite pavement performance models using the Mississippi Department of Transportation (MDOT) database is presented. The developed models in this doctoral study can be utilized by agencies as a tool for predicting future pavement conditions and incorporating the M&R scheduling effectively to prioritize the resources. Furthermore, the developed models do not use any distress data as an input variable, which will help transportation agencies to save time from data collection/processing to assess the condition of the pavement.

### 1.3. Research Objectives

The overall objective of this dissertation is to study, analyze, and identify key variables that affect pavement performance to develop ANN models that can contribute to the state-of-the-art by providing a more inclusive, reliable, and user-friendly tool. This tool can support objective

decisions regarding maintenance and rehabilitation interventions and budget plans allowing agencies to prioritize the resources for critical pavement sections. According to the stated problems in Section 1.2, the following tasks are included:

1. Complete literature review of asphalt, concrete, and composite pavement performance models,
2. Develop performance models for composite pavements using the LTPP database and verify the model performance,
3. Develop performance models for asphalt, concrete, and composite pavements using the MDOT database and verify the model performance,
4. Develop a GUI for implementation of the best-developed models,
5. Evaluate the developed models using GUI for the enhancement of pavement asset management.

As stated before, this study explores the use of an advanced modeling technique to enhance pavement performance modeling. The use of new approaches to incorporate M&R history and other significant variables are explored and explained throughout the dissertation. The following section explains how the dissertation is organized and discussed in detail.

#### 1.4. Dissertation Organization

Chapter I presents an overview of the main aspects of the dissertation followed by the problem statement, research objectives, and the dissertation organization sections.

Chapter II describes the background of machine learning and artificial neural networks followed by an overview of pavement performance indicators, the LTPP program, and the MDOT database.

Chapter III provides an extensive literature review of pavement performance models for flexible, rigid, and composite pavements.

The following chapters of the dissertation are divided into peer-reviewed papers that were submitted, published, and presented at national and international conferences.

Chapter IV presents a paper that discusses the lack of research on composite pavement performance models and the benefits of using the ANN modeling technique. It introduces a new approach incorporating M&R history and actions into model development and considering all four climate zones of the LTPP database. The paper did not include all data points of the database in the analysis and recommended that future studies should use the entire database and also investigate the use of specific climate zones in the model development.

Chapter V presents a paper that discusses the effects of climate and traffic on composite pavements and develops performance models using all data points of the LTPP composite pavement database. It uses the same approach as the previous paper to include M&R history and introduces the use of Cumulative Equivalent Single Axle Load (CESAL) to account for the traffic load repetition history.

Chapter VI presents a paper that studies the use of a specific climate zone and the inclusion of M&R history in the development of performance models using a traditional (multiple regression) and an advanced (ANN) modeling technique. The paper compares the results of both methods utilizing statistical measures to identify the most accurate approach. The use of a specific climate zone is also studied to identify if it helps the models to achieve better accuracy results and explore it in future studies.

Chapter VII shows a paper that develops pavement performance models for a specific climate zone different than the previous. It includes M&R history and climatological factors to identify which model offers the most accurate results. The use of models that do not include any

distress data can help agencies to reduce costs and offer a more effective M&R plan.

Chapter VIII presents a paper that discusses and develops a new method for incorporating M&R history into the performance model development. It compares two models for composite pavements located in a specific climate zone of the LTPP database and utilizes two different approaches to identify which one delivers better accuracy. This new approach can be used in different climate zones to verify it will also enhance other pavement performance models.

Chapter IX presents a paper that utilized the two different approaches for M&R history developed earlier, for a different climate zone of the LTPP composite pavement database. A combination of climatological and traffic variables is also used to improve model accuracy and a comparison of all models is performed to identify the most accurate model.

Chapter X presents the results of a research project for the MDOT. This study develops a new M&R approach that can be used to identify rehabilitation needs and predict future pavement conditions. The models also account for the effects of several key factors such as geographic location, pavement structure, drainage, and traffic. A two-output model approach is considered and resulted in promising results that can be further studied in rigid and composite pavement types.

Chapter XI presents a paper focused on the model development of rigid pavements for the MDOT database using the ANN modeling technique. This paper utilizes some of the approaches developed in the previous chapter and applies them to a rigid pavement type. It also compares models with one and two outputs to identify the most accurate model.

Chapter XII presents a paper that develops ANN performance models for composite pavements. The paper utilizes different input variables to evaluate which model would give the most accurate predictions. The use of M&R history and actions, traffic repetition, and a

combination of both variables generate promising results. A two-output model approach is used and the best model is selected based on three accuracy measurements.

Chapter XIII shows a graphical user interface (GUI) tool developed in this dissertation to implement the best-developed models. This low-cost and easy-access tool can be used by federal and state agencies to predict future pavement conditions, identify rehabilitation needs, and prioritize the resources for the critical pavement sections.

Chapter XIV presents a summary of the research work followed by major conclusions of each chapter and recommendations for future research.

## CHAPTER II: BACKGROUND

### 2.1. Machine Learning

#### 2.1.1. Overview of Machine Learning

Machine Learning (ML) is the science of making computers learn and act intelligently and improving their learning over time by feeding them data and information with observations and real-world interactions. The fundamental goal of the ML algorithms is to generalize beyond the training samples to successfully interpret data that it has never seen before [38]. Several types of machine learning algorithms (i.e., K-nearest neighbor, support vector machines, naive Bayes, logistic regression, decision trees, artificial neural networks, Bayesian networks, conditional random fields, etc.) have been developed and used to process large volumes of data with high degrees of accuracy, handle noisy and complex data, solve non-linear problems, and once trained, make predictions and generalizations at any time [20,39].

The machine learning techniques hold significant potential for building a modern and robust pavement system management due to the excellence in automation and pattern recognition [20]. The literature review shows that artificial neural networks are not only one of the first machine learning techniques to be used but also the most used technique in civil and pavement engineering [40,41]. Because ML has a data-driven approach, IRI appears as a suitable indicator for modeling, since it is widely available in pavement databases (e.g., LTPP database), measured by objective means (e.g., laser profilometer), and known as one of the most common indicators

for pavement performance evaluation [42]. In this dissertation, an artificial neural network technique was used for the development of performance prediction models for flexible, rigid, and composite pavements.

## 2.1.2. Artificial Neural Networks

### 2.1.2.1. Overview of ANN

An artificial neural network is an information-processing system based on mathematical models that use the concept of human cognition and neural biology [43]. The ANN method attempts to emulate the structure and/or functional aspects of biological neural networks [44]. It consists of several simple processing elements called neurons (or nodes) and connecting links between them. When the information is processed, the connection links are used to transfer signals between neurons [43]. Each neuron evaluates its input signals to determine its output signal and transmits it to all neurons that are on the receiving side of the connection links originating in the transmitting neuron. Each connection has an associated weight that multiplies the signal transmitted [43]. Complex relationships that are difficult to reproduce using traditional sequential, logic-based modeling and computation techniques can be successfully represented by neural networks. However, the accuracy of ANN models is highly dependent on the accuracy of the database used to train the neural network. For this reason, the database cannot contain a significant amount of erroneous data or be too small, otherwise, the ANN model will generate significantly inaccurate or wrong predictions [44]. There are many types of neural networks characterized by their architecture, training algorithm, and activation function [45] as explained in the next sections.

### 2.1.2.2. ANN Elements and Architecture

The most simple and essential element of a neural network is called a neuron, which imitates the biological neurons from the nervous system. These neurons are a part of the ANN architecture that



also consists of four main elements: input layer, hidden layer(s), an output layer, and connection weights [28]. Figure 1 shows an example of a typical ANN architecture.

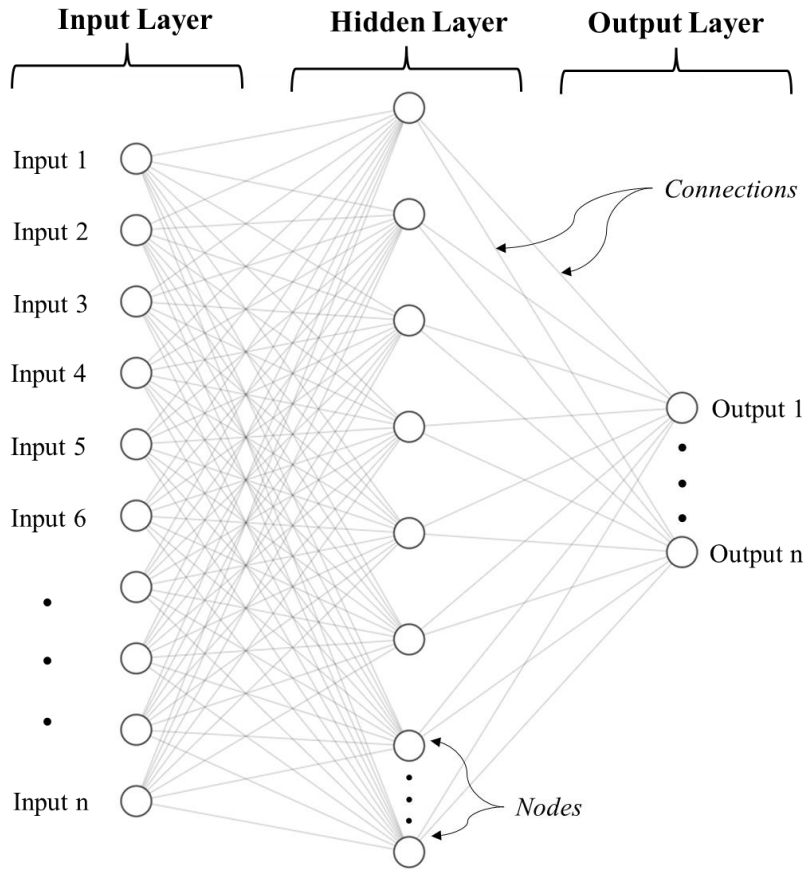


Figure 1. Example of Typical ANNs Architecture

The layers presented in Figure 1 are described as follows [9]:

- **Input Layer:** It consists of independent variables that are used in the model.
- **Hidden Layer(s):** The hidden layer(s) can consist of one or more layers, and each layer can contain a different number of hidden nodes.
- **Output Layer:** It consists of the dependent variable used in the model. It can contain one or more output nodes.

### 2.1.2.3. Feed-Forward Network and Backpropagation Learning Algorithm

In this study, a feed-forward neural network with a back-propagation training algorithm

was used for the development of performance prediction models for flexible, rigid, and composite pavements. The neural network gains its knowledge through a trained feed-forward network that uses a set of training data consisting of inputs (independent variables) and output(s) (dependent variable(s)). The resulted output is compared to the target values and the back-propagation process adjusts the connection weight to reduce the error between actual and target values [2]. After training, the network offers an approximate functional mapping of any input pattern onto its corresponding output pattern. Then, the validation process was carried out using datasets that were excluded from the model database [2]. After the validation process, it is necessary to retrain the best-performing network using all experimental data to increase the prediction accuracy and account for all patterns in the database [44].

This study used different databases that contain both categorical and continuous variables. For this reason, the model development considered only one hidden layer. The use of more than one hidden layer combined with an insufficient number of databases may cause the network to memorize the data in the training phase. Therefore, the developed model used only one hidden layer to maintain the generalization capability of the network [46].

#### 2.1.2.4. Learning Algorithm

##### 2.1.2.4.1. Nodal Input Values

The nodes from the input layer are connected to the hidden layer nodes and subsequently to the output layer nodes as shown in Figure 1. Node values are multiplied by the specific connection weights added to calculate a total sum of weights that will be transferred for the next node. A bias is also added as an additional set of weights and carried in the calculation. The sum of weights along with a bias is used to adjust the output of the hidden node, which will be the new feedforwarded value for the next node [9]. Sultana exemplified the calculation of an arbitrary node

“A” at a hidden layer; the node value is the sum of the value of the weights from the input layer.

Equation 1 expresses the input value for a node “A” [28,44]:

$$Node_A = \sum_{i=1}^n [(Input\ Node\ value)_i \times (connection\ weight)_i] + bias \quad Eq. 1$$

#### 2.1.2.4.2. Activation Function: Sigmoidal Function

Many activation functions can be used to introduce non-linearity in artificial neural networks. The use of non-linear functions allows the model to learn complex relationships from the database and turn the model into a universal approximator. Bipolar sigmoidal, logistic sigmoidal, and binary steps are an example of some available functions. Specific applications might require the use of specific functions with different ranges and properties. However, the activation function must be continuous, differentiable, and monotonically non-decreasing to be applied in the backpropagation neural network [9,47].

The feed forwarded information at the nodes in the hidden layer(s) and output layer need to pass through the activation function to introduce the nonlinearity into the network. Nonlinear transformations that occur in all nodes of the hidden and output layer(s) can be simplified using Equation 2 for an arbitrary node “A” [9,44,47]:

$$Out_A = f(Net_j^l)_A \quad Eq. 2$$

Where:

- f: activation function
- (input)<sub>A</sub>: input for node A, computed using Equation 3.

A sigmoidal function was used as the activation function in this study. The sigmoidal function is especially advantageous for use in backpropagation networks because the simple relationship between the value of the function at a point and the value of the derivative at that point reduce the computational load during the training phase [45]. An output value with a specific

interval between 0 and 1 is expected for this function [44]. Figure 2 shows the graphical representation of the sigmoidal activation function that can be mathematically expressed using Equation 3.

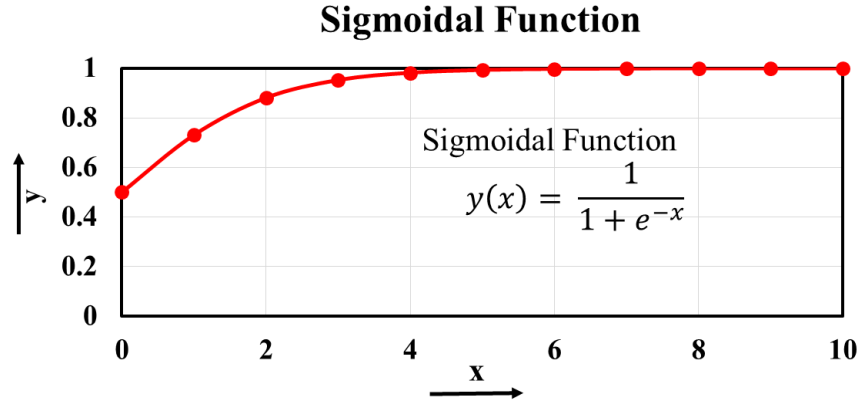


Figure 2. Sigmoidal Function

$$y(x) = \frac{1}{1 + e^{-x}} \tag{Eq. 3}$$

#### 2.1.2.4.3. Weight Adjustment

The predicted values resulting from the output node are compared to the actual (targeted) value and the error calculated from this comparison is used to adjust the connection weights. Different propagating error methods can be used to adjust the connection weights. The most commons are Levenberg-Marquardt, Perceptron's, and Gradient Descent [9,48]. In this study, the gradient descent method was used due to its simplicity, stability, and effectiveness. The gradient descent method propagates the error from the output layer to the preceding layers using the derivatives of the activation function [9,28]. The weight's incremental adjustments can be calculated using Equation 4 [44].

$$\Delta w_{ji}^L = w_{ji}^{L(new)} - w_{ji}^{L(old)} \tag{Eq. 4}$$

Where:

- New: current iteration
- Old: previous iteration

Gasteiger and Zupan [49] used the Delta rule to calculate the backpropagation neural network algorithm's incremental change (Equation 5).

$$\Delta w_{ji}^L = n\delta_j^L Out_i^{L-1} \quad Eq. 5$$

Where [49]:

- n: learning rate
- $\delta$ : represents the weighted error of the connection ji
- $Out_i^{L-1}$ : outcome from the  $i^{th}$  neuron in the  $(L-1)^{th}$  layer

#### 2.1.2.4.4. Learning Process

The learning process can be summarized in six steps [44]:

1. Input vectors are identified as Input 1, Input 2, ..., Input n, where n indicates the total number of variables (Figure 1).
2. Propagate the input vectors, Input 1, Input 2, ..., Input n via the connection weights to generate the output vectors.
3. Itemize the initial weights and update the connection weights on the output layer.
4. Update all weights in the hidden layer(s).
5. Repeat steps 1 through 4 for each training dataset.
6. Repeat steps 1 through 5 until the predicted output meets the corresponding target output within a predetermined tolerance or the training iterations reach the maximum limit.

#### 2.1.2.4.5. Number of Hidden Nodes

The user is responsible to specify the number of initial and maximum hidden nodes in the ANN model development. The ANN process begins with the user-specified initial hidden node and goes up to the maximum allowed number predetermined. At the end of this process, the ANN

structures with the least number of hidden nodes and the best statistical accuracy errors are selected to be re-evaluated in terms of statistical accuracy measures as well as graphical accuracy measures. Equation 6 can be used to calculate the maximum number of hidden nodes [44].

$$\text{Max. Number of Hidden Nodes} \leq \frac{(\text{number of training datasets}) - (\text{number of output variables})}{(\text{number of input variables}) + (\text{number of output variables}) + 1} \quad \text{Eq. 6}$$

Yasarer pointed out that choosing too many hidden nodes may lead to an overtraining situation. On the other hand, a few numbers of hidden nodes may not be sufficient to capture the behavior of complex phenomena. To utilize the generalization capability of the neural networks approach, this study uses networks with one hidden layer [44].

#### 2.1.2.5. Model Selection Criteria

Three statistical accuracy measures were used to compare the performance of the developed networks and to select the best performing network. The three measures are the Average Square Error (ASE), the Mean Absolute Relative Error (MARE), and the Coefficient of Determination ( $R^2$ ). During the evaluation process, the training, testing, validation, and overall performance statistics need to be considered. The best-performing model is chosen based on the lowest ASE, lowest MARE, and highest  $R^2$ , which indicates the level of agreement between predicted and actual output values. Equation 7 shows the ASE calculation [50].

$$ASE = \frac{\sum_{i=1}^N \sum_{j=1}^n (Y_{ij}^P - Y_{ij}^O)^2}{N.n} \quad \text{Eq. 7}$$

Equation 8 expresses the MARE calculation [50].

$$MARE = \frac{\sum_{i=1}^N \sum_{j=1}^n \left| \frac{Y_{ij}^P - Y_{ij}^O}{Y_{ij}^O} \right|}{N.n} \quad \text{Eq. 8}$$

Where:

- $Y_{ij}^P$  = Predicted output
- $Y_{ij}^O$  = Actual output
- N = Number of datasets
- n = Number of outputs

Normalization of the input values is performed to prevent the ANN models from being biased towards a specific input. Equations 9 and 10 show the data normalization formula for input variables, while Equations 11 and 12 show the output variables [9].

$$\frac{X_{Max} - ANN_{XMin}}{ANN_{XMax} - ANN_{XMin}} = 0.8 \quad Eq. 9$$

$$\frac{X_{Min} - ANN_{XMin}}{ANN_{XMax} - ANN_{XMin}} = 0.2 \quad Eq. 10$$

$$\frac{Y_{Max} - ANN_{YMin}}{ANN_{YMax} - ANN_{YMin}} = 0.9 \quad Eq. 11$$

$$\frac{Y_{Min} - ANN_{YMin}}{ANN_{YMax} - ANN_{YMin}} = 0.1 \quad Eq. 12$$

Where:

- X = Value of each independent variable
- $X_{max}$  = Maximum X
- $X_{min}$  = Minimum X
- Y = Value of dependent variable
- $Y_{max}$  = Maximum Y
- $Y_{min}$  = Minimum Y
- $ANN_{xmax}$  = Maximum X value normalized concerning the value on the right side of the equation
- $ANN_{xmin}$  = Minimum X value normalized concerning the value on the right side of the

equation

- $ANN_{Y_{max}}$  = Maximum Y value normalized concerning the value on the right side of the equation
- $ANN_{Y_{min}}$  = Minimum Y value normalized concerning the value on the right side of the equation

#### 2.1.2.6. Summary of ANN Model Development Stages

The ANN model development and the desired criteria to choose the optimal network structures can be described in four successive stages [51], as follows:

- Stage 1: Determine the ANN architecture. Decide input and output categories based on problem characteristics and ANN knowledge. Classify the datasets as training, testing, and validation sets.
- Stage 2: Train and test the network on the experimental data to obtain the optimum number of hidden nodes and iterations for the ANN architecture defined in the previous stage. Determine the best-performing networks based on the lowest ASE, lowest MARE, and highest  $R^2$  values.
- Stage 3: Validate the best-performing network from the second stage using the validation database. Check if the accuracy results from the training, testing and validation database are comparable. If they are, then stage four maybe not be necessary.
- Stage 4: Retrain the best performing network from stage 2 using all experimental data to increase prediction accuracy and account for all patterns in the database.

Typically, retraining the selected final network with all experimental data is expected to deliver reliable predictions and overall better accuracy measures since all the knowledge in the



database are incorporated into the final network [51]. Research studies by Najjar and co-workers [43,51–54] recommend that stage four is necessary to arrive at a better-performing network model. In this study, the TR-SEQ1 computer program [55] was used to develop the ANN models.

#### 2.1.2.7. Dynamic-Sequential ANN Modeling

A dynamic-sequential ANN modeling technique was also used in this study to develop pavement performance models. The dynamic-sequential ANN-based training technique adopted by Najjar [55] and Yasarer [46] is used to model the time-dependent pavement performance. The dynamic-sequential technique uses the framework of the conventional feed-forward error-backpropagation neural network approach [46,56]. According to the feedback approach, the futuristic (i.e., year (n+1)) desired value (i.e.,  $(A)^{n+1}$ ) is determined from some predetermined input parameters. This logic is mathematically represented by Equation 13 [56].

$$\{(A)^{n+1}\} = ANN_{(m+1)-k-1}\{x_1, x_2, \dots, x_m(A)^n\} \quad Eq\ 13$$

Where ANN denotes the neural network model that best relates a given number of inputs (m+1) [i.e.,  $x_1, x_2, \dots, x_m, (A)^n$ ] to the desired output [i.e.,  $(A)^{n+1}$ ]. Note that  $\{x_1, x_2, \dots, x_m\}$  is a vector of (m) parameters used to represent all static input parameters that might affect the desired output. The (m+1)-k-1 notation represents the architecture of the selected network. In this case, (m+1) represents the (m) static inputs, and the one additional feedback parameter, k is the optimal number of hidden nodes, which needs to be determined through the training and testing processes, and (1) is the desired number of outputs, namely, the futuristic desired value [i.e.,  $(A)^{n+1}$ ] [56]. An important component of dynamic-sequential modeling is that the datasets must be in sequential order and equal time steps. For the dynamic procedure is assumed that each data is recorded at the same intervals [28]. Figure 3 shows an example of dynamic network architecture with one output.

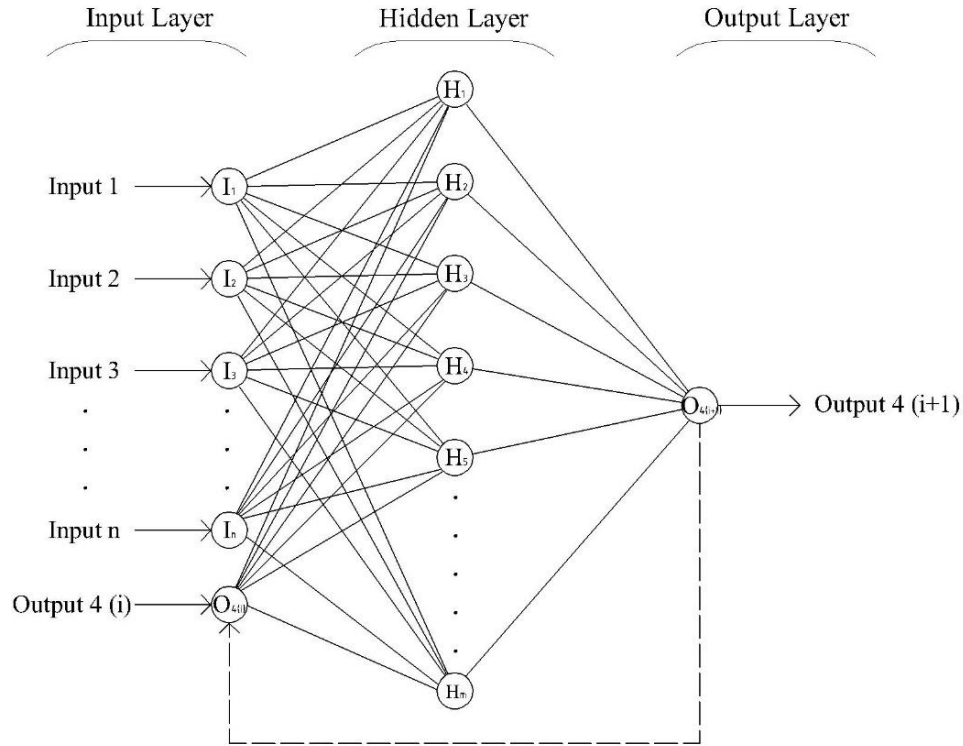


Figure 3. Dynamic-Sequential Network Structure [46]

## 2.2. Overview of Pavement Performance Indicators

Among the most important measures of pavement performance, roughness is an indicator of road conditions and is used for making objective decisions related to the management of road networks [8]. Pavement roughness describes the irregularities in the pavement surfaces that affect the ride quality experienced by daily road users [2]. Of several pavement condition indices used to assess pavement surface conditions, the PCR and IRI are the most used and well-recognized pavement performance indicators.

The PCR is a rating method based upon visual inspection of pavement distress. Although the relationship between pavement distress and performance is hard to be understood, there is evidence that the ability of pavement to sustain traffic loads safely and smoothly is adversely affected by the incidence of observable distress. The PCR method presents a procedure for

uniformly identifying and describing, in terms of severity and extent, pavement distress. The mathematical expression for PCR gives an index reflecting the composite effects of varying distress types, severity, and extent upon the overall condition of the pavement. The PCR calculation is based upon the summation of deducting points for each observable kind of distress. Deduct values are a function of distress type, severity, and extent [57]. The weights of distresses, severity, and extent are multiplied to find the deduction for each distress type. Appendix A presents the forms that aid field personnel in establishing distress severity and extent while performing the PCR surveys [57]. Equation 14 shows the PCR mathematical expression [28].

$$PCR = 100 - \sum_{1}^{n} Deduct_i \quad Eq. 14$$

Where:

- PCR = Pavement Condition Rating
- n = number of observable distresses
- $Deduct_i$  = multiplication of the weight of distress, the weight of severity, and weight of extent for distress I (Appendix A).

The Ohio Department of Transportation [57] developed a PCR scale to describe the pavement condition using the PCR numbers calculated from Equation 14. This scale has a range from 0 to 100; a perfect pavement with no observable distress has a PCR of 100 and pavement with all distress present at their “High” levels of severity and “Extensive” levels of extent have a PCR of 0. Figure 4 illustrates the PCR Scale and the explanatory condition of a pavement associated with the various ranges of the PCR values.

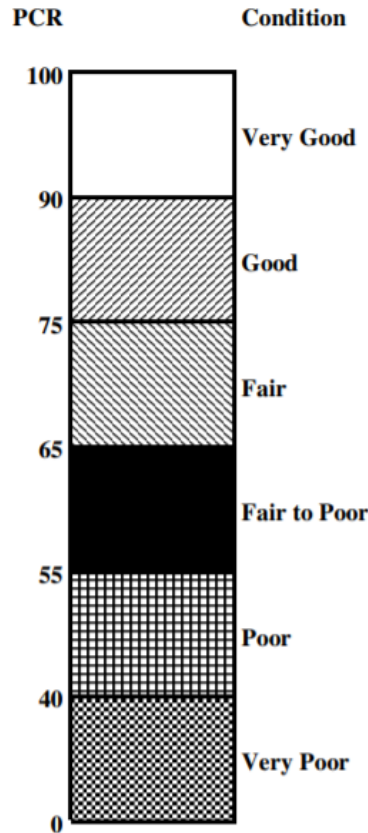


Figure 4. Pavement Condition Rating (PCR) Scale [57]

Several methods were created to measure pavement roughness which turned difficult the use of roughness data since they were obtained by different methods. For this reason, there was a need to establish a standard roughness index to eliminate possible problems caused by using different roughness indices, methods, and data collection [8]. In 1982, the World Bank and the government of Brazil proposed the International Road Roughness Experiment (IRRE) to find a standard roughness index appropriate for the many types of roughness to offer a basis for comparing roughness measures obtained by different procedures. Forty-nine road test sites were measured using different test equipment and measurement conditions. A full roughness range of asphaltic concrete, surface treatment, gravel, and earth roads was included in the study. The results from the IRRE showed that a standard roughness index was practical, and an index was proposed

that is measurable by most of the equipment, including road meters and profilometers. This selected measure has been denoted as IRI. The IRI is based on the quarter-car analysis method, a mathematical model of a vehicle that represents a body and a single wheel [7], with standardized parameter values and a reference simulation speed of 80 km/h [8].

The IRI measurement can be expressed in two types of units, in/mile or m/km. A higher IRI value indicates a rough pavement profile, which affects the ride quality experienced by road users. A lower IRI value indicates a smooth pavement profile, causing a better ride quality for the road users. Using high-speed vans equipped with laser equipment, accelerometers, and a computer, the pavement profile is measured generating the IRI values. The surface profiles are measured at traffic speed and the onboard accelerometer provides the data to calculate the changes in the vertical position. The distance between the vehicle and the surface of the road is measured by laser and the collected data is stored in the computer periodically. Since the change in longitudinal pavement profile over time is directly related to the change in roughness with time, it becomes an important indicator of pavement performance. The MEPDG [5] designed to update the 1993 AASHTO [6] uses the IRI measurements of longitudinal roughness to indicate pavement smoothness. The IRI measurements are stable, easy to be reproduced from longitudinal profile elevation, highly correlated with other roughness measuring devices, and offer good correlations with important user serviceability ratings, like present serviceability rating.

### 2.3. Long-Term Pavement Performance

The mission to study pavement performance and promote extended pavement life across the United States had been advanced since the late 1950s. However, just with the passage of the Surface Transportation and Uniform Relocation Assistance Act of 1987, Congress authorized the

LTPP program as part of the first Strategic Highway Research Program (SHRP) in 1987. A 5-year applied research program funded by the 50 States through a dedicated share of the Highway Trust Fund [58]. A total of 51 States including Washington D.C. and ten Canadian provinces' highway agencies that joined the program seeking to advance highway research and planning were considered in the initial selection of the test sections in the LTPP program [2].

The objectives of the LTPP program are to collect and store performance data from a large number of in-service highways over an extended period to support analysis and product development. Analyze the collected data to describe pavements' performance and translate these insights into usable engineering products related to pavement design, construction, rehabilitation, maintenance, preservation, and management [58]. The data collection started in 1989 and 2,509 pavement test sections were selected or constructed for the study. In 1992 with the end of the SHRP, the LTPP program continued under the U.S. Department of Transportation's Federal Highway Administration (FHWA) and continues until the present day. New experiments and studies are being added to monitor the performance of pavement materials and new technologies that were not yet in use when the LTPP program began [58].

The design of experiments was an integral part of the planning and preparation for the LTPP program. In the mid-1980s, two study types were considered, the GPS and the SPS. The GPS experiments used in-service pavement sections to examine general performance by pavement type. In contrast, the SPS experiments were designed to investigate the influence of specific features on pavement performance, these sections were to be constructed specifically for the LTPP study [58].

The LTPP data collection has different spatial and temporal locations throughout the U.S.

A climatic zone classification was created during the initial recruitment phases of the LTPP test sections and is divided into four different climates zones identified as wet-freeze, wet-non freeze, dry-freeze, and dry-non freeze zones as shown in Figure 5 [2,59]. This climatic zone map was altered in some places to adhere to State boundaries to ease data collection processes.



Figure 5. LTPP Climate Zone Map [59]

Over the years the LTPP program has accumulated a vast repository of data, documentation, and related tools, which compose LTPP's comprehensive Information Management System (IMS). The LTPP IMS is the premier product of the LTPP program and is used for research, pavement design, and product development [60]. The data and information in the LTPP IMS became available via the web through the data portal system, LTPPInfoPave™ [61], in January 2014. InfoPave is the public gateway to access data and other information about

the LTPP program.

#### 2.4. Mississippi Department of Transportation Performance Models

One of the main problems for state transportation agencies is to perform timely and proper maintenance and rehabilitation of roads and highways to meet the public's needs and safety concerns. To ensure the safe passage of not only passengers, but goods, and products, the safety of the roadways cannot be understated [29]. Current investigations of the mechanical properties of existing pavements and their M&R management are crucial for the continuation of an uninterrupted transportation system. These studies require well-coordinated field measurements and a complete and reliable decision-making process to overcome future issues in pavement systems [62].

The improvement of data collection equipment allowed a faster and more complete data acquisition, providing thousands of datasets with higher resolution. During a typical one-year survey, approximately 27,250 miles of survey data are collected from state, interstate, and non-interstate highways, as well as freeway expressways or other principal arterial routes [62]. At this time, the MDTO database has over 40 million records of data that includes condition, distress, friction, curve and grade, mean roughness index, global positioning system (GPS) location, 360-degree images, and roadway images[62].

In the late 1980s, the MDOT started a research collaboration with the University of Mississippi resulting in sets of models that are currently in use for decision-making. The existing system utilizes Markov probabilistic models to estimate pavement distresses. It estimates the probability of a pavement section moving from one state of distress to a state of more severe distress within one year given a specific pavement preservation action [62]. Among several



pavement condition indices used to assess pavement surface conditions, MDOT utilizes the PCR and IRI, which are the most widely used and well-recognized pavement performance indicators to make a timely decision and maintenance schedule [63]. However, existing models have developed over 30 years ago. Since that time, the design methods, resources, and construction procedures have been updated to the latest technology based on cutting-edge research in material science and pavement design. Therefore, previously developed models should be updated to be valid for the new design methods and material changes.

In transportation systems, decision support systems must work rapidly to ensure the correct maintenance without delay and reduce issues regarding traffic operations. The development of advanced pavement performance models using modeling techniques that are more intelligent, inclusive, reliable, and accurate, will lead MDOT to plan more efficient M&R actions saving time and money. To achieve this objective, new pavement performance models utilizing the ANN technique are developed in this dissertation for the MDOT database.

The pavement database utilized in this research is a part of Mississippi's pavement survey performed by the MDOT. Every two years, MDOT collects data to monitor the current pavement conditions and predict M&R for the Mississippi road network. Flexible, rigid, and composite pavements are part of the database and were utilized to develop performance prediction models. Due to the new methods in the data collection system, only recorded datasets from 2010 to 2020 were utilized, which resulted in 6 usable years of data. To characterize the behavior of pavement deterioration in a one-year time increment, a continuous database was needed to be used for developing reliable models. Since MDOT collects data every even year to develop prediction models that are applicable for a 1-year increment, the odd-year data were generated by averaging

consecutive years from 2010 to 2020. This approach was successfully utilized for the MDOT database in previous studies [28–30,63–65].

By assessing the quality of databases, sections with missing or illogical data have been excluded as the ANN model development process needs a complete dataset. This includes instances of negative IRI and the sections without the recorded length. The data collection, processing, and modeling are explained in detail in the upcoming chapters of this dissertation.

## CHAPTER III: LITERATURE REVIEW

### 3.1. Pavement Performance Models

Understanding the role that pavement performance plays in the maintenance and rehabilitation activities of the highway network system has pointed transportation agencies to the need for developing intelligent and efficient pavement performance models. These performance models can help to make decisions regarding pavement maintenance and rehabilitation priorities.

The design premise included in the AASHTO MEPDG [5] for predicting smoothness degradation is that the occurrence of surface distress will result in increased roughness, rising the IRI value, or in other words, reducing smoothness. The MEPDG developed an equation to predict the IRI over time for composite pavements using data collected from the LTPP program. The equation is embedded in the MEPDG, and it is presented in Equation 15.

$$IRI = IRI_0 + 0.00825 (SF) + 0.575 (FC_{Total}) + 0.0014 (TC) + 40.8 (RD) \quad \text{Eq. 15}$$

Where:

- $IRI_0$ : Initial IRI after construction, in/mi
- $SF$ : site factor, refer to Equation 15
- $FC_{Total}$  is the area of fatigue cracking (combined alligator, longitudinal, and reflection cracking in the wheel path), percent of total lane area. All load-related areas combined on an area basis-length of cracks are multiplied by 1 ft. to convert length into area basis.)
- $TC$ : length of transverse cracking (including the reflection of transverse cracks in existing HMA pavements), ft/mile
- $RD$ : average rut depth.

The site factor ( $SF$ ) is calculated using Equation 16.

$$SF = Age [0.02003 (PI + 1) + 0.007947 (Precip + 1) + 0.000636 (FI + 1)] \quad \text{Eq. 16}$$

Where:

- $Age$ : pavement age, year
- $PI$ : percent plasticity index of the soil
- $FI$ : average annual freezing index, °F
- $Precip$ : average annual precipitation or rainfall, in.

Sandra and Sarkar [22] developed a model based on the relationship between roughness and noticeable distresses commonly observed on Indian roads such as cracking, potholes, patching, rutting, and raveling. The data was collected over 39.5 km length of roads with different functional classes such as national highways, state highways, and major district roads in the Rajasthan state of India. All the distresses were considered in terms of extent and severity. The developed model showed that besides the usual distresses such as rutting, patching, and cracking considered by various studies, the contributions of potholes and raveling were quite predominant on Indian roads. Also, the level of severity of particular distress had a differential impact on the roughness, and thus, both severity and extent need to be considered for developing such models. The author also recommended that different models should be developed since the quality of construction, maintenance management, and traffic volume and composition vary substantially with the class of road.

Rahim et al. [66] evaluated the IRI for asphalt pavement overlaid over concrete slab treated with crack, seat, and overlay (CS&O) rehabilitation technique. Two LTPP regions were used for the development of non-linear regression models, wet-freeze, and wet-non-freeze. An additional model was developed for pavement sections in California. Asphalt overlay thickness and base type

(bound or unbound) were factors evaluated in the study. The independent variables were age, annual Equivalent Single Axle Load (ESAL), CESAL, base type, asphalt, and concrete pavement thicknesses. The study obtained an  $R^2$  of 0.55 for the wet-freeze model,  $R^2$  of 0.50 for the wet-non-freeze model, and an  $R^2$  of 0.62 for the California model.

Cheng and Zhang [24] investigated the applicability of IRI-based pavement deterioration prediction models, including four deterministic models for pavement performance prediction (i.e., the National Cooperative Highway Research Program (NCHRP) model, Al-Omari–Darter model, Dubai model, and the New Mexico Department of Transportation (NMDOT) model). Comparisons of these models were made using the data from both the NMDOT pavement management system and the LTPP sites in New Mexico. The Al-Omari-Darter model and the Dubai model were used to predict IRI values while the NCHRP and NMDOT models were used to predict performance based on IRI values. Al-Omari and Darter [67] found that pavement rut depth and the standard deviation of rut depth were the most significant factors affecting IRI and developed Equation 17 and Equation 18.

$$IRI = 57.56RD - 334.28, R^2 = 0.93 \quad \text{Eq. 17}$$

$$IRI = 136.19SD - 116.36, R^2 = 0.94 \quad \text{Eq. 18}$$

The Dubai model used Age as the predictor for the IRI predictions as shown in Equation 19.

$$IRI = 0.796 \exp(0.0539Age), R^2 = 0.801 \quad \text{Eq. 19}$$

The NCHRP model used an exponential regression equation as shown in Equation 20.

$$PSI = 5 \exp(-0.29IRI), R^2 = 0.703 \quad \text{Eq. 20}$$

The NMDOT model devised the PSI by combining IRI and eight other pavement distresses. After the analysis and comparison of the models, it was found that the Dubai model and the NCHRP model stand a better chance of being consistent and reasonably precise than the Al-Omari–Darter model and the NMDOT model.

Khattak et al. [68] developed IRI prediction models using regression analysis for composite (asphalt overlaid concrete) and flexible pavements in the state of Louisiana as a result of a three-phased study by the Louisiana Department of Transportation and Development. In this study, the performance of overlay treatment of about 751.5 km (467 miles) of composite pavements and 2027.7 km (1260 miles) of flexible pavements in the state of Louisiana was analyzed to develop the IRI models. For the composite pavement IRI regression model an  $R^2$  of 0.63 was found using nine input variables (IRI value before treatment (m/km), CESAL, thickness of HMA overlay, thickness of Portland Cement Concrete (PCC) layer, functional classification, cumulative temperature index, age of treatment, precipitation index, and a variable delta). For the flexible pavement IRI regression model an  $R^2$  of 0.47 was found using seven input variables (functional classification, cumulative equivalent single-axle load, thickness of overlay, temperature index, the age of treatment, cumulative precipitation index, and a variable delta created in the study). The study concludes that the developed IRI models presented good agreement between the measured and predicted IRI values with most data within 5% of prediction error and the models could be used as a good pavement management tool for pavement maintenance and rehabilitation actions.

However, pavement performance modeling is not a simple task due to the complex relations between pavement structure and its responses to climate and traffic variables, and the interaction between all these elements together [11]. The modeling of asphalt and concrete

pavement performance has been investigated in many studies over the years. However, composite pavements have not been well investigated. Also, several performance models have used distress data which are costly to collect and not easily available for all agencies. Traditionally, linear, non-linear, multiple linear regression analysis, Markov chains, mechanistic-empirical, survivor curves, semi-Markov, and Bayesian models have been used for predicting pavement roughness value [20]. However, due to the complexity of the relations between each one of the variables that affect pavement roughness, the use of simple statistical approaches such as linear regression does not seem appropriate to develop performance-prediction models [21].

Advanced modeling techniques using machine learning appear as an alternative for predicting pavement deterioration, offering significant improvements over traditional techniques. By feeding data in the form of observations and real-world interactions to computers and making them learn and act intelligently to find complex connections between variables, machine learning can process large volumes of data with a high degree of accuracy [20]. It can also handle noisy and complex data, solve non-linear problems, and once trained, it can make predictions and generalizations at any time [39]. Machine learning techniques hold significant potential for building a modern and robust pavement system due to the excellence in automation and pattern recognition [20]. The literature review shows that a remarkable number of researchers have used ANN to predict pavement performance.

### 3.1.1. Flexible Pavement Performance Models Using ANN

Attoh-Okine [69] used a backpropagation neural network algorithm to develop an IRI prediction model for asphalt pavements using data from the LTPP database and apply a

sensitivity analysis to find the relative significance of the material and construction variables on the roughness. Asphalt content, asphalt layer thickness, cumulative equivalent single axle load, structural number (SN), and the percentage of fines passing through the No. 200 sieve were used as independent variables. The IRI was used as the output variable. The study concluded that the ANN technique is feasible when a large database on pavement conditions is available. This technique could form the basis for developing a generic intelligent pavement deterioration process. However, it is also important to explore different preprocessing of input data, learning rules, and transfer functions to perform more successful predictions.

Kargah-Ostadi et al. [70] developed an ANN model for flexible pavements using a specific pavement study (SPS-5) from the LTPP database. The objective of the study was to use the model to predict and compare pavement roughness variation trends after various rehabilitation alternatives. The optimum ANN structure had eight input variables, five hidden nodes within one hidden layer, and one output. Model testing resulted in the prediction of IRI with minimal errors and future roughness prediction trends that match perfectly with the observed values. These findings indicate that the ANN model performs successfully in predicting IRI trends following each kind of treatment in the SPS-5 experiment.

Hossain et al. [16] developed an ANN prediction model for flexible pavements using climate and traffic data collected from the LTPP database. The study compared the ANN-predicted IRI and measured IRI for flexible pavements under specific climatic zones (wet freeze) with a two hidden-layered ANN structure with seven independent variables, nine hidden nodes for the first and second hidden layers, and one output (7-9-9-1), using a nonlinear transfer function. A Root Mean Squared Error (RMSE) of 0.027 was found for the flexible ANN model, indicating that the



IRI prediction was reasonable for both short-term and long-term predictions using only climate and traffic data.

Solatifar and Lavasani [27] developed an ANN flexible pavement deterioration model based on IRI utilizing Back-Propagation Neural Networks (BPNN) technique with the LTPP database for two GPS sections (GPS-1 and GPS-2). After training and testing the final developed model, the results were compared with a polynomial model developed with a nonlinear regression method. Several statistical error calculations were used to compare the results of both models. The ANN model showed an RMSE of 0.2750 and 0.2120 for the GPS-1 and GPS-2 respectively while the polynomial model showed 0.2751 and 0.2120. Similar results were obtained but the ANN model showed more accurate results. By using the ANN model, a more precise decision in choosing an M&R policy can reduce the costs of pavement management.

Jaafar [2] developed mechanistic-empirical models using ANN and multiple linear regression (MLR) techniques for predicting IRI, rutting, and cracking for asphalt pavements using the LTPP database. For the IRI modeling, the ANN architecture used seven independent variables, five hidden nodes within a single hidden layer, and one output (i.e., 7-5-1 ANN structure). The independent variables used for IRI modeling were initial IRI, pavement age, structural number (SN), CESAL, air temperature, precipitation, and construction number (CN) (an indicator of major maintenance and/or rehabilitation). The ANN model showed a coefficient of correlation (R) of 0.72, which is considered reasonably accurate for IRI prediction in asphalt pavements. Sollazzo et al. [18] also developed an ANN model and compared it with linear regression, obtaining better accuracy when using the ANN model compared to the MLR model.

Choi [14] developed an ANN prediction model for flexible pavements on a granular base

from three states: Texas, New Mexico, and Arizona. The results show that the ANN model could deliver a reasonable explanation for their predictive behavior and model the relationship between input variables and pavement performance.

Duckworth [28] and Duckworth et al. [29] developed pavement performance prediction models using the ANNs approach for flexible pavements based on the MDOT database. A two-output model for predicting PCR and IRI was found to be the most promising. The ANN model successfully characterized the deterioration behavior with statistical measures in a suitable range.

Yamany [13] developed pavement performance models for flexible pavements using data from eight Midwestern states, and Zeiada [12], developed prediction models for warm climate regions in the LTPP database. Both studies found that by specifying these characteristics their prediction models performed better since the data gather the same characteristics and helped the model to understand the variability of the datasets.

Barros et al. [71] developed ANN performance models for flexible pavements considering traffic and climate loads, pavement age, initial roughness condition, and M&R interventions using the LTPP database. The developed models efficiently characterized the deterioration behavior of asphalt pavements over time, and effectively capture the effect of M&R interventions. The predicted IRI values were in good agreement with observed values and the developed models ( $R^2=0.61$  and  $R^2=0.67$  for Model 1 and Model 2, respectively). Barros highlights that even though the development of the ANN model requires a good understanding of the roughness phenomena, the developed models are simple, fast, and do not require the user to have any prior knowledge of IRI or ANN.

### 3.1.2. Rigid Pavement Performance Prediction Models Using ANN

Relatively fewer studies have been conducted in recent years to predict rigid pavement performance when compared to flexible pavements.

Hossain et al. [17] developed a prediction model for IRI for rigid pavement using climate and traffic data by employing Artificial Neural Network (ANN) modeling. The climate and traffic data are collected from the LTPP database. The ANN model is trained, tested, and validated using 70%, 15%, and 15% of data respectively. The trained model and the validated model are compared by calculating the RMSE and Mean Absolute Percentage Error (MAPE) of ANN predicted IRI and measured IRI. The study developed a model for rigid pavement located in the wet non-freeze climate zone, employing a 7-9-9-1 ANN structure and using a hyperbolic tangent sigmoidal transfer function, the RMSE, and MAPE values generated are 0.01 and 0.01 (1% error) respectively.

Yasarer et al. [30] developed a new set of ANN models that contain daily traffic volume, IRI, soil condition, pavement thickness, and mean roughness index ( $IRI_{Mean}$ ) for the Jointed Concrete Pavements (JCP) in Mississippi. The best performing ANN model had an  $R^2$  of 0.93 and was integrated into a Microsoft Excel spreadsheet to generate an application that is simple, user-friendly, and allows the user to visualize the future projections of the pavement section. The authors recommended that MDOT personnel can employ this application to predict the condition of the JCP and prioritize the maintenance and rehabilitation schedule.

Yasarer et al. [65] developed a performance model for CRCP pavement using the ANN modeling technique for Mississippi. This study used maintenance and rehabilitation actions as an input in the model. The database used in this study contained 69 CRCP pavement sections that

resulted in 212 datasets from 2010 to 2018. The ANN model was trained using 25% data, then tested with 25% data, and the other 25% of data was employed to validate the model by comparing ANN predicted IRI and measured IRI. The study developed a model employing an 11-18-1 ANN structure with the accuracy of 0.0012 for ASE, 5.923 for MARE, and 0.872 for  $R^2$  statistical measures.

Sultana [9] developed performance models for Jointed Plain Concrete Pavement (JPCP), Jointed Reinforced Concrete Pavement (JRCP), and Continuously Reinforced Concrete Pavement (CRCP) using MLR and ANN techniques considering the effects of M&R history in the model development. The input and output variables were similar for all the models and retrieved from the LTPP database. The ANN models showed better accuracy in predicting pavement performance compared to the multiple regression models for all types of concrete pavements. A high  $R^2$  of 0.94, 0.95, and 0.95 were obtained for the JPCP, JRCP, and CRCP, respectively, presenting a significant improvement over models that currently use mechanistic-empirical pavement design.

Sultana et al. [33] developed an ANN pavement deterioration model for jointed plain concrete pavement (JPCP). The models were developed using LTPP data for the wet, freeze climatic region. The input variables were initial pavement condition (i. e., initial IRI), pavement structural and mechanical properties (i.e., age, concrete pavement thickness, base/subbase thickness, average contraction spacing, base/subbase materials type), traffic (CESAL), and climate attributes (i.e., average annual air temperature, total annual precipitation, annual freezing index, annual freeze-thaw), and IRI as the output variable. The developed ANN model had an  $R^2$  of 0.92, an ASE, and a MARE value of 0.00103 and 9.93, respectively. The total data points used to develop the ANN model were 636 and the final model structure was 13-19-1, where 13 is the

number of input variables, 19 hidden nodes, and 1 output variable. The best model was used to simulate extreme climate conditions by developing a GUI. IRI values gradually increased, and pavement conditions deteriorated over time when climate conditions change to the extreme. The study addressed a few gaps in the literature including the scarcity of studies on long-term IRI prediction using LTPP data and studies on the effect of climate attributes on pavement deterioration.

Sultana et al. [32] exhibited a methodology to determine pavement performance incorporating maintenance and rehabilitation history using the LTPP database and ANN modeling approach. The study incorporated the M&R history as CN in the LTPP database and the hypothesis testing demonstrated that M&R treatment has a significant effect on pavement performance. Several ANN models were attempted to evaluate the best way to include M&R history and resulted in a more realistic prediction of pavement conditions. A continuous CN approach resulted in an  $R^2$  of 0.901 compared to the categorical CN approach of  $R^2$  of 0.878.

Sultana et al. [35,36] utilized the CN variable for developing IRI prediction models for JPCP. Three ANN models were developed using variables such as initial IRI, pavement age, concrete pavement thickness, ESAL, climatic region, and CN. The best model had an  $R^2$  of 0.87 and successfully estimated the increase of IRI values with time and decrease of IRI value after maintenance and rehabilitation.

Sultana et al. [34] studied climate attributes such as precipitation, extreme temperature, and freeze-thaw cycles along with traffic loads that cause pavement distresses. Sultana developed IRI models that successfully estimated the IRI values for JPCP considering the M&R history of the pavements using the ANN approach. The variables used for the ANN model development are  $IRI_0$ ,

pavement age, concrete pavement thickness, ESAL, climate zone (wet-freeze, wet non-freeze, dry-freeze, dry non-freeze), CN, and several climatological data. The best-performing ANN model resulted in promising statistical measures (i.e.  $R^2= 0.87$ ).

Abd El-Hakim and El-Badawy [72] developed an ANN model to predict IRI values for JPCP sections using the LTPP database. The model inputs were  $IRI_0$  value, pavement age, transverse cracking, percent joints spalled, flexible and rigid patching areas, total joint faulting, freezing index, and percent subgrade passing No. 200 U.S. sieve. The data included a total of 184 IRI measurements and the results show that the ANN model yielded a higher prediction accuracy ( $R^2$  of 0.83, and ratio of standard error of estimate (predicted) to standard deviation of the measured IRI values:  $Se/Sy =0.414$ ) compared to the MEPDG model ( $R^2$  of 0.584,  $Se/Sy =0.643$ ). In addition, the bias in the predicted IRI values using the ANN model was significantly lower compared to the MEPDG regression model.

### 3.1.3. Composite Pavement Performance Prediction Models Using ANN

Literature review to date shows that ANN models performed successfully in predicting IRI values for asphalt and concrete pavements. However, performance prediction models for composite pavements have not been well investigated. A few studies are available using composite pavements data and fewer studies utilized M&R history in the model development.

Kaya et al. [10] developed pavement performance models for flexible and composite (asphalt concrete over the jointed plain concrete pavement) pavement systems in Iowa. ANN-based models were found to be good tools for modeling pavement deterioration when there were many pavement sections with various traffic, thickness, and other various deterioration trends.

Abdelaziz et al. [15] develop an IRI prediction model for both original and composite pavements using general pavement studies (GPS-1, GPS-2, and GPS-6) and the specific pavement studies (SPS-1, SPS-3, and SPS-5) of the LTPP database. MLR and ANN techniques predict IRI as a function of pavement age,  $IRI_0$ , transverse cracks, alligator cracks, and standard deviation of the rut depth. The ANN model resulted in better results compared to the regression model,  $R^2 = 0.75$  and  $R^2 = 0.57$ , respectively. The network consisted of five inputs, three hidden layers with ten nodes each, and one output, with a Logarithmic-Sigmoidal (LOGSIG) as the transfer function.

Barros et al. [31] developed pavement roughness models for composite pavements using the LTPP database and the feed-forward ANN approach. A total of 592 data points from 52 pavement sections were analyzed. Five models were developed and the best performing model had an ASE of 0.002, a MARE of 12.936, and an  $R^2$  of 0.88. It utilized 14 input variables (i.e., Initial  $IRI_{Mean}$ , Age, Wet-Freeze, Wet Non-Freeze, Dry-Freeze, Dry Non-Freeze, Asphalt Thickness, Concrete Thickness, CN Code, ESAL, Annual Air Temperature, Freeze Index, Freeze-Thaw, and Precipitation) and one output variable ( $IRI_{Mean}$ ).

This doctoral research presents several papers that explore the development of pavement performance models for flexible, rigid, and composite pavements embedding M&R history and interventions to provide more realistic performance predictions. Since one of the main concerns of state and federal agencies is prioritizing the decisions for M&R actions, the developed models can be used to properly assess the condition of the pavements and predict future scenarios to perform more effective and timely M&R interventions. Furthermore, the models developed in this dissertation do not use distress variables as an input, which makes the use of the prediction models easier for agencies.

## CHAPTER IV: ROUGHNESS MODELING FOR COMPOSITE PAVEMENTS USING MACHINE LEARNING

Rulian Barros, Hakan Yasarer, Waheed Uddin, Salma Sultana

Department of Civil Engineering, University of Mississippi, University, MS 38677, United States  
rferrei1@go.olemiss.edu

(Paper published at the IOP Conference Series: Materials Science and Engineering, 1203(3), 032034. <https://doi.org/10.1088/1757-899x/1203/3/032034>)

### 4.1. Abstract

A large number of paved highway surfaces comprises composite pavements as a result of concrete pavement rehabilitation that uses an asphalt overlay on top of the concrete surface. Annually, billions of dollars are spent on the maintenance and rehabilitation of road networks. Roughness is one of the several indicators of road conditions used to make objective decisions related to road network management. The irregularities in the pavement surface affecting the ride quality of road users can be described by a standard roughness index defined as the International Roughness Index (IRI). Roughness prediction models can identify rehabilitation needs, analyze rehabilitation effects, and estimate future pavement conditions to implement different Maintenance and Rehabilitation (M&R) activities to extend the pavement life cycle and provide a smooth



surface for road users. This study intended to develop pavement performance models to predict roughness for asphalt overlay on concrete pavement sections using the Long-Term Performance Pavement (LTPP) program database. The artificial Neural Networks (ANNs) approach was used to develop roughness prediction models. A total of 52 pavement sections with 592 data points were analyzed. Five models were developed, and the best performing model, Model 5 was found with an average square error (ASE) of 0.0023, mean absolute relative error (MARE) of 12.936, and coefficient of determination ( $R^2$ ) of 0.88. Model 5 utilized one output variable (IRIMean) and 14 input variables (i.e., Initial IRIMean, Age, Wet-Freeze, Wet Non-Freeze, Dry-Freeze, Dry Non-Freeze, Asphalt Thickness, Concrete Thickness, CN Code, ESAL, Annual Air Temperature, Freeze Index, Freeze-Thaw, and Precipitation). The ANN model structure utilized for Model 5 was 14-9-1 (14 inputs, 9 hidden nodes, and 1 output). Environmental impacts and traffic repetitions can cause severe damage to the pavement if timely maintenance and rehabilitation are not performed. By considering the effects of the M&R history of the pavement, it is possible to obtain realistic prediction models for future planning. Therefore, the developed ANN roughness performance models in this paper can be used as a prediction tool for IRI values and guide decision-makers to develop a better M&R plan. Local and state agencies can use available historical traffic and climatological data in the developed models to estimate the change in IRI values. Utilizing these prediction models eliminates time-consuming data collection and post-processing, and consequently, a cost reduction. This low-cost tool will improve the condition assessment and effective M&R scheduling.

#### 4.2. Introduction

An efficient and safe transportation network for public mobility and freight transportation

is an important part of a nation's economy and prosperity [2]. From the 2.6 million miles of paved roads and highways in the United States, 93 percent of them are surfaced with asphalt [2]. However, a large portion of the paved highway surfaces comprises composite pavements, which are made of an asphalt overlay on concrete pavements. Most of the composite pavements are a result of concrete pavement rehabilitation [3]. When concrete pavements start to fail, they are overlaid with a hot mix of asphalt (HMA) [3]. The use of composite pavements compared to flexible or rigid pavements, can provide better levels of performance both structurally and functionally and accordingly can be a more cost-effective alternative [3]. Annually, billions of dollars are required for the maintenance and rehabilitation of road networks. If timely maintenance and rehabilitation are not performed, the pavement damages inflicted by environmental impacts and traffic repetitions may lead the pavement to poor conditions that can cause life-threatening for road users [2].

Pavement performance modeling is an important part of pavement management systems (PMS), which allows decision-makers a better budget allocation plan for future pavement maintenance and rehabilitation (M&R) actions [37]. However, current pavement performance prediction models do not account for the influence of M&R activities during the service life of the pavement, which can affect the accuracy of the predictions [37]. Pavement roughness models are necessary to identify rehabilitation needs, analyze rehabilitation effects, and estimate future pavement conditions to implement different M&R activities to extend the pavement life cycle and provide good surface quality for road users [27,73]. The International Roughness Index (IRI) is accepted as an important indicator of pavement performance and used as the standard for pavement roughness [74]. The objective of this paper is to develop a pavement roughness model using ANNs

approach for asphalt overlay on concrete pavement sections in the LTPP database. An IRI prediction method was proposed based on the analysis of the influence of pavement structure, climate, and traffic data.

#### 4.2.1. Objectives

The main objectives of this paper are to:

- (1) Analyze roughness data for asphalt overlay on concrete pavements in the U.S. territories using the Long-Term Performance Pavement (LTPP) database.
- (2) Develop a roughness model for asphalt overlay on concrete pavements using the Artificial Neural Networks (ANNs) approach on the LTPP database.
- (3) Perform a sensitivity analysis on one section of the database.

#### 4.2.2. Scope

The scope is limited to asphalt overlay on concrete pavements

### 4.3. Literature Review

#### 4.3.1. Literature Review of LTPP Program

The mission to study pavement performance and promote extended pavement life across the United States had been advanced since the late 1950s. Congress authorized the LTPP program as part of the first Strategic Highway Research Program (SHRP) in 1987 [58]. A 5-year applied research program funded by the 50 States through a dedicated share of the Highway Trust Fund [58]. The objectives of the LTPP program were to collect and store performance data from a large number of in-service highways over an extended period to support analysis and product development. Also, analyze the collected data to describe pavements' performance and translate these insights into usable engineering products related to pavement design, construction, rehabilitation, maintenance, preservation, and management [58]. The data collection started in

1989 and 2,509 pavement test sections were selected or constructed for the study.

#### 4.3.2. Literature Review of International Roughness Index

Roughness is an indicator of road conditions and is useful for making objective decisions related to the management of road networks [75]. Pavement roughness describes the irregularities in the pavement surfaces that affect the ride quality experienced by daily road users [2]. In 1982, the World Bank and the government of Brazil proposed the International Road Roughness Experiment (IRRE) to find a standard roughness index appropriate for the many types of roughness to provide a basis for comparing roughness measures obtained by different procedures. The IRRE results showed that a standard roughness index was practical, and an index was proposed, the IRI. The IRI is based on the quarter-car analysis method, a mathematical model of a vehicle that represents a body and a single wheel [7], with standardized parameter values and a reference simulation speed of 80 km/h [75]. The IRI measurement can be expressed in two types of units, in/mile or m/km. A higher IRI value indicates a rough pavement profile, which results in a lower ride quality experienced by road users. A lower IRI value indicates a smooth pavement profile, causing a better ride quality for the road users.

#### 4.3.3. Literature Review of Roughness Models

Recently, several studies showed interest in developing pavement roughness prediction models for both flexible and concrete pavements.

Kargah-Ostadi [70] developed an ANN model for IRI prediction of flexible pavements using a specific pavement study (SPS-5) from the LTPP database. The objective of the study was to use the model to predict and compare pavement roughness variation trends after various rehabilitation alternatives. The optimum ANN structure had eight input variables, five hidden

nodes within one hidden layer, and one output. Model testing resulted in the prediction of IRI with minimal errors and future roughness prediction trends that match perfectly with the observed values. These findings indicate that the ANN model performs successfully in predicting IRI trends following each kind of treatment in the SPS-5 experiment.

Hossain et al. [16,17] developed an ANN prediction model for IRI for both flexible and concrete pavements using climate and traffic data collected from the LTPP database. Seven independent variables were considered as input parameters for predicting IRI. Both models compared the ANN predicted IRI and measured IRI for flexible and rigid pavements under specific climatic zones (wet-freeze for flexible pavement and wet non-freeze for rigid pavement). Both ANN models used a two hidden-layered ANN structure with seven independent variables, nine hidden nodes for the first hidden layer, nine hidden nodes for the second hidden layer, and one output (7-9-9-1), using a nonlinear transfer function. Both studies indicated that the IRI prediction was reasonable for both short-term and long-term predictions using only climate and traffic data.

Mohamed Jaafar [2] developed mechanistic-empirical models using ANN and multiple linear regression techniques for predicting IRI, rutting, and cracking for asphalt pavements using the LTPP database. For the IRI modeling, the ANN architecture used seven independent variables, five hidden nodes, one hidden layer, and one output (i.e. 7-5-1 ANN structure). The ANN model showed a high coefficient of correlation (R) of 0.72. A multiple linear regression model was also developed. An R-value of 0.63 was found using multiple linear regression. The results show that both ANN and multiple regression models were reasonably accurate for IRI prediction in asphalt pavements.

Khattak et al. [68] developed IRI prediction models using regression analysis for overlay

treatment of composite and flexible pavements in the state of Louisiana. For the composite pavement, an  $R^2$  of 0.63 was found using nine input variables. For the flexible pavement, an  $R^2$  of 0.47 was found using seven input variables. The study concludes that the developed IRI models provided good agreement between the measured and predicted IRI values with most of the predictions within 5%.

Literature review to date indicates that most roughness prediction models did not consider M&R history as an independent variable. This study proposes the use of CN as a categorical variable in the IRI prediction model for composite pavements. This approach was recently used in an asphalt highway pavement performance study at the University of Mississippi [2]. This paper developed a pavement roughness prediction model using the ANNs approach for asphalt overlay on concrete pavement sections in the LTPP database.

#### 4.4. Model Development

##### 4.4.1. Data Collection

Using the LTPPInfoPave<sup>TM</sup> database [60], a total of 311 sections were identified with asphalt and concrete in the same section. The asphalt thickness varies from 0.1 to 13.3 inches. The concrete thickness varies from 6.4 to 20.5 inches. Sections that have an asphalt layer thickness equal to or greater than three inches were considered as composite pavement sections. Following this criterion, 272 sections were identified as composite pavement sections with a total of 16,842 IRI measurements from 1989 to 2018. Each section has two types of IRI measurements,  $IRI_{Left}$  and  $IRI_{Right}$ . A mean roughness index ( $IRI_{Mean}$ ) was calculated by averaging the  $IRI_{Left}$  and  $IRI_{Right}$  measurements. On each visit date, several IRI measurement runs were done for each section. By averaging the IRI measurement runs, a single IRI measurement was obtained for  $IRI_{Left}$ ,  $IRI_{Right}$ ,

and  $IRI_{Mean}$  for each visit date. A total of 3,304 IRI measurements for 272 sections were obtained. For this study, a total of 592 datasets from 52 different sections were used to develop the ANN IRI prediction model. Figure 6 shows the 592  $IRI_{Mean}$  measurements for the 52 sections.

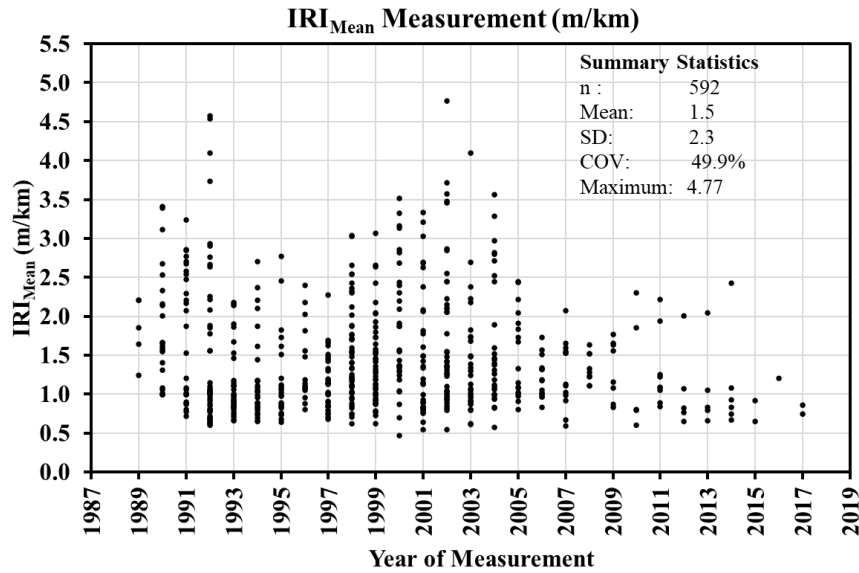


Figure 6.  $IRI_{Mean}$  Measurement (m/km)

#### 4.4.2. Consideration of M&R Treatment in IRI Roughness Prediction Model

The CN is the attribute that LTPP uses to monitor and identify M&R in each section of the database. A CN1 is assigned when the pavement section was opened to the traffic. When an M&R is conducted, the CN number will change from CN1 to CN2. Thus, the CN factor indicates that a major M&R treatment was conducted on the pavement section. The treatment intervention generally improves the pavement condition and performance for roughness, cracking, faulting, joint deterioration, and other surface defects. For this reason, it is imperative to consider CN as a factor for a more realistic and accurate model. For the ANN model development, CN will be used as a categorical variable with a value of zero or one. A zero value is assigned if no M&R was implemented in that section and a value of one is assigned if there was an M&R intervention. The

use of M&R actions in the model development was expected to result in more realistic models considering that M&R actions affect the future condition of the pavement. As an illustration, Figure 7 shows different CN values for section 06-7455 located in California. This section has three construction numbers (CN1, CN2, and CN3), which were assigned in 1989, 2001, and 2010.

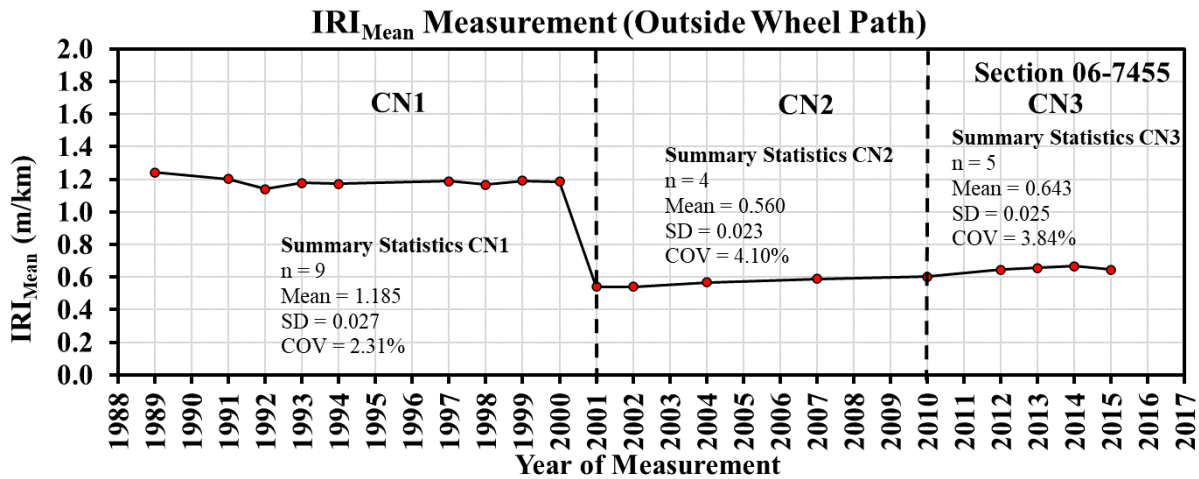


Figure 7. IRI<sub>Mean</sub> Measurement, Section 06-7455, California, 1989-2015

It is evident from Figure 7 that the M&R treatments improved the composite pavement condition, which contributed to lower IRI values. The IRI values decreased 54% from CN1 (1.185 in 2000) to CN2 (0.541 in 2001) when maintenance and rehabilitation (M&R) were performed in the section. To support this statement an independent sample t-test was performed to determine whether there are statistically significant differences between the means of IRI measurements between CN1 and CN2. The results show that the difference in the means of CN1 IRI<sub>Mean</sub> and CN2 IRI<sub>Mean</sub> are statistically significant at  $\alpha$  0.05 probability of chance error. This implies that both IRI<sub>Mean</sub> samples (CN1 and CN2) are from different populations. Thus, M&R treatments significantly improved the pavement surface condition and contributed to lower IRI values.



## 4.5. ANN Model

### 4.5.1. Overview of ANN

Artificial Neural Networks is a predictive modeling technique based on mathematical operations that use the concept of human cognition and neural biology [43]. The ANNs approach attempts to emulate the structure and/or functional aspects of biological neural networks [44]. It consists of several simple processing elements called neurons (or nodes) and connection links between them [43]. When the information is processed, the connection links are used to transfer signals between neurons [43]. Complex relationships that are difficult to be identified using traditional sequential, logic-based modeling and computational technics can be successfully represented by neural networks [43]. There are many types of neural networks characterized by their architecture, training algorithm, and activation function [45]. In this study, a feed-forward neural network with a back-propagation training algorithm was used for the development of the roughness prediction model. Different variable types that contain both categorical and continuous variables were used, and one hidden layer was considered in the model development. The use of more than one hidden layer combined with an insufficient number of databases may cause the network to memorize the data in the training phase [46]. Therefore, the developed models used only one hidden layer to maintain the generalization capability of the network [46]. The TR-SEQ1 computer program [55] was used to develop the ANN models in this study. A sigmoidal function is used for data generalization purposes.

### 4.5.2. ANN Model Variables and Architecture

IRI was used as a dependent variable (i.e. output) and several independent variables (i.e. inputs) were used in this study. Some variables were included based on previous literature studies

and new variables were introduced in this research. For this paper, five models were tried using different independent and dependent variables. Table 1 shows the variables used for each ANN model in this study.

Table 1. Independent and Dependent Variable Configuration for Five ANN Models.

Models	Model 1	Model 2	Model 3	Model 4	Model 5
Independent Variables	I <sub>0</sub> Left	I <sub>0</sub> IRI <sub>Mean</sub>	I <sub>0</sub> Left	I <sub>0</sub> IRI <sub>Mean</sub>	I <sub>0</sub> IRI <sub>Mean</sub>
	I <sub>0</sub> Right	Age	I <sub>0</sub> Right	Age	Age
	Age	Wet-Freeze	Age	Wet-Freeze	Wet-Freeze
	Wet-Freeze	Wet, Non-Freeze	Wet-Freeze	Wet, Non-Freeze	Wet, Non-Freeze
	Wet, Non-Freeze	Dry-Freeze	Wet, Non-Freeze	Dry-Freeze	Dry-Freeze
	Dry-Freeze	Dry, Non-Freeze	Dry-Freeze	Dry, Non-Freeze	Dry, Non-Freeze
	Dry, Non-Freeze	h <sub>asphalt</sub>	Dry, Non-Freeze	h <sub>asphalt</sub>	h <sub>asphalt</sub>
	h <sub>asphalt</sub>	h <sub>concrete</sub>	h <sub>asphalt</sub>	h <sub>concrete</sub>	h <sub>concrete</sub>
	h <sub>concrete</sub>	CN Code	h <sub>concrete</sub>	CN No Action	CN Code
	CN Code	ESAL	CN No Action	CN Any Action	ESAL
	ESAL		CN Any Action	ESAL	Temperature
		ESAL		Freeze Index	
				Freeze-Thaw	
				Precipitation	
Dependent Variables	IRI <sub>Left</sub>	IRI <sub>Mean</sub>	IRI <sub>Left</sub>	IRI <sub>Mean</sub>	IRI <sub>Mean</sub>
	IRI <sub>Right</sub>		IRI <sub>Right</sub>		

Model 1 used eight input variables; however, the climatic region is used as a categorical variable with four categories. The CN code has a value of 0 for no CN changes and 1 for any changes. Therefore, the first model had 11 input variables and 2 output variables (IRI<sub>Left</sub> and IRI<sub>Right</sub>). Model 2 used I<sub>0</sub> IRI<sub>Mean</sub> instead of I<sub>0</sub> Left and I<sub>0</sub> Right for the input variable. A total of 10 input variables and 1 output variable (IRI<sub>Mean</sub>) were used. Model 3 used a total of 12 input variables and 2 output variables. The CN was considered as two categorical inputs, CN No action (1 or 0) and CN Any Action (1 or 0). Model 4 used 11 input variables and 1 output variable. Model 5 included climatological factors using a total of 14 input variables and 1 output variable (IRI<sub>Mean</sub>).

All variables used in the model developing for this paper are not related to distresses data, which needs a lot of work, equipment, time, and money to be measured. The ANN models developed in this study used easily available variables that normally most agencies have the records of.

#### 4.5.3. ANN Model Selection

The best model was selected based on the lowest average square error (ASE), lowest mean absolute relative error (MARE), and highest coefficient of determination ( $R^2$ ). Table 2 shows statistical measures of the ANN model development stages (i.e., training, testing, validation, and all-data) for the five developed models. The final structure of each model is written at the bottom row in an order that depicts the number of inputs, hidden nodes, and output(s), respectively.

Table 2. ANN Model Results

Model		MODEL 1	MODEL 2	MODEL 3	MODEL 4	<b>MODEL 5</b>
Network		1-9-20000	1-9-20000	5-9-20000	1-7-20000	<b>1-9-11100</b>
Training	MARE	13.540	16.539	14.058	15.481	<b>15.076</b>
	$R^2$	0.83629	0.83525	0.85762	0.84083	<b>0.84757</b>
	ASE	0.002775	0.003288	0.002362	0.003064	<b>0.002899</b>
Testing	MARE	17.772	18.878	17.629	24.457	<b>22.209</b>
	$R^2$	0.70008	0.62038	0.68056	0.54602	<b>0.66829</b>
	ASE	0.004578	0.006705	0.005393	0.009307	<b>0.0072</b>
Validation	MARE	20.199	21.822	18.133	21.763	<b>23.282</b>
	$R^2$	0.5498	0.62044	0.52444	0.50833	<b>0.47758</b>
	ASE	0.007772	0.007641	0.008425	0.008617	<b>0.009887</b>
All Data	MARE	16.069	16.137	15.096	18.212	<b>12.936</b>
	$R^2$	0.77346	0.78018	0.78704	0.7188	<b>0.87741</b>
	ASE	0.003659	0.004087	0.003466	0.005285	<b>0.002272</b>
Final Structure		11-9-2	10-9-1	12-9-2	11-7-1	<b>14-9-1</b>

All the 592 datasets were used to retrain the network at its optimal structure and iteration to obtain the generalized response throughout the complete database. All-data stage for Model 5 outperformed all other models with significant improvements on model accuracy measures. For this reason, Model 5 was chosen as the best performing ANN model. Figure 8 shows the network

structure of Model 5.

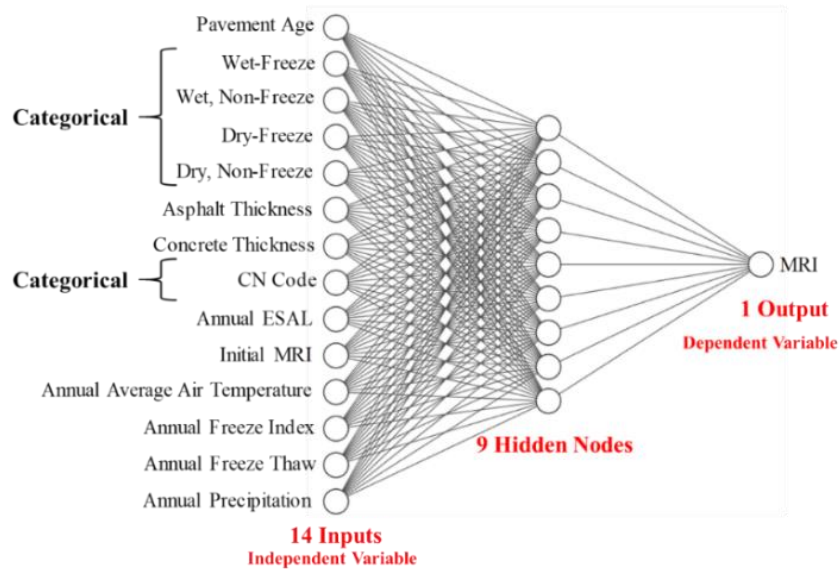


Figure 8. Network Architecture of the Best Performing ANN Model (Model 5)

#### 4.6. Results and Discussions

Figure 9 shows the accuracy measures of all the developed ANN models for the All-Data stage. The accuracy measures show reliable results for all models developed. However, Model 5 results outperform all other models developed. Model 5 has an ASE value (0.0023) 38% lower than the second-lowest ASE value (0.0035, Model 3); a MARE value (12.94) 14% lower than the second-lowest MARE value (15.10, Model 3); a  $R^2$  value (0.88) 11% higher than the second-highest  $R^2$  value (0.79, Model 3).

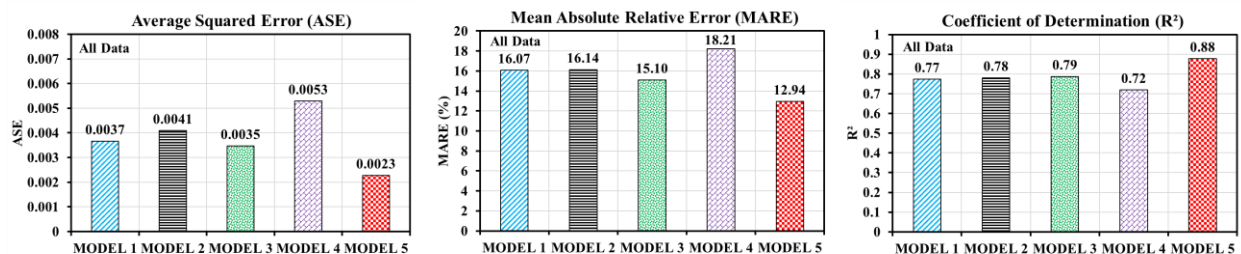


Figure 9. Graphical Comparison of Accuracy Measures between Developed ANN Models

Observed  $IRI_{Mean}$  values collected from the LTPP database and the predicted  $IRI_{Mean}$  predictions using Model 5 are presented in Figure 10. The plot shows the  $IRI_{Mean}$  (m/km) values in the y-axis and a section sequence number (generated to identify the data points in the database) on the x-axis.

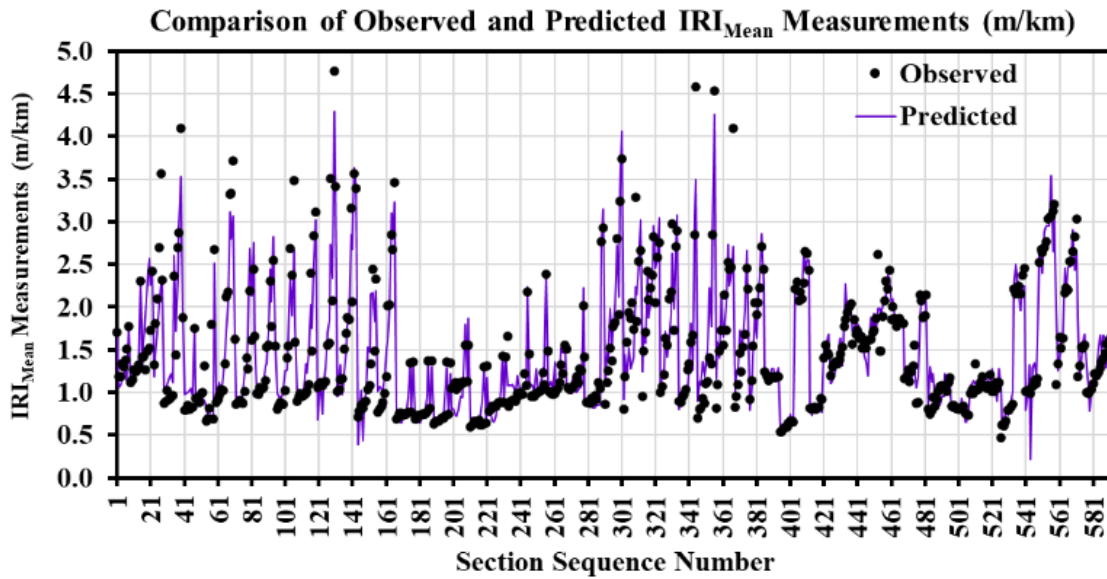


Figure 10. Observed and Predicted  $IRI_{Mean}$

Lower  $IRI_{Mean}$  values were better predicted than higher values. Figure 11 shows that Model 5 predictions clustered around the line of equality, but the predicted values are closer to the observed values until the  $IRI_{Mean}$  value is equal to 3 m/km. When the observed  $IRI_{Mean}$  value is greater than 3 m/km, the model was not as accurate as it was for lower  $IRI_{Mean}$  values. Nevertheless, a high  $R^2$  of 0.88 was obtained for the ANN Model 5.

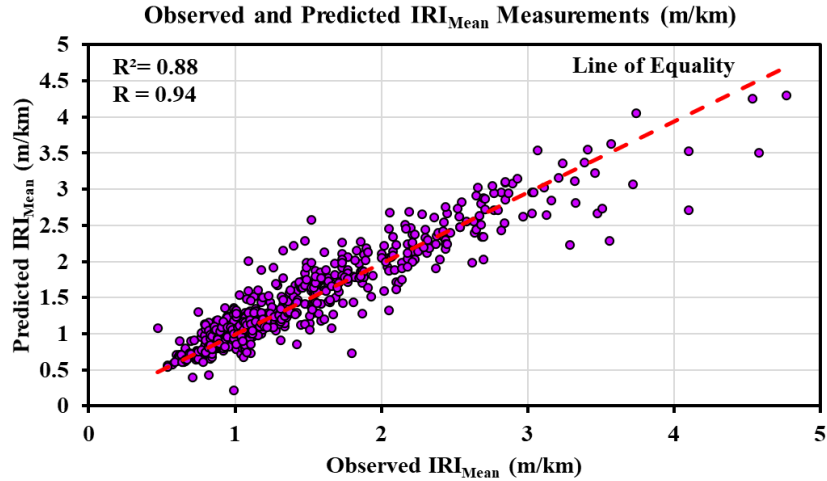


Figure 11. Observed vs. Predicted  $IRI_{Mean}$  for ANN Model 5

#### 4.7. Sensitivity Analysis

To simulate the performance of the developed IRI model, a random section was selected. IRI prediction values were generated for seven different years and compared with the observed IRI values. Figure 12 shows the observed vs. predicted plot of IRI for Section 01-0604 in Alabama. Predicted values were close to observed values. The predicted mean  $IRI_{Mean}$  (1.23) is 10.2% lower than the observed  $IRI_{Mean}$  (1.37). The projected values showed the roughness model behavior was captured by the developed model and the results were accurate and reliable for this section.

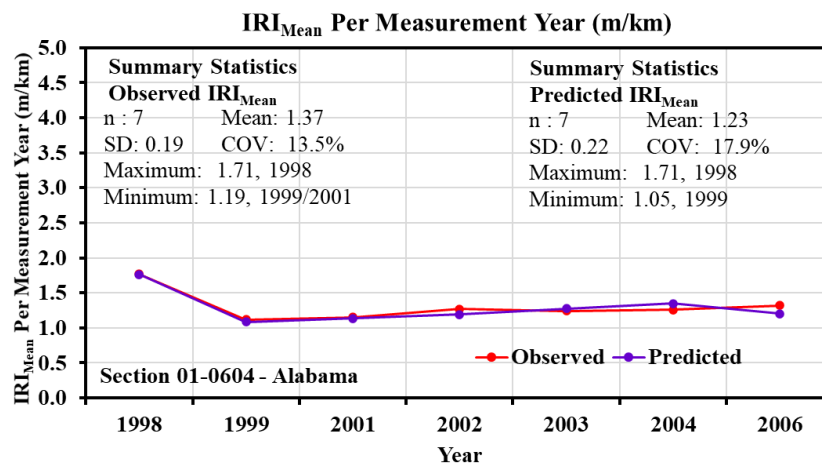


Figure 12. Observed and Predicted Plot of  $IRI_{Mean}$  for Section 01-0604

Another sensitivity analysis was generated to evaluate future predictions using the developed model. The prediction was performed over 10 years from the last measurement. The variable that controls the time factor is the input variable “age”. For the predicted years, previous years’ climatological variables were averaged to be used. The ESAL values for the upcoming years were calculated by assuming an annual growth rate of 1%. No M&R intervention was assumed for this section. The model was able to predict future  $IRI_{Mean}$  values successfully for Section 01-0604. As the road deteriorates over time, the  $IRI_{Mean}$  value will increase without any M&R action. Accordingly, the sensitivity analysis results shown in Figure 13 presents promising predictions for this section.

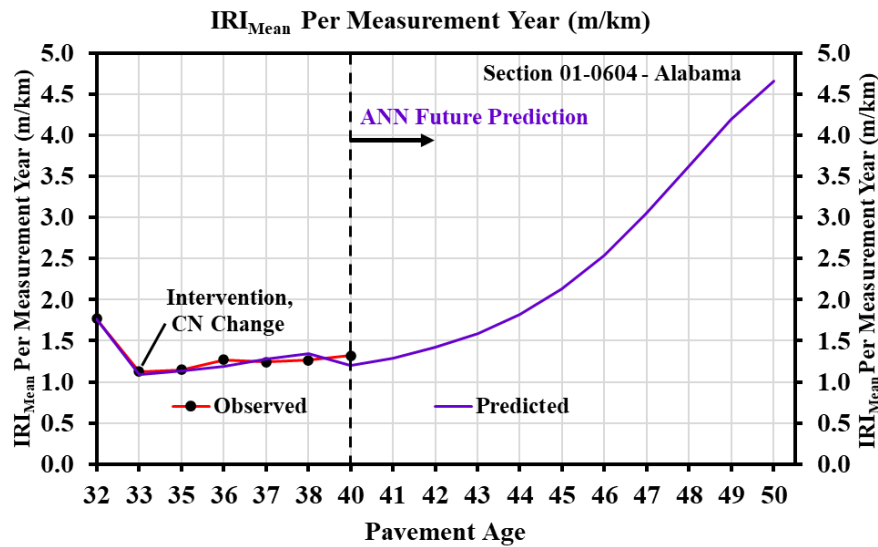


Figure 13. ANN Future Prediction  $IRI_{Mean}$  for Asphalt Overlay on Concrete Section 01-0604

The observed  $IRI_{Mean}$  values start on the pavement age of 32 years and continue until 40 years. The ANN model was used to predict  $IRI_{Mean}$  values from 41 to 50 years. As expected, the predicted  $IRI_{Mean}$  values increase with time and an M&R intervention needs to be performed to maintain the ride quality and road safety for the users. The developed model can be used to identify

in which year the section will need an intervention. Also, if no M&R is performed, the deterioration of the pavement occurs exponentially as can be seen from the slope of the  $IRI_{Mean}$  prediction curve after the pavement age of 43 years. For section 01-0604 an M&R intervention is recommended before the pavement age of 46 years to maintain an acceptable roughness value for the pavement. Therefore, the sensitivity analysis shows that the developed model can be used as a powerful tool by visualizing effective solutions for the future condition of the roads and their M&R planning.

#### 4.8. Conclusions

In this paper, an artificial neural network approach with a backpropagation learning algorithm was utilized to develop IRI prediction models for asphalt overlay on concrete pavements. The best performing ANN model was selected based on the accuracy measures shown in Table 2. Model 5 showed better prediction accuracy for IRI values compared to the other ANN models developed. However, all the developed models are acceptable and can be used for generating reliable predictions. The developed ANN models have efficiently characterized the roughness phenomena on composite pavements. Most of the studies in the literature developed roughness models for asphalt or concrete pavements. This paper can be considered as a unique study that composite pavement roughness models were developed using ANNs approach in the LTPP database. Since asphalt overlay on concrete pavements is a large part of the LTPP database, this study can be employed by the transportation agencies and stakeholders. Therefore, the developed ANN model can be used as a prediction tool for IRI values and guide decision-makers to develop a better M&R plan. Furthermore, the developed model will predict future IRI values without the need for distress data. This will allow local and state agencies to save time from data collection and processing, resulting in cost reductions by providing a tool for better condition



assessment and effective M&R scheduling.

#### 4.8.1. Acknowledgment(s)

The research described in this paper was conducted at the University of Mississippi. The contents of this paper reflect the views of the authors who are responsible for the facts, findings, and data presented herein.

## CHAPTER V: THE EFFECT OF CLIMATE AND TRAFFIC ON COMPOSITE PAVEMENT ROUGHNESS MODELING USING MACHINE LEARNING

Rulian Barros, Hakan Yasarer, Salma Sultana, Yacoub Najjar

Department of Civil Engineering, University of Mississippi, University, MS 38677, United States  
rferrei1@go.olemiss.edu

(Paper submitted to the International Journal of Pavement Engineering, Taylor and Francis)

### 5.1. Introduction

Roadways are responsible for providing mobility, accessibility, and connectivity throughout a country, playing a key role in the nation's quality of life, economic growth, and national security [13]. In the United States, there are more than 4 million miles of roadways where 2.6 million miles are paved roads and 93% of them are surfaced with asphalt [2]. However, some of these paved roads are a result of pavement rehabilitation where concrete pavement is overlaid with asphalt and turned into a composite pavement [3]. These pavements can be more cost-effective and provide better levels of performance both structurally and functionally [3]. The collection and assessment of current pavement information, prediction of future conditions, and decisions regarding reconstruction, rehabilitation, and maintenance strategies to reach a predetermined level of performance are vital elements of Pavement Management Systems (PMS)

[76]. Road pavements continuously deteriorate under the effects of traffic, climate, and a combination of both traffic and climate loads. The capacity of the road to meet the demands of traffic and the environment throughout its design life at an acceptable level is known as performance [77]. Over the last decades, state and federal transportation agencies have established several performance indicators to evaluate the effectiveness and efficiency of their service provision [4]. The International Roughness Index (IRI) is the most well-recognized pavement performance indicator, and it is used by most transportation agencies throughout the world as a standard to determine road surface roughness [78]. The IRI reflects not only the pavement condition but also the ride quality and comfort level of road users [15]. IRI assesses the pavement surface deviations along the road that affect vehicle suspension movement vertically [20]. To ensure superior standards, agencies have to monitor the performance parameters of the entire network frequently to adopt the most proper maintenance operations where needed [18].

Pavement performance prediction models are imperative to highway agencies because they provide decision support for their overall maintenance and budget plan [13]. Pavement performance models are also required to optimize maintenance and rehabilitation (M&R) policies over a planning horizon [76]. However, the use of pavement performance models to estimate the pavement deterioration process is a difficult task strongly related to the serviceability level and assessment of pavement condition [79,80]. Because of the large number of variables and the complex relations between each one of them, the use of simple statistical approaches such as linear regression does not seem appropriate to develop performance-prediction models [21]. The type of algorithms, number of measured data, and effective variables will directly influence the accuracy of the prediction models [81]. Several algorithms and statistical methods have been developed to

predict pavement conditions. In this paper, an Artificial Neural Network (ANN) approach was selected to develop the composite pavement roughness prediction models. The ANN is a predictive modeling technique that uses the concept of human cognition and neural biology [43]. Neural networks can successfully represent complex relationships that are difficult to identify using traditional sequential methods [52]. The Long-Term Pavement Performance (LTPP) program database was used to provide the required data for modeling. The LTPP program is a part of the first Strategic Highway Research Program (SHRP), created in 1987, and funded by the Highway Trust Fund to collect and store performance data from a large number of in-service highways over an extended period to support analysis and product development [58]. In this research, the prediction models were used to assess the influence of climate and traffic data on composite pavement performance and identify the most accurate ANN model.

#### 5.1.1. Objectives

The major objectives of this paper are to:

- (1) Use the LTPP database to analyze roughness data for composite pavements.
- (2) Use the ANN approach to develop roughness models for composite pavements using different independent variables.
- (3) Assess the influence of climate and traffic data on composite pavement performance and identify the best ANN model developed.
- (4) Perform an in-depth analysis using different sections of the database to identify how the developed models are performing and identify the best performance model.

#### 5.2. Literature Review of IRI Prediction Models

Pavement roughness is one of the major parameters to describe pavement irregularities, ride quality, and hence the user perspective about the road. Rough pavements cause an increase in fuel consumption, and a decrease in vehicle efficiency, and result in traffic safety issues that can

lead to the loss of lives and millions of dollars per year [82]. Among other pavement condition indices used to assess pavement surface condition, the IRI is the most used and well-recognized pavement performance indicator. Developed in 1982 by the World Bank and the government of Brazil at the International Road Roughness Experiment (IRRE), the IRI was created to be a standard roughness index and provide a basis for comparing roughness measurements obtained by different procedures. The IRI is based on the quarter-car analysis method, a mathematical model of a vehicle that represents a body and a single wheel [8], with standardized parameter values and a reference simulation speed of 80 km/h [7]. Higher IRI values represent a rough pavement surface that indicates a lower ride quality for road users, while lower IRI values indicate smooth pavements with better ride quality.

Understanding the role that pavement performance plays in the maintenance and rehabilitation activities of the highway network system, transportation agencies have pointed to the need for developing intelligent and efficient pavement performance models. These performance models can help to make decisions regarding pavement maintenance and rehabilitation priorities. However, pavement performance modeling is not a simple task due to the complex relations between pavement structure and its responses to climate and traffic variables, and the interaction between all these elements together [11]. The modeling of asphalt and concrete pavement performance has been investigated in many studies over the years. However, composite pavements have not been well investigated. Also, several performance models have used distress data which are costly to collect and not easily available for all agencies. Traditionally, linear, non-linear, multiple linear regression analysis, Markov chains, mechanistic-empirical, survivor curves, semi-Markov, and Bayesian models have been used for predicting pavement roughness value [20].

However, due to the complexity of the relations between each one of the variables that affect pavement roughness, the use of simple statistical approaches such as linear regression does not seem appropriate to develop performance-prediction models [21]. Advanced modeling techniques using machine learning are appearing as an alternative for predicting pavement deterioration, offering significant improvements over traditional techniques. By feeding data in the form of observations and real-world interactions to computers and making them learn and act intelligently to find complex connections between variables, machine learning can process large volumes of data with a high degree of accuracy [20]. It can also handle noisy and complex data, solve non-linear problems, and once trained, it can make predictions and generalizations at any time [39]. Machine learning techniques hold significant potential for building a modern and robust pavement system due to the excellence in automation and pattern recognition [20]. The literature review shows that a remarkable number of researchers have used ANN to predict pavement roughness.

Lin et al. [74] developed ANNs models to predict IRI using a back-propagation neural network with seven independent variables, age, initial IRI, AC thickness, climatic conditions, pavement distresses, SN, and cumulative ESALs.

Kargah-Ostadi et al. [70] developed an ANN model for the IRI prediction of flexible pavements using a specific pavement study (SPS-5) from the LTPP database. The optimum ANN structure had eight input variables, five hidden nodes within one hidden layer, and one output. The ANN model performed successfully in predicting IRI trends following each kind of treatment in the SPS-5 experiment.

Sollazzo et al. [18] developed an ANN model using input parameters, related to traffic, weather, and structural aspects. The ANN approach was effectively used as a powerful tool for

estimating structural performance using roughness data, especially if compared to linear regression.

Mazari and Rodriguez [83], proposed a methodology that included the application of a hybrid technique that combines gene expression programming (GEP) and ANN. The results were improved when using a hybrid GEP-ANN approach. The hybrid method was found to effectively predict the IRI, and the results were satisfactory compared to similar prediction models from the literature.

Hossain et al. [16] and Hossain [17] developed ANN prediction models for IRI for flexible and concrete pavements, respectively, using climate and traffic data collected from the LTPP database. Both models compared the ANN predicted IRI and measured IRI for flexible and rigid pavements under specific climatic zones (wet freeze for flexible pavement and wet non-freeze for rigid pavement). A two hidden-layered ANN structure with seven independent variables, nine hidden nodes for the first and second hidden layers, and one output (7-9-9-1), using a nonlinear transfer function. An RMSE of 0.027 and 0.01 were found for the flexible and rigid ANN models, respectively, indicating that the IRI prediction was reasonable for both short-term and long-term predictions using only climate and traffic data.

Jaafar [2] developed mechanistic-empirical models using ANN and multiple linear regression techniques for predicting IRI, rutting, and cracking for asphalt pavements using the LTPP database. For the IRI modeling, the ANN architecture used seven independent variables, five hidden nodes within a single hidden layer, and one output (i.e. 7-5-1 ANN structure). The ANN model showed a high coefficient of correlation (R) of 0.72. The results show that ANN models were reasonably accurate for IRI prediction in asphalt pavements.

Duckworth [28] developed pavement performance prediction models using the ANNs approach for flexible pavements based on the Mississippi Department of Transportation (MDOT) database. A two-output model for predicting Performance Condition Rating (PCR) and IRI was found to be the most promising. The ANN model successfully characterized the deterioration behavior with statistical measures in a suitable range.

Kaya et al. [10] developed pavement performance models for flexible and composite (asphalt concrete over the jointed plain concrete pavement) pavement systems in Iowa. ANN-based models were found to be good tools for modeling pavement deterioration when there were many pavement sections with various traffic, thickness, and other various deterioration trends.

Abdelaziz et al. [15] develop an IRI prediction model for both original and overlaid flexible pavements using general pavement studies (GPS-1, GPS-2, and GPS-6) and the specific pavement studies (SPS-1, SPS-3, and SPS-5) of the LTPP database. Multiple linear regression and ANN techniques predict IRI as a function of pavement age, initial IRI, transverse cracks, alligator cracks, and standard deviation of the rut depth. The ANN model resulted in better results compared to the regression model,  $R^2 = 0.75$  and  $R^2 = 0.57$ , respectively. The network consisted of five inputs, three hidden layers with ten nodes each, and one output, 5-10-10-10-1 with LOGSIG as the transfer function.

Yamany et al. [13] used condition data of interstate flexible pavements from eight Midwestern states to estimate three models: fixed-parameters regression, random- parameters regression, and ANN. The ANN model was found to statistically outperform the regression models when estimating pavement roughness across all states.

Jaafar [19] developed IRI prediction models for asphalt pavements using multiple linear



regression and ANN modeling approaches for the Western region on the LTPP database. The variables included in the models were IRI, pavement age, design structural number, ESAL, and a dummy variable for construction number. The feedback ANN predictions using all data showed a higher R of 0.85 compared to the enhanced dummy regression equation.

Solatifar and Lavasani [27] developed an ANN flexible pavement deterioration model based on IRI using Back-Propagation Neural Networks (BPNN) technique using LTPP data for two GPS sections (GPS-1 and GPS-2). The ANN model showed more accurate results leading to a more precise decision in choosing M&R policy.

Ziari et al. [81] investigated the capabilities of ANNs and group method of data handling (GMDH) methods in predicting flexible pavement conditions. Results indicated that ANN models predicted the future condition of pavement with high accuracy in the short and long term, while GMDH models did not have the same accuracy.

Bashar and Torres-Machi [20] assessed the overall performance of three machine learning algorithms (ANN, Random Forest (RF), and Support Vector Machine (SVM)) and compared them to traditional techniques used to predict IRI. The authors recommend the use of ANN to model IRI since its performance was very accurate ( $R=0.930$ ) over a significant number of studies with both small and larger sample sizes.

Literature review to date shows that ANN models performed successfully in predicting IRI values for asphalt and concrete pavements. However, roughness prediction models for composite pavements have not been well investigated. Furthermore, there is a lack of studies exploring the effects of climate and traffic variables on composite pavements. This paper developed a pavement roughness prediction model using the ANNs approach for composite pavement sections and

assessed the effects of climate and traffic variables in the performance models using the LTPP database.

### 5.3. Model Development

#### 5.3.1. Methodology

Pavement performance models can be used as a valuable prediction tool for IRI. It can estimate future pavement conditions, timely maintenance, and major rehabilitation actions. This allows local and state agencies to develop better condition assessment and effective M&R scheduling saving time and resulting in cost reductions. The model development process used in this paper is described, as follows:

- (1) Conduct a literature review of previous studies to identify independent variables responsible for affecting pavement performance.
- (2) Compile databases for composite pavements model development from the LTPP database, which must include all variables identified in step (1).
- (3) Identify missing/erroneous data and evaluate the quality of databases.
- (4) Develop procedures for estimating important missing data from step (3).
- (5) Develop pavement performance models using the ANN approach technique.
- (6) Evaluate the accuracy of the developed models for composite pavements using statistical measures.
- (7) Select the most accurate model based on statistical indicators.
- (8) Perform in-depth analysis for developed models.
- (9) Implement selected performance models for composite pavements.

#### 5.3.2. LTPP Data Collection for Composite Pavements

The LTPP program was created in 1987 as a part of the first SHRP. Five years later the LTPP program was handed to the U.S. Department of Transportation's Federal Highway Administration (FHWA) [58]. A total of 51 States, including Washington D.C., and ten Canadian

provinces were considered in the initial selection of the test sections in the LTPP program seeking to advance highway research and planning [2].

The objectives of the LTPP program are to collect and store performance data from a large number of in-service highways over an extended period to support analysis and product development, analyze the collected data to describe pavement performance, and translate these insights into usable engineering products related to pavement design, construction, rehabilitation, maintenance, preservation, and management. The data collection started in 1989, and 2,509 pavement test sections were selected or constructed for the study. New experiments and studies are being added to monitor the performance of pavement materials and new technologies that were not yet in use when the LTPP program began [58].

The design of experiments was an integral part of the planning and preparation for the LTPP program. In the mid-1980s, two study types were considered, the General Pavement Studies (GPS) and the Specific Pavement Studies (SPS). The GPS experiments used in-service pavement sections to examine general performance by pavement type. In contrast, the SPS experiments were designed to investigate the influence of specific features on pavement performance, these sections were to be constructed specifically for the LTPP study [58]. The LTPP database includes inventory data, monitoring data, and maintenance and rehabilitation data organized in seven modules.

The LTPP data collection has different spatial and temporal locations throughout the U.S. A climatic zone classification was created during the initial recruitment phases of the LTPP test sections and was divided into four different climate zones, identified as wet freeze, wet non-freeze, dry-freeze, and dry non-freeze zones, as shown in Figure 14 [2,59]. This climatic zone map was altered in some places to adhere to State boundaries to ease data collection processes.

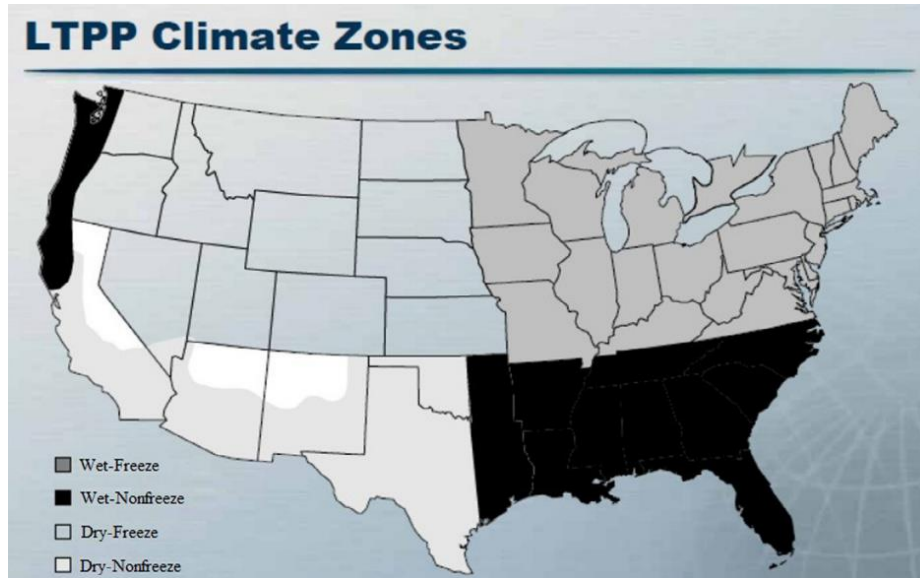


Figure 14. LTPP Climate Zones

Wet climate zones contained in the LTPP database correspond to values of precipitation per year higher than 508 mm (20 inches) while dry climate zones are lower than this value. For freeze/non-freeze zones, the threshold contained in the LTPP database is based on an annual average freezing index of 83 °C (150 °F) days. Locations with an index over this threshold are classified in the freeze zone and those under the threshold in a non-freeze climate zone [84]. Table 3 shows the number of sections in each LTPP climate zone used in this study. A total of 264 sections were identified as composite pavement sections, where the asphalt layer thickness over the concrete layer is equal to or greater than three inches.

Table 3. Summary of Climate Zones

Climate Zones	Number of Sections	Percentage of Sections
Wet, Freeze	189	72%
Wet, Non-Freeze	52	20%
Dry, Freeze	21	8%
Dry, Non-Freeze	2	1%
Total	264	100%

### 5.3.3. Data Processing

Several variables were retrieved from the LTPP database to construct the database used in the study. The data processing for output and input variables is described in this section.

#### 5.3.3.1. Output Variables

As the standard measurement of pavement roughness, the IRI was considered the output variable in the modeling process for this research. For the 264 sections, a total of 16,842 IRI measurements from 1989 to 2018 were found in the database. Each section has two types of IRI measurements, IRI inside wheel path ( $IRI_{Left}$ ) and IRI outside wheel path ( $IRI_{Right}$ ). A mean roughness index ( $IRI_{Mean}$ ) is calculated by averaging the  $IRI_{Left}$  and  $IRI_{Right}$  measurements. On each visit date, several IRI measurement runs were done for each section. By averaging the IRI measurement runs, a single IRI measurement was obtained for  $IRI_{Left}$ ,  $IRI_{Right}$ , and  $IRI_{Mean}$  for each visit date. By doing this, a total of 3,304 IRI measurements were obtained for the 264 sections. In this study, a total of 2,487 datasets from 255 different sections were used to develop the ANN roughness prediction model. Table 4 shows the descriptive statistics of output variables and Figure 15 shows the measurements for  $IRI_{Left}$  and  $IRI_{Right}$ .

Table 4. Descriptive Statistics of Output Variables

Variables	Description	Mean	SD	COV (%)	Max.	Min.
$IRI_{Left}$	IRI Inside Wheel Path, m/km	1.18	0.45	38.2%	3.93	0.39
$IRI_{Right}$	IRI Outside Wheel Path, m/km	1.26	0.51	40.3%	6.35	0.37
$IRI_{Mean}$	Mean of IRI Inside and Outside Wheel Path, m/km	1.22	0.46	37.6%	4.22	0.41

IRI measurements were greater on the outside wheel path compared to the inside wheel path. The average IRI measured for the  $IRI_{Right}$  is 1.26 m/km (79.8 in./mile), which is 6.3% greater than the  $IRI_{Left}$  of 1.18 m/km (74.8 in./mile).

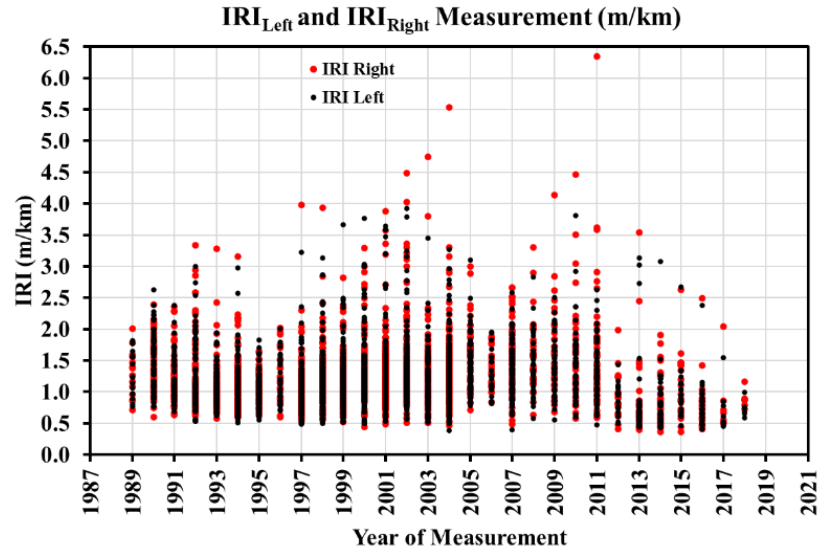


Figure 15.  $IRI_{Left}$  and  $IRI_{Right}$  Measurements

Mean differences were assessed using the independent samples t-test to determine whether  $IRI_{Right}$  and  $IRI_{Left}$  differ on average from each other. The step-by-step procedure is described as follows:

- **Step 1:** Set up the null hypothesis and alternative hypothesis.

Null Hypothesis:  $H_0: \mu_1 = \mu_2$

Where:

$\mu_1$  = Mean of Population 1 for  $IRI_{Left}$

$\mu_2$  = Mean of Population 2 for  $IRI_{Right}$

The population means of the two samples ( $IRI_{Left}$  and  $IRI_{Right}$ ) are equal. This implies that both samples are from the same population.

Alternative Hypothesis:  $H_A: \mu_1 \neq \mu_2$

The population means of the two samples ( $IRI_{Left}$  and  $IRI_{Right}$ ) are not equal. This implies that both samples are from different populations.

- **Step 2:** Select  $\alpha$  probability of Type 1 chance error for  $\alpha$  level of statistical significance.  
 $\alpha = 0.05$  ;  $\alpha/2 = 0.025$  (for two-tailed test)

Figure 16 shows the two-tailed t-test probability distribution.

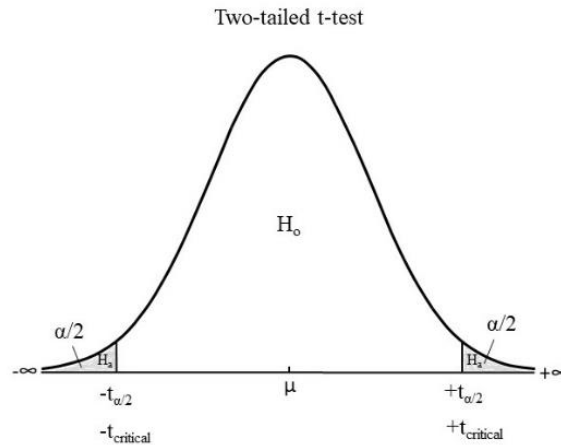


Figure 16. Two-tailed t-test Probability Distribution Graph

- **Step 3:** Define test criteria and the decision rule for rejecting  $H_0$ .

Test criteria:  $t_{critical} = 1.96$  for degree of freedom (dof) = 4,972 and  $\alpha/2 = 0.025$

Decision Rule: Reject  $H_0$  if t-test statistics ( $t_{test}$ ) exceeds the absolute value of  $t_{critical}$  ( $t_{test} > t_{critical}$ ) and probability of significance value,  $p \leq$  Probability of Type-1 chance error,  $\alpha$ .

- **Step 4:** Calculate t test statistics,  $t_{test}$ , and p-significance value.

$$t_{test} = -6.13$$

Probability of significance, p-value  $< 0.001$

- **Step 5:** Interpret the results.

$$t_{test} (-6.13) > t_{critical} (1.96) \text{ and } p (< 0.001) < \alpha (0.05)$$

Therefore, the t-test rejects the null hypothesis. The results show that the difference in the means of  $IRI_{Left}$  and  $IRI_{Right}$  is statistically significant at  $\alpha 0.05$  probability of chance error. This

implies that both samples ( $IRI_{Left}$  and  $IRI_{Right}$ ) are from different populations. For this reason, this study will utilize  $IRI_{Right}$  as a dependent variable since it shows the highest value for pavement roughness, considering the worst-case scenario.

#### 5.3.3.2. Input Variables

An input or independent variable is a variable considered to be the cause of some effect. Pavement deterioration is known to occur due to several factors such as pavement structure, climate, traffic, and maintenance over the years. In this study, an extensive literature review of input variables used in previous studies was done to identify key variables that influence pavement performance. In addition, other input variables were found to assist the understanding of the pavement deterioration process and were included in this paper. To understand the relationship between input and output variables, it is necessary to plot each one of the input variables with the desired output variable. Figure 17 shows plots of each input variable versus  $IRI_{Right}$ .

The “initial  $IRI_{Right}$ ” represents the first IRI value measured in the outside wheel path for a specific section of the LTPP database. The first measurement is usually done when the pavement is built and opened to traffic and indicates the road surface condition at the beginning of the analysis period.

The variable “age” is calculated by subtracting the year when the section was opened to traffic from the IRI measurement year. This variable was selected since it reflects the effects of pavement exposure time to climate and traffic loads. Age is also used to predict pavement performance for future years.

The “cumulative ESAL” (CESAL) variable is a sum of annual ESAL data over the years. The ESAL represents a mixed stream of traffic of different axle loads and axle configurations predicted over the design or analysis period and then converted into an equivalent number of



18,000-lb. single axle loads summed over that period. In some years, the LTPP did not have ESAL information, and an interpolation and extrapolation procedure was applied to calculate ESAL for the missing year. The CESAL variable represents the effects of traffic loads in the model.

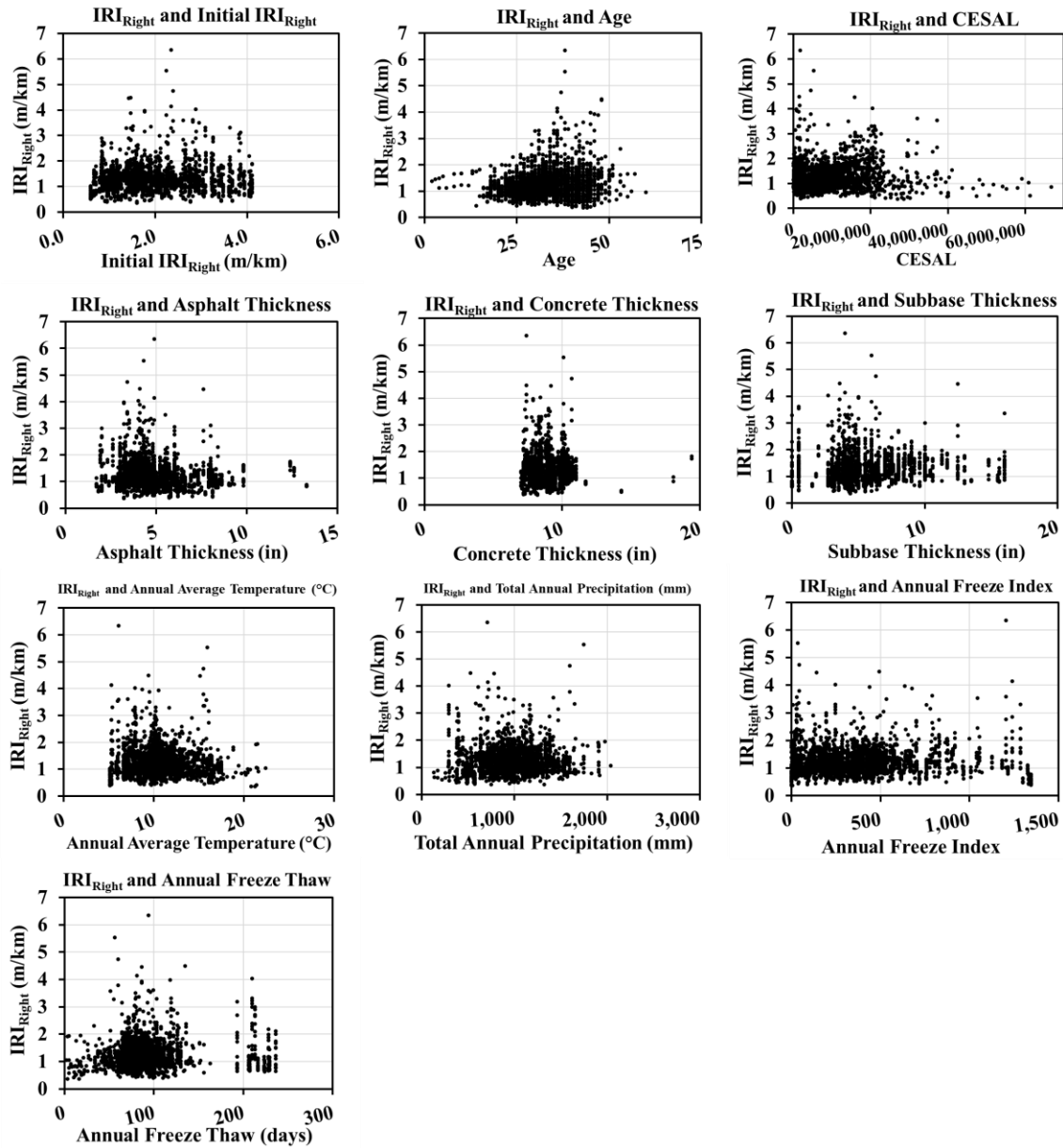


Figure 17. Independent Variables Versus  $IRI_{Right}$

The “asphalt thickness” variable represents the surface asphalt thickness overlaid on a

concrete pavement. The “concrete thickness” variable represents the overall thickness of concrete used for the pavement section. The subbase thickness variable represents the subbase thickness used for the pavement, consisting of an unbound (granular) or bound (treated) subbase. The asphalt thickness, concrete thickness, and subbase thickness represent the structural properties of pavements.

The “annual average temperature” variable represents the average daily mean air temperatures for the year. The temperature changes affect the material properties of pavements and contribute to the deterioration process. The “total annual precipitation” variable represents the sum of monthly precipitation for the year. The “annual freeze index” variable represents the sum of the difference between 0 and the mean daily air temperature when the mean daily air temperature is less than 0 °C for each day of the month. The “annual freeze-thaw” variable represents the number of days in the year when the air temperature goes from less than 0 °C to greater than 0 °C, assuming the minimum daily temperature occurs before the maximum daily temperature. The annual average temperature, total annual precipitation, annual freeze index, and annual freeze-thaw represent the effect of climate variables in the model. Asphalt and concrete pavements, base/subbase, and subgrade materials are susceptible to both temperature and moisture variations, which makes necessary the use of climate variables for more accurate pavement performance models.

Correlation analysis of the available datasets was performed by running a correlation test to obtain the Pearson correlation coefficient ( $r$ ) between all variables. The correlation matrix is capable of summarizing large amounts of data where the goal is to identify patterns and observe whether the variables are correlated with each other. It provides a linear association between the

output and the proposed input variables. However, the correlation matrix only presents linear correlations; if the data is not linearly correlated, it will have a low correlation value. Table 5 summarizes the descriptive statistics of input variables used in this study and the correlation between each input variable and the output variable ( $IRI_{Right}$ ).

Table 5. Descriptive Statistics of Input Variables

Variables	Mean	SD	COV (%)	Max.	Min.	$r$ ( $IRI_{Right}$ )
Initial $IRI_{Right}$ (m/km)	1.88	0.84	44.8%	4.1	0.6	0.13
Age	32.48	7.65	23.6%	60.0	2.0	0.19
Cumulative ESAL	8105128	8183586	1.01	66776707	35566	0.08
Asphalt Thickness (in)	5.02	1.88	37.4%	13.3	1.7	-0.15
Concrete Thickness (in)	9.14	1.08	11.8%	19.4	7.0	-0.02
Subbase Thickness (in)	5.28	2.85	54.0%	16.0	0.0	0.08
Annual Average Temperature ( $^{\circ}C$ )	10.56	2.72	25.8%	22.4	5.1	-0.03
Total Annual Precipitation (mm)	967.61	285.11	29.5%	2036.4	131.8	0.03
Annual Freeze Index	382.04	282.55	74.0%	1345.0	0.0	0.04
Annual Freeze-Thaw (days)	92.98	37.59	40.4%	236.0	2.0	0.03

Notes: Standard Deviation (SD); Coefficient of Variation (COV); Pearson Correlation with

$IRI_{Right}$  ( $r$  ( $IRI_{Right}$ ))

The three highest correlations between input variables and  $IRI_{Right}$  were observed with the variable “age” (0.19) followed by a negative correlation with “asphalt thickness” (-0.15), and “initial  $IRI_{Right}$ ” (0.13). Other variables show low linear correlation values, which means they are not linearly correlated with the dependent variable. However, these variables might follow a non-linear correlation with  $IRI_{Right}$  that cannot be identified by the correlational analysis. Therefore, even though the correlation between independent variables and the dependent variable was low in some cases, all independent variables were used in this study for the development of ANN models.

Some variables used in this study are categorical variables with fixed values of 0 or 1 for modeling purposes and were not included in Figure 17 and Table 5. The categorical variables used in this study are summarized in Table 6.

Table 6. Categorical Variables Summary

Categorical Variables	Description	Categorical Variable Components	r (IRI Right)
Construction Number Code	Categorical variable for M&R	No Intervention	0.28
		Intervention	
Climate Zone	Categorical variable for LTPP climate region	Wet, Non-Freeze	0.05
		Dry, Non-Freeze	-0.09
		Dry, Freeze	-0.06
		Wet, Freeze	0.01
Season	Categorical variable for the month of the measurement	Spring (March, Apr, May)	0.03
		Summer (June, July, Aug)	0.08
		Autumn (Sep, Oct, Nov)	-0.02
		Winter (Dec, Jan, Feb)	-0.12
Subbase Type	Categorical variable for subbase type (stabilized or granular)	Granular	0.12
		Treated	-0.09

The “construction number code” represents M&R interventions performed in the pavement structure. When an M&R intervention is performed, the variable receives a value of 1. On the other hand, if there is no M&R intervention, a 0 value is assigned for the variable. This variable represents the effect of maintenance and rehabilitation activities in pavement roughness modeling. A correlation of 0.28 was observed between this variable and  $IRI_{Right}$ , which was the highest correlation found in the analysis.

The “climate zone” variable represents the climatic zones defined by the LTPP, which consists of four different regions: wet non-freeze, dry non-freeze, dry freeze, and wet freeze. Correlation analysis shows low values for this variable. The “season” variable represents the month in which the measurement was performed, and it consists of four seasons, spring (March, April, May), summer (June, July, August), autumn (September, October, November), and winter (December, January, February), which was found to have the highest correlation (-0.12). The climate region and season variables represent the effect of climate and season on pavement deterioration.

The “subbase type” variable defines the type of subbase used in the pavement section, which consists of two types, unbound (granular) or bound (stabilized), and represents the effect of material properties in the model.

#### 5.4. ANN Development

ANNs are mathematical models built to simulate the neural structure of a human brain [85]. The structure of ANNs is formed from elements called neurons (processing units) and the connection weights between them. The connection links are used to transfer signals between neurons when the information is processed. Input signals are evaluated, and the output signal is determined and transmitted to all neurons that are on the receiving side of the connection links [43]. ANNs help estimate functions or patterns through their learning ability from a large body of data sets. Feed-forward back-propagation, radial basis function, recurrent, and modular neural networks are some of the many types of neural networks [86]. In this paper, a feed-forward neural network with a back-propagation training algorithm was used for the development of the roughness prediction model. Figure 18 shows an example of a simple ANN architecture.

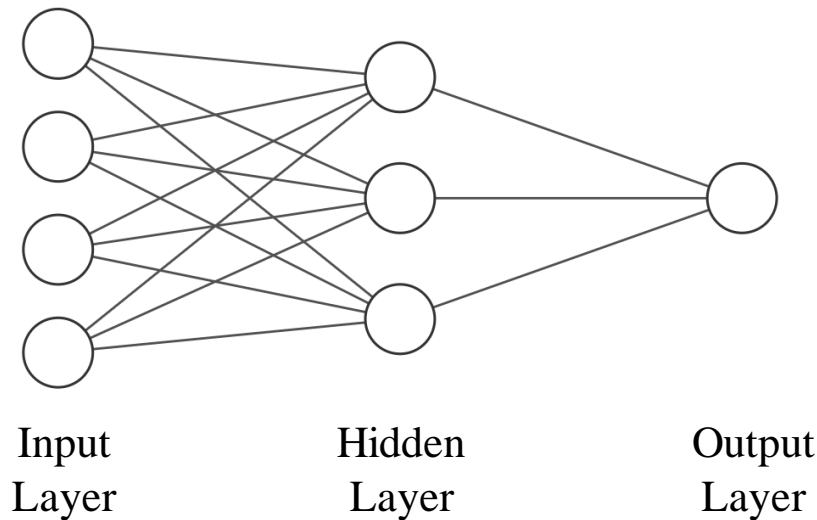


Figure 18. ANN Network Plot

The neural network gains its knowledge through a trained feed-forward network that uses a set of training data consisting of inputs and output(s). The resulting output is compared to the target values, and the back-propagation process adjusts the connection weight to reduce the error between actual and target values. After training, the network provides an approximate functional mapping of any input pattern onto its corresponding output pattern. Then, the validation process is carried out using datasets that were excluded from the model database [2]. This study used different variable types that contain both categorical and continuous variables. A one-hidden layer structure was considered since the use of more than one hidden layer may cause the network to memorize the data in the training phase [46]. The TR-SEQ1 computer program [55] was used to develop the ANN models in this study. Figure 19 shows the curve for the sigmoidal function used in this study for data generalization purposes.

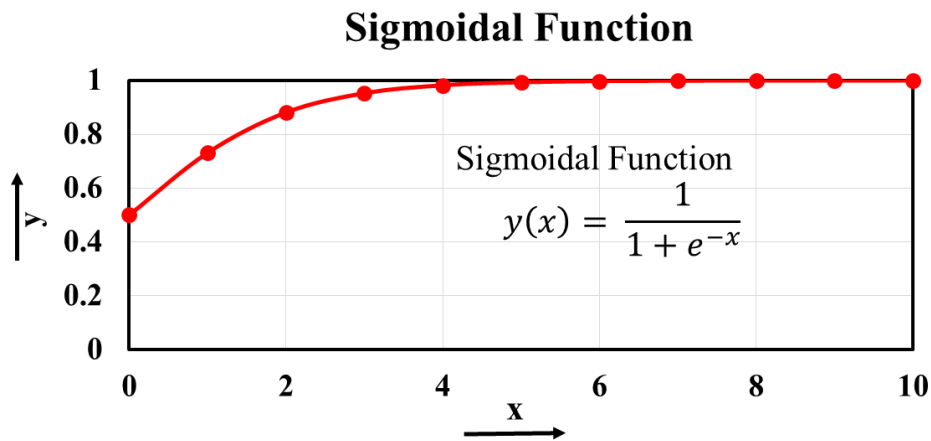


Figure 19. Sigmoidal Function

#### 5.4.1. ANN Methodology

Yasarer and Najjar [51,55] described four successive stages for ANN model development, as follows:

- Stage 1: Determine the ANN architecture. Decide input and output categories based on problem characteristics and ANN knowledge. Classify the datasets as training (50%), testing (25%), and validation (25%) sets.
- Stage 2: Train and test the network on the experimental data to obtain the optimum number of hidden nodes and iterations for the ANN architecture defined in the previous stage.
- Stage 3: Validate the best-performing network from the second stage using the validation database. Check if the accuracy results from the training, testing, and validation databases are comparable. If they are, then stage four maybe not be necessary.
- Stage 4: Retrain the best performing network from stage two using all experimental data to increase prediction accuracy and account for all patterns in the database [44].

Typically, retraining the network with all experimental data is expected to provide reliable predictions and overall better accuracy measures [51]. Several research studies [43,87] recommended that stage four is necessary to arrive at a better-performing network model.

#### 5.4.2. ANN Model Architecture

For this paper, three models were developed using different independent variables. Table 7 shows the variables used for each ANN model.

Model 1 used 10 input variables; however, the “climate region” and “seasons” are categorical variables with four categories each. The “construction number code” is a categorical variable that represents M&R interventions performed in the pavement structure. When an M&R intervention is performed, the variable receives a value of 1. If there is no M&R intervention, a 0 value is assigned for the variable. Therefore, model 1 consists of 17 input variables and one output variable ( $IRI_{Right}$ ). Model 2 did not use the variable responsible for traffic loads, CESAL, but included all climate variables, “air temperature,” “precipitation,” “freeze index,” and “freeze-thaw.” A total of 20 input variables were used for model 2. Model 3 used a total of 21 input variables which include both traffic and climate loads. All ANN models developed in this study

did not use any input variables related to distress data, which requires a lot of time and money to be measured. Therefore, the variables used in this study are easily available for transportation agencies.

Table 7. Input and Output Variables Configuration

No.	Input Variables	Model 1	Model 2	Model 3
1	IRI <sub>0</sub> (Initial IRI Outside Wheel Path) (m/km)	Initial IRI Outside Wheel Path (m/km)	Initial IRI Outside Wheel Path (m/km)	Initial IRI Outside Wheel Path (m/km)
2	Age (Pavement age, years)	Age	Age	Age
3	Climate Zone (Categorical variable for LTPP climatic zone)	Wet, Non-Freeze	Wet, Non-Freeze	Wet, Non-Freeze
		Dry, Non-Freeze	Dry, Non-Freeze	Dry, Non-Freeze
		Wet, Freeze	Wet, Freeze	Wet, Freeze
		Dry, Freeze	Dry, Freeze	Dry, Freeze
4	Seasons (Categorical variable for the season)	Winter (Dec-Feb)	Winter (Dec-Feb)	Winter (Dec-Feb)
		Spring (Mar-May)	Spring (Mar-May)	Spring (Mar-May)
		Summer (June-Aug)	Summer (June-Aug)	Summer (June-Aug)
		Autumn (Sept-Nov)	Autumn (Sept-Nov)	Autumn (Sept-Nov)
5	Subbase Materials (Categorical variable for Subbase materials)	Unbound (Granular) Subbase	Unbound (Granular) Subbase	Unbound (Granular) Subbase
		Bound (Treated) Subbase	Bound (Treated) Subbase	Bound (Treated) Subbase
6	Construction Number Code (Construction Number, variable for M & R)	CN Code: (0,1...) No Intervention 0 Interventions 1...	CN Code: (0,1...) No Intervention 0 Interventions 1...	CN Code: (0,1...) No Intervention 0 Interventions 1...
7	CESAL (Cumulative Equivalent Single Axle Load)	CESAL	-	CESAL
8	$h_{asphalt}$ (Asphalt pavement thickness, in)	$h_{asphalt}$	$h_{asphalt}$	$h_{asphalt}$
9	$h_{concrete}$ (Concrete pavement thickness, in)	$h_{concrete}$	$h_{concrete}$	$h_{concrete}$
10	$h_{base/subbase}$ (Base/Subbase thickness, in)	$h_{subbase}$	$h_{subbase}$	$h_{subbase}$
11	Air Temperature (°C)	-	Air Temperature	Air Temperature
12	Precipitation (mm)	-	Precipitation	Precipitation
13	Freeze Index	-	Freeze Index	Freeze Index
14	Freeze-Thaw (days)	-	Freeze-Thaw	Freeze-Thaw
No.	Output Variables	Model 1	Model 2	Model 3
1	IRI Outside Wheel Path (IRI Right)	IRI <sub>outside</sub>	IRI <sub>outside</sub>	IRI <sub>outside</sub>



### 5.4.3. ANN Model Network Selection

To choose the best-performing network for each model, three statistical accuracy measures were used. The Coefficient of Determination ( $R^2$ ), the Mean Absolute Relative Error (MARE), and the Average Square Error (ASE). Each ANN model development stage (i.e., training, testing, validation, and all-data) was considered, and the best-performing network was selected based on the lowest ASE, lowest mean MARE, and highest  $R^2$ . Table 8 shows the summary network statistics for model 1. The network analyzed can be identified in the first row of the table in an order that depicts the initial hidden node, final hidden node, and the number of iterations. The final structure of each model is shown in the bottom row of the table in an order that represents the number of inputs, number of the final hidden node, and number of output(s), respectively.

Table 8. Summary of Model 1 Network Statistics

Network		4-15-20000	<b>5-19-20000</b>	6-17-20000	7-19-20000	10-13-20000
Training	MARE	19.69	<b>20.69</b>	21.15	21.40	21.81
	$R^2$	0.72	<b>0.73</b>	0.72	0.70	0.64
	ASE	0.00143	<b>0.00150</b>	0.00163	0.00167	0.00188
Testing	MARE	25.66	<b>28.14</b>	27.18	27.14	26.15
	$R^2$	0.26	<b>0.36</b>	0.35	0.35	0.38
	ASE	0.00402	<b>0.00368</b>	0.00374	0.00354	0.00323
Validation	MARE	24.41	<b>27.14</b>	27.90	27.23	25.11
	$R^2$	0.38	<b>0.37</b>	0.33	0.33	0.34
	ASE	0.00340	<b>0.00427</b>	0.00455	0.00406	0.00389
All Data	MARE	19.50	<b>18.34</b>	19.28	18.13	20.38
	$R^2$	0.65	<b>0.70</b>	0.66	0.67	0.62
	ASE	0.00162	<b>0.00142</b>	0.00162	0.00152	0.00181
Chosen		5-19-20000				
Final Structure		17-19-1				

The training development stage used 50% of the total datasets in the database, while testing and validation used 25% each. After developing the possible networks, the all-data stage used all

datasets from the database to retrain the network at its optimal structure and iteration to obtain the generalized response throughout the complete database. For this reason, the all-data stage statistics are used as the main comparison between each network. For model 1, even if it did not have the best results for training, testing, and validation, the 5-19-20000 network outperformed all other networks in the all-data stage with lower ASE (0.00142) and higher R<sup>2</sup> (0.70). The final network structure of model 1 includes 17 input variables, 1 hidden layer with 19 hidden nodes, 20,000 iterations, and 1 output.

Table 9 shows the summary network statistics for model 2.

Table 9. Summary of Model 2 Network Statistics

Models		3-19-20000	4-9-20000	<b>4-19-20000</b>	8-14-20000	12-19-20000
Training	MARE	17.96	21.68	<b>18.00</b>	20.50	20.60
	R <sup>2</sup>	0.78	0.65	<b>0.78</b>	0.71	0.76
	ASE	0.00121	0.00183	<b>0.00123</b>	0.00164	0.00149
Testing	MARE	30.47	26.26	<b>28.42</b>	27.93	28.48
	R <sup>2</sup>	0.27	0.41	<b>0.37</b>	0.39	0.38
	ASE	0.00443	0.00294	<b>0.00392</b>	0.00333	0.00354
Validation	MARE	29.28	26.75	<b>27.93</b>	29.44	28.99
	R <sup>2</sup>	0.24	0.30	<b>0.31</b>	0.37	0.35
	ASE	0.00511	0.00393	<b>0.00455</b>	0.00402	0.00378
All Data	MARE	17.47	20.48	<b>16.54</b>	17.61	17.42
	R <sup>2</sup>	0.73	0.61	<b>0.75</b>	0.70	0.75
	ASE	0.00128	0.00185	<b>0.00120</b>	0.00137	0.00122
Chosen		4-19-20000				
Final Network		20-19-1				

For model 2, climatological variables were included, and the traffic variable (CESAL) was excluded from the model. The 4-19-20000 network showed the lowest ASE (0.00120) and MARE (16.54) values and the highest R<sup>2</sup> (0.75) for the all-data stage compared to the other four networks. Therefore, the 4-19-20000 network was chosen as the best network for model 2. The final network

structure of model 2 includes 20 input variables, 1 hidden layer with 19 hidden nodes, 20,000 iterations, and 1 output.

Table 10 shows model 3 summary network statistics. Model 3 included all climatological and traffic variables in the modeling process, which resulted in better results for the networks. The 14-19-20000 network showed the lowest ASE (0.00107) and MARE (15.67) values and the highest R<sup>2</sup> (0.77) for the all-data stage compared to the other three networks. Hence, the 14-19-20000 network was chosen as the best network for model 3. The final network structure of model 3 includes 21 input variables, 1 hidden layer with 19 hidden nodes, 20,000 iterations, and 1 output.

Table 10. Summary of Model 3 Network Statistics

Models		7-19-20000	10-19-20000	13-19-20000	<b>14-19-20000</b>
Training	MARE	17.67	19.84	19.29	<b>17.35</b>
	R <sup>2</sup>	0.80	0.77	0.77	<b>0.80</b>
	ASE	0.00117	0.00148	0.00147	<b>0.00117</b>
Testing	MARE	28.79	28.89	27.97	<b>27.33</b>
	R <sup>2</sup>	0.28	0.34	0.29	<b>0.39</b>
	ASE	0.00532	0.00447	0.00480	<b>0.00404</b>
Validation	MARE	28.06	30.10	27.49	<b>26.83</b>
	R <sup>2</sup>	0.32	0.35	0.43	<b>0.40</b>
	ASE	0.00503	0.00474	0.00380	<b>0.00369</b>
All Data	MARE	16.24	17.32	16.91	<b>15.67</b>
	R <sup>2</sup>	0.76	0.73	0.75	<b>0.77</b>
	ASE	0.00111	0.00129	0.00116	<b>0.00107</b>
Chosen		14-19-20000			
Final Network		21-19-1			

## 5.5. ANN Model Results

### 5.5.1. ANN Best Model Selection

After choosing the best networks for each model, it is necessary to compare all three models developed and their networks to choose the best model performance. Table 11 shows the summary

of the best network performance for each ANN model. The accuracy measures indicate reliable results for all models developed; however, model 3 outperformed all other models with significant improvements in training, testing, validation, and all-data stages. Compared to models 1 and 2, model 3 the all-data stage showed a MARE 15% and 5% lower, an  $R^2$  10% and 3% higher, and an ASE 21% and 8% lower, respectively. These results show a better performance for the model that uses both climate and traffic variables (model 3) compared to the ones that just use traffic or climate variables.

Table 11. Summary of Best Network Performance for each ANN Model

Models		1	2	3
Structure (Initial Hidden Node - Final Hidden Node – Iterations)		5-19-20000	4-19-20000	<b>14-19-20000</b>
Training	MARE	20.69	20.60	<b>17.35</b>
	$R^2$	0.73	0.76	<b>0.80</b>
	ASE	0.0015	0.0015	<b>0.0012</b>
Testing	MARE	28.14	28.48	<b>27.33</b>
	$R^2$	0.36	0.38	<b>0.39</b>
	ASE	0.0037	0.0035	<b>0.0040</b>
Validation	MARE	27.14	28.99	<b>26.83</b>
	$R^2$	0.37	0.35	<b>0.40</b>
	ASE	0.0043	0.0038	<b>0.0037</b>
All Data	MARE	18.34	16.54	<b>15.67</b>
	$R^2$	0.70	0.75	<b>0.77</b>
	ASE	0.0014	0.0012	<b>0.0011</b>
Final Network		17-19-1	20-19-1	<b>21-19-1</b>
		# of Inputs - Final Hidden Node - # of Outputs		

Figure 20 shows a graphical comparison of the accuracy measures for the all-data stage between all three models.

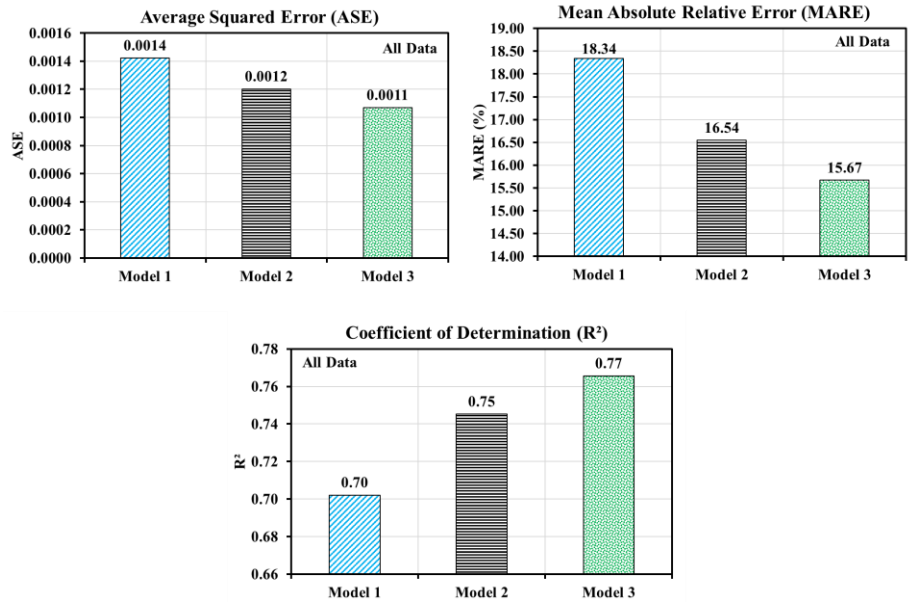


Figure 20. Graphical Comparison of Accuracy Measures

The graphical comparison confirms that model 3 had the best statistical measures, and therefore it was chosen as the best performing model. Figure 21 shows the network architecture of the best-performing ANN model (model 3).

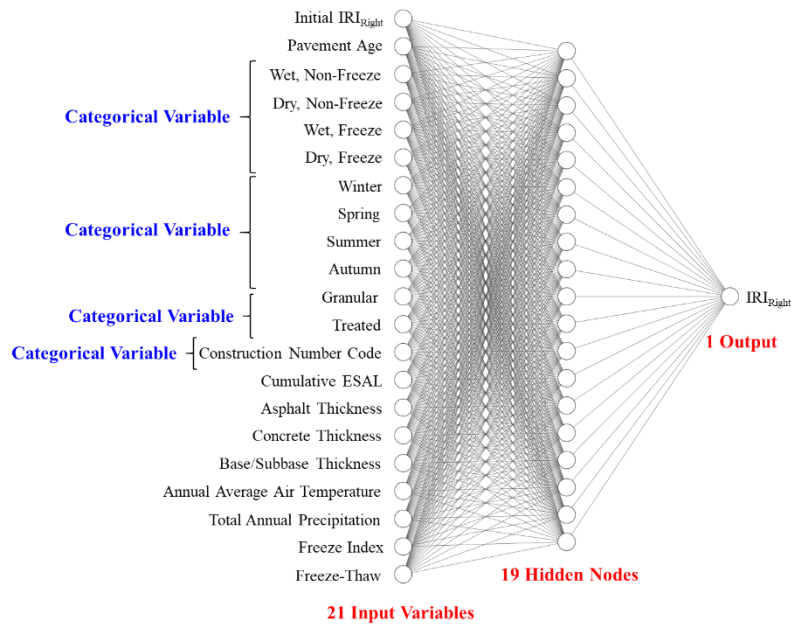


Figure 21. Network Architecture of the Best Performing ANN Model (model 3)

Figure 22 shows the observed  $IRI_{Right}$  values collected from the LTPP database and the predicted  $IRI_{Right}$  using model 3. A section sequence number was generated to identify the data points in the database and to help the visualization of the graph. The ANN model was capable to capture the pavement deterioration behavior. However, some higher IRI values were not captured by the model, which predicted lower values compared to the observed IRI.

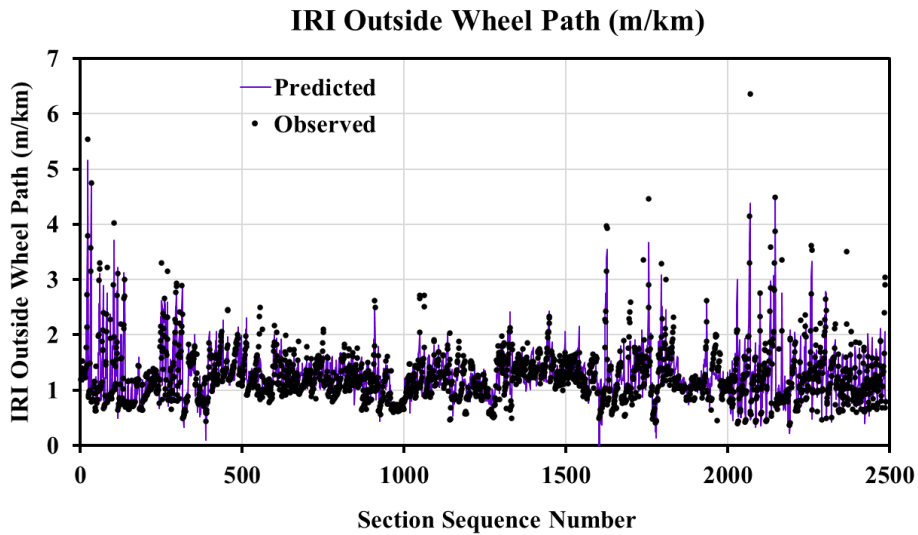


Figure 22. Model 3 - Observed vs Predicted Values for  $IRI_{Right}$

Figure 23 shows that predicted values clustered around the line of equality until IRI values were equal to 3m/km. When the observed IRI values were greater than 3 m/km, the model was not as accurate as it was for lower IRI values. Nevertheless, a high  $R^2$  of 0.77 was obtained for the ANN model 3.

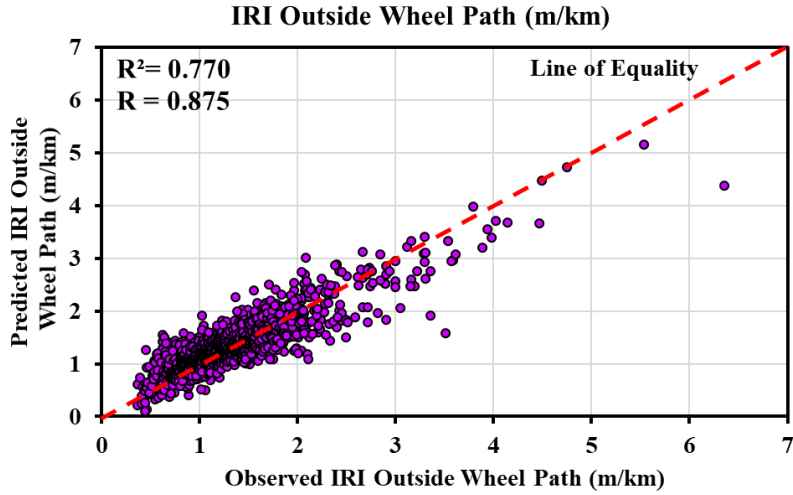


Figure 23. Observed vs. Predicted  $IRI_{Right}$  for ANN Model 3

### 5.6. In-depth Analysis

An in-depth analysis was performed to evaluate how the developed models performed in four random sections of the database. Each plot shows the observed and predicted IRI values for each model. Figure 24 shows the observed vs. predicted plot of IRI for Section 04-0607 in Arizona.

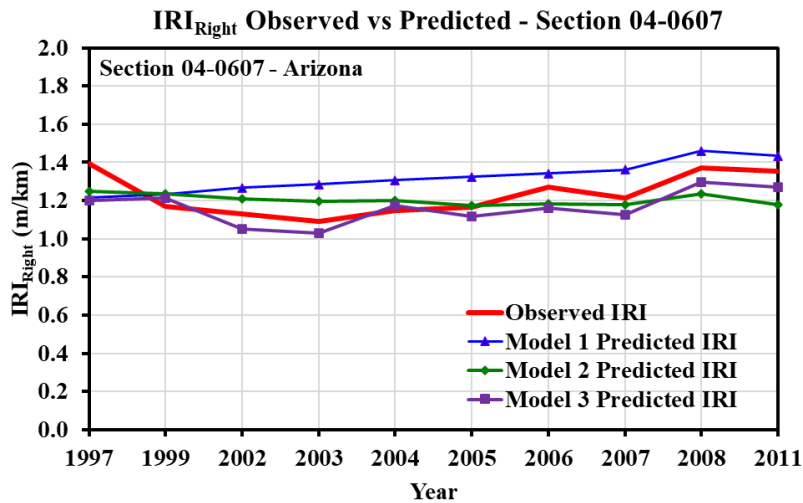


Figure 24.  $IRI_{Right}$  Observed vs Predicted for Section 04-0607 - Arizona

Section 04-0607, located in Arizona, is part of the dry freeze climate region, which

indicates precipitation per year lower than 508 mm (20 inches) and an annual average freezing index lower than 83 °C (150 °F) days. Predicted values were close to observed values for all models. The mean difference percentage with observed values was lower for model 3 (1.80%) compared to model 2 (8.15%) and model 1 (13.61%). Also, model 3 showed to be the best in capturing the pavement deterioration behavior and closely followed the observed values. Model 1 and model 2 were not capable of obtaining the variations of IRI and, for this reason, could not achieve better IRI predictions for each analyzed year.

Figure 25 shows the observed vs. predicted plot of IRI for Section 05-A607 in Arkansas.

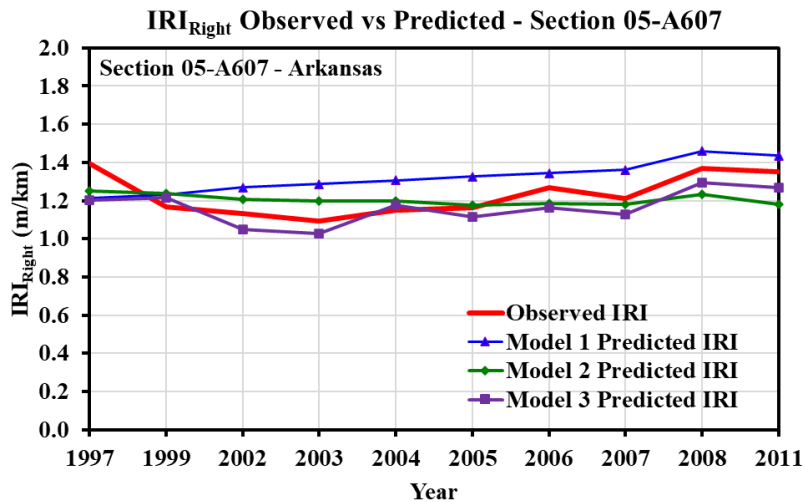


Figure 25. IRI<sub>Right</sub> Observed vs Predicted for Section 05-A607 - Arkansas

Section 05-A607, located in Arkansas, is part of the wet non-freeze climate region, which indicates precipitation per year higher than 508 mm (20 inches) and an annual average freezing index lower than 83 °C (150 °F) days. The mean difference percentage with observed values was lower for model 2 (-2.10%) compared to model 3 (-5.37%) and model 1 (7.54%). However, models 1 and 2 were not able to follow the IRI variations from the observed data; instead, they showed straight lines upwards and downwards. Model 3 was the only model capable of capturing the



variations observed in the IRI behavior over time. Therefore, model 3 had the best performance compared to other models.

Figure 26 shows the observed vs. predicted plot of IRI for Section 18-0607 in Indiana.

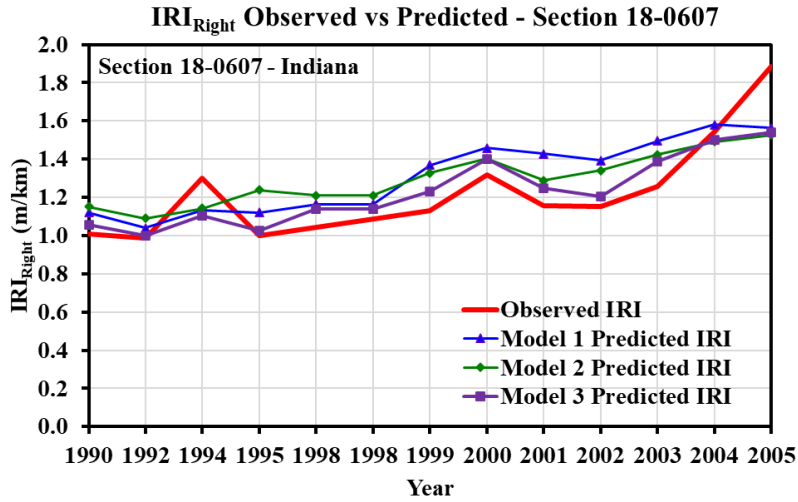


Figure 26. IRI<sub>Right</sub> Observed vs Predicted for Section 18-0607 – Indiana

Section 18-0607, located in Indiana, is part of the wet freeze climate region, which indicates precipitation per year higher than 508 mm (20 inches) and an annual average freezing index higher than 83 °C (150 °F) days. The mean difference percentage indicates lower values for model 3 (0.37%) compared to model 2 (5.53%) and model 1 (6.98%). Different from Figure 24 and Figure 25, all three models in section 18-0607 were able to capture the variations in IRI behavior but model 3 predictions showed to be closer to the observed values. Figure 27 shows the observed vs. predicted plot of IRI for Section 48-7165 in Texas.

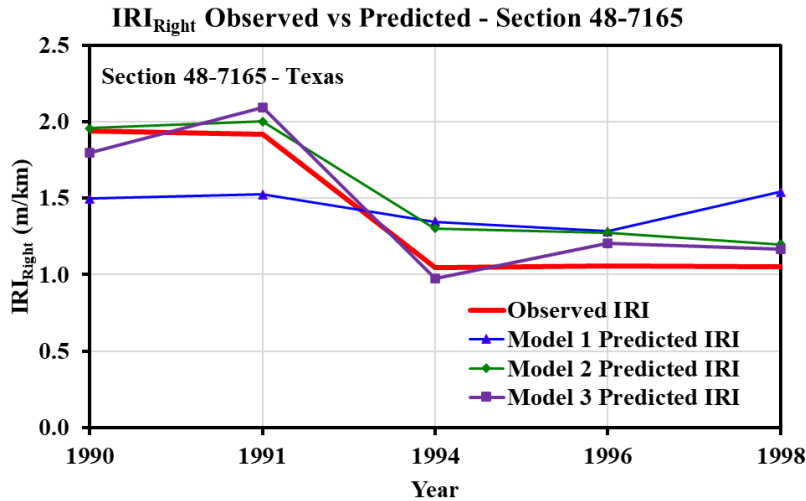


Figure 27. IRI<sub>Right</sub> Observed vs Predicted for Section 48-7165 - Texas

Section 48-7165, located in Texas, is part of the wet-non freeze climate region, which indicates precipitation per year higher than 508 mm (20 inches) and an annual average freezing index lower than 83 °C (150 °F) days. The mean difference percentage indicates lower values for model 1 (2.60%) compared to model 3 (3.34%) and model 2 (10.33%). However, model 1 predictions did not follow the same behavior as the observed values. Model 3 showed to be the best to capture the pavement roughness behavior.

Therefore, the in-depth analysis showed that all developed models were reasonably accurate and reliable for each section analyzed. However, model 3 showed better IRI predictions and better capacity to recognize IRI variations over time for the composite pavement sections.

### 5.7. Conclusion

In this study, the LTPP database was used to analyze roughness data for composite pavements. An artificial neural network approach with a backpropagation algorithm was used to develop roughness prediction models and assess the influence of climate and traffic data using different input variables for three models. Model 1 used the CESAL variable to represent the

effects of traffic loads in the model while model 2 used air temperature, precipitation, freeze index, and freeze-thaw to represent the effects of climate variables in the pavement deterioration process. Model 3 used all variables to assess the importance of both traffic and climate in the modeling process.

The results indicate that all developed models provided reliable results with good accuracy between observed and predicted values, as shown in Table 11. Model 3 outperformed all other models in training, testing, validation, and all-data stages. The final network structure of model 3 (Figure 21) includes 21 input variables, 1 hidden layer with 19 hidden nodes, 20,000 iterations, and 1 output.

An in-depth study was performed for different sections of the database and the analysis reveals that model 1 and model 2 were not able to capture all the variations of IRI over time. These variations were better predicted in model 3 when both traffic and climate variables were used together in the model. Due to the complexity of the deterioration phenomena, traffic and climate variables must be used together to assist the model to recognize the roughness behavior, identify the influence of each variable and the relation between them, and predict reasonable values according to the data provided. This study proves the importance of traffic and climate variables in the modeling process and demonstrates that the model provides better prediction values when both traffic and climate variables are used together, compared to the use of each one of them separately. Furthermore, the models developed in this study used the composite pavements database, which the literature lacks in research studies and modeling development. The prediction models also do not use any distress data, which can save time and money from data collection and processing for transportation agencies.

### 5.7.1. Recommendation

The recommendations for better model development and further study of pavement roughness modeling are, as follows:

- In this study, the categorical variable of construction number code was used to identify M&R interventions on the pavement. Further study of this variable is necessary to improve model performance.
- Do further study on climate variables to identify new variables able to improve model predictions.
- Develop different models according to each climate region to verify if IRI predictions will be more accurate.
- Develop ANN models using the dynamic approach.
- Develop a graphical user interface (GUI) to create a product that can be used for agencies to predict IRI.

### 5.7.2. Acknowledgments

The research described in this paper was conducted at the University of Mississippi. The findings of the Highway Pavement Condition Deterioration Modeling Considering Maintenance History project sponsored by the National University Transportation Center (UTC) and National Center for Transportation Infrastructure Durability & Life-Extension, led by Washington State University (WSU), were very beneficial to this research. The authors would like to acknowledge Dr. Waheed Uddin for his significant contribution to this research with his guidance and recommendations. The contents of this paper reflect the views of the authors who are responsible for the facts, findings, and data presented herein.

CHAPTER VI: INTERNATIONAL ROUGHNESS INDEX MODEL FOR COMPOSITE  
PAVEMENTS IN THE LTPP WET NON-FREEZE CLIMATE REGION: MACHINE  
LEARNING AND REGRESSION APPROACHES

Rulian Barros  
Ph.D. Student, Department of Civil Engineering  
The University of Mississippi, University, MS 38677-1848  
Email: rferrei1@go.olemiss.edu

Hakan Yasarer  
Assistant Professor, Department of Civil Engineering  
The University of Mississippi, University, MS 38677-1848  
Email: hy@go.olemiss.edu

Salma Sultana  
Ph.D. Candidate, Department of Civil Engineering  
The University of Mississippi, University, MS 38677-1848  
Email: ssultana@go.olemiss.edu

(Paper accepted for poster presentation at the 101<sup>st</sup> Annual Meeting of the Transportation  
Research Board (TRB) 2022, Washington D.C., 9-13 January 2022. )

### 6.1. Abstract

Pavement performance prediction models are an important part of the Pavement Management System (PMS), allowing decision-makers to do better budget allocation plans for future pavement maintenance and rehabilitation actions. This study intends to analyze roughness data for composite pavements (asphalt overlay on concrete) in the wet non-freeze climate zone of

the Long-Term Performance Pavement (LTPP) database, develop pavement roughness prediction models using Artificial Neural Network (ANN) and Multiple Linear Regression (MLR) approaches, evaluate the accuracy of developed models, and compare their results to identify the best performance model. A total of 49 sections with 353 datapoints were used for the analysis. The ANN and MLR models included 11 input variables and 1 output variable. The results indicated the ANN model outperformed the MLR model with a MARE (13.14) 53% lower and an ASE (0.00182) 99% lower, compared to the MLR model. The  $R^2$  value improved from 0.37, obtained by the MLR model, to 0.86, obtained by the ANN model. This translates into 132% better prediction accuracy by using the ANN-based model. Therefore, the ANN model showed to be more accurate than the MLR model and was hence chosen as the best-performance model. The use of a specific climate region helped the model to capture almost 90% of the variability, which may be not viable when using data from all climate zones together. Furthermore, the developed models did not use any distress data for input variables, which can help transportation agencies to save time and money from data collection and processing.

Keywords: International Roughness Index (IRI), Artificial Neural Network (ANN), Composite Pavement, Long-Term Pavement Performance (LTPP), Wet Non-Freeze

## 6.2. Introduction

Reliable and precise assessment of the existing and future pavement conditions are key components of a successful pavement management system (PMS) [82]. One of the goals of PMS is to increase pavement life by considering the effects of pavement material and environment on pavement performance. Over the last two decades, transportation agencies have established several performance indicators to evaluate the effectiveness and efficiency of their service provision [4].

The most well-recognized pavement performance indicator is the International Roughness Index (IRI), which is a standard to determine road surface roughness [20]. Developed in 1982 by the World Bank and the government of Brazil at the International Road Roughness Experiment (IRRE), the IRI was created to be a standard roughness index. The IRI describes the irregularities in the pavement surfaces that affect the ride quality experienced by road users, and it is useful for making objective decisions related to the management of road networks [2,7,8]. Higher IRI values represent a rough pavement surface that indicates a lower ride quality for users, while lower IRI values indicate smooth pavements with better ride quality.

Pavement performance prediction models are an important part of PMS since they allow decision-makers to develop better budget allocation plans for future pavement maintenance and rehabilitation (M&R) actions [37]. Pavement performance models are necessary to identify rehabilitation needs, analyze rehabilitation effects, and estimate future pavement conditions to implement different M&R activities to extend the pavement life cycle and provide good surface quality for road users [27]. However, the use of prediction models to estimate the pavement deterioration process is a difficult task related to the serviceability level and assessment of pavement condition [79,80]. A growing body of literature explores different modeling techniques for pavement performance prediction. Several studies explore the use of regression models to predict pavement roughness [12–15,68]. However, due to the complex relations between pavement structure and its responses under climate and traffic variables, and the interaction among all of these elements [11], more advanced modeling techniques using machine learning proved to be promising. The Artificial Neural Network (ANN) technique was used in several studies [2,10,12–20], offering significant improvements over traditional techniques, such as regression, by

processing large volumes of data with a higher degree of accuracy. However, most of the studies developed predictions models for asphalt or concrete pavements, which resulted in a lack of research on composite pavements that mainly consist of asphalt overlaid concrete pavements. Compared to flexible or rigid pavements, composite pavements can provide better performance measures both structurally and functionally, and accordingly, can be considered a cost-effective alternative [2]. Therefore, developing performance models for composite pavements can be very beneficial to understanding the contribution of this road infrastructure in the pavement management system.

Several pavement performance prediction models in the literature that uses the Long-Term Performance Pavement (LTPP) database try to account for the influence of climate using data from different climate zones in the same model. The LTPP climate zones include four different zones identified as wet freeze, wet non-freeze, dry-freeze, and dry non-freeze zones [2,59]. The LTPP used the precipitation per year to identify wet (higher than 508 mm) and dry (lower than 508 mm) climate zones. For freeze and non-freeze zones, the threshold used was based on an annual average freezing index of 83 °C (150 °F) days. Locations with an index over this threshold are classified in the freeze zone and those under the threshold in a non-freeze climate zone [84]. The use of data from different climate zones in the same model might affect pavement performance prediction accuracy since the model needs to explain a higher degree of variability to perform well. An alternative to this would be to develop one model for each climate zone instead of using a unique model for all climate zones together.

Therefore, the objectives of this paper are to develop pavement roughness prediction models using Multiple Linear Regression (MLR) and ANN approaches and compare which



modeling technique provides more accurate results for composite pavement sections in the wet non-freeze climate region of the LTPP database. The prediction models were proposed based on the influence of pavement structure, climate, and traffic data.

#### 6.2.1. Objectives

The main objectives of this paper are to:

1. Analyze roughness data for composite pavement sections in the wet non-freeze region using the LTPP database.
2. Identify key parameters responsible for affecting pavement performance.
3. Develop roughness prediction models using ANN and MLR approaches.
4. Evaluate the accuracy of ANN and MLR models using statistical measurements.
5. Perform a comparison between ANN and MLR modeling predictions.

#### 6.3. Literature Review

The literature shows that a significant number of researchers have used MLR and ANN to predict pavement roughness. Some studies used only the MLR approach to develop prediction models, such as Khattak et al. [68] who developed IRI prediction models using regression analysis for overlay treatment of composite and flexible pavements in the state of Louisiana. An  $R^2$  of 0.63 and 0.47 were found for the composite and flexible pavements, respectively. However, most of the studies compare the accuracy of MLR and ANN modeling techniques to identify the best approach. Prediction models for flexible pavements are the most common in the literature.

Sollazzo et al. [18] developed an ANN model for flexible pavement using input parameters related to traffic, weather, and structural aspects. The ANN approach was effectively used as a powerful tool for estimating structural performance using roughness data, especially when compared to

linear regression.

Jaafar [2] developed prediction models using ANN and MLR techniques for predicting IRI, rutting, and cracking for asphalt pavements using the LTPP database. For the IRI modeling, the ANN architecture used seven independent variables, five hidden nodes within a single hidden layer, and one output (i.e., 7-5-1 ANN structure). A high coefficient of correlation (R) of 0.72 and 0.63 was found for the ANN and MLR models, respectively. The results show that both models were reasonably accurate for IRI prediction in asphalt pavements, but the ANN model outperformed the MLR with higher accuracy.

Jaafar [19] also developed IRI prediction models for asphalt pavements using multiple linear regression and ANN modeling approaches for the Western region on the LTPP database. The variables included in the models were IRI, pavement age, design structural number, ESAL, and a dummy variable for construction number. Correlation values of 0.85 and 0.57 were found for the ANN and MLR models, respectively. Therefore, the ANN model showed to be more accurate than the MLR approach.

Choi [14] developed an ANN prediction model for flexible pavements on a granular base from three states: Texas, New Mexico, and Arizona. The results show that the ANN model could provide a reasonable explanation for their predictive behavior and model the relationship between input variables and pavement performance.

Hossain et al. [16,17] developed an ANN prediction model for flexible pavements using climate and traffic data collected from the LTPP database. The study compared the ANN-predicted IRI and measured IRI for flexible pavements under specific climatic zones (wet freeze) with a two hidden-layered ANN structure with seven independent variables, nine hidden nodes for the first

and second hidden layers, and one output (7-9-9-1), using a nonlinear transfer function. An RMSE of 0.027 was found for the flexible ANN model, indicating that the IRI prediction was reasonable for both short-term and long-term predictions using only climate and traffic data.

Some studies from the literature developed models for both flexible and composite pavements.

Kaya et al. [10] developed ANN-based models for flexible and composite pavement systems in Iowa. ANN pavement performance prediction models were found to be good tools for modeling pavement deterioration when there were many pavement sections with various traffic, thickness, and other various deterioration trends.

Abdelaziz et al. [15] develop IRI prediction models for flexible and composite pavements using general pavement studies (GPS-1, GPS-2, and GPS-6) and specific pavement studies (SPS-1, SPS-3, and SPS-5) of the LTPP database. The MLR and ANN techniques predict IRI as a function of pavement age, initial IRI, transverse cracks, alligator cracks, and standard deviation of the rut depth. The ANN model resulted in better accuracy measures compared to the MLR model,  $R^2 = 0.75$  and  $R^2 = 0.57$ , respectively.

Some studies specified some states or specific climate zones, as can be observed in Yamany [13], that developed pavement performance models for flexible pavements using data from eight Midwestern states, and Zeiada [12], that developed prediction models for warm climate regions in the LTPP database. Both studies found that by specifying these characteristics their prediction models performed better since the data gather the same characteristics and helped the model to understand the variability of the datasets.

The literature review to date shows that MLR and ANN models performed successfully in predicting IRI values for asphalt pavements. However, prediction models for composite pavements

have not been well investigated as well as the use of specific climate regions. Therefore, this paper develops pavement roughness prediction models using the ANN and MLR approaches for composite pavement sections in the wet non-freeze climate zone from the LTPP database.

#### 6.4. Model Development

##### 6.4.1. Methodology

Pavement performance models can estimate future pavement conditions, timely maintenance, and major rehabilitation actions. The prediction models allow transportation agencies to develop better condition assessment and effective M&R scheduling, saving time and subsequently reducing costs. Figure 28 shows the pavement performance modeling methodology flowchart for the LTPP database.

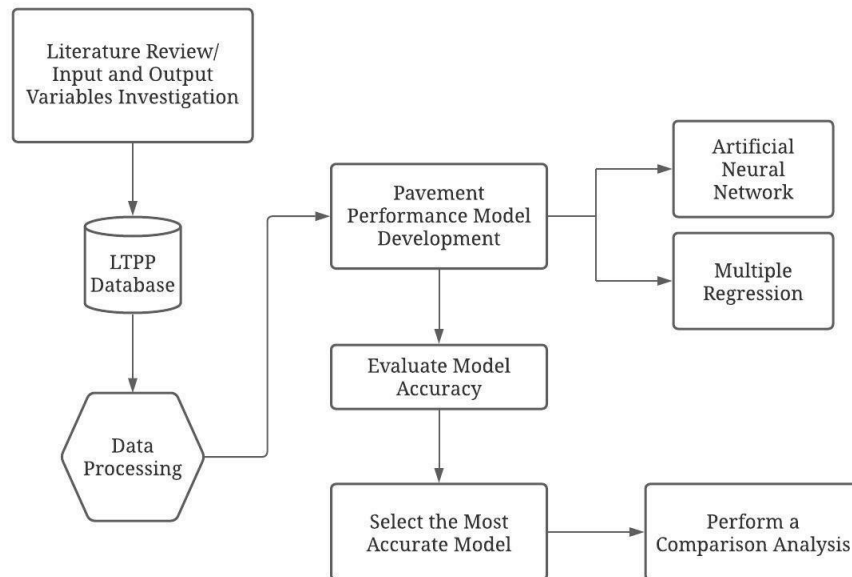


Figure 28. Pavement Performance Modeling Methodology

The model development methodology used in this paper is described as follows:

1. Conduct a literature review of previous studies to identify independent variables responsible for affecting pavement performance.
2. Compile databases for composite pavements in the wet non-freeze region of the LTPP database, which must include all variables identified in step (1).
3. Identify missing/erroneous data and evaluate the quality of databases.
4. Develop pavement performance models using the ANN and MLR approach techniques.
5. Evaluate the accuracy of the developed models using statistical measures.
6. Select the most accurate models based on statistical indicators.
7. Perform comparison analysis for ANN and MLR developed models.

#### 6.4.2. Data Collection

The LTPP program was created in 1987 to collect and store performance data over an extended period to support analysis and product development, analyze the collected data to describe pavement performance, and translate these insights into usable engineering products related to pavement design, construction, rehabilitation, maintenance, preservation, and management. The initial selection of the test sections of the LTPP includes 50 states and Washington D.C., and ten Canadian provinces. The data collection started in 1989, and 2,509 pavement test sections were selected or constructed for the study. New experiments and studies are being added to monitor the performance of pavement materials and new technologies that were not yet in use when the LTPP program began [58]. A climate zone classification was developed since the data collection had different spatial and temporal locations throughout the U.S. The LTPP climate zones include four different zones, identified as wet freeze, wet non-freeze, dry-freeze, and dry non-freeze zones, as shown in Figure 29 [2,59]. The LTPP used the precipitation per year

to identify wet (higher than 508 mm) and dry (lower than 508mm) climate zones. For freeze and non-freeze zones, the threshold used was based on an annual average freezing index of 83 °C (150 °F) days. Locations with an index over this threshold are classified in the freeze zone and those under the threshold in a non-freeze climate zone [84].

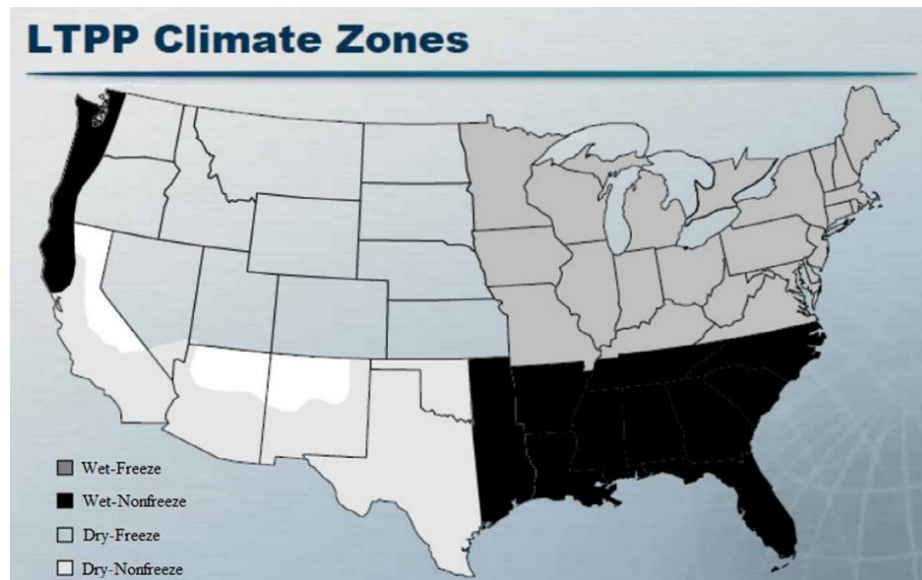


Figure 29. LTPP Climate Zones [2,59]

Using the LTPP database, a total of 264 sections were identified as composite pavement sections, where the asphalt layer thickness over the concrete layer is equal to or greater than three inches. In this study, only composite pavement sections located in the wet non-freeze climate zone were used. A total of 49 sections fit the criteria and were used for the analysis.

#### 6.4.3. Data Processing

Several variables were retrieved from the LTPP database to construct the database used in the study. The output and input variables are described in this section.

##### 6.4.3.1. Output Variables

The IRI is considered the standard measurement of pavement roughness, and it was used as the output variable for modeling. Each section had two types of IRI measurements; IRI inside

wheel path ( $IRI_{Left}$ ) and IRI outside wheel path ( $IRI_{Right}$ ). A mean roughness index ( $IRI_{Mean}$ ) was calculated by averaging the  $IRI_{Left}$  and  $IRI_{Right}$  measurements. On each visit date, several IRI measurement runs were done for each section. By averaging the IRI measurement runs, a single IRI measurement was obtained for  $IRI_{Left}$ ,  $IRI_{Right}$ , and  $IRI_{Mean}$  for each visit date. By doing this, a total of 353 IRI measurements from 1989 to 2018 were found for the 49 sections. Figure 30 shows the measurements for  $IRI_{Left}$  and  $IRI_{Right}$ .

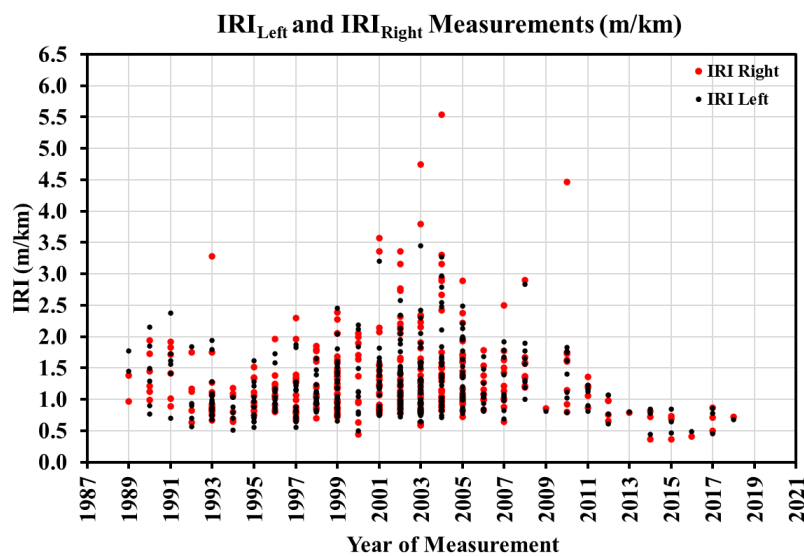


Figure 30.  $IRI_{Left}$  and  $IRI_{Right}$  Measurements

IRI measurements for  $IRI_{Right}$  (1.31 m/km or 83.0 in./mile) were 8% greater than the  $IRI_{Left}$  (1.22 m/km or 77.3 in./mile). Mean differences were assessed using the independent samples t-test to determine whether  $IRI_{Right}$  and  $IRI_{Left}$  differ on average from each other. The results show that the difference in the means of  $IRI_{Right}$  and  $IRI_{Left}$  is statistically significant at  $\alpha$  0.05 probability of chance error. This implies that both samples are from different populations. Therefore, this study used  $IRI_{Right}$  as a dependent variable since it shows the highest value for pavement roughness.

#### 6.4.3.2. Input Variables

An extensive literature review was conducted to identify key input variables to explain

pavement deterioration over the years. Plots with each input variable and the output variable are recommended to assist the understanding of the relationship between them. Figure 31 shows plots of each input variable versus  $IRI_{Right}$ .

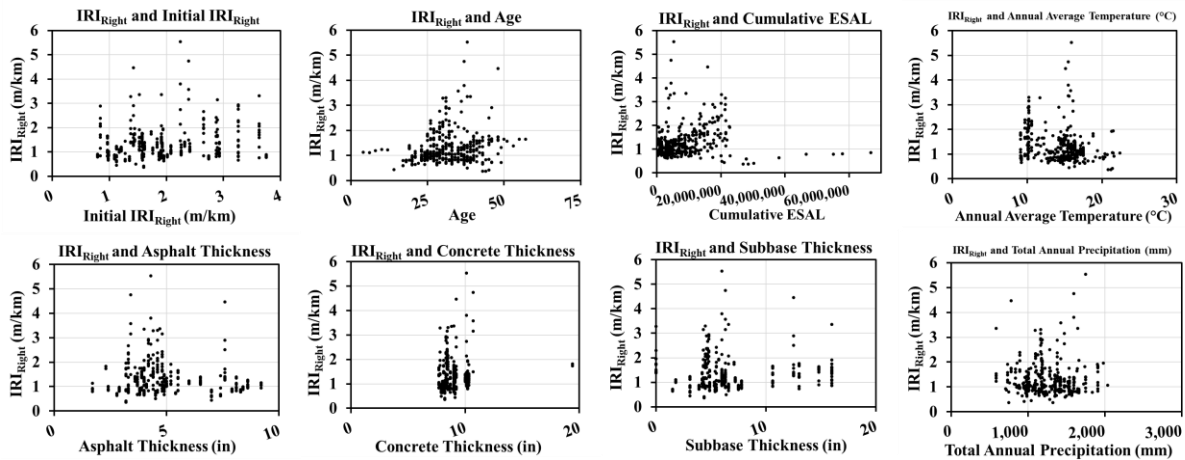


Figure 31. Independent Variables Versus  $IRI_{Right}$

Pavement structure, climate, traffic, and maintenance are known as key variables for pavement performance. In addition to these variables, this study used other input variables to assist the model's understanding of the pavement deterioration process to improve the accuracy of the predictions. Three variables used in this study are categorical variables with fixed values of 0 or 1 for modeling purposes and were not included in Figure 31.

A correlation analysis was performed by running a correlation test to obtain the Pearson correlation coefficient ( $r$ ) between all variables. The correlation matrix provides a linear association between the output and the proposed input variables. Therefore, if the data is not linearly correlated, it will have a low correlation value. Table 12 summarizes the descriptive statistics of input variables used in this study and the correlation between each input variable and the output variable ( $IRI_{Right}$ ).



Table 12. Descriptive Statistics of Input Variables

Input Variables	Mean	SD	COV (%)	Max.	Min.	r (IRI <sub>Right</sub> )
Initial IRI <sub>Right</sub> (m/km)	1.94	0.74	0.38	3.75	0.79	0.16
Age	32.24	8.23	0.26	57.00	4.00	0.17
CN <sub>Code</sub>	0.43	0.50	1.15	1.00	0.00	0.45
Granular Base/ Subbase	0.38	0.49	1.29	1.00	0.00	0.14
Treated Base/ Subbase	0.59	0.49	0.83	1.00	0.00	-0.19
CESAL	8800963.08	8344342.56	0.95	66776706.5	256280	0.13
Surface Asphalt Thickness (in)	4.99	1.75	0.35	9.20	1.70	-0.13
Concrete Thickness (in)	9.03	1.14	0.13	19.40	7.70	0.09
Subbase Thickness (in)	6.85	3.71	0.54	16.00	0.00	0.07
Annual Average Temperature (°C)	14.47	2.93	0.20	22.40	9.10	-0.24
Total Annual Precipitation (mm)	1252.52	289.40	0.23	2036.40	583.60	-0.03

Notes: Construction Number Code (CN<sub>Code</sub>); Standard Deviation (SD); Coefficient of Variation (COV); Pearson Correlation with IRI<sub>Right</sub> (r (IRI<sub>Right</sub>))

The three highest correlations between input variables and IRI<sub>Right</sub> were observed with the variable “CN<sub>Code</sub>” (0.45), where CN<sub>Code</sub> denotes the maintenance and rehabilitation history in the LTPP database, followed by a negative correlation with “Annual Average Temperature” (-0.24), and “Treated Base/ Subbase” (-0.19). Other variables show low linear correlation values, indicating they are not linearly correlated with the dependent variable. However, these variables might follow a non-linear correlation with IRI<sub>Right</sub> that cannot be identified by the correlational analysis. Hence, all independent variables were used in this study for the development of ANN and MLR models.

### 6.5. ANN Development

ANN is a predictive modeling technique based on mathematical operations that use the concept of human cognition and neural biology [43]. The ANNs approach attempts to emulate the structure and/or functional aspects of biological neural networks [44]. Complex relationships that are difficult to be identified using traditional sequential, logic-based modeling and computational

techniques can be successfully represented by neural networks [43]. There are many types of neural networks characterized by their architecture, training algorithm, and activation functions. In this study, a feed-forward neural network with a back-propagation training algorithm was used for the development of the roughness prediction model. The neural network gains its knowledge through a trained feed-forward network that uses a set of training data consisting of inputs (independent variables) and output(s) (dependent variable(s)). The resulting output is compared to the target values, and the back-propagation process adjusts the connection weight to reduce the error between actual and target values [2]. After training, the network provides an approximate functional mapping of any input pattern onto its corresponding output pattern. One hidden layer was considered in the model development. The use of more than one hidden layer combined with an insufficient number of databases may cause the network to memorize the data in the training phase [44]. The TR-SEQ1 computer program [55] was used to develop the ANN models in this study. A sigmoidal function is used for data generalization purposes.

#### 6.5.1. ANN Methodology

Four successive stages were used for ANN model development and the desired criteria to choose the optimal network structures [51], as follows:

- Stage 1: Determine the ANN architecture. Decide input and output variables. Classify the datasets as training (50%), testing (25%), and validation (25%) sets.
- Stage 2: Train and test the network on the experimental data to obtain the optimum number of hidden nodes and iterations for the ANN architecture.
- Stage 3: Validate the best-performing network from the second stage using the validation database. Check if the accuracy results from the training, testing, and validation database

are comparable. If they are, then stage four may not be necessary.

- Stage 4: Retrain the best performing network from stage 2 using all experimental data to increase prediction accuracy and account for all patterns in the database.

Typically, retraining the selected final network with all experimental data is expected to provide reliable predictions and overall better accuracy measures since all the knowledge in the database are incorporated in the final network [43,51].

### 6.5.2. ANN Model Network Selection

Three statistical accuracy measures were used to select the best-performing network for the model. The Coefficient of Determination ( $R^2$ ), the Mean Absolute Relative Error (MARE), and the Average Square Error (ASE). Each ANN model development stage (i.e., training, testing, validation, and all-data) was considered, and the best performing network was selected based on the lowest ASE, lowest mean MARE, and highest  $R^2$ . Table 13 shows the summary of ANN model network statistics.

Table 13. Summary of ANN Model Network Statistics

Networks		1-19-20000	2-5-20000	3-19-1200	<b>13-19-2000</b>	16-19-7100
Training	MARE	7.29	15.27	10.72	<b>9.98</b>	9.60
	$R^2$	0.97	0.87	0.96	<b>0.94</b>	0.94
	ASE	0.00029	0.00146	0.00058	<b>0.00072</b>	0.00070
Testing	MARE	30.92	25.38	21.13	<b>20.71</b>	20.69
	$R^2$	0.44	0.60	0.62	<b>0.72</b>	0.67
	ASE	0.01476	0.00522	0.00713	<b>0.00425</b>	0.00533
Validation	MARE	30.68	25.32	21.32	<b>20.24</b>	20.45
	$R^2$	0.43	0.59	0.62	<b>0.71</b>	0.67
	ASE	0.01492	0.00528	0.00721	<b>0.00429</b>	0.00538
All Data	MARE	14.48	18.36	13.98	<b>13.14</b>	12.94
	$R^2$	0.70	0.77	0.82	<b>0.86</b>	0.83
	ASE	0.00479	0.00264	0.00262	<b>0.00182</b>	0.00214
Chosen Network		<b>13-19-2000</b>				
Final Model Structure		<b>11-19-1 (inputs-hidden nodes-output)</b>				

The networks analyzed can be identified in the first row of Table 13 in an order that depicts the initial hidden node, final hidden node, and the number of iterations. The final structure of each model is shown in the bottom row of the table in an order that represents the number of inputs, number of the final hidden node, and number of output(s), respectively.

Following the ANN methodology, after developing the possible networks, the all-data stage used all datasets from the database to retrain the network at its optimal structure and iteration to obtain the generalized response throughout the complete database. Therefore, the all-data stage statistics are used as the main comparison between each network. The 13-19-2000 network outperformed all other networks in the all-data stage with lower ASE (0.00182) and higher  $R^2$  (0.86). Thus, the 13-19-2000 was chosen as the best-performing network. The final model structure includes 11 input variables, one hidden layer with 19 hidden nodes, 2,000 iterations, and one output. Figure 32 shows the architecture of the best-performing network for the ANN model.

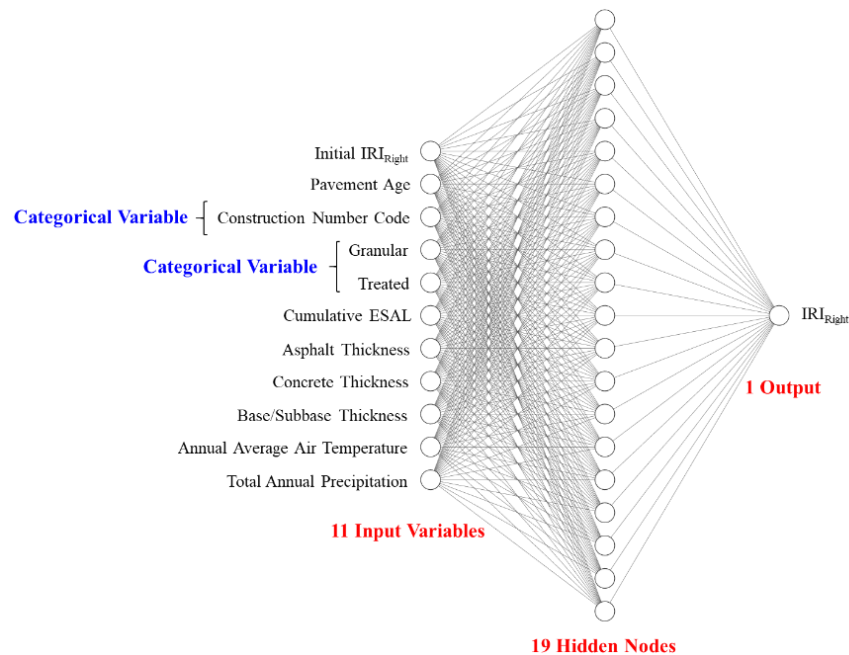


Figure 32. Architecture of the Best-Performing Network for the ANN Model

## 6.6. Regression Model

A multiple linear regression analysis was conducted to develop the roughness prediction model for composite pavements in the LTPP wet non-freeze region. The MLR model was developed using the same input and output variables used in the ANN model. The input variables included initial  $IRI_{Right}$ , age,  $CN_{Code}$ , granular base/subbase, treated base/subbase, CESAL, surface asphalt thickness, concrete thickness, subbase thickness, annual average temperature, and total annual precipitation. The  $IRI_{Right}$  was used as the output variable.

### 6.6.1. Multiple Regression Model Development

To develop a multiple linear regression, the data should follow some assumptions, as shown:

- The sample should be random.
- The output variable should be normally distributed.
- Observations should be independent.
- The relationship between the output variable and each input variable should be linear.
- The correlation between each input variable should not be more than 0.8.
- Error prediction should be normally distributed.

The output variables need to be checked for model assumptions before developing any multiple regression models. Depending on the relation and pattern among output and input variables, data transformations might be needed for dependent and/or independent variables. In this study, the output variable ( $IRI_{Right}$ ) was normally distributed, standard residuals were normally distributed, and scatterplots did not show curvilinear relationships between  $IRI_{Right}$  and independent variables. A statistically significant relationship between input variables and the

output variable was found,  $F(11, 341) = 17.91$ ,  $p < 0.001$ . The MLR model accounted for 37% of the variance in the model.

## 6.7. Discussion

### 6.7.1. ANN Model Results

Figure 33 shows a scatter plot of observed vs. predicted  $IRI_{Right}$  values for the ANN model.

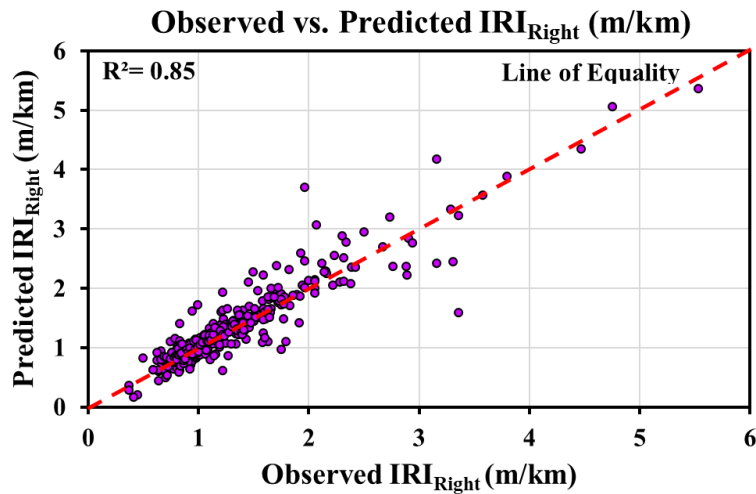


Figure 33. Observed vs. predicted  $IRI_{Right}$  for ANN model

The ANN model was able to capture the roughness changes and pavement deterioration behavior; however, when the observed IRI values were greater than 3 m/km, the model was not as accurate as for lower IRI values. Nevertheless, the ANN model resulted in a high  $R^2$  of 0.85.

### 6.7.2. Regression Model Results

The developed MLR model equation is given in Equation 21, as follows:

$$\begin{aligned}
 IRI_{Right} = & 0.572 + 0.093(Initial\ IRI_{Right}) + 0.006(Age) + 0.679(CN_{Code}) - \\
 & 0.712(Granular_{Base/Subbase}) - 0.771(Treated_{Base/Subbase}) - 6.85 \times 10^{-9}(CESAL) - \\
 & 0.022(Asphalt_{Thickness}) + 0.171(Concrete_{Thickness}) - 0.007(Subbase_{Thickness}) - \\
 & 0.046(Annual\ Average\ Temperature) + 0.00011(Total\ Annual\ Precipitation) \quad Eq. 21
 \end{aligned}$$

Figure 34 shows the graphical comparison of observed and predicted values accuracy for the MLR model.

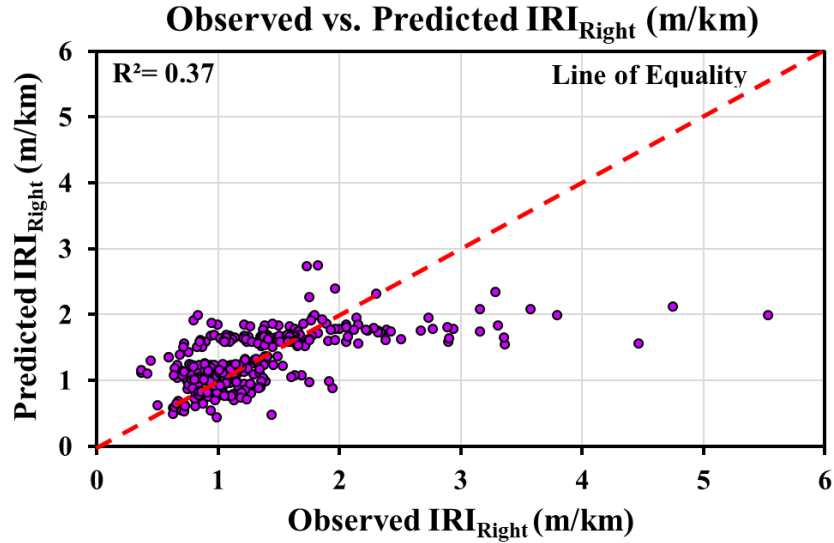


Figure 34. Observed vs. predicted  $IRI_{Right}$  for MLR model

The MLR model prediction is scattered and does not follow the line of equality, which shows that the MLR model was not able to capture most of the roughness changes and pavement deterioration behavior.

The model predicted lower values compared to the observed  $IRI_{Right}$ . The MLR model resulted in a low  $R^2$  of 0.37. The MLR model was verified using three sections that were not used in the model development. Figure 35 shows the observed vs. predicted plot for the verification sections.

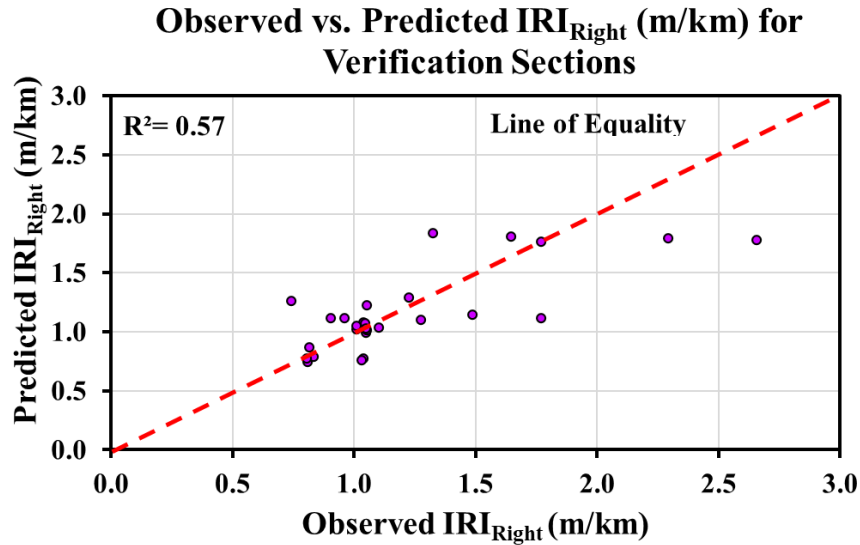


Figure 35. Observed vs. Predicted  $IRI_{Right}$  Values for Verification Sections

The MLR model showed better results for the verification sections compared to the full database. A total of 27 data points were used for the verification. The MLR model accounted for 57% of the variability in the model, compared to 37% when using the full database. Therefore, the verification shows that the MLR model performs reasonably when exposed to sections outside the original database used for the model development.

#### 6.8. Comparison Between ANN and MLR Model Results

The prediction accuracy measures for the ANN model and the MLR model are graphically depicted in Figure 33 and Figure 34, respectively. It is very clear from the predicted versus observed plots that the ANN model outperformed the MLR model with more accurate predictions. The ANN model had a MARE (13.14) 53% lower and an ASE (0.00182) 99% lower than the MLR model. The  $R^2$  value has improved from 0.37 obtained by the MLR model to 0.86 obtained by the ANN model. This translates into a 132% improvement in accuracy using the ANN-based model. It is possible to increase the accuracy of the regression model by using non-linear regression;



however, several studies show that even the best non-linear regression models will not deliver better accuracy measures than those obtained via an appropriately developed ANN-based model [52].

A comparison between the ANN and MLR model predictions for two random sections of the database was performed to identify the most accurate model approach and evaluate the performance of the developed models. Figure 36 shows the observed and predicted values for both the ANN and MLR models of section 05-A608 in Arkansas.

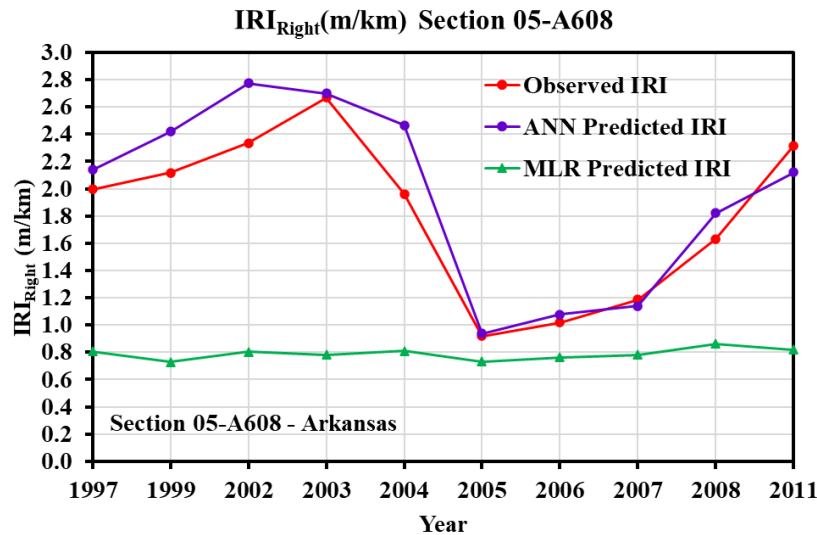


Figure 36. IRI<sub>Right</sub> Section 05-A608 - ANN and MLR Comparison

Predicted values were close to observed values only for the ANN model. The MLR model underpredicted the IRI values for all years, following almost a straight line with not much variation. The mean difference percentage with observed values was much lower for the ANN model (8%) compared to the MLR model (-57%). Therefore, the ANN model proved to be better in capturing the pavement deterioration behavior over time. Figure 37 shows the observed and predicted values for ANN and MLR models of section 41-7018 in Oregon.

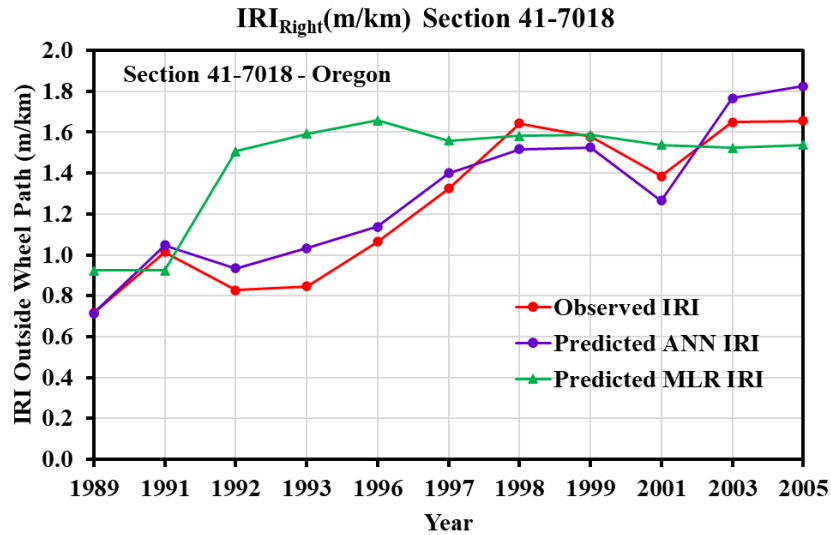


Figure 37. IRI<sub>Right</sub> Section 41-7018 - ANN and MLR Comparison

For section 41-7018, it is possible to observe that the MLR prediction did not follow the pattern of observed values. Predicted values were close to observed values only for the ANN model. The mean difference percentage with observed values was a lot lower for the ANN model (3%) compared to the MLR model (16%). The ANN model was able to capture the pavement deterioration behavior and IRI variations over time, closely following the observed values. Therefore, the ANN model showed to be more accurate than the MLR model and, hence, was chosen as the best-performance model.

## 6.9. Conclusions

In this study, multiple linear regression and artificial neural network approaches were used to develop roughness prediction models for composite pavements in the wet non-freeze climate region from the LTPP database. Both models used 11 input variables (i.e., initial IRI<sub>Right</sub>, age, CN<sub>Code</sub>, granular Base/ Subbase, treated Base/ Subbase, CESAL, surface asphalt thickness, concrete thickness, subbase thickness, annual average temperature, and total annual precipitation) and 1

output variable ( $IRI_{Right}$ ). The best-performing network for the ANN model was selected based on three accuracy measures shown in Table 13. The 13-19-2000 network outperformed all other networks in the all-data stage with lower ASE (0.00182) and higher  $R^2$  (0.86). The final ANN model includes 11 input variables, one hidden layer with 19 hidden nodes, 2,000 iterations, and one output. The developed MLR model accounted for only 37% of the variability. Predictions for the MLR model were scattered and did not follow the line of equality. However, verification sections for the MLR showed better results with an  $R^2$  of 0.57 compared to the 0.37 when using the complete database.

Figure 33 and Figure 34 show graphically the prediction accuracy measures for the ANN and MLR models. The results showed that the ANN model outperformed the MLR model with 53% lower MARE, 99% lower ASE, and 132% better  $R^2$  prediction accuracy. The developed ANN model was able to capture the roughness changes and pavement deterioration behavior over time. Predicted values clustered around the line of equality, showing good accuracy. However, when observed values were higher than 3 m/km, the model was not as accurate as for lower IRI values. It might be interesting to develop different models for lower and higher IRI values. A comparison between the ANN and MLR model predictions was performed for two random sections of the database to identify the most accurate model approach and evaluate the performance of the developed models. The ANN model performed better than the MLR for both sections, being able to capture the pavement deterioration behavior and follow closely the IRI variations over time.

Therefore, the ANN model showed to be more accurate than the MLR model and, hence, chosen as the best-performance model. The complex relationship between pavement structure, climate, and traffic variables is better explained when using the ANN model. The developed ANN

model can contribute to the state-of-the-art by providing a more inclusive, reliable, and user-friendly tool to support objective decisions regarding maintenance and rehabilitation interventions and budget plans allowing agencies to prioritize the resources for critical pavement sections. Additionally, the use of a specific climate region helped the model to capture 86% of the variability, which may not be viable when using data from all climate zones together. Moreover, the models developed in this study did not use any distress data for input variables, which can save time and money from data collection and processing for transportation agencies.

#### 6.10. Acknowledgments

The research described in this paper was conducted at the University of Mississippi. The contents of this paper reflect the views of the authors who are responsible for the facts, findings, and data presented herein. The authors would like to acknowledge Dr. Waheed Uddin for his significant contribution to this research with his guidance and recommendations.

#### 6.11. Author Contributions

The authors confirm contribution to the paper as follows: study conception and design: R. Barros, H. Yasarer, and S. Sultana; data collection: R. Barros; analysis and interpretation of results: R. Barros, H. Yasarer; draft manuscript preparation: R. Barros and H. Yasarer. All authors reviewed the results and approved the final version of the manuscript.

## CHAPTER VII: COMPOSITE PAVEMENT ROUGHNESS MODELING FOR LTPP WET FREEZE CLIMATE REGION USING MACHINE LEARNING

R. Barros, H. Yasarer, W. Uddin, S. Sultana & Y. Najjar

Department of Civil Engineering, University of Mississippi, Oxford, Mississippi, United States

(Paper accepted for session presentation and published at the Eleventh International Conference on the Bearing Capacity of Roads, Railways and Airfields (BCRRA), Trondheim, Norway, 27-30 June 2022)

### 7.1. Abstract

Objective decisions related to the management of road networks are based on important measures of pavement performance. The International Roughness Index (IRI) is a critical indicator of pavement performance, and it is considered the standard for pavement roughness. A reliable pavement performance prediction model is needed to predict future pavement conditions and identify maintenance and rehabilitation (M&R) needs. This study intends to develop pavement roughness models using the Artificial Neural Networks (ANNs) approach for composite (asphalt overlay on concrete) pavements using the Long-Term Performance Pavement (LTPP) program database for the Wet-Freeze climate region. A total of 186 pavement sections with 1,930 data points were analyzed. Five models were developed using different independent variables (i.e.,

Initial IRI<sub>Right</sub>, Age, Seasons, Asphalt Thickness, Concrete Thickness, Subbase Thickness, Subbase Type, Construction Number (CN), Cumulative Equivalent Single Axle Load (CESAL), Air Temperature, Freeze Index, Freeze-Thaw, and Precipitation) and one dependent variable (i.e., IRI<sub>Right</sub>). The best-performing model was selected based on the lowest average square error (ASE), lowest mean absolute relative error (MARE), and highest coefficient of determination (R<sup>2</sup>). Results showed that the developed models had satisfactory results with a good fit of observed and predicted data. Therefore, local and state agencies can use the developed ANN roughness models as a tool for better condition assessment and effective M&R scheduling. Furthermore, the use of available climatological and historical traffic data to predict IRI changes will also eliminate time-consuming data collection and processing, accordingly, reducing costs.

Key words: International Roughness Index (IRI), Artificial Neural Network (ANN), Long-Term Pavement Performance (LTPP), composite pavement, wet freeze.

## 7.2. Introduction

A large portion of the paved highways in the U.S. comprises composite pavements. This type of pavement is commonly a result of concrete pavement rehabilitation, where an asphalt layer is overlaid on a concrete surface [3]. In the U.S., pavement performance models are required for state highway agencies to assist in their pavement management decision-making processes [10]. Performance models bring key features to a successful pavement management system (PMS) [82] providing an estimation of pavement conditions and rehabilitation needs and allowing agencies to prioritize road sections that are in the worst conditions. Performance models are easy to understand and can provide deeper insights converting performance indices into operational measures to inform how long and how well the road will continue to serve the users [10]. Numerous pavement

performance indices have been developed in the last three decades; however, the international roughness index (IRI) is the most well-recognized performance index [12,20]. The IRI expresses the irregularities in the pavement surfaces that affect the ride quality, and it is useful for making objective decisions related to the management of road networks [2,8].

Pavement performance modeling is a challenging task due to the complexity of the pavement structure and its responses under traffic loading, dynamic weather and climate changes, variability in construction activities, and the interaction among all these elements [11]. Advanced modeling techniques such as artificial neural networks (ANNs) have been used successfully in several studies offering significant improvements over traditional techniques (i.e., linear regression) by processing large volumes of data with excellent accuracy. However, several performance models in the literature use the Long-Term Performance Pavement (LTPP) database to develop models without considering specific conditions of local climate and geography, which makes the model less accurate.

Therefore, this paper proposes the development of pavement roughness prediction models to study the performance of composite pavements using the ANN approach according to the specific climate and geographical conditions (wet freeze climate zone), pavement structure, and traffic data using the LTPP database.

#### 7.2.1. Objectives

The major objectives of this paper are to:

- (1) Analyze roughness data for composite pavement sections in the wet freeze region using the LTPP database.
- (2) Use the ANN approach to develop roughness models for composite pavements in the wet

freeze region of the LTPP database using different independent variables.

- (3) Evaluate the accuracy of the ANN models using statistical measurements to identify the most accurate model.
- (4) Perform a model comparison among the developed models using two random sections of the database to identify the best-performing model.

### 7.3. Literature Review

The literature review shows that a remarkable number of researchers have used ANN to predict pavement roughness. Duckworth [28] and Yasarer et al. [65] developed pavement performance prediction models for flexible and continuously reinforced concrete pavement (CRCP) sections, respectively, using the ANNs approach based on the Mississippi Department of Transportation (MDOT) database. Both papers concluded the ANN model successfully predicted roughness values and could be used for pavement performance prediction.

Hossain et al. [16] and Hossain et al. [17] also developed ANN roughness prediction models for flexible and concrete pavements, respectively, but using climate and traffic data collected from the LTPP database. The studies used data from the wet freeze climate zone for flexible pavements and wet non-freeze for rigid pavements. An RMSE of 0.027 and 0.01 were found for the flexible and rigid ANN models, respectively, indicating that the IRI prediction was reasonable for both short-term and long-term predictions.

Jaafar [2] and Jaafar [19] developed IRI prediction models for asphalt pavements using multiple linear regression and ANN modeling approaches using all the LTPP database and the Western region of the LTPP database, respectively. The ANN model showed a high coefficient of correlation (R) of 0.72 and 0.85, respectively, showing promising results for IRI predictions in



asphalt pavements. For flexible and composite pavements, Kaya et al. [10] used the Iowa database to develop ANN-based models that were found to be good tools for predicting pavement deterioration when there were many pavement sections with various traffic, thickness, and other various deterioration trends.

The literature shows ANN models successfully predicted IRI values for asphalt, concrete, and composite pavements. However, composite pavements have not been well investigated, especially the effects of traffic and climate variables for a specific climate region of the LTPP database. Since each location has its local climate and specific characteristics, it is necessary to use data from an individual region to assist the performance models to be more accurate. Hence, this paper develops a pavement roughness prediction model using the ANNs approach for composite pavement sections using data from the wet freeze climate zone of the LTPP database.

## 7.4. Model Development

### 7.4.1. Data Collection

The LTPP program created in 1987 to collect and store performance data over several years was used as the database for the data analysis in this study. The LTPP program was developed to support analysis and product development, analyze the collected data to describe pavement performance, and translate these insights into usable engineering products related to pavement design, construction, rehabilitation, maintenance, preservation, and management. The data collection started in 1989, and 2,509 pavement test sections from 51 U.S. states and ten Canadian provinces were selected or constructed for the program. The LTPP developed a climate zone classification that included four different zones identified as wet freeze, wet non-freeze, dry-freeze, and dry non-freeze zones [2,84]. The LTPP used the precipitation per year to identify wet (higher

than 508 mm) and dry (lower than 508mm) climate zones. For freeze and non-freeze climate zones, the LTPP used a threshold based on an annual average freezing index of 83 °C (150 °F) days. Locations with an index over this threshold were classified in the freeze zone and those under the threshold in a non-freeze climate zone [84].

A total of 264 composite sections where the asphalt layer thickness over the concrete layer is equal to or greater than three inches were found using the LTPP database. For this study, only composite pavement sections located in the wet freeze climate zone were used. A total of 186 pavement sections with 1,930 data points were used for the analysis.

#### 7.4.2. Data Processing

To construct the database used in this study, all the variables were retrieved from the LTPP database. The data processing for output and input variables is described in this section.

##### 7.4.2.1. Output Variables

The IRI is accepted as one of the most important indicators of pavement performance and used as the standard for pavement roughness, and it was used as the output variable for modeling. Each section in the database had two types of IRI measurements: IRI inside wheel path ( $IRI_{Left}$ ) and IRI outside wheel path ( $IRI_{Right}$ ). Several IRI measurements were done on each visit date for each section. By averaging the IRI measurement runs, a single IRI measurement was obtained for  $IRI_{Left}$  and  $IRI_{Right}$  for each visit date. By doing this, a total of 1,930 IRI measurements from 1989 to 2018 were found for the 186 sections.

IRI measurements for  $IRI_{Right}$  (1.27 m/km or 80.5 in./mile) were 8% greater than the  $IRI_{Left}$  (1.18 m/km or 74.8 in./mile), which agreed with previous research [2,31,36,88,89]. An independent samples t-test was used to determine whether  $IRI_{Right}$  and  $IRI_{Left}$  differ on average from each other. The t-test result shows that the difference in the means of  $IRI_{Right}$  and  $IRI_{Left}$  is

statistically significant at  $\alpha$  0.05 probability of chance error implying that both samples are from different populations. Therefore, this study used  $IRI_{Right}$  as a dependent variable since it shows the highest value for pavement roughness.

#### 7.4.2.2. Input Variables

For the input variables, the literature shows that pavement structure, climate, traffic, and maintenance are key variables to account for pavement deterioration. This research used some variables that were already studied in the literature but introduced new variables that were found to help the model to achieve better results. An important step of modeling is to identify how the input variables relate to the output variable. Figure 38 shows plots of each input variable versus  $IRI_{Right}$ . Two variables used in this study are categorical variables ( $CN_{Code}$ , subbase type), which have fixed values of 0 or 1 for modeling purposes and were not included in Figure 38.

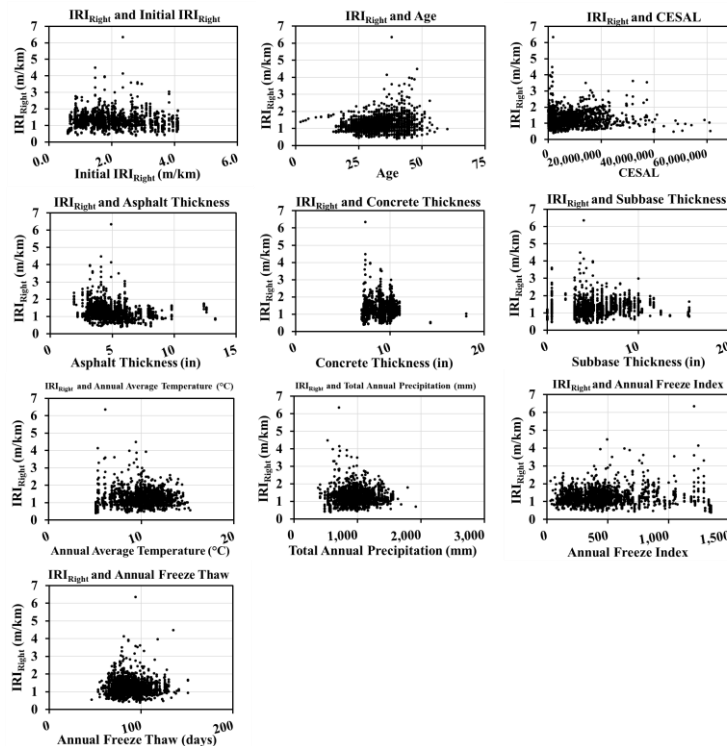


Figure 38. Independent Variables Versus  $IRI_{Right}$

Table 14 shows a summary of the descriptive statistics of input variables and the correlation between each input variable and the output variable ( $IRI_{Right}$ ). A correlation test was performed to obtain the Pearson correlation coefficient ( $r$ ) among all variables. The correlation analysis provides the linear association between the output and the proposed input variables. Hence, if the data analyzed do not have a good linear correlation, the coefficient will show low values.

Table 14. Descriptive Statistics of Input Variables

Input Variable	Mean	Standard Deviation	COV (%)	Max.	Min.	$r$ ( $IRI_{Right}$ )
$IRI_0$ , Initial IRI Right (m/km)	1.90	0.83	43.7%	4.1	0.60	0.058
Age	32.63	7.77	23.8%	60.0	2.00	0.199
CESAL	7649125	8113176	1.06	6127620 1	35566.0 0	0.059
Surface Asphalt Thickness (in)	5.13	1.88	36.6%	13.3	1.90	-0.214
Concrete Thickness (in)	9.25	1.05	11.4%	18.1	7.00	-0.079
Subbase Thickness (in)	5.18	2.67	51.5%	15.6	0.00	0.058
Annual Average Temperature ( $^{\circ}C$ )	10.07	1.93	19.2%	15.3	5.10	-0.008
Total Annual Precipitation (mm)	966.58	218.80	22.6%	1919.5	377.40	-0.015
Annual Freeze Index	458.10	270.40	59.0%	1345.0	28.00	0.067
Annual Freeze-Thaw (days)	87.33	15.00	17.2%	151.0	46.00	0.006
$IRI_{Right}$ (m/km)	1.25	0.45	35.6%	6.4	0.39	1.000

Notes: Coefficient of Variation (COV); Pearson Correlation with  $IRI_{Right}$  ( $r$  ( $IRI_{Right}$ ))

The three highest correlations were observed with the variable “ $CN_{Code}$ ” (0.21) followed by a negative correlation with “surface asphalt thickness” (-0.21), and “age” (0.20). Other variables indicate low linear correlation values, implying they are not linearly correlated with  $IRI_{Right}$ . However, these variables might follow a non-linear correlation that cannot be identified by the correlational analysis. Therefore, even though the correlation was low in some cases, all independent variables were used for the development of ANN models.

### 7.5. Ann Development

ANNs are a predictive modeling technique based on mathematical models built to simulate the neural structure of a human brain using the concept of human cognition and neural biology

[43]. The neural networks can successfully represent complex relationships that are difficult to be identified using traditional sequential, logic-based modeling, and computational techniques [43]. A feedforward neural network with a back-propagation training algorithm was used for the development of the roughness prediction model in this study. A one hidden layer network was considered in the model development since the use of more than one hidden layer may cause the network to memorize the data in the training phase [44]. A sigmoidal function was used for data generalization purposes and the TR-SEQ1 computer program [55] was used to develop the ANN models.

#### 7.5.1. ANN Model Architecture

For this paper, five models were developed using different independent variables and  $IRI_{Right}$  as the dependent variable. Table 15 shows the independent variables used for each ANN model.

Table 15. Input Variables Configuration for Five ANN Models

No.	Model 1	Model 2	Model 3	Model 4	Model 5
1	$IRI_0$	$IRI_0$	$IRI_0$	$IRI_0$	$IRI_0$
2	Age	Age	Age	Age	Age
3	-	-	Winter (Dec-Feb)		
	-	-	Spring (Mar-May)		
	-	-	Summer (June-Aug)		
	-	-	Autumn (Sept-Nov)		
4	Granular Subbase	Granular Subbase	Granular Subbase	Granular Subbase	Granular Subbase
5	Treated Subbase	Treated Subbase	Treated Subbase	Treated Subbase	Treated Subbase
6	$CN_{Code}$	$CN_{Code}$	$CN_{Code}$	$CN_{Code}$	$CN_{Code}$
7	-	CESAL	CESAL	CESAL	CESAL
8	$h_{asphalt}$	$h_{asphalt}$	$h_{asphalt}$	$h_{asphalt}$	$h_{asphalt}$
9	$h_{concrete}$	$h_{concrete}$	$h_{concrete}$	$h_{concrete}$	$h_{concrete}$
10	$h_{subbase}$	$h_{subbase}$	$h_{subbase}$	$h_{subbase}$	$h_{subbase}$
11	-	-	-	Air Temp.	Air Temp.
12	-	-	-	Precipitation	Precipitation
13	-	-	-	-	Freeze Index
14	-	-	-	-	Freeze-Thaw

Model 1 used 8 input variables; however, the “granular subbase,” “treated seasons,” and “CN<sub>Code</sub>” are categorical variables. The “CN<sub>Code</sub>” is a categorical variable that represents M&R interventions in the pavement structure. A value of 1 is assigned when an M&R intervention was performed. If there was no M&R intervention, a 0 value was assigned. Model 2 used 9 input variables, including an input variable responsible for traffic loads, CESAL. Model 3 used 13 input variables including four more variables that represent the season in which each test was performed in the section. It consists of four seasons spring (March, April, May), summer (June, July, August), autumn (September, October, November), and winter (December, January, February), and represents the effect of climate and season on pavement deterioration. Model 4 used 11 input variables including “air temperature” and “precipitation” to account for the climate effects on the pavement. Model 5 used 13 input variables and included “freeze index” and “freeze-thaw” to account for the effect of lower temperatures on pavement performance. All variables used in this study are not related to distress data and are easily available for federal and state transportation agencies.

#### 7.5.2. ANN Model Selection

The best ANN model was chosen based on the lowest average square error (ASE), lowest mean absolute relative error (MARE), and highest coefficient of determination ( $R^2$ ). Table 16 demonstrates the statistical measures of each ANN model development stage (i.e., training, testing, validation, and all-data) for all developed models.

Training, testing, and validation stages used 972, 479, and 479 data points, respectively. After developing the possible networks for each model, the all-data stage used all 1,930 data points from the database to retrain the network at its optimal structure and iteration. Therefore, the all-

data stage statistics were used as the main comparison between each model. The all-data stage for model 5 outperformed the other models with more accurate measures. Hence, model 5 was chosen as the best-performing ANN model. The final model structure includes 13 input variables, one hidden layer with 19 hidden nodes, 20,000 iterations, and one output.

Table 16. ANN Model Results

Model		1	2	3	4	5
Structure		4-19-20000	6-18-20000	4-19-20000	7-19-20000	<b>8-19-20000</b>
Training	MARE	20.09	20.19	17.39	17.03	<b>18.11</b>
	R <sup>2</sup>	0.58	0.58	0.69	0.69	<b>0.71</b>
	ASE	0.00097	0.0010	0.0007	0.00068	<b>0.00070</b>
Testing	MARE	23.81	19.85	22.73	20.22	<b>20.84</b>
	R <sup>2</sup>	0.34	0.51	0.32	0.53	<b>0.44</b>
	ASE	0.00140	0.0011	0.0016	0.00099	<b>0.00136</b>
Validation	MARE	26.65	20.90	25.00	19.21	<b>22.29</b>
	R <sup>2</sup>	0.27	0.45	0.22	0.50	<b>0.37</b>
	ASE	0.00195	0.0012	0.0023	0.00099	<b>0.00157</b>
All Data	MARE	17.94	17.94	17.29	17.14	<b>15.77</b>
	R <sup>2</sup>	0.57	0.59	0.62	0.68	<b>0.71</b>
	ASE	0.00087	0.00082	0.00076	0.00066	<b>0.00059</b>

## 7.6. Discussion

Figure 39 shows a graphical comparison of the accuracy measurements for the all-data stage among all developed models.

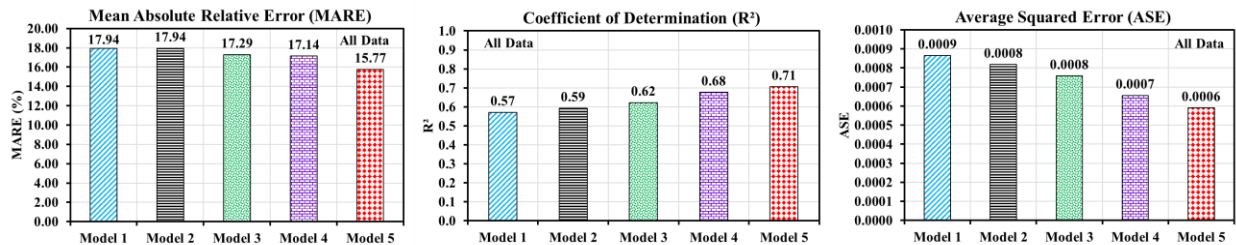


Figure 39. Graphical Comparison of Accuracy Measures

The accuracy measures show reliable results for all models developed. However, model 5

outperformed all other models developed in this study. Compared to models 1, 2, 3, and 4, the all-data stage of model 5 showed an ASE 32%, 28%, 22%, and 10% lower, a MARE 12%, 12%, 9%, and 8% lower, and an  $R^2$  24%, 19%, 13%, and 4% higher, respectively. Figure 40 illustrates the observed  $IRI_{Right}$  values collected from the LTPP database and the predicted values using model 5. The ANN model could imitate the pavement deterioration behavior. However, some higher IRI values were not captured by the model, which predicted lower values compared to the observed IRI. Predicted values tried to cluster around the line of equality but some values were underpredicted, especially when the observed  $IRI_{Right}$  values were higher than 3 m/km; however, all values were used in the model. Nevertheless, a good  $R^2$  of 0.71 was obtained for the ANN model 5.

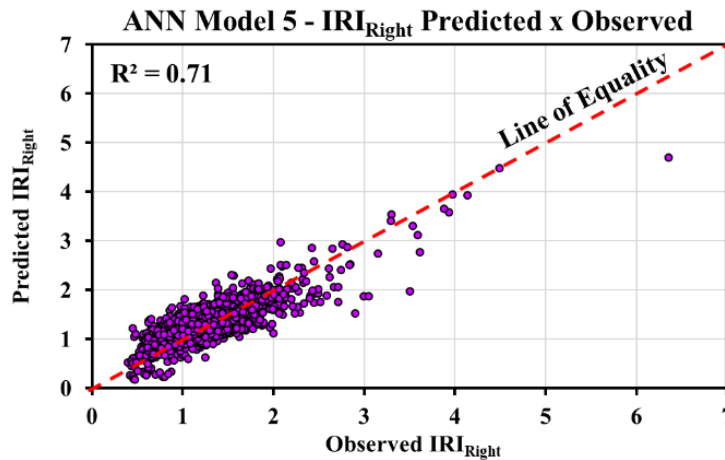


Figure 40. Observed Versus Predicted  $IRI_{Right}$  for ANN Model 5

### 7.7. Model comparison

Model comparison was conducted to evaluate the performance of the developed models in two random sections of the database. IRI prediction values were generated using all models and compared with observed values. Figure 41 shows the observed versus predicted plots of  $IRI_{Right}$  for section 29-0661 in Missouri and section 46-7049 in South Dakota.



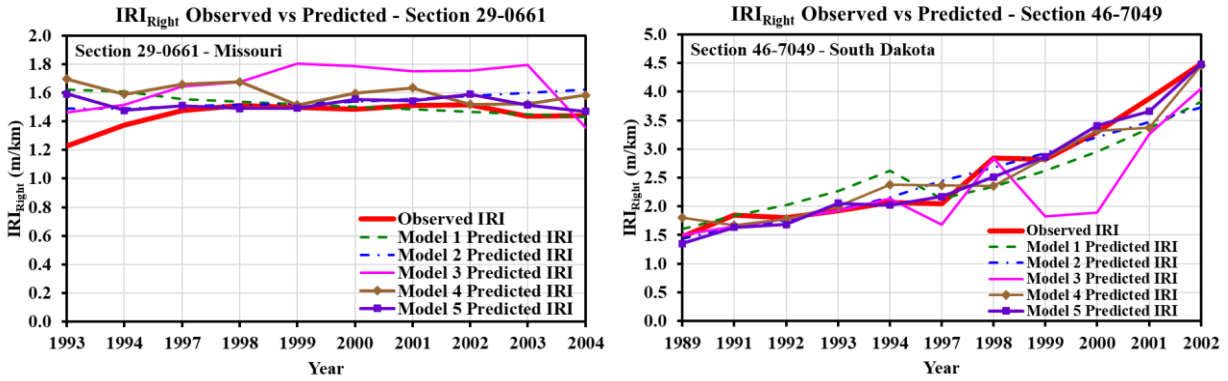


Figure 41. Observed and Predicted Plots of  $IRI_{Right}$  for Sections 29-0661 and 46-7049

Predicted values were close to observed values for almost all models in both sections. Model 3 was the only one to not closely follow the observed  $IRI_{Right}$  values. This might happen because model 3 was the only model to use the input variable “season,” which could cause the model to predict less accurate values. Model 5 showed the most accurate results for both sections, being the best to explain the variability in the model and capture the pavement deterioration behavior. The use of variables that represents the pavement initial condition ( $IRI_0$ ), effects of pavement exposure time (age), pavement structure (thickness of asphalt, concrete, subbase, and type of subbase), the effect of maintenance and rehabilitation ( $CN_{Code}$ ), the effect of traffic loads (CESAL), and climatological effects of temperature, moisture, and freeze (air temperature, precipitation, freeze index, freeze-thaw) in the same model might be the reason why model 5 had a better performance than other models that did not use all variables at the same time.

Therefore, the model comparison showed that all developed models were reasonably accurate and reliable for both analyzed sections and can be used to predict pavement roughness. Furthermore, the analysis showed that by combining all input variables in the same model, the predictions were more accurate since the model had a better understanding of the pavement deterioration behavior and translated this into more precise  $IRI_{Right}$  values. Therefore, the

developed models in this study can be used as a powerful tool to predict pavement's future condition and better M&R planning.

## 7.8. Conclusions

The concluding remarks are summarized, as follows:

- (1) The pavement roughness models for composite pavements were developed using the artificial neural networks modeling technique. The models used data collected from the LTPP database for the wet freeze climate zone.
- (2) The best-performing ANN model was selected based on the accuracy measures presented in Table 16. Model 5 outperformed all other models accounting for 71% of the variability with an  $R^2$  of 0.71 with 13 input variables, one hidden layer with 19 hidden nodes, 20,000 iterations, and one output.
- (3) Compared to models 1, 2, 3, and 4, the all-data stage of model 5 showed an ASE 32%, 28%, 22%, and 10% lower, a MARE 12%, 12%, 9%, and 8% lower, and an  $R^2$  24%, 19%, 13%, and 4% higher, respectively.
- (4) The ANN model could replicate the pavement deterioration behavior with reasonable accuracy. Predicted values cluster around the line of equality but some values were underpredicted, especially when the observed  $IRI_{Right}$  values were higher than 3 m/km.
- (5) Model comparison was performed for sections 29-0661 and 46-7049. Predicted values were close to observed values for almost all models in both sections. Model 5 showed to be the most accurate.
- (6) The study showed that the use of a specific climate zone combined with input variables

that represent the pavement's initial condition, effects of pavement exposure time, pavement structure, effects of maintenance and rehabilitation, effects of traffic loads, and climatological effects of temperature, moisture, and freeze in the same model increased the model accuracy.

Therefore, the developed ANN model efficiently characterized the pavement roughness behavior on composite pavements and can be used by agencies as a prediction tool for IRI values and guide decision-makers to develop a better M&R plan. Furthermore, the developed model can predict IRI values without using distress data, which will result in cost reductions and more effective M&R scheduling

CHAPTER VIII: PAVEMENT PERFORMANCE MODELING FOR COMPOSITE  
PAVEMENTS IN THE LTPP WET FREEZE CLIMATE REGION INCORPORATING  
MAINTENANCE AND REHABILITATION

Barros, R., Yasarer, H., Sultana, S.

rferrei1@go.olemiss.edu, University of Mississippi, United States

(Paper accepted for session presentation and published at the 1<sup>st</sup> Data Science in Pavement Symposium (DSPS), Turner-Fairbank Highway Research Center, McLean, Virginia, 22-24 March 2022)

### 8.1. Introduction

One of the main goals of the Pavement Management System (PMS) is to enhance pavement condition by considering the effects of pavement structure and material, environment, and maintenance and rehabilitation (M&R) interventions on pavement performance [26]. A growing body of studies has been exploring different modeling techniques and variables to achieve accurate predictions. Advanced modeling techniques that use machine learning have shown promising results when predicting pavement roughness.

The Artificial Neural Network (ANNs) technique has been used in several studies offering significant improvements over traditional techniques, such as regression, by processing large volumes of data with a higher degree of accuracy. However, current pavement performance

prediction models did not account for the influence of M&R history in the model development, which can affect the accuracy of the predictions [88]. In addition, most studies developed performance models for flexible or concrete pavements, which resulted in a lack of research on composite pavements. Composite pavements are normally a result of concrete pavement rehabilitation, when concrete pavements start to fail, they are overlaid with Hot Mix Asphalt (HMA). Compared to flexible or rigid pavements, composite pavements can provide better performance measures both structurally and functionally, and accordingly, can be considered a cost-effective alternative [3].

Therefore, the objectives of this paper are to (1) develop pavement roughness prediction model for composite pavement sections in the wet non-freeze climate region of the LTPP database using the ANN technique and (2) utilize a new approach to incorporate M&R history in the model development.

## 8.2. Model Development

### 8.2.1. Methodology

The model development methodology used in this paper is described as follows:

1. Perform a literature review of previous studies to identify independent variables responsible for impacting pavement performance.
2. Compile databases for composite pavements in the wet freeze region of the LTPP database.
3. Develop pavement performance models using the ANN technique.
4. Evaluate the accuracy of the developed models using statistical measures.
5. Select the most accurate model based on statistical indicators.

### 8.2.2. Data Collection

Using the LTPP database, a total of 264 sections were identified as composite pavement

sections. In this study, only composite pavement sections located in the wet freeze climate zone were used. A total of 186 pavement sections with 1,930 data points were used for the analysis.

### 8.2.3. Data Processing

Several variables were retrieved from the LTPP database to develop the database used in this paper. The output and input variables are described in this section.

#### 8.2.3.1. Output Variables

The International Roughness Index (IRI) is recognized as one of the most important indicators of pavement performance. A total of 1,930 IRI measurements were found for the 186 sections. The mean of IRI outside wheel path ( $IRI_{Right}$ ) (1.27 m/km) measurements were 8% greater than the mean of IRI inside wheel path ( $IRI_{Left}$ ) (1.18 m/km) measurements, which agreed with previous research [2,31,36,88,89]. Therefore, the worst scenario was chosen and the  $IRI_{Right}$  was defined as the dependent variable for this study.

#### 8.2.3.2. Input Variables

Table 17 shows the variables used for each ANN model in this study.

Table 17. Independent and Dependent Variable Configuration for Five ANN Models.

Models	Model 1	Model 2
Independent Variables	$I_0$ $IRI_{Right}$	$I_0$ $IRI_{Right}$
	Age	Age
	CN <sub>Code</sub>	CN <sub>Continuous</sub>
	Unbound Base/Subbase	Unbound Base/Subbase
	Bound Base/Subbase	Bound Base/Subbase
	Cumulative ESAL	Cumulative ESAL
	$h_{asphalt}$	$h_{asphalt}$
	$h_{concrete}$	$h_{concrete}$
	$h_{base/subbase}$	$h_{base/subbase}$
Dependent Variable	IRIRight	IRIRight

Model 1 and Model 2 used nine input variables explained, as follows:

- $I_0$   $IRI_{Right}$ : initial  $IRI_{Right}$  when the pavement section was opened to traffic (m/km)

- Age: time in years from the opening date to the last measurement
- Unbound Base/Subbase: categorical variable to indicate unbound base/subbase (0 or 1), 1 if unbound
- Bound Base/Subbase: categorical variable to indicate bound base/subbase (0 or 1), 1 if bound
- Cumulative ESAL: cumulative equivalent single axle load from the opening date until the last measurement
- $h_{\text{asphalt}}$ : asphalt thickness (in.)
- $h_{\text{concrete}}$ : concrete thickness (in.)
- $h_{\text{base/subbase}}$ : base/subbase thickness (in.)
- $CN_{\text{Code}}$  and  $CN_{\text{Continuous}}$ : variables created for incorporating M&R actions

#### 8.2.4. Incorporating M&R Interventions in the Model Development

The construction number (CN) is the variable used by the LTPP to identify M&R actions in each section of the database. An initial CN1 is assigned when the pavement section is opened to the traffic. When an M&R is conducted, the CN number changes from CN1 to CN2. Hence, the CN factor indicates that an M&R treatment was conducted on the pavement section. The treatment intervention normally improves the pavement condition and performance with respect to roughness and other surface defects. Therefore, it is essential to consider CN as a factor for a more inclusive, realistic, and accurate performance prediction model. Two approaches were used for incorporating the M&R interventions in the ANN model development:

- The first approach was to use the CN as a categorical variable with a value of zero or one. A zero value is assigned if no M&R was implemented in that section and a value of one is assigned if there was an M&R intervention. This variable was called  $CN_{\text{Code}}$ .
- The second approach was to use the CN as a continuous variable with values from 1 to 9. An initial CN1 was assigned when the section became a composite section and increased

by one unit whenever an M&R action was performed in the pavement section. By doing this the CN values became CN2, CN3, CN4, CN5, CN6, CN7, CN8, and CN9 according to how many M&R interventions were performed in the same section. This variable was called CN<sub>Continuous</sub>.

- The use of a CN variable for incorporating M&R actions in the model development was expected to result in more realistic models considering that M&R interventions affect the future condition of the pavement.

#### 8.2.5. ANN Model Selection

A feedforward neural network with a back-propagation training algorithm was used for the development of the performance model in this study. A one hidden layer network was considered in the model development and a sigmoidal function was used for data generalization purposes. The TR-SEQ1 computer program [55] was used to develop the ANN models. Table 18 shows statistical measures of the ANN model development stages for the two developed models.

Table 18. ANN Model Results

Model		1	2
Structure (Initial Hidden Node – Final Hidden Node – Iterations)		6-18-20000	<b>6-19-20000</b>
Training	MARE	20.19	<b>18.52</b>
	R <sup>2</sup>	0.58	<b>0.69</b>
	ASE	0.0010	<b>0.00072</b>
Testing	MARE	19.85	<b>22.34</b>
	R <sup>2</sup>	0.51	<b>0.38</b>
	ASE	0.0011	<b>0.00135</b>
Validation	MARE	20.90	<b>20.98</b>
	R <sup>2</sup>	0.45	<b>0.43</b>
	ASE	0.0012	<b>0.00121</b>
All Data	MARE	17.94	<b>17.25</b>
	R <sup>2</sup>	0.59	<b>0.65</b>
	ASE	0.00082	<b>0.00072</b>
Final Network		9-19-1	
		# of Inputs - Final Hidden Node - # of Outputs	



The best model was selected based on the lowest average square error (ASE), lowest mean absolute relative error (MARE), and highest coefficient of determination ( $R^2$ ).

Training, testing, and validation stages used 973, 479, and 478 data points, respectively. After developing the possible networks for each model, the all-data stage used all 1,930 data points from the database to retrain the network at its optimal structure and iteration. Therefore, the all-data stage statistics were used as the main comparison between each model. The accuracy measures show reliable results for both developed models. However, Model 2 outperformed Model 1 in the all-data stage with a 12% lower ASE, 4% lower MARE, and 10% higher  $R^2$ . Therefore, Model 2 was chosen as the best-performing model. The final model structure includes 9 input variables, one hidden layer with 19 hidden nodes, 20,000 iterations, and one output.

### 8.3. Discussion

#### 8.3.1. Model Results

Figure 42 shows a scatter plot of observed versus predicted  $IRI_{Right}$  values for Model 1 and Model 2.

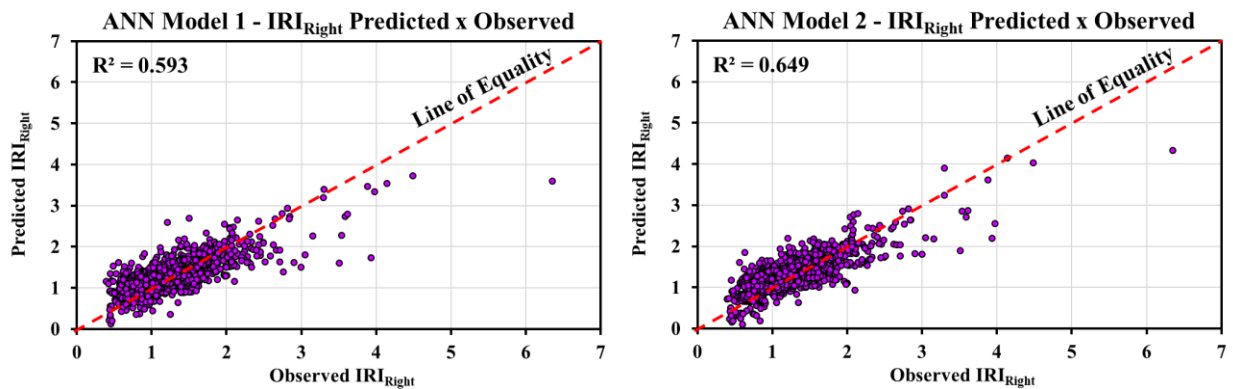


Figure 42. Observed vs. Predicted  $IRI_{Right}$  for ANN Model 1 and ANN Model 2, respectively

Both ANN models were able to capture the pavement deterioration behavior. However, Model 2 showed better accuracy with an  $R^2$  (0.649) 10% higher than Model 1 (0.593). Predicted

values for both models clustered around the line of equality; however, lower  $IRI_{Right}$  values were better predicted than higher values. When the observed  $IRI_{Right}$  values were greater than 3 m/km, both models were not as accurate. Nevertheless, both models resulted in reliable predictions explaining 59.3% and 64.9% of the variability in the model. Therefore, the use of a new approach utilizing a continuous variable for incorporating M&R history in the model development provided better results compared to the use of a categorical variable as observed in Table 18 and Figure 42.

#### 8.4. Conclusions

Model 2 showed better accuracy for predicting IRI values compared to Model 1. The use of a continuous variable ( $CN_{Continuous}$ ) for incorporating M&R actions showed results 12% lower for ASE, 4% lower for MARE, and 10% higher for  $R^2$ , which indicates a significant improvement over the model that used a categorical variable ( $CN_{Code}$ ). This new approach for maintenance and rehabilitation actions can also be implemented for other climate regions to verify if it can enhance other pavement performance models by providing more accurate predictions. The continuous approach developed in this study will assist transportation agencies to support objective decisions regarding maintenance and rehabilitation actions and budget plans permitting agencies to prioritize the resources for critical pavement sections. Furthermore, the developed models are user-friendly and can be easily utilized for predicting future pavement roughness conditions without the need for any distress data, saving time and money for federal and state agencies.

CHAPTER IX: PAVEMENT PERFORMANCE MODELING CONSIDERING  
MAINTENANCE AND REHABILITATION FOR COMPOSITE PAVEMENTS IN THE LTPP  
WET NON-FREEZE REGION USING NEURAL NETWORKS

Rulian B., Ph.D. Candidate, Hakan Y., Ph.D., Salma S., Ph.D., and Yacoub N., Ph.D.

Department of Civil Engineering, University of Mississippi, University, MS 38677, United States  
rferrei1@go.olemiss.edu

(Paper published and accepted for poster presentation at the American Society of Civil Engineers (ASCE) International Conference on Transportation & Development (ICTD) 2022, Seattle, Washington, May 31-June 3, 2022.)

### 9.1. Abstract

Efficient and well-maintained pavement systems are crucial to ensure appropriate conditions for the road networks. If timely maintenance and rehabilitation (M&R) is not performed, the pavement deterioration may lead to poor conditions that affect the comfort and safety of road users. The effectiveness of any M&R actions essentially depends on the time of treatment. This paper presents the development of pavement roughness models using the Artificial Neural Networks (ANNs) approach for composite pavements using the Long-Term Performance Pavement (LTPP) program database for the wet non-freeze climate region. A total of 49 composite pavement sections with 353 data points were analyzed. The use of an M&R variable in the model

development resulted in more realistic and accurate models to predict future pavement conditions, identify M&R actions, and simulate interventions for future years. The developed models could be used by transportation agencies as a valuable tool for more effective M&R scheduling prioritizing worst condition pavement sections.

## 9.2. Introduction

Efficient and well-maintained pavement systems are vital to guarantee proper conditions for the road networks. Pavement performance is a general term that intends to describe pavements' conditions over time indicating valuable information to keep the roads in acceptable levels of service to the users. Transportation agencies have used numerous performance indicators to assess the effectiveness and efficiency of their service provision [4]. Among others, the International Roughness Index (IRI) became the standard indicator to determine road surface roughness [20]. A well-constructed newly laid pavement has an initial roughness that tends to be lower, and it starts to increase with time as the pavement deteriorates due to climate and traffic loads [22]. With the lack of adequate and timely maintenance and rehabilitation (M&R) interventions, the rate of deterioration starts to rapidly increase leading the pavement to poor conditions.

Many funds are necessary to be invested in maintaining road networks within satisfactory conditions. However, if these funds are not sufficient, it is not possible to perform adequate and timely M&R activities, which is one of the main problems that agencies have been dealing [29,90]. For this reason, the development of pavement performance models becomes a vital tool that allows agencies to implement a better budget allocation plan for future M&R interventions. The performance models can identify rehabilitation needs, analyze rehabilitation effects, and estimate future pavement conditions to implement different M&R activities to extend the pavement life cycle leading agencies to develop a more efficient and effective Pavement Management System

(PMS) [27,88].

The Long-Term Pavement Performance (LTPP) program is the largest pavement performance research program ever undertaken, gathering data from more than 2,000 pavement test sections over a 20-year test period [84], and for this reason, is one of the most used databases for developing performance prediction models. Machine learning modeling techniques such as the Artificial Neural Network (ANN) have been proving to offer significant modeling improvements by processing large volumes of data with a higher degree of accuracy [26].

Therefore, this paper utilizes the ANN modeling technique to develop pavement performance models for composite pavements in the LTPP wet non-freeze region using a new approach to consider the effects of M&R treatments on IRI predictions.

### 9.3. Literature Review

In the last decades, significant research efforts have been shifted toward the use of machine learning algorithms for pavement performance modeling. The literature shows that the ANNs are not only one of the first machine learning techniques to be used but also the most used technique in civil and pavement engineering [40,41]. The ANN method attempts to emulate the structure and/or functional aspects of biological neural networks [46]. It consists of several simple processing elements called neurons (or nodes) and connecting links between them. When the information is processed, the connection links are used to transfer signals between neurons [43]. Each neuron evaluates its input signals to determine its output signal and transmitted to all neurons that are on the receiving side of the connection links originating in the transmitting neuron. Each connection has an associated weight that multiplies the signal transmitted. Complex relationships that are difficult to reproduce using traditional sequential, logic-based modeling and computation technics can be successfully represented by neural networks [43].

Since machine learning has a data-driven approach, IRI appears as a suitable indicator for modeling, since it is widely available in pavement databases (e.g., LTPP database), measured by objective means (e.g., laser profilometer), and known as one of the most common indicators for pavement performance evaluation [42]. Numerous studies have successfully used the ANN technique to predict IRI values for flexible pavements [10,13,15,16,28,29,68,69,91]. Other authors explored the use of ANNs for rigid pavements and obtained reasonable to high accuracy on IRI predictions [9,17,30,34,36,65,72]. The literature shows relatively less number of studies for composite pavements, but the ANN continues to be the most accurate and promising technique according to the results presented [10,15,26,68,88].

However, literature to date shows that several models did not include M&R history in modeling development. Incorporating M&R activities is a challenging task due to the complex relationships between pavement roughness and its responses before and after rehabilitation. Recent studies have shown promising results after using different approaches to introduce M&R variables into their modeling procedures [26,29–31,34,36,65,88]. Therefore, the development of performance prediction models for composite pavements incorporating M&R history is a must. This paper uses different approaches to generate M&R variables that assist the models to understand the roughness behavior before and after rehabilitation activities resulting in more reliable, inclusive, and accurate prediction models.

## 9.4. Model Development

### 9.4.1. Methodology

Figure 43 shows the pavement performance modeling methodology flowchart for the LTPP database.

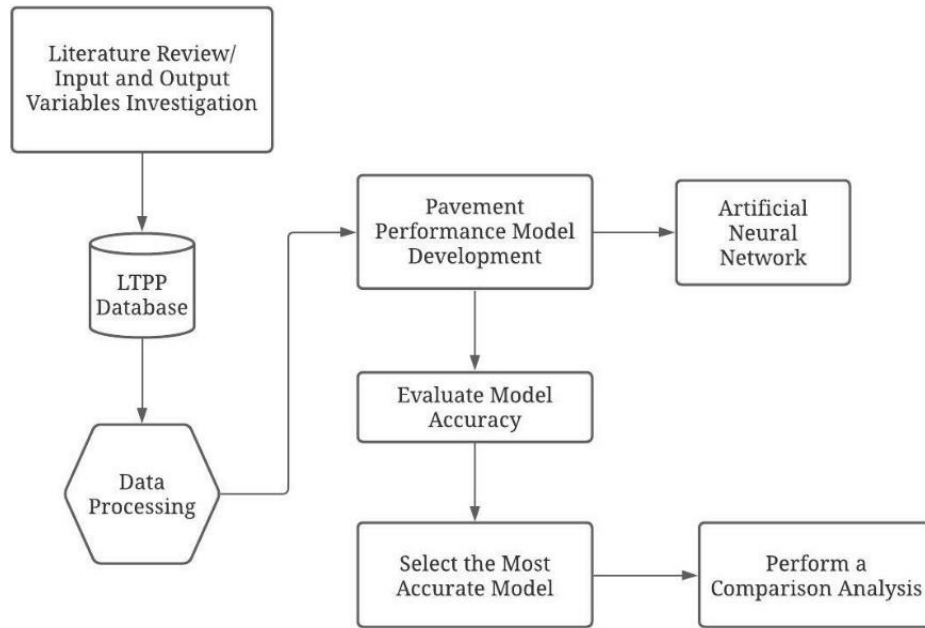


Figure 43. Pavement Performance Modeling Methodology

The description of the methodology is, as follows:

1. Conduct an extensive literature review to identify key input and output variables.
2. Compile databases for composite pavements in the wet non-freeze region of the LTPP database.
3. Develop pavement performance models using the ANN technique.
4. Evaluate the accuracy of the developed models using statistical measures.
5. Select the most accurate models based on statistical indicators.
6. Perform comparison analysis for the developed ANN models.

#### 9.4.2. Data Collection

The data used in this study were retrieved from the LTPP database. A total of 264 sections were identified as composite pavement sections, where the asphalt overlay thickness over the concrete layer was equal to or greater than three inches. In this study, only composite pavement sections located in the wet non-freeze climate zone were utilized, which resulted in a total of 49

sections available for the analysis.

The climate zone classification was developed at the beginning of the LTPP study since the data collection had different spatial and temporal locations throughout the U.S. Four climate zones were defined as wet freeze, wet non-freeze, dry-freeze, and dry non-freeze [84]. Table 19 shows the states located in the wet non-freeze climate region. The LTPP used the precipitation per year to identify wet (higher than 508 mm) and dry (lower than 508mm) climate zones. For freeze and non-freeze zones, the threshold used was based on an annual average freezing index of 83 °C (150 °F) days. Locations with an index over this threshold are classified in the freeze zone and those under the threshold in a non-freeze climate zone [84].

Table 19. States Located in the Wet Non-freeze LTPP Climate Zone

State Code	State Name	State Code	State Name
1	Alabama	24	Maryland
5	Arkansas	28	Mississippi
6	California	37	North Carolina
10	Delaware	40	Oklahoma
11	District of Columbia	41	Oregon
12	Florida	45	South Carolina
13	Georgia	47	Tennessee
15	Hawaii	48	Texas
16	Idaho	53	Washington
22	Louisiana	72	Puerto Rico

#### 9.4.3. Data Processing

To build the modeling database for this study, input and output variables were retrieved from the LTPP database. These variables are described in this section.

##### 9.4.3.1. Output Variables

The IRI is considered the standard measurement of pavement roughness, and it was used as the output variable for modeling. Each section had two types of IRI measurements; IRI inside wheel path ( $IRI_{Left}$ ) and IRI outside wheel path ( $IRI_{Right}$ ). A mean roughness index ( $IRI_{Mean}$ ) was



calculated by averaging the  $IRI_{Left}$  and  $IRI_{Right}$  measurements. On each visit date, several IRI measurement runs were done for each section. By averaging the IRI measurement runs, a single IRI measurement was obtained for  $IRI_{Left}$ ,  $IRI_{Right}$ , and  $IRI_{Mean}$  for each visit date. By doing this, a total of 353 IRI measurements from 1989 to 2018 were found for the 49 sections.

IRI measurements for  $IRI_{Right}$  (1.31 m/km or 83.0 in./mile) were 8% greater than the  $IRI_{Left}$  (1.22 m/km or 77.3 in./mile). Mean differences were assessed using the independent samples t-test to determine whether  $IRI_{Right}$  and  $IRI_{Left}$  differ on average from each other. The results show that the difference in the means of  $IRI_{Right}$  and  $IRI_{Left}$  is statistically significant at  $\alpha$  0.05 probability of chance error. This implies that both samples are from different populations. Therefore,  $IRI_{Right}$  was selected as the dependent variable since it shows the highest value for pavement roughness.

#### 9.4.3.2. Input Variables

The input variables were selected after an extensive literature review and several inputs were tried in a preliminary study using a trial-and-error method. The selected variables showed in this section were found to be the most significant variables considering the practical point of view. The selected input variables used in this study are explained, as follows:

- $IRI_0$  (m/km): represents the first IRI value measured in the outside wheel path for a specific pavement section of the LTPP database. The first measurement is usually done when the pavement was built and opened to traffic, or the pavement was first included in the LTPP study. It indicates the road surface condition at the beginning of the analysis period.
- Age (years): calculated by subtracting the year when the section was opened to traffic from the year that the IRI measurement was collected. This variable represents the time pavement was exposed to climate and traffic loads. Age is also a fundamental variable to be used as an input variable to predict pavement performance for future years.
- Granular Base/ Subbase (0 or 1): categorical variable to represent the use of a granular

base/subbase in the pavement structure. A value of “1” was used when granular base/subbase was utilized and a value of “0” when there was no granular base/subbase.

- Treated Base/ Subbase (0 or 1): categorical variable to represent the use of treated base/subbase in the pavement structure. A value of “1” was used when treated base/subbase was utilized and a value of “0” when there was no treated base/subbase.
- CESAL: The Cumulative Equivalent Single Axle Load (CESAL) is the sum of annual Equivalent Single Axle Load (ESAL) data over the years. The ESAL represents the effects of traffic loads on the pavement over time. In some years, the LTPP database did not have ESAL information corresponding to the IRI measurements data. Interpolation and extrapolation procedures were applied using known data points to compute ESAL for the missing years. Cumulative ESAL represents the cumulative traffic load that was endured by the pavement over pavements’ life.
- Surface Asphalt Thickness (in.): the surface asphalt thickness represents the thickness of asphalt overlaid on a concrete base.
- Concrete Thickness (in.): represents the thickness of concrete under the asphalt layer. The concrete layer serves as a base for the asphalt layer on top of the pavement.
- Subbase Thickness (in.): represents the thickness under the concrete base. Thicker subbases tend to provide more support for the pavement structure.
- Annual Average Temperature (°C): represents the average daily mean air temperatures for the year. The temperature changes impact the material properties of pavements and influence the pavement deterioration process. Thus, this climatological variable is necessary for more realistic performance models predictions.
- Total Annual Precipitation (mm): represents the sum of monthly precipitation for the year.

The amount of precipitation affects the material properties of pavements' base/ subbase and subgrade layers, which also impacts the pavement deterioration process and should also be used as a climatological variable in the modeling development.

For the consideration of M&R treatment in the ANN model development two variables ( $CN_{Code}$  and  $CN_{Continuous}$ ) were created and are explained, as follows:

The construction number (CN) is the attribute that LTPP uses to monitor and identify M&R in each section of the database. A CN1 is assigned when the pavement section was opened to the traffic. When an M&R is conducted, the CN number will change from CN1 to CN2. Thus, the CN factor indicates that a major M&R treatment was conducted on the pavement section. The treatment intervention normally improves the pavement condition concerning roughness, cracking, faulting, joint deterioration, and other surface defects. Hence, it is essential to consider CN as a factor for a more realistic and precise model. As an illustration, Figure 44 shows different CN values for section 06-0661 located in California. This test section has three construction numbers CN1, CN2, and CN3.

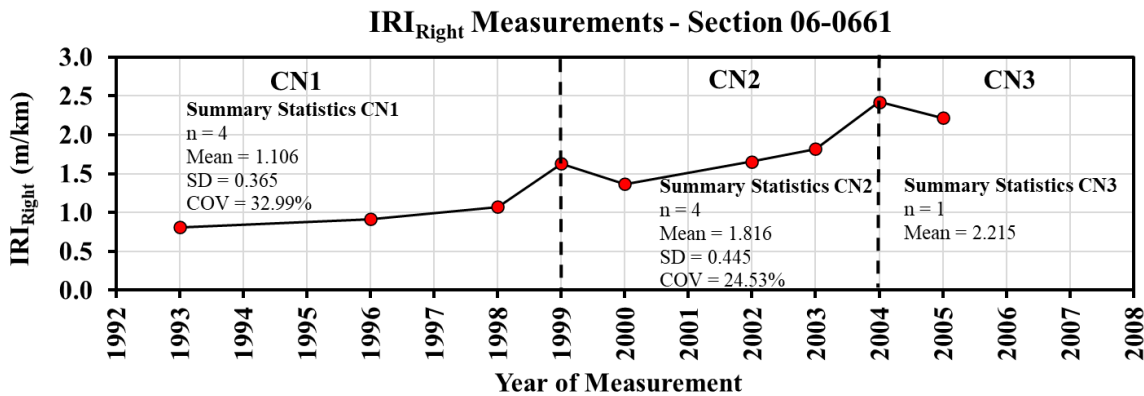


Figure 44. IRI<sub>Right</sub> measurements, Section 06-0661, California, 1993-2015

It is evident from Figure 44 that M&R treatments improved the composite pavement condition resulting in lower IRI values. The IRI values decreased 16% from CN1 (1.63 in 1999)

to CN2 (1.37 in 2000) when maintenance and rehabilitation (M&R) were performed. To support this statement an independent sample t-test was performed to determine whether there are statistically significant differences between the means of IRI measurements between CN1 and CN2. The results show that the difference in the means of CN1 IRI<sub>Right</sub> and CN2 IRI<sub>Right</sub> are statistically significant at  $\alpha$  0.05 probability of chance error. This implies that both IRI<sub>Right</sub> samples (CN1 and CN2) are from different populations. This confirmed the IRI<sub>Right</sub> value for CN2 (1.37 m/km) is statistically significant and lower, compared to the IRI<sub>Right</sub> value for CN1 (1.63 m/km). Thus, M&R treatments significantly improved the pavement surface condition and contributed to lower IRI values.

Therefore, two variables were created using different approaches.

- CN<sub>Code</sub>: the CN<sub>Code</sub> was developed based on the original CN collected from the LTPP database corresponding to each IRI data point. If no M&R action (CN1 in the original LTPP database) was done, the CN<sub>Code</sub> assumed a value of “0”, and continued as “0” until a new M&R action was performed. The first M&R action (CN2 in the original LTPP database) assumed a value of “1”, and this “1” continued until the end of IRI data points for this pavement section.
- CN<sub>Continuous</sub>: the CN<sub>Continuous</sub> was also developed based on the original CN collected from the LTPP database corresponding to each IRI data point. If no M&R action (CN1 in the original LTPP database) was done, the CN<sub>Code</sub> assumed a value of “1”, and continued as “1” until a new M&R action was performed. The next M&R action (CN2 in the original LTPP database) assumed a value of “2”, and this “2” continued until the next M&R intervention. The next M&R action (CN3 in the original LTPP database) assumed a value of “3”, and this “3” continued until the next M&R action. This CN<sub>Continuous</sub> assignment

continued until all the M&R actions performed on a pavement section were categorized.

#### 9.4.4. ANN Modeling and Structure

For the ANN modeling procedure, this study utilized a feed-forward neural network with a back-propagation training algorithm to develop performance prediction models for asphalt pavements in the wet non-freeze region of the LTPP database. One hidden layer was implemented to maintain the generalization capability of the network without causing the network to memorize the data in the training phase [46]. The modeling database was divided into training (50%), testing (25%), and validation (25%). After achieving the optimum model networks, the model was retrained in the All-data stage using 100% of the data points. The TR-SEQ1 computer program [55] with a sigmoidal function (for data generalization purposes) was used to develop the ANN models in this study.

A total of three models were developed using different input and output variables from the modeling database. Table 20 summarizes the structure of each model used in this paper.

Table 20. Input and Output Variables for ANN Models

Models	Model 1	Model 2	Model 3
Input Variables	IRI <sub>0</sub>	IRI <sub>0</sub>	IRI <sub>0</sub>
	Age	Age	Age
	CN <sub>Code</sub>	CN <sub>Code</sub>	CN <sub>Continuous</sub>
	Granular Base/ Subbase	Granular Base/ Subbase	Granular Base/ Subbase
	Treated Base/ Subbase	Treated Base/ Subbase	Treated Base/ Subbase
	CESAL	CESAL	CESAL
	Surface Asphalt Thickness	Surface Asphalt Thickness	Surface Asphalt Thickness
	Concrete Thickness	Concrete Thickness	Concrete Thickness
	Subbase Thickness	Subbase Thickness	Subbase Thickness
		Annual Average Temperature	
	Total Annual Precipitation		
Output Variables	IRI <sub>Right</sub>	IRI <sub>Right</sub>	IRI <sub>Right</sub>

Models 1 and 3 used nine input variables and one output variable. However, Model 1 used

the  $CN_{Code}$  to incorporate M&R actions in the model development, while Model 3 used the  $CN_{Continuous}$ . This approach was used to identify which M&R variable would provide more accurate predictions. Model 2 utilized 11 input variables and one output variable. The two additional input variables (Annual Average Temperature and Total Annual Precipitation) in Model 2 were included to take into consideration the climatological effects on pavement deterioration. Model 2 also used the  $CN_{Code}$  as the main variable for M&R actions. All models have used the same output model ( $IRI_{Right}$ ).

#### 9.4.5. ANN Model Selection

The best model was chosen based on the lowest Average Square Error (ASE), Mean Absolute Relative Error (MARE), and highest Coefficient of Determination ( $R^2$ ). Table 21 shows the summary of the ANN model results. The model network is written in an order that depicts the number of initial hidden nodes, the number of final hidden nodes, and iterations. The final structure of each model is written at the bottom row of the table, in an order that depicts the number of inputs, hidden nodes, and output(s), respectively.

Table 21. ANN Model Results

Model		<b>Model 1</b>	Model 2	Model 3
Network		<b>1-19-20000</b>	13-19-2000	8-17-20000
Training	MARE	<b>8.08</b>	9.98	10.68
	$R^2$	<b>0.96</b>	0.94	0.93
	ASE	<b>0.0005</b>	0.0007	0.0008
Testing	MARE	<b>17.15</b>	20.71	20.66
	$R^2$	<b>0.44</b>	0.72	0.60
	ASE	<b>0.0121</b>	0.0043	0.0058
Validation	MARE	<b>14.73</b>	20.24	25.90
	$R^2$	<b>0.83</b>	0.71	0.62
	ASE	<b>0.0020</b>	0.0043	0.0057
All data	MARE	<b>10.45</b>	13.14	12.94
	$R^2$	<b>0.90</b>	0.86	0.87
	ASE	<b>0.0011</b>	0.0018	0.0013
Final Structure		<b>9-19-1</b>	11-19-1	9-17-1

Training stage results show that Model 1 had better accuracy measures in terms of ASE, MARE, and  $R^2$ . On the other hand, for the testing stage, Model 2 showed better results for ASE and  $R^2$  while Model 1 had a better MARE outcome. In the validation stage, Model 1 showed lower ASE, MARE, and higher  $R^2$  results compared to other models. In the All-data stage, all the 353 data points were used to retrain the network at its optimal structure and iterations to obtain the generalized response throughout the complete database. Results for the All-data stage show that Model 1 outperformed all other models with better accuracy measures for ASE, MARE, and  $R^2$ . Statistical measures of Model 1 retrained with all data showed an ASE of 0.0011, a MARE of 10.45, and a high  $R^2$  of 0.90.

Therefore, Model 1 was selected as the best-performing ANN model for the composite pavements located in the LTPP wet non-freeze climate zone. The network structure of Model 1 includes 9 input variables, 1 hidden layer with 19 hidden nodes, 20,000 iterations, and 1 output.

## 9.5. Discussion

### 9.5.1. ANN Model Results

Figure 45 shows a graphical comparison of statistical measures for all models using the All-data stage.

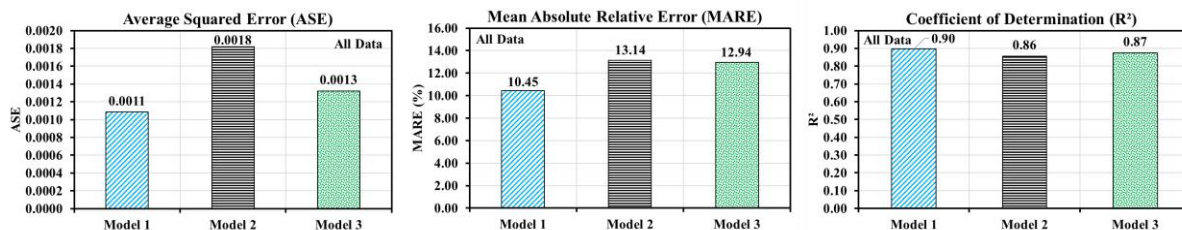


Figure 45. Graphical comparison of accuracy measures for all developed models

All three models presented accurate, reliable, and consistent results. However, Model 1 outperformed all other models with an ASE 40% and 18% lower, a MARE 20% and 19% lower, and an  $R^2$  5% and 3% higher than Models 2 and 3, respectively. This analysis supports the previous

selection of Model 1 as the best-performing model. Figure 46 shows the observed IRI values collected from the LTPP database versus the predicted IRI values using Model 1.

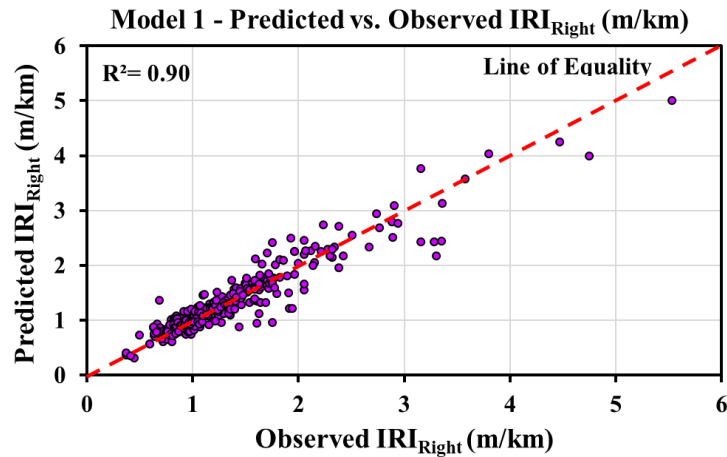


Figure 46. Model 1 Predicted vs. Observed IRI<sub>Right</sub>

Model 1 predictions clustered around the line of equality showing an excellent agreement ( $R^2 = 0.90$ ) between observed and predicted values IRI values. The model was able to capture the composite pavement deterioration behavior and understand the roughness changes over time explaining 90% of the variability. The complex relationships between the input variables and IRI<sub>Right</sub> were successfully modeled and translated into reliable, consistent, and accurate predictions.

The use of the CN<sub>Code</sub> variable showed more efficiency in assisting the network to incorporate the effects of M&R in the model development than the CN<sub>Continuous</sub>. This might occur because when no rehabilitation is performed the use of “0” demonstrates to the model that there is no need to assign any weight for that variable at that moment. After an M&R intervention, the CN<sub>Code</sub> receives a value of “1”, which indicates to the model that some action was performed and a drop in the IRI<sub>Right</sub> value needs to occur. Another important point of discussion is the use of climatological variables (Model 2). Since the database was already developed for sections located in the wet non-freeze climate region and this region was already classified according to the



precipitation and freeze data, it is not necessary to utilize additional climatological variables that contain the same information. The accuracy comparison between Model 1 (no climatological variables) and Model 2 (with climatological variables) confirms this statement and this approach should be applied in different LTPP climate zones for future studies.

Therefore, statistical measures and prediction plots showed that all developed models were accurate and reliable tools that could be used by transportation agencies to assist their PMS providing more efficient and timely effective M&R interventions. However, Model 1 showed to provide the most accurate results and, thus, is recommended to be used as the best-performing model for predicting IRI values for composite pavements in the LTPP wet non-freeze climate region.

### 9.6. Implementation of Model 1 for Random Sections

This section presents the implementation of Model 1 for two random sections of the database to visualize and evaluate predicted and observed roughness. This comparison is important to identify the model's capability to predict IRI values for some sections under different input variable values. Figure 47 shows the comparison between predicted and observed IRI values using Model 1 for Sections 40-0603 and 48-5154.

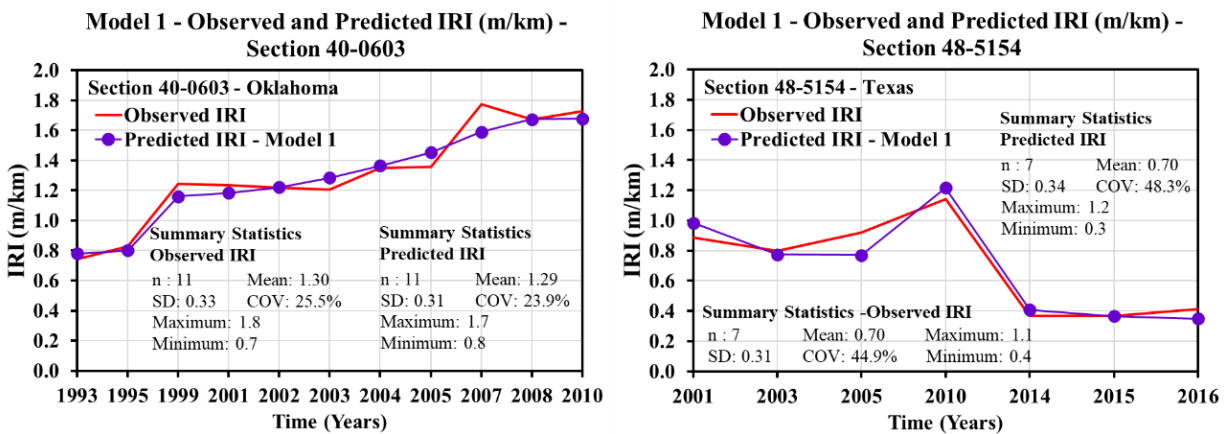


Figure 47. Model 1 Observed and Predicted Plots of IRI for Sections 40-0603 and 48-5154

Predicted IRI values were very close to the observed values for both sections. The predicted mean of IRI (1.29) for section 40-0603 was only 1% lower than the observed mean (1.30), while for section 48-5154 predicted values were 0.5% lower. The model predicted successfully the IRI increase over time and the IRI decrease right after an M&R action was applied in the pavement section. Therefore, the implementation analysis showed that agencies can use the developed model for predicting future roughness values and incorporate M&R interventions effectively to prioritize funds for the worst condition sections.

### 9.7. Conclusion

In this study, an ANN modeling technique was used to develop roughness performance prediction models for composite pavements in the LTPP wet non-freeze climate region. Even though all developed models showed accurate and reliable results, Model 1 outperformed other models with an ASE 40% and 18% lower, a MARE 20% and 19% lower, and an  $R^2$  5% and 3% higher than Models 2 and 3, respectively. The network structure of Model 1 includes 9 input variables, 1 hidden layer with 19 hidden nodes, 20,000 iterations, and 1 output. The predicted vs. observed plot (Figure 46) showed that Model 1 predictions clustered around the line of equality showing an excellent agreement ( $R^2 = 0.90$ ) between observed and predicted values IRI values. The use of the  $CN_{Code}$  variable efficiently incorporated M&R actions in the model development and assisted the network to achieve higher accuracy compared to the  $CN_{Continuous}$  variable. The results also showed that since the database was already developed for a specific climate region, the inclusion of climatological variables related to precipitation and temperature did not improve model accuracy and for this reason should not be included when modeling specific climate regions. Implementation of Model 1 for two random sections of the database showed that the model had excellent performance predicting the IRI increase and decrease due to time exposure and M&R

actions.

Therefore, this study successfully developed an ANN performance model that can contribute to the state-of-the-art by providing a reliable, accurate and consistent tool to support objective decisions regarding M&R interventions allowing agencies to develop more efficient PMS by prioritizing resources for critical composite pavement sections.

#### 9.7.1. Acknowledgments

The research described in this paper was conducted at the University of Mississippi. The contents of this paper reflect the views of the authors who are responsible for the facts, findings, and data presented herein.

## CHAPTER X: DEVELOPMENT OF PERFORMANCE MODELS FOR FLEXIBLE PAVEMENTS USING THE MDOT DATABASE

Rulian Barros and Hakan Yasarer

Department of Civil Engineering, University of Mississippi, Oxford, Mississippi, United States

(Research submitted for the Mississippi Department of Transportation (MDOT) Research Project, Development of Pavement Performance Models for MDOT: A Neural Network Approach)

### 10.1. Model Development

#### 10.1.1. Methodology

PCR and IRI are the two most important indicators of pavement smoothness. State agencies must predict these two measures accurately to estimate the future condition of the pavement sections. Frequent maintenance, rehabilitation, and eventually resurfacing are necessary to maintain the pavement in acceptable conditions. Therefore, pavement performance models are necessary to accurately estimate pavement deterioration to prioritize budget allocation for a better management system. Figure 48 presents the pavement performance modeling methodology flowchart for the MDOT database.

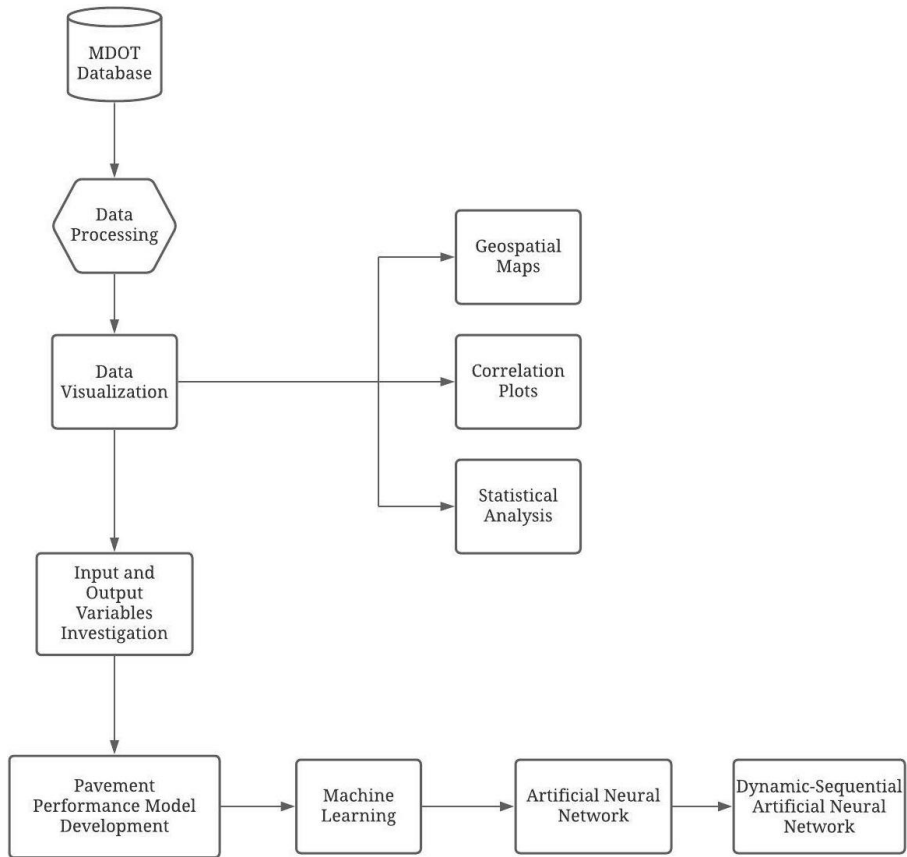


Figure 48. Pavement Performance Modeling Methodology for the MDOT Database

The model development methodology for the MDOT database used in this study is presented, as follows:

- (1) Compile databases for asphalt, concrete, and composite pavements from the MDOT database, including variables that affect pavement performance.
- (2) Assess the quality of databases and identify missing/erroneous data items.
- (3) Develop procedures for estimating important missing data in the time series.
- (4) Develop pavement performance models for asphalt, concrete, and composite pavements using the ANN modeling technique.
- (5) Evaluate the accuracy of the developed performance prediction models.
- (6) Select the best-performing model based on statistical measures and verify the prediction behavior.
- (7) Implement the selected performance models via GUI.

(8) Evaluate the selected models using GUI for the enhancement of pavement asset management.

#### 10.1.2. Data Collection

The data used in this study is collected as a part of the MDOT's survey. Mississippi has four different pavement types: Flexible, JCP, CRCP, and composite pavement. Like other states, the MDOT operates a pavement management system that includes PCR, IRI, and distress data. Typically, PCR values decrease and IRI values increase with the deterioration of pavement over time. IRI is measured as  $IRI_{Mean}$  and classified as low, medium, and high severity levels. The PCR values are calculated based on measured IRI and the deduction factors that are determined as functions of distresses and severity level [28,50].

#### 10.1.3. Data Processing

The database for the ANN model development is obtained after cleansing and reorganizing the raw data files. Due to new data acquisition methods, materials, and tools used by the MDOT, only data collected from 2010 to 2020 are included in the model development. Sections with missing or illogical data have been excluded as the ANN model development process needs a complete dataset. These exclusions reduced the number of asphalt, concrete, and composite pavement sections. Each section is comprised of five different datasets based on PCR, IRI, and rehabilitation actions. In Mississippi, the distress data is collected every even year. To develop prediction models that are applicable for a 1-year increment, the odd-year data needed to be generated by averaging consecutive years from 2010 to 2020 [28,50].

It is known that MDOT did not keep track of all rehabilitation actions and for this reason, another approach for assigning rehabilitation actions was proposed based on the discussions with the state agency. Improvement of PCR and IRI values without any rehabilitation action was

considered irrational. Some uncertainty due to the calibration of the profilometer, systematic errors, and the environmental conditions on the day of the survey may have resulted in some of the irrational condition measures. To incorporate the effect of the rehabilitation on PCR and IRI, artificial rehabilitation actions based on the significant changes in PCR and IRI have been assigned to the database [50]. Threshold values for PCR and IRI were assigned based on the evaluation of data history. If PCR increased 8% to 12% and IRI decreased 5% to 16% in a year compared to the previous measurement, a minor rehabilitation was assumed to take place in that year. If PCR increased above 12% and IRI decreased more than 16%, a major rehabilitation was assumed. If none of these situations occurred, it was assumed no rehabilitation. The models with two outputs (i.e., PCR and IRI) were modified to be used with the complementary PCR (i.e., 100-PCR) since the outputs need to be directly proportional in the ANN modeling. Because PCR and IRI usually change inversely over time, it was necessary to utilize the complementary PCR and IRI since both change proportionally over time. Therefore, the use of complementary PCR assisted the network to optimize the model with better accuracy and establishing a superior correlation between actual and predicted outputs.

## 10.2. MDOT Flexible Pavement Performance Model

For the development of the asphalt pavement performance models, 35,712 data points from 3,968 pavement sections throughout the state of Mississippi were used. Figure 49 shows a spatial map of the 3,968 flexible pavement sections from the MDOT database included in the model development.

## MDOT Flexible Pavement Sections

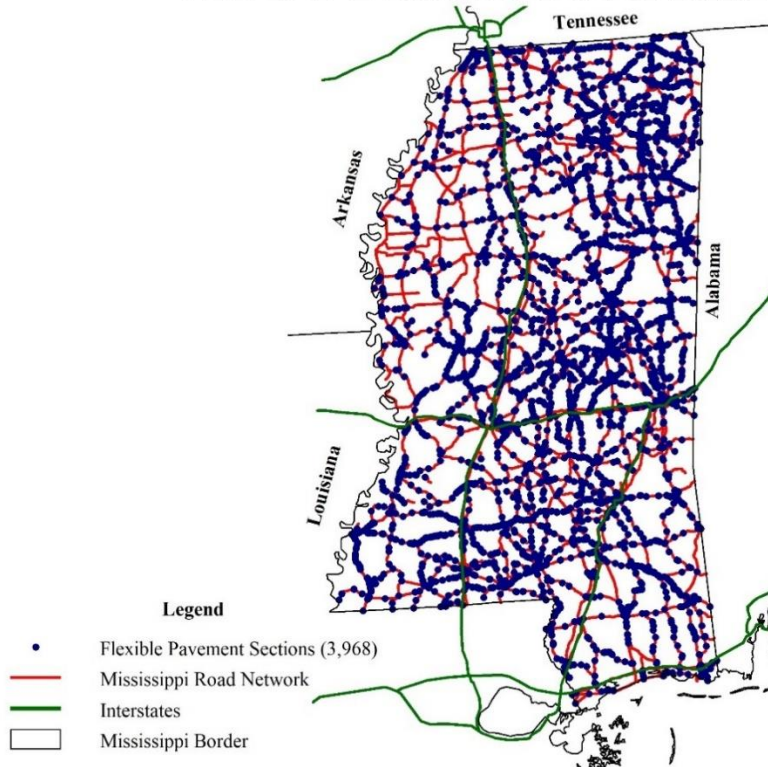


Figure 49. Spatial Map of MDOT Flexible Pavement Sections

### 10.2.1. ANN Model Variables and Architecture for MDOT Asphalt Pavements

For the MDOT asphalt pavement modeling research, different models were developed varying the numbers of independent and dependent variables. These variables are explained, as follows:

- Beginning Longitude and Latitude: coordinates to indicate the initial location of the roadway section.
- Ending Longitude and Latitude: coordinates to indicate the end of the roadway section.
- Structural Number: indicates the strength of the roadway when factoring material properties, thickness, and drainage in each layer.
- Length of the Section: length of the section recorded in miles.
- Age of the Section in 2010: shows the section's age since the earliest available pavement



measurement was recorded.

- PCR in 2010: shows the initial PCR value in 2010 to indicate the base starting value.
- IRI in 2010: shows the initial IRI in 2010 to indicate the baseline value to represent the pavement initial condition.
- Time since 2010: represents the time since 2010 to the desired prediction year. This value is associated with the effects of pavement aging.
- Drainage: categorical variable to indicate whether the section has drainage or not. The categorical value of “0” indicates no drainage while the sections with “1” indicate the existence of drainage.
- Minor Rehabilitation: categorical variable to represent minor rehabilitation. Use “1” if PCR increased 8% to 12% and IRI decreased 5% to 16% in a year compared to the previous measurement. If not, use “0”.
- Major Rehabilitation: categorical variable to represent major rehabilitation. Use “1” if PCR increased above 12% and IRI decreased more than 16%. If not, use “0”.
- ESAL: Equivalent Single Axle Load in that specific year.
- CESAL: Cumulative Equivalent Single Axle Load in that specific year.
- PRE PCR: variable used for dynamic ANN models to indicate the PCR from the previous years that will be used to predict the actual year.
- PRE IRI: variable used for dynamic ANN models to indicate the IRI from the previous years that will be used to predict the actual year.
- IRI: International Roughness Index measured in that year.
- Complementary PCR: this variable is calculated by subtracting 100-PCR. This parameter is generated because the outputs need to be directly proportional in the ANN modeling. Because PCR and IRI usually change inversely over time, it was necessary to utilize complementary PCR. The use of this variable assisted the network to optimize the model with better accuracy and correlation between observed and predicted outputs.

In this study, five models were developed for the MDOT asphalt pavements. Table 22 shows the variables used for each model.

Table 22. Independent and Dependent Variables for MDOT Asphalt Pavements

Models	Model 1	Model 2	Model 3	Model 4	Model 5
Independent Variables	Begin Lat.	Begin Lat.	Begin Lat.	Begin Lat.	Begin Lat.
	Begin Long.	Begin Long.	Begin Long.	Begin Long.	Begin Long.
	End Lat.	End Lat.	End Lat.	End Lat.	End Lat.
	End Long.	End Long.	End Long.	End Long.	End Long.
	SN	SN	SN	SN	SN
	Length	Length	Length	Length	Length
	Age in 2010	Age in 2010	Age in 2010	Age in 2010	Age in 2010
	PCR in 2010	PCR in 2010	PCR in 2010	PCR in 2010	PCR in 2010
	IRI in 2010	IRI in 2010	IRI in 2010	IRI in 2010	IRI in 2010
	Time (t)	Time (t)	Time (t)	Time (t)	Time (t)
	Drainage	Drainage	Drainage	Drainage	Drainage
	IRI Minor	IRI Minor	IRI Minor	IRI Minor	IRI Minor
	IRI Major	IRI Major	IRI Major	IRI Major	IRI Major
		ESAL	ESAL	ESAL	CESAL
	PRE PCR	PRE PCR	PRE PCR	PRE PCR	PRE PCR
PRE IRI	PRE IRI	PRE IRI	PRE IRI	PRE IRI	
Dependent Variables	Complementary PCR	Complementary PCR	Complementary PCR		Complementary PCR
	IRI <sub>Mean</sub>	IRI <sub>Mean</sub>		IRI <sub>Mean</sub>	IRI <sub>Mean</sub>

Model 1 utilized 15 independent variables and was the only one that did not have any traffic variable included in its development. A two-output model with IRI<sub>Mean</sub> and complementary PCR was utilized. Model 2, Model 3, and Model 4 included 16 independent variables and used the ESAL variable to incorporate the effects of traffic in the pavement deterioration process. The output variables with IRI<sub>Mean</sub> and complementary PCR were also utilized for Model 2 while Model 3 used only complementary PCR and Model 4 only IRI. Model 5 also used 16 independent variables but utilized CESAL to embed the history of traffic loads for inclusive predictions. Model 5 utilized the same two-output model with IRI<sub>Mean</sub> and complementary PCR. The use of variables that are not related to distress data makes the models more accessible for transportation agencies since most of these variables are easily available in their databases. Therefore, this model brings a valuable tool for the MDOT’s asphalt pavement management system.

### 10.2.2. ANN Model Selection for MDOT Asphalt Pavements

The best model was selected based on the lowest ASE, MARE, and highest R<sup>2</sup>. A total of 35,712 data points from 3,968 flexible pavement sections were used to build the ANN modeling database. The maximum and minimum values of each independent variable were included in the training phase for the network to represent the characteristics of the response. The maximum and minimum ranges of each input/output variable for ANN model development were chosen on purpose to be wider than their actual ranges for better mathematical mapping [51]. The final structure of each network is written at the bottom row in an order that depicts the number of inputs, hidden nodes, and output(s), respectively. Table 23 shows the comparison of each model for the complementary PCR output.

Table 23. ANN Model Results for MDOT Flexible Pavements (Complementary PCR Output)

Model		Model 1	Model 2	Model 3	<b>Model 5</b>
Complementary PCR		9-19-20000	16-19-20000	6-12-20000	<b>3-16-20000</b>
Training	MARE	17.52	17.25	17.46	<b>17.69</b>
	R <sup>2</sup>	0.659	0.668	0.660	<b>0.664</b>
	ASE	0.00380	0.00376	0.00380	<b>0.00373</b>
Testing	MARE	17.29	17.21	16.96	<b>17.15</b>
	R <sup>2</sup>	0.545	0.539	0.551	<b>0.551</b>
	ASE	0.00388	0.00391	0.00384	<b>0.00377</b>
Validation	MARE	19.22	19.17	18.77	<b>19.10</b>
	R <sup>2</sup>	0.513	0.503	0.516	<b>0.519</b>
	ASE	0.00491	0.00512	0.00496	<b>0.00490</b>
All data	MARE	17.34	17.80	17.32	<b>17.57</b>
	R <sup>2</sup>	0.622	0.615	0.616	<b>0.617</b>
	ASE	0.00434	0.00449	0.00429	<b>0.00441</b>

For the complementary PCR output, the training stage showed that Model 5 had the lowest ASE value while Model 2 had the lowest MARE and highest R<sup>2</sup> values. In the testing stage, Model 5 had the lowest ASE while Model 1 had lowest MARE, the highest R<sup>2</sup> was achieved by both

Model 5 and Model 1. In the validation stage, Model 5 showed better accuracy measures in terms of ASE and  $R^2$  while Model 3 had the lowest MARE. In the all-data stage, the 35,712 data points were used to retrain the network at its optimal structure and iteration to obtain the generalized response throughout the complete database. The all-data stage results show that Model 3 had the lowest ASE and MARE, while Model 1 had best  $R^2$ . However, all models showed similar performance showing reasonable accuracy. Considering the results of all stages, Model 5 was selected as the most promising performance model. Statistical measures of Model 5 trained with all data showed an ASE of 0.00441, a MARE of 17.57, and an  $R^2$  of 0.617.

Table 24 shows the comparison of each model for the IRI output.

Table 24. ANN Model Results for MDOT Flexible Pavements (IRI Output)

Model		Model 1	Model 2	Model 4	<b>Model 5</b>
IRI		9-19-20000	16-19-20000	5-16-19200	<b>3-16-20000</b>
Training	MARE	23.39	22.68	24.46	<b>23.35</b>
	$R^2$	0.397	0.403	0.361	<b>0.394</b>
	ASE	0.0000235	0.0000233	0.0024900	<b>0.0000236</b>
Testing	MARE	22.42	22.08	22.12	<b>22.31</b>
	$R^2$	0.360	0.385	0.380	<b>0.390</b>
	ASE	0.0000311	0.0000303	0.0030300	<b>0.0000303</b>
Validation	MARE	27.27	27.30	27.02	<b>26.75</b>
	$R^2$	0.122	0.108	0.130	<b>0.132</b>
	ASE	0.0000375	0.0000416	0.0036700	<b>0.0000369</b>
All data	MARE	23.39	23.37	23.83	<b>23.61</b>
	$R^2$	0.370	0.362	0.347	<b>0.365</b>
	ASE	0.0000275	0.0000286	0.0029100	<b>0.0000275</b>

For the IRI output, the training stage showed Model 2 had better accuracy measurements for ASE, MARE and  $R^2$ . The testing stage showed that Model 2 and Model 5 had the lowest ASE value, while Model 2 had the lowest MARE and Model 5 the highest  $R^2$  value. In the validation stage, Model 5 showed better results for all three statistics measurements. The all-data stage

showed that all models had similar results with moderate confidence. Model 1 and Model 5 showed the lowest ASE values while Model 2 had the lowest MARE and Model 1 highest  $R^2$  values. Considering the results of all stages, Model 5 was selected as the best model. Statistical measures of Model 5 trained with all data showed an ASE of 0.0000275, a MARE of 23.61, and an  $R^2$  of 0.365.

Therefore, Model 5 was chosen as the best-performing ANN model for MDOT flexible pavement. The network structure of Model 5 includes 16 input variables, 1 hidden layer with 16 hidden nodes, 20,000 iterations, and 2 outputs. Figure 50 shows the network architecture of the best-performing ANN model.

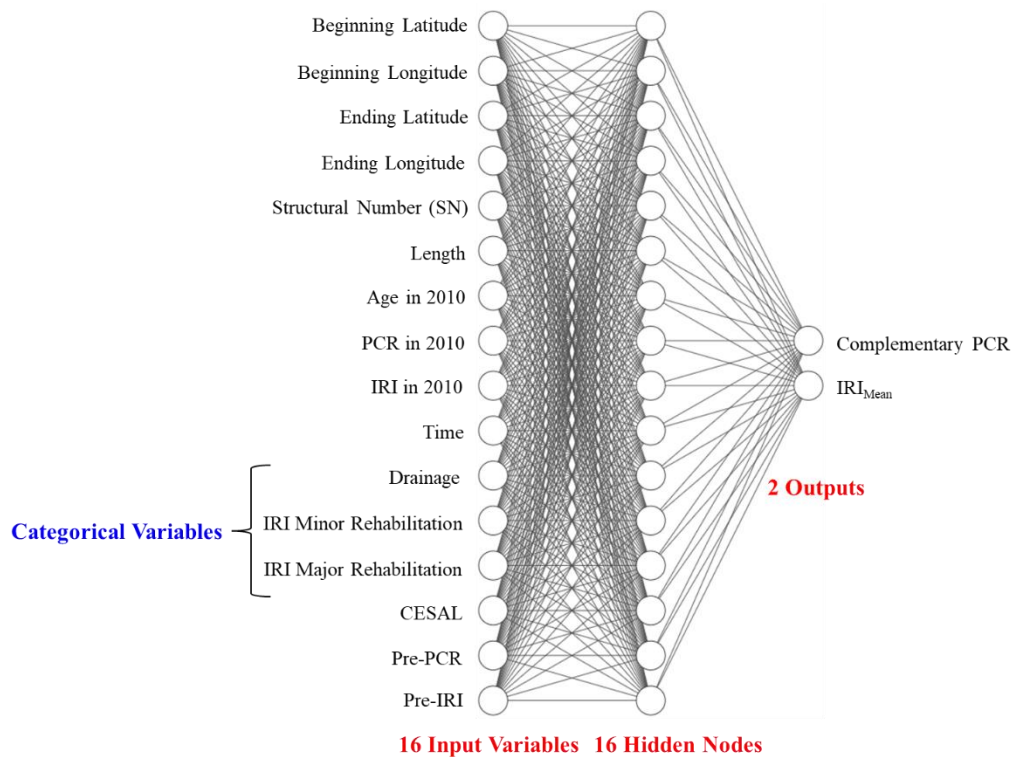


Figure 50. ANN Architecture for the Best Performing Model for MDOT Asphalt Pavement

### 10.2.3. ANN Model Results for MDOT Asphalt Pavements

Model 5 was chosen as the best performing network based on statistical measures (ASE,

MARE, and  $R^2$  value) of all the stages. Figure 51 shows the all-data stage comparison of the developed MDOT asphalt pavement ANN models for  $IRI_{Mean}$  and complementary PCR.

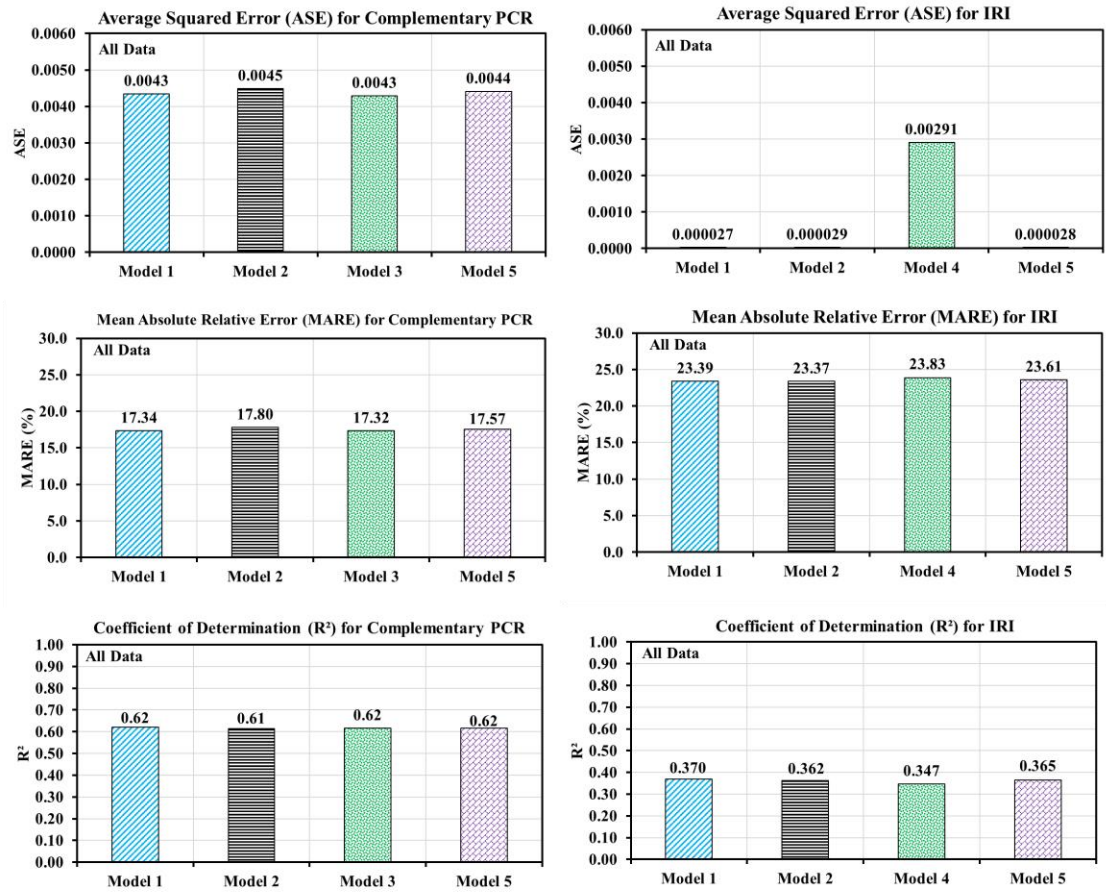


Figure 51. Graphical Comparison of Accuracy Measures for MDOT ANN Models

The accuracy measures show reasonable results for all five developed models. However, Model 5 results were more consistent considering training, testing, validation, and all-data stages. The complementary PCR output results show that all four models had similar results with reasonable accuracy. For the IRI output, Model 4 was the only model with higher ASE value, all other models had similar accuracy and error measurements.

Figure 52 shows the observed PCR values collected from the MDOT asphalt database and the predicted PCR values using Model 5.

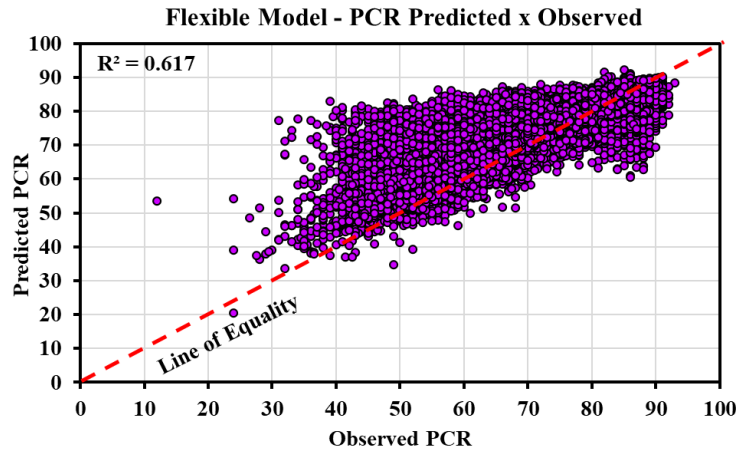


Figure 52. Observed vs. Predicted Complementary PCR for MDOT Flexible ANN Model 5

Model 5 predictions clustered around the line of equality but the model overpredicted some of the PCR values, however, the model was able to capture the PCR behavior. Hence, the model showed a reasonable  $R^2$  of 0.617. Figure 53 shows the observed  $IRI_{Mean}$  values from the MDOT asphalt database and the predicted  $IRI_{Mean}$  values using Model 5.

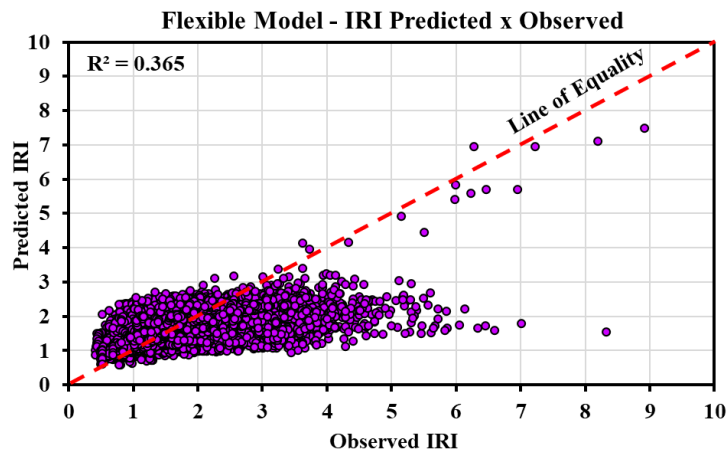


Figure 53. Observed vs. Predicted  $IRI_{Mean}$  for MDOT Flexible ANN Model 5

Predicted values tried to get close to the line of equality but most values were underpredicted by the model.  $IRI_{Mean}$  values that were between 0 and 2.5 were predicted with good accuracy, however, when the observed values were greater than 2.5 m/km the model could not

capture the  $IRI_{Mean}$  behavior showing an  $R^2$  of 0.365.

Therefore, to improve the model accuracy for the IRI output, it is necessary to further study the ANN model architecture and variables used for the modeling development. However, the model can be considered reasonably accurate if the user goal is to obtain PCR values for the asphalt pavement section.

### 10.3. Implementation of the Developed Models via GUI for the ANN MDOT Asphalt Model

A preliminary GUI was developed in Microsoft Excel using the connections weights and threshold values for the ANN MDOT asphalt model. The developed excel application has 16 input variables and 2 output variables. In this application, by entering the input variables of beginning latitude and longitude, ending latitude and longitude, structural number, section length, age in 2010, PCR in 2010, IRI in 2010, time, drainage, IRI minor, IRI major, CESAL, Pre-PCR, and Pre-IRI in the Excel interface shown in Figure 54, the  $IRI_{Mean}$  is automatically calculated by the application. The appropriate ranges for the input variables are also shown in Figure 54. If any of the provided input values are outside the applicable range, it may cause the model to generate unreliable predictions.

INPUTS										OUTPUTS	
Section ID	Begin Lat.	Begin Long.	End Lat.	End Long.	SN	Length	Age at 2010	PCR @ 2010		PCR	IRI
13	34.121	-90.826389	34.211141	-90.710163	2.79	14.586	46.0	68.0		70	1.44
	Iri @ 2010	Time (t)	Drainage	IRI Minor	IRI Major	CESAL	PRE PCR	PRE IRI			
	1.45	1	0	0	0	229.95	32	1.45			

USER RANGES								
	Begin Lat.	Begin Long.	End Lat.	End Long.	SN	Length	Age at 2010	PCR @ 2010
MINIMUM	30.2397	-91.5479	30.3021	-91.5214	0.20	0.062	1	35
MAXIMUM	34.9954	-88.1168	34.9957	-88.1160	9.17	22.363	110	92
	Iri @ 2010	Time (t)	Drainage	IRI Minor	IRI Major	CESAL	PRE PCR	PRE IRI
MINIMUM	0.59	1	0	0	0	31	8	0.42
MAXIMUM	7.49	9	1	1	1	3307831	88	8.92

Figure 54. Preliminary GUI Application for ANN Model for MDOT Asphalt Pavement



### 10.3.1. Sensitivity Analysis for the ANN MDOT Asphalt Model

A sensitivity analysis was carried out to evaluate the performance of the developed ANN MDOT asphalt model. Using the preliminary GUI application, a random section was selected from the database and IRI prediction values were generated for nine consecutive years and compared with the observed IRI values. Figure 61 shows the observed vs. predicted plot of PCR for the MDOT asphalt pavement section #14.

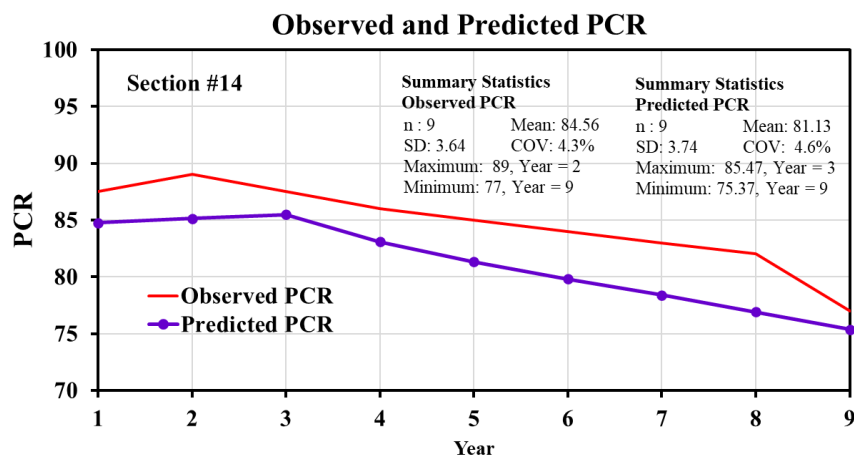


Figure 55. Observed and Predicted Plot of PCR for MDOT Asphalt Pavement Section #14

Predicted PCR values were close to the observed values. The predicted mean of PCR (81.13) was 4% lower than the observed mean of PCR (84.56), which showed that the PCR trend was captured by the developed model and the results were accurate and reliable for this section. However, the model performance may vary based on each section's characteristics, but the model will still follow a similar trend. Figure 56 shows the observed vs. predicted plot of  $IRI_{Mean}$  for the MDOT asphalt pavement section #14.

Predicted  $IRI_{Mean}$  values followed the trend embedded within the actual IRI data, increasing with time. The predicted mean of  $IRI_{Mean}$  (1.31) was 26.2% higher than the observed mean of  $IRI_{Mean}$  (1.04), which showed that the model was able to understand the deterioration behavior but

overpredicted most  $IRI_{Mean}$  values. However, the prediction results were reasonably accurate and reliable since the complex relationships among all variables are not easy to be predicted. It is also important to note that the model performance for  $IRI_{Mean}$  may vary based on each section.

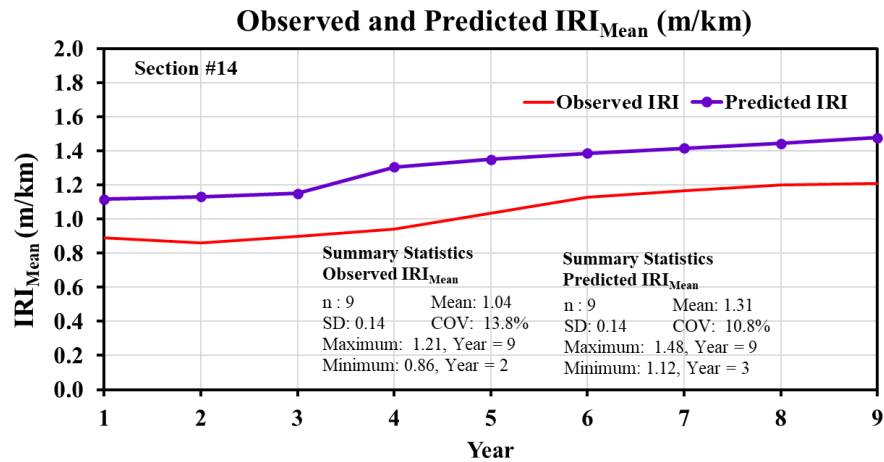


Figure 56. Observed and Predicted Plot of  $IRI_{Mean}$  for MDOT Asphalt Pavement Section #14

Therefore, the sensitivity analysis showed that MDOT can use the developed model as a tool for predicting future conditions of asphalt pavement sections and incorporate the M&R scheduling effectively to prioritize the resources. If no M&R actions are performed, the rate of deterioration increases exponentially, reducing the pavement life and leading to poor conditions that can be life-threatening for road users.

#### 10.4. Recommendations from the ANN MDOT Model

1. The developed model showed high accuracy for PCR predictions. However, further study on exploring independent variables and model architectures needs to be performed for improving IRI prediction accuracy.
2. A preliminary GUI was developed in this study. However, a more complete GUI needs to be developed for better visualization of asphalt pavement deterioration and to help users to

simulate different scenarios by utilizing different input values.

3. Use the developed GUI to perform another sensitivity analysis by modifying input values for selected variables to evaluate their effects on PCR and IRI predictions.

The MDOT study developed ANN models for asphalt pavement sections in Mississippi.

Further models for concrete and composite pavements are also developed in this doctoral research.

## CHAPTER XI: PERFORMANCE PREDICTION MODEL FOR JOINTED CONCRETE PAVEMENTS IN MISSISSIPPI USING MACHINE LEARNING

Rulian B., Ph.D. Candidate, Hakan Y., Ph.D., Salma S., Ph.D., and Yacoub N., Ph.D.

Department of Civil Engineering, University of Mississippi, University, MS 38677, United States  
rferrei1@go.olemiss.edu

(Paper published and accepted for poster presentation at the American Society of Civil Engineers (ASCE) International Conference on Transportation & Development (ICTD) 2022, Seattle, Washington, May 31-June 3, 2022.

### 11.1. Introduction

The Mississippi Department of Transportation (MDOT) is responsible for maintaining about 400 miles of jointed concrete pavement (JCP) statewide. An efficient Pavement Management System (PMS) is essential for transportation agencies to maintain adequate pavement conditions and allocate resources for future Maintenance and Rehabilitation (M&R) planning. One of the main problems that state transportation agencies encounter is the proper M&R of roads and highways to meet the public's needs and safety concerns [29]. Pavement performance models become a key part of PMS allowing decision-makers a better budget allocation plan for future M&R interventions [92]. Although pavement prediction models can forecast many variables, the International Roughness Index (IRI) and Pavement Condition Rating (PCR) are the most

significant variables in determining the condition of the pavement [30].

Traditionally, linear, non-linear, multiple linear regression analysis, Markov chains, mechanistic-empirical relations, survivor curves, semi-Markov, and Bayesian models have been used for predicting pavement performance [20]. The MDOT currently uses Markov probability matrices to support their M&R decisions; however, it does not consider M&R history, and key factors that affect pavement performance (i.e., traffic load, geographic location, and pavement design). Advanced modeling techniques using artificial neural networks (ANN) appear as an alternative for predicting pavement deterioration, offering significant improvements over traditional techniques [20,26]. Although the MDOT PMS includes all types of pavements, this paper will focus on the model development of JCP pavements using an ANN approach to develop a more, inclusive, reliable, and accurate prediction of future pavement conditions to perform effective and timely M&R interventions. The performance models will consider the effects of geographical location, traffic loads, pavement design, and M&R history. Furthermore, the developed models do not use any distress variables as an input, which makes the use of the prediction models easier for transportation agencies.

#### 11.1.1. Objectives

The main objectives of this paper are to:

- Analyze pavement performance data for JCP sections in the MDOT database.
- Identify key parameters responsible for affecting pavement performance.
- Develop pavement performance prediction models using the ANN approach.
- Evaluate the accuracy of ANN models using statistical measurements.
- Implement the best ANN model developed in the study for a random section.

## 11.2. Literature Review

Among the most important measures of pavement performance, the IRI and PCR are the most used and well-recognized pavement performance indicators [20]. Developed in 1982 by the World Bank and the government of Brazil at the International Road Roughness Experiment (IRRE), the IRI was created to be a standard roughness index. The IRI describes the irregularities in the pavement surfaces that affect the ride quality experienced by road users, and it is useful for making objective decisions related to the management of road networks [2,7,8]. Higher IRI values represent a rough pavement surface that indicates a lower ride quality for users, while lower IRI values indicate smooth pavements with better ride quality.

The PCR method provides a procedure for uniformly identifying and describing, in terms of severity and extent, pavement distress. The MDOT uses PCR as the main performance indicator to assess pavement conditions. Represented with a number from 0 to 100, higher PCR numbers represent better pavement conditions. MDOT groups pavements in good, fair, or poor condition (Figure 57) based on their PCR value [93]. The mathematical expression for PCR (Equation 22) gives an index reflecting the composite effects of varying distress types, severity, and extent upon the overall condition of the pavement [29].

$$PCR = 100 - \sum_1^n Deduct_i \quad Eq. 22$$

Where:  $n$  = number of observable distresses;  $Deduct_i$  = multiplication of the weight of distress, the weight of severity, and weight of extent for distress  $i$ .

**Poor Condition** - PCR of 71 or below.



**Fair Condition** - PCR of 72 – 81.



**Good Condition** - PCR of 82 or better.



Figure 57. MDOT PCR Scale [93]

Different modeling techniques such as linear, non-linear, multiple linear regression analysis, and Markov chains have been used for predicting pavement performance. The Hidden Markov was one of the most popular probabilistic models in the 1990s in pavement management systems [65]. Elhadidy et al. [90] used the Markov chain model for predicting pavement performance over the life span and developing a genetic algorithm to achieve optimum cost-benefit actions while Zhao and Guo [94] used the Markov chain to predict IRI values from 2005 to 2011 using data from test sections located on Highway 281 in the US. The use of simple statistical approaches such as linear regression has been used for modeling but does not seem appropriate to develop prediction models due to the complexity of the relations between each one of the variables that affect pavement performance. Advanced modeling techniques such as ANN appear as an alternative for predicting pavement deterioration, offering significant improvements over traditional techniques. The ANN is a predictive modeling technique that emulates the structure and/or functional aspects of biological neural networks. The ANN is capable of identifying complex relationships that are difficult for traditional sequential, logic-based modeling and computational techniques [43,44].

The ANN technique was successfully utilized to predict IRI values for flexible [2,18,69,70,83], concrete [9,17,34,36,65], and composite [10,26,31,68,92] pavements. Literature

review to date shows that ANN models performed successfully in predicting PCR and IRI by processing large volumes of data with a higher degree of accuracy. Therefore, there is a need for state transportation agencies to develop reliable, inclusive, and accurate pavement performance models using the state's database to create more advanced prediction tools. The prediction models need to be able to incorporate key factors such as geographical location, traffic loads, pavement design, and especially M&R history to assess the development of better budget allocation plans for future M&R actions. Furthermore, models that do not use distress variables are recommended for ease of use by federal and state agencies.

### 11.3. Model Development

#### 11.3.1. Methodology

The methodology for the model development is presented, as follows:

- Compile databases for JCP pavements from the MDOT database, including variables that affect pavement performance.
- Assess the quality of databases and identify missing/erroneous data items.
- Develop pavement performance models for JCP pavements using the ANN modeling technique.
- Evaluate the accuracy of the developed performance prediction models.
- Select the best-performing model based on statistical measures and verify the prediction behavior.
- Implement the selected performance models via Graphical User Interface (GUI).
- Evaluate the selected model using GUI for the enhancement of pavement asset management.

#### 11.3.2. Data Collection and Processing

The data used in this paper is a part of a pavement survey in the state of Mississippi performed by MDOT. Every two years, MDOT collects data to monitor the current pavement



conditions and predict M&R for the Mississippi road network. Four different pavement types are found in the database: Flexible, JCP, Continuous Reinforced Concrete Pavement (CRCP), and composite pavement. In this study, only JCP pavements were utilized to develop prediction models. Like other states, the MDOT operates a pavement management system that includes PCR, IRI, and distress data. Only data collected between 2010 and 2018 were used since the data collection method before 2010 was not consistent. A total of 101 JCP sections with 909 data points were utilized in the model development. Sections with missing or illogical data have been excluded as the ANN model development process needs a complete dataset. Since MDOT collects data every even year, to develop prediction models that are applicable for a 1-year increment, the odd year data were generated by averaging consecutive years from 2010 to 2020. This approach was successfully utilized for the MDOT database in previous studies [29,30,65].

The purpose of developing an ANN pavement performance model was to predict when M&R actions were needed and how it affects the roadway. It is known that not all rehabilitation actions were properly recorded and for this reason, a different approach for assigning rehabilitation actions was proposed based on the discussions with the state agency. To incorporate the effect of the rehabilitation on PCR and IRI, artificial rehabilitation actions based on significant changes in PCR and IRI have been assigned to the database. Threshold values for PCR and IRI were assigned based on the evaluation of data history. Several threshold values were studied [29,30,65] and optimum threshold values were found. If PCR increased 8% to 12% and IRI decreased 5% to 16% in a year compared to the previous measurement, a minor rehabilitation was assumed to take place in that year. If PCR increased above 12% and IRI decreased more than 16%, a major rehabilitation was assumed. If none of these situations occurred, it was assumed no rehabilitation.

### 11.3.3. ANN Structure and Modeling Process

In this study, a single layer feed-forward neural network with a back-propagation training algorithm was used for the development of prediction models for JCP pavements. The neural network gains its knowledge through a trained feed-forward network that uses a set of training data consisting of inputs and output(s). The resulting output is compared to the target values and the back-propagation process adjusts the connection weight to reduce the error between actual and target values [44]. One hidden layer was considered in the model development. After training, the network provides an approximate functional mapping of any input pattern onto its corresponding output pattern. Then, the validation process was carried out using datasets that were excluded from the model database. After the validation process, it is necessary to retrain the best-performing network using all experimental data to increase the prediction accuracy and account for all patterns in the database [44]. All ANN models were trained with 50%, tested with 25%, validated with 25% of the data, and finally, retrained using the best-performing network with the full dataset. The TR-SEQ1 computer program [55] was used to develop the ANN models and a sigmoidal function was utilized for data generalization purposes.

For the JCP performance prediction modeling, several input variables were selected after an extensive literature review and consultation with MDOT personal to identify what parameters were significant to the agency. Different inputs were tried in a preliminary study using a trial-and-error method to select the most significant variables considering the practical point of view. Different models were developed varying the numbers of independent and dependent variables in a preliminary study using a trial-and-error method to select the most significant variables considering the practical point of view and identify the optimum modeling structure. The four best developed models are presented in this paper and the variables used in these models are explained,

as follows:

- Beginning Longitude and Latitude: coordinates to indicate the initial location of the roadway section.
- Ending Longitude and Latitude: coordinates to indicate the end of the roadway section.
- Thickness: indicates the concrete thickness in the pavement section in inches.
- Length of the Section: length of the section recorded in miles.
- Age of the Section in 2010: shows the section's age since the earliest available pavement measurement was recorded.
- PCR in 2010: shows the initial PCR value in 2010 to indicate the base starting value.
- IRI in 2010: shows the initial IRI in 2010 to indicate the baseline value to represent the pavement initial condition.
- Time since 2010: represents the time since 2010 to the desired prediction year. This value is associated with the effects of pavement aging.
- Minor Rehabilitation: categorical variable to represent minor rehabilitation. Use "1" if PCR increased 8% to 12% and IRI decreased 5% to 16% in a year compared to the previous measurement. If not, use "0".
- Major Rehabilitation: categorical variable to represent major rehabilitation. Use "1" if PCR increased above 12% and IRI decreased more than 16%. If not, use "0".
- Equivalent Single Axle Load (ESAL): ESAL in that specific year.
- Cumulative Equivalent Single Axle Load (CESAL): CESAL in that specific year.
- PRE PCR: variable used for dynamic ANN models to indicate the PCR from the previous years that will be used to predict the actual year.
- PRE IRI: variable used for dynamic ANN models to indicate the IRI from the previous years that will be used to predict the actual year.
- IRI: International Roughness Index measured in that year.
- Complementary PCR: this variable is calculated by subtracting 100 - PCR. This parameter is generated because the outputs need to be directly proportional in the ANN modeling. Since PCR and IRI are normally inversely proportional over time, it was necessary to utilize the complementary PCR approach. The use of this variable assisted the model to optimize the network calculations providing better accuracy between observed and

predicted values.

A total of four models were developed using the same JCP database. All the inputs and outputs used in these models are listed in Table 25.

Table 25. Independent and Dependent Variables for ANN Models

Models	Model 1	Model 2	Model 3	Model 4
Independent Variables	Begin Lat.	Begin Lat.	Begin Lat.	Begin Lat.
	Begin Long.	Begin Long.	Begin Long.	Begin Long.
	End Lat.	End Lat.	End Lat.	End Lat.
	End Long.	End Long.	End Long.	End Long.
	Thickness	Thickness	Thickness	Thickness
	Length	Length	Length	Length
	Age in 2010	Age in 2010	Age in 2010	Age in 2010
	PCR in 2010	PCR in 2010	PCR in 2010	PCR in 2010
	IRI in 2010	IRI in 2010	IRI in 2010	IRI in 2010
	Time (t)	Time (t)	Time (t)	Time (t)
	IRI Minor	IRI Minor	IRI Minor	IRI Minor
	IRI Major	IRI Major	IRI Major	IRI Major
	ESAL	CESAL	ESAL	ESAL
	PRE PCR	PRE PCR	PRE PCR	
	PRE IRI	PRE IRI		PRE IRI
Dependent Variables	Complementary PCR	Complementary PCR	Complementary PCR	
	IRI	IRI		IRI

The independent and dependent variables for the models in this study were obtained after examining previous studies that have also used the MDOT database [29,30,65]. Models 1 and 2 utilized 15 independent variables and two outputs (IRI and complementary PCR). To incorporate the effects of traffic in the pavement deterioration process, Model 1 used ESAL while Model 2 used CESAL. The idea of using CESAL instead of ESAL is to introduce the cumulative history of traffic loads since the first recorded measurement for more inclusive predictions. Models 3 and 4 included 14 independent variables and one output variable each, Complementary PCR and IRI respectively. The purpose of using separate outputs is to identify if the model will have a better performance when using a two-output model or individual outputs. The use of variables that are not related to distress data makes the models more accessible for transportation agencies since

most of these variables are easily available in their databases. Therefore, the models developed in this paper bring a new and valuable tool for the MDOT’s pavement management system.

#### 11.3.4. ANN Model Selection

The best model was selected based on the lowest Average Square Error (ASE), Mean Absolute Relative Error (MARE), and highest Coefficient of Determination ( $R^2$ ). The maximum and minimum values of each independent variable were included in the training phase for the network to represent the characteristics of the response. The maximum and minimum ranges of each input/output variable for ANN model development were chosen on purpose to be wider than their actual ranges for better mathematical mapping [44]. The model network is written in an order that depicts the number of initial hidden nodes, the number of final hidden nodes, and iterations. The final structure of each model is written at the bottom row in an order that depicts the number of inputs, hidden nodes, and output(s), respectively. Table 26 shows the comparison of Models 1, 2, and 4 for the complementary PCR output and Models 1, 2, and 3 for the IRI output.

Table 26. ANN Model Results for JCP Pavements (Complementary PCR and IRI Outputs)

Output		Complementary PCR			IRI		
Model		Model 1	Model 2	Model 4	Model 1	Model 2	Model 3
Network		5-19-4100	8-8-6200	12-19-20000	5-19-4100	8-8-6200	13-15-20000
Training	MARE	3.15	3.99	3.45	4.34	7.58	4.84
	$R^2$	0.96	0.89	0.95	0.97	0.93	0.97
	ASE	0.00018	0.00048	0.00021	0.0000017	0.0000045	0.0002055
Testing	MARE	7.59	5.05	10.37	8.75	7.24	8.02
	$R^2$	0.65	0.86	0.52	0.72	0.83	0.77
	ASE	0.00146	0.00052	0.00255	0.0000100	0.0000092	0.0008246
Validation	MARE	6.97	6.72	10.10	11.53	12.32	12.47
	$R^2$	0.65	0.67	0.41	0.79	0.76	0.76
	ASE	0.00118	0.00139	0.00199	0.0000112	0.0000135	0.0012898
All data	MARE	3.51	4.16	4.42	5.73	7.93	5.97
	$R^2$	0.93	0.85	0.89	0.95	0.91	0.94
	ASE	0.00029	0.00061	0.00048	0.0000031	0.0000056	0.0003290
Final Structure		15-19-2	15-8-2	14-19-1	15-19-2	15-8-2	14-15-1

For the complementary PCR output, the training stage showed that Model 1 had the lowest

ASE and MARE values and the highest  $R^2$ . In the testing stage, Model 2 showed better accuracy measures in terms of ASE, MARE, and  $R^2$ . In the validation stage, Model 1 showed better accuracy measures in terms of ASE, but Model 2 showed better MARE and  $R^2$  values. In the-all data stage, the 909 data points were used to retrain the network at its optimal structure and iteration to obtain the generalized response throughout the complete database. The all-data stage results show that Model 1 outperformed all other models with better accuracy measures for ASE, MARE, and  $R^2$ . Statistical measures of Model 1 retrained with all-data showed an ASE of 0.00029, a MARE of 3.51, and an  $R^2$  of 0.93.

For the IRI output, Model 1 showed better accuracy measures for training and validation stages, while Model 2 showed better results for testing. In the all-data stage, Model 1 outperformed all other models with better accuracy measures for ASE, MARE, and  $R^2$ . Statistical measures of Model 1 trained with all-data showed an ASE of 0.0000031, a MARE of 5.73, and an  $R^2$  of 0.95.

Therefore, Model 1 was selected as the best-performing ANN model for MDOT JCP pavement. The network structure of Model 1 includes 15 input variables, 1 hidden layer with 19 hidden nodes, 4,100 iterations, and 2 outputs.

#### 11.4. Discussion

##### 11.4.1. ANN Model Results

Figure 58 shows a graphical comparison of statistical measures for Complementary PCR and IRI outputs using all-data stage.

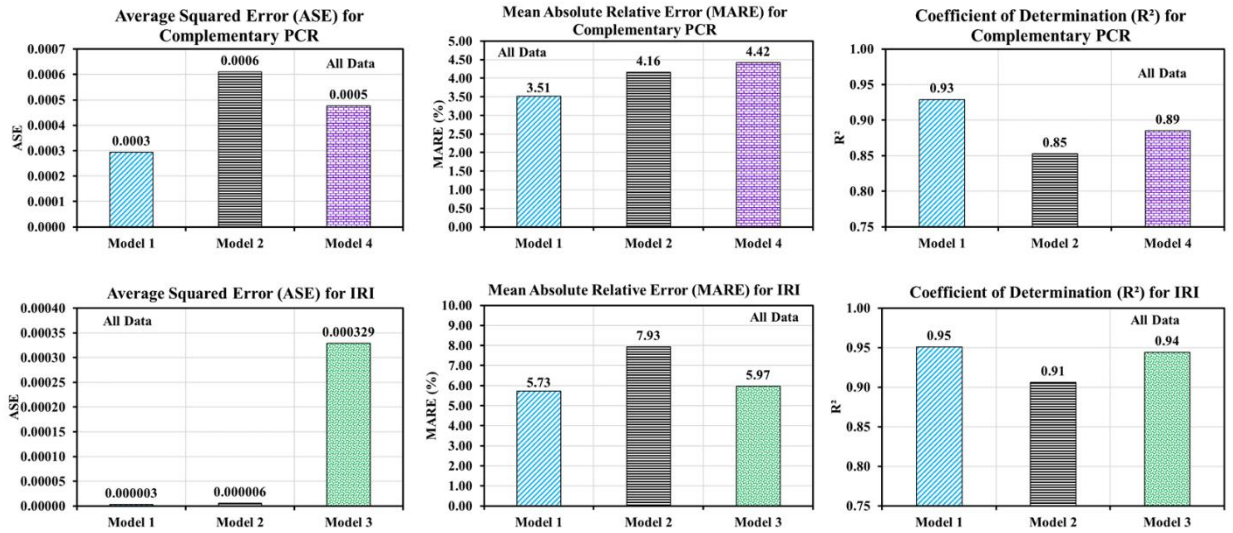


Figure 58. Graphical Comparison of Accuracy Measures for All Developed Models

Figure 58 shows that all four developed models presented reliable results. However, Model 1 outperformed all other models with more accurate predictions. For the complementary PCR output, Model 1 had an ASE 52% and 38% lower, a MARE 16% and 20% lower, and an R<sup>2</sup> 9% and 5% higher than Models 2 and 4, respectively. For the IRI output, Model 1 had an ASE 45% and 99% lower, a MARE 28% and 4% lower, and an R<sup>2</sup> 5% and 1% higher than Models 2 and 3, respectively. These results reaffirm the selection of Model 1 as the best-performing model for the MDOT JCP pavement. Figure 59 shows the observed PCR and IRI values collected from the MDOT database and the predicted PCR and IRI values using Model 1.

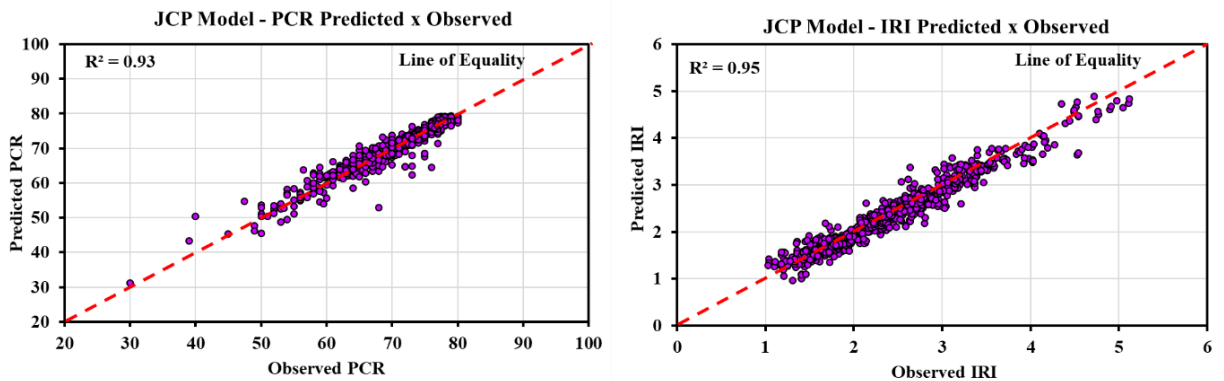


Figure 59. Observed vs. Predicted PCR and IRI for Model 1

Model 1 predictions clustered around the line of equality and showed high accuracy when predicting PCR and IRI. The model was able to capture the deterioration behavior of the JCP pavements resulting in an excellent agreement ( $R^2 = 0.93$  and  $R^2 = 0.95$ ) between observed and predicted values for PCR and IRI, respectively. These results indicate that the model was capable to understand the complex relationships between input and output variables translating them into reliable predictions. The use of CESAL did not improve the model accuracy since Model 2 could not outperform Model 1, which has used ESAL to incorporate traffic history in the model development. The two-output model approach used in Model 1 also proved to be more efficient for PCR and IRI predictions than using the traditional one-output model. The use of both PCR and IRI data assisted the model to achieve higher accuracy. However, changing the PCR output to Complementary PCR was important for the model's success. Since PCR and IRI performance measures are negatively correlated, the ANN technique performs better when the outputs are positively correlated. Thus, the use of this approach was essential to assist the network to optimize the model predictions.

Therefore, all developed models showed to be accurate and reliable tools to be used by MDOT, but Model 1 showed to be the best-performing model for predicting PCR and IRI values for JCP pavement.

#### 11.5. Implementation of Model 1 via GUI

A GUI was developed in Microsoft Excel using the connections weights and threshold values for Model 1. The user enters the input variables, and the application automatically computes the PCR and IRI values for the selected year as shown in Figure 60. The appropriate ranges for the input values are shown in the application, if any of the provided input values are outside the applicable range, it may cause the model to generate unreliable predictions.



Section #	INPUTS								OUTPUTS	
	Begin Lat.	Begin Long.	End Lat.	End Long.	Thickness	Length	Age @ 2010	PCR @ 2010	PCR	IRI
903	32.353	-90.629146	32.342155	-90.604198	203.2	1.662	41.0	75.0	<b>75.4</b>	<b>2.01</b>
	<b>IRI @ 2010</b>	<b>Time (t)</b>	<b>Minor</b>	<b>Major</b>	<b>ESAL</b>	<b>Dynamic PCR</b>	<b>Dynamic IRI</b>			
	2.010	1	0	0	289737	25	2.0			

USER RANGES								
Variables	Begin Lat.	Begin Long.	End Lat.	End Long.	Thickness	Length	Age @ 2010	PCR @ 2010
<b>Minimum</b>	30.299	-91.066853	30.298663	-91.067662	76.2	0.099	3.0	46.0
<b>Maximum</b>	34.824559	-88.11399	34.82477	-88.113945	558.8	10.097	78	79
Variables	IRI @ 2010	Time (t)	Minor	Major	ESAL	Dynamic PCR	Dynamic IRI	
<b>Minimum</b>	1.33	1	0	0	76.65	20	1.03	
<b>Maximum</b>	4.53	9	1	1	390110.175	61	5.12	

Figure 60. GUI Application for Model 1

To implement Model 1, the GUI application was utilized to predict PCR and IRI values for a random section of the database. This analysis intends to compare the accuracy of the prediction values generated by the ANN model and the observed values collected from the database. This comparison is important to identify the model's capability to understand pavement responses under different input variables. Figure 61 shows the comparison between age predicted and observed PCR and IRI values for Section #903.

Predicted PCR values were very close to the observed values. The predicted mean of PCR (75.73) was 0.08% higher than the observed mean of PCR (75.67), which confirmed that the ANN model was able to capture the PCR changes over time and generate reliable predictions for Section #903. Predicted IRI values also followed the increase and decrease of IRI data over time. The predicted mean of IRI (1.97) was 0.38% lower than the observed mean of IRI (1.98), showing that the roughness deterioration behavior was successfully identified by the ANN model. However, it is known that the model performance may vary based on each section's characteristics, but a similar trend will be followed by the model.

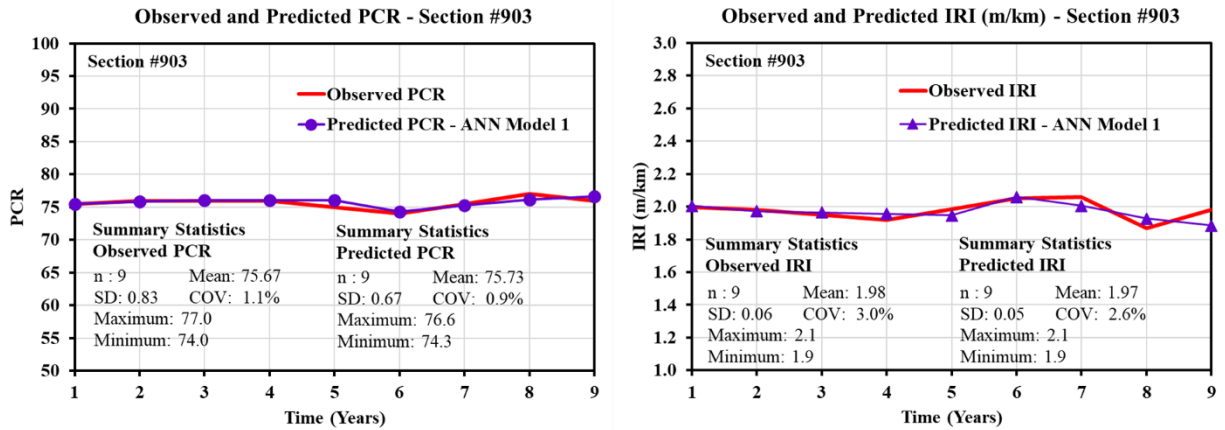


Figure 61. Observed and Predicted Plot of PCR and IRI for JCP Pavement Section #903

Therefore, the implementation of Model 1 for a random section of the database showed that MDOT can use the developed model as a key tool for predicting future conditions of JCP pavement sections and incorporating the M&R scheduling effectively to prioritize the resources for the most needed sections. Timely M&R interventions are necessary to keep roads in acceptable conditions, prevent early pavement deterioration, increase pavement life, and provide a more comfortable ride for road users.

## 11.6. Conclusion

This study attempted to develop pavement performance models utilizing the ANN approach for JCP sections in the MDOT database. Four models were developed using different input and output variables. Model 1 outperformed all other models for the complementary PCR output with an ASE 52% and 38% lower, a MARE 16% and 20% lower, and an  $R^2$  9% and 5% higher than Models 2 and 4, respectively. For the IRI output, Model 1 also showed to be the best-performing model with an ASE 45% and 99% lower, a MARE 28% and 4% lower, and an  $R^2$  5% and 1% higher than Models 2 and 3, respectively. Model 1 was able to explain 93% and 95% of the variability for complementary PCR and IRI, respectively. These results indicate that the developed model was able to understand the complex relationships between input and output

variables capturing the deterioration behavior of JCP pavements for both PCR and IRI performance indicators. The use of ESAL and a two-output model (Complementary PCR and IRI) helped the ANN model to generate more accurate predictions compared to the use of CESAL and the traditional one output models. Furthermore, the use of variables that are not related to distress data makes the models more accessible for agencies since these variables are easily available in their databases. The final network structure for Model 1 includes 15 input variables, one hidden layer with 19 hidden nodes, 4,100 iterations, and two outputs.

A GUI application was developed in Microsoft Excel for the implementation of Model 1 in a random section of the database. Predicted mean values for PCR and IRI were very close to the observed values for Section #903. Predicted PCR and IRI values were only 0.08% higher and 0.38% lower compared to their observed means, respectively. These results confirm the reliability of Model 1 to generate PCR and IRI predictions. Therefore, the ANN performance prediction model developed in this paper provides a new, more reliable, inclusive, and accurate tool to assist MDOT to predict future JCP pavement conditions and incorporate the M&R scheduling effectively to develop a better budget allocation plan and a more effective PMS.

## CHAPTER XII: PERFORMANCE EVALUATION OF COMPOSITE PAVEMENTS ON MISSISSIPPI HIGHWAYS VIA MACHINE LEARNING

R. Barros, H. Yasarer, W. Uddin, S. Sultana & Y. Najjar

Department of Civil Engineering, University of Mississippi, Oxford, Mississippi, United States

(Paper accepted for session presentation and published at the Eleventh International Conference on the Bearing Capacity of Roads, Railways and Airfields (BCRRA), Volume 2–Hoff, Mork & Saba(eds), ISBN 978-1-032-12049-2, Trondheim, Norway, 27-30 June 2022)

### 12.1. Abstract

Local, state, and federal highway agencies run some form of maintenance and rehabilitation program during the design life of highways. Due to budgetary restrictions, maintenance and rehabilitation actions must be prioritized based on the future condition of the highway section. There are important factors that affect the performance of highways. To properly assess the condition of the pavement and operate maintenance, prediction models with significant condition variables are essential. Mississippi Department of Transportation (MDOT) utilizes probability-based prediction models to determine which sections of the highway and when they need rehabilitation. The current probability models predict the performance without the section-specific parameters. The goal of this study is to develop a new set of performance prediction models for the composite Pavements in Mississippi by using machine learning. The

best-performing model can be used as a simple and user-friendly tool to allow the user to visualize the future projections of the pavement section. MDOT personnel can employ this application to predict the condition of the composite pavement section and prioritize the maintenance and rehabilitation schedule.

Keywords: Artificial Neural Network (ANN), machine learning, Mississippi Department of Transportation (MDOT), pavement performance, composite pavement.

## 12.2. Introduction

Road networks are one of the largest public infrastructure assets of a country, they provide public mobility and freight transport to secure the nation's economy and prosperity [36]. Annually, transportation agencies spend billions of dollars for the maintenance and rehabilitation (M&R) of their road networks. One of the most difficult tasks for state transportation agencies, such as the Mississippi Department of Transportation (MDOT), is to maintain roads and highways in acceptable conditions to meet the public's needs and safety concerns [28]. If timely M&R is not performed, it may lead the pavement to poor conditions, causing an uncomfortable ride experience for road users [31]. To assess pavement surface condition and ride quality, Performance Condition Rating (PCR) and International Roughness Index (IRI) are the two widely used measures worldwide. Pavement performance models assist agencies to predict how a pavement deteriorates over time due to traffic, environmental conditions, and M&R history, being an important part of the pavement management system (PMS). The use of prediction models allows decision-makers to develop a better budget allocation plan and M&R schedule [26].

Flexible and rigid pavements are the main concern of most studies in the literature. However, a portion of the paved surfaces is comprised of composite pavements, which are made

of an asphalt overlay on concrete pavements [88]. When concrete pavements start to fail, they are overlaid with a layer of a hot mix of asphalt (HMA) to provide better levels of performance being a more cost-effective alternative [3]. Although a PMS include all types of pavements, the study of composite pavements has not been well investigated in the literature, which is the focus of this paper.

A growing body of studies has been exploring several modeling techniques and variables to reach accurate predictions. The use of an Artificial Neural Network (ANNs) technique has shown significant improvements over traditional methods, such as regression, by processing large volumes of data with a higher degree of accuracy. However, present pavement performance prediction models did not account for the influence of some essential parameters such as pavement design, rehabilitation interventions, and traffic in the model development, which affects the accuracy of the predictions [88].

In this study, composite pavement performance models were developed for the Mississippi road network using the ANN approach to predict the future condition of pavement sections incorporating the influence of pavement structure and design, traffic, and M&R interventions. The data used for this project was provided by the MDOT, and only data collected between 2010 and 2020 were analyzed. Several models were created to predict PCR and IRI. Over time, this model can be improved further in the future for the prediction of pavement conditions as more data is accurately measured and added to the existing database.

#### 12.2.1. Objectives

The major objectives of this paper are to:

- (1) Analyze data for composite pavement sections in the MDOT database.
- (2) Use the ANN technique to develop performance models for composite pavements in the MDOT database using different independent variables.

- (3) Evaluate the accuracy of the developed models using statistical measurements to identify the most accurate model.

### 12.3. Literature Review

There has been an increased interest in the use of machine learning approaches in different fields of engineering. Several types of machine learning algorithms have been developed and used to process large volumes of data with high degrees of accuracy, handle noisy and complex data, solve non-linear problems, and once trained, make predictions and generalizations at any time. The machine learning techniques hold significant potential for building a modern and robust pavement system management due to the excellence in automation and pattern recognition [20,39]. The literature review shows that ANN is not only one of the first machine learning techniques to be used, but also the most used technique in civil and pavement engineering [40,41]. The ANN method attempts to emulate the structure and/or functional aspects of biological neural networks [44]. Complex relationships that are difficult to reproduce using traditional sequential, logic-based modeling and computation techniques can be successfully represented by neural networks. However, the accuracy of ANN models depends on the accuracy of the database used to train the neural network.

Several studies have used ANN to predict pavement performance. Prediction of IRI values for flexible pavements in the wet freeze region and rigid pavements in the wet non-freeze region of the Long-Term Pavement Performance (LTPP) database achieved reasonable accuracy for both short-term and long-term predictions [16,17]. Models for Jointed Plain Concrete Pavement (JPCP), Jointed Reinforced Concrete Pavement (JRCP), and Continuously Reinforced Concrete Pavement (CRCP) were developed with a high degree of accuracy presenting better accuracy compared to models that currently use mechanistic-empirical pavement design or multiple regression [9]. ANNs were also used to predict pavement performance of flexible and rigid roads in Mississippi

achieving satisfactory results [29,30]. Because machine learning has a data-driven approach, PCR and IRI appear as suitable indicators for modeling, as they are widely available in pavement databases and known as one of the most common indicators for pavement performance evaluation.

## 12.4. Model Development

### 12.4.1. Methodology

The database for the ANN model development is obtained after cleansing and reorganizing the raw data files. Due to new data acquisition methods, materials, and tools used by the MDOT, only data collected from 2010 to 2020 are included in the model development. Sections with missing or illogical data have been excluded as the ANN model development process needs a complete dataset. Each section is comprised of five different datasets based on PCR, IRI, and rehabilitation actions. In Mississippi, the distress data is collected every even year. To develop prediction models that are applicable for a 1-year increment, the odd year data was needed to be generated by averaging consecutive years from 2010 to 2020 [28,65]. The sections of the roads with JCP varied in length from 0.04 miles to 15.9 miles and includes variables such as beginning and ending latitudes and longitudes, the thickness of pavement layer, initial PCR, initial IRI, minor and major rehabilitation, and traffic. The pavement age was included in the model development.

MDOT conducts a survey to monitor the actual condition of Mississippi roads every two years. A database was generated with all data collected over the years and combined to develop the performance prediction model for composite pavements. Additional information regarding maintenance and rehabilitation history were also incorporated into the model's database. An independent variable is added to indicate that an M&R was performed on a particular section so it can be utilized in the performance model. By doing this, the ANN can learn the effect of a maintenance and rehabilitation intervention in the pavement section and provide more accurate



PCR and IRI predictions. However, the MDOT did not keep track of all rehabilitation actions and for this reason, they were assigned based on a few criteria. Improvement of PCR and IRI values without any rehabilitation action was considered irrational. Some uncertainty due to the calibration of the profilometer, systematic errors, and the environmental conditions on the day of the survey may have resulted in some of the irrational condition measures. To incorporate the effect of the rehabilitation on PCR and IRI, artificial rehabilitation actions based on the significant changes in PCR and IRI have been assigned to the database [65]. Threshold values for PCR and IRI were assigned based on the evaluation of data history. If PCR increased 8% to 12% and IRI decreased 5% to 16% in a year compared to the previous measurement, a minor rehabilitation was assumed to take place in that year. If PCR increased above 12% and IRI decreased more than 16%, a major rehabilitation was assumed. If none of these situations occurred, it was assumed no rehabilitation.

#### 12.4.2. ANN Development

This study uses a feedforward neural network with a back-propagation training algorithm for the development of the performance prediction model. A one hidden layer network was considered in the model development, as the use of more than one hidden layer may cause the network to memorize the data in the training phase [44]. For the development of the composite performance models 10,305 data points from 1,145 pavement sections throughout the state of Mississippi were used. In this study, the TR-SEQ1 computer program [55] was used to develop the ANN models. A total of three different models have been developed by changing selected variables to identify the optimum model. For each of the three models that were created for this study, one network was chosen to be the best performing using three statistical measures.

#### 12.4.3. ANN Model Architecture

For this paper, three models were developed using different independent variables and PCR

and IRI as dependent variables. Table 27 shows the variables used for each ANN model.

Table 27. Independent and Dependent Variables Configuration

Models	Model 1	Model 2	Model 3
Independent Variables	Begin Lat.	Begin Lat.	Begin Lat.
	Begin Long.	Begin Long.	Begin Long.
	End Lat.	End Lat.	End Lat.
	End Long.	End Long.	End Long.
	Thickness	Thickness	Thickness
	Length	Length	Length
	Age in 2010	Age in 2010	Age in 2010
	PCR in 2010	PCR in 2010	PCR in 2010
	IRI in 2010	IRI in 2010	IRI in 2010
	Time (t)	Time (t)	Time (t)
	Drainage	Drainage	Drainage
		Minor Rehabilitation	Minor Rehabilitation
		Major Rehabilitation	Major Rehabilitation
			ESAL
		PRE PCR	PRE PCR
	PRE IRI	PRE IRI	
Dependent Variables	Complementary PCR	Complementary PCR	Complementary PCR
	IRI	IRI	IRI

The three models created in this study have the same core of inputs: beginning latitude and longitude, ending latitude and longitude, the thickness of the top layer of pavement, length of pavement, pavement age at the year 2010, PCR at the year 2010, IRI at the year 2010, time counted from 2010, drainage, pre-PCR, and pre-IRI. Minor and major rehabilitations and a traffic variable were utilized in different models to study the significance of each variable. The time ranges from 1 to 9 because there are only 9 years of past data available for this study.

Model 1 utilized 13 independent variables and did not utilize any M&R history and traffic variables in its development. Model 2 included 15 independent variables, incorporating M&R history, but not using a traffic variable. Model 3 included 16 independent variables incorporating the effects of minor and major rehabilitation and traffic in the model. A two-output model with IRI and complementary PCR was utilized for all developed models. The two outputs (i.e., PCR and IRI) were modified to be used with the complementary PCR (i.e., 100-PCR), as the outputs

need to be directly proportional in the ANN modeling. Because PCR and IRI usually change inversely over time, it was necessary to utilize the complementary PCR and IRI, as both change proportionally over time. Therefore, the use of complementary PCR assisted the network to optimize the model with better accuracy and to establish a superior correlation between actual and predicted outputs. The use of variables that are not related to distress data makes the models more accessible for transportation agencies, as most of these variables are easily available in their databases. Therefore, this model brings a valuable tool for the MDOT's pavement management system.

## 12.5. Results And Discussion

### 12.5.1. ANN Model Selection

The best model was chosen based on how close the prediction of the model is with the existing data. The three best-performing models were selected based on the lowest average square error (ASE), lowest mean absolute relative error (MARE), and highest coefficient of determination ( $R^2$ ). Table 28 shows the statistical measures of each ANN model development stage (i.e., training, testing, validation, and all-data) for the complementary PCR output.

For the complementary PCR output, the training stage showed that Model 3 indicated better accuracy measures in terms of ASE, MARE, and  $R^2$ . In the testing stage, Model 2 had the lowest ASE and MARE values and the highest  $R^2$ . In the validation stage, Model 2 had the lowest ASE and highest  $R^2$  while Model 3 showed the lowest MARE. In the all-data stage, the 1,930 data points were used to retrain the network at its optimal structure and iteration to obtain the generalized response throughout the complete database. The all-data stage results show that Model 2 and Model 3 had good accuracy with low ASE and MARE values and an  $R^2$  around 0.83. However, Model 1, the only model that did not use any variable related to M&R actions demonstrated lower

accuracy compared to the other models that incorporated the effects of rehabilitation interventions. Statistical measures of Model 2 trained with all data showed to be the best model for the complementary PCR with an ASE of 0.0007808, a MARE of 7.566, and an R<sup>2</sup> of 0.833.

Table 28. ANN Model Results for Complementary PCR Output

		Model 1	<b>Model 2</b>	Model 3
Complementary PCR		9-12-20000	<b>16-19-20000</b>	8-18-16000
Training	MARE	20.84	<b>8.13</b>	7.66
	R <sup>2</sup>	0.672	<b>0.819</b>	0.838
	ASE	0.0029880	<b>0.0008506</b>	0.0007635
Testing	MARE	21.89	<b>8.22</b>	8.35
	R <sup>2</sup>	0.575	<b>0.805</b>	0.785
	ASE	0.0035952	<b>0.0009134</b>	0.0010196
Validation	MARE	19.19	<b>8.55</b>	8.34
	R <sup>2</sup>	0.693	<b>0.779</b>	0.739
	ASE	0.0026957	<b>0.0010106</b>	0.0011974
All data	MARE	21.971	<b>7.566</b>	7.926
	R <sup>2</sup>	0.654	<b>0.833</b>	0.827
	ASE	0.0034965	<b>0.0007808</b>	0.0008096

Table 29 shows the comparison of each model for the IRI output.

Table 29. ANN Model Results for IRI Output

		Model 1	<b>Model 2</b>	Model 3
IRI		9-12-20000	<b>16-19-20000</b>	8-18-16000
Training	MARE	23.84	<b>12.85</b>	12.11
	R <sup>2</sup>	0.805	<b>0.825</b>	0.836
	ASE	0.0000171	<b>0.0000092</b>	0.0000084
Testing	MARE	25.97	<b>13.96</b>	13.89
	R <sup>2</sup>	0.715	<b>0.763</b>	0.757
	ASE	0.0000249	<b>0.0000110</b>	0.0000113
Validation	MARE	21.80	<b>14.03</b>	13.25
	R <sup>2</sup>	0.763	<b>0.747</b>	0.726
	ASE	0.0000183	<b>0.0000112</b>	0.0000123
All data	MARE	19.823	<b>12.147</b>	12.333
	R <sup>2</sup>	0.781	<b>0.819</b>	0.816
	ASE	0.0000157	<b>0.0000088</b>	0.0000090

For the IRI output, the training stage showed that Model 3 demonstrated better accuracy measures in terms of ASE, MARE, and  $R^2$ . In the testing and validation stages, Model 2 had the lowest ASE and highest  $R^2$ , while Model 3 had the lowest MARE. The all-data stage results show that all models had good accuracy with really low ASE values and high  $R^2$  values that vary from 0.781 to 0.819. It is important to notice that the least accurate model (Model 1) was the only one that did not incorporate M&R history in its development. Model 2 and Model 3 demonstrated better accuracy when using minor and major rehabilitation variables. Statistical measures of Model 2 trained with all data showed to be the best model for IRI with an ASE of 0.0000088, a MARE of 12.147, and an  $R^2$  of 0.819.

Therefore, Model 2 was chosen as the best performing ANN model for composite pavements in the MDOT database. The network structure of Model 2 includes 15 input variables, 1 hidden layer with 19 hidden nodes, 20,000 iterations, and 2 outputs. Figure 62 shows the observed PCR values collected from the MDOT database versus the predicted values using all developed models.

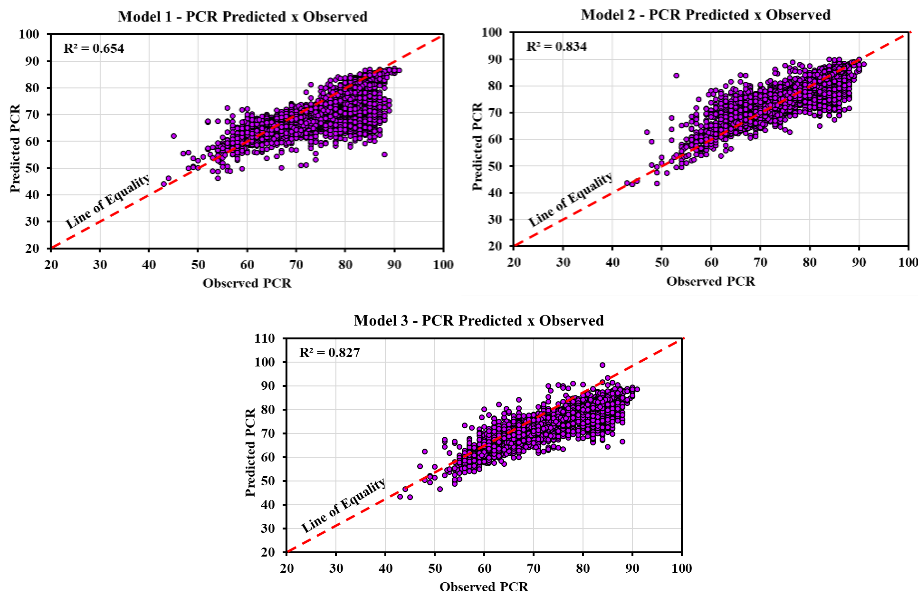


Figure 62. Observed Versus Predicted PCR for All Developed Models

PCR predictions clustered around the line of equality and showed that the PCR behavior was captured by all the prediction models. However, some prediction values for Model 1 did not follow closely the line of equality, which explains the lower  $R^2$  value compared to other models. Model 2 and Model 3 showed a high  $R^2$  of 0.834 and 0.827, respectively, but it is possible to observe that Model 2 prediction values were closer to the line of equality. Figure 63 shows the observed IRI values collected from the MDOT database versus the predicted values using all developed models.

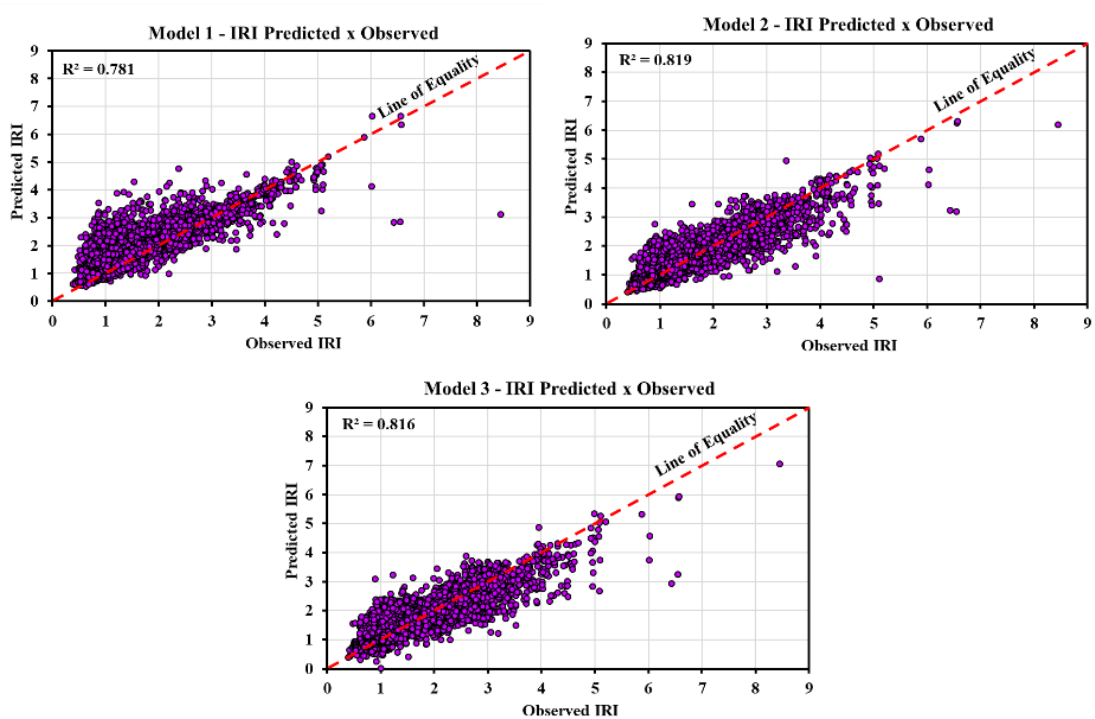


Figure 63. Observed Versus Predicted IRI for All Developed Models

IRI predictions followed the line of equality and showed reliable results for all developed models. However, Model 1 demonstrated that several predicted values were overpredicted, showing higher IRI values, which reduced the model accuracy. Model 2 and Model 3 showed similar results where the models could capture the IRI behavior, but Model 2 showed a better trend for predicted values and was chosen as the best performance model for composite pavements in

the MDOT database.

Therefore, all developed models were able to imitate the pavement deterioration behavior. However, the model that did not use variables that incorporate the effect of M&R actions showed lower accuracy compared to other models. The results show the importance of considering maintenance and rehabilitation history when developing performance models, especially for composite pavements, which is the focus of this study.

## 12.6. Conclusions

In this study, an artificial neural network approach was used to develop performance prediction models for composite pavements in Mississippi from the MDOT database. Three different models were developed using different input variables and a two-output structure. All models had satisfactory results and could be used for generating reliable predictions for pavement performance. However, the only model that did not consider the effects of maintenance and rehabilitation history (Model 1) showed to be the least accurate. Models that considered M&R (Model 2 and Model 3) showed better predictions. The best-performing model was selected based on three accuracy measures shown in Table 28 and Table 29. Model 2 outperformed all other models in the all-data stage with lower ASE, MARE, and higher  $R^2$  for IRI and PCR outputs. The final ANN model includes 15 input variables, one hidden layer with 19 hidden nodes, 20,000 iterations, and two outputs. Figure 62 and Figure 63 show predicted versus observed values for PCR and IRI for all developed models. It is possible to observe that the developed ANN models were able to capture the pavement deterioration behavior over time, but Model 2 and Model 3 had better accuracy than Model 1. Predicted values clustered around the line of equality, showing good agreement between observed and predicted values. Therefore, the development of performance models that include maintenance and rehabilitation variables provided a more accurate, inclusive,

reliable, and realistic model for transportation agencies. The developed model can support objective decisions regarding maintenance and rehabilitation interventions and budget plans permitting state agencies to prioritize the resources for critical pavement sections. Furthermore, the best-performing model can be used as a simple and user-friendly tool to allow the user to visualize the future projections of the pavement section. MDOT personnel can employ this application to predict the condition of the composite pavement section and prioritize the maintenance and rehabilitation schedule.



## CHAPTER XIII: GRAPHICAL USER INTERFACE (GUI)

### 13.1. Graphical User Interface Development

Several papers presented in this dissertation was developed with GUIs to implement the utilization of the best ANN models. In this section, a complete GUI is presented as an example to show the utilization of the model and better visualization of pavement deterioration, understanding of the effects of M&R, simulation of different scenarios by utilizing different input values, and support objective decisions regarding M&R interventions for a better budget allocation.

An example of a developed GUI for the best flexible model of Chapter X is shown in Figure 64.

A	B	C	D	E	F
1	<b>FLEXIBLE PAVEMENT - PERFORMANCE PREDICTION MODEL</b>				
2					
3					
4					
5	Section ID		18		
6	Begin Lat.	32.763156	Drainage	No	
7	Begin Long.	-91.083889	Age at 2010	55	
8	End Lat.	32.800877	PCR @ 2010	75	
9	End Long.	-91.102137	Iri @ 2010	2.18	
10	Structural Number	2.79	<b>PROJECTIONS</b>		
11	Length	6.92			
12	Projection Year	2030			
13	Rehab Year	2015			
14					

Figure 64. GUI for Flexible Pavement

The database, connection weights, threshold values, and coefficients for the model are imported into excel worksheets and the GUI utilizes programming codes from Excel’s Visual Basic to perform the calculations necessary for generating pavement performance predictions. The steps for the use of the developed GUI (Figure 64) are described as follows:

- The first step is to select which Section ID the user wants to analyze. The program will collect the information of the given section and complete the fields that are constant (Begin Lat, Begin Long, End Lat, End Long, Structural Number, Length, Drainage, Age at 2010, PCR @ 2010, and IRI @ 2010).
- The second step is to decide until what year the user wants the model to predict. The program will generate predictions until reaching the date inputted by the user.
- The third step is to input the rehabilitation year that the user wants to perform in the pavement section. The program will collect this data and use it to generate three different scenarios for the section. A no rehabilitation, a minor rehabilitation, and a major rehabilitation using the year inputted by the user.
- The fourth step is to click on the “Projections” button. A user form window will pop up and the user will need to choose the approximate ESAL increase per year and click on the “Submit” button to initiate the analysis.

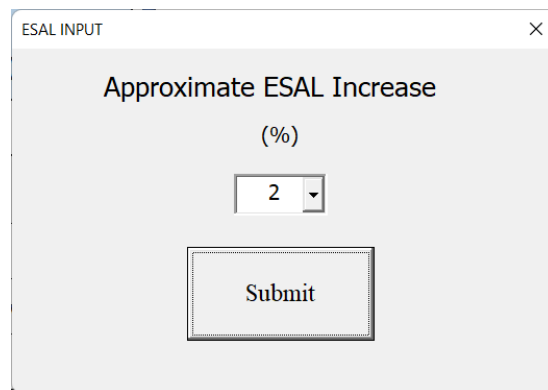


Figure 65. User Form for Approximate ESAL Increase

- The program will perform all calculations using the database, connection weights, threshold values, and coefficients to generate a table with the pavement performance predictions for the three scenarios described previously (Figure 66).

No Rehab			Minor Rehab			Major Rehab		
Year	PCR	IRI (mm/m)	Year	PCR	IRI (mm/m)	Year	PCR	IRI (mm/m)
2010	75	2.18	2010	75	2.18	2010	75	2.18
2011	73.81	2.01	2011	73.81	2.01	2011	73.81	2.01
2012	73.96	1.99	2012	73.96	1.99	2012	73.96	1.99
2013	74.21	2.01	2013	74.21	2.01	2013	74.21	2.01
2014	74.75	2.02	2014	74.75	2.02	2014	74.75	2.02
2015	75.63	2.03	2015	75.63	2.03	2015	75.63	2.03
2016	76.72	2.02	2016	76.72	2.02	2016	76.72	2.02
2017	77.41	1.98	2017	77.41	1.98	2017	77.41	1.98
2018	77.06	1.93	2018	77.06	1.93	2018	77.06	1.93
2019	76.16	1.99	2019	76.16	1.99	2019	76.16	1.99
2020	83.24	1.53	2020	83.24	1.53	2020	83.24	1.53
2021	78.99	2.23	2021	78.99	2.23	2021	78.99	2.23
2022	88.62	0.6	2022	88.62	0.6	2022	88.62	0.6
2023	80.65	1.5	2023	80.65	1.5	2023	80.65	1.5
2024	87.48	0.52	2024	87.48	0.52	2024	87.48	0.52
2025	89.66	0.37	2025	94.11	0.39	2025	94.51	0.35
2026	91.63	0.25	2026	93	0.21	2026	93.06	0.21
2027	92.15	0.22	2027	92.42	0.21	2027	92.43	0.21
2028	92.21	0.21	2028	92.26	0.21	2028	92.26	0.21
2029	92.22	0.2	2029	92.23	0.2	2029	92.23	0.2
2030	92.22	0.2	2030	92.23	0.2	2030	92.23	0.2

Figure 66. Pavement Performance Predictions

- The program will also generate a plot for each output variable including the predictions from all three scenarios.

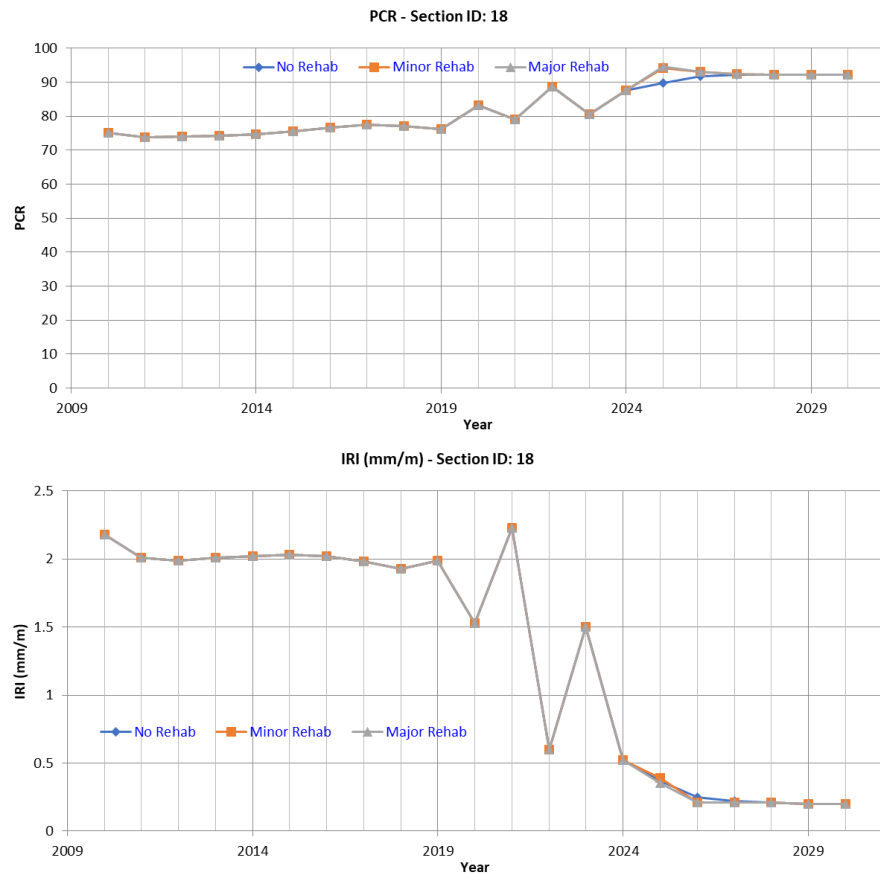


Figure 67. Predictions for PCR and IRI

- To perform another analysis, just return to step one and repeat the steps to obtain new predictions.

By providing a tool that generates predictions considering the effects of no rehabilitation, minor and major rehabilitations, agencies can analyze different scenarios and identify which sections would need to be prioritized in the M&R decision plan. This tool can also be used to simulate different years of M&R actions to evaluate the effects of early or late interventions on the pavement performance condition.

Therefore, this simple, low-cost, and easy-access tool brings a significant contribution to the enhancement of agencies' PMS by predicting future pavement conditions, identifying rehabilitation needs, and allowing a better budget allocation for critical pavement sections.

## CHAPTER XIV: SUMMARY, CONCLUSIONS, AND RECOMMENDATIONS

### 14.1. Summary

In this study, a set of pavement performance models were developed utilizing an ANN modeling technique and considering the effects of M&R history and interventions to offer a more inclusive, reliable, and accurate prediction tool. The developed models can be utilized by state and federal transportation agencies to predict future pavement conditions and incorporate the M&R scheduling effectively to prioritize the resources for the critical pavement sections. Additionally, the developed models did not use any distress data as an input variable, which will help agencies to save time in data collection and processing.

A new approach for considering the influence of M&R history was introduced in Chapter IV for the development of performance models for composite pavements utilizing the ANN technique within the LTPP database. All four climate zones of the LTPP database were utilized but the paper did not include all data points of the database in the model development. Five models were developed using different input and output variables and the best ANN model structure consisted of 14 inputs, 9 hidden nodes, and 1 output. By developing a new approach that considered the effects of the M&R history and interventions it is possible to obtain realistic and accurate prediction models for future planning.

Chapter V used the same approach established in the previous chapter to develop

performance prediction models to assess the influence of climate and traffic data on composite pavements. The study utilized all data points available for composite pavements on the LTPP database and developed three models using the ANN approach. The results indicate that all models resulted in reliable predictions but the model that included all climatological and traffic variables was more accurate. The final ANN structure for the best model included 21 inputs, 19 hidden nodes, and 1 output. An in-depth study was performed for random sections and demonstrated the ability of the developed model to capture the variations of roughness over time. The study showed that due to the complexity of the deterioration phenomena, M&R, traffic, and climate variables must be used together in the modeling process to deliver better prediction values.

The use of a specific climate zone in the development of pavement performance models was explored in Chapter VI. This study utilized a traditional (multiple regression) and an advanced (ANN) modeling technique to develop prediction models for the wet non-freeze climate zone of the LTPP database. The complex relationship between pavement structure, climate, and traffic variables was investigated and the results showed that the ANN model presented an accuracy 132% higher than the MLR model. Also, the use of a specific climate zone helped the model to capture 86% of the variability, which would not be viable when using data from all climate zones together. Therefore, further study with different climate zones should also be explored to verify if it provides more accurate predictions.

Chapter VII developed pavement performance models for a specific climate zone different than the one used in the previous chapter. The study used all data points available in the wet-free climate zone of the LTPP database. This chapter showed that the use of a specific climate zone combined with input variables that represent the pavement's initial condition, effects of pavement exposure time, pavement structure, effects of maintenance and rehabilitation, effects of traffic

loads, and climatological effects of temperature, moisture, and freeze in the same model increased the model accuracy. The developed models efficiently characterized the pavement roughness behavior on composite pavements and could be used as a prediction tool to develop better M&R plans.

A new approach for incorporating M&R history into the performance model development is discussed in Chapter VIII. Two models considering two different methods to account for the influence of M&R history were developed for the wet non-freeze climate region of the LTPP database using the ANN modeling technique. The new approach for M&R consisted to use a continuous variable instead of a categorical variable for the CN and resulted in an accuracy 10% higher when predicting roughness values compared to the previous method. This approach was then recommended to be used for other climate zones to verify if it would enhance other performance models and assist agencies to support objective decisions regarding M&R schedule and budget plans.

Chapter IX utilized the two developed approaches for M&R to develop new performance models for the wet non-freeze region of the LTPP database, a climate zone different than the one used in the previous chapter. Three models were developed and showed reliable results, however, the most promising model appeared to be the one using a categorical variable for M&R and without climatological variables. This result showed that the use of the  $CN_{Code}$  variable efficiently incorporated M&R actions in the model development and assisted the network to achieve higher accuracy compared to the  $CN_{Continuous}$  variable, different than the results from the previous chapter. Therefore, each climate zone might have different behaviors when exposed to different M&R approaches and variables and for this reason, a specific performance model needs to be developed for each one of the climate zones separately.

The use of a different database and model development was implemented in Chapter X. This study utilized the MDOT database and an ANN dynamic sequential modeling technique to develop performance prediction models for flexible pavements. Five models were developed considering the effects of several key factors such as geographic location, pavement structure, drainage, and traffic. Additionally, an innovative approach for M&R actions was developed to generate more realistic predictions and to be useful for the state agency. A two-output model structure was considered and resulted in promising results that could be further studied for rigid and composite pavements.

Chapter XI used the approach developed in the previous chapter to develop performance models utilizing the dynamic sequential ANN technique for rigid pavements of the MDOT database. The study developed four models utilizing different input and output variables including a discussion of the use of a one-output or two-output model to enhance pavement performance predictions. Model 1 showed the best results for PCR and IRI predictions with high accuracy, indicating that the deterioration behavior was captured by the model and translated in good agreement between predicted and observed values.

Chapter XII developed a new set of performance prediction models for composite pavements in Mississippi utilizing the dynamic sequential ANN method. Three models were developed with the same core of inputs but differing M&R and traffic variables to study the significance of each variable. A two-model output structure was utilized and all models had satisfactory results. However, the models with M&R variables performed better than the model without the effect of M&R, which shows the importance of utilizing an M&R variable. The study also showed that the best models can be used by MDOT personnel since they can be employed in a simple and user-friendly tool to predict future pavement conditions.



Chapter XIII shows the development of a graphical user interface for the implementation of developed models in a simple, fast, low-cost, and user-friendly tool to support objective decisions regarding maintenance and rehabilitation interventions and budget plans. Even though the development of pavement performance prediction models requires a good understanding of the pavement deterioration phenomena, and ANN modeling techniques, the developed GUI does not require the user to have any prior knowledge of pavement performance or ANN. This tool brings an enormous contribution to the enhancement of PMS.

## 14.2. Conclusions

Based on the results presented in the previous chapters, the key conclusions are listed in the following sections.

### 14.2.1. Roughness Modeling For Composite Pavements Using Machine Learning

- Most studies in the literature developed roughness models for flexible or rigid pavements, which resulted in a lack of research for composite pavements. This study contributes to the community by studying, discussing, and developing ANN performance models to better understand the deterioration behavior of composite pavements.
- The development of a new approach to consider M&R history in the model development and account for the effects of interventions resulted in an improvement in roughness prediction accuracy.
- Lower IRI values were better predicted than higher values, this might happened because the deterioration behavior for pavements of good quality is different from the pavements in worst conditions.
- Sensitivity analysis showed that predicted values were in good agreement with the observed values, which confirms the efficiency of the developed model. Additional analysis also showed that for future predictions the model successfully predicted the pavement deterioration over time when no M&R actions were performed.
- The developed model can be used to identify the specific year that a section would need an intervention to keep the road in an acceptable condition, which offers a powerful tool to

visualize effective solutions for future pavement needs.

#### 14.2.2. The Effect of Climate And Traffic On Composite Pavement Roughness Modeling Using Machine Learning

- The use of the ANN approach with a backpropagation algorithm for the development of performance prediction models for composite pavements of the LTPP database resulted in reliable and accurate predictions, which confirms the ANN capability of generalization studied in the literature.
- The use of traffic and climate variables applied together in the same model resulted in better prediction accuracy. This happens because due to the complexity of the pavement deterioration phenomena, these variables bring important information to assist the model in better recognizing the changes in roughness behavior by identifying the influence of each variable on the output and the relation between each of them. This resulted in an improvement in roughness prediction according to the data given.
- An in-depth study was performed for random sections of the database and showed that models that did not include M&R, climate, and traffic variables were not able to capture all the variations of IRI over time. These variations were better predicted when applying all variables in the same model.

#### 14.2.3. International Roughness Index Model For Composite Pavements In The LTPP Wet Non-Freeze Climate Region: Machine Learning And Regression Approaches

- Mean differences were assessed using the independent samples t-test to determine whether  $IRI_{Right}$  and  $IRI_{Left}$  differ on average from each other. The results show that the difference in the means of  $IRI_{Right}$  and  $IRI_{Left}$  is statistically significant at  $\alpha 0.05$  probability of chance error, which implies that both samples are from different populations. Therefore, the  $IRI_{Right}$  was used as the dependent variable since it shows the highest value for pavement roughness.
- Advanced and traditional modeling techniques were utilized for the development of

pavement performance models. The ANN model outperformed the MLR model with a MARE 53% lower and an ASE 99% lower. For the  $R^2$  value, the ANN model showed an improvement of 132% when compared to the MLR model. This confirms the selection of the ANN model as the best prediction model.

- A comparison between the ANN and MLR model predictions was also performed for two random sections of the database and the ANN model also proved to be better in capturing the pavement deterioration behavior over time following closely the observed values.
- This study was performed for a specific climate zone of the LTPP (wet non-freeze) to analyze if the prediction accuracy would be improved since the data points share the same climate characteristics. Results showed that the use of a specific climate zone helped the developed model to capture almost 90% of the variability, which may be not viable when using data from all climate zones together.

#### 14.2.4. Composite Pavement Roughness Modeling For LTPP Wet Freeze Climate Region Using Machine Learning

- The developed ANN model could replicate the pavement deterioration behavior with reasonable accuracy for composite pavements in the LTPP wet freeze climate zone. Predicted values cluster around the line of equality but some values were underpredicted, especially when the observed  $IRI_{Right}$  values were higher than 3 m/km.
- The use of a specific climate zone combined with input variables that represent the pavement's initial condition, effects of pavement exposure time, pavement structure, effects of maintenance and rehabilitation, effects of traffic loads, and climatological effects of temperature, moisture, and freeze in the same model increased the model accuracy.
- Model comparison was performed for two random sections of the database and showed that four of the five developed models performed reasonably, however, the only model that included the season variable was the only one that did not follow closely the observed values. Therefore, the season variable should not be used since it did not help to improve prediction accuracy and might be causing confusion in the model network.

#### 14.2.5. Pavement Performance Modeling For Composite Pavements in the LTPP Wet Freeze Climate Region Incorporating Maintenance and Rehabilitation

- A new approach ( $CN_{\text{Continuous}}$ ) to incorporate M&R history in the model development was developed and compared to the previous approach ( $CN_{\text{Code}}$ ) used in earlier chapters. The new approach consists in using a continuous variable instead of the categorical variable utilized before.
- Both ANN-developed models were able to predict pavement roughness values with reasonable accuracy. However, the model that utilized the  $CN_{\text{Continuous}}$  approach resulted in better accuracy when compared to the  $CN_{\text{Code}}$  approach.
- Therefore, the use of a new approach ( $CN_{\text{Continuous}}$ ) presented a significant improvement in prediction accuracy. However, this approach should be applied in other climate zones and datasets to verify its performance.

#### 14.2.6. Pavement Performance Modeling Considering Maintenance And Rehabilitation For Composite Pavements in the LTPP Wet Non-Freeze Region Using Neural Networks

- The  $CN_{\text{Continuous}}$  approach was applied for composite pavements located in the LTPP wet non-freeze climate zone and compared to the  $CN_{\text{Code}}$  approach to verify which M&R method would perform better.
- Even though all developed models showed accurate and reliable results, the model utilizing the  $CN_{\text{Code}}$  approach achieved higher accuracy compared to the model utilizing the  $CN_{\text{Continuous}}$ . Additionally, the most accurate model did not use any climatological factors.
- The use of climatological variables related to precipitation and temperature did not improve model performance and should not be used when modeling specific climate zones. This might happen because since the climate zone is already specified, there is no reason to overfeed the model with more data that has the same meaning as another one that has already been fed to the model.
- The implementation of Model 1 for two random sections of the database resulted in an excellent performance for predicting IRI variations due to time exposure and M&R actions.

#### 14.2.7. Development of Performance Models For Flexible Pavements Using The MDOT Database

- A dynamic sequential ANN technique was successfully implemented for the development of flexible pavements for the MDOT database.
- To incorporate the effect of the rehabilitation on PCR and IRI, a new M&R approach was utilized, and artificial rehabilitation actions based on the significant changes in PCR and IRI have been assigned to the database.
- The use of a two-output model was implemented to enhance model prediction accuracy when compared to models that used only one output.
- Five models were developed utilizing different input and output models. The best model showed to be the one utilizing CESAL for traffic consideration, and a two-output model with complementary PCR and IRI.
- The idea of using CESAL instead of ESAL is to introduce the cumulative history of traffic loads since the first recorded measurement for more inclusive predictions. This approach should be promising for flexible pavements in the MDOT database.
- PCR predictions were more accurate than IRI predictions. Further study is necessary to improve the model accuracy for the IRI output. However, the model can be considered highly accurate if the user goal is to obtain PCR values for the asphalt pavement sections.

#### 14.2.8. Performance Prediction Model For Jointed Concrete Pavements In Mississippi Using Machine Learning

- A dynamic sequential ANN technique was successfully implemented for the development of JCP pavements for the MDOT database.
- The same approach of artificial rehabilitation actions based on the significant changes in PCR and IRI has been utilized and applied to the database.
- The use of ESAL and a two-output model (Complementary PCR and IRI) helped the ANN model to generate more accurate predictions compared to the use of CESAL and the traditional one-output models.

- The model had an excellent prediction accuracy explaining more than 90% of the variability for both PCR and IRI, which indicates a good understating of complex relationships between input and output variables.
- A GUI application was developed for the model implementation in a random section of the database and resulted in accurate predictions, which confirms that the developed model can be considered a reliable, inclusive, and accurate tool to assist MDOT to predict JCP pavement conditions and incorporate a more effective M&R scheduling.

#### 14.2.9. Performance Evaluation Of Composite Pavements On Mississippi Highways Via Machine Learning

- A dynamic sequential ANN technique was successfully implemented for the development of composite pavements for the MDOT database.
- The same approach of artificial rehabilitation actions based on the significant changes in PCR and IRI has been utilized and applied to the database.
- All models had satisfactory results and could be used for generating reliable predictions for pavement performance. However, when not considering the effects of M&R history, model prediction accuracy significantly decreased.
- For the composite pavements, the most accurate model did not use the ESAL variable to account for the influence of traffic repetitions, which shows different behavior from other pavement types.

#### 14.2.10. Graphical User Interface

- The GUI developed in this dissertation for the implementation of the best models is a unique tool that enables a simple, fast, low-cost, and user-friendly tool to help agencies to apply the models in real life.
- The GUI can support objective decisions regarding maintenance and rehabilitation interventions and budget plans, allowing better visualization of the pavement deterioration behavior and helping state and federal agencies to simulate different scenarios according to their need.
- The GUI translates a specific and deep knowledge of pavement deterioration phenomena

and advanced modeling techniques into a tool that can be used by a user without any prior knowledge, which contributes to the agencies' PMS.

Based on the previously stated conclusions, this study has demonstrated that the developed performance models utilizing the ANN approaches have successfully predicted pavement deterioration over time with accurate and promising results. The inclusion of new variables and the development of new methodologies to consider the key effects of M&R history have been shown to improve model performance. Additionally, the development of a GUI for the implementation of the performance models offered a unique tool that can be utilized by agencies to support objective M&R decisions and develop better budget plans according to their needs. Furthermore, all models developed in this research do not require the use of any variables related to distress data, which makes the implementation of these models more accessible, helping agencies to save time and money in data collection and processing.

Therefore, this dissertation has successfully contributed to the enhancement of pavement performance modeling and the ANN's state-of-the-art by providing an efficient, reliable, and accurate tool that can be used in the scientific, academic, and practical fields.

#### 14.3. Recommendation for Future Research

- Even though this dissertation developed several models utilizing different pavement types, databases, and input and output variables, further studies exploring different variables and model architectures should be performed. These new models should attempt to achieve higher accuracy measurements.
- This study developed performance models including all climate zones and two specific climate zones. Further study should develop additional models for other climate zones that were not explored in this dissertation and evaluate the model's accuracy.
- The ANN models developed for the LTPP database utilized the static approach. The use of

the dynamic sequential approach should be explored to evaluate if it will enhance model accuracy.

- The ANN models were able to capture the performance changes and pavement deterioration behavior over time. However, some models were not as accurate when the observed IRI values were greater than 3 m/km. New performance models should be developed separating lower and higher IRI values to investigate if it improves model accuracy.
- Use the developed GUI to perform additional sensitivity analysis or to simulate different scenarios by modifying input values for selected variables to evaluate their effects on pavement performance predictions.
- Utilize the developed GUI to simulate pavement extreme conditions by inputting extreme values for climatological factors to evaluate the effects of climate changes on pavement performance.
- Verify the models' predictions with newly available measurements from the database to evaluate the model performance. As soon as new data points are available, the models can be updated to generate more accurate predictions for future years.



## LIST OF REFERENCES

## REFERENCES

- [1] Hudson WRonald, Uddin Waheed, Haas RC. Infrastructure Management: Integrating Design, Construction, Maintenance, Rehabilitation and Renovation. McGraw-Hill; 1997.
- [2] Jaafar ZFM. Computational Modeling and Simulations of Condition Deterioration to Enhance Asphalt Highway Pavement Design and Asset Management. University of Mississippi, 2019.
- [3] Chen C, Williams RC, Marasinghe MG, Ashlock JC, Smadi O, Schram S, et al. Assessment of composite pavement performance by survival analysis. Journal of Transportation Engineering 2015;141:1–9. [https://doi.org/10.1061/\(ASCE\)TE.1943-5436.0000784](https://doi.org/10.1061/(ASCE)TE.1943-5436.0000784).
- [4] Georgiou P, Plati C, Loizos A. Soft Computing Models to Predict Pavement Roughness: A Comparative Study. Advances in Civil Engineering 2018;2018. <https://doi.org/10.1155/2018/5939806>.
- [5] AASHTO. Mechanistic Empirical Pavement Design Guide: A Manual Practice. 2008.
- [6] AASHTO. AASHTO Guide for Design of Pavements Structures 1993:624.
- [7] Sayers MW. Two quarter-car models for defining road roughness. IRI and HRI. Transportation Research Record 1989:165–72.
- [8] Sayers MW, Gillespie TD, Queiroz CAV. International Road Roughness Experiment: a Basis for Establishing a Standard Scale for Road Roughness Measurements. Transportation Research Record 1986:76–85.

- [9] Sultana S. Computational Modeling of Climate Attributes and Condition Deterioration of Concrete Highway Pavements. Ph.D. Dissertation. University of Mississippi, 2021.
- [10] Kaya O, Ceylan H, Kim S, Waid D, Moore BP. Statistics and Artificial Intelligence-Based Pavement Performance and Remaining Service Life Prediction Models for Flexible and Composite Pavement Systems. *Transportation Research Record* 2020;2674:448–60. <https://doi.org/10.1177/0361198120915889>.
- [11] Gupta, Ankit; Kumar, Praveen; Rastogi R. A critical review of flexible pavement performance models developed for indian perspective. *Indian Highway* 2012;40:41–60.
- [12] Zeiada W, Dabous SA, Hamad K, Al-Ruzouq R, Khalil MA. Machine Learning for Pavement Performance Modelling in Warm Climate Regions. *Arabian Journal for Science and Engineering* 2020;45:4091–109. <https://doi.org/10.1007/s13369-020-04398-6>.
- [13] Yamany MS, Saeed TU, Volovski M, Ahmed A. Characterizing the Performance of Interstate Flexible Pavements Using Artificial Neural Networks and Random Parameters Regression. *Journal of Infrastructure Systems* 2020;26:04020010. [https://doi.org/10.1061/\(asce\)is.1943-555x.0000542](https://doi.org/10.1061/(asce)is.1943-555x.0000542).
- [14] Choi JH, Adams TM, Bahia HU. Pavement roughness modeling using back-propagation neural networks. *Computer-Aided Civil and Infrastructure Engineering* 2004;19:295–303. <https://doi.org/10.1111/j.1467-8667.2004.00356.x>.
- [15] Abdelaziz N, Abd El-Hakim RT, El-Badawy SM, Afify HA. International Roughness Index prediction model for flexible pavements. *International Journal of Pavement Engineering* 2020;21:88–99. <https://doi.org/10.1080/10298436.2018.1441414>.
- [16] Hossain MI, Gopiseti LSP, Miah MS. International roughness index prediction of flexible

pavements using neural networks. *Journal of Stomatology* 2019;145:1–10.  
<https://doi.org/10.1061/JPEODX.0000088>.

[17] Hossain M, Gopiseti LSP, Miah MS. Artificial neural network modelling to predict international roughness index of rigid pavements. *International Journal of Pavement Research and Technology* 2020;13:229–39. <https://doi.org/10.1007/s42947-020-0178-x>.

[18] Sollazzo G, Fwa TF, Bosurgi G. An ANN model to correlate roughness and structural performance in asphalt pavements. *Construction and Building Materials* 2017;134:684–93.  
<https://doi.org/10.1016/j.conbuildmat.2016.12.186>.

[19] Jaafar ZFM. Pavement Roughness Modeling Using Regression And Ann Methods For LTPP Western Region 2019:536–48. <https://doi.org/10.15405/epms.2019.12.53>.

[20] Bashar MZ, Torres-Machi C. Performance of Machine Learning Algorithms in Predicting the Pavement International Roughness Index. *Transportation Research Record: Journal of the Transportation Research Board* 2021:036119812098617.  
<https://doi.org/10.1177/0361198120986171>.

[21] Hunt PD, Bunker JM. Study of Site-Specific Roughness Progression for a Bitumen-Sealed Unbound Granular Pavement Network. *Transportation Research Record* 2003;I:273–81.  
<https://doi.org/10.3141/1819a-40>.

[22] Sandra AK, Sarkar AK. Development of a model for estimating International Roughness Index from pavement distresses. *International Journal of Pavement Engineering* 2013;14:715–24.  
<https://doi.org/10.1080/10298436.2012.703322>.

[23] Rahim AM, Fiegel G, Ghuzlan K, Khumann D. Evaluation of international roughness index for asphalt overlays placed over cracked and seated concrete pavements.

<https://doi.org/10.1080/10298430802342773>.

[24] Chen C, Zhang J. Comparisons of IRI-Based Pavement Deterioration Prediction Models Using New Mexico Pavement Data. *Geo-Frontiers Congress 2011* 2011:4594–603. [https://doi.org/10.1061/41165\(397\)470](https://doi.org/10.1061/41165(397)470).

[25] Mohamed Jaafar ZF bin, Uddin W, Najjar Y. Asphalt Pavement Roughness Modeling Using the Artificial Neural Network and Linear Regression Approaches for LTPP Southern Region. *Transportation Research Board* 2016.

[26] Barros R, Yasarer H, Sultana S. International Roughness Index Model for Composite Pavements in the LTPP Wet Non-Freeze Climate Region: Machine Learning and Regression Approaches. *101st Annual Meeting of the Transportation Research Board (TRB) 2022*, 2022.

[27] Solatifar N, Lavasani SM. Development of an artificial neural network model for asphalt pavement deterioration using LTPP data. *Journal of Rehabilitation in Civil Engineering* 2020;8:121–32. <https://doi.org/10.22075/JRCE.2019.17120.1328>.

[28] Tennant Duckworth P. *Evaluation of Flexible Pavement Performance Models in Mississippi: A Neural Network Approach*. University of Mississippi, 2020.

[29] Duckworth P, Yasarer H, Najjar Y. *Evaluation of Flexible Pavement Performance Models in Mississippi: A Neural Network Approach*. vol. 164. Springer International Publishing; 2022. [https://doi.org/10.1007/978-3-030-77230-7\\_15](https://doi.org/10.1007/978-3-030-77230-7_15).

[30] Yasarer H, Andrews W, Najjar Y. Performance Evaluation of Jointed Concrete Pavements on Mississippi Highways via Artificial Neural Network. *Tran-SET 2021*, Reston, VA: American Society of Civil Engineers; 2021, p. 86–92. <https://doi.org/10.1061/9780784483787.010>.

[31] Barros R, Yasarer H, Uddin W, Sultana S. Roughness Modeling for Asphalt Overlay on

Concrete Pavements Using Neural Networks. Transportation Research Board 2021.

[32] Sultana S, Yasarer H, Barros R, Najjar Y. Incorporating Maintenance and Rehabilitation History into Pavement Performance Modeling for Jointed Plain Concrete Pavement. ASCE International Conference on Transportation & Development (ICTD) 2022 2022:1–12.

[33] Sultana S, Yasarer H, Barros R, Najjar Y. Evaluating the Effect of Climate Change in Pavement Performance Modeling using Artificial Neural Network Approach. ASCE International Conference on Transportation & Development (ICTD) 2022 2022:1–12.

[34] Sultana S, Yasarer H, Uddin W, Barros R. International Roughness Index Modeling For Jointed Plain Concrete Pavement Using Artificial Neural Network. IOP Conference Series: Materials Science and Engineering 2021;1203:032034. <https://doi.org/10.1088/1757-899X/1203/3/032034>.

[35] Sultana S, Yasarer HI, Uddin W, Barros R. International Roughness Index Model for Jointed Plain Concrete Highway Pavements: An Artificial Neural Network Application. BCRRRA 2022 2022.

[36] Sultana S, Yasarer HI, Uddin W, Barros R. International Roughness Index Model for Jointed Plain Concrete Highway Pavements: An Artificial Neural Network Application. Transportation Research Board 2021.

[37] Qian J, Jin C, Zhang J, Ling J, Sun C. International Roughness Index Prediction Model for Thin Hot Mix Asphalt Overlay Treatment of Flexible Pavements. Transportation Research Record 2018;2672:7–13. <https://doi.org/10.1177/0361198118768522>.

[38] Faggella D. What is Machine Learning? 2020. <https://emerj.com/ai-glossary-terms/what-is-machine-learning/> (accessed October 4, 2021).

- [39] Darko A, Chan APC, Adabre MA, Edwards DJ, Hosseini MR, Ameyaw EE. Artificial intelligence in the AEC industry: Scientometric analysis and visualization of research activities. *Automation in Construction* 2020;112:103081. <https://doi.org/10.1016/j.autcon.2020.103081>.
- [40] Adeli H. Neural networks in civil engineering: 1989-2000. *Computer-Aided Civil and Infrastructure Engineering* 2001;16:126–42. <https://doi.org/10.1111/0885-9507.00219>.
- [41] Ceylan H, Bayrak MB, Gopalakrishnan K. Neural networks applications in pavement engineering: A recent survey. *International Journal of Pavement Research and Technology* 2014;7:434–44. <https://doi.org/10.6135/ijprt.org.tw/2014>.
- [42] Marcelino P, de Lurdes Antunes M, Fortunato E, Gomes MC. Machine learning approach for pavement performance prediction. *International Journal of Pavement Engineering* 2021;22:341–54. <https://doi.org/10.1080/10298436.2019.1609673>.
- [43] Najjar YM, Huang C. Simulating the stress-strain behavior of Georgia kaolin via recurrent neuronet approach. *Computers and Geotechnics* 2007;34:346–61. <https://doi.org/10.1016/j.compgeo.2007.06.006>.
- [44] Yasarer H. Characterizing the Permeability of Concrete Mixes Used in. Kansas State University, 2010.
- [45] Fausett L. Fundamentals of Neural Networks: Architectures, Algorithms and Applications Fundamentals of Neural Networks: Architectures, Algorithms and Applications. vol. 814. Prentice-Hall; 2005.
- [46] Yasarer H. Decision Making in Engineering Prediction Systems. PhD Dissertation. Kansas State University, 2013.
- [47] Al-masaeid HH. Utilizing of Magnesium Oxysulfate Binders for Additive Construction

Applications 2019.

[48] Qatu K. Optimizing the performance of complex engineering systems aided by artificial neural networks. 2019.

[49] Gasteiger J, Zupan J. Neural Networks in Chemistry. *Angewandte Chemie International Edition in English* 1993;32:503–27. <https://doi.org/10.1002/ANIE.199305031>.

[50] Yasarer H, Oyan MNS, Najjar Y. A Performance Prediction Model for Continuously Reinforced Concrete Pavement Using Artificial Neural Network. *Lecture Notes in Civil Engineering* 2020;76:771–82. [https://doi.org/10.1007/978-3-030-48679-2\\_72](https://doi.org/10.1007/978-3-030-48679-2_72).

[51] Yasarer H, Najjar Y. Development of a mix-design based Rapid Chloride Permeability assessment model using neuronets. *Proceedings of the International Joint Conference on Neural Networks* 2011:2756–60. <https://doi.org/10.1109/IJCNN.2011.6033580>.

[52] Yasarer H, Najjar YM. Characterizing the Permeability of Kansas Concrete Mixes Used in PCC Pavements. *International Journal of Geomechanics* 2014;14:04014017. [https://doi.org/10.1061/\(asce\)gm.1943-5622.0000362](https://doi.org/10.1061/(asce)gm.1943-5622.0000362).

[53] Yasarer H, Najjar Y. Assessing the auto associative network approach for prediction in civil engineering databases. *Procedia Computer Science* 2014;36:618–22. <https://doi.org/10.1016/j.procs.2014.09.064>.

[54] Najjar YM, Basheer IA. Utilizing computational neural networks for evaluating the permeability of compacted clay liners. *Geotechnical and Geological Engineering* 1996;14:193–212. <https://doi.org/10.1007/BF00452947>.

[55] Najjar Y. *Quick Manual for the Use of ANN program TRSEQ1*. Manhattan, Kansas: 1999.

[56] Najjar Y, Felker V. *Characterizing the roughness progression on Kansas Pavements: A*



sequential neuronet approach. GeoCongress 2006: Geotechnical Engineering in the Information Technology Age 2006;2006:127. [https://doi.org/10.1061/40803\(187\)127](https://doi.org/10.1061/40803(187)127).

[57] Ohio Department of Transportation. Pavement Condition Rating System. 2006.

[58] FHWA FHA. Performance Program. 2015. <https://doi.org/10.4135/9781412950558.n409>.

[59] FHWA FHA. Chapter 7. Recommendations - Evaluation of Long-Term Pavement Performance (LTPP) Climatic Data for Use in Mechanistic-Empirical Pavement Design Guide (MEPDG) Calibration and Other Pavement Analysis , May 2015 - FHWA-HRT-15-019. 2015.

[60] FHWA FHA. LTPP INFOPAVE Brochure 2018.

[61] FHWA FHA. Long Term Pavement Performance (LTPP) Info Pave: Data 2020. <https://infopave.fhwa.dot.gov/Data/DataSelection>.

[62] Yasarer H, Barros R, Mississippi Department of Transportation M. Development of Pavement Performance Models for MDOT: A Neural Network Approach. Oxford, Mississippi: 2022.

[63] Barros R, Yasarer H, Najjar Y. Performance Evaluation of Composite Pavements on Mississippi Highways via Machine Learning. BCRRA 2022 2022.

[64] Barros R, Yasarer H, Sultana S, Najjar Y. Performance Prediction Model for Jointed Concrete Pavements in Mississippi using Machine Learning. ASCE International Conference on Transportation & Development (ICTD) 2022 2022:1–13.

[65] Yasarer H, Oyan MNS, Najjar Y. A Performance Prediction Model for Continuously Reinforced Concrete Pavement Using Artificial Neural Network. Proceedings of the 9th International Conference on Maintenance and Rehabilitation of Pavements—Mairepav9, vol. 76, 2020, p. 771–82. [https://doi.org/10.1007/978-3-030-48679-2\\_72](https://doi.org/10.1007/978-3-030-48679-2_72).

- [66] Rahim AM, Fiegel G, Ghuzlan K, Khumann D. Evaluation of international roughness index for asphalt overlays placed over cracked and seated concrete pavements. *International Journal of Pavement Engineering* 2009;10:201–7. <https://doi.org/10.1080/10298430802342773>.
- [67] Al-Omari B, Darter MI. Effect of pavement deterioration types on IRI and rehabilitation. *Transportation Research Record* 1995:57–65.
- [68] Khattak MJ, Nur MA, Bhuyan MRUK, Gaspard K. International roughness index models for HMA overlay treatment of flexible and composite pavements. *International Journal of Pavement Engineering* 2014;15:334–44. <https://doi.org/10.1080/10298436.2013.842237>.
- [69] Attoh-Okine N O. Predicting Roughness Progression in Flexible Pavements Using Artificial Neural Networks. *3rd International Conference on Managing Pavements* 1994;1:55–62.
- [70] Kargah-Ostadi N, Stoffels SM, Tabatabaee N. Network-level pavement roughness prediction model for rehabilitation recommendations. *Transportation Research Record* 2010:124–33. <https://doi.org/10.3141/2155-14>.
- [71] Barros R, Yasarer H, Sultana S, Jaafar ZFM, Najjar Y. Performance Model Development for Flexible Pavements via Neural Networks. *ASCE International Conference on Transportation & Development (ICTD) 2022* 2022:1–13.
- [72] Abd El-Hakim R, El-Badawy S. International roughness index prediction for rigid pavements: An Artificial Neural Network Application. *Advanced Materials Research* 2013;723:854–60. <https://doi.org/10.4028/www.scientific.net/AMR.723.854>.
- [73] Shahin MY. *Pavement Management for Airports, Roads, And Parking Lots*. Springer; 1994.
- [74] Lin J, Yau J-T, Hsiao L-H. *Correlation Analysis Between International Roughness Index (*

IRI ) By. Transportation Research Board 82th Annual Meeting 2003:1–21.

[75] FHWA. LTPP Manual for Profile Measurements and Processing 2008.

[76] Meegoda JN, Gao S. Roughness progression model for asphalt pavements using long-term pavement performance data. *Journal of Transportation Engineering* 2014;140:1–7. [https://doi.org/10.1061/\(ASCE\)TE.1943-5436.0000682](https://doi.org/10.1061/(ASCE)TE.1943-5436.0000682).

[77] Chandra S, Sekhar CR, Bharti AK, Kangadurai B. Relationship between pavement roughness and distress parameters for indian highways. *Journal of Transportation Engineering* 2013;139:467–75. [https://doi.org/10.1061/\(ASCE\)TE.1943-5436.0000512](https://doi.org/10.1061/(ASCE)TE.1943-5436.0000512).

[78] Michigan Department of Transportation (MDOT). International Roughness Index. 2017.

[79] Yang J, Lu JJ, Gunaratne M, Xiang Q. Forecasting Overall Pavement Condition with Neural Networks: Application on Florida Highway Network. *Transportation Research Record* 2003:3–12. <https://doi.org/10.3141/1853-01>.

[80] Bianchini A, Bandini P. Prediction of pavement performance through neuro-fuzzy reasoning. *Computer-Aided Civil and Infrastructure Engineering* 2010;25:39–54. <https://doi.org/10.1111/j.1467-8667.2009.00615.x>.

[81] Ziari H, Sobhani J, Ayoubinejad J, Hartmann T. Prediction of IRI in short and long terms for flexible pavements: ANN and GMDH methods. *International Journal of Pavement Engineering* 2015;17:776–88. <https://doi.org/10.1080/10298436.2015.1019498>.

[82] Elhadidy AA, El-Badawy SM, Elbeltagi EE. A simplified pavement condition index regression model for pavement evaluation. *International Journal of Pavement Engineering* 2021;22:643–52. <https://doi.org/10.1080/10298436.2019.1633579>.

[83] Mazari M, Rodriguez DD. Prediction of pavement roughness using a hybrid gene

expression programming-neural network technique. *Journal of Traffic and Transportation Engineering (English Edition)* 2016;3:448–55. <https://doi.org/10.1016/j.jtte.2016.09.007>.

[84] FHWA FHA. LONG-TERM PAVEMENT PERFORMANCE Information Management System User Guide. vol. 088. 2019.

[85] Sinshaw TA, Surbeck CQ, Yasarer H, Najjar Y. Artificial Neural Network for Prediction of Total Nitrogen and Phosphorus in US Lakes. *Journal of Environmental Engineering* 2019;145:04019032. [https://doi.org/10.1061/\(asce\)ee.1943-7870.0001528](https://doi.org/10.1061/(asce)ee.1943-7870.0001528).

[86] Sibanda W, Pretorius P. Artificial Neural Networks-A Review of Applications of Neural Networks in the Modeling of HIV Epidemic. *International Journal of Computer Applications* 2012;44:975–8887.

[87] Najjar Y, Mryyan S. Characterization of a Perchlorate Contaminated Site. *Proceedings of World Environmental and Water Resources Congress 2009 - World Environmental and Water Resources Congress 2009: Great Rivers 2009*;342:1–11. [https://doi.org/10.1061/41036\(342\)240](https://doi.org/10.1061/41036(342)240).

[88] Barros R, Yasarer H, Uddin W, Sultana S. Roughness Modeling for Composite Pavements using Machine Learning. *IOP Conference Series: Materials Science and Engineering* 2021;1203:032035. <https://doi.org/10.1088/1757-899X/1203/3/032035>.

[89] Sultana S, Yasarer H, Uddin W, Barros R. International Roughness Index Modeling For Jointed Plain Concrete Pavement Using Artificial Neural Network. *IOP Conference Series: Material Science and Engineering 6th World Multidisciplinary Civil Engineering-Architecture-Urban Planning Symposium - WMCAUS 2021, Prague (Czech Republic), 30 August - 3 September 2021* 2021.

[90] Elhadidy AA, Elbeltagi EE, Ammar MA. Optimum analysis of pavement maintenance

using multi-objective genetic algorithms. *HBRC Journal* 2015;11:107–13.  
<https://doi.org/10.1016/j.hbrcj.2014.02.008>.

[91] Madeh Pirayonesi S, El-Diraby TE. Using Machine Learning to Examine Impact of Type of Performance Indicator on Flexible Pavement Deterioration Modeling. *Journal of Infrastructure Systems* 2021;27:04021005. [https://doi.org/10.1061/\(asce\)is.1943-555x.0000602](https://doi.org/10.1061/(asce)is.1943-555x.0000602).

[92] Barros R, Yasarer H, Uddin W, Sultana S. Roughness Modeling for Composite Pavements using Machine Learning. *IOP Conference Series: Material Science and Engineering* 6th WMCAUS 2021 2021. <https://doi.org/10.1088/1757-899X/1203/3/032035>.

[93] Mississippi Department of Transportation M. MDOT Pavement Condition (PATH) 2021. [https://path.mdot.ms.gov/pavement\\_condition](https://path.mdot.ms.gov/pavement_condition) (accessed January 17, 2022).

[94] Zhao Z, Guo Z. Prediction of Asphalt Pavement International Roughness Index (IRI) by Combined Approach of Empirical Regression and Markov. *ICTE 2013 - Proceedings of the 4th International Conference on Transportation Engineering* 2013:2106–13. <https://doi.org/10.1061/9780784413159.307>.

## APPENDIX

## Appendix A Pavement Condition Rating Forms and Key Forms [57]

Section: \_\_\_\_\_ Date: \_\_\_\_\_

Log Mile: \_\_\_\_\_ to \_\_\_\_\_ **KEY** Rated by: \_\_\_\_\_

Sta: \_\_\_\_\_ to \_\_\_\_\_ **FLEXIBLE PAVEMENT CONDITION** # of Utility Cuts \_\_\_\_\_

**RATING FORM**

DISTRESS	Distress Weight	SEVERITY*			EXTENT**			STR ***
		L	M	H	O	F	E	
RAVELING	10	Slight Loss of Sand	Open Texture	Rough or pitted	<20%	20-50%	>50%	
BLEEDING	5	not rated	Bit and Agg visible	Black Surface	<10%	10-30%	>30%	
PATCHING	5	<1 ft <sup>2</sup>	<1 yd <sup>2</sup>	>1 yd <sup>2</sup>	<10/mile	10-20/mile	>20/mile	
DEBONDING	5	depth <1" area <1 yd <sup>2</sup>	<1", >1 yd <sup>2</sup>	>1" and >1 yd <sup>2</sup>	<5/mile	5-10/mile	>10/mile	
CRACK SEALING DEFIC.	5	Not considered			<50%	>50%	No Sealant	
RUTTING	10	1/8" - 3/8"	3/8" - 3/4"	> 3/4"	<20%	20-50%	>50%	✓
SETTLEMENTS	0	Noticeable effect on ride	Some Discomfort	Poor Ride	<2/mi	2-4/mi	>4/mi	
POTHoles	10	depth <1" area <1 yd <sup>2</sup>	<1", >1 yd <sup>2</sup>	>1" and >1 yd <sup>2</sup>	<5/mile	5-10/mile	>10/mile	✓
WHEEL TRACK CRACKING	15	Single/multiple cracks <1/4"	Multiple cracks >1/4"	Alligator >1/4" Spalling with	<20%	20-50%	>50%	✓
BLOCK & TRANSVERSE CRACKING	10	> 6' X 6' or Transverse Crk.	6' x 6' to 3' x 3'	< 3' x 3'	<20%	20-50%	>50%	
LONGITUDINAL CRACKING	5	Single, <1/4", no Spalling	single/multiple 1/4-1", some Spalling	Multiple, >1", Spalling	< 50' per 100'	50-150' per 100'	>150' per 100'	✓
EDGE CRACKING	10	Tight, <1/4"	>1/4", some Spalling	>1/4", moderate Spalling	<20%	20-50%	>50%	✓
THERMAL CRACKING	10	<1/4"	1/4-1"	>1"	CS > 200'	CS 75-200'	CS <75'	

\*L = LOW      \*\*O = OCCASIONAL      \*\*\*STR = DISTRESS INCLUDED IN STRUCTURAL DEDUCT CALCULATIONS.  
M = MEDIUM      F = FREQUENT  
H = HIGH      E = EXTENSIVE

Figure A1. Key Flexible Pavement Condition Rating Form

Section: \_\_\_\_\_ Date: \_\_\_\_\_

Log mile: \_\_\_\_\_ to \_\_\_\_\_ **FLEXIBLE** Rated by: \_\_\_\_\_

Sta: \_\_\_\_\_ to \_\_\_\_\_ # of Utility Cuts \_\_\_\_\_

**PAVEMENT CONDITION RATING FORM**

DISTRESS	DISTRESS WEIGHT	SEVERITY WT.*			EXTENT WT.**			DEDUCT POINTS***
		L	M	H	O	F	E	
RAVELING	10	0.3	0.6	1	0.5	0.8	1	
BLEEDING	5	0.8	0.8	1	0.6	0.9	1	
PATCHING	5	0.3	0.6	1	0.6	0.8	1	
DEBONDING	5	0.4	0.7	1	0.5	0.8	1	
CRACK SEALING DEFICIENCY	5	1	1	1	0.5	0.8	1	
RUTTING	10	0.3	0.7	1	0.6	0.8	1	✓
SETTLEMENT	0	0.0	0.0	0.0	0.0	0.0	0.0	
POTHoles	10	0.4	0.8	1	0.5	0.8	1	✓
WHEEL TRACK CRACKING	15	0.4	0.7	1	0.5	0.7	1	✓
BLOCK AND TRANSVERSE CRACKING	10	0.4	0.7	1	0.5	0.7	1	
LONGITUDINAL CRACKING	5	0.4	0.7	1	0.5	0.7	1	✓
EDGE CRACKING	10	0.4	0.7	1	0.5	0.7	1	✓
THERMAL CRACKING	10	0.4	0.7	1	0.5	0.7	1	

\*L = LOW      \*\*O = OCCASIONAL      TOTAL DEDUCT = \_\_\_\_\_  
M = MEDIUM      F = FREQUENT      SUM OF STRUCTURAL DEDUCT (✓) = \_\_\_\_\_  
H = HIGH      E = EXTENSIVE      100 - TOTAL DEDUCT = PCR = \_\_\_\_\_

\*\*\* DEDUCT POINTS = DISTRESS WEIGHT x SEVERITY WT. x EXTENT WT.

Figure A2. Flexible Pavement Condition Rating Form

Section: \_\_\_\_\_

**KEY**

Date: \_\_\_\_\_

Log Mile: \_\_\_\_\_ to \_\_\_\_\_

**ASPHALT SURFACE LOCAL**

Rated by: \_\_\_\_\_

Sta: \_\_\_\_\_ to \_\_\_\_\_

**RATING FORM**

# of Utility Cuts \_\_\_\_\_

DISTRESS	Distress Weight	SEVERITY*			EXTENT**			STR ***
		L	M	H	O	F	E	
RAVELING	10	Slight Loss of Sand	Open Texture	Rough or Pitted	<20%	20-50%	>50%	
BLEEDING	5	not rated	Bitumen & Agg. Visible	Black Surface	<10%	10-30%	>30%	
PATCHING	5	<1 ft <sup>2</sup>	<1 yd <sup>2</sup>	>1 yd <sup>2</sup>	<10/mile	10-20/mile	>20/mile	
SURFACE DISINTEGRATION / DEBONDING /POTHOLES	5	depth <1" area <1 yd <sup>2</sup>	<1", >1 yd <sup>2</sup>	>1" and >1 yd <sup>2</sup>	<5/mile	5-10/mile	>10/mile	
RUTTING	10	1/8" - 3/8"	3/8" - 3/4"	>3/4"	<20%	20-50%	>50%	✓
MAP CRACKING	5	5' x 5' to 9' x 9'	1' x 1' to 5' x 5'	< 1' x 1' or alligator	<20%	20-50%	>50%	
BASE FAILURE	10	Barely Noticeable Pitch & Roll	Noticeable Pitch & Roll, Jarring Bump	Severe Distortion, Poor Ride	<2/mi	2-5/mi	>5/mi	✓
SETTLEMENTS	5	Noticeable effect on ride	Some Discomfort	Poor Ride	<2/mi	2-4/mi	>4/mi	
TRANSVERSE CRACKS	10	<1/4", no spalling	1/4 - 1", >.5 spalled	>1", >.5 spalled	CS>100'	100'<CS<50'	CS<50'	✓
WHEEL TRACK CRACKING	15	Single/multiple cracks <1/4"	Multiple cracks >1/4"	Alligator >1/4" Spalling	<20%	20-50%	>50%	✓
LONGITUDINAL CRACKING	5	<1/4", no spalling	1/4 - 1", >.5 spalled	>1", >.5 spalled	<50' per 100	50 -150' per 100'	>150' per 100	✓
EDGE CRACKING	5	Tight, <1/4"	>1/4", some Spalling	>1/4", moderate Spalling	<20%	20-50%	>50%	✓
PRESSURE DAMAGE/ UPHEAVAL	5	bump <1/2", Barely Noticeable	1/2" - 1", Fair Ride	>1", Poor Ride	<5/mile	5-10/mile	>10/mile	
CRACK SEALING DEFIC.	5	Not considered			<50%	>50%	No Sealant	

\*L = LOW                      \*\*O = OCCASIONAL                      \*\*\*STR = DISTRESS INCLUDED IN STRUCTURAL DEDUCT CALCULATIONS.  
M = MEDIUM                      F = FREQUENT  
H = HIGH                              E = EXTENSIVE

Figure A3. Key Asphalt Surface Local Rating Form

Section: \_\_\_\_\_

**LOCAL**

Date: \_\_\_\_\_

Log mile: \_\_\_\_\_ to \_\_\_\_\_

Rated by: \_\_\_\_\_

Sta: \_\_\_\_\_ to \_\_\_\_\_

# of Utility Cuts \_\_\_\_\_

**PAVEMENT CONDITION RATING FORM**

DISTRESS	DISTRESS WEIGHT	SEVERITY WT.*			EXTENT WT.**			DEDUCT POINTS***
		L	M	H	O	F	E	
RAVELING	10	0.3	0.6	1	0.5	0.8	1	
BLEEDING	5	0.8	0.8	1	0.6	0.9	1	
PATCHING	5	0.3	0.6	1	0.6	0.8	1	
SURFACE DISINTEGRATION or DEBONDING	5	0.3	0.6	1	0.6	0.8	1	
RUTTING	10	0.3	0.7	1	0.6	0.8	1 ✓	
MAP CRACKING	5	0.2	0.6	1	0.4	0.8	1	
BASE FAILURE	10	0.6	0.8	1	0.7	0.9	1 ✓	
SETTLEMENTS	5	0.4	0.7	1	0.6	0.8	1	
TRANSVERSE CRACKS	10	0.4	0.7	1	0.5	0.7	1 ✓	
WHEEL TRACK CRACKING	15	0.4	0.7	1	0.5	0.7	1 ✓	
LONGITUDINAL CRACKING	5	0.2	0.6	1	0.4	0.8	1 ✓	
EDGE CRACKING	5	0.4	0.7	1	0.5	0.7	1 ✓	
PRESSURE DAMAGE/UPHEAVAL	5	0.4	0.6	1	0.5	0.8	1	
CRACK SEALING DEFICIENCY	5	1	1	1	0.5	0.8	1	

\*L = LOW                      \*\*O = OCCASIONAL                      TOTAL DEDUCT = \_\_\_\_\_  
M = MEDIUM                      F = FREQUENT                      SUM OF STRUCTURAL DEDUCT (✓) = \_\_\_\_\_  
H = HIGH                              E = EXTENSIVE                      100 - TOTAL DEDUCT = PCR = \_\_\_\_\_

\*\*\* DEDUCT POINTS = DISTRESS WEIGHT X SEVERITY WT. X EXTENT WT.  
REMARKS:

Figure A4. Local Pavement Condition Rating Form



Section: \_\_\_\_\_ **KEY** Date: \_\_\_\_\_  
 Log Mile: \_\_\_\_\_ to \_\_\_\_\_ **COMPOSITE PAVEMENT CONDITION** Rated by: \_\_\_\_\_  
 Sta: \_\_\_\_\_ to \_\_\_\_\_ **RATING FORM** # of Utility Cuts: \_\_\_\_\_

DISTRESS	Distress Weight	SEVERITY*			EXTENT**			STR***
		L	M	H	O	F	E	
RAVELING	10	Slight Loss of Sand	Open Texture	Rough or Pitted	<20%	20-50%	>50%	
BLEEDING	5	not rated	Bitumen & Agg. Visible	Black Surface	<10%	10-30%	>30%	
PATCHING	5	<1 ft <sup>2</sup>	<1 yd <sup>2</sup>	>1 yd <sup>2</sup>	<10/mile	10-20/mile	>20/mile	
SURFACE DISINTEGRATION/DEBONDING	5	depth <1" area <1 yd <sup>2</sup>	<1", >1 yd <sup>2</sup> >1", <1 yd <sup>2</sup>	depth >1" area >1 yd <sup>2</sup>	<5/mile	5-10/mile	>10/mile	
RUTTING	10	1/8" - 3/8"	3/8" - 3/4"	>3/4"	<20%	20-50%	>50%	
PUMPING	10	Rater is certain of pumping			<10%	10-25%	>25%	✓
SHATTERED SLAB (Jointed Base)	10	Little Spall, No Faults	Moderate Faults	Distortion, Poor Ride	<2/mi	2-5/mi	>5/mi	✓
SETTLEMENTS	0	Noticeable effect on ride	Some Discomfort	Poor Ride	<2/mi	2-4/mi	>4/mi	
TRANSVERSE CRACKS (Unjointed Base)	20	<1/4", no spalling	1/4 - 1", >.5 spalled	>1", >.5 spalled	CS>15'	10'<CS<15'	CS<10'	✓
JOINT REFLECTION CRACKS (Jointed Base)	12	<1/4", no spalling	1/4 - 1", >.5 spalled	>1", >.5 spalled	<20%	20-50%	>50%	
INTERMEDIATE TRANSVERSE CRACKS (Jointed Base)	8	<1/4", no spalling	1/4 - 1", >.5 spalled	>1", >.5 spalled	< 10% of the Slabs	10% - 30% of the Slabs	> 30% of the Slabs	✓
LONGITUDINAL CRACKING	5	<1/4", no spalling	1/4 - 1", >.5 spalled	>1", >.5 spalled	<50' per 100'	50 - 150' per 100'	>150' per 100'	✓
PRESSURE DAMAGE/UPHEAVAL	5	bump <1/2", Good Ride	1/2-1", Fair Ride	>1", Poor Ride	<5/mile	5-10/mile	>10/mile	
CRACK SEAL DEFICIENCY	5	Not considered			< 50%	> 50%	no sealant	
CORNER BREAKS (Jointed Base)	10	NO DEPRESSION	DEPRESSION < 2"	DEPRESSION > 2"	< 4/mile	4 - 10/mile	> 10/mile	✓
PUNCHOUTS (Unjointed Base)	15	NOT RATED	DEPRESS < 2"	DEPRESS > 2"	< 4/mile	4 - 10/mile	> 10/mile	✓

\*L = LOW      \*\*O = OCCASIONAL      \*\*\*STR = DISTRESS INCLUDED IN STRUCTURAL DEDUCT CALCULATIONS.  
 M = MEDIUM      F = FREQUENT  
 H = HIGH      E = EXTENSIVE

Figure A5. Key Composite Pavement Condition Rating Form

Section: \_\_\_\_\_ **COMPOSITE** Date: \_\_\_\_\_  
 Log mile: \_\_\_\_\_ to \_\_\_\_\_ Rated by: \_\_\_\_\_  
 Sta: \_\_\_\_\_ to \_\_\_\_\_ # of Utility Cuts: \_\_\_\_\_

**PAVEMENT CONDITION RATING FORM**

DISTRESS	DISTRESS WEIGHT	SEVERITY WT.*			EXTENT WT.**			DEDUCT POINTS***
		L	M	H	O	F	E	
RAVELING	10	0.3	0.6	1	0.5	0.8	1	
BLEEDING	5	0.8	0.8	1	0.6	0.9	1	
PATCHING	5	0.3	0.6	1	0.6	0.8	1	
SURFACE DISINTEGRATION or DEBONDING	5	0.3	0.6	1	0.6	0.8	1	
RUTTING	10	0.3	0.7	1	0.6	0.8	1	
PUMPING	10	1	1	1	0.3	0.7	1	✓
SHATTERED SLAB (Jointed Base)	10	0.6	0.8	1	0.7	0.9	1	✓
SETTLEMENTS	0	0.0	0.0	0.0	0.0	0.0	0.0	
TRANSVERSE CRACKS, (Unjointed Base)	20	0.2	0.6	1	0.4	0.8	1	✓
JOINT REFLECTION CRACKS (Jointed Base)	12	0.2	0.6	1	0.4	0.8	1	
INTERMEDIATE TRANSVERSE CRACKS (Jointed Base)	8	0.2	0.6	1	0.4	0.8	1	✓
LONGITUDINAL CRACKING	5	0.2	0.6	1	0.4	0.8	1	✓
PRESSURE DAMAGE/UPHEAVAL	5	0.4	0.6	1	0.5	0.8	1	
CRACK SEALING DEFICIENCY	5	1	1	1	0.5	0.8	1	
CORNER BREAKS (JOINTED BASE)	10	0.4	0.8	1	0.5	0.8	1	✓
PUNCHOUTS (UNJOINTED BASE)	15	.8	.8	1	0.5	0.8	1	✓

\*L = LOW      \*\*O = OCCASIONAL      TOTAL DEDUCT = \_\_\_\_\_  
 M = MEDIUM      F = FREQUENT      SUM OF STRUCTURAL DEDUCT (✓) = \_\_\_\_\_  
 H = HIGH      E = EXTENSIVE      100 - TOTAL DEDUCT = PCR = \_\_\_\_\_

\*\*\* DEDUCT POINTS = DISTRESS WEIGHT X SEVERITY WT. X EXTENT WT.

Figure A6. Composite Pavement Condition Rating Form

Section: \_\_\_\_\_

**KEY**

Date: \_\_\_\_\_

Log Mile: \_\_\_\_\_ to \_\_\_\_\_

**JOINTED CONCRETE PAVEMENT**

Rated by: \_\_\_\_\_

Sta: \_\_\_\_\_ to \_\_\_\_\_

**CONDITION RATING FORM**

# of Utility Cuts \_\_\_\_\_

DISTRESS	Distress Weight	SEVERITY WEIGHT*			EXTENT WEIGHT**			STR ***
		L	M	H	O	F	E	
SURFACE DETERIORATION	10	Aggregate visible	Loss of fine aggregate	Surface rough & pitted	< 20%	20 - 50%	> 50%	
LONGITUDINAL JOINT SPALLING	5	< 3"	3" - 6"	> 6"	<10%	10 - 30%	>30%	
PATCHING	10	<1 ft <sup>2</sup> , no deterioration	<1 ft <sup>2</sup> , deterioration	>1 ft <sup>2</sup>	<10/mi	10-20/mi	>20/mi	
PUMPING	15	Rater is certain of pumping			<10%	10-25%	>25%	✓
FAULTING (Joints & Cracks)	10	<1/4"	1/4-1/2"	>1/2"	<20%	20-50%	>50%	✓
SETTLEMENTS	0	Noticeable effect on Ride	Some discomfort	Poor Ride	2/mi.	2-4/mi.	>4/mi.	
TRANSVERSE JOINT SPALLING	10	<4" wide	4-9" wide	>9" wide	<25%	25-75%	75%	
TRANSVERSE CRACKING (Plain Concrete)	15	Hairline	Hairline - 3/16"	> 3/16"	<10% of the slabs	10 to 30%	>30%	✓
PRESSURE DAMAGE	5	Not considered			<1/mi	1-3/mi	>3/mi	
TRANSVERSE CRACKING (Reinforced Concrete)	15	No failed Cracks	If Severity is Low, use <input type="checkbox"/>		<25% slabs	25 to 75%	>75% slabs	✓
			1 failed Crack	>1 failed Crack	5 to 10%	10 to 30%	>30% slabs	
LONGITUDINAL CRACKING	10	Hairline	1/4-1"	>1"	<5%	5-20%	>20%	✓
CORNER BREAKS	10	<1/4"	1/4-1"	>1"	<4/mi	4-10/mi.	>10 mi.	✓

\*L = LOW  
M = MEDIUM  
H = HIGH

\*\*O = OCCASIONAL  
F = FREQUENT  
E = EXTENSIVE

\*\*\*STR = DISTRESS INCLUDED IN STRUCTURAL DEDUCT CALCULATIONS.

Figure A7. Key Jointed Concrete Pavement Condition Rating Form

Section: \_\_\_\_\_

**JOINTED**

Date: \_\_\_\_\_

Log mile: \_\_\_\_\_ to \_\_\_\_\_

Rated by: \_\_\_\_\_

Sta: \_\_\_\_\_ to \_\_\_\_\_

# of Utility Cuts: \_\_\_\_\_

**PAVEMENT CONDITION RATING FORM**

DISTRESS	DISTRESS WEIGHT	SEVERITY WT.*			EXTENT WT.**			DEDUCT POINTS***
		L	M	H	O	F	E	
SURFACE DETERIORATION	10	0.4	0.7	1	0.6	0.8	1	
LONGITUDINAL JOINT SPALLING	5	0.4	0.7	1	0.6	0.8	1	
PATCHING	10	0.4	0.7	1	0.5	0.8	1	
PUMPING	15	1	1	1	0.3	0.7	1 ✓	
FAULTING (JOINTS AND CRACKS)	10	0.4	0.7	1	0.5	0.8	1 ✓	
SETTLEMENTS	0	0.0	0.0	0.0	0.0	0.0	1	
TRANSVERSE JOINT SPALLING (CIRCLE IF D-CRACKED)	10	0.4	0.7	1	0.5	0.8	1	
TRANSVERSE CRACKING (Plain Concrete)	15	1	1	1	0.5	0.8	1 ✓	
PRESSURE DAMAGE	5	1	1	1	0.5	0.8	1	
TRANSVERSE CRACKING (Reinforced Concrete)	15	0.1	0.8	1	0.4	0.8	1 ✓	
LONGITUDINAL CRACKING	10	0.5	0.7	1	0.4	0.9	1 ✓	
CORNER BREAKS	10	0.4	0.8	1	0.5	0.8	1 ✓	

\*L = LOW  
M = MEDIUM  
H = HIGH

\*\*O = OCCASIONAL  
F = FREQUENT  
E = EXTENSIVE

TOTAL DEDUCT = \_\_\_\_\_  
SUM OF STRUCTURAL DEDUCT (✓) = \_\_\_\_\_  
100 - TOTAL DEDUCT = PCR = \_\_\_\_\_

\*\*\* DEDUCT POINTS = DISTRESS WEIGHT X SEVERITY WT. X EXTENT WT.

Figure A8. Jointed Concrete Pavement Condition Rating Form

Section: \_\_\_\_\_  
 Log Mile: \_\_\_\_\_ to \_\_\_\_\_  
 Sta: \_\_\_\_\_ to \_\_\_\_\_

**KEY**  
**CRC PAVEMENT CONDITION**  
**RATING FORM**

Date: \_\_\_\_\_  
 Rated by: \_\_\_\_\_  
 # of Utility Cuts \_\_\_\_\_

DISTRESS	DISTRESS WEIGHT	SEVERITY WEIGHT*			EXTENT WEIGHT**			STR ***
		L	M	H	O	F	E	
SURFACE DETERIORATION	10	Aggregate visible	Loss of fine aggregate	Surface rough or pitted	<20%	20-50%	>50%	
POPOUTS	5	Not considered			<20%	20-50%	>50%	
PATCHING	5	<1 ft <sup>2</sup> , no deterioration	<1 ft <sup>2</sup> , deterioration	>1 ft <sup>2</sup>	<10/mi	10-20/mi	>20/mi	✓
PUMPING	15	some staining, rater is certain of pumping		excessive staining	<10%	10-25%	>25%	✓
SETTLEMENTS & WAVES	10	Noticeable effect on Ride	Some discomfort	Poor Ride	<2/mi. (<20%)	2-4/mi 20-50%	>4/mi (>50%)	✓
TRANSVERSE CRACK SPACING	10	CS 3-5'	CS <3'	CS <3' Many cracks intersect	<20%	20-50%	>50%	✓
LONGITUDINAL CRACKING	10	Hairline	>1/4" - 1"	>1"	<5%	5-15%	>15%	✓
PUNCHOUTS & EDGE BREAKS	15	Not rated	cracks <1/4" depress <1/2"	depress >1/2" Breaking up	<2/mi.	2-5/mi	>5/mi	✓
SPALLING	15	<1", few pieces missing	1 - 4" wide, most pieces missing	>4" wide, most pieces missing	<20%	20-50%	>50%	
PRESSURE DAMAGE	5	Not considered			<1/mi.	1 - 3/mi.	>3/mi.	

\*L = LOW                      \*\*O = OCCASIONAL                      \*\*\*STR = DISTRESS INCLUDED IN STRUCTURAL DEDUCT CALCULATIONS.  
 M = MEDIUM                F = FREQUENT  
 H = HIGH                      E = EXTENSIVE

Figure A9. Key CRC Pavement Condition Rating Form

Section: \_\_\_\_\_  
 Log mile: \_\_\_\_\_ to \_\_\_\_\_  
 Sta: \_\_\_\_\_ to \_\_\_\_\_

**C R C**

Date: \_\_\_\_\_  
 Rated by: \_\_\_\_\_  
 # of Utility Cuts \_\_\_\_\_

**PAVEMENT CONDITION RATING FORM**

DISTRESS	DISTRESS WEIGHT	SEVERITY WT.*			EXTENT WT.**			DEDUCT POINTS***
		L	M	H	O	F	E	
SURFACE DETERIORATION	10	0.4	0.7	1	0.5	0.8	1	
POPOUT	5	1	1	1	0.4	0.6	1	
PATCHING	5	0.4	0.7	1	0.5	0.8	1 ✓	
PUMPING	15	0.7	0.7	1	0.3	0.7	1 ✓	
SETTLEMENTS & WAVES	10	0.3	0.7	1	0.4	0.7	1 ✓	
TRANSVERSE CRACK SPACING	10	0.4	0.7	1	0.4	0.8	1 ✓	
LONGITUDINAL CRACKING	10	0.4	0.8	1	0.5	0.8	1 ✓	
PUNCHOUTS OR EDGE BREAKS	15	0	0.8	1	0.6	0.9	1 ✓	
SPALLING	15	0.3	0.6	1	0.5	0.8	1	
PRESSURE DAMAGE	5	1	1	1	0.7	0.9	1	

\*L = LOW                      \*\*O = OCCASIONAL  
 M = MEDIUM                F = FREQUENT  
 H = HIGH                      E = EXTENSIVE

TOTAL DEDUCT = \_\_\_\_\_  
 SUM OF STRUCTURAL DEDUCT (✓) = \_\_\_\_\_  
 100 - TOTAL DEDUCT = PCR = \_\_\_\_\_

\*\*\* DEDUCT POINTS = DISTRESS WEIGHT X SEVERITY WT. X EXTENT WT.

Figure A10. CRC Pavement Condition Rating Form

Section: \_\_\_\_\_

**KEY**

Date: \_\_\_\_\_

Log mile: \_\_\_\_\_ to \_\_\_\_\_

**BRICK PAVER CONDITION**

Rated by: \_\_\_\_\_

Sta: \_\_\_\_\_ to \_\_\_\_\_

**RATING FORM**

# of Utility Cuts: \_\_\_\_\_

DISTRESS	Distress Weight	SEVERITY*			EXTENT**			STR***
		L	M	H	O	F	E	
BRICK DETERIORATION	15	Some disintegration	Open texture / Rough. Area < 1 yd <sup>2</sup> exhibits breaks or loss of bricks	Most of surface is worn away. Areas > 1 yd <sup>2</sup> exhibit breaks or loss of bricks	<10%	10-30%	>30%	
DISCOLORATION	5	Rater is certain of Discoloration			<10%	10-30%	>30%	
PATCHING	10	<1 ft <sup>2</sup>	<1 yd <sup>2</sup>	>1 yd <sup>2</sup>	<3/500ft	3-6/500ft	>6/500ft	
PUMPING	15	Rater is Certain of Pumping			<10%	10-25%	>25%	✓
RUTTING	20	<1"	1-3"	>3"	<20%	20-50%	>50%	✓
CORRUGATIONS	5	Noticeable effect on ride	Some Discomfort	Poor Ride	<10%	10-30%	>30%	
JOINT EROSION	10	<1/2"	1/2-3/4"	>3/4"	<20%	20-50%	>50%	
BRICK SETTLEMENT	20	Area < 2 ft <sup>2</sup> , depth < 3", & Noticeable effect on ride	Area between 2 ft <sup>2</sup> & 1 yd <sup>2</sup> , depth > 3", & Some discomfort	Area > 1 yd <sup>2</sup> , depth > 3", & poor ride; or pumping	<2/500ft	2-5/500ft	>5/500ft	✓

\*L = LOW  
M = MEDIUM  
H = HIGH

\*\*O = OCCASIONAL  
F = FREQUENT  
E = EXTENSIVE

\*\*\*STR = DISTRESS INCLUDED IN STRUCTURAL DEDUCT CALCULATIONS.

Figure A11. Key Brick Paver Condition Rating Form

Section: \_\_\_\_\_

Date: \_\_\_\_\_

Log mile: \_\_\_\_\_ to \_\_\_\_\_

**Brick**

Rated by: \_\_\_\_\_

Sta: \_\_\_\_\_ to \_\_\_\_\_

# of Utility Cuts: \_\_\_\_\_

**PAVER CONDITION RATING FORM**

DISTRESS	DISTRESS WEIGHT	SEVERITY WT.*			EXTENT WT.**			DEDUCT POINTS***
		L	M	H	O	F	E	
BRICK DETERIORATION	15	0.4	0.7	1	0.5	0.8	1	
DISCOLORATION	5	1	1	1	0.4	0.7	1	
PATCHING	10	0.3	0.6	1	0.6	0.8	1	
PUMPING	15	1	1	1	0.3	0.7	1 ✓	
RUTTING	20	0.3	0.7	1	0.6	0.8	1 ✓	
CORRUGATIONS	5	0.3	0.7	1	0.4	0.7	1	
JOINT EROSION	10	0.4	0.7	1	0.5	0.8	1	
BRICK SETTLEMENT	20	0.4	0.7	1	0.6	0.8	1 ✓	

\*L = LOW  
M = MEDIUM  
H = HIGH

\*\*O = OCCASIONAL  
F = FREQUENT  
E = EXTENSIVE

TOTAL DEDUCT = \_\_\_\_\_

SUM OF STRUCTURAL DEDUCT (✓) = \_\_\_\_\_

100 - TOTAL DEDUCT = PCR = \_\_\_\_\_

\*\*\* DEDUCT POINTS = DISTRESS WEIGHT X SEVERITY WT. X EXTENT WT.

Figure A12. Brick Paver Condition Rating Form

## VITA

Rulian Ferreira de Almeida Barros was born in Rio de Janeiro, Brazil. He earned his bachelor's degree in civil engineering from the Centro Federal de Educação Tecnológica do Rio de Janeiro (CEFET-RJ), Rio de Janeiro, Brazil, in 2018. His current field of study includes transportation engineering, asphalt, concrete, and composite pavements research, and computational modeling utilizing machine learning. He began his graduate study in transportation engineering under Dr. Waheed Uddin's supervision at the University of Mississippi. After Dr. Uddin passed away, Dr. Hakan Yasarer became his advisor and contributed to several publications in the pavement and artificial neural networks fields. Rulian published and presented at several national and international conferences such as the Transportation Research Board (TRB) and American Society of Civil Engineers (ASCE) International Conference on Transportation & Development (ICTD). Besides his job as a researcher, he also worked as a teaching assistant from 2018 to 2020 and as a graduate instructor at the beginning of 2021. He received the Fall Dissertation Fellowship award in 2021 from the University of Mississippi and two student travel grants from the National Center for Transportation Infrastructure Durability and Life Extension Center (TriDurLE). He also received the Outstanding Ph.D. Student of the Civil Engineering Department at the University of Mississippi in 2022. He is a student member of the ASCE and a member of the Chi Epsilon Civil Engineering Honor Society since February of 2020. After graduation, he will start to work in a private company as a transportation engineer.