The American University in Cairo

Graduate Studies

Prediction of Distresses in Pavement Networks: A Machine Learning Approach

A THESIS SUBMITTED BY

Mahmoud Mostafa Kotb

TO THE

Construction Engineering Department

SUPERVISED BY

Dr. Maram Saudy

January 3rd, 2024

in partial fulfillment of the requirements for the degree of Master of Science in Construction Management

Declaration of Authorship

I, Mahmoud Kotb, declare that this thesis titled, "***Prediction of Distresses in Pavement Networks: A Machine Learning Approach***" and the work presented in it are my own. I confirm that:

- This work was done wholly or mainly while in candidature for a research degree at this University.
- Where any part of this thesis has previously been submitted for a degree or any other qualification at this University or any other institution, this has been clearly stated.
- Where I have consulted the published work of others, this is always clearly attributed.
- Where I have quoted from the work of others, the source is always given. With the exception of such quotations, this thesis is entirely my own work.
- I have acknowledged all main sources of help.
- Where the thesis is based on work done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed myself.

Signed:

____

Date:

____

Abstract

The quality of pavement networks is greatly affected by different distresses. These distresses appear in many forms, such as cracking, potholes, rutting and different types of deformation. As a result, to ensure effective pavement management, accurate modeling of these different distresses has become essential. Moreover, machine learning models have shown great potential in modeling pavement performance in recent years. The objective of this research is to develop machine learning models for modeling key parameters of pavement distress, specifically the International Roughness Index (IRI), fatigue and longitudinal cracking. Data for this investigation were extracted from the Long-Term Pavement Performance (LTPP) database, with a focus on areas exhibiting environmental conditions similar to those in Egypt. By doing so, the models would be applicable to Egyptian settings. The dataset comprised of 8537 datapoints on 221 different pavement sections. The variables collected include IRI, temperature, precipitation, Equivalent Single Axle Loads (ESALs), pavement age, time since last maintenance, asphalt concrete layer thickness, average asphalt content, bulk specific gravity, granular base thickness, percentage of fatigue cracking, and percentage of longitudinal cracking.

Six machine learning algorithms were used for modeling each output variable: XGBoost, Random Forest, K-Nearest Neighbors (KNN), Bayesian Regression, Ridge Regression, and Decision Trees. Model performance was assessed using Mean Absolute Error (MAE) and R^2 as evaluation metrics. Comparative analysis revealed that the XGBoost algorithm demonstrated superior performance in modeling all three output variables. The results showed a MAE of 0.17 and R^2 of 0.729 for modeling IRI. For modeling fatigue cracking and longitudinal cracking, the model produced a MAE of 4.92% and 2.96%, respectively, with an R^2 of 0.672 and 0.692 respectively.

The findings are significant for many reasons. Firstly, they offer a framework for modeling pavement distress parameters, which is crucial for effective pavement management and maintenance strategies. Secondly, the study confirms the efficacy of machine learning algorithms in modeling pavement performance indicators, especially when using ensemble models. Lastly, the exceptional performance of the XGBoost algorithm indicates its reliability as a tool for both future research and practical applications in pavement management. Importantly, the models are tailored to be applicable in Egypt, providing a data-driven approach to improve the quality of road infrastructure in the region.

Acknowledgements

I express my profound gratitude to Allah for the strength and perseverance bestowed upon me during this academic journey. I extend my sincere appreciation to my advisor, Dr. Maram Saudy, for her invaluable guidance, expertise in pavement management, and unwavering support throughout the course of our research. Special thanks are due to Dr. Ossama Hosny, Chair of the Construction Engineering Graduate Program, for his support and encouragement from the very beginning of my master's studies, providing both inspiration and valuable insights that have been instrumental to my academic growth. I am also grateful to the members of the examining committee, whose insightful comments and suggestions significantly contributed to the refinement of this thesis. My heartfelt thanks go to my friends and colleagues for their support and encouragement. Lastly, I would like to thank my family for their endless love, patience, and understanding, which were the bedrock of my resilience and success.

Contents

List of Figures/Images/Sound Files

List of Tables

List of Abbreviations

Name	Abbreviation
Adaptive Network-Based Fuzzy Interference System	ANFIS
American Association of State Highway and Transportation Officials	AASHTO
Area under receiver operating characteristics curve	AUC-ROC
Artificial Neural Networks	ANN
Asphalt Concrete	AC
Average Annual Daily Traffic	AADT
Average Annual Daily Truck Traffic	AADTT
Centered Mean Square Difference	CMSD
Centered Root Mean Square Difference	CRMSD
Composite	COM
Department of Transportation	DOT
Department of Transportation and Development	DOTD
Extreme Gradient Boosting	XGBoost
Federal Highway Administration	FHWA
Firefly Algorithm	FA
Genetic Algorithm	GA
Group Method of Data Handling	GMDH
International Roughness Index	IRI
K-Nearest Neighbors	KNN
Locally Weighted Polynomials	LWP
Long Short-Term Memory	LSTM
Long-Term Pavement Performance	LTPP
Mean Absolute Error	MAE
Mean Bias	MB
Mean Error	ME
Mean Normalized Bias	MNB
Mean Squared Error	MSE
Ministry of Transportation and Infrastructure	MoIT
Multiple Linear Regression	MLR
Particle Swarm Optimization	PSO
Pavement Condition Index	PCI
Pavement Management System	PMS
Portland Cement Concrete	PCC
Road Management System	RMS
Root Mean Squared Error	RMSE
Support Vector Machines	SVM
Support Vector Regression	SVR

Chapter 1

Introduction

1.1 Research Context

1.1.1 Importance of Pavement Networks

Pavement networks are considered one of the most important types of infrastructure, affecting multiple dimensions of society, including economics, social welfare, and environmental sustainability. This includes a 31.5% investment in pavement structures compared to other infrastructure investments in the United States of America, according to a report by the Federal Highway Administration (FHWA) (Smith, 1994). As forms for transportation, these networks serve critical roles in linking markets, influencing trade, fostering economic union, and contributing to overall economic progress (Banister & Berechman, 2001). In addition to this, the construction and upkeep of such networks generate diverse employment opportunities, thereby impacting labor markets both directly and indirectly (Ke et al., 2020).

With regard to social benefits, well-constructed and adequately maintained pavement networks enhance the availability and accessibility of essential services. For instance, they facilitate more efficient access to healthcare facilities, educational institutions, and emergency services, which in turn has a direct positive impact on the welfare of society (Litman, 2017). Environmentally, the quality of pavement networks is inversely correlated with fuel consumption and vehicular emissions, substantiating their environmental importance (Streimikiene et al., 2013). Safety considerations are also noticeable; superior pavement quality is associated with a reduced incidence of road accidents (Chan et al., 2010).

In addition to the direct benefits, pavement networks exert influence on several secondary sectors. One example relates to public health; high-quality networks encourage active transportation modes like cycling and walking, thereby promoting healthier lifestyles (Pucher & Dijkstra, 2003). Effective management of these networks constitutes a valuable asset for any country. Infrastructure asset management strategies, encompassing consistent monitoring and maintenance, can lead to substantial monetary savings by obviating the need for expensive rehabilitation projects (Burningham & Stankevich, 2005).

Lastly, pavement networks hold implications for the tourism industry. Adequate road infrastructure increases the likelihood of attracting tourists, thereby improving the economic state of a nation (Alkheder, 2015). Furthermore, the resilience of pavement networks is particularly crucial in locations that are susceptible to natural disasters, serving as critical evacuation routes and playing an instrumental role in the logistics of relief and recovery operations (Boakye et al., 2022).

1.1.2 Current State of Pavement Management

Over the years, subjective evaluation models, as well as reactive measures were the main tools of infrastructure asset management. That is specifically true when it comes to pavement deterioration management. These methods are advantageous in terms of simplicity; however, their disadvantages include a lack of quantitative accuracy, reproducibility, and scalability. Disadvantages also include subjectivity and human error, which is especially included when it comes to manual surveys and visual inspections. These errors might result in inconsistent data leading to ineffective management decisions (Li et al., 1997).

Over the years, computer models have been used to create more rigorous, data-driven methodologies in infrastructure asset management. In the field of pavement management, the international roughness index (IRI) has been a commonly used metric to evaluate pavement performance and ride quality. The reason for that is because IRI can provide an objective and quantitative means of evaluation. This coupled with data analytics resulted in an evolutionary shift in pavement management. Moreover, machine learning models made it possible to transition from reactive to proactive infrastructure management strategies (Hanandeh, 2022). These machine learning models have proved to produce predictions on the rate of pavement deterioration influenced by different environmental and load conditions by their ability to analyze large datasets. This in turn, would result in both financial and sustainable effectiveness in pavement networks. Another benefit of the usage of machine learning models in pavement management is the resulting proper resource allocation and preventative maintenance strategies (Bashar & Torres-Machi, 2021).

However, there are different aspects that should be addressed due to this transition. On one hand, these models will provide a reliable and effective framework for pavement management due to their predictive modeling capability. However, further investigation is required on means to incorporate these techniques into the current practice as well as re-evaluate the current models. Moreover, due to their link to computer sciences, this transition would also require the need to integrate between the conventional engineering fields and computer science. This in return, would produce a wider range of research opportunities with the goal of developing more advanced prediction models to be used in infrastructure management (Justo-Silva et al., 2021).

Even though predictive modeling has advanced over the years, there are still some issues that should be addressed. These issues include the need for consistent data collection procedures, the need for large datasets in order to train prediction models, and the need to validate the predictive capabilities of these models against actual scenarios (Papadimitroulas et al., 2021). Another potential issue is the potential of cybersecurity issues associated with the use of digital, datacentric strategies. This issue should be addressed to maintain the integrity and safety of different infrastructure management systems (Chen & Hoang, 2012).

1.1.3 Pavement Condition

There are many reasons pavement integrity is of a great priority; this includes safety and economic productivity. That is why attention to pavement condition is an area of great concern. This is especially important when it comes to load-related distresses such as fatigue/alligator cracking and longitudinal cracking. Moreover, a significant variable used to represent pavement conditions is IRI. It was introduced in order to provide a standardized quantitative approach for evaluating the accumulated vehicle suspension motion to distance traveled (Sayers et al., 1986).

Alligator cracking, which is also referred to as fatigue or flexural fatigue cracking is one of the most common types of load related cracks. It usually appears as interconnected cracks in a pattern similar to an alligator's skin. Due to its link to traffic loads, it is usually directed along the direction of loading. The main reason this type of cracking occurs is because of recurrent or cyclical loading on pavement. Over time, these cracks deteriorate the tensile strain at the asphalt layer's base. It is very common to see this type of crack in parking spaces and narrow lanes which are subjected to constant loading and unloading. Fig. 1 shows an example of a pavement section with fatigue cracking (Cong et al., 2017).

Figure 1: Fatigue/Alligator Cracks (Ahmad & Khawaja, 2018)

Another type of load related crack is the wheel path cracking. This type, however, is produced in areas of pavement sections which are subjected to direct tire pressure. Wheel path cracking often appears as longitudinal cracks and usually indicates structural collapse in pavement. It is usually the result of strong traffic loads and insufficient drainage (Mulungye et al., 2007). Fig. 2 shows a pavement section with wheel path cracking.

Figure 2: Wheel Path Cracking (Ahmad & Khawaja, 2018)

1.1.4 The Use of Machine Learning in Predicting Pavement Condition

Historically, traditional approaches such as visual inspections, manual data collection, and empirical models have been used to forecast and evaluate specific types of pavement distresses such as cracking (Zhang et al., 2022). Even though these equations and models have helped shed light on the mechanisms of these distresses, they usually require a lot of input variables. Moreover, these parameters did not only require a lot of laborbut also may result in human error data. (Majidifard et al., 2020).

On the other hand, machine learning models have been shown to be a promising alternative to model pavement distresses. Unlike conventional approaches which might not detect the complex, non-linear interactions between different variables that affect pavement deterioration, machine learning algorithms have been able to tackle this issue. This is because of their fundamentally data-driven nature. This is specifically true when it comes to predicting fatigue cracking and longitudinal cracking due to them being influenced by different variables such as material characteristics, traffic loading, and environmental conditions (Nguyen et al., 2019).

Furthermore, machine learning models can produce more accurate, timely, and comprehensive distress predictions due to their ability to integrate multidimensional data sets. These data sets include sensor network outputs, traffic monitoring systems, and meteorological data. This, in return, would result in optimized resource allocation for maintenance and rehabilitation which would ultimately increase pavement network service life (Tantalaki et al., 2019).

1.1.5 Problem Statement

When addressing the potential of modeling pavement deterioration and IRI, there are several challenges that are being faced. The first challenge is linked to the standardization and quality of the collected data. Due to the different data collection methods used by different entities around the world, inconsistent datasets are commonly produced. As a result, it becomes challenging to develop models that can be used on a global scale (Zimmerman et al., 2010). Moreover, models can be subjected to overfitting and reduced generalizability resulting from irregular data collection. This has a negative impact of the accuracy of the predicted results (Cano-Ortiz et al., 2022).

Another challenge related to modeling pavement deterioration is related to computational complexity. As datasets increase, the need for large computing resources for the machine learning models also increases. This might result in scalability problems along with inefficiencies (Nguyen et al., 2019). Another aspect that results in demanding computational needs is that pavement conditions are affected by many factors such as environmental factors, traffic related loads, and maintenance history. These factors must be incorporated into the model to ensure higher accuracy (Niazi et al., 2017). As a result, incorporating real scenarios into these machine learning models requires a balance of both computational efficiency and robustness, which is sometimes difficult to acquire (Fang et al., 2020).

To conclude, there are different challenges related to modeling pavement deterioration and IRI. These challenges can be categorized into issue related to computational complexity and data quality. However, this leaves room for potential research to be done in this area. With the involvement of different disciplines in this area, like computer science and civil engineering, advance prediction models can be developed to tackle these challenges. These challenges inspired the author to use different machine learning algorithms to model and predict various pavement performance indicators. Then, compare the accuracy and robustness of the various applied machine learning techniques to select the most promising one.

1.2 Research Objectives

The main objectives of this research are:

1. To investigate the machine learning applications, in terms of predicting pavement condition, for proper management of the pavement through a comprehensive review of the literature

2. To specify the main factors/variables including environmental conditions, pavement age, maintenance history, pavement material characteristics, and traffic loading, that influence the pavement performance prediction.

3. To construct and provide a systematic approach for extracting and organizing different variables from large datasets such as the LTPP in order to develop robust data-driven models.

4. To select the most promising machine learning algorithm in terms of accuracy and robustness to model and predict pavement condition in terms of IRI and load related cracks.

1.3 Research Framework

To conduct this research, the methodology has been depicted in Figure 3 below.

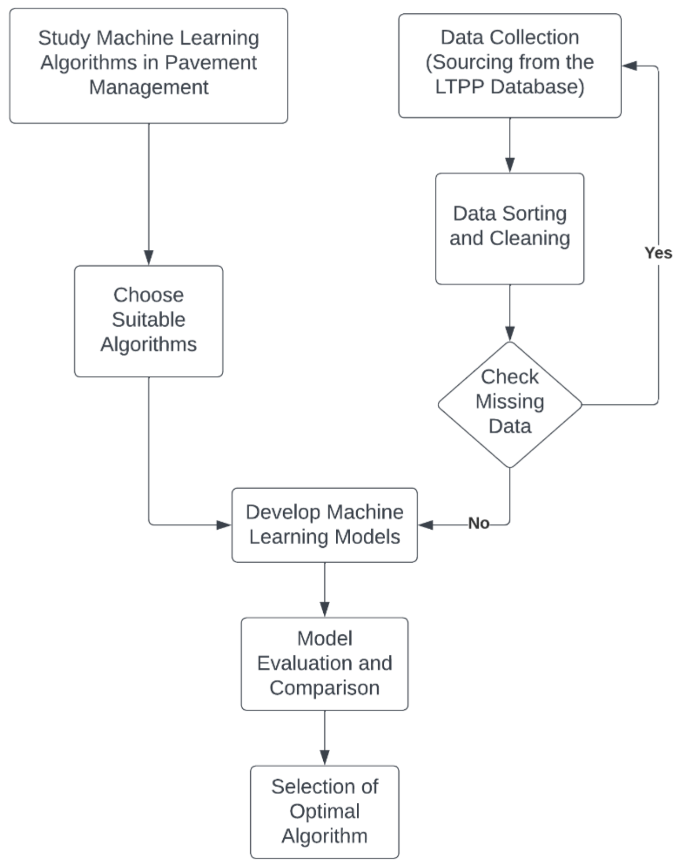


Figure 3: Research Methodology Flowchart

It is clear from Figure 3 that the research comprised three different phases. The first phase is data collection. The second phase is applying different ML algorithms, recommended on a comprehensive review of the current practices from literature, on the previously collected data. The third phase is assessing the performance of the different applied ML algorithms to select the most promising technique.

1.3.1 Data Collection

The effectiveness of any machine learning model is highly dependent on the dataset used for developing the model. This is especially true when it comes to modeling pavement conditions/distresses. For this research, the Long-Term Pavement Performance (LTPP) database will be used as the main source of data. The LTPP, which was developed by FHWA, contains an immense collection of data regarding different aspects of pavement sections around the United States of America and some areas in Canada. These aspects include measurements of pavement performance indicators, traffic loads, environmental conditions and more.

1.3.2 Sourcing from the LTPP Database

The LTPP database is organized in the form of different tables with four main categories: Pavement Structure and Construction, Climate, Traffic, and Performance. For this study, the extracted tables used will provide information specific to pavement age, maintenance and rehabilitation history, environmental conditions, such as temperature and precipitation, traffic loading, IRI, cracking conditions, and pavement material characteristics.

1.3.3 Aligning with Egyptian Conditions

Even though the LTPP database offers data regarding the different factors affecting pavement conditions in the United States of America and Canada, this work aims to develop a model tailored and applicable to conditions found in Egypt. As a result, the extracted data from the LTPP database will be of areas that reflect conditions similar to those found in Egypt, such as high temperatures and dry weather. It will be assumed that the traffic loading and pavement materiality is similar in pavement sections extracted from the LTPP database and those in Egypt.

1.3.4 Ensuring Model's Adaptability

The data collection process was a repetitive process, as during the data collection and model training, some data gaps appeared and features which are specific to Egyptian conditions were sometimes overlooked. As a result, this was corrected by revisiting the LTPP and updating the collected dataset. This iterative process improved the model significantly and made it produce more accurate results.

1.3.5 Selection of Machine Learning Algorithms

There are many machine learning algorithms that have been used to model pavement performance and pavement distresses. This research proposes using six different machine learning algorithms to model IRI and both fatigue and longitudinal cracking: Extreme Gradient Boosting (XGBoost), Random Forest, Ridge Regression, K-Nearest Neighbors (KNN), Decision Tree Regression, and Bayesian Regression.

The reason for choosing several machine learning algorithms is to conduct a thorough study on the performance of these models and to cross-verify and validate their results.

As the models are developed, their results will be compared. The comparative analysis conducted will be done using several statistical tests such as the accuracy, mean square error and the mean absolute error. This will be done to determine the efficiency of each model and determine the model which produces the most accurate results related to pavement performance and cracking.

1.3.6 Model Development and Training

The development and training stage is one of most important stages in developing a reliable machine learning model. When developing the six mentioned machine learning models, this step will be conducted after the comprehensive data collection stage.

The first step will be to train each model separately. This will be done by providing historical data of different pavement sections regarding IRI, current condition of both wheel path and fatigue cracking, environmental conditions, traffic loading, age, maintenance history, and different material characteristics. The models will then start to "learn" the patterns and relationships between different variables which will help in predicting the condition of IRI and pavement cracking. However, a certain percentage of the dataset will not be provided in the training stage and will be used to test the model after it has been trained.

The final step will be to determine which of the six models has produced the most reliable results. This will be done by comparing the results of the testing stage by using statistical measurements. These measures include mean absolute errors, and R^2. The main goal is to focus on the best performing model, in order to ensure reliable results when modeling pavement IRI and load-related cracks.

Chapter 2

Literature Review

2.1 Current Prediction Models

2.1.1 Background and Importance of Pavement Condition Modelling

One of the most important aspects of managing transportation infrastructure is predicting pavement deterioration and distresses. Reliable forecasts allow for efficient planning of rehabilitation and maintenance strategies, lowering the overall lifecycle costs of pavement assets and improving the road user experience (Abu Samra et al., 2017). Planning and managing transportation infrastructure requires a thorough assessment of the evaluation and monitoring of pavement condition. Visual inspections and manual collection of data are commonly used with conventional techniques for pavement condition monitoring; however, those techniques are labor-extensive and sensitive to human errors (Majidifard et al., 2020). Pavement condition prediction models form a major part of any pavement management system (PMS) and are utilized by many transportation agencies worldwide (Haider et al., 2011).

Traditional pavement condition prediction models, like the American Association of State Highway and Transportation Officials (AASHTO) Road Test model, are generally based on empirical data (Hu et al., 2022). Although those models have been essential in improving pavement management strategies, they have got many drawbacks. The used models won't completely capture the complex relationships between many influencing aspects such as traffic loads, weather conditions, and pavement material characteristics, even though those elements tend to have a great impact on pavement overall performance. These models frequently require human involvement, and they generally have issues in dealing with large datasets and non-linear relationships (Rezapour et al., 2022).

Pavement deterioration modeling became developed over many years from simple methods to more complex models. These models may be divided into parametric and non-parametric models. While parametric models depend upon explanatory variables to forecast pavement deterioration, non-parametric techniques depend on machine learning and artificial intelligence. Models which might be parametric or non-parametric can be stochastic (based totally on probability) or deterministic, primarily based on mechanistic or mechanistic-empirical data (Haas et al., 2015).

The Markov chain is one of the modern prediction models for pavement deterioration. On both the project and program levels, it allows the prediction of future conditions given the current condition without the requirement for substantial data sets (Swei et al., 2019). In order to predict the future deterioration of pavements using a probabilistic method, it is assumed that the existing condition of the pavement is capable of comprehensively capturing the influence of all-important elements that impact the pavement condition

(Rose et al., 2016).

For their potential to simulate the variable nature of pavement deterioration over the years, stochastic models have obtained a lot of interest. Although Markov Chain models are regularly used to forecast pavement conditions, other stochastic models have also been useful. To realize the complicated interactions between the different elements which impact pavement overall performance, methods like Bayesian networks and Monte Carlo simulations have been utilized. For instance, Bayesian Networks use probabilistic graphical techniques to depict the conditional relationships amongst many variables, giving a comprehensive view of the variables affecting pavement distress. On the other hand, Monte Carlo simulations use random sampling techniques to forecast different results primarily based on the probability distributions of the input variables, offering an outline of probable pavement conditions (Mills et al., 2012).

The complexity and dynamic nature of pavement deterioration are better captured through those novel stochastic models compared to conventional techniques. They enable the incorporation of a wider variety of variables, ranging from material characteristics to environmental conditions, generating more thorough and accurate forecasts (Swei et al., 2019). As a result of the increasing computational power and data availability, researchers and practitioners now have more access and capability to use these models.

On the other hand, machine learning algorithms have shown promise in providing a more reliable method for modeling pavement conditions when compared to traditional and stochastic models. Different machine learning algorithms such as Random Forests and Support Vector Machines (SVMs) have been able to deal with large, multidimensional datasets and are able to perform more accurately than the other techniques. The reason for that is because machine learning algorithms can capture complex, nonlinear relationships between variables, instead of stochastic models, which mostly depend on probabilistic results, or traditional models, which rely on linear assumptions (Nagalla et al., 2017; Zhang et al., 2023). To conclude, machine learning algorithms greatly surpass the capabilities of both traditional and stochastic models when modeling pavement deterioration. This offers a more robust, and adaptive framework for modeling pavement deterioration.

2.1.2 Importance of Accurate Prediction Models

Deterioration of pavement networks naturally has a substantial cost and safety impact. So, the development of accurate models for pavement deterioration is a very important area of research. As mentioned earlier, conventional techniques' reliance on visual inspections and manual surveys add to the techniques' limitations, and subjectivity to human error. This is also associated with variability in the data collection stage (Li et al., 1997). Empirical techniques are also used to model pavement conditions and distresses. These models include the Mechanistic-Empirical Pavement Design Guide (Li et al., 2009). However, the geographic and climatic specificity of those models limits their applicability to different locations such as Egypt.

Premature pavement collapse from faulty prediction models can lead to increased repair costs and endanger road safety along with leading to inefficient resource allocation for maintenance. For instance, failing to account for significant factors like climate, current IRI, and pavement material characteristics may result in large wastes of funds (Hosseini & Smadi, 2021). Therefore, while considering the range of variables impacting pavement distress, starting from age and maintenance histories along with the effect of traffic loads, the need for strong, dependable prediction models is emphasized.

Accordingly, due to their proven ability to manage complex, multidimensional data, machine learning offers a distinctive approach for developing prediction models with higher accuracy.

2.1.3 Machine Learning Algorithms to Predict Pavement Condition

The objective of conducting this literature review is to thoroughly study existing work done on developing accurate and reliable prediction models which aim to forecast pavement conditions. To be specific, the review focuses on studying both algorithms and methodologies which have shown potential in producing prediction models with high predictive accuracy. The focus will be on machine learning algorithms that have been developed to predict both IRI and the state of both fatigue and longitudinal cracks in pavement sections.

To thoroughly study the relevant information related to the topic firstly, the applications of machine learning algorithms in relation to monitoring the condition of different types of infrastructure was investigated. The evolution of algorithms used in infrastructure, and transitions into the specific machine learning algorithms, specifically used in monitoring and forecasting pavement conditions, was thoroughly reviewed. Also, the performance analysis and adaptability of these machine learning algorithms were presented. Secondly, the different ways to evaluate prediction models, by starting from the traditional accuracy, precision, and recall metrics, and then discussing more advanced metrics like the F1-Score and AUC-ROC were illustrated. Then, the relevance of these metrices in the context of infrastructure and pavement management was discussed.

Thirdly, the localized conditions that affect pavement quality were addressed by discussing the impact of each condition. The main conditions considered were geographical factors relating to the location of the pavement sections, environmental factors like weather, precipitation, and icing. Lastly were the socio-economic variables which may indirectly influence pavement quality. This was done to tailor the produced prediction models to Egypt. Then, an overview of the commonly used data sources related to pavement networks, emphasizing on the LTPP database, was given.

Finally, a comparative analysis of the different machine learning algorithms used in previous studies was provided. This was followed by outlining the challenges and limitations found in current literature in terms of using machine learning in pavement prediction models. The literature review was concluded by a summary of the key findings in the literature, along with the research gaps.

2.2 Machine Learning Algorithms in Infrastructure Condition Monitoring

2.2.1 Evolution of Algorithms in the Infrastructure Management

The use of machine learning algorithms to create prediction models has been considered to be a significant turning point in infrastructure management. However, that was not always the case. Before using these models, conventional models were commonly used (Zhu & Wang, 2021). Most of these rely on mechanistic-empirical methods, which consider different factors in generating forecasts such as environmental conditions, traffic loading effects and material characteristics (Chen et al., 2004). However, over time, those techniques showed a lack of adaptability to various conditions and a regular need for recalibration. These limitations showed that conventional models, in the long run, could not be relied on (Gardiner et al., 2008).

However, over time, stochastic models started to gain popularity in infrastructure management and specifically in modeling pavement performance. These models were considered to be more reliable than conventional approaches. Stochastic models include Markov Chains and Bayesian networks. By using probabilistic factors, stochastic models showed potential with adapting to complex data and unforeseen conditions (Badr et al., 2021; El-Awady & Ponnambalam, 2021; Weber et al., 2012). Nevertheless, limitations started to appear related to scaling. This was mainly due to the increased complexity of datasets and variables considered when developing stochastic prediction models (Garí et al., 2021).

As a result, researchers started to explore the application of machine learning algorithms in infrastructure management. This was especially doable as computational power advanced. Machine learning algorithms that are commonly used in developing models related to infrastructure management include neural networks, decision trees, KNN, and SVMs (Nagalla et al., 2017; Zhang et al., 2023). Machine learning algorithms greatly surpass the capabilities of both traditional and stochastic models when used in infrastructure management. This results from their ability to capture complex, non-linear relationships between variables along with their ability to adapt to large datasets (Nguyen et al., 2019).

Further research in machine learning algorithms introduced more developed algorithms to develop more accurate models. These algorithms include ensemble models such as random forests and XGBoost. The idea of ensemble models is that they combine more than one base model. Other researchers introduced deep learning models, like convolutional neural networks, which incorporate image-based assessments. By doing so, this allows the addition of a new dimension to be considered in data interpretation (Nhat-Duc & Van-Duc, 2023).

To summarize, machine learning was not always used in infrastructure management. Computational models started with conventional techniques, which were later followed by stochastic models, and lastly transitioned to machine learning models. The main limitations of conventional techniques were their lack of adaptability and the constant need for recalibrating (Zhu & Wang, 2021; Chen et al., 2004; Gardiner et al., 2008). These limitations were tackled by stochastic models such as Markov Chain and Bayesian networks which did so by developing probabilistic forecasts. However, other limitations appeared related to stochastic models which were associated with data complexity and scaling (Badr et al., 2021; El-Awady & Ponnambalam, 2021; Weber et al., 2012; Garí et al., 2021). These issues were resolved with the emergence of machine learning models which were able to capture complex relationships between different variables and handle large datasets (Nagalla et al., 2017; Zhang et al., 2023; Nguyen et al., 2019). However, machine learning algorithms are still improving with the development of ensemble and deep learning algorithms. These algorithms were proved to better interpret large datasets along with improved predictive capabilities (Nhat-Duc & Van-Duc, 2023). This evolution of computational models shows the importance of applying machine learning algorithms in infrastructure management.

2.2.2 Machine Learning Algorithms Specifically Employed in Pavement Monitoring

Over the years, IRI has been commonly used to assess pavement conditions and has been vital in pavement monitoring and planning. Not only is IRI used to represent ride quality and pavement conditions, but since it is considered a function of pavement distresses, it is a very important aspect of pavement management (Elhadidy et al., 2021; Hossain et al., 2019; OBrien et al., 2018).

As a result, many researchers have thrived to develop computational models to better predict pavement IRI, given its importance in pavement management. Traditional regression models were previously used, such as Multiple Linear Regression (MLR), for modeling IRI. This was specifically done by Gharieb and Nishikawa (2021) in their work to predict IRI using data from the Laos Road Management System (RMS) by incorporating an MLR model. Other work was done by Pérez-Acebo et al. (2020) using the same technique, and they were able to achieve a coefficient of determination R^2 of 0.48 for flexible pavements in Spain.

Moreover, advancements in research and machine learning algorithms introduced the usage of ANNs in modeling IRI. Research showed that the results of ANN models yielded more accurate results when compared to traditional regression models. This was shown by Abdelaziz et al. (2020) in their work in modeling pavement IRI with their ANN model yielding an R^2 of 0.75 which was more than the 0.57 produced by the regression model. ANN was also compared to another technique called Group Method of Data Handling (GMDH) and the ANN model gave better results on the long term and short-term predictions (Ziari et al. 2016).

Further advancements introduced other algorithms such as the Adaptive Network-Based Fuzzy Interference System (ANFIS) which has also been used in modeling IRI. Research done by Terzi (2013) showed promising results when using ANFIS to model pavement IRI. This algorithm was further optimized by introducing hybrid models which involved the integration of other algorithms with ANFIS. These integrated algorithms include Genetic Algorithm (GA), Particle Swarm Optimization (PSO), and Firefly Algorithm (FA). Nguyen et al. (2019) integrated these three algorithms with an ANFIS model and concluded that the PSO-ANFIS hybrid model performed better than the other hybrid models.

The creation of hybrid machine learning models for modeling IRI paved the way for conducting comparative studies between the usage of single algorithm models and hybrid models. Mazari and Rodriguez (2016) conducted this comparison using a hybrid GA-ANN model, and comparing its results with an ANFIS model, an ANN model, and a GA model to forecast IRI in flexible pavement. Their results showed that the hybrid GA-ANN's results were more accurate than the other three models. As a result, it can be concluded that the increased applications of machine learning algorithms have led to a significant advancement in using data-driven models for pavement condition assessment (Sholevar et al., 2022; Kheirati and Golroo, 2022).

Moreover, Kumar & Sowmya (2021) highlighted the pros and cons of some of the major algorithms used in developing machine learning models as shown in table 1.

Table 1: Pros and Cons of Solo Models (Kumar & Sowmya, 2021)

Algorithm	Pros	Cons
Ridge Regression	• Reducing variance in models, decreasing overfitting • Improved generalization performance	• Computational intensity • Difficulty interpreting the model
KNN	• Simplicity in implementation • Offer multi-modal classification	• Issues when categorizing unknown records
Decision Trees	• Can be used in either regression or classification problems • Ability to fill missing data with most likely values	• Susceptible to sampling errors • Issues with over-fitting
Bayesian Regression	• Prevents data from overfitting • Ability to deal with various datasets	• Issues with prioritization • Can be computationally intensive

In addition, work using ensemble machine learning models such as Random Forest and XGBoost has been instrumental due to their advantages. One of the advantages of an ensemble machine learning model such as XGBoost is its ability to create no assumptions related to data distribution. This is done by using several decision trees, and as a result, not be affected by multicollinearity (Montomoli et al., 2021). Moreover, Random Forest is considered a significant ensemble model due to its ability to also utilize more than one decision tree, and as a result, is also unaffected by multicollinearity (Fawagreh et al., 2014).

However, one of the major drawbacks of machine learning models is their black-box nature and the difficulty interpreting their results. This issue was tackled by models such as decision trees and random forests which provided better model transparency (Chi et al., 2014; Gong et al., 2018; Piryonesi & El-Diraby, 2020; Piryonesi & El-Diraby, 2018). Another issue with using machine learning algorithms is the fear of overfitting when using a small training set while training the model. These issues leave room for more research and work to be done in developing machine learning models for assessing pavement conditions.

2.3 Evaluation Metrics for Prediction Models

2.3.1 Statistics and Error Metrics

It is well-known that any machine learning prediction model is never completely accurate when it comes to predictive capabilities. This allowance is made to avoid overfitting when training and developing these models. As a result, it has become necessary to quantify the performance of machine learning models to properly evaluate and in most cases, compare the results of more than one model. In this section, several metrics for evaluating model performance will be discussed, along with the means of calculating these metrics. Finally, the section will be concluded by discussing the application of these different statistical measures in modeling pavement performance.

To begin, as done by Plevris et al. (2022), assume p is an N×1 vector representing the predicted values of a certain model, and r to be another N×1 vector representing the true or observed values of a specific quantity that has been measured or calculated N times. N can represent the entire dataset of a subset like the training data or testing data. An independent evaluation can be done by using another separate dataset.

However, in certain situations where variables are categorical, different metrics are used to evaluate results such as precision, recall, confusion matrices, and accuracy. For datasets with positive values only, Plevris et al. (2022) compiled the following metrics and their means of calculation which can be used to evaluate the results of machine learning models.

The Bias Error is defined as the difference between the predicted values and the true or observed values as shown in Eq. 1:

$$e_i = p_i - r_i \tag{1}$$

Where is e_i represent the bias error, the p_i represents the predicted values, and the r_i represents the observed values.

The Mean Bias (MB) is calculated as the average of the individual bias errors and is illustrated in Eq. 2:

$$MB = \bar{e} = \frac{1}{N}\sum_{i=1}^{N} e_i = \frac{1}{N}\sum_{i=1}^{N}(p_i - r_i) = \bar{p} - \bar{r} \tag{2}$$

Where both $\bar{p}$ and $\bar{r}$ represent the averages of p and r, respectively. This can be shown in equations 3 and 4:

$$\bar{p} = \frac{1}{N}\sum_{i=1}^{N} p_i \tag{3}$$

$$\bar{r} = \frac{1}{N}\sum_{i=1}^{N} r_i \tag{4}$$

Next is the Mean Absolute Error (MAE), which is the average of the sum of errors, without considering the signs of each error. It is calculated by adding up the absolute values of the errors measured and then dividing the product by the number of observations. The MAE can be calculated using Eq. 5 as shown below:

$$MAE = \frac{1}{N}\sum_{i=1}^{N} |e_i| = \frac{1}{N}\sum_{i=1}^{N} |p_i - r_i| \tag{5}$$

The Mean Squared Error (MSE) is also commonly used to evaluate the results of regression models. Similar to the MAE, the MSE does not consider the signs of the errors, but this is done by squaring the errors and adding them together. The product is then divided by the number of observations. This metric is calculated using Eq. 6:

$$MSE = \frac{1}{N}\sum_{i=1}^{N}(p_i - r_i)^2 \tag{6}$$

However, one of the issues with using MSE is their sensitivity to outliers. Due to the squaring process, a single outlier could exponentially affect the resulting MSE.

The Root Mean Squared Error (RMSE) is another measure for assessing machine learning model performance. It is calculated by square rooting of the MSE. It can be calculated using Eq. 7:

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(p_i - r_i)^2} \tag{7}$$

The Centered Mean Square Difference (CMSD) can be calculated using Eq. 8:

$$CMSD = \frac{1}{N}\sum_{i=1}^{N}[(p_i - \bar{p}) - (r_i - \bar{r})]^2 \tag{8}$$

The square root of the CMSD, known as the Centered Root Mean Square Difference (CRMSD) can be used to evaluate model performance and is calculated by Eq. 9:

$$CRMSD = \sqrt{CMSD} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}[(p_i - \bar{p}) - (r_i - \bar{r})]^2}$$

(9)

Finally, the Mean Normalized Bias (MNB), which is expressed as a percentage, is calculated as the average of the normalized bias errors. This is given by Eq. 10:

$$MNB = \frac{1}{N}\sum_{i=1}^{N}\frac{(p_i - r_i)}{r_i}$$

(10)

Table 1 summarizes the errors mentioned with their units, ranges, and their values in case of a perfect match.

Table 2: Metrics for Evaluating Model Performance (Plevris et al., 2022)

ID	Metric	Abbreviation	Units	Range	Perfect match value
1	Mean bias	*MB*	Units of *x*, *p*	[-∞,+∞]	0
2	Mean Absolute Error	*MAE*	Units of *x*, *p*	[0, +∞]	0
3	Mean Square Error	*MSE*	Units of *x*, *p*	[0, +∞]	0
4	Root Mean Square Error	*RMSE*	Units of *x*, *p*	[0, +∞]	0
5	Centered Root Mean Square Difference	*CRMSD*	Units of *x*, *p*	[0, +∞]	0
6	Mean Normalized Bias	*MNB*	Unitless	[-1, +∞]	0

2.3.2 Accuracy, Precision and Recall

When assessing the results of prediction models, researchers have commonly used accuracy, precision and recall as different metrics for assessing their results in different fields (Alakus & Turkoglu, 2020; Jishan et al., 2015; Mehdiyev et al., 2016; Zimmermann et al., 2009).

According to Tatbul et al. (2018), precision and recall can be defined using Eq. 11 and Eq. 12, respectively:

$$Precision = TP \div (TP + FP) \tag{11}$$

$$Recall = TP \div (TP + FN) \tag{12}$$

Where *TP, FP,* and *FN* represent true positives, false positives, and false negatives, respectively. Moreover, accuracy of prediction models can be displayed using Eq. 13 as illustrated by Ampomah et al. (2020):

$$Accuracy = (TP + TN) \div (TP + TN + FN + FP) \tag{13}$$

Where *TN* represents true negatives.

To clarify, precision is used to identify the percentage of true positives in the results compared to the sum of all the positives in the data. Recall, however, quantifies the proportion of true negatives that have been correctly identified. Both these metrics are considered complimentary and in some cases are integrated for a more comprehensive evaluation. One type of integrated metric of evaluation is the F_β-Score, where the β represents the relative weight of recall versus precision. These metrics can be used to evaluate the abnormalities found in machine learning models (Tatbul et al., 2018). Moreover, accuracy, which is another metric for evaluating machine learning models, represents the percentage of correct predictions developed by the model when compared to the actual results (Ampomah et al., 2020).

2.3.3 Advanced Metrics: F1-Score, AUC-ROC

Over the years, more advanced means of evaluating machine learning models were developed and widely used in different industries. These metrics include the F1-Scoe and the area under receiver operating characteristics curve (AUC-ROC) (AlSaad et al., 2022; Ampomah et al., 2020; Kumar et al., 2021; Mancini et al., 2020).

The AUC-ROC represents the ability of a model to differentiate between positive and negative outputs. Ideally, if the model can completely make this differentiation, it would yield an AUC-ROC value of 1. Otherwise, if the model cannot make this differentiation whatsoever, it would yield an AUC-ROC value of 0.5. The main benefit of using the AUC-ROC is its capability of evaluating a model in terms of its ability to identify the rate of both true and negative positives (AlSaad et al., 2022; Ampomah et al., 2020).

Moreover, the F1-Score is one of the most used metrics of evaluation when dealing with categorical variables. It is considered the average of precision and recall, while considering false positive and false negative rates. As mentioned earlier, precision is measured as the percentage of true positives compared to the total positives. Ideally, the F1-Score would be 1, whereas the least favorable value of the F1-Score would be 0. By considering both recall and precision, the F1-Score includes the importance of both metrics in a single number (AlSaad et al., 2022; Ampomah et al., 2020).

According to Ampomah et al. (2020) the F1-Score can be calculated using Eq. 14:

$$f1\_Score = \frac{2 \times precision \times recall}{precision + recall} \tag{14}$$

2.3.4 Relevance of Metrics in Pavement Context

Machine learning models are used in many areas in the field of pavement management. Even though modeling pavement performance is one of those fields, others include modeling stiffness of different high-modulus asphalt concretes, pavement temperature prediction, crack severity classification, and pavement roughness (Baldo et al., 2022; Bashar & Torres-Machi, 2022; Liu et al., 2022; Milad et al., 2021).

However, different metrics have been used in different studies. But the result is always an evaluation of the machine learning model's performance. For example, in their work, Torres-Machi (2022), used MAE, RMSE, and R2 to evaluate the performance of their model for estimating pavement roughness. Moreover, Liu et al. (2022) assessed their results using accuracy and F1-Score in modeling asphalt pavement crack severity. Milad et al. (2021) used deep learning to model pavement temperature. To evaluate their results, they used R2, MAE, MSE, and mean absolute percentage error. Additionally, Ali et al. (2022) used RMSE, R2, MAE to evaluate their model which was used to predict the pavement condition index of pavement sections using Fuzzy Logic.

However, research was done on the evaluation metric R^2 to determine the reliability of the metric. Accordingly, Anscombe's quarter was developed as shown in Fig. 4.

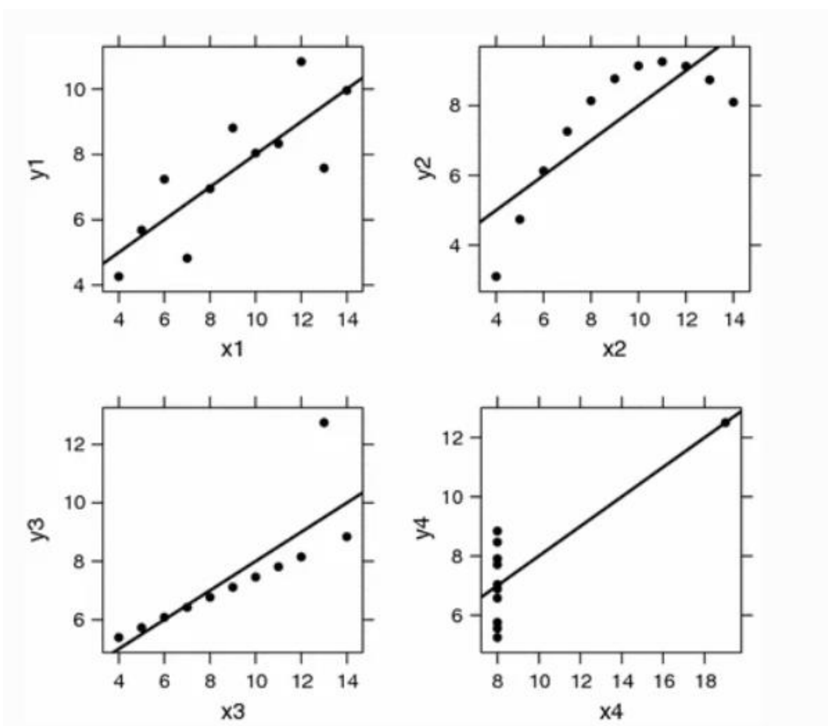


Figure 4: Anscombe's Quartet (Dayal, 2015)

To illustrate, the four graphs shown in Figure 4 have equal R^2, even though the predictive accuracy of each graph is completely different than the others. As a result, it is arguable that R^2 can sometimes be a misleading evaluation metric (Bu & Clemente, 2022; Dayal, 2015; Matejka & Fitzmaurice, 2017).

To summarize, even though there are several machine learning algorithms applied in different areas of pavement management, every machine learning model requires any means of model performance evaluation. Different studies in pavement management have shown the usage of different metrics like MAE, MSE, RMSE, R^2, accuracy, and F1-Score to evaluate model performance. These models include models for pavement performance, stiffness measurement, pavement temperature prediction, and classification of crack severity. When modeling pavement roughness and pavement conditions, the metrics which were commonly used to evaluate the machine learning models are MAE, MSE, RMSE, and R^2. On the other hand, when it comes to evaluating other factors such as pavement crack severity, the F1-Score was used. That is due to its ability to handle categorical data. In order to evaluate the performance of any machine learning model, the choice of evaluation metric is a vital step. Moreover, the usage of these evaluation metrics in a crucial way of comparing the performance of different models and allowing the advancement in developing newer models.

2.4 Factors that Affect Pavement Performance

Traditionally, researchers have focused on very specific factors that affect pavement performance. These factors include pavement material characteristics and the different impacts of traffic loads (Saad et al., 2005). However, it has been increasingly recognized that other factors, such as environmental conditions, soil properties, and age, have a significant impact on pavement conditions (Gupta et al., 2013; Stoner et al., 2019).

Furthermore, Fortney et al. (2022) addressed the different factors that affect pavement quality for airport management systems. In their work, they highlighted that current airport pavement management systems do not account for localized conditions such as climate and traffic conditions when assessing pavement performance. To tackle this issue, they propose a bias-reduced statistical model which incorporates the mentioned localized conditions in order to model pavement deterioration in airports. After validating their model on various climate zones and airport traffic conditions, the authors noted that the effect of environmental conditions had a higher impact on airport pavement deterioration compared to traffic loads. Lastly, they recommend incorporating these local factors into current airport pavement management systems for more accurate and sustainable pavement management.

Additionally, Saha et al. (2012) focused on the impact of climatic factors on the performance of flexible pavement in Canada, as assessed using the mechanistic-empirical pavement design guide. The study employs 206 Canadian climatic files for analysis, comparing specific metrics like the freezing index and frost depth against data from Canadian databases. One key finding is that various performance indicators, such as the total pavement rutting and IRI, are sensitive to climatic changes. In essence, the study aids in the adaptation of the mechanistic-empirical pavement design guide within the Canadian context by highlighting the role of climate in pavement degradation.

One of the factors studied that affect pavement performance is design factors. Zeiada et al. (2020) conducted a study on the effect of design factors on pavement performance in regions with warm climates. Their work involved the usage of ANNs and a forward sequential features selection algorithm. To develop their model, the variables considered were mainly focused on capturing the effect of environmental and structural conditions on the pavement sections. The environmental factors include initial IRI, relative humidity, average wind velocity, albedo, and average emissivity. Moreover, the structural factors were comprised of traffic volume and pavement structural capacity. Their results showed that the seven studied variables have a significant impact on pavement deterioration in warm regions. While comparing the results of the two models, the ANN model outperformed the forward sequential features selection model. The results were also compared to a regression model in terms of accuracy, and the ANN again outperformed. These results indicate that as opposed to what was considered previously in the literature, which was just the traffic and materiality of the pavement, environmental conditions have a significant impact on pavement performance.

As has been demonstrated, when modeling pavement performance the main factors considered when developing computational models have shifted from simple materiality and traffic conditions to incorporating additional factors such as environmental conditions. Recent studies have shown that environmental conditions such as precipitation, climate conditions, humidity, albedo, wind velocity, and emissivity have a significant impact of pavement deterioration (Gupta et al., 2013; Stoner et al., 2019; Fortney et al., 2022; Saha et al., 2012; Zeiada et al., 2020). As mentioned previously, machine learning algorithms have been able to capture the complex relationships between these environmental conditions and pavement performance. This would help in developing more robust and accurate prediction models. To conclude, these variables have been found to be essential in developing any machine learning model for modeling pavement performance, and their complex relationships should be further studied to further refine research in this area.

2.5 Overview of Data Sources for Pavement Monitoring

One of the main factors that influence the success and accuracy of machine learning models is the quality of the data provided to the model. As mentioned earlier, pavement conditions are influenced by both environmental factors and others related to materiality and traffic loads. One of the methods used for data collection is visual inspections and testing done on-site (Ragnoli et al., 2018). However, with the increased number of roads, this method would not be very efficient. With advanced technologies, other techniques like digital cameras, line scan cameras, laser imaging, and terrestrial laser scanners were used in detecting distresses in pavement (Ragnoli et al., 2018). This collected data would be later stored in different data sources in order to be utilized in improving Pavement Management Systems (PMS). This subsection discusses the different data sources used in pavement management, while emphasizing the use of the LTPP given that it is the main data source for this research.

According to Pierce et al. (2013), different transportation agencies around the world regularly collect and store data related to pavement networks in order to be used in pavement monitoring. Each of these agencies collect their own type of data, and at set internals which are also unique to each entity.

The first entity is the British Colombia Ministry of Transportation and Infrastructure (MoTI). They collect data on surface distresses, rut depths, and IRI. The frequency of collecting data varies between 2 to 4 years depending on the location and use of the studied roads. Secondly is Colorado's Department of Transportation (DOT), which collects data on cracking, rut depths, and IRI on an annual basis. This is followed by Idaho's DOT and Indiana's DOT which both separately collect data on surface distresses, rut depths, and IRI once a year. Next is Florida's DOT and Kentucky Transportation which both collect data on surface distresses, faulting, rut depth, and IRI on an annual basis. Moreover, Iowa's DOT collects data on cracking, rut depth, faulting, D-cracking, joints spalling, and IRI once every 2 years. Other highway agencies also collect data regarding pavement distresses. These agencies include Louisiana's Department of Transportation and Development (DOTD), LTPP, Maryland State Highway Administration, Nebraska's Department of Roads, New Mexico's DOT, North Carolina's DOT, Oklahoma's DOT, Oregon's DOT, Pennsylvania's DOT, Viriginia's DOT, and Washington's DOT.

Table 2 illustrate the different entities in the United States of America that collect data regarding pavement distresses and includes the type of data collected along with the frequency of collecting this data.

Table 3: Data Collection Agencies and Frequency of Data Collection (Pierce et al., 2013)

Highway Agency	**Collected Data**	**Frequency of Data Collection**
British Colombia MoTI	Surface distress, rut depth, and IRI	Once every 2 to 4 years depending on the type of road
Colorado DOT	Cracking, rut depth, and IRI	Yearly
Florida DOT	Surface distress, faulting, rut depth, and IRI	Yearly
Idaho DOT	Surface distress, rut depth, and IRI	Yearly
Indiana DOT	Surface distress, rut depth, and IRI	Yearly
Iowa DOT	Cracking, rut depth, faulting, D-cracking, joints spalling, and IRI	Every 2 years
Kentucky Transportation Cabinet	Surface distress, faulting, rut depth, and IRI	Yearly
Louisiana DOTD	Cracking, patching, faulting, rut depth, and IRI	Yearly
LTPP	Surface distress, faulting, rut depth, and longitudinal profile	Every 2 years
Maryland State Highway Administration	Cracking, rut depth, and IRI	Yearly
Nebraska Department of Roads	Surface distress, faulting, rut depth, and IRI	Yearly
New Mexico DOT	Surface distress and faulting	Yearly
North Carolina DOT	Surface distress, faulting, rut depth, and IRI	Yearly
Oklahoma DOT	Surface distress, faulting, rut depth, and IRI	Once every 1 to 2 years depending on type of system
Oregon DOT	Surface distress, faulting, rut depth, and IRI Surface distress, faulting, rut depth, and IRI	Yearly
Pennsylvania DOT	Surface distress, faulting, rut depth, and IRI	Yearly
Virginia DOT	Surface distress, rut depth, and IRI	Yearly

Washington DOT	Surface distress, faulting, rut depth, and IRI	Yearly

However, in the area of pavement management and modeling pavement conditions, the LTPP database has been commonly utilized in the literate as a robust data source. The LTPP has been used in not only modeling pavement deterioration, but also in pavement design, detecting flushing of thin-sprayed seal pavements, and developing models for prioritizing PMS's on a network level (Ekramnia & Nasimifar, 2022; García-Segura et al., 2020; Hall et al., 2011; Kodippily et al., 2012; Mamlouk & Zapata, 2010; Radwan et al., 2020). This is why the LTPP was the source of data used for this research.

The LTPP database, which is run by the FHWA, contains data on over 2500 pavement sections across USA and Canada. Data includes location of sections, environmental conditions such as precipitation, temperature, humidity, and wind, pavement layer characteristics, including materials, thicknesses, and test results. It also includes historical data on maintenance and rehabilitation of different pavement sections. Moreover, the LTPP database includes traffic information such as average annual daily traffic, average annual daily truck traffic, traffic loading conditions, vehicle classifications and other.

2.6 Comparative Analysis of Machine Learning Algorithms in Previous Studies

As mentioned earlier, over the past years, machine learning algorithms have become an instrumental tool in modeling pavement performance. These algorithms include ANNs, decision trees, KNNs, MLR, ANFIS, PFO, FA, and random forests. In this section, research done on modeling pavement performance will be analyzed. The models will be compared in terms of machine learning algorithm used, variables considered, pavement performance index produced, number of data points used for training the model, data source used, and the evaluation metrics used for each model.

Elhadidy et al. (2021) utilized a regression model as a machine learning algorithm to study the relationship between the Pavement Condition Index (PCI) and IRI of different pavement sections. To do so, the variables considered for the study were the IRI, and distress related data like the area of fatigue cracking, edge cracking, block cracking, longitudinal cracking, transverse cracking, patching, potholes, shoving, bleeding, rutting, and reveling. A total of 12,744 data points was considered. The data source for this research was the LTPP database, and the evaluation metric used to evaluate the performance of the model was the R^2.

Hossain et al. (2019) worked on modeling IRI for flexible pavement using an ANN model. The variables considered for their research were annual average temperature, freezing index, maximum and minimum humidity, precipitation, average annual daily traffic (AADT), and average annual daily truck traffic (AADTT). These variables were chosen to capture the effect of both traffic and climate conditions. The ANN model was represented using a 7-9-9-1 model, as shown in Fig. 5, where the input layer for the ANN model consists of the 7 variables mentioned. This was followed by 2 hidden layers with

9 neurons each, giving the output layer which is the predicted IRI. To conduct their research, the data source was the LTPP, where they collected a total number of 200 data points. Finally, in order to evaluate their results, they used the RMSE as the main metric for evaluating the results of the model.

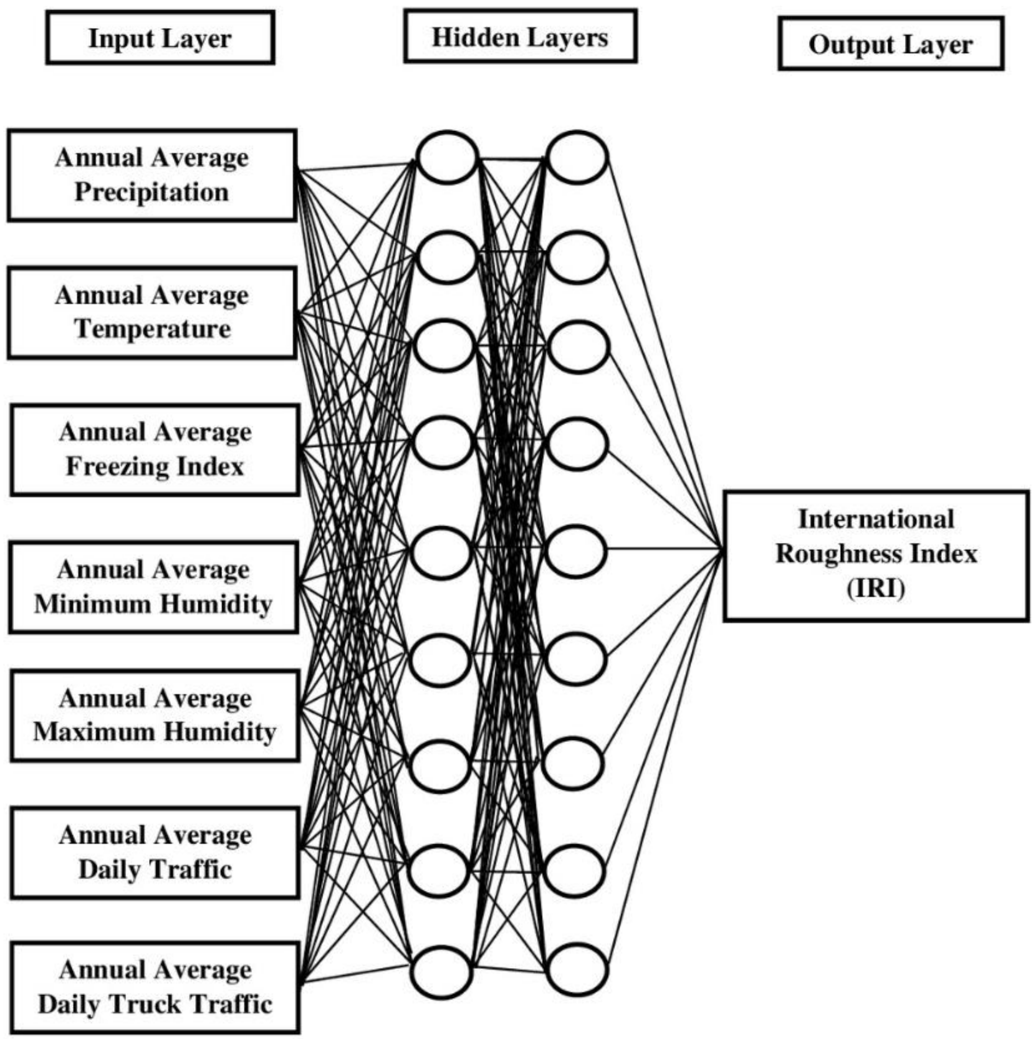


Figure 5: 7-9-9-1 ANN Model (Hossain et al., 2019)

Other researchers were focused on developing different machine learning models and conducting a comparative analysis to identify the better performing algorithm. Kaloop et al. (2023) did this by developing four separate prediction models for modeling IRI. The algorithms used for this research were Gaussian Process Regression (GPR) model, Locally Weighted Polynomials (LWP), a PSO-ANFIS model, and a PSO-ANN model. To develop their models, the variables considered were the pavement age, initial IRI, alligator, longitudinal, and transverse cracks, standard deviation of rutting, and subgrade plasticity index. These variables were chosen to capture both structural and material characteristics that affect pavement sections and performance. The data source for their work was the LTPP where they collected data on 126 different flexible pavement sections, with a total of 925 data points. To evaluate and compare the results of the different machine learning models, they used the R^2 as the main metric for evaluation.

Furthermore, Ali et al. (2022) used the fuzzy logic algorithm to model PCI of several pavement sections. To do so, several variables were analyzed and used to develop the machine learning model. These variables included rutting, fatigue cracking, block cracking, longitudinal cracking, transverse cracking, potholes, patching, bleeding, and raveling. These variables were used to cover a range of factors that contribute to pavement distress. The main source of data for this work was the LTPP. Ali et al. prepared two separate datasets of 120 and 150 different sections making a total number of 270. To assess the output of the model, they used the R^2 metric along with the RMSE and MAE.

Another research done which incorporated the use of several machine learning algorithms to model pavement performance was done by Damirchilo et al. (2021). Damirchilo et al. (2021) developed three machine learning models to predict IRI. The models were constructed using XGBoost, Support Vector Regression (SVR), and Random Forest. Out of the three algorithms, XGBoost was found to develop the most accurate results. Their research involved the use of age, days above 32^{o} C, freezing index, freeze thaw, hydraulic conductivity, ESALs, pavement thickness, and precipitation as input variables for each of the three machine learning models. Data was collected from the LTPP database on 12,637 data points. Finally, the results were compared between the results of the different models using the MAE and R^2.

Furthermore, Sharma & Kumar (2022) also developed more than one machine learning model in order to compare the performance of different algorithms. In their work, they constructed three models: a Logistic Regression model, a Naive Bayes model, and a KNN model. Their results showed that the Logistic Regression model yielded the most accurate results in predicting PCI out of the three models. The variables considered were alligator cracking, bleeding, block cracking, edge cracking, longitudinal cracking, transverse cracking, patching, pothole, shoving and raveling. Data was collected from the LTPP database with a total of over 10,000 data points. To assess the results and compare the three models, the authors used accuracy, precision, recall, and F1-Score as the metrics for model assessment.

Other research done in modeling pavement conditions is done by Hosseini et al. (2020), where their work was focused on modeling PCI of pavement sections. The researchers employed two modeling techniques. The first was a Long Short-Term Memory (LSTM) model, and the second model was a regression model. In their work, they considered the following variables: rutting, transverse cracking, longitudinal cracking, alligator cracking, wheel-path cracking, patching, age, traffic levels, and IRI. The data was collected on asphalt concrete (AC), Portland cement concrete (PCC), and composite (COM) pavements with a total of 31,045, all collected from the LTPP database. In order to evaluate their results, and compare between the LSTM and regression model, Hosseini et al. (2020) used the R^2 and mean error (ME) as evaluation metrics.

To conclude, existing research done in modeling pavement performance is wide and varied. Research was done on using both machine learning and traditional approaches to model pavement performance. Different data sources have been used in the literature, but the most comprehensive and commonly used data source was found to be the LTPP. Different performance indicators are used to evaluate pavement performance, but the two most studied indicators are the IRI and PCI of pavement sections. Even though the number of data points used in previous research varies greatly, starting from hundreds to thousands of data points, the more data points used, the more robust and inclusive the model becomes.

A wide array of machine learning algorithms has been applied in the literature, including regression models, ANN, GPR, LWP, PSO, PSO-ANFIS, PSO-ANN, Fuzzy Logic, XGBoost, SVR, Random Forests, Naive Bayes, KNN, and LSTM. Moreover, to evaluate these models, different evaluation metrics have been used, such as the R^2, RMSE, MAE, ME, accuracy, precision, recall, and F1-Score.

Furthermore, some research is not only focused on developing one machine learning model, but also performing a comparative analysis between two or more machine learning models. These analyses were done by also incorporating statistical indicators to assess the results of the different models.

Given the diversity and complexity of this research, table 3 was constructed to summarize the previous work done in this area. The table contains a summary of the machine learning model used, performance indicators modeled, variables considered, sources of data collection for the research and the total number of data points considered in the research. The purpose of this table is to serve as an overview of the significant and recent work done in the area of modeling pavement performance using machine learning algorithms.

Table 4: Previous Pavement Performance Models

Model	Model Used	Performance Index	Variables	Data Source	Number of data points	Evaluation Metrics
Elhadidy et al. (2021)	Regression Model	PCI	IRI, Cracks, patching, potholes, shoving, bleeding, rutting, reveling	LTPP	12,744	R^2
Hossain et al. (2019)	ANN	IRI	Temperature, freezing index, maximum and minimum humidity, precipitation, AADT, AADTT	LTPP	200	RMSE
Kaloop et al. (2023)	GPR, LWP, PSO-ANFIS, PSO-ANN	IRI	Age, initial IRI, alligator, longitudinal, and transverse cracks,	LTPP	925	R^2

			standard deviation of rutting, subgrade plasticity index			
Ali et al. (2022)	Fuzzy Logic	PCI	Rutting, cracks, potholes, patching, bleeding, raveling	LTPP	270	R^2, RMSE, MAE
Damirchilo et al. (2021)	XGBoost, SVR, Random Forests	IRI	Age, days above 32° C, freezing index, freeze thaw, hydraulic conductivity, ESALs, pavement thickness, precipitation	LTPP	12,637	MAE, R^2
Sharma & Kumar (2022)	Regression Model, Naive Bayes, KNN	PCI	Alligator cracking, bleeding, block cracking, edge cracking, longitudinal cracking, transverse cracking, patching, pothole, shoving raveling	LTPP	>10,000	Accuracy, precision, recall, F1-Score
Hosseini et al. (2020)	LSTM, Regression Model	PCI	rutting, transverse cracking, longitudinal cracking, alligator cracking, wheel-path cracking, patching, age, traffic levels, IRI	LTPP	31,045	R^2, ME

2.7 Summary and Gaps in Existing Literature

2.7.1 Summary of Key Findings

The previous sections showed a comprehensive review of the current research done in modeling pavement performance, with a specific emphasis on models that incorporate machine learning algorithms. That is because even though there are other approaches, such as deterministic and stochastic approaches, the literature showed that machine learning models tend to provide better results in terms of model accuracy and the ability to adapt with the variability of pavement.

As a result, the application of machine learning in pavement condition modeling has evolved greatly. Machine learning algorithms commonly include XGBoost, random forests, regression models, naive bayes, ANN, among other algorithms. Machine learning models have shown promise compared to traditional and stochastic models previously developed.

Moreover, when evaluating machine learning models, previous researchers have used a variety metrics such as RMSE, MAE, ME, MSE, R^2, accuracy, precision, recall and other advanced metrics such as F1-Score and AUC-ROC to evaluate model performance. These metrics have shown that they can be a reliable tool to assess the performance of a single machine learning model or compare the results of different models in modeling pavement performance.

In terms of factors to be considered when developing pavement performance models, it was found that pavement age, environmental conditions, state of different cracks, maintenance history, the effect of traffic loading and material properties have a significant impact on pavement performance. For collecting data regarding these variables, the LTPP database is one of the most commonly used sources of data.

2.7.2 Research Gaps

After studying the literature, a few research gaps were found that require additional work around modeling pavement performance, especially when it comes to utilizing machine learning models in this area:

- **Comprehensive Variable Analysis:** In most of the research done, only a specific set of variables are considered when modeling model performance. Researchers focus on specific types of variables like environmental conditions, traffic loading, or material properties separately. Minimal research is done that focuses on the combination of the preceding types of variables and analyzing the most significant ones.

- **Multi Objective Modeling:** Previous models showed a main focus on modeling pavement performance indicators or the state of specific distresses like rutting or certain types of cracks. However, there was a lack of models that combined the modeling of performance indicators like IRI and PCI and the state of distresses such as fatigue cracking and longitudinal cracking.

- **Machine Learning Algorithm Comparison:** Even though there has been research done on comparing the results of different machine learning algorithms, most of the literature showed a focus on developing a single machine learning model for pavement performance. There is a relative scarcity on conducting a comparative analysis on different machine learning algorithms in modeling pavement performance.

- **Localized Context:** The literature showed that environmental and geographic conditions have a significant impact on pavement deterioration. However, most studies utilize data from the LTPP database without any specific research done on the environmental conditions similar to those in Egypt.

Chapter 3

Methodology

3.1 Introduction to the Methodology

This section of the research highlights the approach used to tackle the research gaps found in the literature. It is known that pavement conditions have a direct impact on user costs, road safety, and the economic well-being of society, with pavement deterioration negatively affecting these aspects. This, coupled with the fact that the IRI is a widely used indicator of pavement conditions and ride quality, makes it necessary to accurately model pavement IRI.

Aside from IRI, a major aspect of pavement deterioration is the creating of cracks in pavement. Load-induced cracks, to be specific, are an area of concern. Even though modeling IRI can provide insight into the overall pavement condition of a given road, it might not fully provide an overview of the conditions of specific distresses in pavement which should be accounted for in planning for maintenance and rehabilitation. As mentioned in the literature, two of the most significant load-induced cracks in pavement are fatigue and longitudinal cracks. As a result, it is vital to model these distresses as well to help in developing maintenance and rehabilitation plans.

To address these matters, this section provides the approach taken to employ several machine learning techniques to model both pavement IRI and the mentioned distresses. This is due to machine learning's capabilities in dealing with large datasets and capturing complex relationships between the different factors that affect pavement performance and lead to the formation of cracks in pavement sections.

This has been done by developing six machine learning models, all with the same objective. By conducting a comparative analysis between the results of these models, the purpose is to identify the model that produces the most accurate results out of the six. The comparative analysis will be done by using statistical evaluation metrics which are commonly used in the literature.

3.2 Data Collection

One of the most important aspects of any machine learning model is the quality of the dataset used to train and develop the model. For this research, the LTPP was the main source of data due to it containing comprehensive data on several factors affecting pavement performance. The database, which is run by the FHWA, contains information on varying traffic and climate conditions, over more than 50 years, which makes it an essential tool for this study.

3.2.1 The Long-Term Pavement Performance Database

The LTPP database contains data about varying factors that have an impact on pavement performance. One of the major contents of the LTPP database is the different IRI measurements over time. These different measurements provide an overview of the condition of different road sections over the years by providing the IRI at different points in time. Other important variables found in the database are related to environmental conditions, traffic loads, road section characteristics, and maintenance history. Figure 6 shows a screenshot of the LTPP InfoPave homepage.

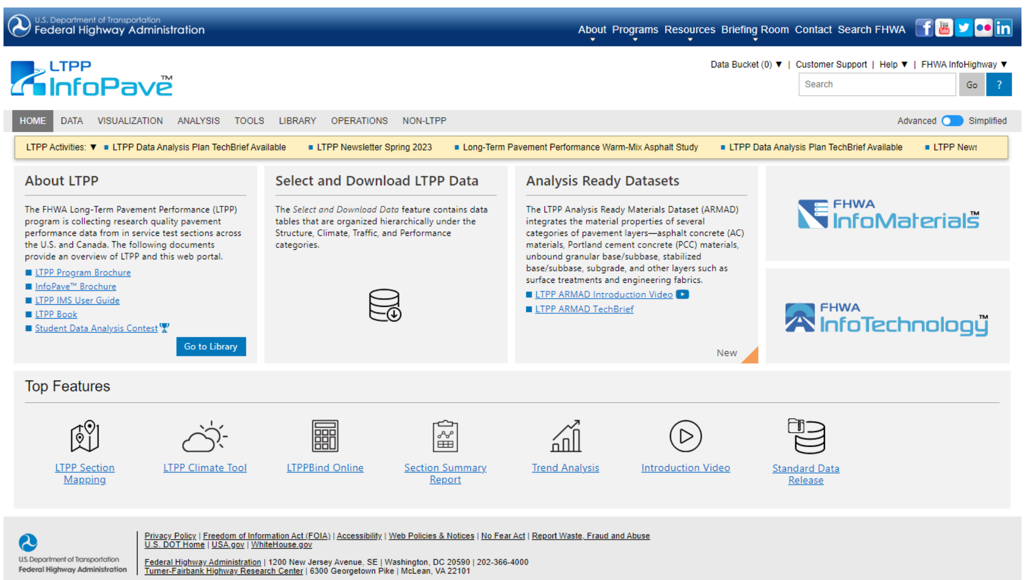


Figure 6. LTPP InfoPave Homepage

3.2.2 Selection Criteria

For this research and in order to align with the conditions similar to those found in Egypt, there were two main selection criteria used:

Climate Region

The climate regions in the LTPP database are divided into four distinct categories:

1) Dry, Freeze

2) Dry, No-Freeze

3) Wet, Freeze

4) Wet, No-Freeze

Since the main purpose was to choose regions where the climate conditions are similar to those found in Egypt, the climate region considered was the Dry, No-Freeze region. Accordingly, data was collected from the following states at times of the year with dry, no-freeze weather conditions:

1) Arizona

2) California

3) Colorado

4) Hawaii

5) Idaho

6) Nevada

7) New Mexico

8) Oregon

9) Texas

10) Utah

11) Washington

Pavement Type

There are four different pavement surface types found in the LTPP database which include:

1) Asphalt Concrete Pavement

2) Continuously Reinforced Concrete Pavement

3) Jointed Plain Concrete Pavement

4) Jointed Reinforced Concrete Pavement

Since deterioration is influenced by the pavement surface type, this study will focus on only one type of pavement which is the Asphalt Concrete Pavement. The resulting dataset comprised of 221 different pavement sections with a total of 8537 different datapoints resulting from measurements starting from the year 1989 till 2021.

3.2.3 Factors Considered

As discussed in the literature, different factors affect pavement performance and influence different distresses in pavement. In order to have a comprehensive overview of the different factors that affect these two, twelve factors were considered. Table 5 summarizes these factors by mentioning the unit of measurement and the factor's source from the LTPP database.

Table 5: Factors Considered for Model Development

Factor Considered	Unit of Measurement	Source from the LTPP database
Pavement Age	Years	PROJECT_HIST_AGE_EXP
Time from last maintenance	Years	EXPERIMENT_SECTION
Precipitation	Mm	MERRA_PRECIP_YEAR
Average Monthly Temperature	Degree Celsius	MERRA_TEMP_YEAR
Equivalent Single Axle Loads	ESALs	TRF_TREND
Asphalt concrete layer thickness	inches	TST_L05B
Average asphalt content	%	TST_AC04
Bulk specific gravity	N/A	TST_AC02
Granular base thickness	inches	TST_L05B
Current IRI	m/km	MON_HSS_PROFILE_SECTION
Percentage of fatigue cracking	%	MON_DIS_AC_CRACK_INDEX
Percentage of longitudinal cracking	%	MON_DIS_AC_CRACK_INDEX

In order to visually display the frequency of each variable across the 8537 datapoints, the following graphs were prepared as shown in Fig. 7.

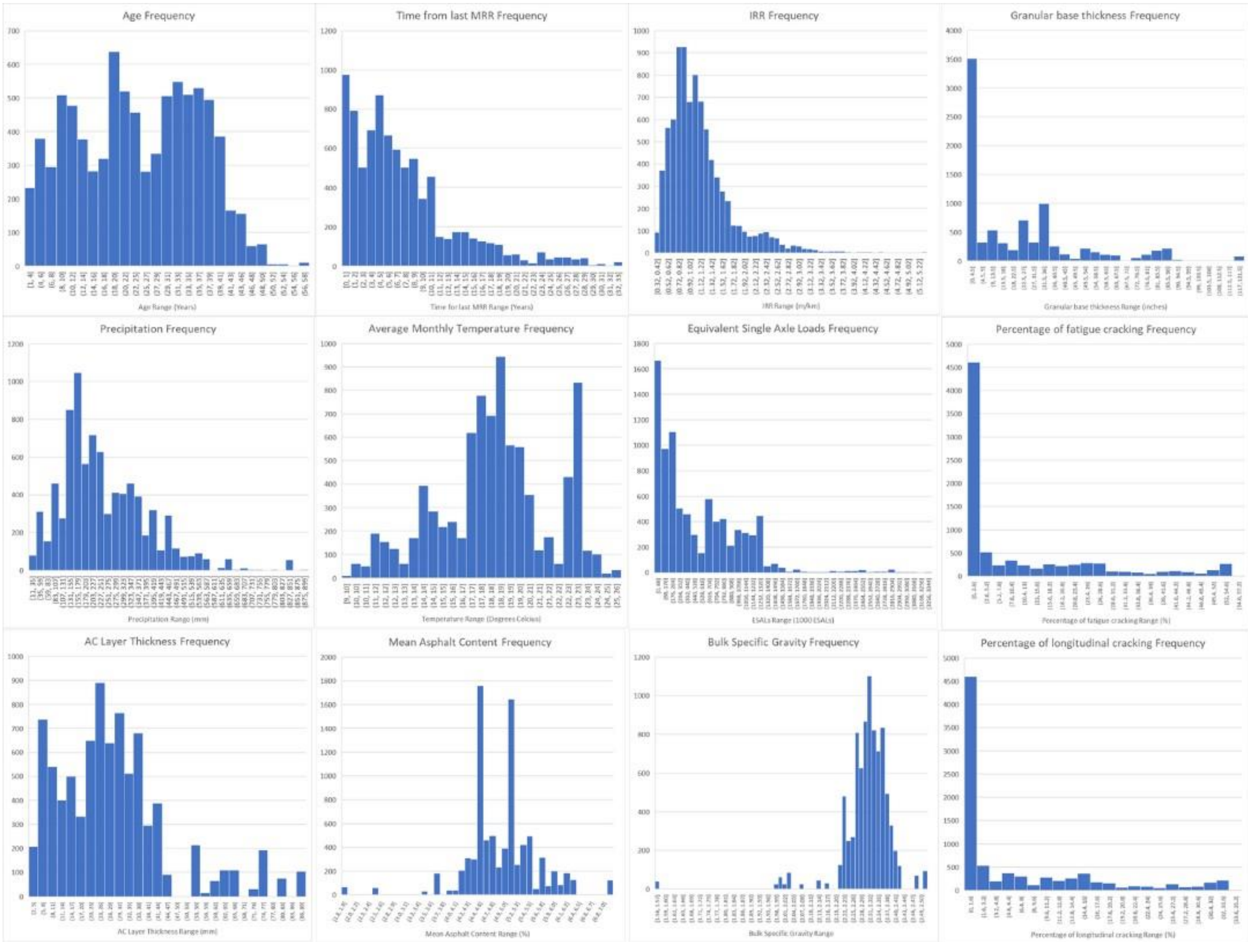


Figure 7: Frequency of used Variables

Moreover, table 6 shows statistical data regarding the dataset, highlighting the minimum, maximum, average, and standard deviation of each variable.

Table 6: Dataset Statistics

Factor Considered	**Min**	**Max**	**Average**	**St. Dev.**
Pavement Age	1.4	57.9	23.9	12.0
Time from last maintenance	0.0	32.9	7.2	6.1
Precipitation	11.48	876.49	252.81	138.29
Average Monthly Temperature	9.30	25.30	18.20	3.28
Equivalent Single Axle Loads (1000)	0.00	3263.26	535.55	511.73
Asphalt concrete layer thickness	2.00	86.70	28.39	17.73
Average asphalt content	1.80	7.06	4.91	0.69
Bulk specific gravity	1.56	2.52	2.31	0.09

Granular base thickness	0.00	120.00	21.90	26.23
Current IRI	0.32	5.28	1.16	0.59
Percentage of fatigue cracking	0.00	55.00	10.78	15.46
Percentage of longitudinal cracking	0.00	34.00	6.54	9.46

3.3 Selection of Machine Learning Algorithms and Model Development

For this research, a set of machine learning algorithms were chosen to model both pavement IRI and percentages of fatigue and longitudinal cracks in pavement. This set included 4 solo machine learning models and 2 ensemble models.

The 4 solo machine learning algorithms chosen we Ridge Regression, KNN, Decision Tree Regression, and Bayesian Regression. Moreover, the ensemble models developed were done using XGBoost and Random Forests.

3.3.1 Model Development and Training

Data was first normalized so that each variable has a range between -1 to 1. This was followed by splitting the data into into 80% for training and 20% for testing the model. Each model is cross validated by splitting the dataset into 10 subsets. Next, 2 of the subsets would be used for testing while the other 8 subsets would be used for training. The training and testing process is repeated using different combinations for training and testing sets. By doing so, a more comprehensive assessment is done on the predictive capabilities of each model. Finally, the average performance is then calculated in order to assess each model's performance. Moreover, a heatmap was developed to test for any collinearity within the variables used. In this process, the correlation between each of the variables used in the models is calculated. If two variables show to be codependent, or colinear, one would be removed from the model in order to produce more accurate results.

3.3.2 Evaluation Metrics

For evaluating the results of the models, two evaluation metrics have been used. The metrics are the MAE and R^2. MAE is used to illustrate the average error of the model represented using the unit of each output variable. R^2 is used to show the goodness of fit of each model. These metrics were chosen as most researchers choose either an error metric to evaluate the closeness of the forecasted values to the actual values, along with the R^2 to represent and compare between the goodness of fit between different models.

Chapter 4

Results and Discussion

4.1 Model Performance and Results

In order to study the correlation between different variables of each model, a heatmap was developed to indicate any possible collinearity, this can be shown in figure 8.

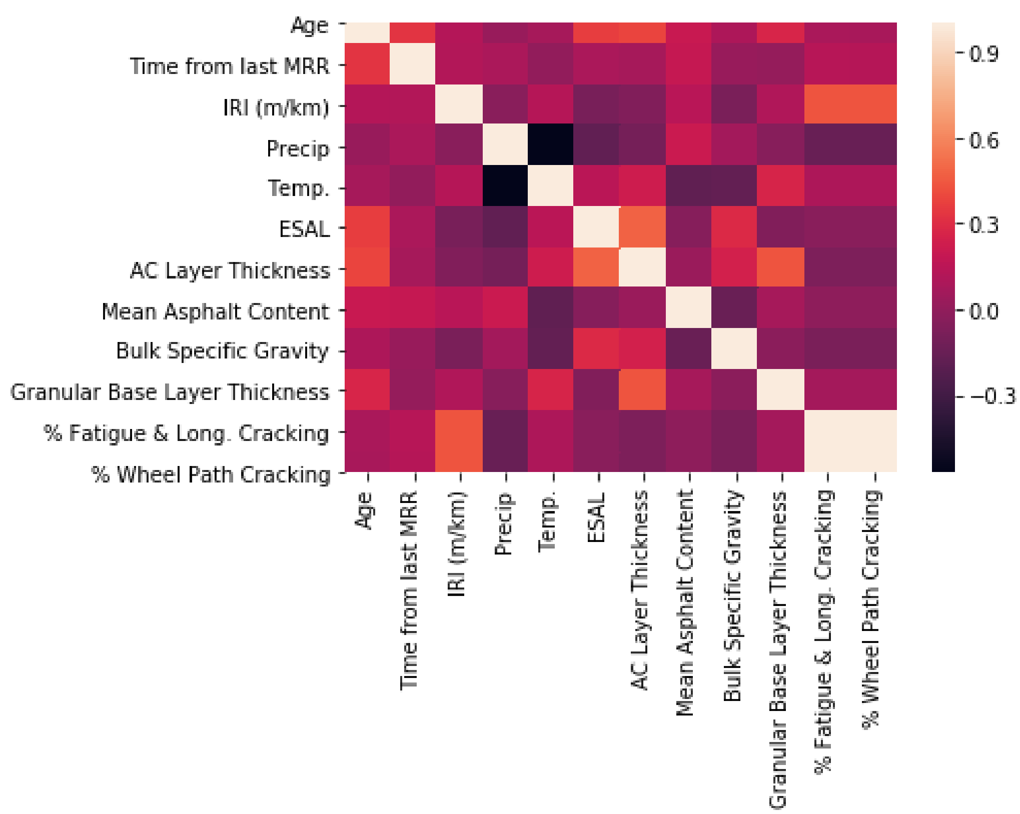
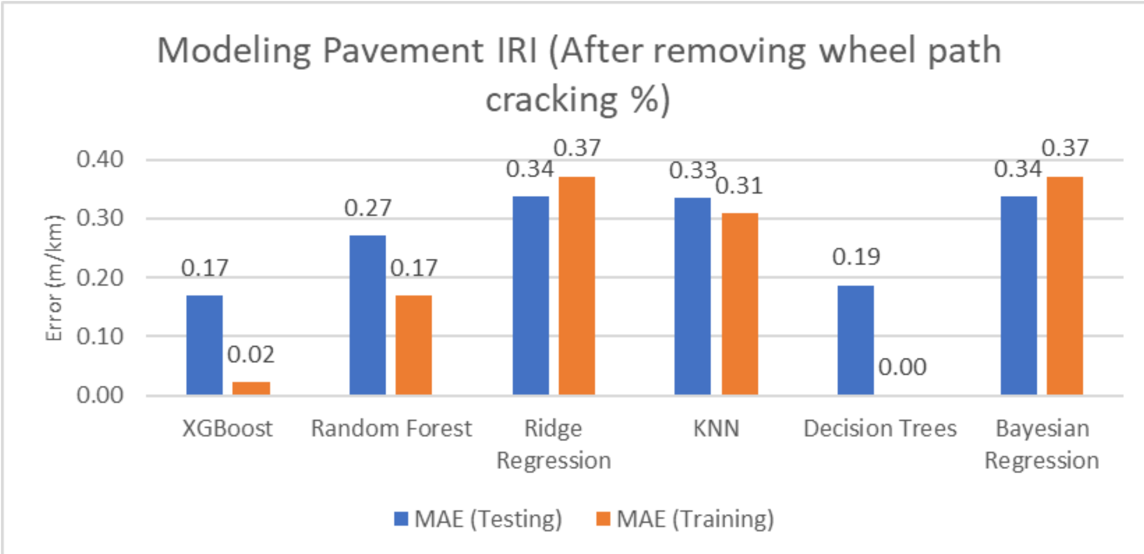


Figure 8: Heatmap to Detect Collinearity

As shown, it was found that the fatigue cracking % and longitudinal cracking % showed high collinearity. As a result, the models were developed by including both variables when modeling IRI, then another iteration was done with removing one of the two variables.

After running the 6 models the results were recorded as shown in figures 9, 10 and 11. Where figure 9 shows the results of the 6 IRI models after removing longitudinal cracking %. Either wheel path or fatigue cracking had to be removed from the input variables due to the collinearity between both variables. Longitudinal cracking was removed because during an initial run which incorporated both these variables, fatigue cracking showed to have a higher influence on the predicted IRI than longitudinal cracking. Figure 10 shows the results of fatigue cracking % models. And lastly, figure 11 shows the results of the longitudinal cracking models.



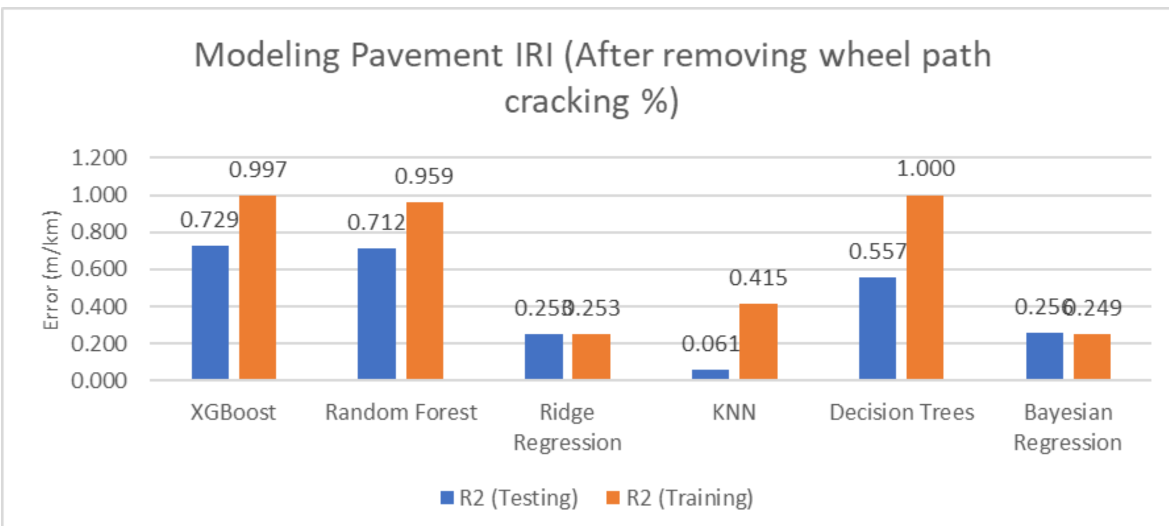
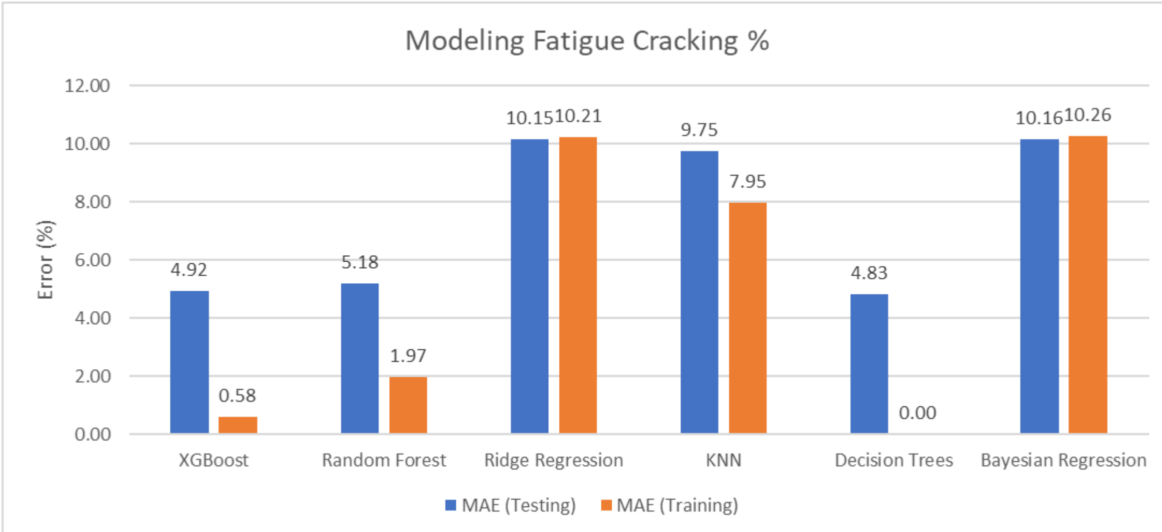


Figure 9: Results of Modified IRI Model (After removing longitudinal cracking %)



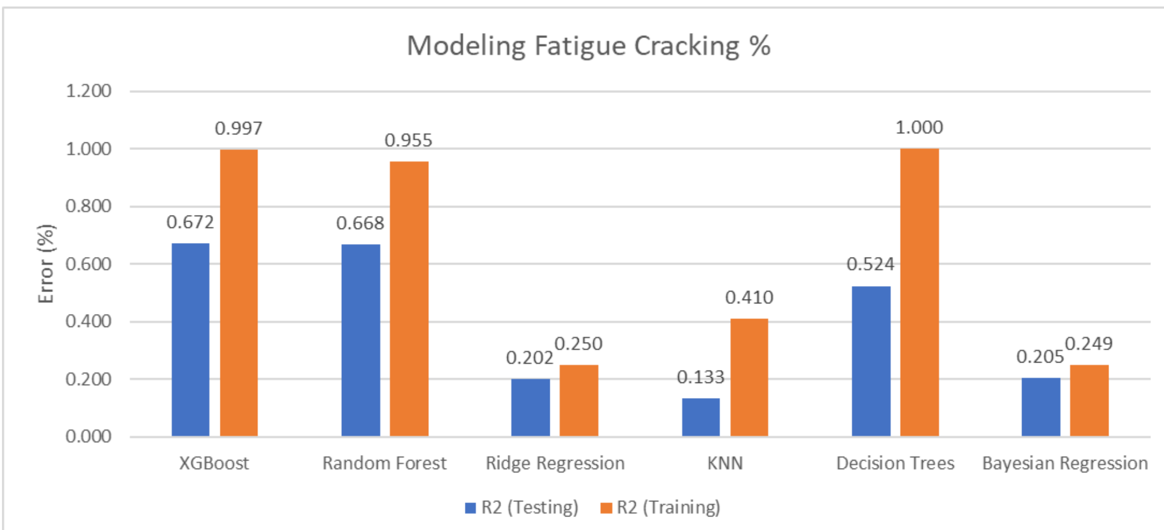
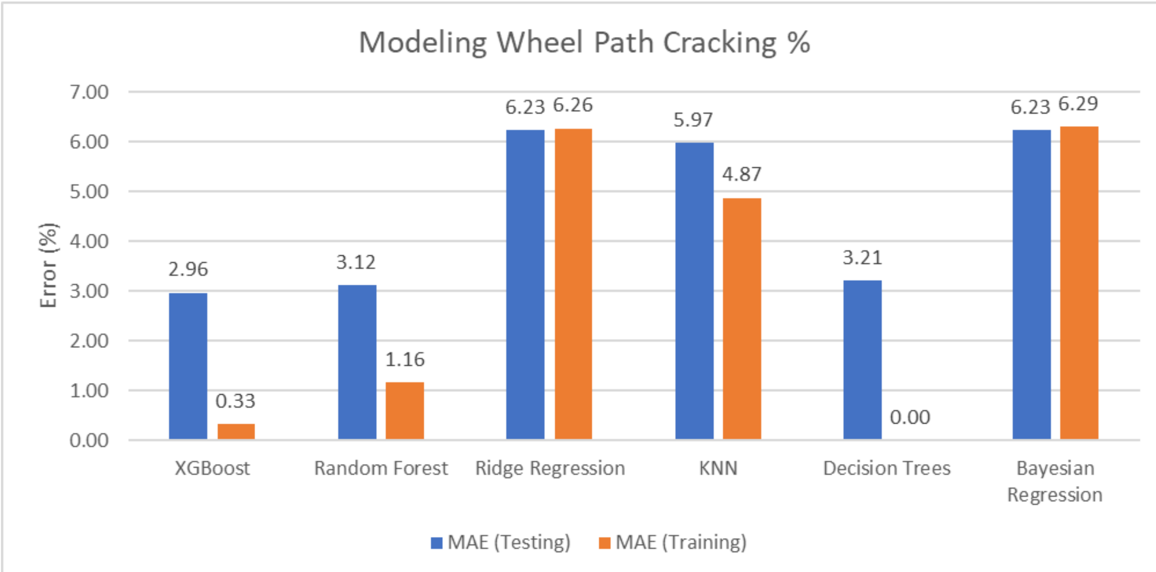


Figure 10: Results of Fatigue Cracking Model



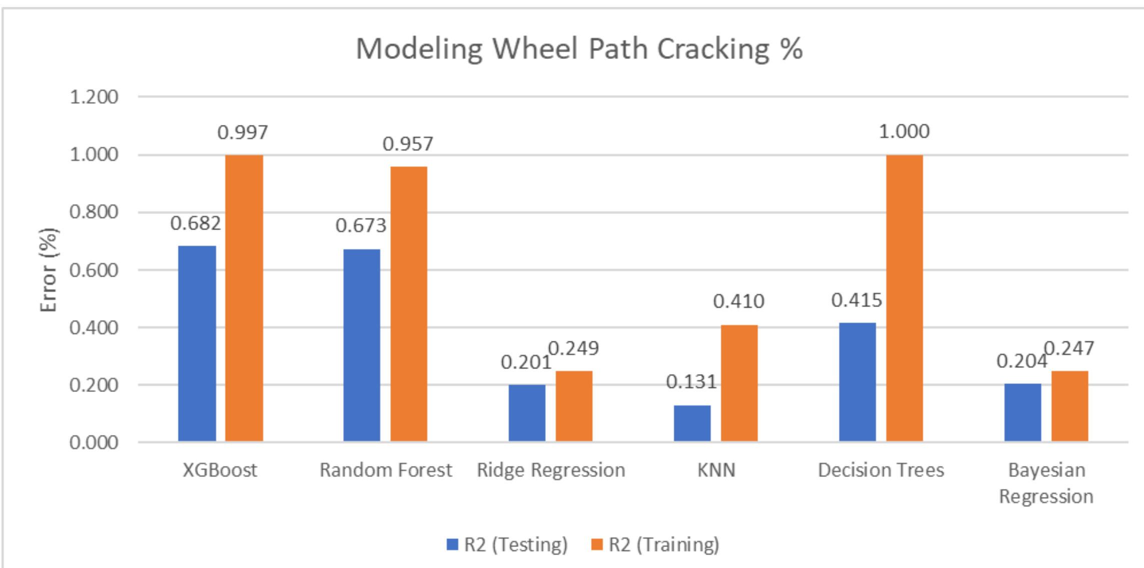


Figure 11: Results of Longitudinal Cracking Model

The MSE for each of the 6 models can be visually displayed using boxplots for modeling each of the three output variables. Figure 12 shows the MSE boxplot for modeling IRI after removing the longitudinal cracking %. Since the fatigue cracking and longitudinal cracking are displayed in percentages, a box plot was made for the negative MAE for each of the two outputs. Figures 13 & 14 show the MAE boxplots for modeling fatigue cracking and longitudinal cracking, respectively.

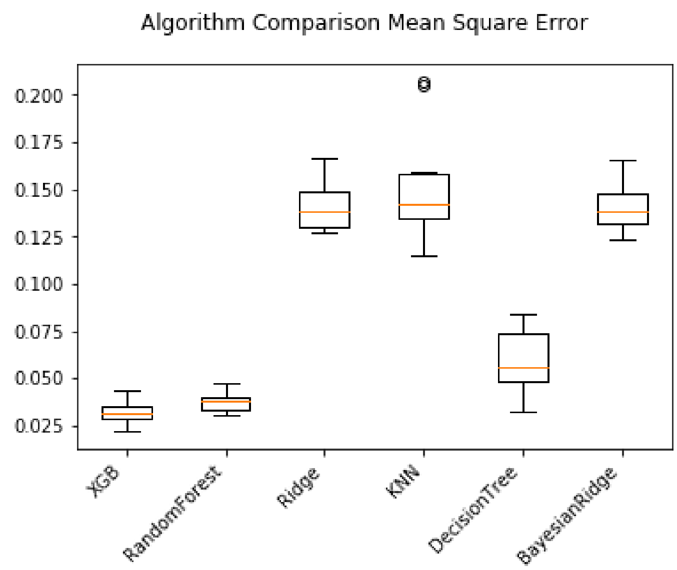


Figure 12: IRI Models' Boxplots (LongitudinalCracking removed)

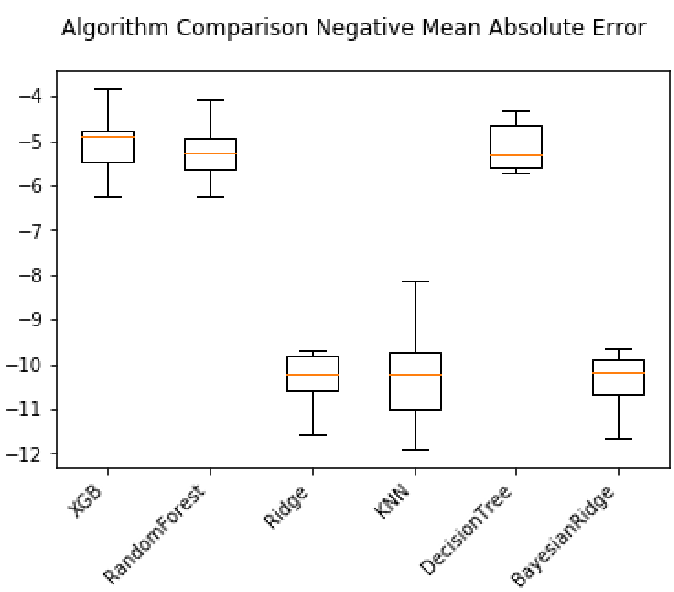


Figure 13: Fatigue Cracking Models' Boxplots

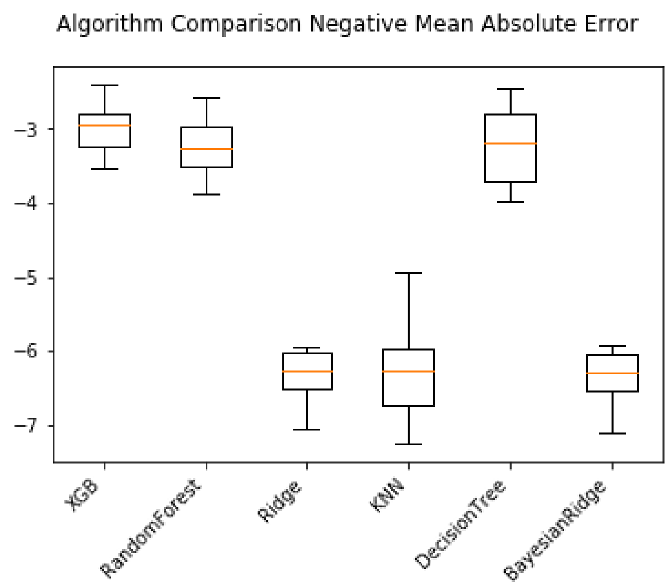


Figure 14: Longitudinal Cracking Models' Boxplots

Following the calculated MAE and R^2 for each model, the results of the models were plotted against the actual values for each of the three output variables. Figures 15 and 16 show the plots for the results of the IRI models, while figure 17 shows the results of the fatigue cracking models, and finally, figure 18 shows the results of the longitudinal cracking models.

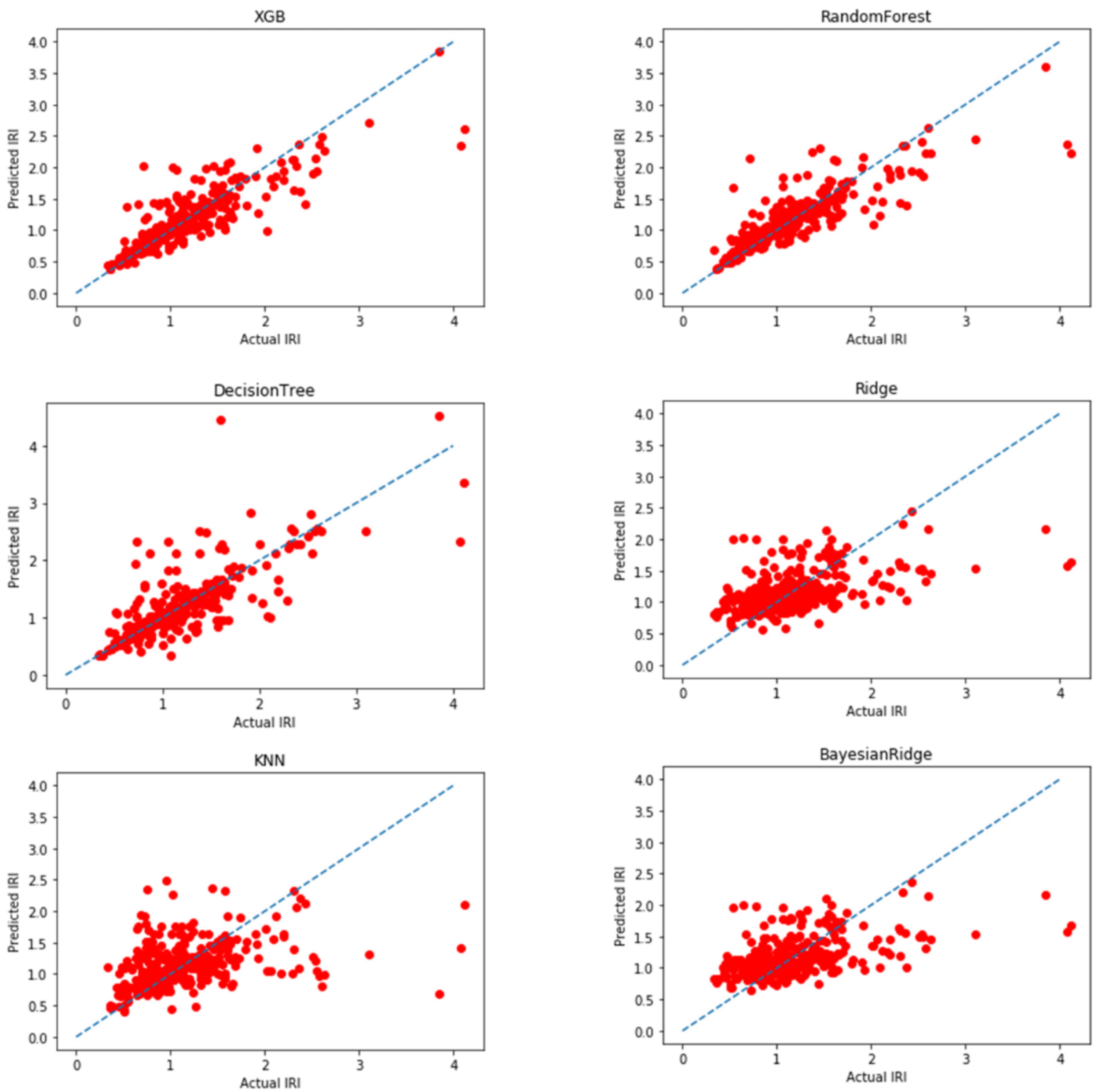


Figure 15: Actual vs. Modeled IRI (All Variables)

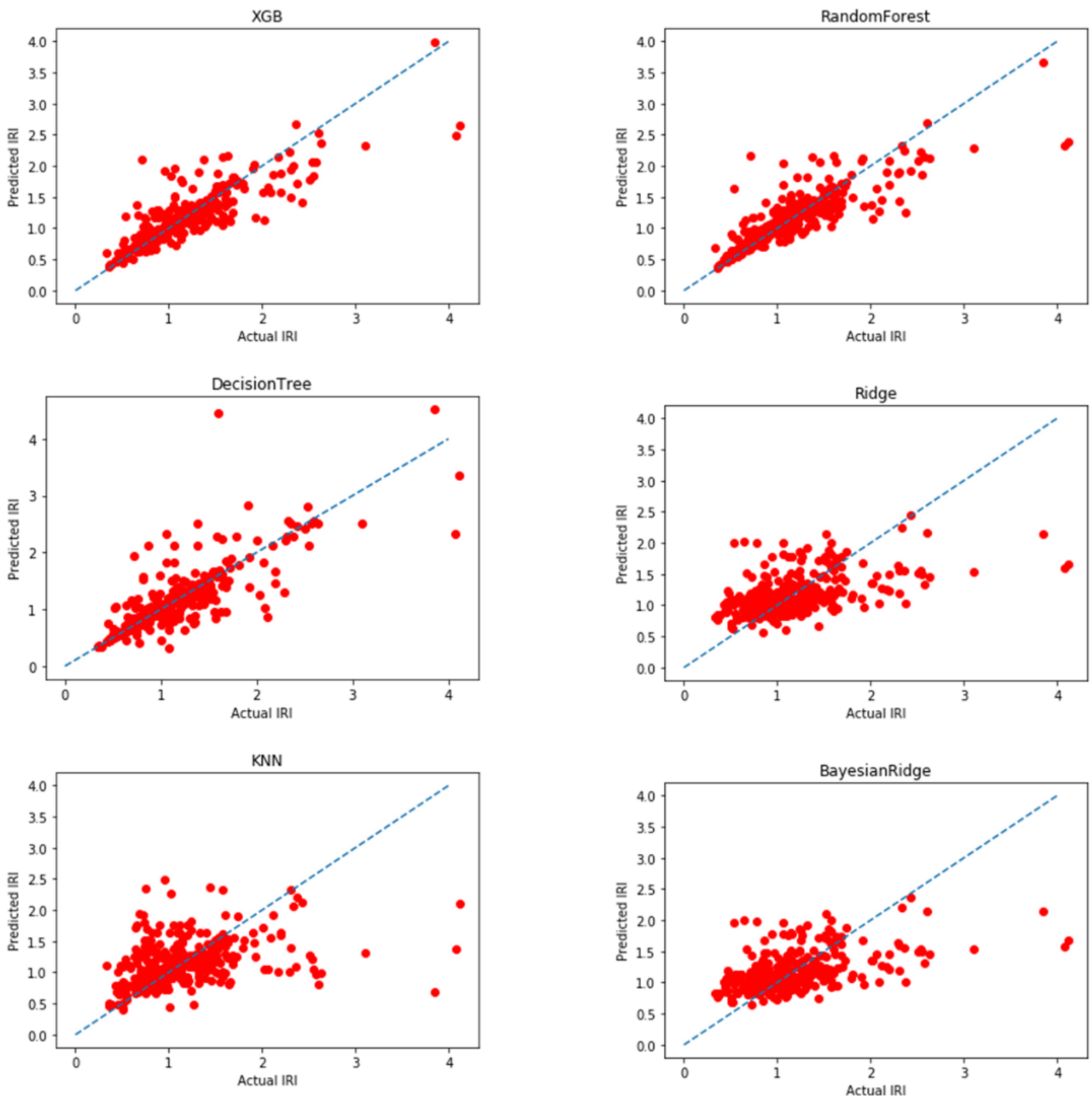


Figure 16: Actual vs. Modeled IRI (Longitudinal Cracking removed)

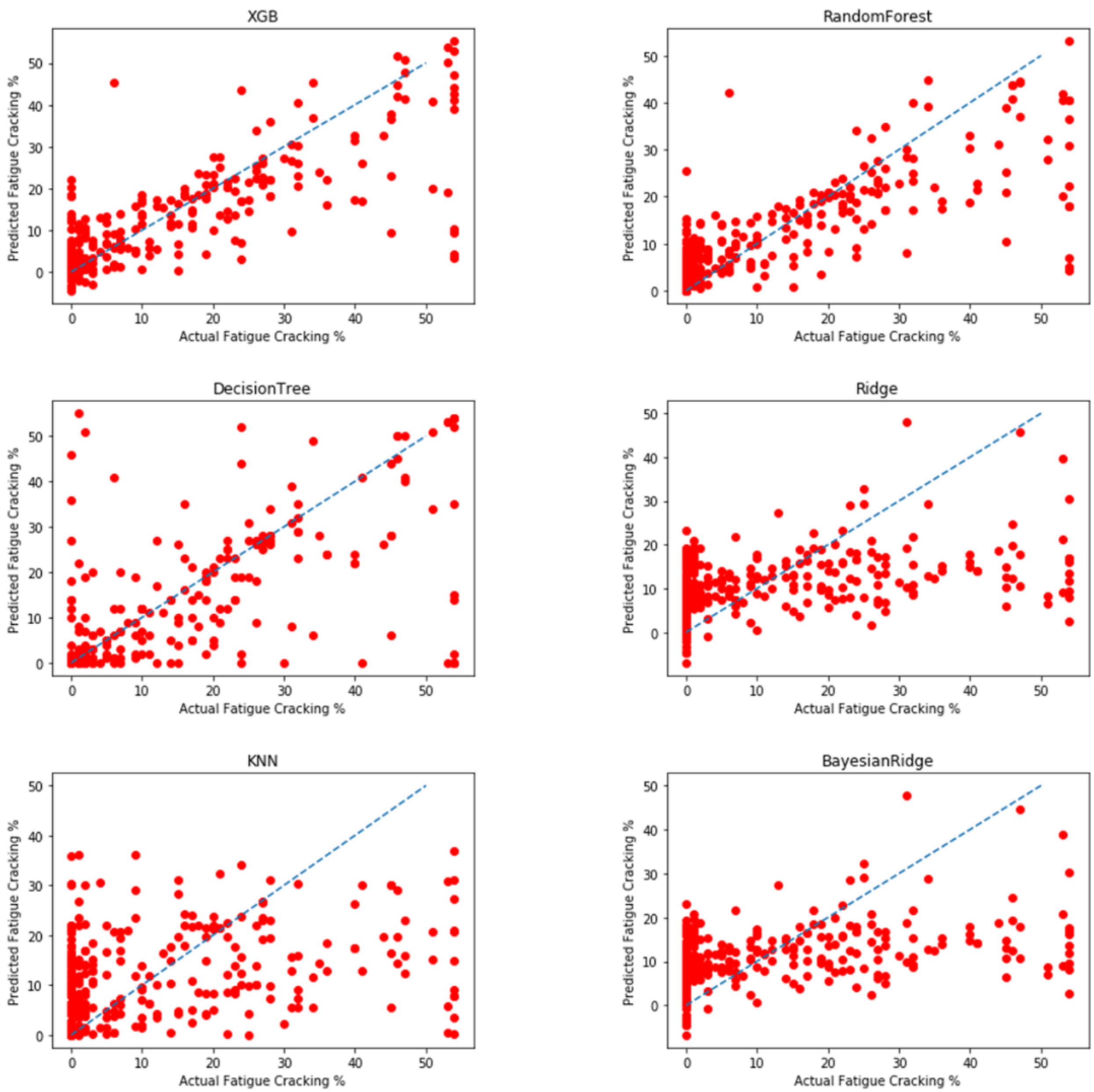


Figure 17: Actual vs. Modeled Fatigue Cracking %

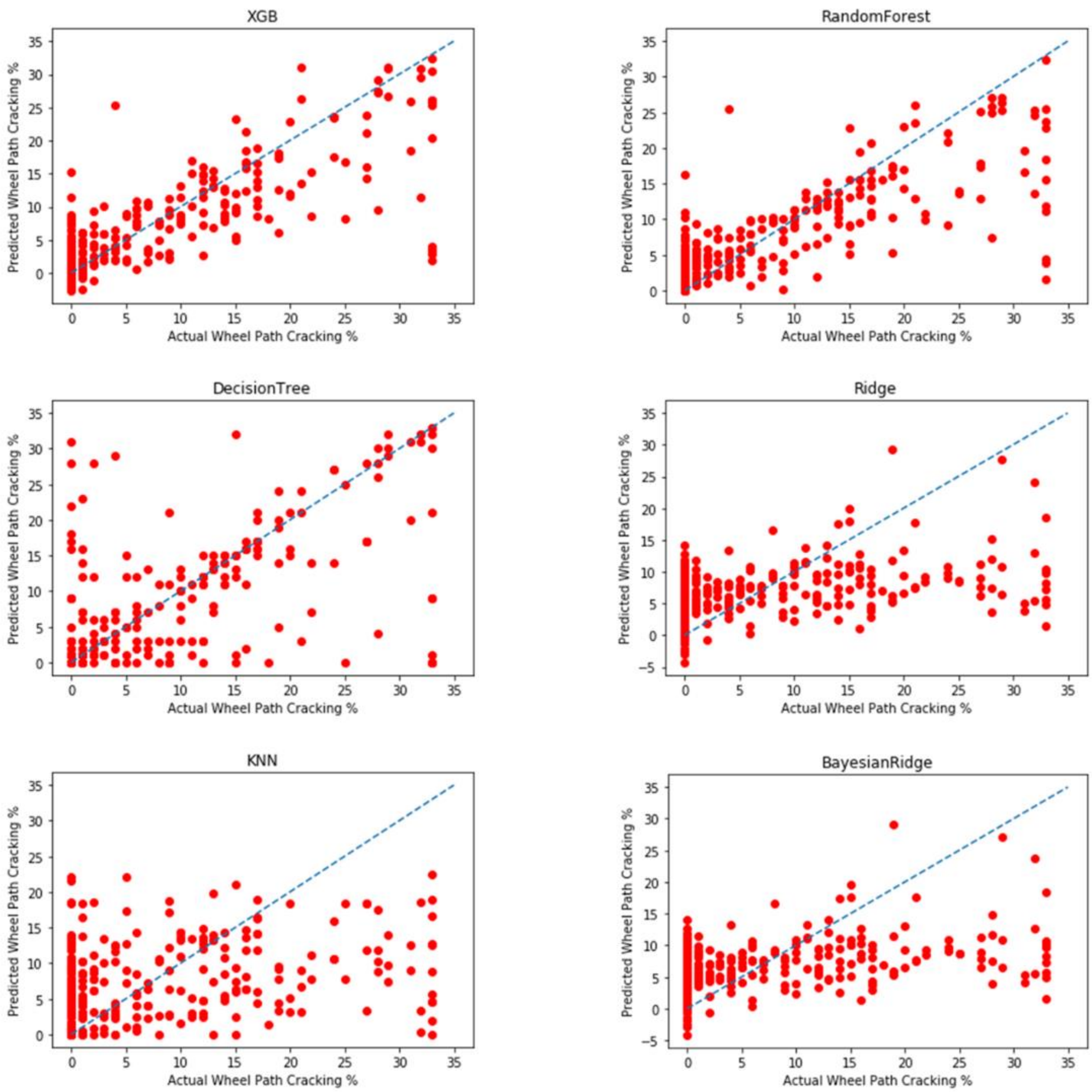


Figure 18: Actual vs. Modeled Longitudinal Cracking %

Out of the 10 variables considered for each of the three output variables, each input variable showed to have a different influence on the results of each model. To analyze the contribution of each variable to the results of each model, the following graphs were plotted for the three machine learning models that showed the least MAE. In this case, the graphs were plotted for the models which utilized the XGBoost, Random Forest, and decision tree algorithms. Figures 19 through 22 show the variable contribution for each model where the x-axis represents each of the variables used and the y-axis represent the variable contribution. The total sum of variable contribution for each model is equal to 1.

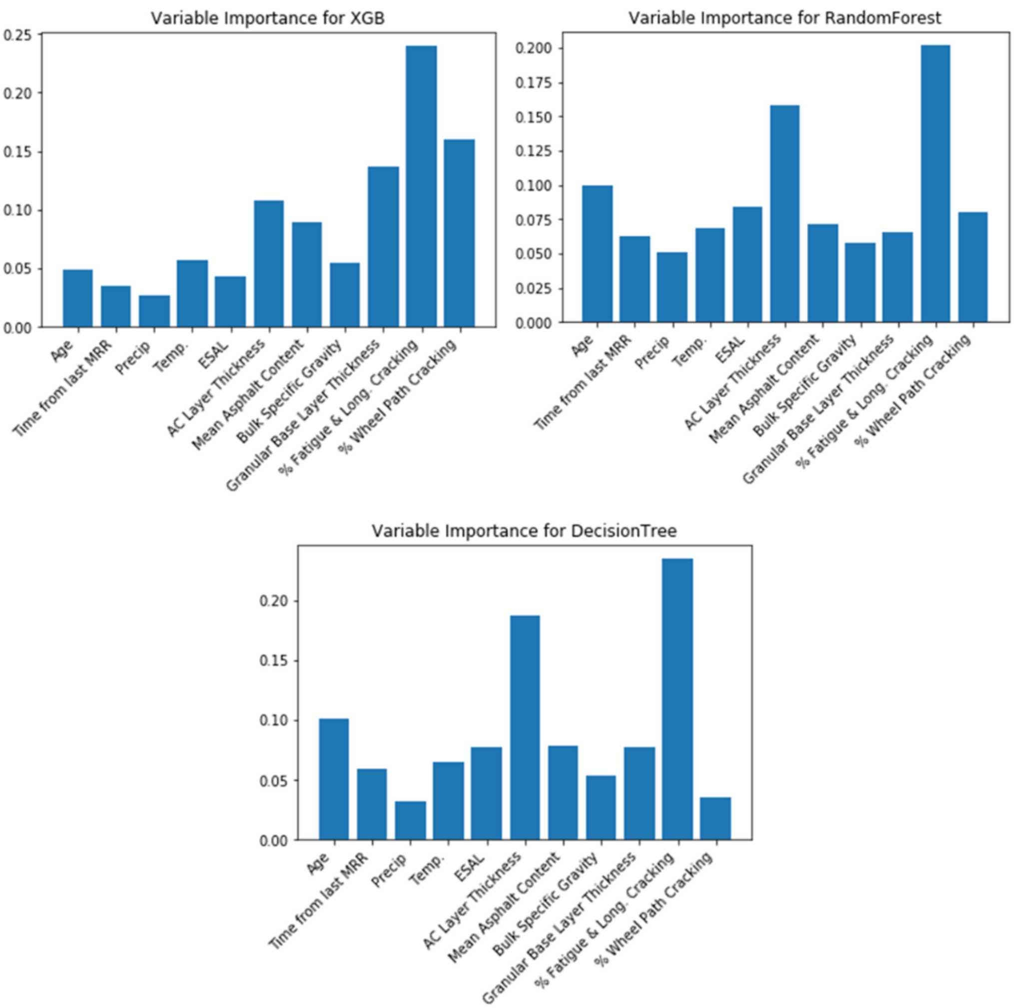


Figure 19: Variable Contribution for IRI Models (All Variables)

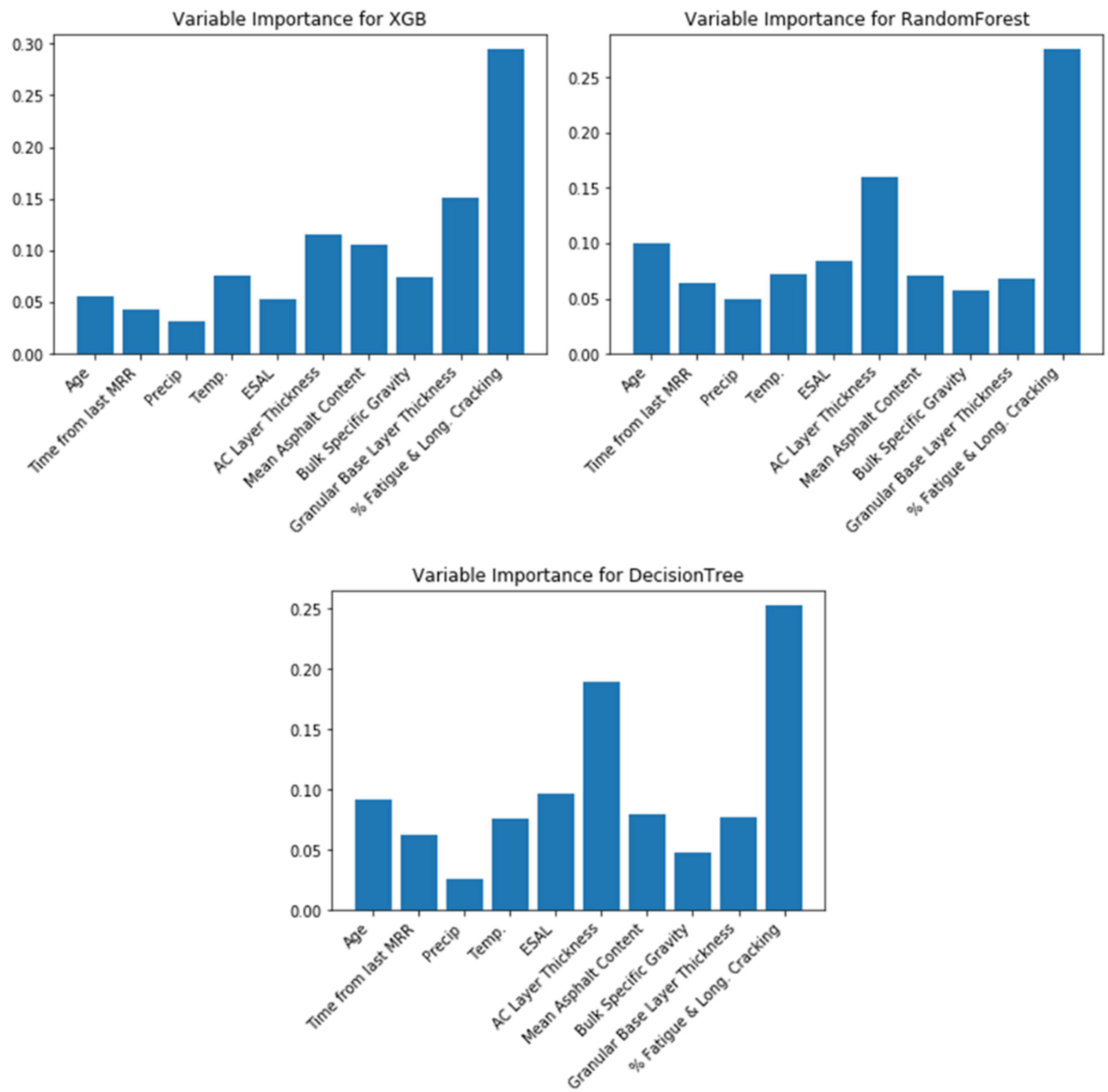


Figure 20: Variable Contribution for IRI Models (Longitudinal Cracking removed)

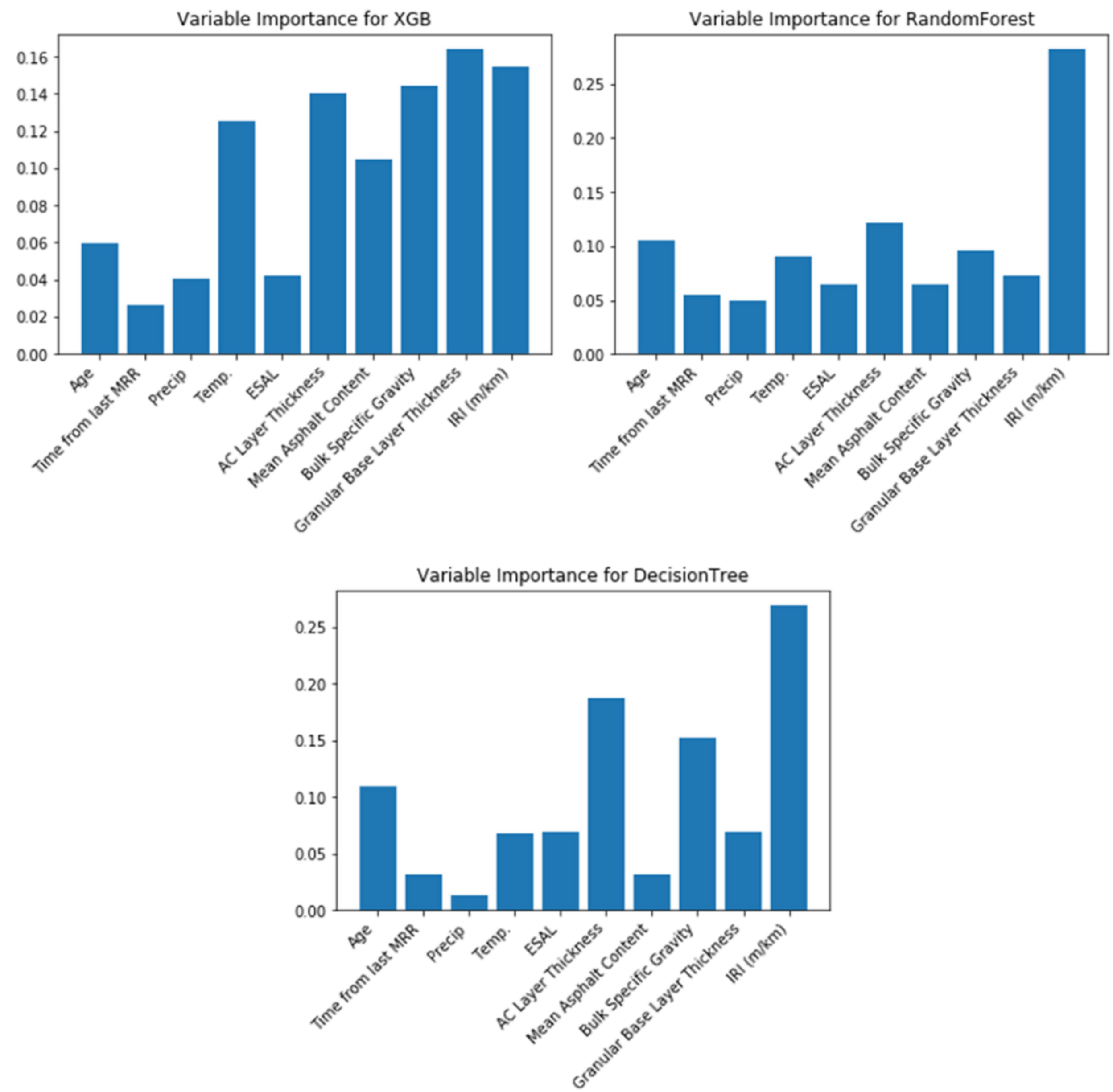


Figure 21: Variable Contribution for Fatigue Cracking Models

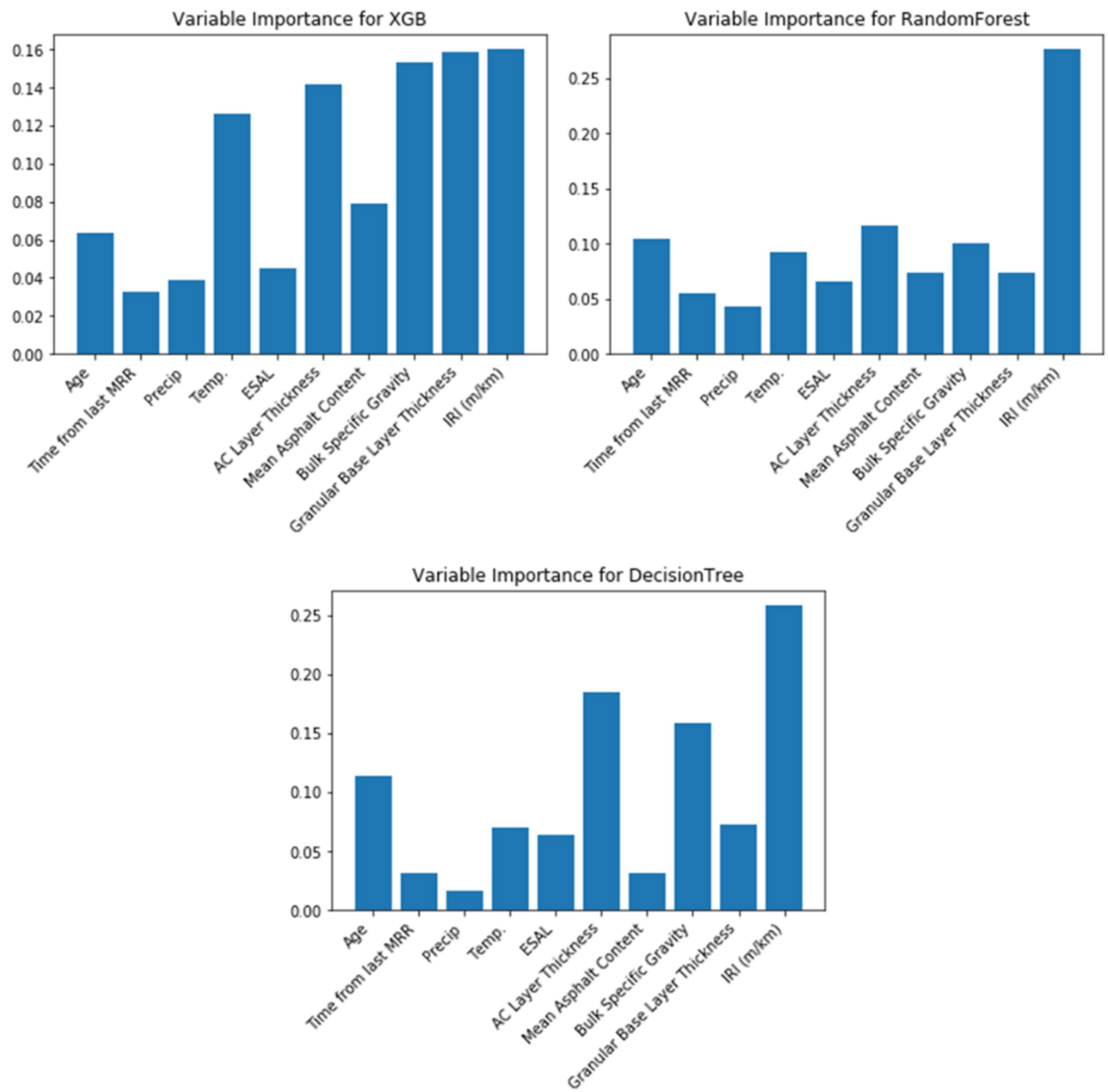


Figure 22: Variable Contribution for Longitudinal Cracking Models

4.2 Analysis and Discussion on the Results

The first step in the analysis process was to determine if any collinearity was present within the variables in the models. As was shown in the developed heatmap, longitudinal cracking and fatigue cracking showed high collinearity. As a result, when modeling IRI, the process was done first with considering all the variables and then with removing the longitudinal cracking %. The reason longitudinal cracking % was removed was because fatigue cracking % showed a higher influence on pavement IRI when both variables were considered in the model.

As a result of removing the longitudinal cracking % from the IRI model, the developed models gave better results compared to the original model which included all 11 variables. Moreover, it was shown that the ensemble models performed better than the models which utilized solo machines learning algorithms. To be more specific, the models which utilized XGBoost showed the least MAE and R^2 when modeling IRI, fatigue cracking % and longitudinal cracking %. This illustrates the superior performance of XGBoost when compared to other machine learning algorithms. When modeling IRI, the Random Forest models were the second-best performing models followed by Decision Trees, Ridge Regression, Bayesian Regression, then KNN. In the IRI model, the XGBoost algorithm resulted in a MAE of 0.02 and R^2 of 0.997 for training and a MAE of 0.17 and R^2 of 0.729 for testing.

Moreover, in terms of variable contribution, the variables which showed the highest influence on IRI in the initial XGBoost model were fatigue cracking %, followed by longitudinal cracking % then granular base layer thickness. However, in the Random Forest and Decision Tree IRI models, the highest contributing variables were found to be fatigue cracking %, then AC layer thickness, followed by pavement age.

After removing the longitudinal cracking %, the top three contributing variables for the XGBoost IRI model were found to be the fatigue cracking %, followed by the granular base layer thickness, then the AC layer thickness. However, for the Random Forest and Decision Tree models, the highest contributing variables stayed the same.

As for the fatigue cracking % models and longitudinal cracking % models, the results were almost the same for both output variables. This can be explained by the collinearity found in the initial developed heatmap. Accordingly, it was found that for both output variables, the XGBoost developed the least MAE of 0.58% and R^2 of 0.997 for the testing set, and a MAE of 4.92% and R^2 of 0.672 for the testing set for fatigue cracking %. Moreover, for the wheel path model, the XGBoost algorithm produced a MAE of 0.33% and R^2 of 0.997 for the training set and MAE of 2.96% and R^2 of 0.682 for the testing set. This was followed by the Decision Tree models, then the Random Forest models, then KNN, Ridge Regression, then Bayesian Regression.

As for variable contribution, both fatigue cracking % and wheel cracking % were observed to be influenced by the same variables. For the XGBoost models, the top three contributing variables were granular base thickness, IRI and bulk specific gravity. As for both the Random Forest and Decision Tree models, the top three contributing variables were found to be IRI, AC layer thickness, and bulk specific gravity.

It is worth noting that the main evaluation metric considered in all the models was the MAE. This is due to the studies done (Bu & Clemente, 2022; Dayal, 2015; Matejka & Fitzmaurice, 2017) which indicate that R^2 can be a misleading evaluation metric.

However, by analyzing the results of the fatigue cracking and longitudinal cracking models, it can be deduced that the results could be improved. This can be done by investigating other variables that directly influence these distresses. As a result, it can be concluded that the variables that affect IRI do not in all cases have a direct effect on either fatigue cracking or longitudinal cracking.

Chapter 5

Conclusion and Future Work

5.1 Conclusion

To conclude, this research introduces a framework for modeling pavement IRI along with fatigue cracking % and longitudinal cracking %. To do so, twelve different variables were studied and considered during model development. Data was collected from the LTPP database for areas with environmental conditions similar to those found in Egypt. The study was conducted on AC pavement sections. Each of the three models was developed using six different machine learning algorithms. The machine learning algorithms were divided into ensemble algorithms and solo algorithms. The ensemble algorithms used were XGBoost and Random Forest, while the solo models used were decision trees, Bayesian Regression, KNN, and ridge Regression. A test of collinearity was done between the different variables of each model and the results were compared using MAE and R^2.

The results showed that the XGBoost models yielded superior results when compared to the other five algorithms for the three output variables with a MAE of 0.17 and an R^2 of 0.729 for modeling IRI, MAE of 4.92% and R^2 of 0.672 for modeling fatigue cracking, and a MAE of 2.96% and R^2 of 0.682 for modeling longitudinal cracking. This was followed by the Random Forest models which produced results with a MAE of 0.27 and R^2 of 0.712 for modeling IRI, MAE of 5.18% and R^2 of 0.668 for modeling fatigue cracking, and MAE of 3.12% and R^2 of 0.673 for modeling longitudinal cracking. This indicated the superior performance of ensemble models when compared to solo models for modeling pavement performance.

5.2 Suggestions for Future Work

This work provides a framework for modeling pavement performance through modeling IRI, fatigue cracking and longitudinal cracking. The main focus was on regions with environmental conditions similar to those in Egypt, with dry and non-freezing conditions. Similar work could be done by exploring other regions with different environmental conditions.

Exploring further variables could allow for better results along with testing other ensemble machine learning algorithms in order to potentially develop more accurate and robust models. This can include other types of distresses other than pavement cracks

Further investigation can be done by collecting actual data from different pavement sections around Egypt and use it to test the developed models and measure the accuracy of the results.

Lastly, the developed models could be integrated with maintenance and rehabilitation strategies for better planning and allocation of relevant funds and maintenance resources in pavement networks in Egypt.

References

Abdelaziz, N., El-Hakim, R., El-Badawy, S., Afify, H. (2020). International roughness index prediction model for flexible pavements. Int. J. Pav. Eng. 21 (1), 88–99. http://dx.doi.org/10.1080/10298436.2018.1441414

Abu Samra, S., Osman, H., & Hosny, O. (2017). Optimal Maintenance and rehabilitation policies for performance-based road maintenance contracts. *Journal of Performance of Constructed Facilities*, *31*(1). https://doi.org/10.1061/(asce)cf.1943-5509.0000928

Ahmad, T., & Khawaja, H. (2018). Review of low-temperature crack (LTC) developments in asphalt pavements. *The International Journal of Multiphysics*, *12*(2). https://doi.org/10.21152/1750-9548.12.2.169

Alakus, T. B., & Turkoglu, I. (2020). Comparison of deep learning approaches to predict COVID-19 infection. *Chaos, Solitons & Fractals*, *140*, 110120. https://doi.org/10.1016/j.chaos.2020.110120

Ali, A., Heneash, U., Hussein, A., & Eskebi, M. (2022). Predicting pavement condition index using fuzzy logic technique. *Infrastructures*, *7*(7), 91. https://doi.org/10.3390/infrastructures7070091

Alkheder, S. A. (2015). Transportation and Tourism Sustainability in major Jordanian tourism cities. *Tourism Planning & Development*, *13*(3), 253–273. https://doi.org/10.1080/21568316.2015.1074098

AlSaad, R., Malluhi, Q., Janahi, I., & Boughorbel, S. (2022). Predicting emergency department utilization among children with asthma using deep learning models. *Healthcare Analytics*, *2*, 100050. https://doi.org/10.1016/j.health.2022.100050

Ampomah, E. K., Qin, Z., & Nyame, G. (2020). Evaluation of Tree-Based Ensemble Machine Learning Models in Predicting Stock Price Direction of Movement. *Information*, *11*(6), 332. https://doi.org/10.3390/info11060332

Badr, A., Yosri, A., Hassini, S., & El-Dakhakhni, W. (2021). Coupled continuous-time Markov chain–bayesian network model for DAM failure risk prediction. *Journal of Infrastructure Systems*, *27*(4). https://doi.org/10.1061/(asce)is.1943-555x.0000649

Baldo, N., Miani, M., Rondinella, F., Valentin, J., Vackcová, P., & Manthos, E. (2022). Stiffness Data of High-Modulus Asphalt Concretes for Road Pavements: Predictive Modeling by Machine-Learning. *Coatings*, *12*(1), 54. https://doi.org/10.3390/coatings12010054

Banister, D. & Berechman, Y. (2001). Transport Investment and the Promotion of Economic Growth. Journal of Transport Geography. 9. 209-218. 10.1016/S0966-6923(01)00013-8

Bashar, M. Z., & Torres-Machi, C. (2022). Deep learning for estimating pavement roughness using synthetic aperture radar data. *Automation in Construction*, *142*, 104504. https://doi.org/10.1016/j.autcon.2022.104504

Bashar, M. Z., & Torres-Machi, C. (2021). Performance of machine learning algorithms in predicting the Pavement International Roughness Index. *Transportation Research Record: Journal of the Transportation Research Board*, 036119812098617. https://doi.org/10.1177/0361198120986171

Boakye, J., Guidotti, R., Gardoni, P., & Murphy, C. (2022). The role of transportation infrastructure on the impact of natural hazards on communities. *Reliability Engineering & System Safety*, *219*, 108184. https://doi.org/10.1016/j.ress.2021.108184

Bu, K., & Clemente, J. (2022). *Solving Anscombe's Quartet Using a Transfer Learning Approach*. https://doi.org/10.1101/2022.10.12.511920

Burningham, S., & Stankevich, N. (2005). Why road maintenance is important and how to get it done.

Cano-Ortiz, S., Pascual-Muñoz, P., & Castro-Fresno, D. (2022). Machine learning algorithms for monitoring pavement performance. *Automation in Construction*, *139*, 104309. https://doi.org/10.1016/j.autcon.2022.104309

Chan, C.Y., Huang, B., Yan, X. and Richards, S. (2010), Investigating effects of asphalt pavement conditions on traffic accidents in Tennessee based on the pavement management system (PMS). J. Adv. Transp., 44: 150-161. https://doi.org/10.1002/atr.129

Chen, J.-S., Lin, C.-H., Stein, E., & Hothan, J. (2004). Development of a mechanistic-empirical model to characterize rutting in flexible pavements. *Journal of Transportation Engineering*, *130*(4), 519–525. https://doi.org/10.1061/(asce)0733-947x(2004)130:4(519)

Chen, Lingfeng and Hoang, Doan, "Active data-centric framework for data protection in cloud environment" (2012). *ACIS 2012 Proceedings*. 71. https://aisel.aisnet.org/acis2012/71

Chi, S., Murphy, M., Zhang, Z., Sustainable road management in texas: network-level flexible pavement structural condition analysis using data-mining techniques, J. Comput. Civ. Eng. 28 (1) (2014) 156–165, https://doi.org/10.1061/(ASCE)CP.1943-5487.0000252

Damirchilo, F., Hosseini, A., Mellat Parast, M., & 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). https://doi.org/10.1061/jpeodx.0000312

Dayal, V. (2015). Anscombe's Quartet: Graphs Can Reveal. In: An Introduction to R for Quantitative Economics. SpringerBriefs in Economics. Springer, New Delhi. https://doi.org/10.1007/978-81-322-2340-5_9

El-Awady, A., & Ponnambalam, K. (2021). Integration of simulation and Markov chains to support Bayesian networks for probabilistic failure analysis of Complex Systems. *Reliability Engineering & System Safety*, *211*, 107511. https://doi.org/10.1016/j.ress.2021.107511

Ekramnia, T., & Nasimifar, M. (2022). Development of a methodological tool for treatment prioritization in network-level Pavement Management System. *Journal of Transportation Engineering, Part B: Pavements*, *148*(1). https://doi.org/10.1061/jpeodx.0000329

Elhadidy, Amr A., El-Badawy, Sherif M., Elbeltagi, Emad E. (2021). A simplified pavement condition index regression model for pavement evaluation. Int. J. Pav. Eng. 22 (5), 643–652. http://dx.doi.org/10.1080/10298436.2019.1633579

Fang, F., Li, L., Gu, Y., Zhu, H., & Lim, J.-H. (2020). A novel hybrid approach for crack detection. *Pattern Recognition*, *107*, 107474. https://doi.org/10.1016/j.patcog.2020.107474

Fawagreh, K., Gaber, M. M., & Elyan, E. (2014). Random forests: From early developments to recent advancements. *Systems Science & Control Engineering*, *2*(1), 602–609. https://doi.org/10.1080/21642583.2014.956265

Fortney, E. M., Schuldt, S. J., Brown, S. L., Allen, J. P., & Delorit, J. D. (2022). A statistical principal component regression-based approach to modeling the degradative effects of local climate and traffic on Airfield Pavement Performance. *Journal of Transportation Engineering, Part B: Pavements*, *148*(2). https://doi.org/10.1061/jpeodx.0000356

García-Segura, T., Montalbán-Domingo, L., Llopis-Castelló, D., Lepech, M. D., Sanz, M. A., & Pellicer, E. (2020). Incorporating pavement deterioration uncertainty into pavement management optimization. *International Journal of Pavement Engineering*, *23*(6), 2062–2073. https://doi.org/10.1080/10298436.2020.1837827

Gardiner, B., Byrne, K., Hale, S., Kamimura, K., Mitchell, S. J., Peltola, H., & Ruel, J.-C. (2008). A review of mechanistic modelling of wind damage risk to forests. *Forestry*, *81*(3), 447–463. https://doi.org/10.1093/forestry/cpn022

Garí, Y., Monge, D. A., Pacini, E., Mateos, C., & García Garino, C. (2021). Reinforcement learning-based application autoscaling in the cloud: A survey. *Engineering Applications of Artificial Intelligence*, *102*, 104288. https://doi.org/10.1016/j.engappai.2021.104288

Gharieb, M., Nishikawa, T., (2021). Development of roughness prediction models for Laos National Road Network. CivilEng 2 (1), 158–173. http://dx.doi.org/10.3390/civileng2010009

Gong, H., Sun, Y., Shu, X., & Huang, B. (2018). Use of random forests regression for predicting IRI of asphalt pavements. *Construction and Building Materials*, *189*, 890–897. https://doi.org/10.1016/j.conbuildmat.2018.09.017

Gupta, A., Kumar, P., & Rastogi, R. (2013). Critical Review of Flexible Pavement Performance models. *KSCE Journal of Civil Engineering*, *18*(1), 142–148. https://doi.org/10.1007/s12205-014-0255-2

Haas, R., Hudson, W. R., & Falls, L. C. (2015). *Pavement Asset Management*. https://doi.org/10.1002/9781119038849

Haider, S. W., Chatti, K., Baladi, G. Y., & Sivaneswaran, N. (2011). Impact of pavement monitoring frequency on pavement management system decisions. *Transportation Research Record: Journal of the Transportation Research Board*, *2225*(1), 43–55. https://doi.org/10.3141/2225-06

Hall, K. D., Xiao, D. X., & Wang, K. C. (2011). Calibration of the mechanistic–empirical pavement design guidefor flexible pavement design in Arkansas. *Transportation Research Record: Journal of the Transportation Research Board*, *2226*(1), 135–141. https://doi.org/10.3141/2226-15

Hanandeh, S. (2022). Introducing mathematical modeling to estimate Pavement Quality Index of flexible pavements based on genetic algorithm and Artificial Neural Networks. *Case Studies in Construction Materials*, *16*. https://doi.org/10.1016/j.cscm.2022.e00991

Hine, J., & Rutter, J. (2000). Roads, personal mobility & poverty: The challenge.

Hossain, M.I., Gopisetti, L.S.P., Miah, M.S., 2019. International roughness index prediction of flexible pavements using neural networks. J. Transp. Eng. Part B: Pavements 145 (1), 04018058. http://dx.doi.org/10.1061/JPEODX.0000088

Hosseini, S. A., & Smadi, O. (2021). How prediction accuracy can affect the decision-making process in Pavement Management System. *Infrastructures*, *6*(2), 28. https://doi.org/10.3390/infrastructures6020028

Hosseini, S. A., Alhasan, A., & Smadi, O. (2020). Use of deep learning to study modeling deterioration of pavements a case study in Iowa. *Infrastructures*, *5*(11), 95. https://doi.org/10.3390/infrastructures5110095

Hu, A., Bai, Q., Chen, L., Meng, S., Li, Q., & Xu, Z. (2022). A review on empirical methods of Pavement Performance Modeling. *Construction and Building Materials*, *342*, 127968. https://doi.org/10.1016/j.conbuildmat.2022.127968

Jishan, S. T., Rashu, R. I., Haque, N., & Rahman, R. M. (2015). Improving accuracy of students' final grade prediction model using optimal equal width binning and synthetic minority over-sampling technique. *Decision Analytics*, *2*(1). https://doi.org/10.1186/s40165-014-0010-2

Justo-Silva, R., Ferreira, A., & Flintsch, G. (2021). Review on machine learning techniques for developing pavement performance prediction models. *Sustainability*, *13*(9), 5248. https://doi.org/10.3390/su13095248

Kaloop, M. R., El-Badawy, S. M., Hu, J. W., & Abd El-Hakim, R. T. (2023). International roughness index prediction for flexible pavements using novel machine learning techniques. *Engineering Applications of Artificial Intelligence*, *122*, 106007. https://doi.org/10.1016/j.engappai.2023.106007

Ke, X., Lin, J. Y., Fu, C., & Wang, Y. (2020). Transport Infrastructure Development and Economic Growth in China: Recent Evidence from Dynamic Panel System-GMM Analysis. *Sustainability*, *12*(14), 5618. MDPI AG. Retrieved from http://dx.doi.org/10.3390/su12145618

Kheirati, A., Golroo, A. (2022). Machine learning for developing a pavement condition index. Autom. Constr. 139 (March), 104296. http://dx.doi.org/10.1016/j.autcon.2022.104296

Kodippily, S., Henning, T. F., & Ingham, J. M. (2012). Detecting flushing of thin-sprayed seal pavements using pavement management data. *Journal of Transportation Engineering*, *138*(5), 665–673. https://doi.org/10.1061/(asce)te.1943-5436.0000354

Kumar, P. S., K, A. K., Mohapatra, S., Naik, B., Nayak, J., & Mishra, M. (2021). CatBoost Ensemble Approach for Diabetes Risk Prediction at Early Stages. *2021 1st Odisha International Conference on Electrical Power Engineering, Communication and Computing Technology (ODICON)*. https://doi.org/10.1109/odicon50556.2021.9428943

Kumar, V. P., & Sowmya, I. (2021). A review on pros and cons of machine learning algorithms. *Journal of Engineering Sciences*, *12*(10), 272-276

Li, J., Pierce, L. M., & Uhlmeyer, J. (2009). Calibration of flexible pavement in mechanistic–empirical pavement design guide for Washington State. *Transportation Research Record: Journal of the Transportation Research Board*, *2095*(1), 73–83. https://doi.org/10.3141/2095-08

Li, N., Haas, R., & Xie, W.-C. (1997). Investigation of relationship between deterministic and probabilistic prediction models in Pavement Management. *Transportation Research Record: Journal of the Transportation Research Board*, *1592*(1), 70–79. https://doi.org/10.3141/1592-09

Litman, T. (2017). *Evaluating public transit criticism: Systematic analysis of political attacks on high quality transit, and how transportation professionals can effectively respond*. Victoria Transport Policy Institute.

Liu, F., Liu, J., & Wang, L. (2022). Deep learning and infrared thermography for asphalt pavement crack severity classification. *Automation in Construction*, *140*, 104383. https://doi.org/10.1016/j.autcon.2022.104383

Majidifard, H., Adu-Gyamfi, Y., & Buttlar, W. G. (2020). Deep Machine Learning Approach to develop a new asphalt pavement condition index. *Construction and Building Materials*, *247*, 118513. https://doi.org/10.1016/j.conbuildmat.2020.118513

Mamlouk, M., & Zapata, C. E. (2010). Necessary assessment of use of state pavement management system data in mechanistic–empirical pavement design guide calibration process. *Transportation Research Record: Journal of the Transportation Research Board*, *2153*(1), 58–66. https://doi.org/10.3141/2153-07

Mancini, A., Vito, L., Marcelli, E., Piangerelli, M., De Leone, R., Pucciarelli, S., & Merelli, E. (2020). Machine learning models predicting multidrug resistant urinary tract infections using "DsaaS." *BMC Bioinformatics*, *21*(S10). https://doi.org/10.1186/s12859-020-03566-7

Mazari, M., Rodriguez, D. (2016). Prediction of pavement roughness using a hybrid gene expression programming-neural network technique. J. Traffic Transp. Eng. (Engl. Ed.) 3 (5), 448–455. http://dx.doi.org/10.1016/j.jtte.2016.09.007

Mehdiyev, N., Enke, D., Fettke, P., & Loos, P. (2016). Evaluating forecasting methods by considering different accuracy measures. *Procedia Computer Science*, *95*, 264–271. https://doi.org/10.1016/j.procs.2016.09.332

Milad, A., Adwan, I., Majeed, S. A., Yusoff, N. I., Al-Ansari, N., & Yaseen, Z. M. (2021). Emerging Technologies of Deep Learning Models Development for Pavement Temperature Prediction. *IEEE Access*, *9*, 23840–23849. https://doi.org/10.1109/access.2021.3056568

Mills, L. N., Attoh-Okine, N. O., & McNeil, S. (2012). Hierarchical markov chain Monte Carlo simulation for modeling transverse cracks in highway pavements. *Journal of Transportation Engineering*, *138*(6), 700–705. https://doi.org/10.1061/(asce)te.1943-5436.0000383

Montomoli, J., Romeo, L., Moccia, S., Bernardini, M., Migliorelli, L., Berardini, D., Donati, A., Carsetti, A., Bocci, M. G., Wendel Garcia, P. D., Fumeaux, T., Guerci, P., Schüpbach, R. A., Ince, C., Frontoni, E., Hilty, M. P., Alfaro-Farias, M., Vizmanos-Lamotte, G., Tschoellitsch, T., … Colak, E. (2021). Machine learning using the extreme gradient boosting (XGBoost) algorithm predicts 5-day delta of SOFA score at ICU admission in COVID-19 patients. *Journal of Intensive Medicine*, *1*(2), 110–116. https://doi.org/10.1016/j.jointm.2021.09.002

Mulungye, R. M., Owende, P. M. O., & Mellon, K. (2007). Finite element modelling of flexible pavements on soft soil subgrades. *Materials & Design*, *28*(3), 739–756. https://doi.org/10.1016/j.matdes.2005.12.006

Nagalla, R., Pothuganti, P., & Pawar, D. S. (2017). Analyzing gap acceptance behavior at unsignalized intersections using support vector machines, decision tree and random forests. *Procedia Computer Science*, *109*, 474–481. https://doi.org/10.1016/j.procs.2017.05.312

Nguyen, G., Dlugolinsky, S., Bobák, M., Tran, V., López García, Á., Heredia, I., Malík, P., & Hluchý, L. (2019). Machine learning and deep learning frameworks and libraries for large-scale data mining: A survey. *Artificial Intelligence Review*, *52*(1), 77–124. https://doi.org/10.1007/s10462-018-09679-z

Nguyen, H., Pham, T., Son, L., Thang, T., Ly, H., Le, T., Ho, S., Le, T., Bui, T. (2019). Adaptive network based fuzzy inference system with meta-heuristic optimizations for international roughness index prediction. Appl. Sci. 9 (21), 4715. http://dx.doi.org/10.3390/app9214715

Nhat-Duc, H., & Van-Duc, T. (2023). Comparison of histogram-based gradient boosting classification machine, random forest, and deep convolutional neural network for Pavement Raveling Severity

Classification. *Automation in Construction*, *148*, 104767. https://doi.org/10.1016/j.autcon.2023.104767

Niazi, M., Nietch, C., Maghrebi, M., Jackson, N., Bennett, B. R., Tryby, M., & Massoudieh, A. (2017). Storm water management model: Performance review and Gap Analysis. *Journal of Sustainable Water in the Built Environment*, *3*(2). https://doi.org/10.1061/jswbay.0000817

OBrien, Eugene J., Taheri, Abdolrahim, Malekjafarian, Abdollah, 2018. An alternative roughness index to IRI for flexible pavements. Can. J. Civil Eng. 45 (8), 659–666. http://dx.doi.org/10.1139/cjce-2017-0443

Papadimitroulas, P., Brocki, L., Christopher Chung, N., Marchadour, W., Vermet, F., Gaubert, L., Eleftheriadis, V., Plachouris, D., Visvikis, D., Kagadis, G. C., & Hatt, M. (2021). Artificial Intelligence: Deep learning in oncological radiomics and challenges of interpretability and Data Harmonization. *Physica Medica*, *83*, 108–121. https://doi.org/10.1016/j.ejmp.2021.03.009

Pérez-Acebo, H., Linares-Unamunzaga, A., Rojí, E., Gonzalo-Orden, H. (2020). IRI performance models for flexible pavements in two-lane roads until first maintenance and/or rehabilitation work. Coatings 10 (2), 97. http://dx.doi.org/10.3390/coatings10020097

Pierce, L. M., McGovern, G., Zimmerman, K. A., & Van, T. (2013). Network-Level Condition Data Analysis. In *Practical Guide for Quality Management of Pavement Condition Data Collection* (pp. 13–14). essay, U.S. Dept. of Transportation, Federal Highway Administration.

Piryonesi, S. M., & El-Diraby, T. E. (2020). Data Analytics in asset management: Cost-effective prediction of the pavement condition index. *Journal of Infrastructure Systems*, *26*(1). https://doi.org/10.1061/(asce)is.1943-555x.0000512

Piryonesi, S.M., El-Diraby, T.E., 2018. Using Data Analytics for Cost-Effective Prediction of Road Conditions: Case of the Pavement Condition Index, Federal Highway Administration, No. FHWA-HRT-18-065. Washington DC

Plevris, V., Solorzano, G., Bakas, N., & Ben Seghier, M. (2022). Investigation of performance metrics in regression analysis and machine learning-based prediction models. *8th European Congress on Computational Methods in Applied Sciences and Engineering*. https://doi.org/10.23967/eccomas.2022.155

Pucher, J., & Dijkstra, L. (2003). Promoting safe walking and cycling to improve public health: Lessons from the netherlands and germany. *American Journal of Public Health (1971), 93*(9), 1509-1516. https://doi.org/10.2105/AJPH.93.9.1509

RADWAN, M., ABO-HASHEMA, M., HASHEM, M., & FAHEEM, H. (2020). Modeling Pavement Performance Based on LTPP database for flexible pavements. *Teknik Dergi*, *31*(4), 10127–10146. https://doi.org/10.18400/tekderg.476606

Rezapour, M., Hafez, M., & Ksaibati, K. (2022). Evaluating the complex relationship between environmental factors and pavement friction based on long-term pavement performance. *Computation*, *10*(6), 85. https://doi.org/10.3390/computation10060085

Rose, S., Mathew, B. S., Isaac, K. P., & Abhaya, A. S. (2016). Risk based probabilistic pavement deterioration prediction models for low volume roads. *International Journal of Pavement Engineering*, *19*(1), 88–97. https://doi.org/10.1080/10298436.2016.1162308

Saad, B., Mitri, H., & Poorooshasb, H. (2005). Three-dimensional dynamic analysis of flexible conventional Pavement Foundation. *Journal of Transportation Engineering*, *131*(6), 460–469. https://doi.org/10.1061/(asce)0733-947x(2005)131:6(460)

Saha, J., Nassiri, S., Bayat, A., & Soleymani, H. (2012). Evaluation of the effects of Canadian climate conditions on the MEPDG predictions for flexible pavement performance. *International Journal of Pavement Engineering*, *15*(5), 392–401. https://doi.org/10.1080/10298436.2012.752488

Sandamal, K., Shashiprabha, S., Muttil, N., & Rathnayake, U. (2023). Pavement roughness prediction using explainable and supervised machine learning technique for long-term performance. *Sustainability*, *15*(12), 9617. https://doi.org/10.3390/su15129617

Sharma, M., & Kumar, P. (2022). Assessment of present pavement condition using machine learning techniques. *Lecture Notes in Civil Engineering*, 71–82. https://doi.org/10.1007/978-3-030-87379-0_5

Sholevar, N., Golroo, A., Sahand Esfahani, R. (2022). Machine learning techniques for pavement condition evaluation. Autom. Constr. 136 (March), 104190. http://dx.doi.org/10.1016/j.autcon.2022.104190

Stoner, A. M., Daniel, J. S., Jacobs, J. M., Hayhoe, K., & Scott-Fleming, I. (2019). Quantifying the impact of climate change on flexible pavement performance and lifetime in the United States. *Transportation Research Record: Journal of the Transportation Research Board*, *2673*(1), 110–122. https://doi.org/10.1177/0361198118821877

Streimikiene, D., Baležentis, T., & Baležentienė, L. (2013). Comparative assessment of road transport technologies. *Renewable and Sustainable Energy Reviews*, *20*, 611–618. https://doi.org/10.1016/j.rser.2012.12.021

Swei, O., Gregory, J., & Kirchain, R. (2019). Embedding flexibility within pavement management: Technique to improve expected performance of roadway systems. *Journal of Infrastructure Systems*, *25*(3). https://doi.org/10.1061/(asce)is.1943-555x.0000496

Tantalaki, N., Souravlas, S., & Roumeliotis, M. (2019). Data-driven decision making in precision agriculture: The rise of big data in Agricultural Systems. *Journal of Agricultural & Food Information*, *20*(4), 344–380. https://doi.org/10.1080/10496505.2019.1638264

Tatbul, N., Lee, T. J., Zdonik, S., Alam, M., & Gottschlich, J. (2018). Precision and recall for time series. *Advances in neural information processing systems*, *31*.

Terzi, S. (2013). Modeling for pavement roughness using the ANFIS approach. Adv. Eng. Softw. 57 (March), 59–64. http://dx.doi.org/10.1016/j.advengsoft.2012.11.013

Weber, P., Medina-Oliva, G., Simon, C., & Iung, B. (2012). Overview on bayesian networks applications for dependability, risk analysis and maintenance areas. *Engineering Applications of Artificial Intelligence*, *25*(4), 671–682. https://doi.org/10.1016/j.engappai.2010.06.002

Zeiada, W., Dabous, S. A., Hamad, K., Al-Ruzouq, R., & Khalil, M. A. (2020). Machine Learning for Pavement Performance Modelling in warm climate regions. *Arabian Journal for Science and Engineering*, *45*(5), 4091–4109. https://doi.org/10.1007/s13369-020-04398-6

Zhang, C., Liu, Y., & Tie, N. (2023). Forest Land Resource Information acquisition with sentinel-2 image utilizing support vector machine, K-nearest neighbor, random forest, decision trees and Multi-Layer Perceptron. *Forests*, *14*(2), 254. https://doi.org/10.3390/f14020254

Zhang, C., Nateghinia, E., Miranda-Moreno, L. F., & Sun, L. (2022). Pavement distress detection using Convolutional Neural Network (CNN): A case study in Montreal, Canada. *International Journal of Transportation Science and Technology*, *11*(2), 298–309. https://doi.org/10.1016/j.ijtst.2021.04.008

Zhu, J., & Wang, Y. (2021). Feature selection and deep learning for deterioration prediction of the Bridges. *Journal of Performance of Constructed Facilities*, *35*(6). https://doi.org/10.1061/(asce)cf.1943-5509.0001653

Ziari, H., Sobhani, J., Ayoubinejad, J., Hartmann, T. (2016). Prediction of IRI in short and long terms for flexible pavements: ANN and GMDH methods Int. J. Pav. Eng. 17 (9), 776–788. http://dx.doi.org/10.1080/10298436.2015.1019498

Zimmerman, K. A., Pierce, L. M., & Krstulovich, J. (2010). *Pavement management roadmap*. U.S. Dept. of Transportation, Federal Highway Administration.

Zimmermann, T., Nagappan, N., Gall, H., Giger, E., & Murphy, B. (2009). Cross-project defect prediction. Proceedings of the 7th Joint Meeting of the European Software Engineering Conference and the ACM SIGSOFT Symposium on The Foundations of Software Engineering. https://doi.org/10.1145/1595696.1595713