2023

WEIGH-IN-MOTION DATA-DRIVEN PAVEMENT PERFORMANCE PREDICTION MODELS

Mohammad Afsar Sujon
mas0130@mix.wvu.edu

WEIGH-IN-MOTION DATA-DRIVEN PAVEMENT PERFORMANCE PREDICTION MODELS

Mohhammad Afsar Sujon

A DISSERTATION

Submitted to the Benjamin M. Statler College of Engineering and Mineral Resources at West Virginia University in partial fulfillment of the requirements for the degree of Doctor of Philosophy in Civil and Environmental Engineering

Fei Dai, Ph.D. (Chair)

Yoojung Yoon, Ph.D.

John P. Zaniewski, Ph.D.

Roger Chen, Ph.D.

Linbing Wang, Ph.D.

Department of Civil and Environmental Engineering

Morgantown, West Virginia

2023

Keywords: Weigh-In-Motion, Pavement Performance Prediction Model, Pavement Condition

ABSTRACT

WEIGH-IN-MOTION DATA-DRIVEN PAVEMENT PERFORMANCE PREDICTION MODELS

Mohhammad Afsar Sujon

The effective functioning of pavements as a critical component of the transportation system necessitates the implementation of ongoing maintenance programs to safeguard this significant and valuable infrastructure and guarantee its optimal performance. The maintenance, rehabilitation, and reconstruction (MRR) program of the pavement structure is dependent on a multidimensional decision-making process, which considers the existing pavement structural condition and the anticipated future performance. Pavement Performance Prediction Models (PPPMs) have become indispensable tools for the efficient implementation of the MRR program and the minimization of associated costs by providing precise predictions of distress and roughness based on inventory and monitoring data concerning the pavement structure's state, traffic load, and climatic conditions. The integration of PPPMs has become a vital component of Pavement Management Systems (PMSs), facilitating the optimization, prioritization, scheduling, and selection of maintenance strategies. Researchers have developed several PPPMs with differing objectives, and each PPPM has demonstrated distinct strengths and weaknesses regarding its applicability, implementation process, and data requirements for development. Traditional statistical models, such as linear regression, are inadequate in handling complex nonlinear relationships between variables and often generate less precise results.

Machine Learning (ML)-based models have become increasingly popular due to their ability to manage vast amounts of data and identify meaningful relationships between them to generate informative insights for better predictions. To create ML models for pavement performance prediction, it is necessary to gather a significant amount of historical data on pavement and traffic loading conditions. The Long-Term Pavement Performance Program (LTPP) initiated by the Federal Highway Administration (FHWA) offers a comprehensive repository of data on the environment, traffic, inventory, monitoring, maintenance, and rehabilitation works that can be utilized to develop PPPMs. The LTPP also includes Weigh-In-Motion (WIM) data that provides information on traffic, such as truck traffic, total traffic, directional distribution, and the number of different axle types of vehicles. High-quality traffic loading data can play an essential role in improving the performance of PPPMs, as the Mechanistic-Empirical Pavement Design Guide (MEPDG) considers vehicle types and axle load characteristics to be critical inputs for pavement design.

The collection of high-quality traffic loading data has been a challenge in developing Pavement Performance Prediction Models (PPPMs). The Weigh-In-Motion (WIM) system, which comprises WIM scales, has emerged as an innovative solution to address this issue. By leveraging computer vision and machine learning techniques, WIM systems can collect accurate data on vehicle type and axle load characteristics, which are critical factors affecting the performance of

flexible pavements. Excessive dynamic loading caused by heavy vehicles can result in the early disintegration of the pavement structure. The Long-Term Pavement Performance Program (LTPP) provides an extensive repository of WIM data that can be utilized to develop accurate PPPMs for predicting pavement future behavior and tolerance. The incorporation of comprehensive WIM data collected from LTPP has the potential to significantly improve the accuracy and effectiveness of PPPMs.

To develop artificial neural network (ANN) based pavement performance prediction models (PPPMs) for seven distinct performance indicators, including IRI, longitudinal crack, transverse crack, fatigue crack, potholes, polished aggregate, and patch failure, a total of 300 pavement sections with WIM data were selected from the United States of America. Data collection spanned 20 years, from 2001 to 2020, and included information on pavement age, material properties, climatic properties, structural properties, and traffic-related characteristics. The primary dataset was then divided into two distinct subsets: one which included WIM-generated traffic data and another which excluded WIM-generated traffic data. Data cleaning and normalization were meticulously performed using the Z-score normalization method. Each subset was further divided into two separate groups: the first containing 15 years of data for model training and the latter containing 5 years of data for testing purposes. Principal Component Analysis (PCA) was then employed to reduce the number of input variables for the model. Based on a cumulative Proportion of Variation (*PoV*) of 96%, 12 input variables were selected. Subsequently, a single hidden layer ANN model with 12 neurons was generated for each performance indicator.

The study's results indicate that incorporating Weigh-In-Motion (WIM)-generated traffic loading data can significantly enhance the accuracy and efficacy of pavement performance prediction models (PPPMs). This improvement further supports the suitability of optimized pavement maintenance scheduling with minimal costs, while also ensuring timely repairs to promote acceptable serviceability and structural stability of the pavement. The contributions of this research are twofold: first, it provides an enhanced understanding of the positive impacts that high-quality traffic loading data has on pavement conditions; and second, it explores potential applications of WIM data within the Pavement Management System (PMS).

ACKNOWLEDGEMENTS

Primarily, I would like to express my sincere gratitude and appreciation to my amazing thesis advisor, Dr. Fei Dai for his continuous support of my Ph.D. study and research, and his motivation, enthusiasm, encouragement, and immense knowledge. His intellectual guidance, insightful comments, constructive criticism, and mentorship in each step of my Ph.D. journey have helped me understand how to be organized, and disciplined, how to proceed step by step with ideas in research, and how to write and improve my academic writing skills. Besides my advisor, I would also like to extend my immense gratitude to Dr. Yoojung Yoon, Dr. John P. Zaniewski, Dr. Roger Chen, and Dr. Linbing Wang for serving as my committee members and for providing valuable suggestions, assistance, constructive and useful comments regarding my research and dissertation.

I would especially like to thank the Department of Civil and Environmental Engineering of West Virginia University for the resources and the logistic support for this study. I am also thankful to my fellow lab mate Linjun Lu, for the research-related discussion that we used to have in the lab and for all the fun we did while working together in the lab. I extend my deep appreciation to my father, Mohhammad Abul Kashem, mother, Ferdushi begum, elder brother, Mohammad Kawsar Sumon, and younger sister, Saidatun Kownine Meem their unconditional love, moral support, encouragement, and faith in me. I would also like to thank my father-in-law, Mohammad Abul Mobarak, and mother-in-law, Nasima Parvin Khan for their immense support.

Most of all, I owe my deepest gratitude to my dear wife, Mashiat Mobarak Bushra, for her affection, love, encouragement, understanding, and patience. She stood by me when I used to go through a challenging time with my research and become impatient, perplexed, and hopeless. She continuously used to try to motivate me and help me by discussing ideas, admiring what I am doing, and pointing out how I can make things better. I would like to thank her for all her help and support throughout my entire Ph.D. journey.

Lastly, I am thankful for being a part of such a beautiful campus of West Virginia University which amazes me with its scenic beauty and colors in the fall and spring seasons.

Table of Contents

List of Tables

List of Figures

Chapter 1. Introduction

1.1 Motivation

The structural stability of the pavement is generally affected by factors such as traffic volume, environmental conditions, material properties, and design considerations, with traffic loads playing a major role in pavement deterioration resulting from overweight vehicles on the road (Salama et al. 2006). An adequate pavement management system is recommended by transportation agencies to effectively allot their insufficient resources to optimally preferred projects at optimal times and to the most cost-effective maintenance treatments. Predicting the future pavement condition accounting for the ambiguity inherently associated with pavement condition data while incorporating the positive and negative effects of relevant influential factors should facilitate the development of an efficient pavement management system.

Pavement Performance Prediction Models (PPPMs) are regarded as an essential tool to provide an optimal allocation of resources in maintenance activities (Mahmood 2015). These models are generated using inventory and monitoring data regarding the state of pavement structure, traffic load, and climate conditions. PPPMs can be classified based on their type of formulation, conceptual format, application level, and type of variables (Justo-Silva et al. 2021). The main advantage of using ML algorithms is the inherent ability to 'learn' from historical data-oriented information fed to it by using computational methods.

ANN models are artificial intelligence-based ML models suited for solving complex problems and can adapt to dynamic environments in real-time. So, ANN provides an excellent tool for dealing with the complexity of pavement structures and the inherent non-linearity of the measured data (Justo-Silva et al. 2021). In the context of data, most researchers used only annual average daily traffic (AADT) data as traffic input while Mechanistic-Empirical Pavement Design

Guide (MEPDG) requires an extensive amount of traffic inputs for design/analysis of pavement systems such as vehicle characteristics and axle characteristics (AASHTO 1993; ARA 2004). As MEPDG uses analysis software (Pavement-ME) to predict pavement life under different traffic loading scenarios. Different traffic distribution patterns considering the overweight and non-overweight traffic in terms of truck classes and axle load had shown a significant difference in pavement performance (Wang et al. 2015). Traffic loading on pavements is generally represented by a collection of specific types of vehicles with variations in load magnitude, number of axles, and axle configuration considered as an important factor in determining the performance of the pavement (Tran and Hall 2007). So, high-quality traffic load data can have a significant impact on the accuracy of PPPMs which can produce better predictions about future pavement conditions.

Weight enforcement plays a significant role in regulating overweight vehicles on the road. There is a great concern that overweight vehicles can cause damage to the roadway system and significantly reduce the performance and service life of the pavement. Trucks circulating excessively overloaded on some roads in Colombia resulted in accelerated deterioration of the pavement structures in terms of fatigue, cracks, and ruts (Fuentes et al. 2012). Moreover, overloaded trucks can result in more traffic accidents and loss of properties and lives (Zhang et al. 2020). Typically, weigh stations are operated to impose weight enforcement. However, overweight trucks around these stations frequently seek to avert weigh stations due to loss of trucking time, possible hassles, and fines (Taylor et al. 2000). WIM can serve as a weight measurement device for vehicles moving at high speeds and reduce needless stops and delays implanted with a stronger invasive type of control enforcement. Further, WIM can cooperate in the process of generating innovative programs to deal with pavement management issues with high-quality data collection (Khalili et al. 2022).

Little work assessed the importance of high-quality traffic load data for developing PPPMs, due to the unavailability of detailed traffic load data and the complexities in collecting high-quality traffic data on highways. This research hypothesized that high-quality traffic load data collected through WIM has an immense potential to understand the detailed traffic loading process. Prudent analysis of the collected high-quality traffic load data can support the development of a sophisticated pavement performance prediction process. Therefore, this research proposed developing PPPMs to better predict the pavement's condition. For this purpose, a set of novel PPPMs were developed utilizing a dataset developed by combining WIM data along with other types of data from LTPP. After cleaning and processing the data, supervised ANN-based PPPMs will be developed utilizing these data for better prediction of specific performance indicators.

1.2 Research Objectives

To achieve the goal of developing WIM data-focused PPPMs, this research work aimed to explore the applicability of machine learning techniques using WIM-generated high-quality traffic load data with other required data collected from the Long-Term Pavement Performance Program (LTPP) for developing PPPMs. Leveraging the application of the data collected from the WIM systems in the United States of America and utilizing traffic characteristics information such as vehicle type to understand the impacts on the improvement of the performance of the PPPMs. The development of PPPMs is aimed to predict the condition of the pavement in terms of the International Roughness Index (IRI), cracking, rutting, potholes, and patch failure utilizing machine learning techniques so that model-generated information can be exploited for further research. The overarching goal of the current research is to develop pavement performance models that support enhanced pavement management systems assisting accurate prediction of pavement conditions and effectively allocating the limited resources of highway agencies to maintain pavements in the desired condition.

1.3 Research Overview

The general framework for the two research objectives of the current research is shown in Figure 1.1. First, the distress types of the flexible pavement and factors affecting the condition of the pavement were investigated through a literature study to understand the significance of pavement condition and find the limitations of current available PPPMs. After identifying the limitations, the potential of WIM for collecting high-quality traffic load data was investigated by understanding the technologies and techniques associated with the development and data generation process of WIM. Later, to understand the need for better prediction of the pavement condition, available pavement evaluation models and techniques that are most suitable for the selection of appropriate pavement performance prediction models based on the effects of traffic, climate, and other factors were analyzed.

After analysis, machine learning algorithm-based Artificial Neural Network (ANN) models were developed to build pavement performance prediction models based on collected and processed LTPP data. The development of PPPMs was implemented into an executable computer program using standard Python programming language. The results of the developed PPPMs were evaluated against the real pavement condition of the LTPP sections for understanding their performance.

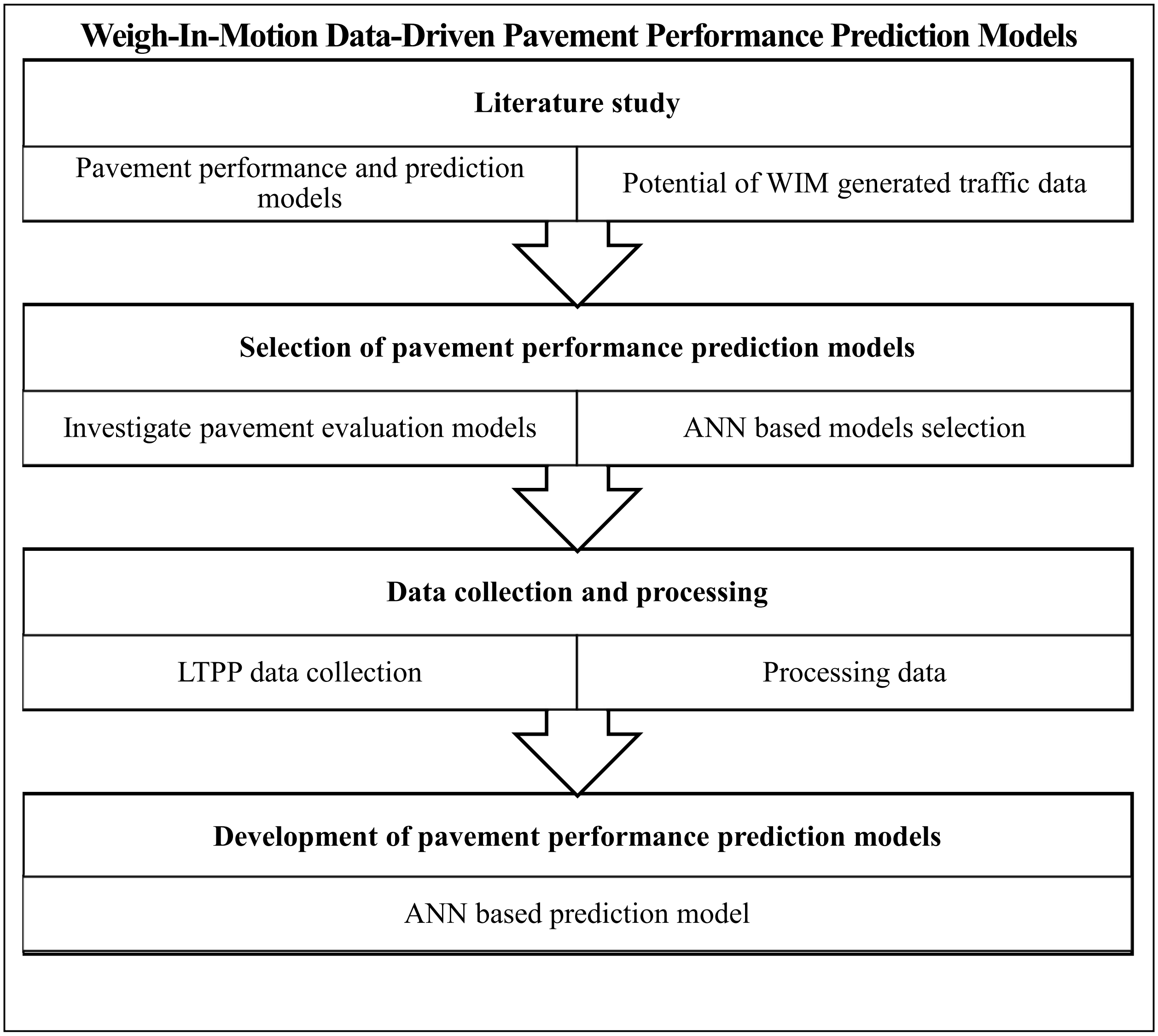


Figure 1.1: Overall framework of the research

1.4 Organization of Dissertation

In *Chapter 1*, the motivation, the overall objective, the specific objectives, and the research overview has been discussed.

In *Chapter 2*, the current state of pavement performance prediction models has been assessed to understand their limitations. In addition, the key aspects of WIM systems were evaluated to leverage the potential applications of high-quality traffic load data generated from these systems.

In *Chapter 3*, the overall methodology of the development of PPPMs with the process of data collection and processing to make them appropriate to be utilized for predicting pavement conditions was discussed.

In *Chapter 4*, the detailed development process of ANN-based PPPMs utilizing the LTPP data with analysis of results was presented to discuss their performance in the prediction of pavement conditions.

In *Chapter 5*, a general discussion based on the findings of the study and recommendations has been provided. Limitations, future extensions, conclusions, and key contributions of the study are also discussed in this chapter.

Chapter 2. Background Study

2.1 Flexible Pavement Condition and Distresses

Pavement systems are constructed to facilitate vehicles' smooth movement over the surface and provide safe passage under diverse climates with adequate serviceability performance. There are three general types of pavement are constructed and they are classified based on their construction materials or surface types specifically flexible or Asphalt, rigid (Portland Cement Concrete, PCC), and composite (Asphalt and PCC). The basic components of the typical flexible pavement system are presented in Figure 2.1.

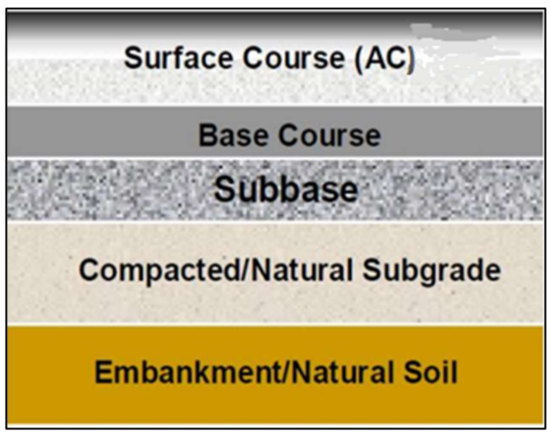


Figure 2.1: Flexible pavement system's basic component (Christopher et al. 2006)

Flexible pavements are constructed using asphalt concrete that is placed over base layers and subgrade layers. High-quality hot mix asphalt (HMA) is used for the construction of the asphaltic surface layer, otherwise, the strength and stiffness of the layer will be lower. The stability and rigidity of the flexible pavements heavily rely on the strength of underlying unbound layers that supplement the load-carrying capability of the asphaltic surface layer. Rigid pavements usually consist of a Portland cement concrete layer that spread over the subgrade layer with or without the existence of an intermediate base layer. Composite pavements are the outcomes of the pavement reconstruction process where the damaged asphalt concrete layer is restored by Portland concrete and vice versa. For flexible pavements, the uniform stresses and nonuniform deflections transfer through the flexible layer whereas, for rigid pavements, this process is carried by the rigid layer. For the improvement of resistance to environmental factors, sometimes seal coats are applied.

Based on the ‘Distress Identification Manual for the Long-Term Pavement Performance Program (Fifth Revised Edition)’ (S. Miller and Y. Bellinger 2014), common distress in flexible pavements are categorized into five groups: 1) cracking, 2) patching, and potholes, 3) surface deformation, 4) surface defects, and 5) miscellaneous distresses. Cracking is subcategorized as fatigue (alligator) cracking, block cracking, edge cracking, longitudinal cracking, and transverse cracking. Surface deformation is also divided into rutting and shoving. Surface defects are subcategorized as bleeding, polished aggregate, and raveling. The data collected for distress conditions are categorized into three severity levels: low, medium, and high. Moreover, the severity level of the crack is a function of the crack opening, which is susceptible to the pavement temperature at the time of data collection. Thus, the crack opening can be considered as “high severity” in one year and “medium severity” in the next year, and vice versa (Baladi et al. 2017). Table 2.1 showed the distress types in the flexible pavement.

Table 2.1: Flexible pavement distress types (S. Miller and Y. Bellinger 2014)

Distress		Area of distress/measurement area	Define severity levels
Category	**Type**		
Cracking	Fatigue (alligator) cracking	Square meters	Yes
	Block cracking	Square meters	Yes
	Edge cracking	Meters	Yes
	Longitudinal cracking	Meters	Yes
	Transverse cracking	Number, meters	Yes
Patching and Potholes	Patch/patch deterioration	Number, square meters	Yes
	Potholes	Number, square meters	Yes
Surface Deformation	Rutting	Millimeters	No
	Shoving	Number, square meters	No
Surface Defects	Bleeding	Square meters	No
	Polished aggregate	Square meters	No
	Raveling	Square meters	No

The descriptions of the major flexible pavement distress types are presented as follows (S. Miller and Y. Bellinger 2014):

1) Fatigue (alligator) Cracking: This type of cracking commonly appears in specific areas that experience repeated traffic loadings, and the crack can form an interconnected series at the initial stages of development. Cracks can be expanded into many-sided, sharp-angled pieces that are typically less than 0.3 meters on the longest side, and alligators' skin or chicken wire appear in later stages. For this, fatigue crack is also referred to as alligator cracking.

Figure 2.2: Fatigue (Alligator) Cracking in flexible pavements (Shatnawi 2008)

2) Block Cracking: This type of cracking shows a pattern of cracks dividing the pavement into rectangular pieces. The size of the rectangular block ranges from 0.1-meter square to 11-meter square.

Figure 2.3: Block Cracking in flexible pavements (Shatnawi 2008)

3) Edge Cracking: This type of cracking builds up in the shape of crescents. Edge Cracking is typically located within 0.6 meters of the pavement edge that borders the unpaved shoulder. There are longitudinal cracks outside of the wheel path and within 0.6 meters of the pavement edge in this type of cracking.

Figure 2.4: Edge Cracking in flexible pavements (Shatnawi 2008)

4) Longitudinal Cracking: This type of cracking is developed parallel to the pavement centerline and mostly takes place within the lane (both wheel path and non-wheel path). The level of the crack in width: low ≤ 6 mm, moderate > 6 mm and ≤ 19 mm, and high > 19 mm.

Figure 2.5: Longitudinal Cracking in flexible pavements (Shatnawi 2008)

5) Transverse Cracking: This type of crack is formed perpendicular to the pavement centerline. The level is the same as the longitudinal crack. Transverse cracking is formed typically for thermal stress on the pavement.

Figure 2.6: Transverse Cracking in flexible pavements (Shatnawi 2008)

6) Patch/patch deterioration: This type of distress appeared when the portion of the pavement surface greater than or equal to 0.1 m^2 is removed and replaced or additional material is applied after the original construction of the pavement. The levels of severity are classified as low < 6 mm, moderate 6 to 12 mm, and high > 12 mm. Patch failure occurs due to heavy traffic and low-quality construction. Distresses that take place in the patched area affect the severity level of the patch.

Figure 2.7: Patch Failure in flexible pavements (Shatnawi 2008)

7) Potholes: Distress type appeared as bowl-shaped holes formed in assorted sizes on the pavement surface. The dimension for the minimum plan is 150 mm as circular potholes should have a minimum diameter of 150 mm. The level of severity is defined as < 25 mm deep for low, 25 to 50 mm deep for moderate, and > 50 mm deep for high.

Figure 2.8: Potholes in flexible pavements (Shatnawi 2008)

8) Rutting: The longitudinal surface depression that develops in the wheel paths of flexible pavement is a result of this type of deformation. It may have had an associated transverse displacement.

Figure 2.9: Rutting in flexible pavements (Shatnawi 2008)

9) Shoving: There is a longitudinal displacement of the pavement surface defying this type of deformation. It can be caused by braking or speeding vehicles and is usually located on hills or curves. It is possible that it also has vertical displacement.

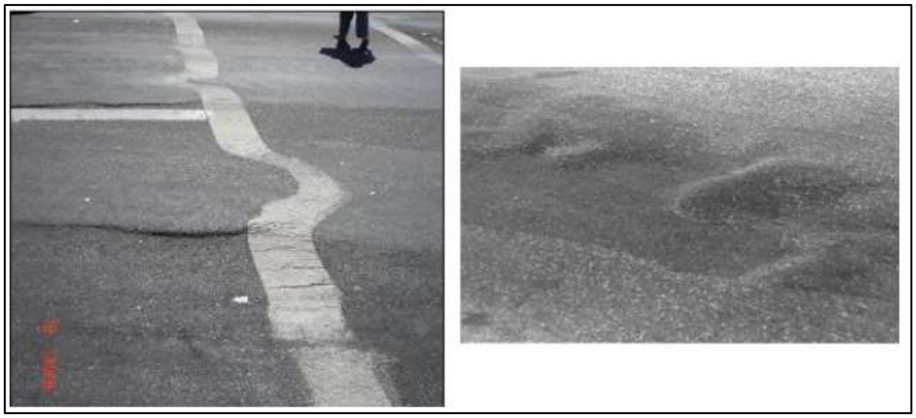


Figure 2.10: Shoving in flexible pavements (Shatnawi 2008)

10) Bleeding: Excess bituminous binder on the pavement surface caused this type of surface defect. It is possible to create a shiny, glass-like reflective surface that is tacky to the touch. It's usually found in the wheel paths.

Figure 2.11: Bleeding in flexible pavements (Shatnawi 2008)

11) Polished Aggregate: This distress is formed by wearing away the surface binder exposing the coarse aggregate of the pavement. The degree of polishing may be reflected in a reduction of

surface friction. Polished aggregates are not rated on the test sections that have received a preventive maintenance treatment that has covered the original pavement surface.

Figure 2.12: Polished Aggregate in flexible pavements (Shatnawi 2008)

12) Raveling: Distress type produced by disintegration of the pavement surface resulting from aggregate particles displacing and asphalt binder loss. The range of raveling is calculated based on fine aggregate to some coarse aggregate loss and eventually forms an extremely rough and pitted surface produced by the loss of aggregates.

Figure 2.13: Raveling in flexible pavements (Shatnawi 2008)

In addition to these distress types, pavement condition is also evaluated based on pavement roughness or riding quality. Pavement roughness is a measurement of pavement distortion of longitudinal profile negatively affecting the comfort level of the vehicle's user. Pavement roughness, as defined by ASTM Standard E867, is ''the deviations of a pavement surface from a true surface with characteristic dimensions that affect vehicle dynamics, ride quality, dynamic loads, and drainage''(Yeganeh et al. 2019). It is an extremely important indicator of pavement condition as it affects not only the quality of the ride but also causes extra fuel consumption, delay cost, and additional vehicle maintenance cost (Gong et al. 2018). The International Roughness Index (IRI) is used to indicate pavement roughness using the average rectified slope (accumulated suspension motion to distance traveled), as derived from a mathematical model of a standard quarter car passing over a measured profile at a speed of 50 mph (Ozbay and Laub 2001). Predicting the progression of roughness during pavement life is important for PMS decision-making.

The degradation of pavements and their concomitant failures constitute a vital concern for the management of transportation infrastructure, with a myriad of factors being known to contribute to this process. These factors can be broadly classified as either intrinsic or extrinsic, and their effect on the pace of pavement deterioration can be considerable. Intrinsic factors pertain to the inherent characteristics of the natural constituents contained within the pavement materials utilized during the construction stage, and these may manifest in a diversity of forms, including plastic deformation, cracking, fatigue, or the structural configuration of the pavement. For instance, the thickness of individual layers can exert a significant influence on the performance of the pavement structure, and it can also impact stress distribution, thereby resulting in pavement failure (Park and Kim 2019).

Numerous antecedent investigations have delineated a range of factors that impinge upon the degradation of pavement performance. These studies have scrutinized the impact of individual factors and have ascertained that pavement age constitutes the most statistically significant variable for prognosticating pavement performance. Age-related factors, such as environmental exposure, traffic load, and maintenance history, can lead to the deterioration of pavement structure over time. Moreover, the interplay between intrinsic and extrinsic factors can engender intricate pavement behavior, thereby impeding precise forecasts of pavement performance. Therefore, a comprehensive grasp of the fundamental factors that contribute to pavement deterioration is indispensable for the effective management and upkeep of pavements (Abaza 2004; Kim and Kim 2006; Rajagopal 2006). Ahmed et al. (2016) used traffic loading and climate conditions to predict pavement deterioration. Subgrade resilient modulus (Abaza 2004; Hong and Somo 2001), pavement treatment expenditure (Montgomery et al. 2018), and construction quality (Rose et al. 2018) are variables that were used to predict pavement performance. Figure 2.14 presents common factors affecting pavement conditions (Plati 2019).

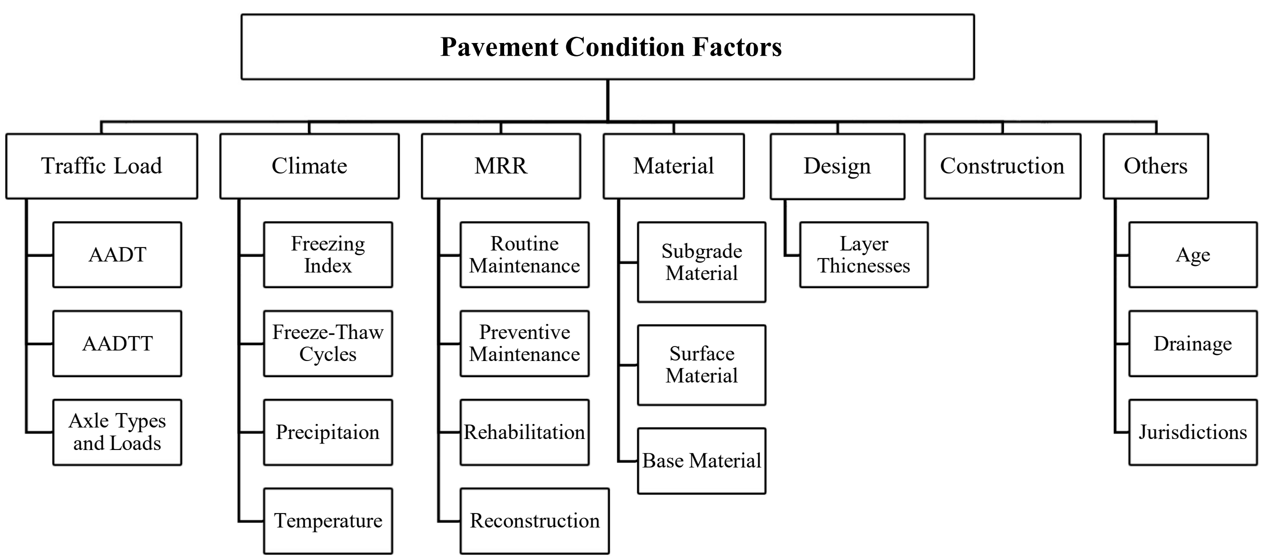


Figure 2.14: Common factors affecting pavement condition (Plati 2019)

2.2 Pavement Performance Prediction Models

To competently manage pavement structures, encompassing project planning, design, construction, maintenance, and rehabilitation, transportation agencies must adopt a methodical approach. Because of the constraints on maintenance budgets and the escalating number of infrastructure elements concluding their design life, cost-effective maintenance strategies are gaining traction for infrastructure asset management. This includes preserving the condition of pavement at a satisfactory level that is commensurate with the demands of traffic while concurrently mitigating environmental pollution throughout its service life (Babashamsi et al. 2016). Pavement management systems (PMS) assume a vital function in adopting cost-effective strategies after comprehending the extant conditions of pavement structures. As expounded upon in the preceding segment, the conditions of pavement structures are typified by surface deformation, roughness/riding quality, cracking, surface friction (skid resistance), or faulting. Owing to the constraints of extant pavement condition prognostication practices, prognosticating distresses based on continually amassed traffic and climatic data via sensors can serve as a cost-effective technique for data collection that can explicate the deterioration process of pavements. Acknowledging its significance as an integral component of PMSs can support dynamic and economical management. To predict pavement conditions, pavement performance prediction models (PPPMs) can constitute an approach for developing cost-effective strategies in PMS. Discerning the potential of PPPMs, these models are employed for the subsequent objectives (Kargah-Ostadi et al. 2019):

- The utilization of pavement performance prediction models (PPPMs) can enable the prediction of both the present and future conditions of pavements. The resultant data is then used to plan, prioritize, schedule, optimize, and select appropriate maintenance treatments.

- Subsequently, the impacts of planned maintenance treatments are analyzed by prognosticating the pavement's future condition. Additionally, the cost associated with the pavement's life cycle under different maintenance scenarios can be estimated.

The assessment of pavement structure conditions utilizing data generated from sensors constitutes a pivotal aspect of the pavement management system. In this context, pavement performance prediction models (PPPMs) are regarded as mathematical functions that are commonly employed to predict the performance of pavements by establishing a correlation between pavement conditions (such as cracking and rutting) and a set of explanatory variables (including traffic loadings, age, environmental factors, and pavement design characteristics). By leveraging these models, transportation agencies can anticipate potential pavement distresses, thereby facilitating the implementation of appropriate maintenance strategies cost-effectively. This underscores the importance of PPPMs as a key component in the overall pavement management system. Details of the classification are presented in Figure 2.15 according to (Marcelino et al. 2021):

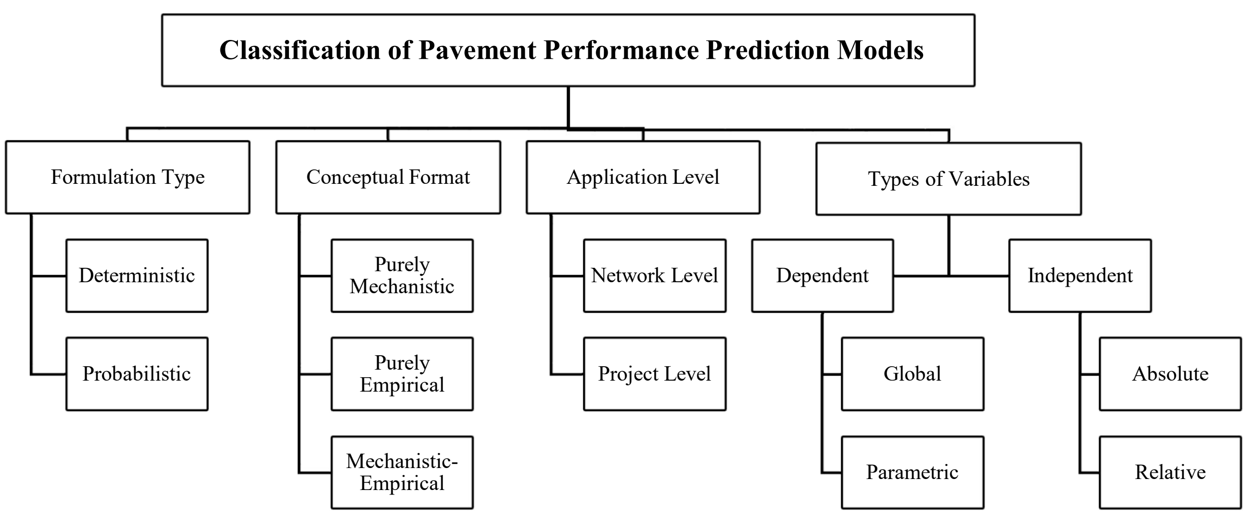


Figure 2.15: Classification of Pavement Performance Prediction Models (Marcelino et al. 2021)

Although machine learning (ML) and statistics employ similar methods, they differ in their underlying objectives. While statistics are primarily concerned with concluding data that has already been collected, machine learning seeks to develop algorithms and models that can be applied to new, previously unseen datasets to make accurate predictions. Therefore, the fundamental goals of these two fields are distinct, despite their methodological similarities (Damirchilo et al. 2021). Machine learning (ML) algorithms utilize computational techniques to learn from historical data or prior experiences, and their performance improves as more training data is incorporated. Through the analysis of large datasets, these algorithms can detect underlying patterns and use this knowledge to make accurate predictions. Machine learning models can be classified into three primary categories:

- Supervised learning—can be used for project-level or network-level pavement management.
- Unsupervised learning—can be used for exploratory and clustering analysis.
- Reinforcement learning—can be used to help decision-makers for both project- and network-level pavement management.

According to Justo-Silva et al. (2021), a summary of the machine-learning algorithms is presented in Figure 2.16.

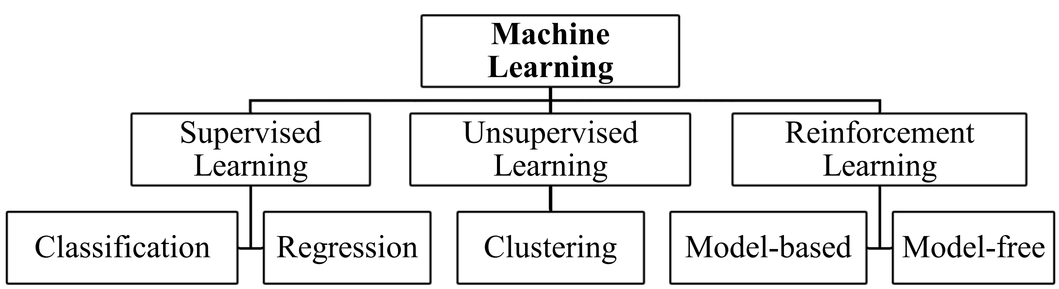


Figure 2.16: Summary of machine learning algorithms (Justo-Silva et al. 2021)

Supervised learning is a type of machine learning that involves the use of input and output data to build a model capable of making predictions. In the context of project management where the goal is to predict a continuous output or target variable, regression machine learning techniques are employed. Linear regression, neural networks, decision trees, and adaptive neural-fuzzy learning are some of the most used regression models. Linear regression models express the relationship between the response variable and one or more predictor variables in a linear fashion, which makes them easy to understand and train. As a result, they are often used as the initial model for new datasets. Nonlinear regression is a statistical method used to model relationships in data, with "nonlinear" referring to the fact that the fitness function is nonlinear concerning the parameters. Nonlinear regression is best suited for datasets with strong nonlinear trends that cannot be easily transformed into a linear space. Gaussian process regression models are used to predict the value of a continuous response variable and are frequently used in spatial analysis when uncertainty is present. Support vector machine (SVM) regression is a variant of SVM classification that is adapted to predict a continuous response variable. Instead of finding a hyperplane that separates the data, SVM regression models identify a model that deviates from the observed data by no more than a predetermined amount, with parameter values that are as small as possible (Justo-Silva et al. 2021).

In machine learning, classification algorithms are utilized when the objective is to predict a categorical or discrete output. Common classification algorithms include Logistic Regression, Decision Trees (bagged and boosted), Support Vector Machine (SVM), K-Nearest Neighbor (KNN), Neural Networks, Naïve Bayes, and Discriminant Analysis. Logistic Regression is a simple algorithm that predicts the probability of a response belonging to one class or another and is often used for binary classification problems. K-Nearest Neighbor (KNN) assigns categories to objects based on the classes of their nearest neighbors in the data set. SVM finds the linear decision boundary that separates all data points of one class from those of the other class. The optimal hyperplane is the one with the largest margin between the two classes when the data is linearly separable. If the data is not linearly separable, a loss function is used to penalize points on the hyperplane's wrong side. Neural Networks consist of interconnected networks of neurons and iteratively adjust the strengths of the connections to map inputs to the correct response. This algorithm is best used for updating the model regularly when data is available, and model interpretability is not a key concern. Naïve Bayes assumes that the presence of a particular feature in a class is independent of the presence of any other feature. Discriminant Analysis is employed to identify features by assuming that different classes generate data. The decision tree algorithm predicts responses by following the tree's decisions from root to leaf, and the values of weights and the number of branches are determined during the training process (Justo-Silva et al. 2021).

Unsupervised learning is a potent methodology for discovering patterns and extracting insights from data, especially for exploratory data analysis and clustering. Various learning techniques are available, including Hierarchical clustering, K-means and k-medoids, Hidden Markov models, and Fuzzy c-means. Clustering methods group data points based on a distance metric, and can be useful for hypothesis development, modeling smaller subsets of data, data reduction, and identifying outliers, even when the goal is supervised learning. Clustering algorithms can be categorized as either hard clustering, in which each data point belongs to a single cluster, or soft clustering, in which each data point may belong to multiple clusters. K-Means is a hard clustering algorithm that divides data into a pre-determined number of mutually exclusive clusters based on the distance of each data point to the center of its assigned cluster. K-Medoids are comparable to K-Means but necessitate that the cluster centers correspond to actual data points, making it suitable for quick clustering of categorical data and scaling large data sets. Hierarchical Clustering generates nested clusters by analyzing similarities between pairs of data points and grouping them into a binary, hierarchical tree. This method is particularly useful when the number of clusters in the data is unknown in advance, and when visualization is needed to guide selection. Finally, Fuzzy c-Means is a soft clustering algorithm that enables data points to belong to more than one cluster, making it a valuable tool when clusters overlap, and the number of clusters is known. Gaussian Mixture Model is used when data points originate from different normal distributions with varying probabilities (Justo-Silva et al. 2021).

Reinforcement learning is a data-driven approach that focuses on dynamic environments, where the objective is to identify the optimal sequence of actions that will maximize long-term rewards. This involves the agent or algorithm learning from the environment to develop the best policy. Reinforcement learning can be classified into two main categories: Model-based and Model-free reinforcement learning, both of which offer distinct advantages and disadvantages. Model-based reinforcement learning employs a model of the environment, which is used to simulate the outcomes of different actions, whereas Model-free reinforcement learning does not use a model and instead relies on direct experience with the environment.

2.3 Previous Work on Pavement Performance Prediction Models

Data mining techniques are gaining popularity because of their ability to overcome the drawbacks of traditional methods because they can yield more reliable and satisfying results in pavement engineering (Ashtiani et al. 2018; Bianchini and Bandini 2010; Kaur and Datta 2007; Nemati and Dave 2018; Nitsche et al. 2014). Data mining can show hidden relationships between input and output variables through the methods of randomizations, clustering, resampling, recursion, visualization, and visualization. Solatifar and Lavasani (2020) and Nian et al. (2022)

demonstrated that neural networks can be used to calculate the contribution of material and construction variables to pavement performance.

Numerous investigations have been undertaken in the domain of pavement performance modeling, encompassing a range of factors that impinge upon pavement durability. However, the majority of extant models have encountered formidable obstacles, such as the intricacies of handling numerous input variables, limitations in the accessibility of certain variables, and interdependence between the variables themselves (Kargah-Ostadi and Stoffels 2015). Recently, ANN models have been widely used to simulate the biological nervous systems in the human brain. The biological nervous system contains billions of neuron cells, and each neuron receives inputs from other neurons, processes them by a transfer function, and sends its output to the next layer (Mehta et al. 2008).

ANN models use data to build prediction models. An ANN can be defined as, "A computational mechanism with an ability to acquire, represent, and compute mapping from one multivariate space of information to another, given a set of data representing that mapping" (Rafiq et al. 2001). ANN techniques can solve complex problems because of the capability of interconnecting neurons between layers to achieve the computation of large data volumes (Basheer and Hajmeer 2000). Engineers often are faced with incomplete or noisy data, so ANN models may be the most appropriate models for recognizing meaningful relationships from data patterns to solve a particular problem (Rafiq et al. 2001). Zhang et al. (1998) reported that ANN models can predict nonlinear relationships between variables as well as traditional models that are usually used to predict these relationships.

ANN models have been widely used in different civil engineering areas with good results because they are accurate and convenient (Karlaftis and Vlahogianni 2011). Adeli (2001) conducted a review of the neural network model literature from 1989 to 2000, with a focus on structural engineering, construction engineering, and management, and reported that ANN models are suitable for modeling complex problems. ANN models can be employed for evaluating current and predicting future pavement conditions as well as for assessing maintenance needs and selecting maintenance strategies (Yamany et al. 2020).

Other, more recent studies have shown the robustness of ANN models compared to regression models. For example, the comparison between ANN and autoregressive time series models for forecasting freeway speeds showed that neural networks provide more accurate predictions than classical statistical approaches (Vlahogianni and Karlaftis 2013). Golshani et al. (2018) compared the prediction capabilities of traditional statistical models and neural network models for modeling two critical trip-related decisions related to travel mode and departure time Their results show that the neural network models offered better performance with an easier and faster implementation process. Felker et al. (2004) reported that the ANN models provided a high R^2 in predicting roughness for jointed Portland cement concrete pavements with $R^2 = 0.90$, while the statistical analysis approach yielded $R^2 = 0.73$ (2004). In Kargah-Ostadi et al. (2010), the ANN model also performed successfully in predicting IRI values using complex input variables. ANN models also have been used to predict the cracking index for Florida's highways and were found to be more accurate than an autoregressive model (Lou et al. 2001).

Further, Attoh-Okine (1994) reported two benefits of using ANN over more traditional statistical prediction models: ANN models can handle unseen data and generalize results and they can solve complex problems because of their massive parallelism and strong interconnectivity. The literature indicates that researchers have used ANN models to predict pavement performance since at least the 1990s and that ANN pavement performance models are powerful modeling tools.

However, most of the existing studies in predicting pavement performance have focused on a specific pavement type at the project management level. Further, many models have not included all the parameters that might impact pavement performance because of a lack of data, and many previous studies do not quantify the impact of input variables such as weather conditions on the ANN model predictions. Table 2.2 presents some of the existing machine learning-based pavement performance prediction models.

Table 2.2: Machine Learning-based Pavement Performance Prediction Models

Study identification	**Models used**	**Pavement type**	**Data source**
(Abdelaziz et al. 2020)	ANN, LR	Flexible	LTPP
(Bayrak et al. 2004)	ANN	Rigid	LTPP
(Chandra et al. 2013)	ANN, NLR	Flexible	Field Data (India)
(Choi and Do 2020)	ANN, MLR	Flexible	LTPP
(El-Hakim and El-Badawy 2013)	ANN, MEPDG	Rigid	LTPP
(Georgiou et al. 2018)	ANN, SVM	NA	Field Data (Greece)
(Gong et al. 2018)	RF, RR	Flexible	LTPP
(Hossain et al. 2020)	ANN	Rigid	LTPP
(Kargah-Ostadi and Stoffels 2015)	ANN, SVM	Flexible	LTPP
(Lin et al. 2003)	ANN	Flexible	Field Data (Taiwan)
(Marcelino et al. 2021)	RF	NA	LTPP

(Marcelino et al. 2020)	RF	Flexible	LTPP and Portuguese Road Administration Database
(Mazari and Rodriguez 2016)	ANN	Flexible	LTPP
(Ozbay and Laub 2001)	ANN	Flexible	LTPP
(Sollazzo et al. 2017)	ANN, LR	Flexible	LTPP
(Yamany et al. 2020)	ANN, LR, RPR	Flexible	LTPP
(Zeiada et al. 2020)	ANN, SVM, LR, QLR, PLSR	Flexible	LTPP
(Ziari et al. 2016)	SVM	Flexible	LTPP
(Ziari et al. 2016b)	ANN	Flexible	LTPP
Note: ANN = artificial neural network; LR = logistic regression; MLR = multiple linear regression; RF = random forest; RR = ridge regression; RPR = random parameter regression; QLR = quadratic linear regression; PLSR = partial least square regression; SVM = support vector machine; SM = sigmoid model; NA = not available; LTPP = Long term pavement performance			

ANN has also been implemented in probabilistic performance modeling. Abdelaziz et al. (2020) used ANN models to estimate the probability that a given level of roughness will occur in the future in bituminous pavements. The models are based on historical pavement condition data and specific data on structural, traffic, and climatic conditions. The output variables are binary; one if the pavement exists within a given roughness level and zero if the pavement is in any other distress level. The success of ANN is contrasted against traditional multiple regression techniques.

However, the binary nature of the response variable creates difficulty for regression analysis that gives the comparison little relevance.

Some advantages of using ANN in pavement performance prediction can be listed as follows (You et al. 2020):

- There is no requirement for the prior specification of the model form as in the regression techniques.
- Many explanatory input variables are dependent on each other, which will result in a combinatorial explosion in regression techniques; however, ANN can manage these complex interactions and nonlinear patterns within the data.
- It has been proven in previous studies (Ghasemi et al. 2018) that ANN can accommodate noisy field data by filtering out the noise and extracting useful information for pattern recognition.
- Once successfully trained, ANN is simple and fast to use.

The following limitations are also noteworthy (You et al. 2020):

- The ANN training process can be computationally intensive, and the use of traditional Backpropagation can cause premature convergence. Therefore, either repeated training with different seed values should be conducted, or other training methods should be implemented.
- Similar to regression equations, ANN performance predictions are only valid within the range of the data used for training, and extrapolation is not recommended.
- Unlike regression equations, the ANN pavement performance model is not amenable to
- the physical interpretation of cause-and-effect relationships. In other words, variability in input parameters would not directly explain any parts of the variability in the output.

2.4 Architecture of Artificial Neural Network

Artificial neural networks are a subcategory of mostly applied machine learning techniques while others include support vector machines (SVM), Bayesian networks, inductive logic programming, decision trees, and radial basis functions (RBF) among others. Computational intelligence techniques include machine learning techniques primarily used for getting definitive information from large sets of data. Machine learning techniques can be used for pattern recognition, classification, regression, and prediction (Marcelino et al. 2021). ANN, SVM, and Kohonen self-organizing networks are popular machine-learning techniques for classification problems (Yuan et al. 2010). Recurrent Neural networks, association rule learning, and logic programming are some of the methods used for prediction. For regression purposes which are more broadly function regression such as function approximation applications, radial basis function networks, support vector machines, Bayesian networks, feed-forward and cascade-forward neural networks, and decision trees are the most often used ML techniques.

Artificial neural networks are similar to biological neural networks and consist of a large number of simple processors with many interconnections. ANN models try to use organizational principles that are believed to be used in the human brain and it is necessary to face challenging problems such as pattern recognition, function approximation, prediction, control, etc. (Yegnanarayana 2009). According to (Da Silva et al. 2017), an ANN is similar to the brain in two ways: 1) the network acquires knowledge through a learning process, and 2) synaptic weights are used to store knowledge. Artificial neurons are the basic processing elements of neural networks.

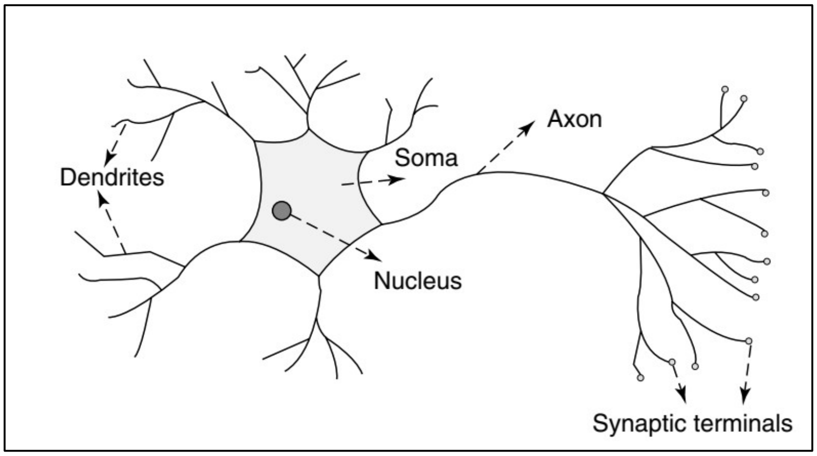


Figure 2.17: Mammalian neuron (Abraham 2005)

A mathematical framework exists that simplifies the complexity of neuronal behavior, whereby synapses' impact on input signals is expressed through connection weights, and a transfer function encapsulates the characteristic shown by neurons. The transformation of the weighted sum of input signals into a neuron impulse is achieved via the transfer function. An artificial neuron's ability to learn can be realized by adjusting the connection weights to optimize the transfer function. This mathematical abstraction has supplied a foundation for the development of neural networks that are capable of learning and performing a wide range of tasks. A typical artificial neuron and the modeling of a multilayered neural network are illustrated in Figure 2.18. Referring to Figure 2.18, the signal flow from inputs $X_1, \ldots, X_n$ is unidirectional, which is shown by arrows, as is a neuron's output signal flow *(O).* The neuron output signal *O* is given by the following relationship:

$$0 = f(net) = f\left(\sum_{j=1}^{n} \omega_j x_j\right) \tag{2.1}$$

where w_j is the weight vector, and the function *f*(net) is referred to as an activation (transfer) function. The variable net is defined as a scalar product of the weight and input vectors,

$$net = \omega^T x = \omega_1 x_1 + \ldots. + \omega_n x_n \tag{2.2}$$

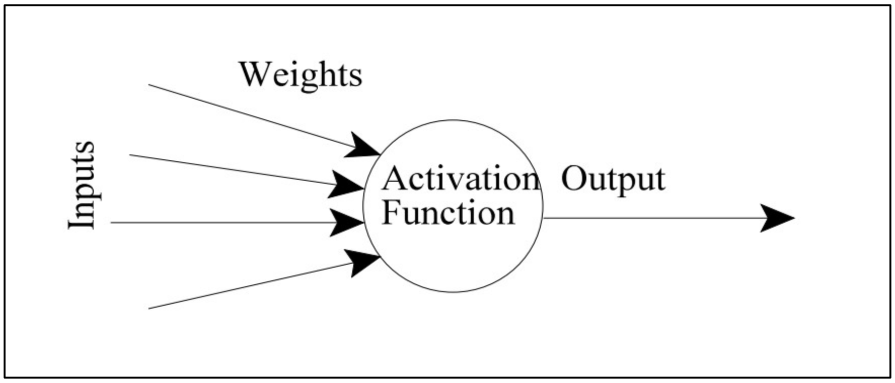
where T is the transpose of a matrix, and, in the simplest case, the output value O is computed as

$$0 = f(x) = \begin{cases} 1, & if\ \omega^T x \geq \theta \\ 0, & otherwise \end{cases} \tag{2.3}$$

where θ is called the threshold level; and this type of node is called a linear threshold unit.

The fundamental architecture of an artificial neural network (ANN) typically consists of three types of neuron layers, namely the input, hidden, and output layers. Data flows from the input layer to the output layer in a unidirectional feed-forward manner, where feedback connections do not exist between units. Recurrent networks incorporate feedback connections, making the dynamic properties of the network more essential than feed-forward networks. The activation values of the units can undergo a relaxation process, leading to a stable state with unchanging activations. In other applications, the changes in the output neuron's activation values can be significant, making the dynamical behavior the network's output (Ghosh et al. 2020).

To cater to the requirements and properties of diverse applications, there are various neural network architectures such as Elman networks, adaptive resonance theory maps, and competitive networks, among others. To ensure that an ANN produces the desired outputs for a given set of inputs, the network must be configured suitably. There are several methods for setting the connection weights, including setting them explicitly based on prior knowledge or training the network by feeding it teaching patterns and allowing it to change its weights (Maind and Wankar 2014). There are three distinct learning paradigms that neural networks can engage in, namely supervised learning, unsupervised learning, and reinforcement learning. In supervised learning, the network is presented with a set of input and desired output pairs, and the output layer includes one node per response. Following a forward pass through the network, discrepancies between the actual and desired output are evaluated. Based on a prescribed learning rule, these discrepancies are used to determine appropriate weight changes in the network (Abraham 2005).



(a) Artificial neuron

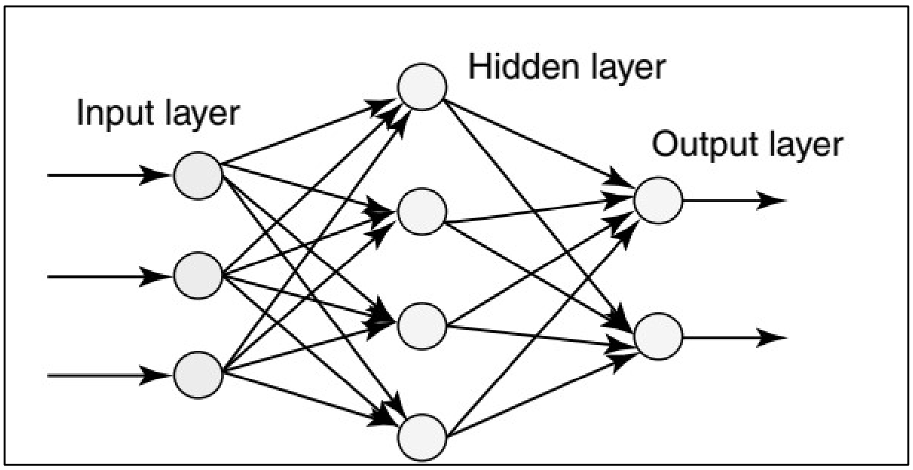


Figure 2.18: The architecture of (a) An Artificial Neuron and
(b) Multilayered Artificial Neural Networks (Maind and Wankar 2014)

The perceptron rule, delta rule, and backpropagation algorithm represent exemplary implementations of the supervised learning approach. Self-organization is a mechanism wherein an output unit is taught to respond to patterns that form clusters within the input data. This process involves the identification of statistically noteworthy features of the input population without a pre-established set of categories. The reinforcement learning approach revolves around developing the ability to map situations to actions that optimize a numerical reward signal. Unlike most machine learning paradigms, the learner is not informed of which actions to take but instead discovers which actions lead to the highest reward by trial and error. This approach is particularly challenging when actions influence both the immediate reward and the next situation. Crucial aspects of reinforcement learning include trial and error search and delayed reward (Van Gerven and Bohte 2017).

One of the most used networks is the backpropagation artificial neural network. They can be learned to copy from one data space to another using a representative set of examples (Duddu et al. 2020). There are two external layers (input and output) and one or more hidden layers in a backpropagation neural network. The network gets data in neurons from the input layer. The network's result is given by the network's output layer (Amin 2020). The hidden layers look at the interdependencies in the model and process the information. Before learning, it is necessary to define the input and output variables and collect the data that will be used to apply the backpropagation algorithm (Roberts and Attoh-Okine 1998). The backpropagation program uses supervised learning to give the network examples of inputs and outputs (Domitrović et al. 2018).

In the training phase of artificial neural networks, information is exchanged in both a forward and backward pass. The initial step involves assigning random weights to the input values, which are then propagated through the network to calculate the output in the forward pass. Next, the calculated output is compared to the desired output, and the difference between the two is computed in the backward pass. The weights are then adjusted to reduce the error. The scaling of the local error and the corresponding increase or decrease in weight are computed for each layer, starting with the layer closest to the output layer and moving back to the first hidden layer. This process is repeated numerous times until the error reaches a previously specified minimum value (Amin 2020). At this point, learning is stopped, and the weights are fixed for the testing phase based on the values obtained in the learning phase. During the testing phase, new input data are presented to the network, which has not been part of the learning process. The network's output is compared to the desired output, and the error is calculated. Based on the size of the calculated error, an assessment is made of the potential application of the artificial neural network (Domitrović et al. 2018). Figure 2.19 presents the architecture of backpropagation ANN.

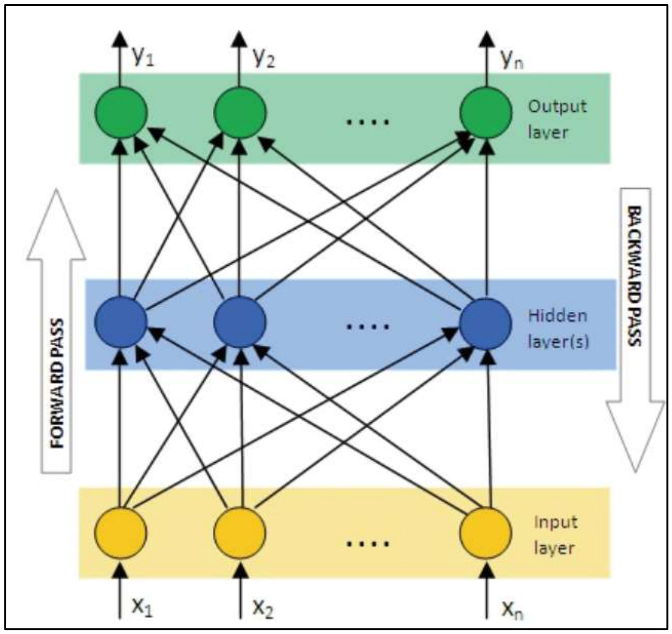


Figure 2.19: Architecture of backpropagation ANN (Domitrović et al. 2018)

In the feedforward back-propagation neural network, the processing follows three distinct steps. Firstly, the signals are transmitted in a forward direction through the neurons until the output layer is reached. Secondly, a comparison between the actual output and the expected output is made to determine the error value. Finally, the error is propagated backward through the network, and weights are adjusted accordingly until the error stops improving. However, the traditional backpropagation algorithm can lead to either overfitting or underfitting issues if it converges slowly. Overfitting occurs when the model captures noise and details from the training data, resulting in deficient performance. The model remembers all the information from the training data instead of learning from it, leading to a poor generalization of testing data. In contrast, underfitting arises when the model fails to perform well on the training data and does not generalize to the testing data (Amin 2020). Figure 2.20 illustrates overfitting and underfitting.

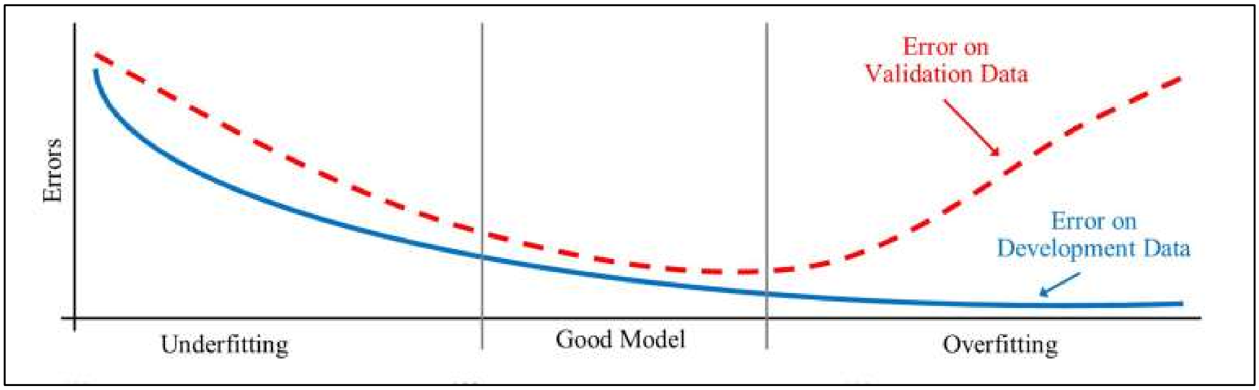


Figure 2.20: Illustration of Overfitting and Underfitting (Kargah-Ostadi et al. 2019)

The noise and overfitting problems in ANN models can be reduced with regularization techniques. In the ANN training process, regularization techniques are used to get a smaller error (Saini 2008). In the realm of regularization, there are three techniques available: the Levenberg-Marquardt method, the conjugate-gradient method, and Bayesian regularization. The Levenberg-Marquardt method is widely accepted and considered the standard in nonlinear optimization. It is known for its superior performance, especially for problems of medium size, and is sometimes referred to as the damped least-squares method. In contrast, the conjugate-gradient method is less commonly used in comparison to the Levenberg-Marquardt method (Roweis 1996). This method is an iterative process that locates the local minimum of a multivariate function that is expressed as the sum of squares of several non-linear relationships (Lourakis 2005). The conjugate gradient method generally produces faster convergence compared to the basic backpropagation algorithm while preserving the error minimization achieved in all the previous steps (Sharma and Venugopalan 2014). The purpose of the Bayesian regularization method is to minimize a combination of squared errors and weights. Bayesian regularization for neural networks is based on probabilistic interpretation to choose optimal sets of weights to minimize estimation error and efficiently avoid overfitting (Kayri 2016). The major advantage of using Bayesian regularization is that it does not require that the test dataset be separated from the training data set. This difference can be noticed when there is little data (Ticknor 2013).

For ANN, the user can get the outputs from the inputs through the many neurons that work in correspondence with weights and bias (Jain et al. 1996). In artificial neural networks, the output values of the neurons can take any value within the range of [−Inf, +Inf] in the absence of activation functions. However, without activation functions, the output values of each layer in the network will be a linear function of the inputs. This means that even in multi-layer networks, the outputs will be a linear combination of the inputs, which can limit the model's ability to learn complex patterns in the data (Golshani et al. 2018). To introduce non-linearity into the output of the neurons, an activation function is used. The activation function decides whether a neuron should be activated or not by calculating a weighted sum of the inputs and adding a bias term. The output of the activation function is then transformed to a value within a specific range, such as [0,1] or [−1,1], which can be interpreted as the probability that the neuron should be activated. The activation function is also referred to as the transfer function since it transfers the input signal to the output signal with non-linear transformations. The activation function can also be applied between two neural networks, which is called a transfer function network (Holmgren et al. 2019).

To ensure that the neural network produces accurate outputs, the weights and biases of the neurons need to be updated during the training process. This process is known as backpropagation, which involves calculating the gradient of the error concerning the weights and biases and using this information to adjust the weights and biases in the opposite direction of the gradient to minimize the error (LeCun et al. 1989). Activation functions make it possible since the gradients are supplied along with the error to update the weights and biases.

The activation function is a crucial component in ANNs as it introduces non-linearity to the model, allowing it to solve more complex problems. For an ANN model to learn and represent non-linear problems, the activation function transforms the input signal into an output signal that can be passed on to the next layer of neurons. There are several common types of activation functions, including the logistic, hyperbolic tangent, and linear functions. The logistic activation function generates output values that fall between 0 and 1 in an S-shaped curve, also known as the sigmoid curve. This function is commonly used in the output layer of a binary classification problem, where the model needs to predict the probability of an event occurring (Yamany et al. 2020).

The hyperbolic tangent activation function, on the other hand, generates output values between -1 and 1 in a similar S-shaped curve. This function is often used in hidden layers of the ANN model as it introduces non-linear transformations to the input, allowing the model to learn more complex relationships between the input and output. The linear activation function, as its name suggests, generates output values that are linearly proportional to the input. This activation function is commonly used in regression problems where the output is a continuous variable. However, the linear activation function is less frequently used in deep learning models as it can result in vanishing or exploding gradients, which can make it difficult for the model to converge during training (Maind and Wankar 2014).

In summary, the choice of activation function can have a significant impact on the performance of an ANN model. The logistic and hyperbolic tangent functions are commonly used due to their ability to introduce non-linearity, while the linear function is typically used for regression problems. The logistic activation function is defined by equation 2.4. The hyperbolic tangent activation function generates all the values between -1 and 1 in an S-shaped curve with

Equation 2.5. The sigmoid curve is remarkably similar to the curve of the hyperbolic tangent function, but it is a little sharper due to the range of outputs. The linear-activation function generates the same values as the input values. Equation 2.6 can identify it.

$$F(X) = \frac{1}{1+e^{-x}} \tag{2.4}$$

$$F(X) = \frac{2}{(1+e^{-2x})} - 1 \tag{2.5}$$

$$F(X) = X \tag{2.6}$$

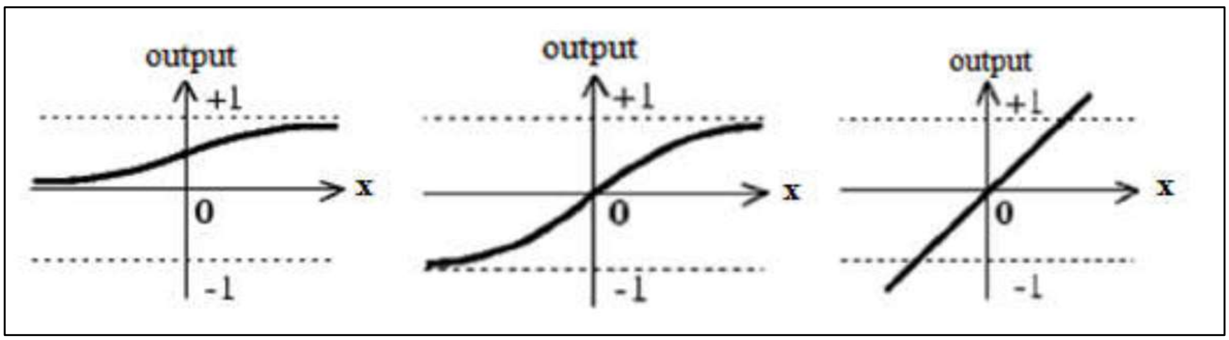


Figure 2.21: Logistic activation function (left), hyperbolic tangent activation function (center), and linear activation function (right). In the x-axis, the x is the incoming value or hidden neuron input.

2.5 Weigh-In-Motion

According to the American Society for Testing and Materials (ASTM), Weigh-In-Motion is defined as the process of determining a moving vehicle's gross weight and the allocation of that weight that is transmitted by each wheel, axle, or a combination thereof, by measurement and analysis of dynamic vehicle tire forces. As per the American Society for Testing and Materials

(ASTM) standards, a Weigh-In-Motion (WIM) system refers to an assemblage of sensors and ancillary equipment designed to measure the motion of a vehicle and the dynamic forces exerted by its tires at fixed locations over time. The WIM system can estimate critical vehicle parameters, such as tire loads, speed, axle spacing, and vehicle class based on axle configuration, and it stores, processes, and displays this information. The system serves to provide accurate and reliable data on vehicle weight and other vital statistics while minimizing the impact of weighing operations on the flow of traffic. In summary, ASTM defines a WIM system as a sophisticated infrastructure that utilizes state-of-the-art sensing technology to gather and process crucial data related to vehicle dynamics, thereby enhancing transportation infrastructure management and maintenance (Katz and Rakha 2002).

2.5.1 Weigh-In-Motion System

A WIM system typically consists of different sensors inserted in the pavement surface to identify, weigh, and classify vehicles (Qin et al. 2018). Such a system also incorporates software and electronics installed to regulate the WIM system sensors and collect, analyze, and save the sensor measurements. Communication hardware is also used to transmit vehicle measurements offsite. The electronics and communications devices are situated in a roadside cabinet adjacent to the WIM site. The entire system is powered by either a direct AC power connection or by batteries commonly charged by a solar panel array. WIM utilizes weight sensors and they are the most fundamental and critical component of the system (Yannis and Antoniou 2005). Weight sensors directly measure the force applied by the vehicles passing over the sensors. Piezoelectric, bending plate, and load cell sensors are the most prevailing sensors for use in this weight-measuring purpose. A piezoelectric WIM system consists of at least one sensor and two inductive loops embedded in a road cut or portable (Burnos et al. 2007). When a mechanical force is applied to a piezoelectric sensor, it generates a voltage that is proportional to the force or weight of the vehicle. As a vehicle passes over the piezoelectric sensor, the system records the electrical charge generated by the sensor and calculates the dynamic load. Piezoceramic sensors, piezopolymer sensors, and piezo quartz sensors are the three major types of piezoelectric sensors available for WIM applications (Jiang et al. 2009). These sensors involve a negligible temperature effect that facilitates protection against age or fatigue issues with accuracy and cost within the load cell range. This system is not suitable for portable WIM applications and is comparatively more expensive than other piezoceramic technologies (Refai 2013). Figure 2.22 represents a typical piezoelectric WIM system layout (Zhang et al. 2007).

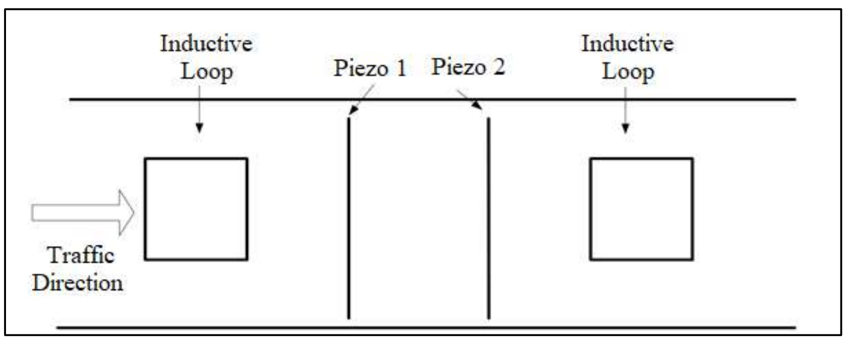


Figure 2.22: A typical piezoelectric WIM system layout

A typical bending plate WIM system consists of two steel platforms for each wheel path of the traffic lane, equipped with two inductive loops (Gaspareto and Gomes 2019). The function of the inductive loop is the same as that of the piezoelectric sensors. Bending plate scales can be portable or installed permanently with excavation into the road structure. When a vehicle passes over the bending plates, the strain gauge on each plate measures the amount of strain, and the WIM system measures the dynamic load that causes it. This sensor is designed for traffic load data collection and weight assessment. Also, it has high accuracy (more than piezoelectric systems) and low cost (lower than load cell systems). It involves minimal maintenance with required refurbishing after four to five years. It is less accurate than load cells and more expressive than piezoceramic sensors (Refai 2013). Figure 2.23 illustrates a typical bending plate WIM system layout (Zhang et al. 2007).

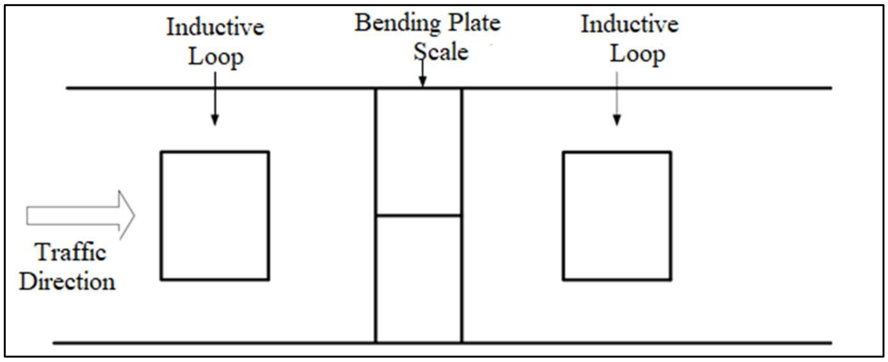


Figure 2.23: A typical bending plate WIM system layout

An ordinary load cell WIM system consists of a single load cell that includes two in-line scales, at least one inductive loop, and one axle sensor. Like the bending plate, the load cell is in the travel lane perpendicular to the travel direction. The purpose of the inductive loop situated upstream of the load cell is to distinguish approaching vehicles and notify the system (Beshears et al. 1998). Load cell WIM employs a single load cell with two scales to identify and weigh the right and left sides of an axle together. As the cell is subject to load, the wire under the strain gauge is compressed slightly and modified. The change in the wire results in a resistance difference from the present. Then, the system measures the variance in the present and determines the weight calculated scale, and then sums them to obtain the axle weight (Cheng et al. 2007). It is the most accurate sensor and can be implemented for traffic load data collection and weight assessment. But it is also the most expensive sensor with the highest maintenance cost and involves restoration after five years of deployment (Refai 2013). Figure 2.24 presents a typical load cell-based WIM system layout (Zhang et al. 2007).

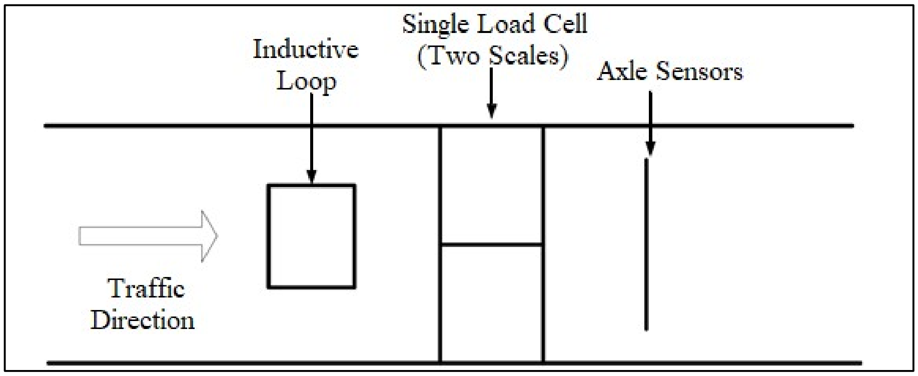


Figure 2.24: A typical load cell-based WIM system layout

The sensor placement affects the performance of the system. Studies indicated that it is feasible to measure the axle weight with embedded strain sensors (Zhang et al. 2008). Optimization of WIM systems based on embedded sensors was performed and the optimization framework established the relationship between the sensor signal and the measurement dynamic range (Xue et al. 2016). Research demonstrated that the depth of sensors embedded, and the road material should be analyzed to increase the capability of dynamic response sensing in the design of a WIM system. The sensors placed in the middle layer of the asphalt concrete layer have better sensing ability for the dynamic response of the pavement, which fits a Gaussian distribution centered at the wheel position (Qin et al. 2018). Most WIM systems are also capable of gathering data applicable to vehicle classification (Sun and Ban 2013). Vehicle classification data are important for pavement and bridge design, and rehabilitation as well as for traffic analysis (Xue et al. 2012). Vehicle classification can be obtained from the data measurements reported by the weight sensors or by applying a combination of the measurements from the weight sensors and a dedicated axle detector also installed in the pavement.

2.5.2 WIM Data Collection Process

The process of collecting data in the Weigh-In-Motion (WIM) system is conducted during the dynamic state of a vehicle, which often leads to inaccuracies in the data. To mitigate the impact of these inaccuracies, it is crucial to establish reference values for calibration purposes and devise a technique for accurately measuring the WIM system's performance (De Wet 2010). The precision of data obtained through Weigh-In-Motion (WIM) technology is contingent upon its specific purpose, be it for enforcement, data collection, or a combination of both (Papagiannakis et al. 2008). WIM calibration pertains to the evaluation of the computed weight, axle spacing, speed values, and overall vehicle length generated by the WIM system. This evaluation involves comparing the WIM system's output against known fixed weights and manually measured parameters such as axle spacing, vehicle length, and speed. Based on the results of this assessment, adjustments to the operating parameters of the WIM system are made to rectify any errors (Dahlin 1992). The primary purpose of implementing the initial calibration in WIM systems is to ensure that the accuracy level of the system aligns with the contract specifications post-installation of the site (Prozzi and Hong 2007; Ramachandran et al. 2011). Regular WIM calibration processes are conducted periodically to ensure that the data accuracy is consistent and maintained within the desired performance specifications (Baker 2019). After suitable calibration, the measurement bias or mean error in WIM measurements for all measured parameters should be reduced to the extent practically possible to achieve an accuracy level as close to zero as possible (Rys 2019).

In general, the calibration of a WIM system is conducted by a qualified technician in accordance with the manufacturer's specifications and guidelines. The primary objective of calibration is to establish parameters that will be utilized in subsequent calculations of the WIM system to establish a correlation between the recorded vehicle speed and tire force signals and the corresponding tire load, axle spacing, and wheelbase values for the stationary vehicle. Calibration is crucial for reducing the impact of various factors such as speed, temperature, truck type (if multiple test trucks are employed), and environmental changes in the supporting pavement structure on the WIM system's accuracy in assessing each measured lane (Hashemi Vaziri et al. 2013). The recommended calibration methods are the ASTM E1318-09 method and LTPP (Long-Term Pavement Performance) method described in LTPP Field Operations Guide for SPS WIM Sites (Haider and Masud 2020). In both techniques, comparable procedures are employed, albeit with slightly different criteria for assessing the calibration outcomes. Once WIM is initially installed and calibrated, it may experience measurement drift in weight and axle detection. Re-calibration is necessary, and there are two primary approaches: (a) on-site calibration, which involves running trucks of known weight over WIM scales, and (b) auto-calibration methods, which involve comparisons to assumed reference weights (Gupta et al. 2018). Auto-calibration can be more cost and time effective than on-site calibration (Durandal and Zhang 2019). The calibration process is uniform for all WIM sensor types, but sensors with lower precision, such as piezo-polymer sensors, may require more truck runs to establish a definitive error measure to determine new equipment compensation factors. Less precise sensors may also necessitate more frequent calibrations to account for seasonal temperature changes. Additionally, special calibration measures may be necessary for piezo-polymer sensor sites experiencing rapid temperature fluctuations between day and night (Selezneva and Wolf 2017).

In the process of collecting WIM data, the collected data is sampled and then converted into a suitable format for subsequent analysis. A computer program is used to analyze the converted data, which involves the calculation of the load spectra of the different four-axle groups, such as steering, single, tandem, and tridem, for distinct types of trucks. Furthermore, information regarding the time, location, and traffic volume distribution can be extracted from this data. By analyzing this data, trends in truck traffic growth, differences in side-wheel loads, and distributions of truck speeds can also be evaluated (Lu et al. 2002). BullPiezo, TrafLoad, Prep-ME, DARWin-ME, and LTPP PLUG have commonly known software to analyze the data (Li et al. 2018).

In the WIM system, there are three distinct software packages, namely, the on-site software, communication software, and in-house software. The on-site software is responsible for analyzing the signals generated by the WIM scale and creating on-site files containing information such as site identification, lane number, vehicle speed and classification, time and date of passage, vehicle sequence number, the weight of all axles or axle groups, ESAL value, code for invalid measurement, optional graphic confirmation, and other relevant details. The communication software offers a range of options for adjusting the on-site software setup, including calibration factors from the in-house computer. The in-house software is responsible for generating hard-copy reports and ASCII files. This software provides the means to develop reports on the collected raw vehicle record files. The communication and in-house software offers features such as real-time vehicle viewing selectable by lane, error report generation and viewing including time down, auto-calibration, system access, improperly completed records, transfer of selected raw data files, report generation from the site system to the office host computer, and other related functions (McCall and Vodrazka Jr 1997).

WIM system implementation can considerably result in a higher quality of traffic data collection that leads to better decision-making and design of transportation systems (Zhang et al. 2007). The benefits of a WIM system implementation include:

- Present a traffic database that can help to analyze the requirements of a local or state's weight enforcement program (Chan et al. 2005),

- Present better-quality traffic data for research programs, such as the LTPP program (Walker and Cebon 2012),
- Enhance the productivity of prescreening overweight/illegal trucks and therefore reduce travel time and lessen delay cost (Zhao and Tabatabai 2012),
- Enhance data utilization for pavement management systems (Wang et al. 2015),
- Enhance safety by effectively reducing overweight trucks on highways (Karim et al. 2014).

To understand the impact of the overweight truck, Figure 2.25 presents the impact of excessive axle loads on damaging roads based on data from Santero et al. (2005) in 10 locations. WIM can be remarkably effective for regulating excessive axle loads to avoid damage to the pavement.

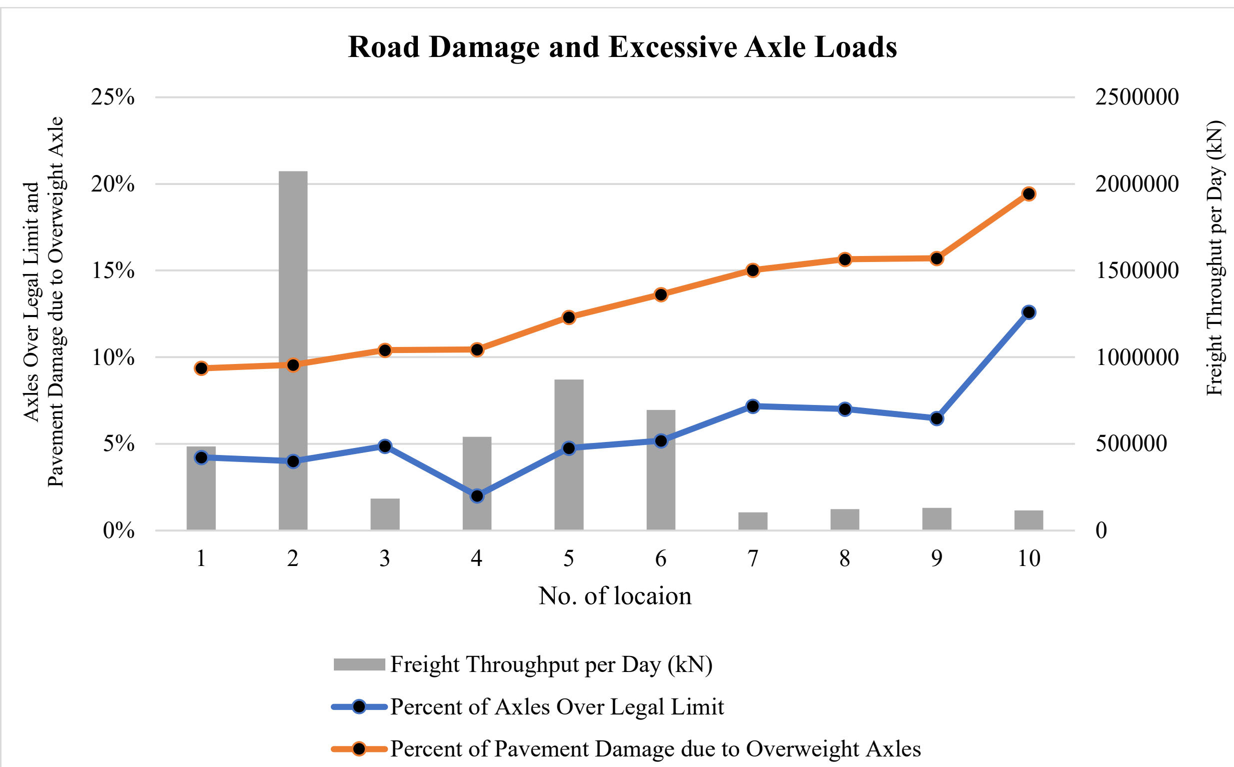


Figure 2.25: Road Damage for Excessive Axle Loads in 10 locations data (Santero et al. 2005)

2.6 Potential of WIM-Generated Traffic load data

Traffic data generated from WIM can be extremely significant for asset design and management, weight enforcement, and freight organization. Most DOTs collected WIM data with a wide variety of sensor types and used them in a variety of applications. Many agencies used WIM data to assist in pavement design, although most were not currently using a Pavement Mechanistic-Empirical Design application. WIM for bridge and asset management purposes was used much less often (Hazlett et al. 2020). In the Twin Cities metropolitan area in Minnesota, truck GPS data were validated with data from WIM sensors and loop detectors to develop reliable freight performance measures that present potential opportunities for freight planners and managers (Liao 2014). In addition, Minnesota DOT used WIM data for traffic forecasting, weight enforcement, and pavement design. Oregon DOT applied WIM data to determine the truck volume as well as axle weight and spacing for input into the state-of-the-art pavement design program (Elkins and Higgins 2008). Extensive WIM data collected on the French main road network were applied to better understand truck loading, overloads, and truck aggressiveness on infrastructure that may lead to policy optimization (Schmidt et al. 2016).

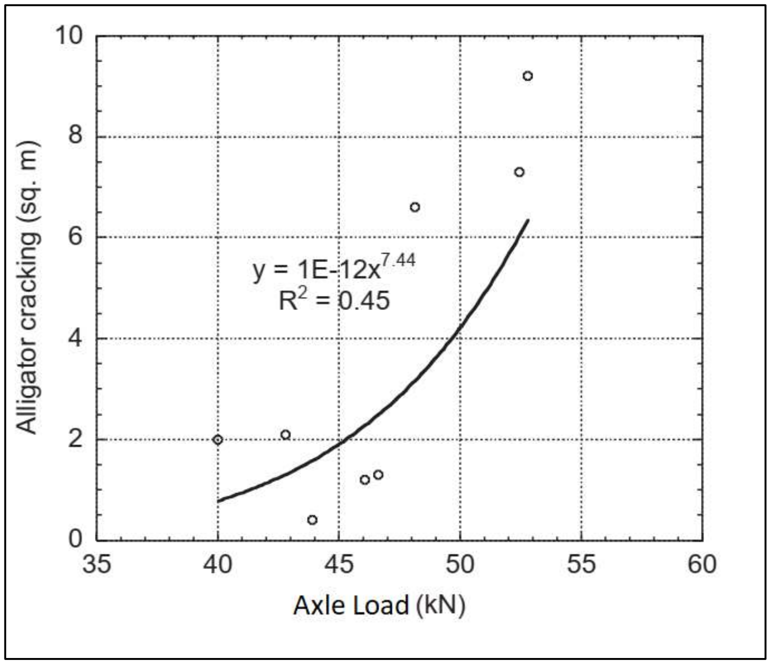
WIM has the potential to cut down the amount of extremely overweight and related costs for pavement resurfacing work and bridge infrastructure repairs (Nassif et al. 2018). Continuous reporting based on WIM data can serve to determine the quantity of pavement damage associated with the weight excess of the vehicles. Excess weight can lead to brake system defects and the truck could become troublesome to maneuver and control. The enhanced productivity of weight management will reduce the costs for weight enforcement resources and operations. Eventually, WIM data will reduce costs related to infrastructure damage (e.g. tunnel, bridge), road and tunnel closures, resurfacing works, and repairs of infrastructure (Haugen et al. 2016). Transportation engineers have access to precise valuable traffic data for transport planning, highway design, construction, and maintenance strategies (Bergan et al. 1998). WIM data in the future can be applied in management systems that may consist of new applications specially derived for traffic management centers such as heavy traffic flow information, real-time traffic data collection, tunnel traffic safety, dangerous good tracking, traffic/congestion monitoring, and diagnosis. Correct, reliable, and up-to-date information about vehicle weight, heavy traffic, and classification are convenient for the management of lane occupancy, traffic volume, and speed (Wang and Nihan 2004). Radar interferometry techniques with the integration of WIM data have been employed for advanced structural health monitoring.

The availability of enriched data collected from this developed WIM system makes it possible and promising to develop an accurate PPPM. Traffic is a key factor influencing the performance of flexible pavements. The new MEPDG uses each axle load distribution to describe traffic loads, while classification and count data are also required to represent load repetitions. There is a direct and rational approach to the analysis and design of pavement structures provided by these load distributions to estimate the effects of traffic on pavement response. Axle load distribution data are also used to calculate hourly and monthly traffic volumes, vehicle class distributions, and growth factors. MEPDG software commonly used for pavement analysis and future pavement performance prediction demonstrated the effects of different traffic level loading characteristics in terms of repetition on pavement performance. The traffic inputs for the MEPDG are presented in Table 2.4 (Jiang et al. 2010).

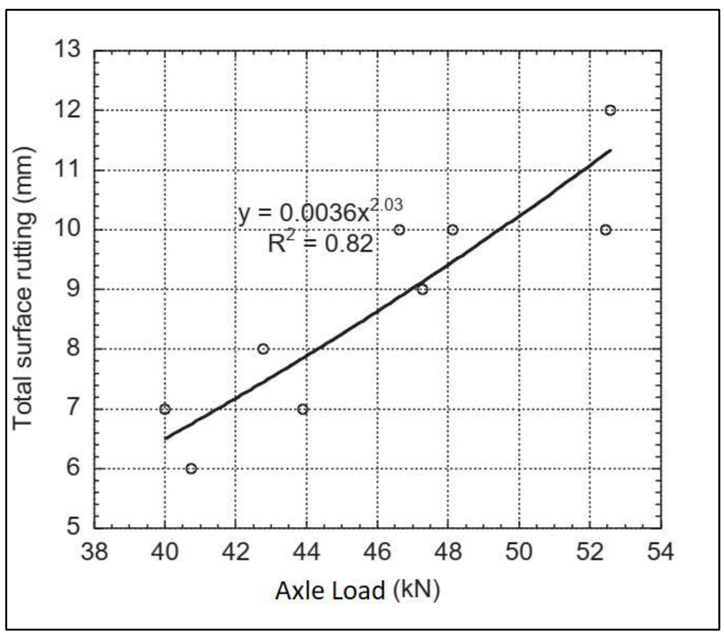
Table 2.3: MEPDG traffic loading input (Jiang et al. 2010)

Category	Design Input
All traffic	Average annual daily truck traffic
Truck traffic	Truck volume monthly adjustment factors
	Truck volume lane distribution factors
	Truck volume directional distribution factors
	Truck volume class distributions
	Traffic volume hourly distribution factors
Axle load distribution	Single-axle load distributions
	Tandem-axle load distributions
	Tridem-axle load distributions
	Quad-axle load distributions
	All-axle load distributions
Axle characteristics	Average axle weight (kips) and average axle spacing (inches) (Note: 1.0 kip = 1,000 pounds)
	Average number of axle types

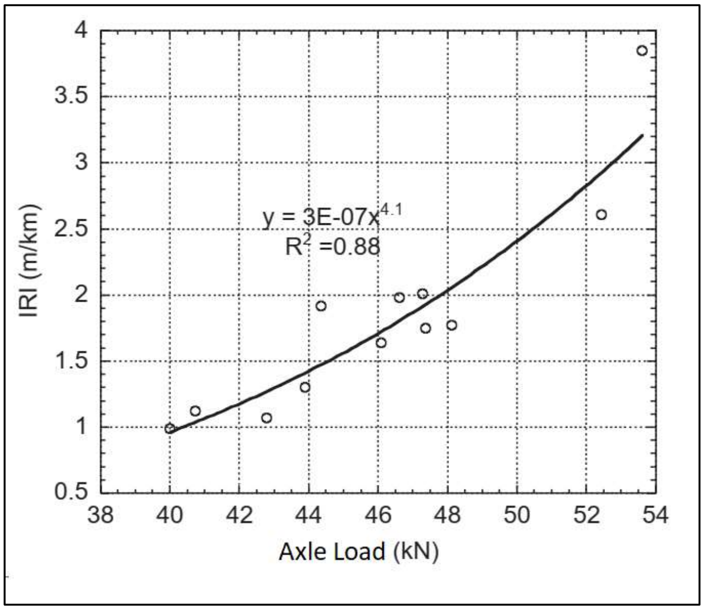
Haider and Harichandran (2009) demonstrated the relation between axle load (kN) with alligator cracking, total surface rutting, and roughness (IRI). Figure 2.26 presents the relations between specific pavement distress and axle loads. R^2 is the proportion of the variance of the dependent variables explained by the MEPDG model.



(a) Relation between alligator cracking and axle load



(b) Relation between total surface rutting and axle load



(c) Relation between roughness (IRI) and axle load

Figure 2.26: Relation between axle load and (a) alligator cracking, (b) total surface rutting, (c) roughness (IRI) (Haider and Harichandran 2009)

Analysis done by Salama et al. (2006) showed performance data from in-service pavements in the state of Michigan, the effect of heavy multiple-axle trucks on flexible pavement damage can be summarized as trucks with single and tandem axles appear to affect pavement cracking Distress Index (DI) more than those with multiple axles (tridem and higher). Conversely, heavier trucks with multiple axles tend to have more effect on rutting than those with single and tandem axles.

In recent years, several research studies have been conducted to evaluate the sensitivity of the MEPDG to these traffic inputs. It was determined that the vehicle class distribution has a significant influence on the design of pavement structures (El-Badawy et al. 2012; Papagiannakis et al. 2006; Romanoschi et al. 2011; Swan et al. 2008; Tran and Hall 2007). Therefore, there is a need to accurately represent the vehicle class distribution at the proposed pavement location to ensure proper pavement design. A study done by Abbas et al. (2014) found that the vehicle class distribution cannot be accurately estimated from functional classification because of large variations in the percentage of each vehicle class.

2.7 Long-Term Pavement Performance Program

The Long-Term Pavement Performance (LTPP) program was initiated as a part of the Strategic Highway Research Project (SHRP) in 1987 under the coordination of the Federal Highway Administration. The LTPP information management system is a pavement database documenting historical performance data for over 2,500 in-service and monitored test sections in North America. Diverse types of information are stored within the database in the form of seven modules: inventory, maintenance, monitoring, rehabilitation, material testing, traffic, and climate data. Collected data at the 2,500 sites is stored in an electronic warehouse called InfoPave. LTPP InfoPave includes creative tools for data viewing, identification, and selection that help users create their own personalized data sets, summary reports, queries, and much more. Over 650 research projects have already been conducted around the world analyzing the collected data to further pavement research. The LTPP program is the world's largest pavement performance monitoring study.

The LTPP program was envisioned as a comprehensive program to satisfy a wide range of pavement information needs. As sufficient data become available, analysis is conducted to provide better performance prediction models for use in pavement design and management; a better understanding of the effects of many variables on pavement performance; and new techniques for pavement design, construction, and rehabilitation. The strategy behind the LTPP program represents a significant shift in the traditional research approach. Figure 2.27 presents an overall view of the LTPP InfoPave website [https://infopave.fhwa.dot.gov/].

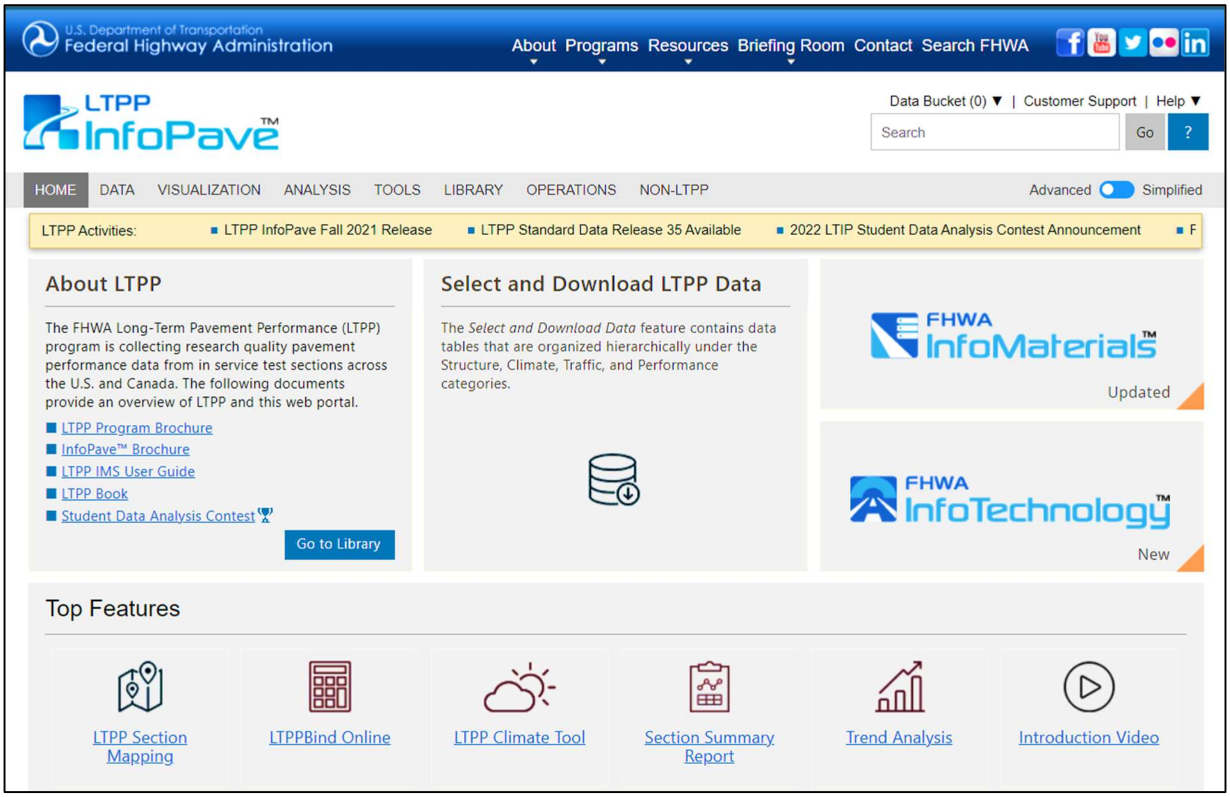


Figure 2.27: View of the LTPP InfoPave website [https://infopave.fhwa.dot.gov/]

Traditionally, pavement performance research was divided into specific topics of limited scope and duration, which started with data collection and ended with recommendations based on analysis of the collected data (Elkins et al. 2003). To overcome some of the challenges posed by the study of pavement behavior in short-term efforts, the LTPP program was established as a long-term national effort. Under the LTPP paradigm, data collection is conducted in advance of the development of many specific data analysis objectives. Since individuals not involved in data collection operations conduct many of the important data analyses, the LTPP program has invested in the development of publicly accessible databases and database use tools. Figure 2.28 illustrates the process of data collection for LTPP. Details traffic characteristics data collected through dynamic load response systems and WIM is the dynamic load response system.

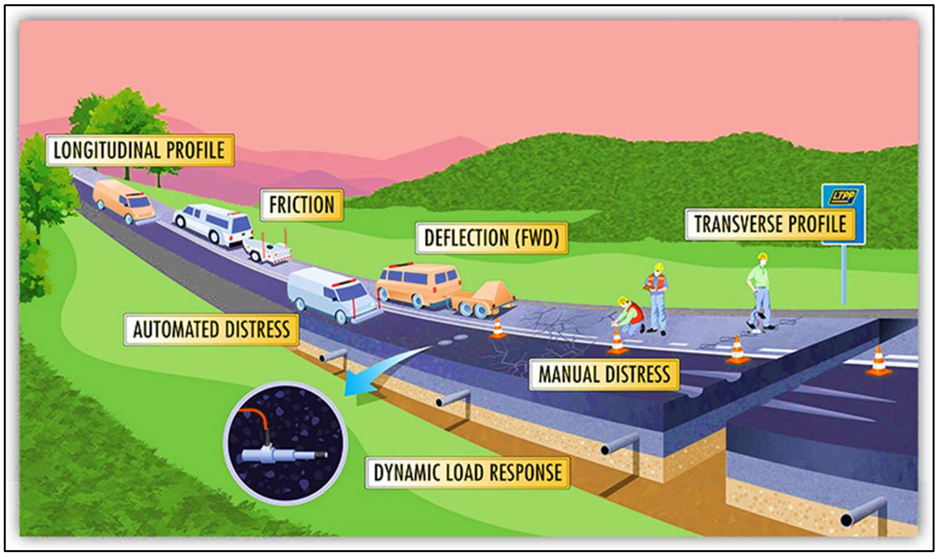


Figure 2.28: Process of the LTPP data collection [https://www.iengineering.com/ltpp-infopave-story/]

2.8 Data Processing

To build performance predictive models, it is important to understand the type of data under analysis. The main data types found in datasets are Numerical data, Categorical data, and Ordinal data. Numerical data represent data/information that is measurable, which can be divided into two subcategories: Discrete integer-based data (e.g., M&R actions, number of pavement sections); and Continuous decimal-based data (e.g., pavement structural capacity, traffic, pavement condition). Categorical data are qualitative data that are used to classify data by categories (e.g., crack initiation = true or false). Ordinal data that represent discretely and ordered data/information (e.g., rank position = 1st, 2nd, 3rd; rutting level = low, medium, high). After collecting the data, it is important to make an exploratory analysis and, if necessary, perform some data preparation. Their goals in data exploration are to fully understand the characteristics of each variable in data and discover any data quality issues. The most common data quality issues are Missing values and Outliers. Missing values where features have missing values, it is necessary to understand why they are missing. For example, road agencies usually do not make pavement inspections every year, but rather every two, three, or four years. Outliers are the values that lie far away from the central tendency and can represent valid or invalid data. There are valid outliers that are vastly different from the rest of the values and should not be removed from the analysis. Noise in the data resulting from invalid outliers must be removed.

Missing values are a widespread problem in many datasets. Missing values can occur for assorted reasons, including human error during data entry or the absence of certain data. Dealing with missing values is an essential step in data preprocessing, as models cannot manage missing data. Missing values can cause models to be biased, inaccurate, or unreliable. Therefore, it is necessary to address missing values before modeling. The first step in handling missing values is to identify the missing data in the dataset. One way to identify missing values is to use summary statistics, such as mean, median, mode, or count. These statistics can identify missing values, as they will be represented by a blank, null, or NaN value. Once the missing data has been identified, the next step is to determine the appropriate method to manage the missing values. One common method for handling missing values is to remove them. Removing missing values is a simple method but can be problematic as it reduces the size of the dataset and may result in the loss of valuable information. Removing missing values is most effective when the number of missing values is insignificant compared to the size of the dataset. In cases where a significant portion of the data is missing, removing missing values may not be a viable option (Dastres and Soori 2021).

Another method for handling missing values is to impute or fill in the missing values. Imputation methods involve estimating the missing values based on other values in the dataset. There are several imputation methods available, including mean imputation, median imputation, mode imputation, and regression imputation. Mean imputation involves replacing the missing values with the mean value of the column. Mean imputation assumes that the missing values are missing at random and are not correlated with other variables in the dataset. Mean imputation can be problematic if the distribution of the data is not normal, as the mean may not be representative of the data. Median imputation involves replacing the missing values with the median value of the column. Median imputation is like mean imputation, but it is less sensitive to outliers and works well for non-normal data distributions. Mode imputation involves replacing the missing values with the mode value of the column. Mode imputation is most effective for categorical data, as it replaces missing values with the most frequent value in the column (Choudhury and Pal 2019).

Regression imputation involves predicting the missing values using a regression model. Regression imputation works by creating a regression model based on the non-missing values in the dataset and then using this model to predict the missing values. Regression imputation is more complex than the other imputation methods but can produce accurate estimates of missing values. After selecting an appropriate imputation method, the next step is to apply it to the missing values in the dataset. Once the missing values have been imputed, it is important to verify the imputation process and ensure that it has not introduced errors or biases into the data. One way to verify the imputation process is to compare the imputed data to the original data and ensure that they are consistent (Lin et al. 2022).

Outliers in a dataset can cause significant problems in data analysis and modeling, and it is essential to preprocess them before proceeding with further analysis. Outliers are data points that are significantly different from the rest of the data points in a dataset. These data points may be due to measurement errors or represent rare occurrences, but they can skew the data distribution and lead to biased results (Zhao et al. 2019). Therefore, it is essential to identify and preprocess outliers before proceeding with further analysis.

The first step in preprocessing outliers is to identify them in the dataset. Outliers can be identified by plotting the data or by using statistical methods such as the Z-score or the Interquartile Range (IQR). The Z-score measures the number of standard deviations a data point is away from the mean. Any data point with a Z-score greater than a specified threshold (e.g., 3 or 4) is considered an outlier. The IQR is the range between the 25th and 75th percentile of the data. Any data point outside of the IQR multiplied by a specified threshold (e.g., 1.5 or 3) is considered an outlier. After identifying outliers in the dataset, the next step is to decide on the treatment strategy. There are several ways to manage outliers, including removing them, transforming the data, or imputing them. The decision on the treatment strategy should be based on the nature of the data, the research question, and the type of analysis to be performed (Nnamoko and Korkontzelos 2020).

The most straightforward treatment strategy for outliers is to remove them from the dataset. Removing outliers can be done by either deleting the data point or imputing the value with a missing value. However, removing outliers can result in a loss of information, especially if the number of outliers is significant (Nowak-Brzezińska and Łazarz 2021). Therefore, this treatment strategy should be used cautiously, and the impact of the removed data points on the results should be carefully considered. Another way to manage outliers is to transform the data to reduce the impact of outliers. Data transformation involves applying a mathematical function to the data to change the data distribution. The most used data transformations include log transformation, square root transformation, and Box-Cox transformation. Data transformation can reduce the impact of outliers by compressing the data range or by spreading the data range (Wang et al. 2019).

Imputing outliers involves replacing the outlier value with a more reasonable value. Imputing outliers can be done by using statistical methods such as mean imputation, median imputation, or regression imputation. Mean imputation involves replacing the outlier value with the mean value of the non-outlier data points. Median imputation involves replacing the outlier value with the median value of the non-outlier data points. Regression imputation involves using regression analysis to predict the missing value based on the other variables in the dataset (Wada 2020). After preprocessing outliers, it is important to assess the data for normality. Normality testing is used to determine whether the data follows a normal distribution. Normality testing can be done by using statistical tests such as the Kolmogorov-Smirnov test or the Shapiro-Wilk test. If the data is not normally distributed, additional data transformation may be required to normalize the data distribution (Mishra et al. 2019). After preprocessing outliers, it is important to verify the results. Verifying the results involves re-analyzing the data and comparing the results to the original analysis. The verification process should include comparing the statistical results, the model fit, and the validity of the research question. In conclusion, outliers can significantly affect the results of data analysis and modeling (Hasan et al. 2021).

The goal of the ANN modeling was to relate the physical causes of stresses in pavement structures and calibrate them with observed pavement performance. An ANN model has no critical rule to determine the optimal number of training data points. The data organization takes outliers out of the dataset otherwise the model's accuracy decreases when there are more outliers in the training data points (Sollazzo et al. 2017). There are significant effects on the quality of an ANN model from outliers. A neural network can receive enough learning if there are a lot of data points to tolerate a substantial number of outliers using the ANN feature of fault tolerance. The model can operate at the fault of its input data if it has fault tolerance. However, when the data points are not sufficient, the neural network won't be able to learn enough, and the model will become more sensitive to outliers (Arimie et al. 2020).

After organizing the data, the normalization of the data is initiated as it is one of the most common data processing techniques. It reduces the adverse impacts of outliers on model development in cases of outliers that can't be detected or reasoned (Vlahogianni and Karlaftis 2013). It's important increasing the error of prediction results, decrease the bias between independent variables, and increase the speed of convergence with the help of the normalization process for training data (Baghirli 2015). When the raw data are directly considered to train the ANN model, it will converge slowly and provide prediction results with large errors. Many normalization techniques scale the data in the same range of values for input and output data values. Among these normalization techniques, ANN models typically use Z-score normalizations and min-max methods. Z-score normalization uses the mean and standard deviation of each feature, and a series of learning data was used to normalize the features included in the input data. The mean and standard deviation are calculated for each feature. The equality used in the method is as below where x' indicates normalized data, x_i input variable, μ_i arithmetic mean of the input variable and σ_i standard deviation of the input variable.

$$x' = \frac{x_i - \mu_i}{\sigma_i} \tag{2.7}$$

The procedure sets the mean and standard deviation of each feature. As part of the procedure, the features in the data set are normalized. The mean and standard deviation are calculated for each feature over the training data, and it is used as a weight in the final system design. Preliminary processing within the artificial neural network structure is what this procedure is all about.

The Min-Max technique is used as an alternative to Z-score Method. This method rescales the features or the outputs in any range into a new range. Usually, the features are scaled between

0-1 or (-1)-1. The equality used in the method is as below where x_{mn} indicates the minimum value, x_m maximum value, x_i input value and x' normalized data:

$$x' = \frac{x_i - x_{mn}}{x_m - x_{mn}} \tag{2.8}$$

Each feature in the new range stays the same when the min-max method is applied. The method keeps all the data's properties. Notably, the min-max normalization technique may limit the normalization of forthcoming data values that exceed the predetermined range. Alternatively, the Z-score normalization method involves calculating the mean and standard deviation of the data values within a given dataset. As a result, the Z-score methodology is well-suited for circumstances where a dataset can be expanded to include additional data points, and where the minimum and maximum values of the data are unknown, and there may be outliers present within the data (Han et al. 2012). Additionally, the Z-score normalization method is beneficial in removing the necessity to assess the uniformity of scales among distinct variables. Such a factor is crucial in the selection of the covariance or correlation matrix in PCA. The normalization of data values by the Z-score approach enables the covariance of normalized data values to be equivalent to the correlation of unprocessed data values. Hence, it allows for the utilization of the correlation matrix in PCA without the need for scale matching.

2.9 Data Dimensionality Reduction using Principle Component Analysis (PCA)

Principle Component Analysis (PCA) is a widely used statistical technique in machine-correlated variables into a set of uncorrelated variables, which are called principle components (Liu and Yoon 2019). The goal of PCA is to reduce the number of dimensions while preserving the essential information in the data. PCA is executed on either the covariance or correlation matrix, with attention to the mathematical constructs that underlie these matrices which are very similar to each other (Kherif and Latypova 2020). Equation (2.9) and Equation (2.10) represent covariance and correlation matrix assuming a set of n observed values, the variables x_i and y_i denote the paired values of a random variable *(X, Y)*, where *i* ranges from 1 to n. The variables $\bar{X}$ and $\bar{Y}$ denote the respective means of the random variables X and Y, while s_X and s_Y represent their corresponding standard deviations. PCA is intrinsically influenced by the magnitude of variables in an unprocessed dataset, with a greater emphasis placed on variables that exhibit higher variances (Jolliffe and Cadima 2016).

$$Cov\ (X,Y) = \frac{\sum_{i=1}^{n}(x_i - \bar{X})(y_i - \bar{Y})}{n-1} \tag{2.9}$$

$$Corr\ (X,Y) = \frac{Cov\ (X,Y)}{s_x s_y} \tag{2.10}$$

The procedure of PCA involves several steps. Firstly, the data should be standardized to ensure that each variable has zero mean and unit variance. Standardization is crucial because variables with larger scales would dominate the analysis otherwise. Secondly, the covariance matrix of the standardized data is computed, which represents the relationships between variables. Thirdly, the eigenvectors and eigenvalues of the covariance matrix are calculated. The eigenvectors describe the directions in which the data varies the most, and the eigenvalues represent the variance explained by each eigenvector. Fourthly, the principle components are selected based on their corresponding eigenvalues. The greater the amount of variance explained by an eigenvector, the more important it is in representing the data. Finally, the data is projected onto the selected principle components, creating a new set of variables that are uncorrelated and ordered by importance (Omuya et al. 2021).

PCA can be applied to diverse types of data, such as numerical, categorical, and mixed data. It is a potent tool for reducing the dimensionality of high-dimensional datasets, enabling easier visualization and analysis of the data. Additionally, PCA can be employed for feature extraction, anomaly detection, and data compression. However, it is noteworthy that PCA presupposes that the data is linearly related and that the relationship between variables is Gaussian. Therefore, PCA may not be suitable for all types of data (Hasan and Abdulazeez 2021).

PCA facilitates the detection of data patterns by leveraging the correlation between various features. In essence, PCA endeavors to discover the directions of greatest variability in multi-dimensional data and transforms it into a fresh subspace comprising either equal or reduced dimensions than its antecedent space (Gewers et al. 2021). The orthogonal axes (also known as principle components) within the novel subspace can be comprehended as the directions of utmost variance while ensuring that the newly established feature axes remain perpendicular to each other, as portrayed in Figure 2.29.

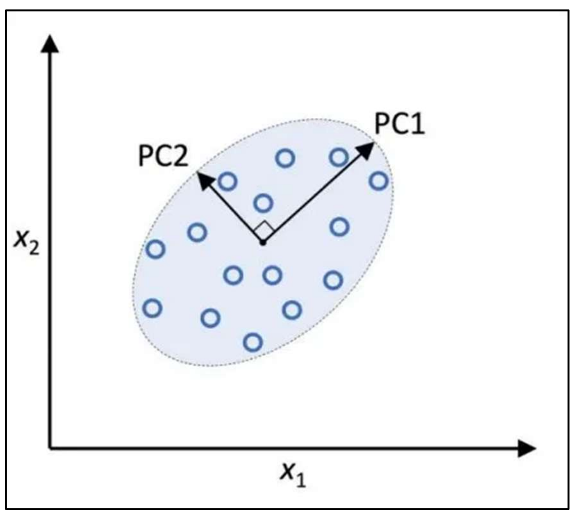


Figure 2.29: Principle Component Analysis (PCA)

In the antecedent illustration, x_1 and x_2 signify the initial feature axes, while PC1 and PC2 denote the principle components. When PCA is utilized for dimensionality reduction, a transformation matrix W with dimensions $d \times k$ is constructed. This matrix facilitates the mapping of a sample vector x onto a new feature subspace that has a reduced dimensionality of k, in contrast to the original d-dimensional feature space.

$$\boldsymbol{x} = [x_{1,} x_2, \dots, x_d], \quad x \in \mathbb{R}^d \tag{2.11}$$

$$\downarrow \boldsymbol{xW}, \boldsymbol{W} \in \mathbb{R}^{d*k}$$

$$\boldsymbol{z} = [z_{1,} z_2, \dots, z_k], \quad z \in \mathbb{R}^k \tag{2.12}$$

Following the transformation of the original d-dimensional data onto the k-dimensional subspace (usually $k \ll d$), the primary principle component will exhibit the maximal achievable variance, whereas all ensuing principle components will possess the greatest variance subject to the restriction that these components remain uncorrelated (orthogonal) to the other principle components (Velliangiri and Alagumuthukrishnan 2019). Furthermore, despite the presence of correlations among the input features, the resultant principle components will remain mutually orthogonal (uncorrelated).

It is worth noting that the PCA directions are sensitive to data scaling. Therefore, it is advisable to standardize the features before performing PCA, particularly if the features were measured on different scales, and aimed to assign equal significance to all the features. Before delving deeper into the PCA algorithm for dimensionality reduction, the approaches are summarized in a few straightforward steps (Kherif and Latypova 2020):

1. Standardize the original d-dimensional dataset.
2. Construct the covariance matrix.
3. Decompose the covariance matrix into its eigenvectors and eigenvalues.
4. Arrange the eigenvalues in descending order to rank the corresponding eigenvectors.

5. Select k eigenvectors that correspond to the k largest eigenvalues, where k represents the dimensionality of the new feature subspace ($k \leqslant d$).
6. Create a projection matrix W using the top k eigenvectors.
7. Transform the d-dimensional input dataset X using the projection matrix W to obtain the new k-dimensional feature subspace.

Upon concluding the mandatory data preprocessing, the subsequent step involves constructing the covariance matrix. The symmetric covariance matrix is d x d dimensional, where d represents the number of dimensions in the dataset. This matrix stores the pairwise covariances between the various features. To illustrate, the covariance between two features, x_j, and x_k, on a population level can be computed by applying the following formula:

$$\sigma_{jk} = \frac{1}{n}\sum_{i=1}^{n}\left(x_j^{(i)} - \mu_j\right)\left(x_k^{(i)} - \mu_k\right) \tag{2.13}$$

In this equation, μ_j and μ_k represent the sample means of features j and k, respectively. It is worth mentioning that if we standardized the dataset, the sample means would be zero. A positive covariance between two features indicates that the features increase or decrease in conjunction, while a negative covariance indicates that the features vary in opposite directions. For instance, the covariance matrix of three features can be expressed as follows.

$$\Sigma = \begin{bmatrix} \sigma_1^2 & \sigma_{12} & \sigma_{13} \\ \sigma_{21} & \sigma_2^2 & \sigma_{23} \\ \sigma_{31} & \sigma_{32} & \sigma_3^2 \end{bmatrix} \tag{2.14}$$

The principle components, also known as the directions of maximum variance, are represented by the eigenvectors of the covariance matrix, whereas their magnitude is determined by the corresponding eigenvalues. The preprocessed dataset contained 38 eigenvectors and eigenvalues from the 38 x 38-dimensional covariance matrix.

In the third step, the eigenpairs of the covariance matrix are obtained. An eigenvector v satisfies the given condition:

$$\sum v = \lambda v \tag{2.15}$$

Here, λ is a scalar: the eigenvalue. The goal is to reduce the dimensionality of the dataset by projecting it onto a new feature subspace, thereby compressing it. To achieve this, only the subset of eigenvectors, i.e., principle components are selected, that contain most of the information or variance. The magnitude of the eigenvectors is determined by the corresponding eigenvalues, so the eigenvalues must be sorted in descending order. Top k eigenvectors that correspond to the highest eigenvalues are preferred since they contain the most information.

Before collecting the k most informative eigenvectors, the variance-explained ratios of the eigenvalues are needed to be examined. The variance-explained ratio of an eigenvalue λj is the ratio of the eigenvalue λ_j to the total sum of eigenvalues. These ratios are plotted to understand the contribution of each eigenvalue to the total variance.

$$\text{The proportion of variance } (PoV)\ (PC = j) = \frac{\lambda_j}{\sum_{j=1}^{d} \lambda_j} \tag{2.16}$$

To reduce the dimensionality of a dataset, it is necessary to determine which principle components (PCs) should be retained or discarded. This can be accomplished by ranking the PCs in descending order of the proportion of variance (PoV) they account for and then computing the cumulative PoV for each successive PC. A scree plot can be used to visualize the PoVs of the PCs. Selecting a larger cumulative PoV will help to ensure that a significant amount of information about the data distribution is not lost. The top p PCs, where p is less than d, are selected based on this ranking. The appropriate number of PCs to consider can be determined by examining changes in the slope of the scree plot, selecting eigenvalues greater than or equal to one, or defining a minimum acceptable percentage of cumulative explained variance.

The decisive step of the PCA algorithm involves calculating the principle component scores for the selected principle components. This is done by computing a linear combination of the raw data values and the coefficient matrix. Specifically, given a matrix Z with n rows (representing data points) and d columns (representing variables), and a coefficient matrix V with p columns (representing the selected principle components), the matrix Y for the linear combination can be computed as follows:

$$Y = ZV = \begin{pmatrix} x'_{1,1} & \cdots & x'_{1,d} \\ \vdots & \ddots & \vdots \\ x'_{n,1} & \cdots & x'_{n,d} \end{pmatrix} \begin{pmatrix} e_{1,1} & \cdots & e_{1,p} \\ \vdots & \ddots & \vdots \\ e_{d,1} & \cdots & e_{d,p} \end{pmatrix} \tag{2.17}$$

Hence, the principle component scores for a given data point i and PC j can be obtained by computing the linear combination function. The resulting value represents the projection of the data point i onto the PC j axis, indicating how much the data point i contributes to the variation along the PC j direction (Jolliffe and Cadima 2016).

2.10 ANN Result Validation Process

Artificial neural networks (ANN) with backpropagation are widely used in various fields of research and industry for prediction and classification tasks. The accuracy of the ANN model in predicting the outcome variable is evaluated using performance measures such as R^2 and root mean squared error (RMSE).

R^2 is a statistical measure that represents the proportion of the variance in the dependent variable that is explained by the independent variables in the model. It ranges from 0 to 1, where 1 indicates that the model perfectly fits the data and 0 indicates that the model does not fit the data at all. RMSE is a measure of the difference between the predicted and actual values of the outcome variable. It is calculated by taking the square root of the average of the squared differences between the predicted and actual values. RMSE is an important measure of accuracy, as it gives an idea of how much the predicted values deviate from the actual values (Chicco et al. 2021). R^2 and RMSE values were determined using Equation 2.18 and Equation 2.19, respectively.

$$R^2 = 1 - \left[\frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2}\right] \tag{2.18}$$

$$RMSE = \sqrt{\frac{\sum_i^n (y_i - \hat{y}_i)^2}{n}} \tag{2.19}$$

where:

- y_i = actual value observation i;
- $\hat{y}$ = predicted value of observation i;
- $\bar{y}$ = average value of observation i; and
- n = the number of observations.

RMSE is a quality measurement for determining how close the estimated values are to real values. An RMSE value closer to zero indicates that the result is useful for prediction. In the ANN model, the values of training RMSE and Train R^2 are used to confirm the learning effect, and Test R^2 is used to check the predictive power; hence, Test R^2 is the most important indicator for the predictive model.

To analyze the performance of the ANN model with backpropagation, first step in analyzing the ANN model's performance is to prepare the data. The data should be checked for missing values, outliers, and inconsistencies. The data should also be split into training and testing sets to evaluate the performance of the model on unseen data. The next step is to train the ANN model on the training data. During training, the model adjusts its weights and biases to minimize the error between the predicted and actual values. The backpropagation algorithm is used to calculate the gradient of the error function for the model parameters, which is then used to update the weights and biases (Wright et al. 2022). Once the model is trained, it is evaluated on the testing data to assess its performance on unseen data. The predicted values are compared with the actual values using R^2 and RMSE. A higher R^2 value indicates a better fit between the predicted and actual values, while a lower RMSE value indicates a lower prediction error. If the model performance is not satisfactory, the model can be optimized by changing the number of neurons in the hidden layer, adjusting the learning rate, or changing the activation functions. The model can be retrained and evaluated to see if the performance has improved (Ashtiani et al. 2018).

In summary, the performance of an ANN model with backpropagation can be evaluated using R^2 and RMSE. These measures can be used to assess the accuracy of the model in predicting the outcome variable. The model can be optimized by adjusting various parameters to improve its performance on unseen data.

There is no perfect way to determine the optimal number of neurons in the hidden layer for the best performance of the ANN model. Instead, the number of neurons in the hidden layer is determined based on the rule of thumb, considering the complexity of a problem or the number of input and output variables. That is, the number of hidden neurons could be increased for more complex problems. Using the simplest way, several hidden neurons can be randomly determined between one and the number of input variables. Using another way, the ANN model utilizes the trial-and-error method, which assesses the number of hidden neurons from one hidden neuron to MAX_{Yn} hidden neurons, where MAX_{Yn} is the maximum number of hidden neurons. ANN model usually adds one additional neuron for bias to input and hidden layers to increase the flexibility of the model (De Veaux and Ungar 2012). If the result is satisfied, the final weights and biases are saved as the result. On the other hand, if the weights and biases are reassigned to the model, it is necessary to assess the model again until a satisfactory model is found. Reassigning the weights and biases means the model is re-trained with the same input combination and number of hidden neurons.

2.11 Hypothesis Testing for Understanding the Impact of WIM Data in PPPMs

A statistical technique is known as the t-test is often employed to assess whether two groups exhibit a significant difference. The t-test primarily examines the difference between the means of the two groups, to establish whether a such difference is statistically significant (Kim and Park 2019). The t-test employs a t-statistic to quantify this difference by taking into consideration the variance of the data points in each group. The t-statistic is then compared to a critical value based on the t-distribution probability distribution, accounting for the degrees of freedom and sample size.

The p-value is another crucial parameter utilized in the t-test. It indicates the probability of achieving the observed results solely by chance, on the assumption that the null hypothesis holds. The null hypothesis asserts that no significant difference exists between the two groups. The p-value is typically compared with a pre-selected significance level (such as 0.05 or 0.01), representing the level of certainty required to reject the null hypothesis (Okunev 2022).

If the computed p-value is smaller than the chosen significance level, the null hypothesis is rejected, implying a significant difference between the two groups. In contrast, if the p-value is larger than the selected significance level, the null hypothesis is not rejected, and one must conclude that the evidence does not support the assertion of a significant difference between the two groups.

To understand the impact of WIM (Weigh-in-Motion) data in PPPMs (Pavement Performance Prediction Models), the following hypothesis will be formulated:

Null hypothesis (H0): WIM data has no significant impact on PPPMs.

Alternative hypothesis (HA): WIM data has a significant impact on PPPMs.

To assess this hypothesis, the t-test statistic will be defined to compare the mean pavement performance prediction model output with and without WIM data. The significance level will be set at 0.05, which means the null hypothesis will be rejected if the p-value is less than 0.05. Collected data will be divided into two different data sets for developing pavement performance prediction models with and without WIM data. Collect data included the pavement condition, traffic volume, and other relevant variables that may affect the pavement performance. Using the collected data, a t-test statistic was generated for the two groups, with and without WIM data. Based on the t-test statistic and the degrees of freedom, the p-value will be calculated. If the p-value is less than 0.05, the null hypothesis will be rejected, and concluded that WIM data has a significant impact on PPPMs. If the p-value is greater than 0.05, it will fail the rejection of the null hypothesis, and it will be concluded that WIM data has a significant impact on PPPMs. This will imply that incorporating WIM data in PPPMs can improve their accuracy and reliability, leading to better pavement management and maintenance decisions.

Chapter 3. Methodology

3.1 Objective

The advent of Weigh-In-Motion (WIM) technology has enabled the generation of high-quality traffic load data, thereby presenting an opportunity to cost-effectively enhance current pavement maintenance practices. Developing mechanistic-empirical Pavement Performance Prediction Models (PPPMs) based on the Long-Term Pavement Performance (LTPP) data is a laudable goal, and to achieve it, this research aims to investigate the suitability of traffic characteristics information in efficiently utilizing them with other pertinent information for PPPM development. The proposed research seeks to explore how the application of traffic characteristics information can contribute to the development of PPPMs, which will provide valuable insights into pavement performance and improve pavement maintenance practices. By leveraging the WIM-generated traffic load data and integrating it with other crucial information, the research endeavors to provide a comprehensive framework for the development of effective PPPMs.

3.2 Methodology

To create effective artificial neural network (ANN) based pavement performance prediction models (PPPMs) for seven different performance indicators, 300 pavement sections with Weigh-In-Motion (WIM) data were carefully selected from various locations across the United States of America. The data was collected over 20 years, ranging from 2001 to 2020, and encompassed a range of factors such as pavement age, material properties, climatic properties, structural properties, and traffic-related characteristics. The primary dataset was then divided into two subsets - one with WIM-generated traffic data and another without. To ensure that the data was suitable for use, thorough data cleaning and normalization procedures were conducted using the Z-score normalization method. Each of the subsets was then further divided into two groups -

one containing 15 years of data for training models and the other containing 5 years of data for testing purposes. The seven distinct performance indicators being predicted by the PPPMs include IRI, longitudinal crack, transverse crack, fatigue crack, potholes, polished aggregate, and patch failure.

This research aimed to develop highly accurate PPPMs using ANN, which would be effective for predicting pavement performance across a range of factors. The inclusion of WIM-generated traffic data can facilitate the efficient use of traffic characteristics information, in combination with other essential information, to improve the accuracy of the models. The use of ANN-based models allowed for the creation of complex models capable of accurately predicting pavement performance across different performance indicators. To accomplish the objectives, an overview of the methodology is presented in Figure 3.1.

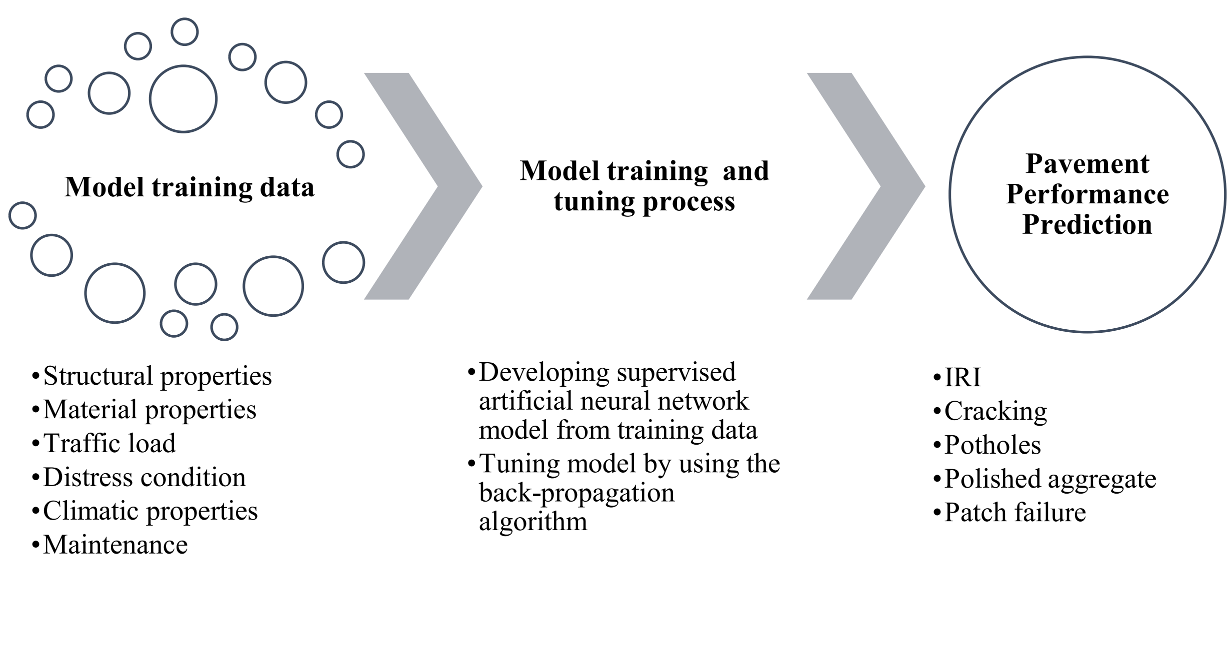


Figure 3.1: Overview of the process for Methodology

3.3 Data Collection

Data play one of the key roles in the development process of PPPMs. The input data used for the development of PPPMS is represented as predictors because these values are used to train the model that aims to predict. The application of machine learning algorithms required a large amount of data to learn the pattern in the data and predict based on this learning. Predictors' data collected from LTPP are categorized as age, material properties, climatic properties, and structural properties. In addition, pavement condition data was also collected from the LTPP database. The description of predictor data collected from LTPP for this research work is presented in Table 3.1.

Table 3.1: Collected LTPP data description

Type of Predictor	Predictor	Description
Age	Age (first construction)	Number of years passed after the first construction of the pavement
	Age (last construction)	Number of years passed after the last construction of the pavement
Material properties	Subgrade liquid limit	Liquid limit of the subgrade material
	Subgrade plastic limit	Plastic limit of the subgrade material
	Subgrade No. 200 passing	Percent of extracted aggregate for subgrade material passing the #200 sieve
	Subgrade unbound specific gravity	The unbound specific gravity of the subgrade material

	Subgrade moisture content	The moisture content of the subgrade material
	Subgrade confining pressure (kPa)	Chamber confining pressure results in the subgrade material
	Subgrade nominal maximum axial stress (kPa)	Nominal maximum axial stress of subgrade materials
	Subgrade resilient modulus average	Resilient modulus average for the subgrade material
	Subgrade resilient strain average	Resilient strain average of the subgrade material
	Base layer liquid limit	Liquid limit of the base layer material
	Base layer No. 200 passing	Percent of extracted aggregate for base layer material passing the #200 sieve.
	Base layer moisture content	The moisture content of the base layer material
	Base layer unbound specific gravity	The unbound specific gravity of the base layer material
	Original surface layer coefficient of thermal expansion (mm/mm/°C)	Coefficient of thermal expansion value
	Original surface layer compressive strength (psi)	Compressive strength of the original surface layer

Climatic properties	Total annual precipitation (mm)	Total precipitation for the year
	Annual total snowfall (mm)	Total snowfall for the year
	Mean annual temperature (°C)	Mean of the annual temperature for the year
	Annual freezing index (°C degree days)	Calculated freezing index for the year
	Annual freeze-thaw	The number of days in the period when the air temperature goes from less than 0°C to greater than zero °C; assumes minimum daily temperature occurs before the maximum daily temperature
	Average humidity	Average daily mean relative humidity for the year
Structural properties	Base layer thickness (in)	The thickness of the base layer
	Subbase layer thickness (in)	The thickness of the subbase layer
	Original surface layer thickness (in)	The thickness of the original surface layer
	Overlay Thickness (in)	Thickness of overlay

Traffic load data for the development of PPPMs are collected from the WIM system installed in three hundred locations for traffic characteristics data collection. WIM system possesses the capacity for high-quality traffic loading data and developing a machine learning-based model requires a large amount of historical data. WIM systems are installed in many locations in these three analyzed states. The details of the traffic characteristics data collected from WIM systems are presented in Table 3.2.

Table 3.2: Traffic characteristics data collected from WIM

Predictor	Description
Vehicle Class (4-13)	No vehicle for FHWA class 4 to class 13

Figure 3.2 presents FHWA vehicle class that was used in LTPP to define vehicle class data.

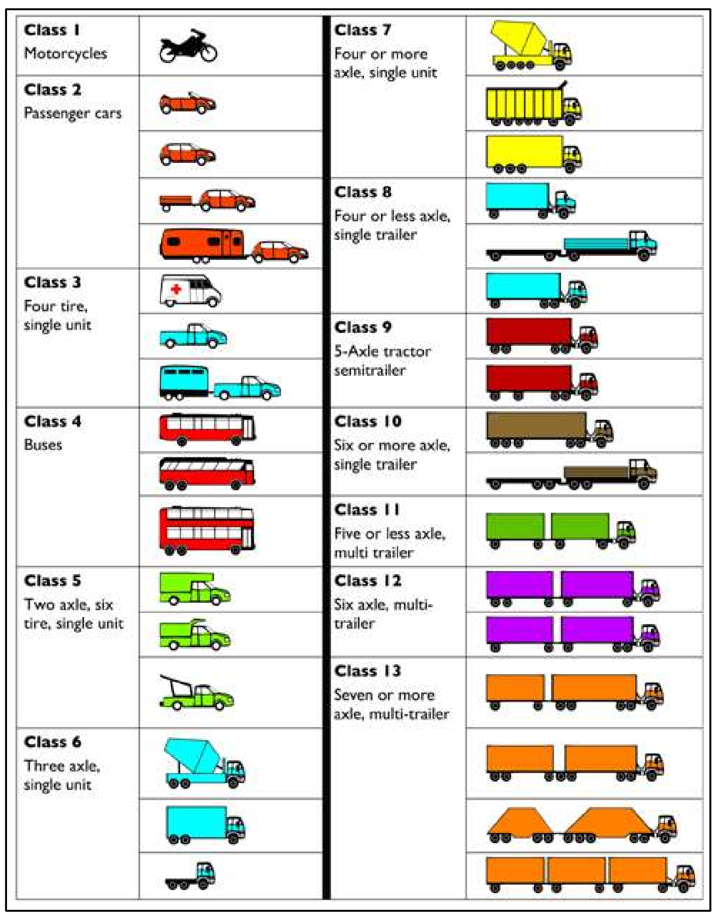


Figure 3.2: FHWA 13 Vehicle Category Classification (FHWA 2001)

3.4 Data Collection Locations

For this study, 300 locations/sections are selected throughout the United States of America. The data was collected for the year 2001 to 2020 from LTPP. The reason for selecting three hundred locations/sections with Weigh-In-Motion (WIM) systems throughout the United States of America is to ensure that the pavement performance prediction models (PPPMs) developed using artificial neural network (ANN) are generalizable and can be used in different regions of the country. By selecting pavement sections from various locations, the models can be trained to account for different pavement materials, climatic conditions, and traffic characteristics that may influence pavement performance. The use of WIM-generated high-quality traffic load data also ensures that the models can accurately predict pavement performance under different traffic volumes and axle loads.

The data collected for the year 2001 to 2020 from the Long-Term Pavement Performance (LTPP) program is to ensure that the models are developed using a wide range of pavement age and condition data. By including data from 20 years, the models can account for the effects of aging on pavement performance and can be used to predict performance for both newer and older pavements. Additionally, using data from the LTPP program ensures that the models are developed using a consistent methodology for collecting and analyzing pavement data, which enhances the reliability and consistency of the models. Two separate datasets were created where one dataset included WIM-generated vehicle types of information and the other dataset excluded WIM-generated data. WIM stored historical traffic load data and data is available for public use through the LTPP website. Each test section consists of a 152-meter (m) monitoring portion with a 15.2-m materials sampling section at each end. On test sections, a maintenance control zone, extending 152 m in front of and 76 m beyond the limits of the monitoring section, is commonly established around each test section. As an example of the data collection locations/sections for this study, the locations for West Virginia, Virginia, and Pennsylvania are presented in Table 3.3.

Table 3.3: Sample Data collection locations/sections

State	SHRP ID	Latitude	Longitude	County
West Virginia	4003	38.14941	-81.84423	Boone
	4004	38.02254	-81.35431	Fayette
	5007	39.28509	-80.42048	Harrison
	1640	38.28388	-81.76463	Kanawha
	7008	38.42770	-81.81807	Kanawha
Pennsylvania	0602	40.97496	-77.7913	Centre
	0603	40.98661	-77.811	Centre
	0604	40.98734	-77.813	Centre
	0608	40.98949	-77.8196	Centre
	0659	40.97982	-77.7945	Centre
	0660	40.99011	-77.8227	Centre
	0662	40.99129	-77.8286	Centre
	1597	41.97236	-77.2385	Tioga
	1598	40.2721	-77.0227	Cumberland
	1599	41.4337	-78.7129	Elk
	1605	41.00081	-76.8332	Northumberland
	1606	40.22314	-78.469	Bedford

	1608	39.99838	-78.5962	Bedford
	1613	39.99621	-75.3474	Delaware
	1614	40.82473	-78.0247	Centre
	1617	40.05779	-75.3301	Montgomery
	1618	39.77067	-78.9127	Somerset
	1623	41.24602	-76.9577	Lycoming
	1627	41.04159	-78.4128	Clearfield
Virginia	0113	36.62313	-79.3651	Pittsylvania
	0159	36.64686	-79.3647	Pittsylvania
	1417	38.60894	-77.7876	Fauquier
	1464	37.33266	-76.7086	York
	2021	36.73346	-80.8029	Carroll
	5009	37.50871	-77.2527	Henrico

The LTPP program is a study of the behavior of in-service pavement sections. These pavement sections have been constructed using highway agency specifications and contractors and subjected to real-life traffic loading. These in-service pavement sections are classified in the LTPP program as General Pavement Studies (GPS) and Specific Pavement Studies (SPS). GPS consists of a series of studies on 800 in-service pavement test sections throughout the United States and Canada. SPS are intensive studies of specific variables involving new construction, maintenance treatments, and rehabilitation activities. The specific 300 locations in the maps are shown in Figure 3.3.

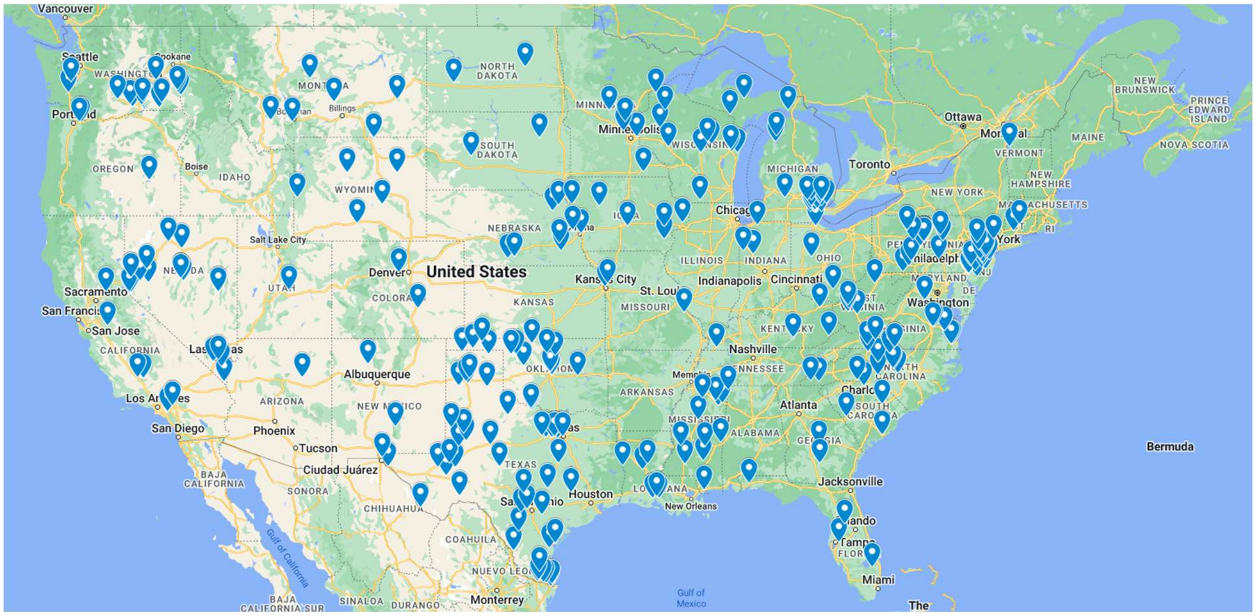


Figure 3.3: The locations of data collection

Artificial Neural Network (ANN) models have become a popular tool for pavement performance prediction due to their ability to model complex relationships and learn from large datasets. However, the accuracy and reliability of the ANN model depend on the quality and quantity of input data used to train the model. Time-series data is one type of data that is commonly used in pavement performance prediction models because it allows for the identification of trends, patterns, and relationships over time.

Time-series data refers to data that is collected over a period at regular intervals. Examples of time-series data in pavement performance prediction include pavement distress data, roughness measurements, and traffic volumes. By using time-series data, an ANN model can be trained to capture the dynamic nature of pavement performance and how it changes over time. This is particularly important for predicting the future performance of pavement, as it allows the model to consider the effects of aging, weather, and other environmental factors.

There are several reasons why time-series data is preferred over other types of data in pavement performance prediction models. Firstly, time-series data provides a more complete picture of the pavement's performance history. By collecting data over time, the model can identify patterns and trends that may not be apparent from a single snapshot in time. For example, if a pavement has experienced a sudden increase in roughness, the model can identify whether this is due to a sudden increase in traffic, changes in weather patterns, or other factors.

Secondly, time-series data allows for the identification of seasonal or cyclical patterns in pavement performance. For example, a pavement may experience more distress during certain times of the year due to changes in weather or traffic volumes. By capturing these patterns, the model can make more accurate predictions of future performance and identify appropriate maintenance and rehabilitation strategies.

Thirdly, time-series data can be used to evaluate the effectiveness of maintenance and rehabilitation strategies over time. By comparing performance data before and after a maintenance intervention, the model can determine the effectiveness of the intervention and refine its predictions for future performance.

Finally, time-series data is often readily available and easy to collect. Many pavement management systems routinely collect time-series data on pavement performance, making it an easily accessible source of input data for ANN models.

In summary, the use of time-series data in ANN-based pavement performance prediction models is important for capturing the dynamic nature of pavement performance and improving the accuracy and reliability of predictions. Time-series data allows for the identification of trends, patterns, and relationships over time, which can help to refine maintenance and rehabilitation strategies and improve the overall performance of pavement networks.

3.6 Data Processing

The execution of dependable data analysis and modeling necessitates the inclusion of data cleaning and normalization procedures. The fundamental objective of data cleaning is to eradicate anomalies from the unprocessed dataset. It is common for real-world data to be tainted by missing values and outliers. Such low-quality data can undermine the precision and trustworthiness of the data analysis and modeling process. Hence, it is imperative to undertake the critical task of identifying and eliminating outliers to achieve more accurate and dependable outcomes.

A rigorous data-cleaning process was implemented to ensure the accuracy and reliability of the data used in the analyses. The data cleaning process involved identifying and resolving various anomalies that could have a negative impact on data analyses, such as noise, missing values, and outliers. To initiate the data cleaning process, a thorough examination of the raw data was conducted to detect any inconsistencies, inaccuracies, or missing information. One of the most prevalent issues with raw data is noise, which refers to irrelevant or random data points that can distort the results of statistical analyses or machine learning models. To mitigate this issue, several techniques, including smoothing, filtering, and clustering was employed.

Smoothing is a technique used to eliminate high-frequency noise from the data while retaining the underlying trends and patterns. This method involved calculating the average of a window of adjacent data points or assigning more weight to recent data points using exponential smoothing. These methods helped reduce the impact of noise on the data, making it easier to detect underlying trends and patterns.

Filtering is another technique used to eliminate noise from the data. Filters are mathematical algorithms that are utilized to selectively remove or retain certain frequencies in data. A low-pass filter was used to remove high-frequency noise, whereas a high-pass filter was used to remove low-frequency noise. Other types of filters include band-pass filters and notch filters, which were used to remove specific frequency ranges or frequencies that interfere with the data.

Clustering is a technique used to identify and eliminate noise from the data. Clustering involves grouping similar data points based on their attributes, such as their distance from each other or their similarity in terms of their values. Data points that differ significantly from the other data points can be recognized as outliers and removed from the dataset.

Another prevalent issue with raw data is the presence of missing values, which can occur due to a variety of reasons such as data entry errors, system failures, or data corruption. To oversee missing values, various methods were employed, including deletion, imputation, or interpolation.

When dealing with missing values, the extent of missing values was evaluated in the dataset. As the missing values were insignificant, deletion was employed to eliminate all the rows or columns containing missing values and use the remaining data for analysis. However, if the missing values were significant, imputation could be employed to replace the missing values with an estimated value based on statistical analysis or modeling.

For numerical data, mean or median imputation involved replacing the missing values with the mean or median value of the corresponding attribute. For categorical data, mode imputation was used, which involved replacing the missing values with the mode value of the corresponding attribute. Another technique that was used was regression imputation, which involved estimating the missing values based on a regression model that is trained on the other attributes in the dataset.

Interpolation is another method that was used to estimate missing values. Interpolation involved estimating the missing values based on the values of neighboring data points. Several methods of interpolation are available, including linear interpolation, spline interpolation, and k-nearest neighbor interpolation. Linear interpolation involves estimating the missing values based on the linear relationship between the neighboring data points, while spline interpolation involves using a piecewise polynomial function to estimate the missing values. K-nearest neighbor interpolation involves estimating the missing values based on the values of the k-nearest neighbors in the dataset.

When analyzing data, it is essential to ensure that the data is accurate and reliable. One common issue that can negatively impact data analysis is the presence of outliers. Outliers are data points that are significantly different from the other data points in the dataset. They can be caused by a range of factors, such as data entry errors, measurement errors, or extreme values that do not fit the expected pattern of the data. Therefore, it is necessary to remove outliers from the dataset to ensure that the data analysis results are dependable.

Various techniques were implemented to remove them from the dataset. One of the methods used was the Z-score method, which is a statistical method that measures the distance between a data point and the mean of the dataset in terms of standard deviations. The Z-score of a data point is calculated as the difference between the data point and the mean divided by the standard deviation of the dataset. If the Z-score of a data point exceeded a certain threshold, which is typically 2.5 or 3, the data point was considered an outlier and removed from the dataset.

Another method used was the interquartile range (IQR) method, which is a non-parametric method that measures the spread of data. The IQR is calculated as the difference between the third quartile and the first quartile of the dataset. Any data point that falls below the first quartile minus 1.5 times the IQR or above the third quartile plus 1.5 times the IQR is considered an outlier and removed from the dataset.

In addition, a visual inspection can be applied to identify outliers that were not detected by statistical methods. The visual inspection involved examining the data plots and manually identifying data points that did not fit the expected pattern of the data. These data points were then removed from the dataset. When removing outliers, it is important to consider the context of the data and the purpose of the analysis.

In conclusion, removing outliers from the dataset is a crucial step in ensuring that the data analysis results are dependable. Employing various techniques such as the Z-score method, IQR method, and visual inspection helped identify and remove outliers from the dataset.

Overall, the data cleaning process is a critical step in ensuring the reliability and accuracy of the data used in my analyses. Through careful examination, the implementation of various techniques, and the utilization of appropriate methods for addressing noise, missing values, and outliers, can provide reliable insights that are useful for making informed decisions.

After cleaning the data, the data was normalized. The implementation of the Z-score normalization method is a widely used technique to standardize data sets, making them more comparable and easier to analyze. The steps involved in implementing the Z-score normalization method are as follows:

1. Collecting and organizing the data set into a matrix

The first step in implementing the Z-score normalization method was to collect and organize the data set into a matrix format. This matrix had rows corresponding to the observations and columns corresponding to the variables.

2. Calculating the mean of each variable

Next, the mean of each variable is calculated. This was done by summing up the values of each variable and dividing them by the total number of observations. The result of this step was a vector containing the mean value for each variable in the data set.

3. Calculating the standard deviation of each variable

The standard deviation of each variable was then calculated. This is done by taking the square root of the variance of each variable, where the variance is the sum of the squared differences between each observation and the mean of that variable, divided by the total number of observations minus one.

4. Subtracting the mean from each value in that variable and dividing by its standard deviation. After computing the mean and standard deviation of each variable, the mean was subtracted from each value in that variable and divided by its standard deviation. This process is called standardization or Z-score normalization. The result of this step was a new matrix where each variable has a mean of zero and a standard deviation of one.

In summary, the implementation of the Z-score normalization method involved collecting and organizing the data set into a matrix, calculating the mean and standard deviation of each variable, standardizing the data set by subtracting the mean from each value in that variable and dividing by its standard deviation, and using the standardized data set for further analysis.

3.7 Principle Component Analysis

Principle Component Analysis (PCA) is a widely used statistical technique that is utilized to explore and reduce the dimensionality of data sets (Liu and Yoon 2019). The process of applying the PCA test for principle components without using any equations involved the following steps:

1. Collecting and organizing the data set into a matrix

The initial step in applying for the PCA test involved collecting and organizing the data set into a matrix format. This matrix had rows corresponding to the observations and columns corresponding to the variables.

2. Calculating the mean of each variable

Next, the mean of each variable was calculated. This was done by summing up the values of each variable and dividing it by the total number of observations. The result of this step was a vector containing the mean value for each variable in the data set.

3. Subtracting the mean from each variable

After computing the mean of each variable, it was subtracted from each value in that variable. This process is called centering the data set. The result of this step was a new matrix where each variable has a mean of zero.

4. Calculating the covariance matrix of the mean-centered matrix

The covariance matrix of the mean-centered matrix was then computed. This matrix showed the covariance between each pair of variables in the data set. The diagonal elements of this matrix represented the variances of the variables.

5. Computing the eigenvalues and eigenvectors of the covariance matrix

The next step was to compute the eigenvalues and eigenvectors of the covariance matrix. Eigenvalues represented the amount of variance explained by each eigenvector, while eigenvectors represented the direction of maximum variance in the data set.

6. Sorting the eigenvectors in descending order of their corresponding eigenvalues

The eigenvectors were sorted in descending order of their corresponding eigenvalues. This step ensures that the eigenvectors with the highest variance are retained, while those with the lowest variance are discarded.

7. Selecting the eigenvectors corresponding to the highest eigenvalues

The eigenvectors corresponding to the highest eigenvalues were then selected. These eigenvectors represented the selected principle components that best explain the variance in the data set.

8. Obtaining the principle components by multiplying the mean-centered matrix by the selected eigenvectors

The principle components were obtained by multiplying the mean-centered matrix by the selected eigenvectors. The resulting matrix had specified selected columns, each corresponding to a principle component.

In summary, the process of applying the PCA test for selected principle components without using any equations involves collecting and organizing the data set into a matrix, calculating the mean of each variable, centering the data set by subtracting the mean from each variable, computing the covariance matrix of the mean-centered matrix, computing the eigenvalues and eigenvectors of the covariance matrix, sorting the eigenvectors in descending order of their corresponding eigenvalues, selecting the eigenvectors corresponding to the highest eigenvalues, obtaining the principle components by multiplying the mean-centered matrix by the selected eigenvectors and using the resulting matrix of principle components for further analysis.

The t-test method is a statistical hypothesis test used to determine whether there is a significant difference between the means of two groups of data. The following are the steps involved in implementing the t-test method:

1. Define the Null and Alternative Hypotheses: The first step in implementing the t-test method was to define the null and alternative hypotheses. The null hypothesis (H0) stated that there is no significant difference between the means of the two groups, while the alternative hypothesis (HA) states that there is a significant difference between the means of the two groups.
2. Collect the Data: The next step was to distribute collected data into two groups that are being compared. It was important to ensure that the data is collected in a random and unbiased manner to avoid any errors in the analysis.
3. Calculate the t-statistic: The t-statistic is a measure of the difference between the means of the two groups, normalized by the standard error of the difference. The formula for calculating the t-statistic is $t = (X_1 - X_2) / SE$, where X_1 and X_2 are the means of the two groups and SE is the standard error of the difference.
4. Determine the Degrees of Freedom: The degrees of freedom are the number of independent observations in the sample. The formula for calculating the degrees of freedom is $df = n_1 + n_2 - 2$, where n_1 and n_2 are the sample sizes of the two groups.
5. Calculate the p-value: The p-value is the probability of obtaining a t-statistic as extreme as the one observed, assuming that the null hypothesis is true. The p-value was calculated using a t-distribution table.

6. Interpret the Results: If the p-value is greater than the significance level (usually 0.05), then the null hypothesis was rejected, and it concluded that there is a significant difference between the means of the two groups. If the p-value was less than the significance level, then the null hypothesis could not be rejected, and it concluded that there is no significant difference between the means of the two groups.

Chapter 4. Development of PPPMs and Analysis

Predicting pavement performance is often considered to be an arduous task because many factors must be considered. Consequently, accurate pavement performance models that include more pavement data are needed as the basis for pavement maintenance and rehabilitation strategy selection. Many causes of pavement deterioration potentially vary from one road section to the next, which makes the modeling of pavement performance a complex process. Therefore, developing pavement performance prediction models requires both obtaining relevant data (e.g., pavement conditions and climate data) and identifying robust performance prediction approaches. In this research, artificial neural network (ANN) models were used to predict pavement performance.

4.1 Development of PPPMs

Traffic loading and environmental factors result in pavement distress. The ability to trace that distress over time allows researchers and agency decision-makers to develop performance prediction models. Predicting pavement performance requires historical data about pavement conditions, traffic loading, structural characteristics, and climate data. These data can be acquired from a single test road or from in-service pavements to obtain data for more practical prediction models. However, constructing and monitoring single test roads is expensive and unrealistic for small and local agencies. Developing accurate prediction models for pavements allows transportation agencies to effectively manage their highways in terms of budget allocation and scheduling maintenance and rehabilitation activities.

In this research, historical traffic loading collected from Weigh–In–Motion (WIM) systems and pavement structural, material, maintenance, condition, and historical climate data was obtained from Long Term Pavement Performance (LTPP) database to include all related variables in the prediction models. The results of this research will improve the understanding of pavement distress and evaluate the impact of traffic characteristics information on predicting pavement performance.

4.1.1 Data Collection

Data are the main building blocks in performance modeling, so obtaining excellent quality data is essential to getting accurate results. In this study, two kinds of data were obtained. Pavement condition, structural, material, and maintenance data were obtained from the LTPP, and traffic characteristics data were obtained from the WIM systems. Data used in this study are described in the following sections, followed by sections that discuss data integration and developing ANN pavement performance models.

The LTPP database includes information about the highway system, including section identification, construction history, pavement type, maintenance history, traffic loading, structure parameters, and pavement distress. The Long-Term Pavement Performance (LTPP) program was established to collect pavement performance data as one of the major research areas of the Strategic Highway Research Program (SHRP) and is currently managed by the Federal Highway Administration (FHWA). Pavement condition data from 2001 through the end of 2020 was used in this research. Each pavement section in the study had the same features (i.e., pavement type, maintenance history, traffic loading, subgrade stiffness, layer thicknesses, and pavement distresses). Rutting, roughness, longitudinal cracking, transverse cracking, fatigue cracking, and patching were pavement distress features of the pavement sections.

Table 4.1: Predictor data collected from LTPP

Type	Age	Material properties	Climatic properties	Structural properties	Traffic Characteristics
No. of predictors	2	16	6	4	10

4.1.2 Data Clean and Processing:

Data cleaning is an essential process in preparing data for analysis by identifying and resolving anomalies such as noise, missing values, and outliers that can negatively affect the analysis. The process began with a thorough examination of the raw data to detect inconsistencies, inaccuracies, or missing information. Noise is a prevalent issue that can distort the results of statistical analyses or machine learning models, and techniques such as smoothing, filtering, and clustering are employed to mitigate it. Smoothing involved eliminating high-frequency noise from the data while retaining underlying trends and patterns, and filtering selectively removes or retains certain frequencies in data. Clustering groups similar data points together to identify and eliminate noise from the data.

Another prevalent issue with raw data is the presence of missing values, which was managed through deletion, imputation, or interpolation. Deletion is used when the missing values are insignificant, whereas imputation is used to replace missing values with an estimated value based on statistical analysis or modeling. Mean or median imputation is used for numerical data, and mode imputation is used for categorical data. Regression imputation involved estimating missing values based on a regression model trained on other attributes in the dataset. Interpolation estimated missing values based on neighboring data points, using methods such as linear interpolation, spline interpolation, or k-nearest neighbor interpolation.

When conducting data analysis, accuracy and reliability are paramount. One potential hindrance to accurate analysis is the presence of outliers, which are data points that differ significantly from most of the dataset due to various factors such as errors in measurement or data entry. To ensure dependable analysis results, outliers were removed from the dataset. Different techniques were employed for this purpose, including the Z-score method, which calculates the difference between a data point and the dataset mean in terms of standard deviations. Any data point with a Z-score exceeding a certain threshold (usually 2.5 or 3) is considered an outlier and removed. The interquartile range (IQR) method, which measures data spread, is another option. The IQR was calculated as the difference between the first and third quartiles of the dataset. Any data point falling outside the range of 1.5 times the IQR below the first quartile or above the third quartile is identified and removed. Additionally, visual inspection was used to identify outliers not detected by statistical methods. It involved examining data plots and identifying data points that do not align with the expected pattern. It is important to consider the data context and analysis purpose when removing outliers. Overall, removing outliers was essential to ensure dependable data analysis results, and various techniques were utilized to achieve this goal.

In conclusion, data cleaning is a critical step in ensuring reliable and accurate analysis results. Employing appropriate techniques to address noise, missing values, and outliers provide trustworthy insights that are valuable for informed decision-making.

The Z-score normalization method is a commonly used technique to standardize data sets for easier analysis and comparability. To implement this method, the data set was organized into a matrix with observations as rows and variables as columns. The mean of each variable is then calculated by dividing the sum of all values in that variable by the total number of observations. The standard deviation of each variable is then determined using the variance, which is the sum of the squared differences between each observation and the variable mean, divided by the total number of observations minus one. The mean is subtracted from each value in the variable, and the result is divided by its standard deviation, resulting in a standardized data matrix where each variable has a mean of zero and a standard deviation of one. This standardized data can then be used for further analysis, such as classification, clustering, or visualization. The interpretation of standardized data depends on the specific data set and context.

4.1.3 PCA

The widely used statistical technique, Principle Component Analysis (PCA), is employed to explore and reduce the dimensionality of data sets. The process of applying the PCA test involved several steps. Firstly, the data set is collected and organized into a matrix format, with rows and columns representing observations and variables, respectively. In the dataset, there were 38 variables presented. So, a 38×38 matrix was generated from the variables.

Next, the mean of each variable is calculated by dividing the sum of its values by the total number of observations. The data set is then centered by subtracting the mean from each variable. The covariance matrix of the mean-centered matrix is then computed to show the covariance between each pair of variables. The eigenvalues and eigenvectors of the covariance matrix are then computed, with eigenvalues representing the variance explained by each eigenvector and eigenvectors representing the direction of maximum variance. These eigenvectors are then sorted in descending order of their corresponding eigenvalues, and the eigenvectors with the highest eigenvalues are selected as principle components. These selected components are obtained by multiplying the mean-centered matrix by the selected eigenvectors, resulting in a matrix with specified columns corresponding to principle components. This matrix of principle components can then be used for further analysis such as clustering, classification, or visualization. To understand the process, Figure 4.1 represents the Proportion of Variation (*PoV*) based on individual and cumulative explained variance.

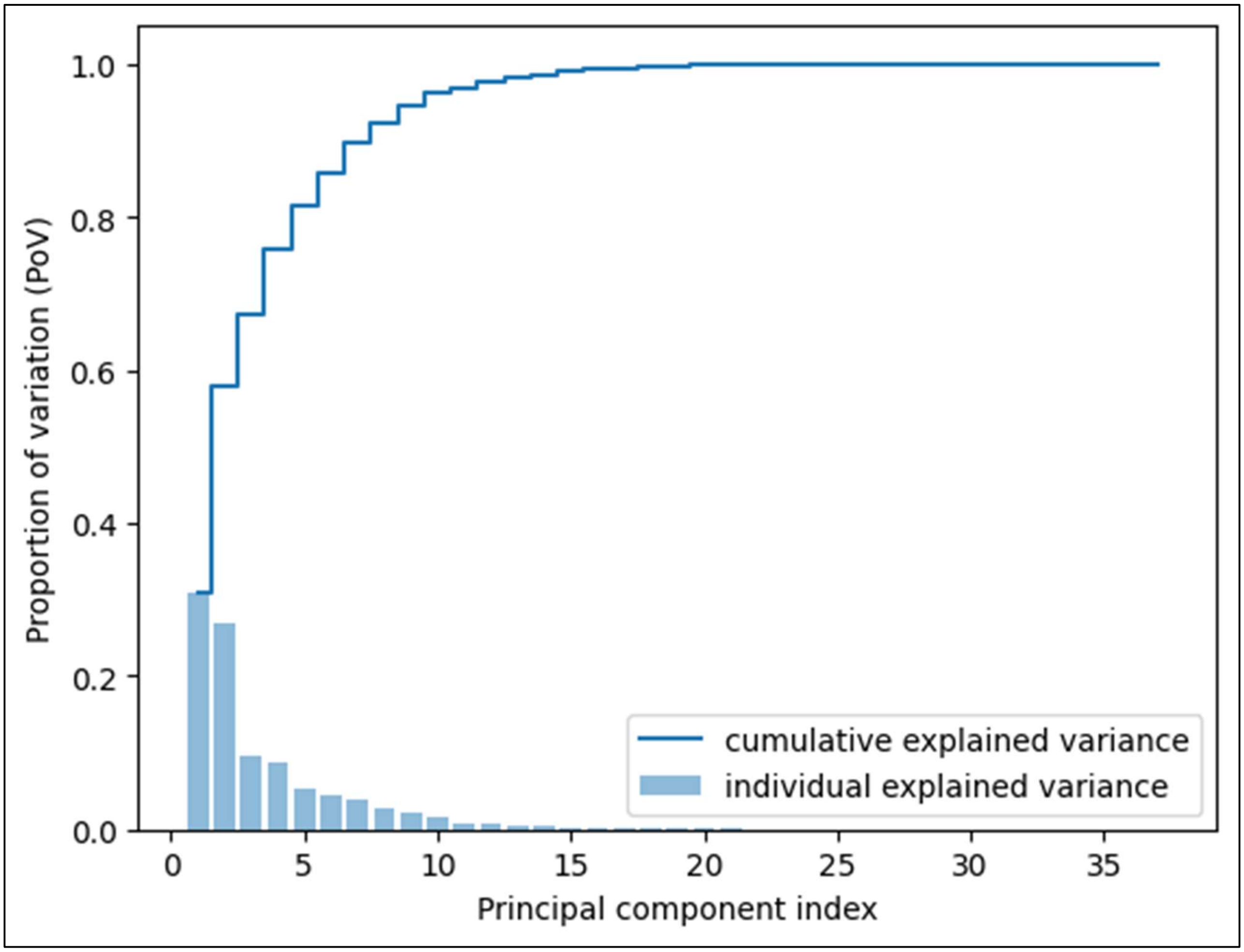


Figure 4.1: Scree plot for Proportion of variation

To incorporate the information of the original input variables in PCA, a cumulative *PoV* of at least 80% is typically considered necessary (Adler and Golany 2001). For this analysis, 12 components cumulatively had around 96% *PoV*. In the present investigation, the chosen approach involved utilizing the twelve Principal Components (PCs) that were identified by meeting the minimum threshold for cumulative Proportion of Variation *(PoV)*. This decision was made in consideration of the case study in question, to achieve an optimal representation of the underlying data structure. The minimum cumulative PoV criterion, which accounts for the proportion of total variance explained by the selected PCs, was employed as a means of identifying the most significant PCs that encapsulated most of the relevant information. Therefore, this study sought to apply a judicious selection of PCs to ensure that the resulting analysis was robust and reflective of the underlying patterns within the data. To compute the principal component scores in this study, a process was undertaken whereby the normalized values of 300 data points, representing locations with data from the year 2020, were used. These scores were computed based on the twelve Principal Components (PCs) that were identified as being significant in capturing the underlying variability in the data. The computation of these scores was conducted to summarize the information contained within the original dataset into a smaller, more manageable set of variables. By utilizing the normalized values, the data points were standardized to ensure that the scores were not influenced by differences in scale or units of measurement. Overall, this approach allowed for a more comprehensive understanding of the relationships between the 300 locations in the dataset and provided a useful framework for analyzing the patterns and trends within the data. Tables 4.2 and 4.3 represent the eigenvalues and eigenvectors for the highest 12 PCs.

Table 4.2: Eigenvalue, Variance percentage, and Cumulative Variance Percentage for the 12 dimensions

	Eigenvalue	Variance Percent	Cumulative Variance Percent
Dim. 1	1.5515	0.263679	0.263679
Dim. 2	1.2546	0.241149	0.504828
Dim. 3	1.1121	0.150701	0.655529
Dim. 4	0.8667	0.088637	0.744166
Dim. 5	0.6213	0.067914	0.81208
Dim. 6	0.5758	0.055463	0.867543
Dim. 7	0.4842	0.046209	0.913752
Dim. 8	0.3912	0.034805	0.948557
Dim. 9	0.3143	0.026298	0.974855
Dim. 10	0.2423	0.015922	0.990777
Dim. 11	0.2013	0.00564	0.996417
Dim. 12	0.1912	0.003583	1.000001

Table 4.3: Eigenvector for the highest PCs

	PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8	PC9	PC10	PC11	PC12
Base Layer Unbound SPEC GRAVITY	0.692344	0.690187	0.68803	0.685873	0.683716	0.681559	0.692344	0.690187	0.686643	0.686335	0.686027	0.685719
Subgrade CON PRESSURE	0.726671	0.50973	0.292789	0.075848	-0.14109	-0.35803	0.726671	0.50973	0.153327	0.122335	0.091344	0.060352
Subgrade NOM MAX AXIAL STRESS	0.524988	0.258102	-0.00878	-0.27567	-0.54256	-0.80944	0.524988	0.258102	-0.18035	-0.21848	-0.25661	-0.29473
Total Annual Precipitation	0.967514	0.057511	-0.85249	-0.7625	-0.6725	-0.2825	0.967514	0.057511	-0.06607	-0.0663	-0.06654	-0.06678
Mean Annual	0.066124	0.055933	0.045742	0.035551	0.02536	0.015169	0.066124	0.055933	0.039191	0.037735	0.036279	0.034823

Temperature												
Subbase layer/interlayer+ Sabgrade layer Thickness	-0.36897	-0.49264	-0.61632	-0.73999	-0.86367	-0.98734	-0.36897	-0.49264	-0.69582	-0.71349	-0.73116	-0.74883
Interlayer BSG	-0.64436	-0.85872	-0.07307	-0.28742	-0.00177	0.183877	-0.64436	-0.85872	-0.36444	-0.35696	-0.34949	-0.34202
AC Layer Below Surface (Binder Course) +Overlay Thickness	0.211365	0.237578	0.263791	0.290004	0.316217	0.34243	0.211365	0.237578	0.280642	0.284387	0.288132	0.291876
VEH CLASS 5 DIST	0.437696	0.522782	0.607868	0.692954	0.77804	0.863126	0.437696	0.522782	0.662566	0.674721	0.686876	0.699032

PERCENT												
VEH CLASS 6 DIST PERCENT	-0.75819	-0.9652	-0.17221	-0.37922	-0.08623	0.106767	-0.75819	-0.9652	-0.45886	-0.45033	-0.44181	-0.43329
VEH CLASS 8 DIST PERCENT	0.276847	0.069839	-0.13717	-0.34418	-0.55119	-0.75819	0.276847	0.069839	-0.27025	-0.29982	-0.32939	-0.35896
VEH CLASS 10 DIST PERCENT	0.804931	0.06208	-0.68077	-0.42362	-0.16647	-0.60932	0.804931	0.06208	-0.0726	-0.08468	-0.09675	-0.10883

In addition, the statistical properties of the selected input variables are presented in Table 4.4.

Table 4.4: Statistical properties of the selected input variables

	Unit	mean	std	min	max
Base Layer Unbound SPEC GRAVITY		2.623189	0.062136	2.526	2.762
Subgrade CON PRESSURE	(kPa)	29.74111	16.15447	13.8	103.4
Subgrade NOM MAX AXIAL STRESS	(kPa)	43.85667	25.99092	13.8	206.8
Total Annual Precipitation	(mm)	1211.67	140.0962	885.7	1679.6
Mean Annual Temperature	(°C)	12.61111	0.794708	11.1	14.3
Subbase layer/interlayer+ Sabgrade layer Thickness	(in)	85.22	2.403705	84	90
Interlayer BSG		2.635798	0.037737	2.589	2.714
AC Layer Below Surface (Binder Course) +Overlay Thickness	(in)	8.408889	6.089328	3.9	20.5
VEH CLASS 5 DIST PERCENT		31.99267	16.6675	9.34	53.47
VEH CLASS 6 DIST PERCENT		9.965778	6.385496	1.93	18.15
VEH CLASS 8 DIST PERCENT		6.436556	1.887951	4.31	9.3
VEH CLASS 10 DIST PERCENT		2.108778	1.101717	0.86	3.59

4.1.4 Developing Artificial Neural Network (ANN) Models

Development of ANN models involve three major components: setting up the architecture of the model designed to devise the structure of connection between the input and output layers, improving the learning method by adjusting the weights of connection, and an activation function to initiate the neurons of the network. The determination process of the ANN architecture plays the leading character in the construction of an optimum ANN model and adequate observation effort in the modeling process is demanded to figure out the architecture that serves the purpose of developing the specialized model. For the development of the PPPMs, the ANN model architecture used in this study consists of three layers: input layer, hidden layer, and output layer as shown in Figure 4.2. The independent variables that connect to the output layer are entered into the input layer and an individual neuron is assigned for each independent variable that served in the next phase of analysis intended to explain the effect of each input variable on the outcomes produced by the output layers. The number of neurons in the hidden layer should be chosen accordingly because that choice impacts model performance. Too many neurons used in the hidden layer can result in complexity in the model (Rafiq et al. 2001). The next challenge was the selection of the number of hidden layers because that choice impacts model performance significantly (Amin 2020). The number of layers decides the deep learning process of the model affecting the accuracy of the developed models. The single hidden layer model removes the limitations of the ambiguousness of the model, but this affects the learning process significantly because the single layer cannot yield reliable results causing drops in the performance of the model.

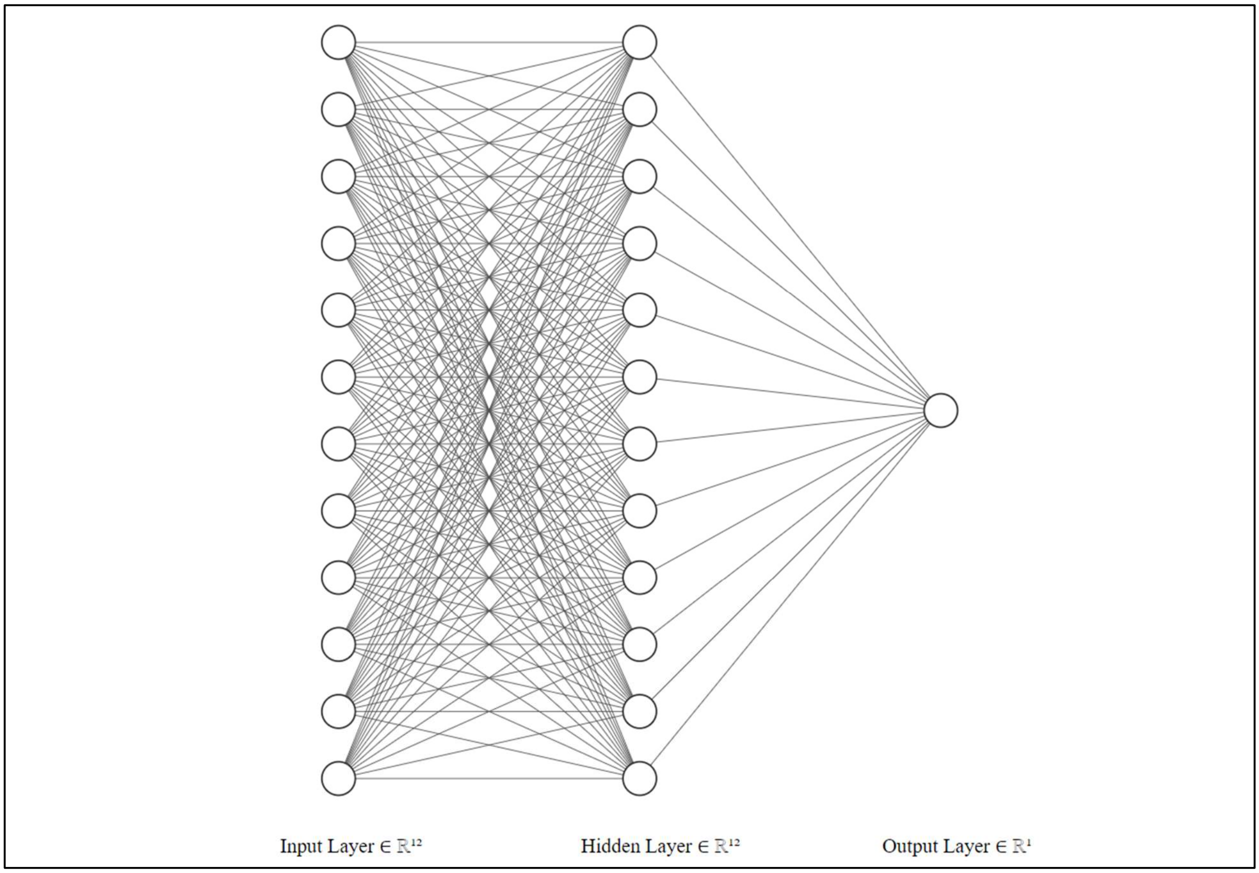


Figure 4.2: Basic structure of ANN utilized in PPPMs development

Artificial Neural Networks (ANNs) are a type of computational model that attempts to emulate the behavior of biological neurons. ANNs consist of an input layer, one or more hidden layers, and an output layer. The hidden layer(s) are responsible for processing the input data and generating output. The optimal number of hidden neurons is a crucial factor that contributes to the success of the ANN model. The process of determining the optimal number of hidden neurons involves several steps:

1. Collecting data: The first step in developing an ANN model was to collect data. The data should be representative of the problem that the ANN is intended to solve. In addition, the data

should be pre-processed and normalized to ensure that the ANN model can effectively learn from it.

2. Defining the architecture of the ANN: The architecture of an ANN refers to the number of layers, the number of neurons in each layer, and the type of activation function used. Determining the optimal number of hidden neurons was the focus of this step. Therefore, the input and output layers were fixed, and the number of neurons in the hidden layer(s) was focused.
3. Choosing a range of values for the number of hidden neurons: The next step was to choose a range of values for the number of hidden neurons. This range should be based on the complexity of the problem that the ANN is intended to solve. A simple problem may require only a few hidden neurons, while a more complex problem may require many more.
4. Training the ANN: Once the architecture of the ANN was defined, the next step was to train the model using the training data. During the training process, the weights of the connections between the neurons were adjusted to minimize the error between the predicted output and the actual output.
5. Evaluating the performance of the ANN: After the ANN was trained, the performance of the model is evaluated using the testing data. The evaluation metric is used to depend on the problem being solved.
6. Varying the number of hidden neurons: The decisive step in determining the optimal number of hidden neurons was to vary the number of neurons in the hidden layer(s) and evaluate the performance of the model at each value. This was done by repeating steps 4 and 5 for different values of the number of hidden neurons.

One important consideration when determining the optimal number of hidden neurons was overfitting. Overfitting occurs when the model is too complex and performs well on the training data but poorly on the testing data. This was because the model memorized the training data instead of learning the underlying patterns. To avoid overfitting, it is important to use techniques such as regularization, early stopping, or dropout.

Regularization involved adding a penalty term to the cost function that is being minimized during training. This penalty term discourages the model from becoming too complex and helps to prevent overfitting. Early stopping involves monitoring the performance of the model on a test set during training and stopping the training process when the performance on the test set starts to degrade. Dropout involved randomly dropping out some of the neurons in the model during training, which helps to prevent the model from relying too heavily on any one neuron. For an ANN-based pavement performance prediction model with six thousand data points, regularization was beneficial in several ways:

- Reducing model complexity: Regularization techniques such as L1 and L2 regularization added a penalty term to the loss function of the model. This penalty term discouraged the model from assigning too much importance to any one feature, thereby reducing the model's complexity.
- Avoiding overfitting: By reducing the model complexity, regularization helped to prevent overfitting, improving the model's ability to generalize to new, unseen data.
- Improving model performance: Regularization can also lead to better model performance, as it helps to remove noise and irrelevant features from the training data, allowing the model to focus on the most prominent features.

Therefore, by using regularization techniques such as L1 or L2 regularization, we can avoid overfitting in an ANN-based pavement performance prediction model and improve its ability to generalize to new data, resulting in more accurate and reliable predictions.

Figure 4.3 presents the flowchart for finding the optimal number of hidden neurons. This part introduces the performance indicators of the training process and the three steps for determining the optimal combination and number of hidden neurons: training, averaging, comparing, and testing.

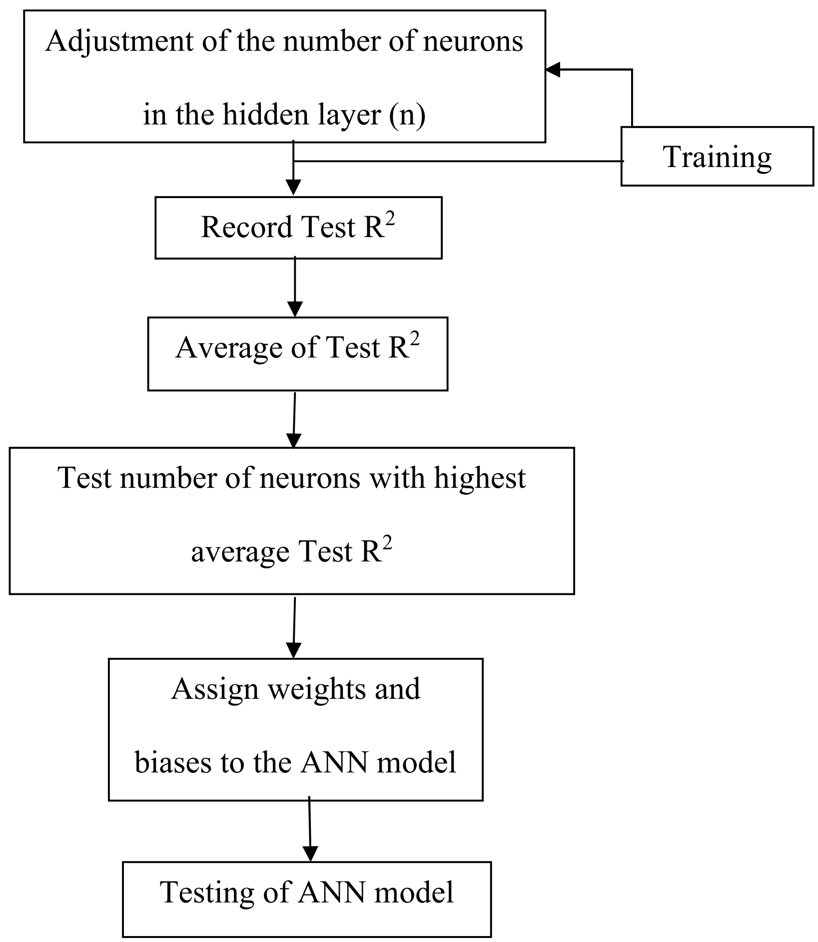


Figure 4.3: Flowchart for finding the optimal number of hidden neurons

For the development of the model, the first 15 years of pavement section data were used as training data, and the remaining 5 years of data served as test data, whereby it produces a mean root square error (RMSE) of training, Train R^2, and Test R^2 values. Figure 4.4 shows the performance of an ANN model for determining the number of neurons in the hidden layer for predicting the surface distress with their average outcomes where the highest average R^2 value was achieved at 12 hidden neurons. The number of optimal numbers of hidden neurons in both normalization methods was 12. The final layer in the structure of the ANN model is the output layer that produces the result from processing provided by the hidden layer.

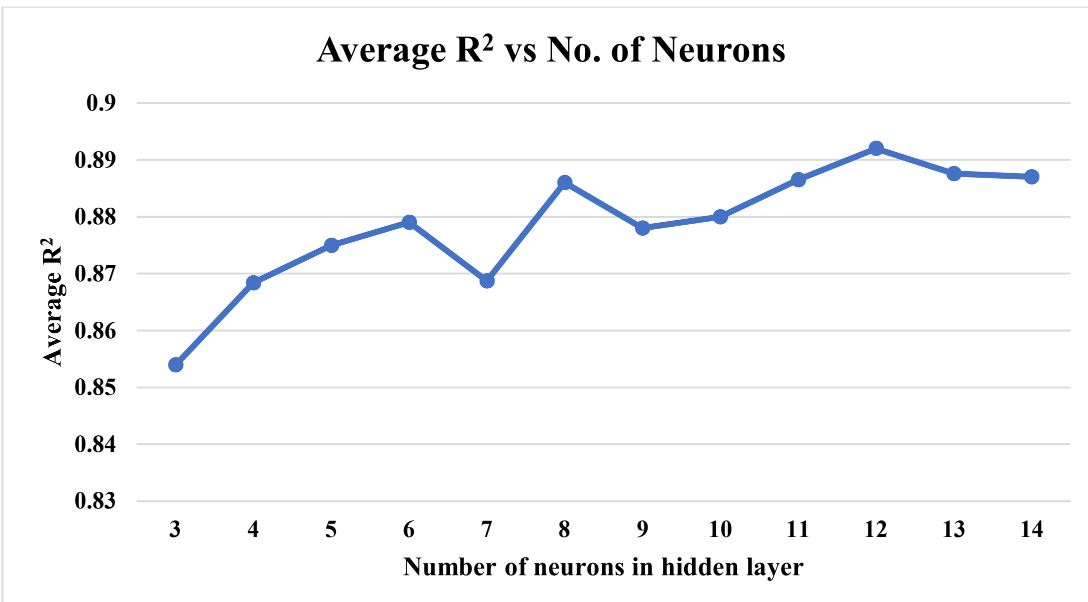


Figure 4.4: Performance of an ANN model based on neuron number

The training data set is used to develop the model, while the test data is used to assess the accuracy of the ANN model and avoid overfitting in the model (Ling et al. 2017). The results of the training process produce the weight matrices that are stored in links between layers and that can also be used to extract information about the contribution of each input in the model output.

In the training process, the connection weights between the layers are adjusted, thereby minimizing the overall mean error by using the back-propagation algorithm.

In an Artificial Neural Network (ANN) backpropagation model, the number of epochs is a crucial hyperparameter that influences the training process and the overall performance of the model. The selection process of the optimal number of epochs is important to achieve a model that is accurate, robust and not overfitting the training data.

The number of epochs refers to the number of times that the entire training dataset is passed through the model during the training process. A single epoch consists of a forward propagation and a backward propagation of the error signal, which updates the weights in the network. Typically, the more epochs the model is trained for, the more it can learn and improve its performance. However, there is a point at which additional epochs lead to overfitting, where the model becomes too specific to the training data and performs poorly on new, unseen data.

The process of selecting the optimal number of epochs for an ANN backpropagation model can be approached in several ways. One of the most common methods is to use a validation dataset, which is separate from the training dataset and is used to evaluate the performance of the model during the training process. The process involves the following steps:

1. Partitioning the data: The first step was to partition the data into three sets: training, validation, and testing. The training set was used to update the weights of the model during each epoch, the validation set was used to evaluate the performance of the model during the training process, and the testing set was used to evaluate the final performance of the model after training. The sample dataset containing thirty locations with 20 years of data was used for this

step. The first 15 years' data was used for training, the next 3 years' data for validation, and the last 2 years' data for testing.

2. Choosing a range of epochs: The next step was to choose a range of epochs to assess the model over. A frequent practice was to start with a small number of epochs and increase gradually, monitoring the validation loss at each epoch.
3. Training the model: The model was trained over the range of epochs while monitoring the validation loss. If the validation loss begins to increase, it is a sign of overfitting, and the training should stop to avoid further overfitting.
4. Evaluating the model performance: Once the model is trained, it is evaluated on the testing dataset to assess its performance. If the performance is satisfactory, the model can be deployed.

Several techniques can be used to prevent overfitting during training, which can influence the optimal number of epochs for the model. These include dropout, regularization, and early stopping. For this step, regularization was used.

In conclusion, the selection process of the optimal number of epochs for an ANN backpropagation model involves a range of techniques, including partitioning the data, defining the model architecture, choosing a range of epochs, training the model, and evaluating the performance of the model. By monitoring the validation loss and implementing techniques to prevent overfitting, a model can be trained that is accurate, robust, and generalizes well to new data.

The proposed ANN model considers the following elements: activation function, training algorithm and regularization, neural-network architecture, and the number of hidden neurons. Table 4.5 summarizes the methods selected for the elements, which are discussed in the following subsections.

Table 4.5: Methods Selected for the Elements

Model Element		**Method of Selection**
Training Algorithm		Back-propagation
Training Regularization		Bayesian regularization
Activation Function	Hidden Layer	Logistic activation function
	Output Layer	Linear activation function
Number of Epochs		1500

The logistic activation function and the hyperbolic tangent activation function can add non-linearity properties to the ANN model so that the model can learn from the data (Godfrey and Gashler 2015). The linear-activation function maps the pre-activation to itself, and the range of output values is between $(-\infty, \infty)$ (Herawan et al. 2017). The logistic and hyperbolic tangent activation functions are used for a hidden layer while the linear activation function is suitable for an output layer where the output values not constrained to any boundary are generated to predict target values. Between the activation functions for non-linearity, the convergence behavior of the logistic activation function shows a more non-linear pattern as the hyperbolic tangent activation function is almost linear at the low absolute values of the input variables. Therefore, PPPMs used the logistic activation function and linear activation function for the hidden layer and output layer, respectively.

Backpropagation is the most common and efficient training algorithm; hence, it was used as a training algorithm in this dissertation. In addition, in traditional neural network training, early stopping regularization is commonly used to avoid overfitting. Early stopping occurs when training is stopped before overfitting, which means that when the test error is at the minimum, the training process should be shut off (see Figure 4.5). However, it is difficult to detect when the test error reaches the minimum it is difficult to stop training at the right point. Most studies point out that Bayesian regularization performs better than early stopping in many cases. Bayesian regularization for neural networks is based on probabilistic interpretation to choose optimal sets of weights to minimize estimation error and efficiently avoid overfitting (Kayri 2016). The major advantage of using Bayesian regularization is that it does not require that the test dataset be separated from the training data set.

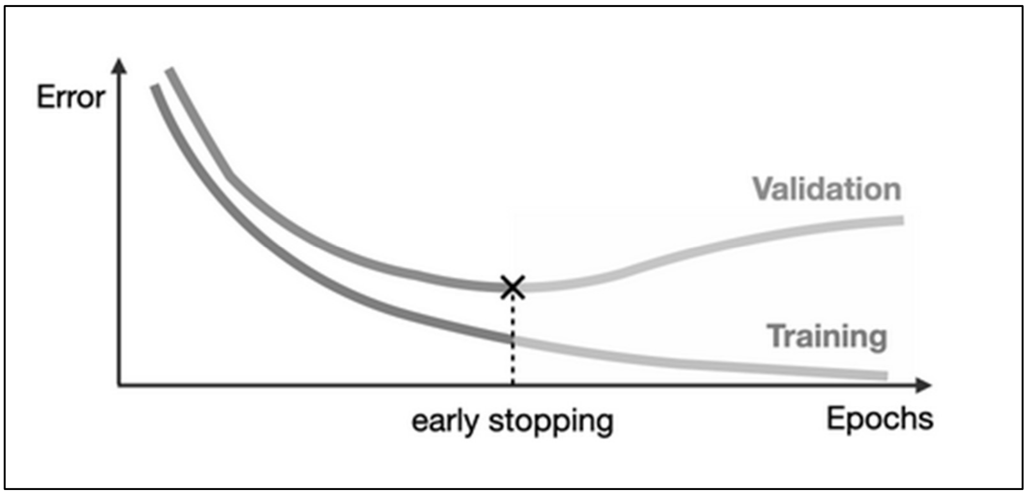


Figure 4.5: Training process of ANN

Despite the vast capability of ANN in prediction modeling, such models are often criticized as “black box” models, because of the difficulty in interpreting the contribution of each variable to the response variable making it hard to gain an understanding of the relationships among variables, which is considered a weakness when compared to traditional statistical models (Olden and Jackson 2002).

4.2 Result

This section presents the results of utilizing ANN to predict pavement performance conditions. To understand the predicted condition of the pavement, pavement conditions of each location was predicted by ANN models using current traffic and other important data input. The ANN model demonstrated the capability for predicting pavement conditions based on several variables and for estimating the relationship between traffic characteristics and pavement conditions at the network management level.

The performance of the ANN models was assessed to understand the accuracy of the models in predicting pavement performance as calculated by riding, cracking, potholes, patch failure, and polished aggregates indices. R^2 and RMSE values were used to measure and compare the performance of the models. Good prediction models should have a high R^2 and low RMSE. Historical data was used on ANN models to predict individual distresses for AC/flexible pavement sections. This individual distress was predicted based on weather factors (i.e., temperature, precipitation, and freeze-thaw cycles), traffic loading (AADT, AADTT, axle types, and loading), pavement age, layer thicknesses, material properties, and maintenance history. By predicting individual distresses, decision-makers can evaluate the individual distress for each pavement section, and determine which distress has more effect on the overall pavement condition.

After determining the architecture of each ANN model, the database for the period from 2001 to 2020 was randomly divided into training (first 15 years data) and test (next 5 years data) datasets. The training data set was used to develop the model whereas the test data was used to assess the accuracy of the ANN model and avoid overfitting in the model (Ling et al. 2017). The results of the training process produced the weight matrices that are stored in links between layers and that can also be used to extract information about the contribution of each input in the model output. The analysis showed the ANN models yield better predictions in terms of R^2 and RMSE.

4.2.1 IRI (Roughness) Model

The result of the IRI training models developed for three states is presented in Figure 4.6.

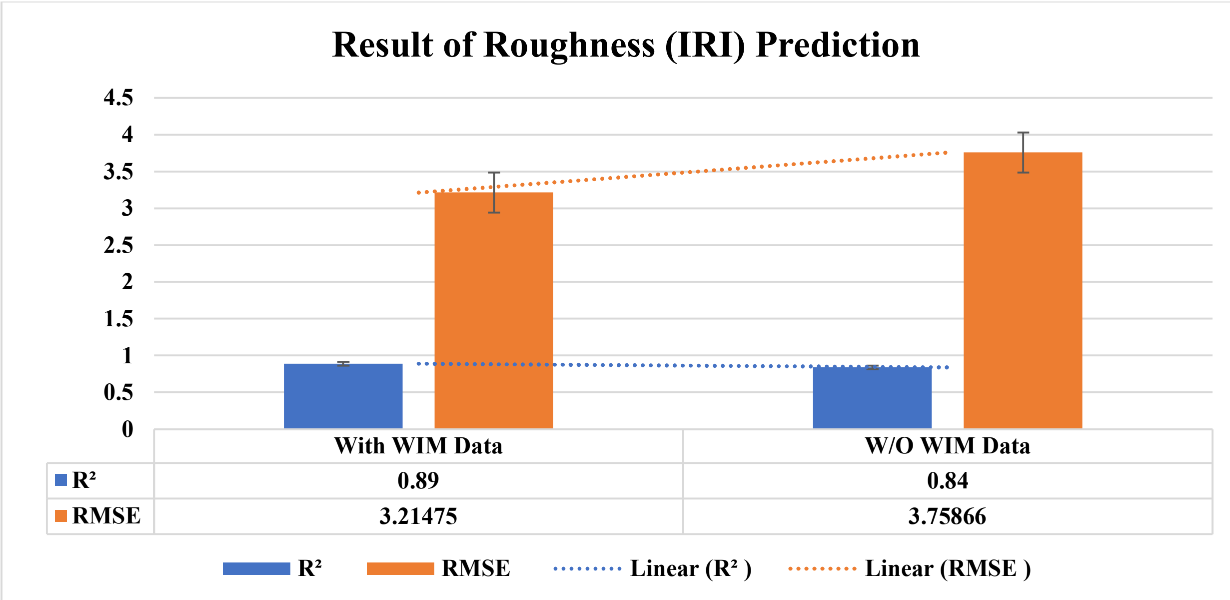


Figure 4.6: Performance for IRI from PPPMs

The result showed that the model generated from WIM data included performed better than without WIM data. This is true for both R^2 and RMSE. Overall, the performance of the PPPMs in terms of IRI prediction looked particularly good as it showed -high values for R^2 and low values for RMSE.

4.2.2 Cracking Models

Cracking is one of the most important measures of deterioration in bituminous pavements. Fatigue and aging have been identified as the principal factors which contribute to the cracking of a bituminous pavement layer. The propagation of cracking is accelerated through the embrittlement resulting from aging and the ingress of water, which can significantly weaken the underlying pavement layers. There are three types of cracking considered: longitudinal, transverse, and fatigue (alligator) cracking. For each type of cracking, separate relationships are given for predicting the time to initiation and then the rate of progression. Figure 4.7, Figure 4.8, and Figure

4.9 illustrate the performance of longitudinal, transverse, and fatigue (alligator) cracking, respectively.

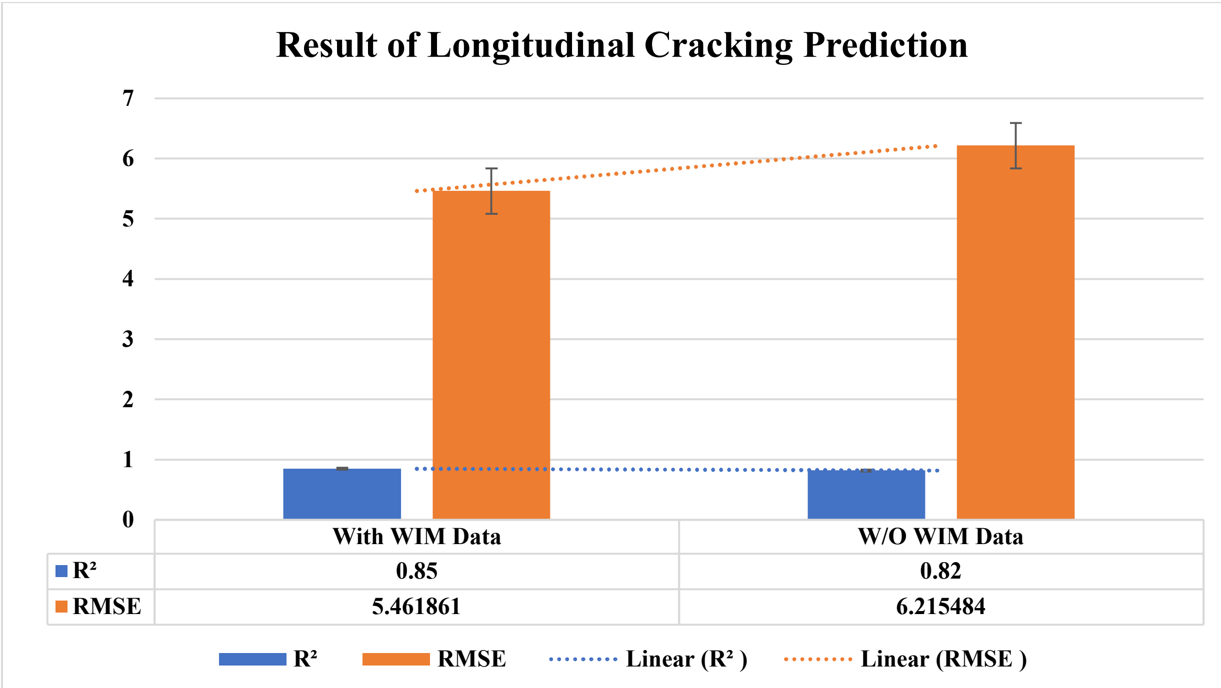


Figure 4.7: Prediction result of longitudinal cracking from PPPMs

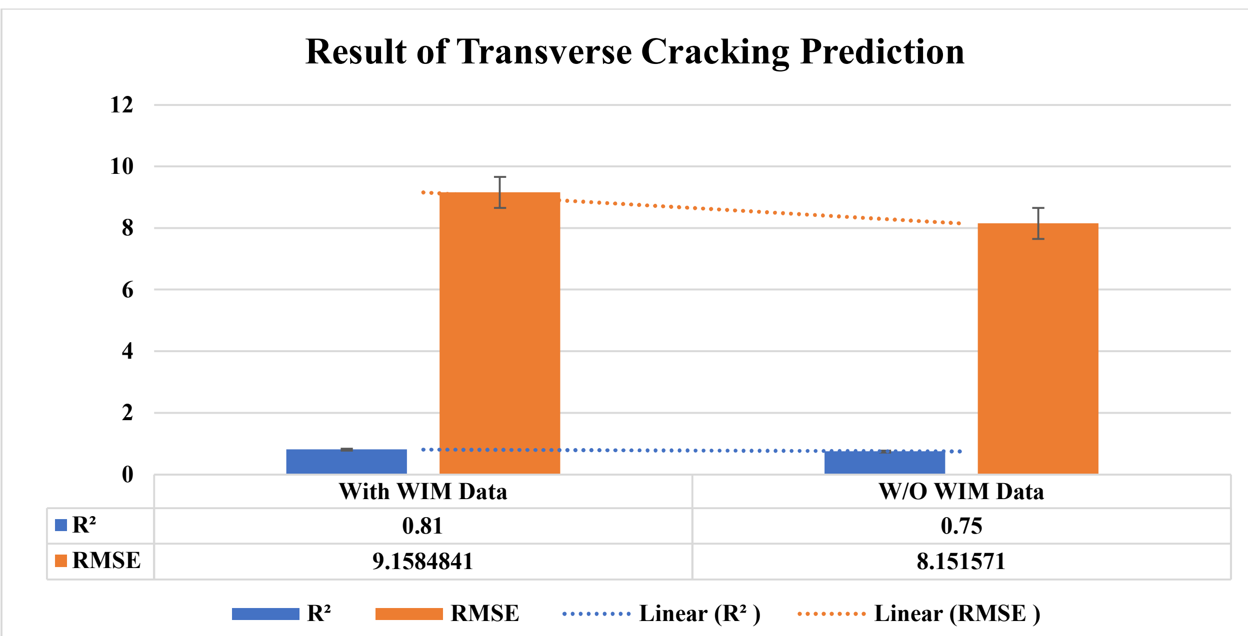


Figure 4.8: Prediction result of transverse cracking from PPPMs

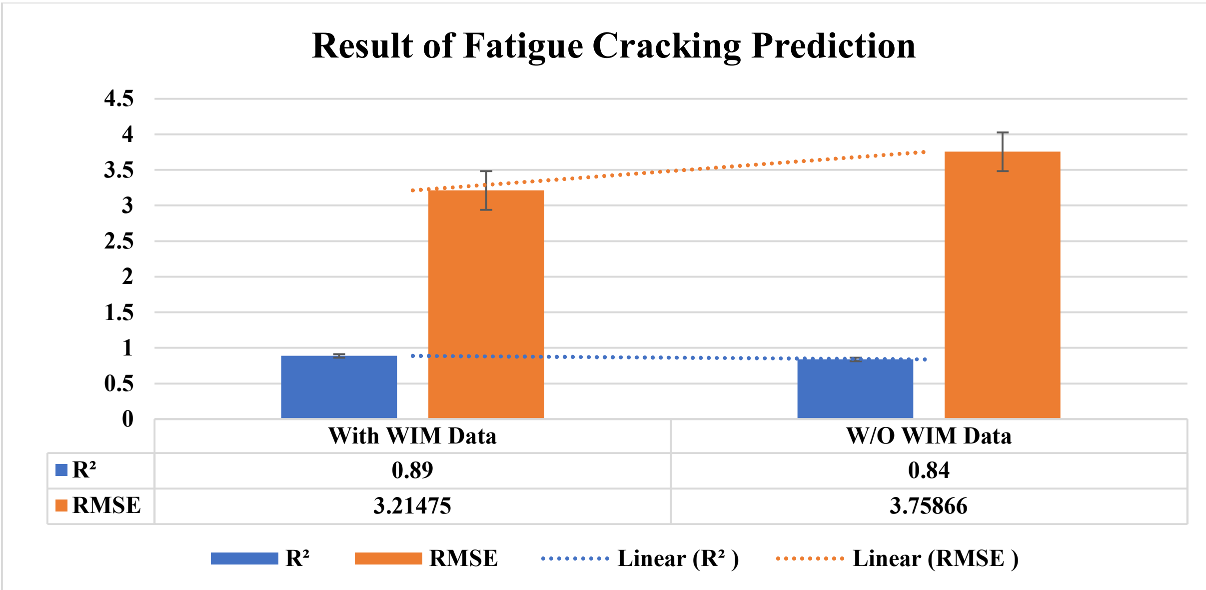


Figure 4.9: Prediction result of fatigue cracking from PPPMs

The result for cracking shows reliable performance for three crack types. For Longitudinal cracking, the best result was generated for the WIM data included model. The performance of PPPMs for longitudinal cracks shows that the generated can be used for different performance indicators with the same data and it can perform well. For the transverse crack, the performance without WIM data was better than the performance of the WIM data included model. The value of R^2 was high and RSME was comparatively low which indicated the performance was better. So, the model outputs for cracking were good considering high R^2 and low RMSE.

4.2.3 Potholes Model

Potholes usually develop on a surface that is either cracked, raveled, or both. The presence of water accelerates pothole formation both through a general weakening of the pavement structure and by lowering the resistance of the surface and base materials to disintegration. Figure 4.10 presents the result of the Pothole generated by PPPMs.

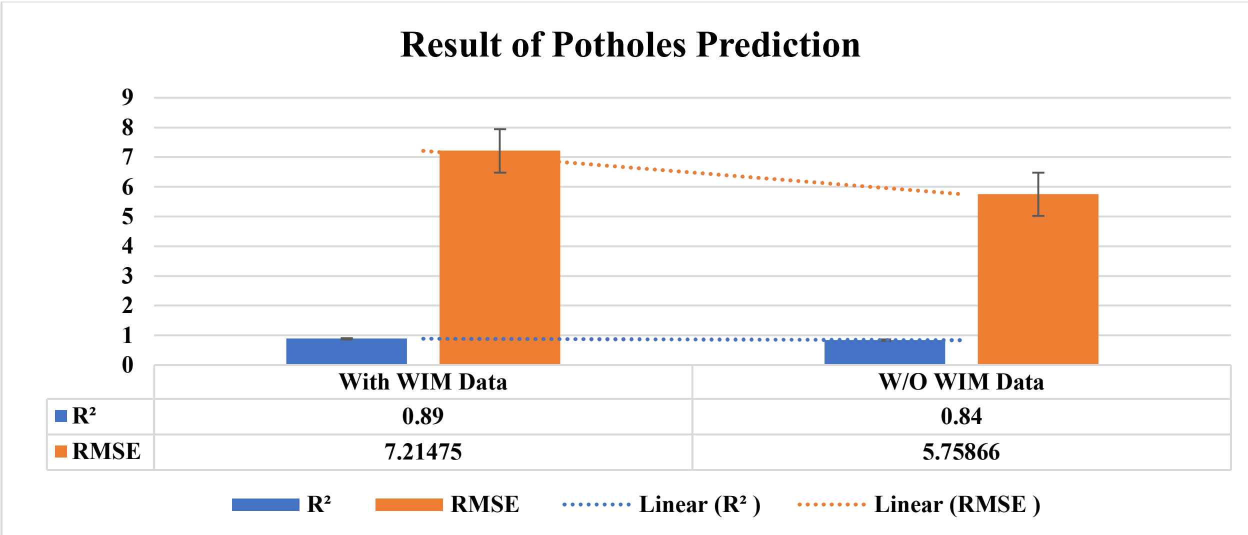


Figure 4.10: Prediction result of potholes from PPPMs

Both WIM-included data and without WIM data-based models had low R^2 values and slightly higher RSME. Among them, the result for WIM data-based model shows a slightly better performance than others.

4.2.4 Polished Aggregate Model

The micro-texture of road aggregates wears away under traffic action over time and gets polished. The result generated by the PPPMs for polished aggregate is presented in Figure 4.10.

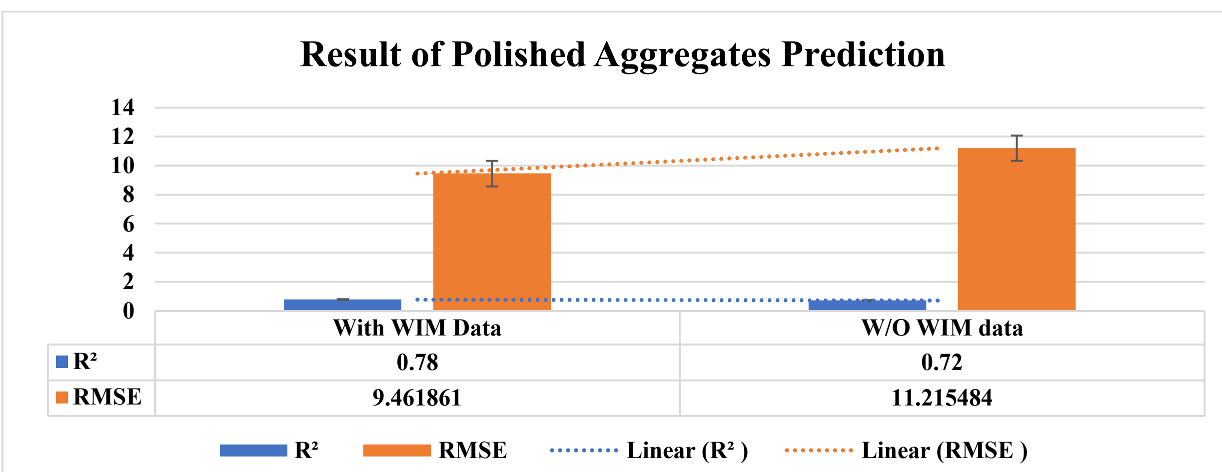


Figure 4.11: Prediction result of polished aggregate from PPPMs

The result of polished aggregates generated by PPPMs shows moderate performance in both WIM and without WIM data-based models. The models of WIM data included model performed better than the other as it has higher R^2, but lower RMSE. The moderate performance of the model can be a consequence of the non-present of polished aggregates on the pavement. So, the models have some limitations regarding zero values. So, further improvement of the PPPMs is needed so that they can adopt distinct types of indicators.

4.2.5 Patch Failure Model

Patching failure can cause serious harm to pavement structure by providing water to enter the pavement at a large volume and the underneath structure swells causing more damage to the pavement. Several reasons can cause the patch to fail. The results generated by PPPMs for patch failure are presented in Figure 4.12.

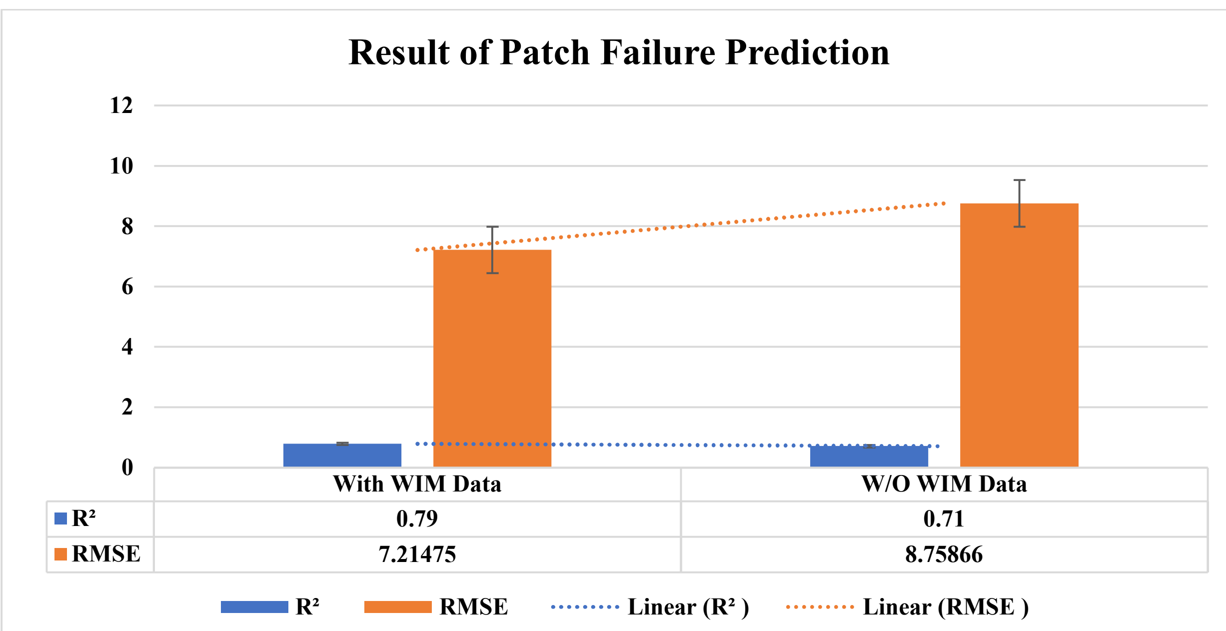


Figure 4.12: Prediction result of patch failure from PPPMs

The result for patch failure models for both normalized data types showed moderate-higher values for R^2 and lower values for RMSE. This indicated the patch failure models performed better, specifically for the WIM data included model which showed a slightly high value for R^2 and low value for RMSE than the other. Future investigation is required to improve the performance of the prediction.

4.3 Improve PPPMs Results Using WIM Data

Pavement performance prediction models are widely used in transportation engineering to assess the performance of roadways and to forecast maintenance needs. These models are developed based on a range of factors, including traffic loads, climate conditions, and pavement material properties. One of the most significant challenges in developing accurate pavement performance prediction models is the limited availability of reliable data. Weigh-In-Motion (WIM) technology has emerged as a promising solution to this problem, as it provides real-time data on traffic loads and vehicle characteristics that can be used to improve the accuracy of pavement performance prediction models.

By incorporating WIM data into pavement performance prediction models, several benefits can be achieved. Firstly, WIM data allows for a more accurate representation of the actual traffic loads that pavements are subjected to, as it captures data on the number of axles, axle weights, and vehicle speed. This information is critical in accurately estimating the damage that pavement will experience due to traffic loads. Secondly, WIM data can provide a more detailed understanding of the distribution of traffic loads on the pavement surface. This information can be used to develop models that consider the impact of distinct types of vehicles and axle configurations, which can improve the accuracy of pavement performance predictions.

Empirical results have demonstrated that the incorporation of WIM data can significantly improve the accuracy of pavement performance prediction models. The coefficient of determination (R^2) and Root Mean Squared Error (RMSE) are commonly used to evaluate the performance of pavement performance prediction models. This study has shown that the use of WIM data increased R^2 values by up to 10% and reduced RMSE by up to 31%, indicating a substantial improvement in model accuracy. The results of the improvements are presented in Figure 4.13.

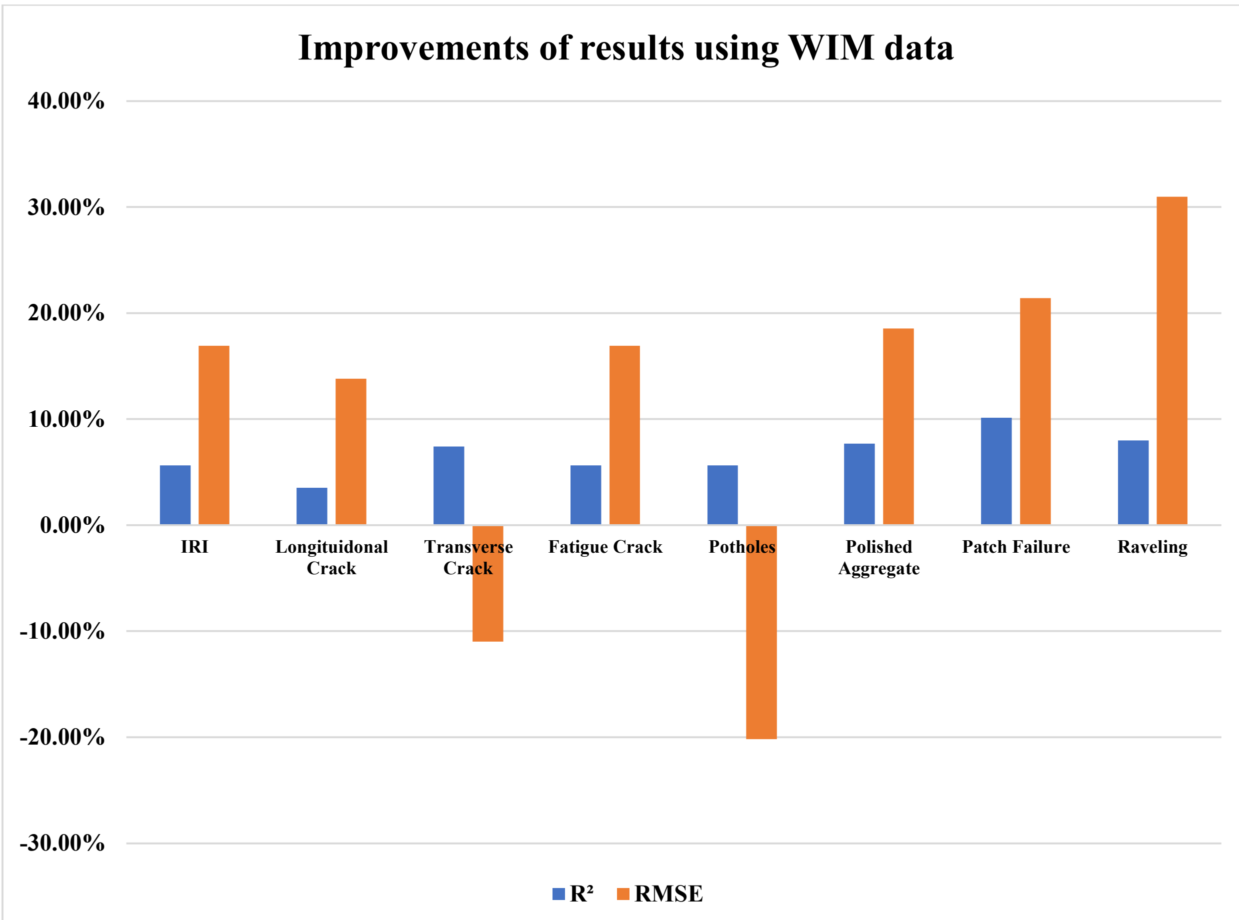


Figure 4.13: Illustration of performance improvements of PPPMs using WIM data

Furthermore, the use of WIM data in pavement performance prediction models can facilitate the identification of pavement distress mechanisms and provide valuable insights into the factors that influence pavement performance. This information can be used to develop more effective pavement management strategies and to prioritize maintenance and rehabilitation activities.

In conclusion, incorporating WIM data into pavement performance prediction models has the potential to significantly improve the accuracy of these models, providing valuable insights into the factors that influence pavement performance, and informing the development of more effective pavement management strategies.

4.4 Hypothesis Testing:

Hypothesis testing is a critical statistical technique used in research studies to determine whether a given hypothesis is supported by the available data or not. In this study, the null hypothesis was that Weigh-In-Motion (WIM) data has no significant impact on the performance of the pavement performance prediction model. To evaluate this hypothesis, a t-test was conducted, and the results showed that the p-value was greater than the significance level. Therefore, the null hypothesis was rejected, indicating that WIM data does have a significant impact on the performance of the Artificial Neural Network (ANN)-based pavement performance prediction model.

The t-test is a statistical test that determines whether the means of two groups are significantly different from each other or not. In this study, the two groups were the pavement performance prediction model without WIM data and the model with WIM data. The t-test was conducted to compare the mean performance of these two groups, and the resulting p-value indicated the probability of obtaining the observed difference in performance by chance. If the p-value is less than the significance level (usually set at 0.05), then the null hypothesis is rejected, indicating that there is a significant difference between the two groups. Table 4.6 showed the two-tailed P-values for RMSEs for eight types of performance indicators.

Table 4.6: P-values for RMSEs for seven types of performance indicators

Comparison of WIM data-based model performance with WITHOUTWIM data-based model	P- Value (RMSE)
IRI	0.010407
Longitudinal Crack	0.00568
Transverse Crack	0.006716
Fatigue Crack	0.008601
Potholes	0.009666
Polished Aggregate	0.007994
Patch Failure	0.006087

In this study, the results of the t-test showed that the p-value was greater than the significance level, indicating that there was a significant difference in the mean performance between the two groups. Since the null hypothesis stated that there was no impact of WIM data on the performance of the pavement performance prediction model, this result was sufficient to conclude the hypothesis testing. Therefore, the null hypothesis was rejected, and it was concluded that WIM data does have a significant impact on the performance of the pavement performance prediction model.

In conclusion, hypothesis testing is a powerful statistical technique that can be used to determine the impact of a range of factors on the performance of a given model or system. In this study, the null hypothesis was rejected, indicating that WIM data does have a significant impact on the performance of the pavement performance prediction model. This finding highlights the importance of considering WIM data in the development of pavement performance prediction models and can inform the development of more accurate and effective pavement management strategies.

Chapter 5. Conclusions and Recommendations

5.1 Summary of the Research

The primary objective of this dissertation is to create Pavement Performance Prediction Models (PPPMs) using the high-quality traffic loading data generated by Weigh-In-Motion (WIM) systems. The research focuses on developing PPPMs for flexible pavements based on the collected LTPP in 300 locations in the USA with WIM systems, intending to assist transportation agencies in managing the maintenance of their road infrastructure.

The dissertation begins with a review of the state-of-the-art machine learning-based modeling of pavement performance. Pavement infrastructure is capital-intensive and requires continuous monitoring to ensure its stability and acceptable serviceability. However, transportation agencies face the challenge of limited resources for maintenance work due to budget constraints. PPPMs have become crucial tools for providing optimal allocation of resources in maintenance activities. They are generated using inventory and monitoring data concerning the state of pavement structure, traffic load, and climate conditions.

PPPMs can be classified based on the formulation type, conceptual format, application level, and type of variables used. Machine learning based PPPMs draw generalizable predictive patterns, while statistical models draw population interpretations from a sample. These models relate pavement conditions (e.g., cracking, rutting) to a set of explanatory variables (e.g., traffic loadings, age, environmental factors, pavement design characteristics).

Among the various machine learning algorithms developed and implemented by the research community for PPPMs, artificial neural networks (ANNs) be the most widely used. Hence, this study employs ANN for the development of PPPMs. Six PPPMs based on distress types and one roughness based PPPM were developed using the data collected from 300 LTPP locations in the USA with the WIM system installed. The LTPP database was used to collect pavement condition and predictor types, including climatic, material, structural, and traffic data, through the InfoPave website.

The use of high-quality traffic load data generated by WIM systems in developing PPPMs holds immense potential. These models can predict the present and future condition of the pavement, which can be used to understand the pavement's performance and prioritize maintenance activities. The models developed in this dissertation can assist transportation agencies in optimizing the allocation of their limited resources for pavement maintenance.

5.2 Summary of the Results

Pavement performance prediction models are widely used in transportation engineering to assess the performance of roadways and to forecast maintenance needs. These models are developed based on a range of factors, including traffic loads, climate conditions, and pavement material properties. One of the most significant challenges in developing accurate pavement performance prediction models is the limited availability of reliable data. Weigh-In-Motion (WIM) technology has emerged as a promising solution to this problem, as it provides real-time data on traffic loads and vehicle characteristics that can be used to improve the accuracy of pavement performance prediction models.

By incorporating WIM data into pavement performance prediction models, several benefits can be achieved. Firstly, WIM data allows for a more accurate representation of the actual traffic loads that pavements are subjected to, as it captures data on the number of axles, axle weights, and vehicle speed. This information is critical in accurately estimating the damage that pavement will experience due to traffic loads. Secondly, WIM data can provide a more detailed understanding of the distribution of traffic loads on the pavement surface. This information can be used to develop models that consider the impact of distinct types of vehicles and axle configurations, which can improve the accuracy of pavement performance predictions.

Empirical results have demonstrated that the incorporation of WIM data can significantly improve the accuracy of pavement performance prediction models. The coefficient of determination (R^2) and Root Mean Squared Error (RMSE) are commonly used to evaluate the performance of pavement performance prediction models. This study has shown that the use of WIM data increased R^2 values by up to 10% and reduced RMSE by up to 31%, indicating a substantial improvement in model accuracy.

The t-test is a statistical test that determines whether the means of two groups are significantly different from each other or not. In this study, the two groups were the pavement performance prediction model without WIM data and the model with WIM data. The t-test was conducted to compare the mean performance of these two groups, and the resulting p-value indicated the probability of obtaining the observed difference in performance by chance. If the p-value is less than the significance level (usually set at 0.05), then the null hypothesis is rejected, indicating that there is a significant difference between the two groups.

The comparison of the WIM data-based model performance with the WITHOUTWIM data-based model showed P-values for RMSEs of IRI, Longitudinal Crack, Transverse Crack, Fatigue Crack, Potholes, Polished Aggregate, and Patch Failure. In this study, the results of the t-test showed that the p-value was greater than the significance level, indicating that there was a significant difference in the mean performance between the two groups. Since the null hypothesis stated that there was no impact of WIM data on the performance of the pavement performance prediction model, this result was sufficient to conclude the hypothesis testing. Therefore, the null hypothesis was rejected, and it was concluded that WIM data does have a significant impact on the performance of the pavement performance prediction model.

5.3 Limitations of the Study

The current research was conducted for interstate flexible pavements using pavement data collected from the LTPP database from three states. Therefore, the results and their implications are specific to the collected data, its quantity, quality, and variation. Since pavement data were collected from various states, data variation and unobserved heterogeneity may be higher because transportation agencies have different standards and specifications for pavement design, construction, and maintenance. Using data from different states may also result in spatial heterogeneity. For potholes and polished aggregates, many data points with zero value may cause inconsistent distribution and may this issue have resulted in the performance drops for PPPMs for these distress types. Moreover, the WIM system is not operated continuously, and this causes data issues. As most transportation agencies do not have a proper understanding of the significance of these systems in generating big data for PMS.

The ANN algorithm utilized for the PPPM development has some disadvantages as this algorithm cannot inherently process time-series data. Also, ANN requires processors with parallel processing power, by their structure. For these, it takes a lot of time to develop the model. There is no specific rule for determining the structure of ANN. For these, it cannot be said our developed models have outperformed others.

5.4 Contributions of the Research

This research makes various contributions to the body of knowledge and the body of practice in pavement infrastructure asset management. The overall contribution of this research is to enhance probabilistic pavement performance modeling that can support the development of sustainable decision-making strategies for PMS.

5.4.1 Contributions to the Body of Knowledge

Artificial Neural Networks (ANN) models have been widely used in various fields of engineering and science due to their ability to model complex systems, predict outcomes, and provide valuable insights into the problem at hand. This dissertation presented a critical analysis of various aspects of ANN-based models as applied in the literature, revealing research gaps in the relevant body of knowledge, and offering suggestions to address these gaps for future research. The study provides important contributions to the body of knowledge by highlighting the potential applications of Weigh-In-Motion (WIM) data in improving the accuracy of pavement performance prediction models (PPPM) and informing pavement management strategies.

The dissertation provided a thorough review of the literature on ANN-based models and highlighted the strengths and weaknesses of different ANN architectures, training algorithms, and input features used in PPPM. The analysis revealed that most of the models in the literature used a feedforward neural network architecture, which was found to be suitable for PPPM due to its ability to model nonlinear relationships between input and output variables. However, the study also showed that other ANN architectures, such as radial basis function networks and recurrent neural networks, can also be used in PPPM to improve model performance.

The study also identified research gaps in the use of ANN-based models in PPPM. One of the significant gaps was the limited availability of reliable data, which makes it challenging to develop accurate PPPM. To address this gap, the study proposed the use of WIM data, which provides real-time data on traffic loads and vehicle characteristics that can be used to improve the accuracy of PPPM. The study also highlighted the need for a comprehensive evaluation of the performance of different ANN-based models and input features in PPPM to identify the best-performing models and improve their accuracy.

The application of WIM data is one of the major contributions to the body of knowledge from this research. The study showed that incorporating WIM data into PPPM has several benefits, including a more accurate representation of actual traffic loads that pavements are subjected to, and a more detailed understanding of the distribution of traffic loads on the pavement surface. This information can be used to develop models that consider the impact of distinct types of vehicles and axle configurations, which can improve the accuracy of PPPM.

The successful application of WIM data has shown the potential for developing more accurate and effective pavement management strategies. Applications of WIM data can be explored in the direction of increasing efficiency of weight enforcement, increasing protection and preservation of pavement structure, improving highway safety, improving traffic operation and management, improving freight monitoring planning, and improving asset tracking. The use of WIM data can also help to reduce the need for expensive and time-consuming pavement inspections, leading to cost savings and more efficient use of resources.

The study's contributions to the body of knowledge can have significant implications for pavement management and transportation engineering. By incorporating WIM data into PPPM, transportation engineers can make more informed decisions regarding pavement design, maintenance, and rehabilitation. This can lead to significant cost savings and a more sustainable pavement infrastructure that meets the needs of the traveling public.

In conclusion, this dissertation provides valuable insights into the use of ANN-based models in PPPM and highlights the potential applications of WIM data in improving the accuracy of PPPM and informing pavement management strategies. The study's contributions to the body of knowledge can help researchers and practitioners better understand the available algorithms and select the appropriate one that can serve their purpose. The use of WIM data presents enormous opportunities for developing more insightful ideas that will support decision-making for PMS, including increasing the efficiency of weight enforcement, improving pavement structure protection, improving highway safety, improving traffic operation and management, improving freight monitoring planning, and improving asset tracking. The findings of this study can inform the development of more accurate and effective pavement management strategies that can bring sustainability to this system.

5.4.2 Contributions to the State of Practice

The successful development of the PPPMs can support policy development for minimizing the cost of the design life of the pavement section. The condition-based pavement performance prediction will open the potential of PPPMs applicability for the optimization of the maintenance activities and can help to generate a more cost-effective maintenance policy. The development of a cost-effective policy will make the service provided to the consumer more productive. The automation of the PMS can be achieved if the outcomes from research work are deeply investigated and further improved to make it applicable to real-life applications.

5.5 Recommendations for Future Research

The data utilized for this research is cleaned and processed only for missing values and outliers. The inconsistent values for two of the predictors showed performance drops due to a significant amount of zero values. In future works, further investigation will be done to know any relation between zero values and performance drop. Also, the data collected from the LTPP are time-series data. Recurrent neural networks (RNNs) are deep learning models, typically used to solve problems with sequential input data such as time series. RNNs are a type of neural network that retains a memory of what it has already processed and thus can learn from previous iterations during its training. It is a class of artificial neural networks where connections between nodes form a directed graph along a temporal sequence. This allows it to exhibit temporal dynamic behavior. Derived from feedforward neural networks, RNNs can use their internal state (memory) to process variable-length sequences of inputs. For accurately predicting future pavement conditions, developing an RNN model based on time-series data can be more advantageous. In the future, this aspect will be explored.

List of References

AASHTO, T. O. (1993). *AASHTO Guide for Design of Pavement Structures, 1993*, AASHTO, Washington D.C.

Abaza, K. A. (2004). "Deterministic performance prediction model for rehabilitation and management of flexible pavement." *International Journal of Pavement Engineering*, 5(2), 111-121.

Abbas, A. R., Fankhouser, A., and Papagiannakis, A. (2014). "Comparison between alternative methods for estimating vehicle class distribution input to pavement design." *Journal of Transportation Engineering*, 140(4), 04014004.

Abdelaziz, N., Abd El-Hakim, R. T., El-Badawy, S. M., and Afify, H. A. (2020). "International Roughness Index prediction model for flexible pavements." *International Journal of Pavement Engineering*, 21(1), 88-99.

Abraham, A. (2005). "Artificial neural networks." *Handbook of measuring system design*.

Adeli, H. (2001). "Neural networks in civil engineering: 1989–2000." *Computer-Aided Civil and Infrastructure Engineering*, 16(2), 126-142.

Adler, N., and Golany, B. (2001). "Evaluation of deregulated airline networks using data envelopment analysis combined with Principal Component Analysis with an application to Western Europe." *European Journal of Operational Research*, 132(2), 260-273.

Ahmed, A., Usman Saeed, T., and Labi, S. (2016). "Estimation of rest periods for newly constructed/reconstructed pavements." *Transport*, 31(2), 183-191.

Amin, S. (2020). "Backpropagation–Artificial Neural Network (BP-ANN): Understanding gender characteristics of older driver accidents in West Midlands of United Kingdom." *Safety science*, 122, 104539.

ARA (2004). "Guide for mechanistic-empirical design of new and rehabilitated pavement structures: NCHRP 1-37A Final Report." *TRB, National Research Council, Washington, DC*.

Arimie, C. O., Harcourt, P., Harcourt, P., and Harcourt, P. (2020). "Outlier detection and effects on modeling." *Open Access Library Journal*, 7(09), 1.

Ashtiani, R. S., Little, D. N., and Rashidi, M. (2018). "Neural network based model for estimation of the level of anisotropy of unbound aggregate systems." *Transportation Geotechnics*, 15, 4-12.

Attoh-Okine, N. O. "Predicting roughness progression in flexible pavements using artificial neural networks." *Proc., Transportation research board conference proceedings*.

Babashamsi, P., Yusoff, N. I. M., Ceylan, H., Nor, N. G. M., and Jenatabadi, H. S. (2016). "Evaluation of pavement life cycle cost analysis: Review and analysis." *International Journal of Pavement Research and Technology*, 9(4), 241-254.

Baghirli, O. (2015). "Comparison of Lavenberg-Marquardt, scaled conjugate gradient and Bayesian regularization backpropagation algorithms for multistep ahead wind speed forecasting using multilayer perceptron feedforward neural network."

Baker, J. (2019). "Auto-Calibration of WIM Using Traffic Stream Characteristics." Master of Science in Civil Engineering, University of Arkansas, Fayetteville.

Baladi, G. Y., Prohaska, M., Thomas, K., Dawson, T., and Musunuru, G. (2017). "Pavement performance measures and forecasting and the effects of maintenance and rehabilitation strategy on treatment effectiveness." United States. Federal Highway Administration.

Basheer, I. A., and Hajmeer, M. (2000). "Artificial neural networks: fundamentals, computing, design, and application." *Journal of microbiological methods*, 43(1), 3-31.

Bayrak, M. B., Teomete, E., and Agarwal, M. (2004). "Use of artificial neural networks for predicting rigid pavement roughness."

Bergan, D. A., Lindgren, N., Berthelot, D. C., and Woytowich, B. (1998). "Preserving highway infrastructure using weigh-in-motion (WIM)." *International Road Dynamics INC, Listopad*.

Beshears, D., Muhs, J., and Scudiere, M. (1998). "Advanced weigh-in-motion system for weighing vehicles at high speed." Oak Ridge National Lab., TN (United States).

Bianchini, A., and Bandini, P. (2010). "Prediction of pavement performance through neuro-fuzzy reasoning." *Computer-Aided Civil and Infrastructure Engineering*, 25(1), 39-54.

Burnos, P., Gajda, J., Piwowar, P., Sroka, R., Stencel, M., and Zeglen, T. (2007). "Measurements of road traffic parameters using inductive loops and piezoelectric sensors."

Chan, T. H., Miao, T., and Ashebo, D. B. (2005). "Statistical models from weigh-in-motion data." *Structural Engineering and Mechanics*, 20(1), 85-110.

Chandra, S., Sekhar, C. R., Bharti, A. K., and Kangadurai, B. (2013). "Relationship between pavement roughness and distress parameters for Indian highways." *Journal of transportation engineering*, 139(5), 467-475.

Cheng, L., Zhang, H., and Li, Q. (2007). "Design of a capacitive flexible weighing sensor for vehicle WIM system." *Sensors*, 7(8), 1530-1544.

Chicco, D., Warrens, M. J., and Jurman, G. (2021). "The coefficient of determination R-squared is more informative than SMAPE, MAE, MAPE, MSE and RMSE in regression analysis evaluation." *PeerJ Computer Science*, 7, e623.

Choi, S., and Do, M. (2020). "Development of the road pavement deterioration model based on the deep learning method." *Electronics*, 9(1), 3.

Choudhury, S. J., and Pal, N. R. (2019). "Imputation of missing data with neural networks for classification." *Knowledge-Based Systems*, 182, 104838.

Christopher, B. R., Schwartz, C. W., Boudreaux, R., and Berg, R. R. (2006). "Geotechnical aspects of pavements." United States. Federal Highway Administration.

Da Silva, I. N., Spatti, D. H., Flauzino, R. A., Liboni, L. H. B., and dos Reis Alves, S. F. (2017). "Artificial neural networks." *Cham: Springer International Publishing*, 39.

Dahlin, C. (1992). "Proposed method for calibrating weigh-in-motion systems and for monitoring that calibration over time." *Transportation Research Record*, 161-161.

Damirchilo, F., Hosseini, A., Mellat Parast, M., and Fini, E. H. (2021). "Machine Learning Approach to Predict International Roughness Index Using Long-Term Pavement Performance Data." *Journal of Transportation Engineering, Part B: Pavements*, 147(4), 04021058.

Dastres, R., and Soori, M. (2021). "Artificial neural network systems." *International Journal of Imaging and Robotics (IJIR)*, 21(2), 13-25.

De Veaux, R., and Ungar, L. (2012). "A Brief Introduction to Artificial Neural Networks." *Artif. Adapt. Syst. Med. New Theor. Models New Appl.*, 5.

De Wet, D. (2010). "WIM calibration and data quality management." *Journal of the South African Institution of Civil Engineering*, 52(2), 70-76.

Domitrović, J., Dragovan, H., Rukavina, T., and Dimter, S. (2018). "Application of an artificial neural network in pavement management system." *Tehnički vjesnik*, 25(Supplement 2), 466-473.

Duddu, V. R., Pulugurtha, S. S., Mane, A. S., and Godfrey, C. (2020). "Back-propagation neural network model to predict visibility at a road link-level." *Transportation research interdisciplinary perspectives*, 8, 100250.

Durandal, Z., and Zhang, F. R. (2019). "Weigh-in-Motion Auto-Calibration Using Automatic Vehicle Identification."

El-Badawy, S., Bayomy, F., and Awed, A. (2012). "Performance of MEPDG dynamic modulus predictive models for asphalt concrete mixtures: local calibration for Idaho." *Journal of Materials in Civil Engineering*, 24(11), 1412-1421.

El-Hakim, A., and El-Badawy, S. "International roughness index prediction for rigid pavements: an artificial neural network application." *Proc., Advanced Materials Research*, Trans Tech Publ, 854-860.

Elkins, G. E., Schmalzer, P., Thompson, T., and Simpson, A. (2003). "Long-term pavement performance information management system: Pavement performance database user reference guide (No. FHWA-RD-03-088)." *Turner-Fairbank Highway Research Center*.

Elkins, L., and Higgins, C. (2008). "Development of truck axle spectra from Oregon weigh-in-motion data for use in pavement design and analysis." Oregon. Dept. of Transportation. Research Unit.

Felker, V., Najjar, Y. M., and Hossain, M. (2004). "Modeling the roughness progression on Kansas portland cement concrete (PCC) pavements."

FHWA (2001). "Federal Highway Administration Traffic Monitoring Guide." *US Department of Transportation, Washington, DC*.

Fuentes, L. G., Macea, L. F., Vergara, A., Flintsch, G. W., Alvarez, A. E., and Reyes, O. J. (2012). "Evaluation of truck factors for pavement design in developing countries." *Procedia-Social and Behavioral Sciences*, 53, 1139-1148.

Gaspareto, D. S., and Gomes, H. M. (2019). "Bending plate WIM system analysis considering the dynamics of the load platform." *ICWIM8*, 304.

Georgiou, P., Plati, C., and Loizos, A. (2018). "Soft computing models to predict pavement roughness: A comparative study." *Advances in Civil Engineering*, 2018.

Gewers, F. L., Ferreira, G. R., Arruda, H. F. D., Silva, F. N., Comin, C. H., Amancio, D. R., and Costa, L. d. F. (2021). "Principal component analysis: A natural approach to data exploration." *ACM Computing Surveys (CSUR)*, 54(4), 1-34.

Ghasemi, P., Aslani, M., Rollins, D. K., Williams, R. C., and Schaefer, V. R. (2018). "Modeling rutting susceptibility of asphalt pavement using principal component pseudo inputs in regression and neural networks." *International Journal of Pavement Research and Technology*, 11(7), 679-690.

Ghosh, A., Sufian, A., Sultana, F., Chakrabarti, A., and De, D. (2020). "Fundamental concepts of convolutional neural network." *Recent trends and advances in artificial intelligence and Internet of Things*, 519-567.

Godfrey, L. B., and Gashler, M. S. "A continuum among logarithmic, linear, and exponential functions, and its potential to improve generalization in neural networks." *Proc., 2015 7th international joint conference on knowledge discovery, knowledge engineering and knowledge management (IC3K)*, IEEE, 481-486.

Golshani, N., Shabanpour, R., Mahmoudifard, S. M., Derrible, S., and Mohammadian, A. (2018). "Modeling travel mode and timing decisions: Comparison of artificial neural networks and copula-based joint model." *Travel Behaviour and Society*, 10, 21-32.

Gong, H., Sun, Y., Shu, X., and Huang, B. (2018). "Use of random forests regression for predicting IRI of asphalt pavements." *Construction and Building Materials*, 189, 890-897.

Gupta, D., Tang, X., and Yuan, L. (2018). "Weigh-in-Motion Sensor and Controller Operation and Performance Comparison." Minnesota. Dept. of Transportation. Research Services & Library.

Haider, S. W., and Harichandran, R. S. (2009). "Effect of axle load spectrum characteristics on flexible pavement performance." *Transportation research record*, 2095(1), 101-114.

Haider, S. W., and Masud, M. M. "Accuracy Comparisons Between ASTM 1318-09 and COST-323 (European) WIM Standards Using LTPP WIM Data." *Proc., Proceedings of the 9th International Conference on Maintenance and Rehabilitation of Pavements—Mairepav9*, Springer, 155-165.

Han, J., Kamber, M., and Pei, J. (2012). "Data mining concepts and techniques third edition." *University of Illinois at Urbana-Champaign Micheline Kamber Jian Pei Simon Fraser University*.

Hasan, B. M. S., and Abdulazeez, A. M. (2021). "A review of Principal Component Analysis algorithm for dimensionality reduction." *Journal of Soft Computing and Data Mining*, 2(1), 20-30.

Hasan, M. K., Alam, M. A., Roy, S., Dutta, A., Jawad, M. T., and Das, S. (2021). "Missing value imputation affects the performance of machine learning: A review and analysis of the literature (2010–2021)." *Informatics in Medicine Unlocked*, 27, 100799.

Hashemi Vaziri, S., Haas, C. T., Rothenburg, L., Haas, R. C., and Jiang, X. (2013). "Investigation of the effects of air temperature and speed on performance of piezoelectric weigh-in-motion systems." *Canadian Journal of Civil Engineering*, 40(10), 935-944.

Haugen, T., Levy, J. R., Aakre, E., and Tello, M. E. P. (2016). "Weigh-in-Motion equipment–experiences and challenges." *Transportation Research Procedia*, 14, 1423-1432.

Hazlett, D., Jiang, N., and Loftus-Otway, L. (2020). "Use of Weigh-in-Motion Data for Pavement, Bridge, Weight Enforcement, and Freight Logistics Applications."

Herawan, T., Ghazali, R., and Deris, M. M. (2017). "Recent advances on soft computing and data mining." *Sl: Springer*.

Holmgren, G., Andersson, P., Jakobsson, A., and Frigyesi, A. (2019). "Artificial neural networks improve and simplify intensive care mortality prognostication: a national cohort study of 217,289 first-time intensive care unit admissions." *Journal of Intensive Care*, 7(1), 44.

Hong, H., and Somo, S. (2001). "Probabilistic assessment of pavement conditions." *Canadian Journal of Civil Engineering*, 28(5), 813-822.

Hossain, M., Gopisetti, L. S. P., and Miah, M. S. (2020). "Artificial neural network modelling to predict international roughness index of rigid pavements." *International Journal of Pavement Research and Technology*, 1-11.

Jain, A. K., Mao, J., and Mohiuddin, K. M. (1996). "Artificial neural networks: A tutorial." *Computer*, 29(3), 31-44.

Jiang, X., Vaziri, S. H., Haas, C., Rothenburg, L., Kennepohl, G., and Haas, R. "Improvements in piezoelectric sensors and WIM data collection technology." *Proc., Proceedings of the 2009 Annual Conference of the Transportation Association of Canada*.

Jiang, Y., Li, S., Nantung, T., Mangold, K., and MacArthur, S. A. "Creation of Truck Axle Load Spectra Using Weigh-in-Motion Data." *Proc., Journal of the transportation Research Forum*.

Jolliffe, I. T., and Cadima, J. (2016). "Principal component analysis: a review and recent developments." *Philosophical transactions of the royal society A: Mathematical, Physical and Engineering Sciences*, 374(2065), 20150202.

Justo-Silva, R., Ferreira, A., and Flintsch, G. (2021). "Review on Machine Learning Techniques for Developing Pavement Performance Prediction Models." *Sustainability*, 13(9), 5248.

Kargah-Ostadi, N., and Stoffels, S. M. (2015). "Framework for development and comprehensive comparison of empirical pavement performance models." *Journal of Transportation Engineering*, 141(8), 04015012.

Kargah-Ostadi, N., Stoffels, S. M., and Tabatabaee, N. (2010). "Network-level pavement roughness prediction model for rehabilitation recommendations." *Transportation Research Record*, 2155(1), 124-133.

Kargah-Ostadi, N., Zhou, Y., and Rahman, T. (2019). "Developing performance prediction models for pavement management systems in local governments in absence of age data." *Transportation Research Record*, 2673(3), 334-341.

Karim, M. R., Ibrahim, N. I., Saifizul, A. A., and Yamanaka, H. (2014). "Effectiveness of vehicle weight enforcement in a developing country using weigh-in-motion sorting system considering vehicle by-pass and enforcement capability." *IATSS research*, 37(2), 124-129.

Karlaftis, M. G., and Vlahogianni, E. I. (2011). "Statistical methods versus neural networks in transportation research: Differences, similarities and some insights." *Transportation Research Part C: Emerging Technologies*, 19(3), 387-399.

Katz, A. J., and Rakha, H. A. (2002). "Field and modeling framework and case study of truck weigh station operations."

Kaur, D., and Datta, D. "Soft computing technique in prediction of pavement condition." *Proc., 6th WSEAS international conference on computational intelligence*.

Kayri, M. (2016). "Predictive abilities of bayesian regularization and Levenberg–Marquardt algorithms in artificial neural networks: a comparative empirical study on social data." *Mathematical and Computational Applications*, 21(2), 20.

Khalili, M., Vishwakarma, G., Ahmed, S., and Papagiannakis, A. T. (2022). "Development of a low-power weigh-in-motion system using cylindrical piezoelectric elements." *International Journal of Transportation Science and Technology*, 11(3), 496-508.

Kherif, F., and Latypova, A. (2020). "Principal component analysis." *Machine Learning*, Elsevier, 209-225.

Kim, S.-H., and Kim, N. (2006). "Development of performance prediction models in flexible pavement using regression analysis method." *KSCE Journal of Civil Engineering*, 10(2), 91-96.

Kim, T. K., and Park, J. H. (2019). "More about the basic assumptions of t-test: normality and sample size." *Korean journal of anesthesiology*, 72(4), 331-335.

LeCun, Y., Boser, B., Denker, J., Henderson, D., Howard, R., Hubbard, W., and Jackel, L. (1989). "Handwritten digit recognition with a back-propagation network." *Advances in neural information processing systems*, 2.

Li, J. Q., Hossain, N., Wang, K. C., and Zaman, M. (2018). "Development of Statewide WIM Data Quality Control and Axle Load Spectra and Traffic Volume Adjustment Factors for Oklahoma."

Liao, C.-F. (2014). "Generating Reliable Freight Performance Measures with Truck GPS Data: Case Study in Twin Cities Metropolitan Area, Minnesota." *Transportation Research Record*, 2410(1), 21-30.

Lin, J.-D., Yau, J.-T., and Hsiao, L.-H. "Correlation analysis between international roughness index (IRI) and pavement distress by neural network." *Proc., 82nd Annual Meeting of the Transportation Research Board*, 12-16.

Lin, W.-C., Tsai, C.-F., and Zhong, J. R. (2022). "Deep learning for missing value imputation of continuous data and the effect of data discretization." *Knowledge-Based Systems*, 239, 108079.

Ling, M., Luo, X., Hu, S., Gu, F., and Lytton, R. L. (2017). "Numerical modeling and artificial neural network for predicting J-integral of top-down cracking in asphalt pavement." *Transportation Research Record*, 2631(1), 83-95.

Liu, T.-H., and Yoon, Y. (2019). "Effects of meteorological factors to reduce large-scale PM10 emission estimation errors on unpaved roads." *Atmospheric Environment*, 217, 116956.

Lou, Z., Gunaratne, M., Lu, J., and Dietrich, B. (2001). "Application of neural network model to forecast short-term pavement crack condition: Florida case study." *Journal of infrastructure systems*, 7(4), 166-171.

Lourakis, M. I. (2005). "A brief description of the Levenberg-Marquardt algorithm implemented by levmar." *Foundation of Research and Technology*, 4(1), 1-6.

Lu, Q., Harvey, J., Le, T., Lea, J., Quinley, R., Redo, D., and Avis, J. (2002). "Truck traffic analysis using weigh-in-motion (WIM) data in California." *University of California, Berkeley, Institute of Transportation Studies, Pavement Research Center, Berkeley, California, USA*.

Mahmood, M. S. (2015). *Network-level maintenance decisions for flexible pavement using a soft computing-based framework*, Nottingham Trent University (United Kingdom).

Maind, S. B., and Wankar, P. (2014). "Research paper on basic of artificial neural network." *International Journal on Recent and Innovation Trends in Computing and Communication*, 2(1), 96-100.

Marcelino, P., de Lurdes Antunes, M., Fortunato, E., and Gomes, M. C. (2020). "Transfer learning for pavement performance prediction." *International Journal of Pavement Research and Technology*, 13(2), 154-167.

Marcelino, P., de Lurdes Antunes, M., Fortunato, E., and Gomes, M. C. (2021). "Machine learning approach for pavement performance prediction." *International Journal of Pavement Engineering*, 22(3), 341-354.

Mazari, M., and Rodriguez, D. D. (2016). "Prediction of pavement roughness using a hybrid gene expression programming-neural network technique." *Journal of Traffic and Transportation Engineering (English Edition)*, 3(5), 448-455.

McCall, B., and Vodrazka Jr, W. C. (1997). "States' successful practices weigh-in-motion handbook." United States. Federal Highway Administration.

Mehta, A. J., Mehta, H. A., Manjunath, T., and Ardil, C. (2008). "A multi-layer artificial neural network architecture design for load forecasting in power systems." *International Journal of Electrical and Computer Engineering*, 5(2), 207-220.

Mishra, P., Pandey, C. M., Singh, U., Gupta, A., Sahu, C., and Keshri, A. (2019). "Descriptive statistics and normality tests for statistical data." *Annals of cardiac anaesthesia*, 22(1), 67.

Montgomery, S. R., Labi, S., and Haddock, J. E. (2018). "Load transfer restoration with diamond grinding on rigid pavements: Short-term and long-term effectiveness." *Journal of Transportation Engineering, Part B: Pavements*, 144(1), 04017019.

Nassif, H., Ozbay, K., Na, C., Lou, P., Fiorillo, G., and Park, C. (2018). "Monitoring and Control of Overweight Trucks for Smart Mobility and Safety of Freight Operations."

Nemati, R., and Dave, E. V. (2018). "Nominal property based predictive models for asphalt mixture complex modulus (dynamic modulus and phase angle)." *Construction and Building Materials*, 158, 308-319.

Nian, T., Li, J., Li, P., Liu, Z., Guo, R., Ge, J., and Wang, M. (2022). "Method to predict the interlayer shear strength of asphalt pavement based on improved back propagation neural network." *Construction and Building Materials*, 351, 128969.

Nitsche, P., Stütz, R., Kammer, M., and Maurer, P. (2014). "Comparison of machine learning methods for evaluating pavement roughness based on vehicle response." *Journal of Computing in Civil Engineering*, 28(4), 04014015.

Nnamoko, N., and Korkontzelos, I. (2020). "Efficient treatment of outliers and class imbalance for diabetes prediction." *Artificial intelligence in medicine*, 104, 101815.

Nowak-Brzezińska, A., and Łazarz, W. (2021). "Qualitative data clustering to detect outliers." *Entropy*, 23(7), 869.

Okunev, R. (2022). "Hypothesis Testing and Interpretation of Results." *Analytics for Retail: A Step-by-Step Guide to the Statistics Behind a Successful Retail Business*, Springer, 77-81.

Olden, J. D., and Jackson, D. A. (2002). "Illuminating the “black box”: a randomization approach for understanding variable contributions in artificial neural networks." *Ecological modelling*, 154(1-2), 135-150.

Omuya, E. O., Okeyo, G. O., and Kimwele, M. W. (2021). "Feature selection for classification using Principal Component Analysis and information gain." *Expert Systems with Applications*, 174, 114765.

Ozbay, K., and Laub, R. (2001). "Models for pavement deterioration using LTPP."

Papagiannakis, A., Bracher, M., and Jackson, N. (2006). "Utilizing clustering techniques in estimating traffic data input for pavement design." *Journal of Transportation Engineering*, 132(11), 872-879.

Papagiannakis, A. T., Quinley, R., and Brandt, S. (2008). *High speed weigh-in-motion system calibration practices*, Transportation Research Board.

Park, S. H., and Kim, J. H. (2019). "Comparative analysis of performance prediction models for flexible pavements." *Journal of Transportation Engineering, Part B: Pavements*, 145(1), 04018062.

Plati, C. (2019). "Sustainability factors in pavement materials, design, and preservation strategies: A literature review." *Construction and Building Materials*, 211, 539-555.

Prozzi, J. A., and Hong, F. (2007). "Effect of weigh-in-motion system measurement errors on load-pavement impact estimation." *Journal of Transportation Engineering*, 133(1), 1-10.

Qin, T., Lin, M., Cao, M., Fu, K., and Ding, R. (2018). "Effects of sensor location on dynamic load estimation in weigh-in-motion system." *Sensors*, 18(9), 3044.

Rafiq, M., Bugmann, G., and Easterbrook, D. (2001). "Neural network design for engineering applications." *Computers & Structures*, 79(17), 1541-1552.

Rajagopal, A. (2006). "Developing pavement performance prediction models and decision trees for the city of Cincinnati."

Ramachandran, A. N., Taylor, K. L., Stone, J. R., and Sajjadi, S. S. (2011). "NCDOT quality control methods for weigh-in-motion data." *Public Works Management & Policy*, 16(1), 3-19.

Refai, H. (2013). "Development of a portable weigh-in-motion system." Oklahoma Transportation Center.

Roberts, C. A., and Attoh-Okine, N. O. (1998). "A comparative analysis of two artificial neural networks using pavement performance prediction." *Computer-Aided Civil and Infrastructure Engineering*, 13(5), 339-348.

Romanoschi, S. A., Momin, S., Bethu, S., and Bendana, L. (2011). "Development of Traffic Inputs for New Mechanistic–Empirical Pavement Design Guide: Case Study." *Transportation research record*, 2256(1), 142-150.

Rose, S., Mathew, B. S., Isaac, K. P., and Abhaya, A. (2018). "Risk based probabilistic pavement deterioration prediction models for low volume roads." *International Journal of Pavement Engineering*, 19(1), 88-97.

Roweis, S. (1996). "Levenberg-marquardt optimization." *Notes, University Of Toronto*, 52.

Rys, D. (2019). "Investigation of weigh-in-motion measurement accuracy on the basis of steering axle load spectra." *Sensors*, 19(15), 3272.

S. Miller, J., and Y. Bellinger, W. (2014). "Distress Identification Manual for the Long-Term Pavement Performance Program (Fifth Revised Edition)." Office of Infrastructure Research and Development, Federal Highway Administration.

Saini, L. M. (2008). "Peak load forecasting using Bayesian regularization, Resilient and adaptive backpropagation learning based artificial neural networks." *Electric power systems research*, 78(7), 1302-1310.

Salama, H. K., Chatti, K., and Lyles, R. W. (2006). "Effect of heavy multiple axle trucks on flexible pavement damage using in-service pavement performance data." *Journal of transportation engineering*, 132(10), 763-770.

Santero, N., Nokes, W., and Harvey, J. T. (2005). *Virtual Weigh Stations in California: A Preliminary Cost-Effectiveness Analysis*, Citeseer.

Schmidt, F., Jacob, B., and Domprobst, F. (2016). "Investigation of truck weights and dimensions using WIM data." *Transportation Research Procedia*, 14, 811-819.

Selezneva, O., and Wolf, D. J. (2017). "Successful Practices in Weigh-in-Motion Data Quality with WIM Guidebook [Volumes 1 & 2]." Arizona. Dept. of Transportation.

Sharma, B., and Venugopalan, K. (2014). "Comparison of neural network training functions for hematoma classification in brain CT images." *IOSR Journal of Computer Engineering*, 16(1), 31-35.

Shatnawi, S. (2008). "Maintenance Technical Advisory Guide." *Volume I-Flexible Pavement Preservation Section Edition*.

Solatifar, N., and Lavasani, S. M. (2020). "Development of an Artificial Neural Network model for asphalt pavement deterioration using LTPP data." *Journal of Rehabilitation in Civil Engineering*, 8(1), 121-132.

Sollazzo, G., Fwa, T., and Bosurgi, G. (2017). "An ANN model to correlate roughness and structural performance in asphalt pavements." *Construction and Building Materials*, 134, 684-693.

Sun, Z., and Ban, X. J. (2013). "Vehicle classification using GPS data." *Transportation Research Part C: Emerging Technologies*, 37, 102-117.

Swan, D., Tardif, R., Hajek, J. J., and Hein, D. K. (2008). "Development of regional traffic data for the mechanistic–empirical pavement design guide." *Transportation research record*, 2049(1), 54-62.

Taylor, B., Bergan, A., Lindgren, N., and Eng, C. B. P. D. P. (2000). "The importance of commercial vehicle weight enformcement in safety and road asset management." *Annual Review pp*, 234, 237.

Ticknor, J. L. (2013). "A Bayesian regularized artificial neural network for stock market forecasting." *Expert systems with applications*, 40(14), 5501-5506.

Tran, N. H., and Hall, K. D. (2007). "Development and influence of statewide axle load spectra on flexible pavement performance." *Transportation Research Record*, 2037(1), 106-114.

Van Gerven, M., and Bohte, S. (2017). "Artificial neural networks as models of neural information processing." Frontiers Media SA, 114.

Velliangiri, S., and Alagumuthukrishnan, S. (2019). "A review of dimensionality reduction techniques for efficient computation." *Procedia Computer Science*, 165, 104-111.

Vlahogianni, E. I., and Karlaftis, M. G. (2013). "Testing and comparing neural network and statistical approaches for predicting transportation time series." *Transportation research record*, 2399(1), 9-22.

Wada, K. (2020). "Outliers in official statistics." *Japanese Journal of Statistics and Data Science*, 3(2), 669-691.

Walker, D., and Cebon, D. "The metamorphosis of LTPP traffic data." *Proc., 6th International Conference on Weigh-In-Motion (ICWIM 6) International Society for Weigh-In-MotionInstitut*

Francais des Sciences et Technologies des Transports, de l'Aménagement et des Réseaux (IFSTARR) International Transport ForumForum of European National Highway Research Laboratories (FEHRL) Transportation Research BoardFederal Highway Administration.

Wang, H., Bah, M. J., and Hammad, M. (2019). "Progress in outlier detection techniques: A survey." *Ieee Access*, 7, 107964-108000.

Wang, H., Zhao, J., and Wang, Z. "Impact of overweight traffic on pavement life using weigh-in-motion data and mechanistic-empirical pavement analysis." *Proc., 9th International Conference on Managing Pavement Assets*.

Wang, Y., and Nihan, N. L. "Dynamic estimation of freeway large-truck volumes based on single-loop measurements." *Proc., Intelligent Transportation Systems*, Taylor & Francis, 133-141.

Wright, L. G., Onodera, T., Stein, M. M., Wang, T., Schachter, D. T., Hu, Z., and McMahon, P. L. (2022). "Deep physical neural networks trained with backpropagation." *Nature*, 601(7894), 549-555.

Xue, W., Wang, D., and Wang, L. (2012). "A review and perspective about pavement monitoring." *Int. J. Pavement Res. Technol*, 5(5), 295-302.

Xue, W., Weaver, E., Wang, L., and Wang, Y. (2016). "Influence of tyre inflation pressure on measured pavement strain responses and predicted distresses." *Road Materials and Pavement Design*, 17(2), 328-344.

Yamany, M. S., Saeed, T. U., Volovski, M., and Ahmed, A. (2020). "Characterizing the performance of interstate flexible pavements using artificial neural networks and random parameters regression." *Journal of Infrastructure Systems*, 26(2), 04020010.

Yannis, G., and Antoniou, C. (2005). "Integration of weigh-in-motion technologies in road infrastructure management." *ITE Journal*, 75(1), 39-43.

Yeganeh, S. F., Mahmoudzadeh, A., Azizpour, M. A., and Golroo, A. (2019). "Validation of smartphone based pavement roughness measures." *arXiv preprint arXiv:1902.10699*.

Yegnanarayana, B. (2009). *Artificial neural networks*, PHI Learning Pvt. Ltd.

You, L., Yan, K., and Liu, N. (2020). "Assessing artificial neural network performance for predicting interlayer conditions and layer modulus of multi-layered flexible pavement." *Frontiers of Structural and Civil Engineering*, 14(2), 487-500.

Yuan, Q., Thangali, A., Ablavsky, V., and Sclaroff, S. (2010). "Learning a family of detectors via multiplicative kernels." *IEEE transactions on pattern analysis and machine intelligence*, 33(3), 514-530.

Zeiada, W., Dabous, S. A., Hamad, K., Al-Ruzouq, R., and Khalil, M. A. (2020). "Machine learning for pavement performance modelling in warm climate regions." *Arabian Journal for Science and Engineering*, 1-19.

Zhang, G., Patuwo, B. E., and Hu, M. Y. (1998). "Forecasting with artificial neural networks:: The state of the art." *International journal of forecasting*, 14(1), 35-62.

Zhang, L., Haas, C., and Tighe, S. L. (2007). "Evaluating weigh-in-motion sensing technology for traffic data collection." *Annual Conference of the Transportation Association of Canada*, 1-17.

Zhang, W., Suo, C., and Wang, Q. (2008). "A novel sensor system for measuring wheel loads of vehicles on highways." *Sensors*, 8(12), 7671-7689.

Zhang, Y., Lu, H., and Qu, W. (2020). "Geographical detection of traffic accidents spatial stratified heterogeneity and influence factors." *International journal of environmental research and public health*, 17(2), 572.

Zhao, J., and Tabatabai, H. (2012). "Evaluation of a permit vehicle model using weigh-in-motion truck records." *Journal of Bridge Engineering*, 17(2), 389-392.

Zhao, Y., Nasrullah, Z., and Li, Z. (2019). "Pyod: A python toolbox for scalable outlier detection." *arXiv preprint arXiv:1901.01588*.

Ziari, H., Maghrebi, M., Ayoubinejad, J., and Waller, S. T. (2016). "Prediction of pavement performance: application of support vector regression with different kernels." *Transportation Research Record*, 2589(1), 135-145.

Ziari, H., Sobhani, J., Ayoubinejad, J., and Hartmann, T. (2016b). "Prediction of IRI in short and long terms for flexible pavements: ANN and GMDH methods." *International journal of pavement engineering*, 17(9), 776-788.