THESIS

INTEGRATING ENVIRONMENTAL CONDITIONS INTO MACHINE LEARNING MODELS FOR PREDICTING BRIDGE MAINTENANCE DETERIORATION

Submitted by

Okohene Papa Ansah

Department of Construction Management

In partial fulfillment of the requirements

For the Degree of Master of Science

Colorado State University

Fort Collins, Colorado

Summer 2025

Master's Committee:

Advisor: Mehmet E. Ozbek
Co-Advisor: Svetlana Olbina

Sarath Sreedharan

ABSTRACT

INTEGRATING ENVIRONMENTAL CONDITIONS INTO MACHINE LEARNING MODELS FOR PREDICTING BRIDGE MAINTENANCE DETERIORATION

Bridge management agencies face mounting pressure to maintain ageing infrastructure amidst intensifying climate extremes and limited budgets. This study develops and tests machine learning (ML) models integrating environmental data with traditional structural and inspection records to predict condition deterioration of Colorado's National Highway System bridges. An extensive database of 75,063 bridge-year observations and 97 features including deck, superstructure, and substructure condition ratings, traffic loads, freeze-thaw cycles, precipitation, temperature extremes, and humidity was compiled from the National Bridge Inventory (NBI) and PRISM climate archives (2014-2024). Decision Tree (DT), Random Forest (RF), and Gradient Boosting (GB) classifiers were trained using SMOTETomek resampling and optimized using cross-validated grid search. GB attained the most consistent top performance, achieving balanced accuracies between 0.972 and 0.978 and Matthews Correlation Coefficient (MCC) values around 0.97 across all bridge components. DTs achieved the single highest metric value for decks (balanced accuracy = 0.9875). The addition of climatic variables significantly improved performance over the baselines which has only structural variables in balanced accuracy and macro F1-scores ranging from 3.0 to 4.5 percentage points. This translated into a drop in the relative error rate by 6-15% for deterioration state forecasting. Feature importance analysis invariably revealed frequency of freeze-thaw cycles, annual rainfall, extreme temperatures, along with sufficiency rating, age, and traffic by trucks to be key predictors. A cost-benefit analysis by simulation indicated optimal allocation using these climate-aided models to generate up to 11.7% savings in life-cycle costs over a span of 30 years. The models resulting from this integration enable more accurate, risk-oriented inspection schedules and data-driven budgets while taking regional climatic stressors into account. This work proposes a structured yet flexible approach to improving climatic resilience in bridge maintenance planning

Keywords: *Machine Learning models, Decision Tree, Random Forest, Gradient Boosting, Bridge Maintenance, Bridge deterioration, Prediction*

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER ONE: INTRODUCTION

1.1 Background

Bridge infrastructure is a critical component of transportation networks that support the transit of people, goods and services. Such base infrastructure is the backbone of economic activities, connecting cities, industries, and rural areas helping to drive not just regional development but also steadying the national economy. Bridges also play an important role in reducing the cost of transport, improving accessibility, and enabling trade and commerce between regions (Wang et al., 2023).

Bridges deteriorate as a result of both loads and time, with the former including those due to various types of vehicles; the latter includes but is not limited to the effects of material aging and environmental conditions such as variations in weather (Srikanth & Arockiasamy, 2020). Heavy traffic, especially freight traffic, places constant loading on structural elements resulting in capacity reduction due to fatigue and damage accumulation (Dinegdae & Birgisson, 2016). Environmental conditions such as temperature fluctuations, moisture ingress and chemical agents (e.g. the use of deicing salts), all cause deterioration of bridge materials in particular concrete and steel (Fujiu et al., 2022). For example, moisture containing chloride ions from deicing salts accelerates steel reinforcing corrosion, while freeze–thaw cycles are responsible for surface cracking and spalling of concrete (Fujiu et al., 2022). In this manner, it impacts the structural integrity of bridges and underlines the necessity for routine maintenance of safety and prevention of catastrophic failures.

Traditionally, bridge maintenance planning has been based on a mix of periodic inspections and expert opinion as well as schedules defined by bridge age or amount of traffic load (FHWA, 2010). The Federal Highway Administration (FHWA) set national standards for bridge inspection, calling for periodic visual inspections and more intensive assessment of areas that appear to have potential problems. Although a visual inspection gives some idea as to the visible condition of bridge components, it is not capable of predicting internal and future structural deterioration. Additionally, manual inspections are resource-intensive, time-consuming and subjective which makes them inconsistent in output and poor from maintenance planning perspective (Graybeal et al., 2002).

Due to the limitations, there is a growing trend in bridge maintenance towards data-driven approaches, especially machine learning (ML) techniques for predictive modeling of bridge deterioration. The ability of ML to analyze large amounts of historical data (e.g., inspection records, structural parameters and traffic loads) allows to find patterns and trends that are not easily observable in a manual inspection (Jaafaru & Agbelie, 2022). Different ML models such as decision trees (DT), random forests (RF), artificial neural networks (ANNs) and deep learning algorithms have been utilized to predict the future state of bridge elements and schedule maintenance appropriately. Such ML algorithms have been capable to analyze large data and discover the latent relationship between many different factors that influence deterioration indirectly, making accurate predictions rather than traditional methods (Liu & El-Gohary, 2022).

The growing availability of data from smart sensors, remote sensing technologies and infrastructure monitoring systems further promoted the application of ML in bridge management (Guan et al., 2018). These technologies gather real-time data on structural health, environmental conditions, and traffic patterns that can be incorporated in ML models to improve the accuracy of predictions. For example, bridge–embedded sensors can monitor the strain, vibration and displacement in real time and remote sensing based on satellites can detect temperature and humidity. This combination of these data provides a real-time adaptive strategy for bridge maintenance.

Despite substantial progress in developing ML models to predict bridge deterioration, a lack of understanding regarding the effect of environmental conditions on the rate of deterioration means that there is still an important gap remaining. Today, available ML models primarily consider structural, inspection data and as well as traffic data such as the Average Annual Daily Traffic (AADT) which helps to capitalize on vehicle loads induced stress in bridge structures. While AADT data is certainly helpful to understand how traffic affects bridges, they typically ignore key environmental variables that have a large impact on the bridge service life. It is obvious that environmental conditions like temperature, precipitations, humidity, and contact with aggressive agents accelerate or delay deterioration processes. For example, bridges in the coastal areas are exposed to saltwater corrosion, while bridges in colder regions are affected by freeze-thaw action (Chyad & Abudayyeh, 2021).

Integrating environmental data into ML models can yield a more secure prediction in the maintenance of these bridge systems. As a result, the ML model is able to provide information on which component of the infrastructure is contributing most to the deterioration rates over time. Studies such as those by Elleathy et al. (2024) utilized an evolutionary computing based framework which has proven that including external environment data to bridge deterioration models significantly improves their predictive performance.

1.2 Problem Statement

Most ML models today in bridge maintenance work on structural parameters and inspection data while treating environmental factors either as a secondary consideration or they are indeed left out of the model completely (Jaafaru & Agbelie, 2022; Liu & El-Gohary, 2020; Srikanth & Arockiasamy, 2020). This method's lack of generalizability limits its application to structures like bridges in regions with intense or diverse climates. Bridges in coastal locations may be afflicted with more intense corrosion due to saltwater exposure, while cold climate bridges are at a high risk for repetitive freeze-thaw damage; conditions poorly captured in structural and traffic-based models. In turn, incomplete information often forces bridge owners and maintenance planners to make poor decisions, which leads to poor maintenance planning and increased long-term costs.

Moreover, adverse effects of climate change will most likely increase the environmental loads on bridges; hence, such conditions should be factored into predictive models (Kumar & Kota, 2024). Increasing temperature, precipitation, and more extreme weather events will increase the stress on bridges, making a comprehensive deterioration modeling approach necessary.

Due to the increasing focus on data-driven decision-making in infrastructure management, it is crucial to improve ML models so they can learn from environmental data more effectively. This will give a comprehensive view of the factors leading to the deterioration of bridges resulting into an integrated prediction and reliable maintenance stratagem. This integration will not only enhance the reliability of predictions over diverse climatic regions but also allow one to develop region-specific maintenance protocols addressing both structural and environmental vulnerabilities.

Based on all of these, the objective of this study is to fill the gap of the integration of environmental conditions into ML models for prediction of bridge deterioration. It is crucial to fill this gap to achieve long-term safety, durability, and sustainability in infrastructure at reasonable life cycle costs. Including environmental factors such as temperature, humidity and precipitation into prediction models will provide predictive tools that are more accurate in predicting bridge conditions leading to optimized resource allocation and better maintenance decisions.

1.3 Aim and Objectives

The overarching aim of this study is integrating environmental conditions into ML models for the prediction of bridge deterioration. More specifically, the study has the following objectives:

1. Identify the key environmental factors responsible for deterioration of bridges.
2. Develop a machine learning model that incorporates environmental conditions along with structural and inspection data.
3. Assess if the integrated model provides an enhanced prediction beyond those models that do not include environmental data.

1.4 Research Questions

The following questions will govern the research in order to achieve these objectives:

1. What are the major environmental factors responsible for bridge deterioration?
2. Which ML modelling framework most effectively integrates environmental, structural, and inspection data?
3. Compared to conventional models, does the predictive performance of bridge deterioration ML models improve by considering environmental conditions?

1.5 Significance of the Study

The significance of this research lies in enhancing bridge infrastructure management by focusing on making the existing bridges safer, more reliable and more sustainable. Bridges are at the heart of a transportation system and their failure can be catastrophic for both public safety and economic well-being. According to the 2025 Report Card for America’s Infrastructure published by the American Society of Civil Engineers, U.S. bridges earned a grade of “C”, underscoring persistent performance gaps and the urgent need for more proactive, data-informed maintenance strategies (ASCE, 2025). Traditional maintenance strategies do not consider the impact of environmental conditions on bridges deteriorating characteristics, leading to a poor characterization of maintenance actions and subsequent an inefficient resource assignment.

One key aspect highlighted in this research is the analysis of factors, like temperature changes, humidity levels and precipitation levels to be included in predictive models. Literature underlines that these factors strongly enhance deterioration processes such as corrosion and material fatigue.

This study further promotes ML methodologies in civil infrastructure management. Compared to traditional statistical methods, ML algorithms have better performances than traditional statistical methods when dealing with complicated datasets possessing hidden structures. The incorporation of environmental variables into predictive analytics allows for the adoption of proactive maintenance strategies, in contrast to merely reactive or inspection-based methodologies. Predictive modeling is used to improve early detection of deterioration, thus optimizing planning of maintenance and reducing both the costs of repairs and operational downtime. The proactive framework will result in better safety, fewer disruptions, and huge cost savings for transportation agencies.

The research adds to the increasing focus on sustainability within civil engineering. Moreover, this study promotes the goal of sustainability as it addresses the issues of climate change together with resilient infrastructure development. Besides the technical significance of this research, it also adds to the growing body of academic work focused on ML applications in civil engineering. These results also form a basis for infrastructure management, which can be extended to roadways and tunnels. In addition, worldwide implementation of the methodology described in this study can be considered, which would facilitate standardization and push towards a paradigm shift to data-driven maintenance. This has great potential to change our approach to bridge maintenance.

1.6 Scope

The goal of this research is to develop a ML model predicting the deterioration of bridges in the state of Colorado using a combination of environmental factors (e.g., humidity, precipitation, and temperature) and traditional structural and inspection data. Data will be obtained from the National Bridge Inventory (NBI) and as well as the Parameter-elevation Regressions on Independent Slopes Model (PRISM) Climate Group datasets for the past 11 years. The analysis will focus on the bridges that are in the National Highway System (NHS) and evaluate the performance of a variety of ML methods. The proposed study will be limited to Colorado, will not include non-NHS bridges, will not involve real-time monitoring, non-environmental variables, or private datasets, and will focus on publicly accessible data and environmental variables only in Colorado.

CHAPTER TWO: LITERATURE REVIEW

Incorporating environmental conditions into ML models for predicting bridge deterioration is a major advance for infrastructure management. Unlike traditional models based on periodic surveys and historical data, the consideration of some environmental variables can help achieve more precise deterioration predictions that lead to a proactive maintenance strategy. In this chapter, a literature review on the factors affecting bridge deterioration, maintenance and data collection practices in practice as well as predictive modelling techniques is carried out. It traces the journey from traditional to ML-based methods, showing the need for integrating environmental conditions and their impacts over time in a coherent manner while also highlighting its scalability constraints.

2.1 Bridge Deterioration Factors

Bridges are at the heart of transportation infrastructure, enabling connectivity and supporting both economic and social activities. Bridges, like all other engineering infrastructures, are prone to deterioration with age, which poses serious safety and functional issues that impact maintenance planning. The deterioration process in bridges may lead to structural failure, increased cost for maintenance, and disruptions in transportation services; therefore, it is important to understand those factors that contribute to degradation (Wu et al., 2021).

Those factors that potentially influence bridge deterioration can be classified into three groups: structural, operational, and environmental. Structural factors include the consequences of material degradation, design limitation, and quality of construction, which all affect the durability and lifespan of bridges. An example is the degradation of concrete and steel, involving a gradual loss in mechanical properties where the aged materials have more tendency to have cracks and corrosion with time (Kim & Song, 2021; Nie et al., 2022). Moreover, design shortcomings, such as low load-carrying capacity or insufficient detailing further deteriorate structures under service conditions (Sgambi et al., 2012). Construction quality also has a great role; poor methods of construction can bring on defects that reduce the structural integrity of the bridge (Nettis et al., 2023).

Operational factors, including traffic loads, vibration, and dynamic stresses caused by heavy vehicles, significantly accelerate wear and fatigue, particularly for bridges with heavy utilization. Bridge components may be damaged by fatigue due to traffic-induced vibrations, particularly in steel structures where repeated cyclic loading initiates and propagates micro-cracks over time (Chen et al., 2012;Wang & Chen, 2012). The dynamic response of bridges under traffic loads is a complex process and is affected by many factors such as vehicle types, speeds, and loading patterns (Anitori et al., 2018; Zhao et al., 2023). More importantly, research has shown that the natural frequencies of bridges can fluctuate much under different traffic states, which indicates that the dynamic effects of traffic should be included in maintenance planning (Ye et al., 2016). Also, the added influx of overweight vehicles crossing the highways imposes additional risks since such vehicles can impose loads greater than those for which many bridges were originally designed (Zandonini et al., 2013).

Environmental factors are increasingly recognized as key determinants for the deterioration of bridges. Critical environmental conditions that contribute to these distresses include temperature, freeze-thaw cycles, humidity, rain, and contact with de-icing salts (Haynes et al., 2019; Pak & Feldmann, 2016). Such is in the case with freeze-thaw cycles, which can lead in expansion and contraction in materials, further leading to concrete surface cracking and spalling (Haynes et al., 2019). Similarly, the application of de-icing salts will accelerate the corrosion of steel reinforcements in concrete, significantly reducing the lifespan of bridge structures (Kim & Song, 2021; Nie et al., 2022). The interaction between environmental conditions and structural materials is quite complex, and understanding these interactions is very important for developing effective maintenance strategies.

It is important to emphasize that bridge deterioration is mostly brought about by an interaction of many factors rather than a single factor. This interaction among the structural, operational, and environmental elements poses a complex problem that needs a holistic approach towards assessment and maintenance. For example, a bridge under heavy traffic loads may be subject to faster deterioration in structural performance when located in a corrosive environment compared to the same structure under less onerous conditions (Kim & Song, 2021; Nie et al., 2022). The different environmental conditions due to geographical location and climatic effects make the deterioration patterns more complicated; hence, more tailored maintenance is needed (Sacconi et al., 2021).

2.2 Maintenance Strategies in Bridge Infrastructure

Bridge infrastructure maintenance is very important in the field of civil engineering for aging structures and increasing traffic loads. Bridges are therefore exposed to deterioration forces that are many-sided in nature, hence their maintenance requires a strong approach that will enable them to maintain their functionality and safety and extend their lifespan, thereby optimizing resource allocation and reducing costly interventions. The American Society of Civil Engineers estimates that about $123 billion in investment in maintenance strategies will be needed to clear the nation's deficient bridge backlog (Liu & El-Gohary, 2020).Therefore, this review aims to consider the three major approaches of maintenance strategies: preventive, predictive, and rehabilitative in the management of bridge deterioration.

Preventive maintenance is routine, scheduled interventions designed to address potential issues before they become critical. These activities may be sealing cracks, protective coatings, or making minor repairs. This is where preventive maintenance is very strong because it acts in a proactive manner, which significantly reduces the chance of serious deterioration and extends the service life of bridge components (Kale et al., 2023a). However, the success rate of preventive maintenance becomes limited while trying to address those unexpected issues popping up through the dynamic environmental conditions like extreme weather events or unpredictable loading scenarios (Liu et al., 2022a). For example, Liu et al., (2022) points out that environmental factors lead to deterioration in bridges and hence require a strategy for maintenance by taking into consideration the localized environmental risks.

Moving to predictive maintenance, this strategy uses data-driven techniques to anticipate deterioration trends and schedule interventions. Predictive maintenance leverages state-of-the-art technologies such as ML algorithms to provide analysis of historical performance data, operational loads and environmental conditions (Jiang et al., 2023a). This approach allows one to make more accurate predictions regarding when the maintenance would take place, hence saving resources and reducing downtime (Gopalakrishnan et al., 2022). For instance, Gopalakrishnan et al. (2022) discusses how data-driven approaches can be used to improve the activities of maintenance decision-making, and productivity while reducing costs.

Rehabilitative maintenance responds to advanced deterioration when major repairs or structural replacements are required. Although rehabilitative strategies are necessary to restore the functionality of bridges, they are normally resource-consuming and may cause tremendous disruptions in traffic and transportation networks (Kim et al., 2020a). The importance of including predictive maintenance in the overall framework of environmental condition monitoring arises from the need for rehabilitative maintenance because proactive interventions can be used to delay or reduce the needs of costly rehabilitative measures (Patel & Joshi, 2023). For example, Shen et al. (2023) presents a bridge network maintenance decision-making method that incorporates environmental factors into the decision-making process for improving the effectiveness of the maintenance strategies. Thus, predictive analytics can help the agencies to prioritize interventions based on the unique environmental and operational challenges of each bridge, leading to more sustainable maintenance practices.

2.3 Standards and Guidelines for Bridge Inspections

Regular inspections are necessary to determine structural integrity, early signs of deterioration, and maintenance strategies. These are the basis for bridge management systems, in which derived data is fundamental to decision-making processes. National Bridge Inspection Standards (NBIS) and FHWA provide wide guidance on the practices of bridge inspection within the United States, mandating frequent inspections at two-year intervals, plus procedures for the evaluation, condition rating, and maintenance needs prioritization of structural components (Abdallah et al., 2021; Mia & Kameshwar, 2023b; Zhang et al., 2023).

The NBIS and FHWA guidelines were created in order to establish a standardized approach to bridge inspection across regions, so that important structural problems can be determined and corrected in a timely manner. Under these guidelines, the minimum requirement is that all public road bridges having at least a 20-foot span must be inspected at least once every two years, with data on 116 NBI items reported annually (Mia & Kameshwar, 2023b; Zhang et al., 2023). These guidelines ensure a systematic approach to assessing bridge conditions, which is paramount for public safety and infrastructure reliability. But periodic visual inspections alone are far from perfect. These inspections are good at identifying gross structural deficiencies, but many subtle or long-term deterioration processes are missed and can be aggravated by environmental conditions (Abdallah et al., 2021; Ahlborn et al., 2012; Kong & Li, 2018).

Conventional techniques of inspection are based mostly on visual examination, which may be subjective and prone to human error. Temperature fluctuations, freeze-thaw cycles, and exposure to corrosive agents may seriously degrade the materials of every bridge. However, the effects of these factors are not well captured using conventional inspection methods (Abdallah et al., 2021; Ahlborn et al., 2012; Kong & Li, 2018). Moreover, there is a growing awareness of developing new inspection methodologies by integrating the environmental data into condition assessments. Such integration would provide a better insight into how environmental factors affect deterioration and thus enable more realistic predictions of the maintenance strategies (Abdallah et al., 2021; Heo, 2020).

The advancements in non-destructive testing (NDT) methods and also sensor-based monitoring technologies look very promising for enhancing the traditional inspection methodologies. Such technologies may provide real-time data on bridge conditions, enabling more detailed assessment results to complement the traditional periodic visual inspection (Abdallah et al., 2021; Heo, 2020). For example, unmanned aerial vehicles (UAVs) with high-resolution cameras and sensors can be used to inspect areas that are difficult to access, and obtain information that otherwise would have been missed during a regular inspection (Azari et al., 2022; Congress et al., 2024). Integration of Bridge Information Modeling (BIM) with inspection data will also contribute to the better management of defect information, which will allow more effective tracking and prioritization of maintenance needs (Zhang, et al., 2023; Mohamed et al., 2023).

2.4 Environmental Conditions Affecting Bridge Deterioration

Environmental conditions contribute to a major part of bridge structure deterioration, which greatly affects their lifespan and structural integrity. Interactions between the environmental and the inherent stresses in bridges can result in accelerated material degradation, thus leading to great challenges in maintenance and management practices. Knowledge of this dynamics is therefore very essential in the creation of an effective predictive model for proactive maintenance. This section synthesizes the existing research on environmental conditions affecting bridge deterioration and points out the need to incorporate these factors into ML models so that the predictions will be more accurate.

One major environmental factor in bridge deterioration is temperature changes. Severe temperature changes may cause thermal expansion and contraction in bridge materials that can lead to cracking and spalling, mainly in regions subjected to big seasonal variations.

Studies show that it is known that freeze-thaw cycles, which occur much more in colder climates, enhance these impacts by making materials degrade more with the cyclic stresses acted upon them (Fom et al., 2015a; Kallias & Imam, 2013). Materials in different temperatures show expansion and contraction, creating stress concentrations that weaken structural components over time and finally compromise the integrity of the bridge (Feng et al., 2023; Nguyen et al., 2017).

Along with temperature, humidity also plays a role in the deterioration process. Higher the humidity, greater the exposure to moisture, and thus corrosion in steel members and weakening of concrete advance at a much faster rate. Research has indicated that moisture not only accelerates the electrochemical reactions that cause corrosion but also affects the rusting of steel reinforcement in reinforced concrete (Li et al., 2022a; Zaki et al., 2017). The interaction between humidity and temperature, therefore, has produced a corrosive ambient environment that further complicates the maintenance management of bridge infrastructures (Garg et al., 2021; Stewart et al., 2012).

Water ingress and precipitation are also key contributors to bridge deterioration. Excessive rain contributes to foundation scouring and destabilizing structures, especially in flood regions. Water infiltration in bridge concrete and structural steel can accelerate the corrosion process, and the deterioration of the materials (Obata et al., 2014). In addition, the accumulation of water can promote biological growth that cause loss of structural integrity of bridges (Wang, et al., 2023; Rizzo & Enshaeian, 2021).

Another important factor is exposure to corrosive agents, which includes chlorides and de-icing salts. In Colorado, where de-icing salts are normally used throughout winter months, the likelihood of corrosion increases significantly. Chlorides may seep through concrete and initiate the onset of corrosion of steel reinforcements, leading to spalling and cracking of the concrete cover (Gao, 2020). This can especially be true in the presence of chlorides at high humidities and temperatures, forming a very corrosive environment that causes further deterioration of the components (Li et al., 2016; Soliman & Frangopol, 2015). This could be more dangerous for bridges located in coastal areas where the salt spray accelerates the processes related to corrosion (Kallias & Imam, 2013; Obata et al., 2014).

Long-term exposure to adverse environmental conditions can cause cumulative effects that result in the progressive weakening of the structural components. Over time, such a combined effect of temperature changes, humidity, precipitation, and corroding agents can significantly shorten the service life of bridges. Research indicates that these deteriorating effects are mostly nonlinear; hence, the damage produced by an individual environmental exposure is cumulative (Bastidas-Arteaga, 2018; Feng et al., 2023). This underlines the importance of continuous monitoring and assessment of environmental conditions in informing maintenance strategies to extend the life span of bridge infrastructures (Li et al., 2022b; Rizzo & Enshaeian, 2021).

The above-discussed environmental factors, i.e., temperature variations, freeze-thaw cycles, humidity, precipitation, and chlorides exposure due to de-icing, are some of the most frequent causes of accelerated bridge deterioration. The stresses have varying effects on bridges depending upon the varying climatic and structural conditions, and their overall effect can severely degrade materials over the years. Several studies have demonstrated that these environmental parameters are one of the significant sources of damage, and it is thus necessary to incorporate them into predictive models in order to achieve improved maintenance planning (Al-Rashed et al., 2023; Fujiu et al., 2022). Although these conditions are some of the most prevalent environmental stressors, they are not the sole conditions influencing bridge deterioration. Some other conditions like ultraviolet radiation, wind-borne debris, air pollution, soil and groundwater chemistry, hydraulic erosion, vegetation growth, ice formation, and excessive heat from wildfires have the potential to influence bridge degradation. These conditions, either singly or in combination with other conditions, have the potential to exacerbate the degradation of bridge materials (Ibrahim et al., 2024a). While this paper does not elaborate on each of these variables, future data collection and modeling effort must be receptive to the inclusion of them if it is demonstrated to have an impact on the performance and life of bridges based on monitoring or sensitivity analysis.

2.5 Data Collection for Bridge Maintenance and Environmental Monitoring

This complex interaction between structural integrity, operational conditions, and environmental factors requires a strong framework in data collection to achieve an effective predictive maintenance model. Recent developments in data collection methodologies, especially in sensor technologies and data integration techniques, have made the potential of monitoring and analyzing in real-time the condition of bridges greater. This evolution gives good ground for the development of predictive maintenance models, which can forecast deterioration with respect to numerous influencing factors, including the environmental conditions, traffic loads, and material properties (Palu & Mahmoud, 2019; Yianni et al., 2016).

Modern sensor technologies, therefore, represent a significant change in the data collection landscape for bridge maintenance. Sensors, such as those for strain, temperature, corrosion, and vibration, enable real-time monitoring of both structural and environmental variables. These technologies allow data collection to continue at localized levels, capturing important information on factors that have a large impact on bridge deterioration. For example, previous studies have already shown their effectiveness in monitoring temperature changes, humidity, and corrosive agents, which are very important in understanding degradation processes in bridge materials (Fom et al., 2015b; Miao et al., 2023a; Sengupta et al., 2021). Integration of sensor networks in the bridge management system should be able to give a more holistic view of how environmental conditions affect structural integrity over service life.

Adding to this, the external sources of data that can be drawn upon include weather stations, satellite imagery, and remote sensing technologies, further enriching the bridge maintenance data landscape. Climatic variables can either be obtained from past archives or real-time monitoring through weather stations. Also, satellite imagery and remote-sensing technologies contribute spatially extensive information on temperature variations and precipitation patterns, as well as land-use changes that might have an impact on bridge performance (Tian et al., 2015; Yuan et al., 2023).

Such integration of data coming from these external sources with those from sensor networks gives bridge management systems a holistic view of localized and regional environmental impacts on bridge conditions. The integration of sensor-based data, together with the records of inspection and other external environmental datasets, is the backbone for ML models deployed in predictive maintenance. ML algorithms can analyze past inspection data together with real-time sensor readings to forecast future deterioration trends; this way, maintenance teams will be able to plan interventions based on the predicted condition of bridge components (Ghosh et al., 2014; Mitterpach et al., 2022; Yuan et al., 2024).

2.6 Studies Related to this Research

Accurate prediction of the deterioration in bridge maintenance is a crucial issue in the long service life and safety of critical infrastructure. More traditional approaches to predicting deterioration in bridges have provided a foundation for knowledge on structural aging, normally based on empirical models, statistical analyses, and condition-based assessments. However, many challenges arise to these methods in cases of real dynamic and complex deterioration due to environmental, operational, and structural factors. Recent advances in ML have proposed new approaches that overcome the existing challenges of the data-driven technique to enable site-specific predictions more accurately. This section discusses how bridge deterioration prediction methodology has evolved from traditional approaches and their fundamental contributions toward the state-of-the-art ML method, with special emphasis on integrating environmental conditions in order to enhance predictive capabilities.

2.6.1 Traditional Approaches in Bridge Deterioration Prediction

Traditional approaches to modeling and forecasting of the deterioration of bridges have based infrastructure management on empirical models, statistical analysis techniques, and heuristic methods. Most of those are based on the adoption of predefined deterioration curves, material degradation rates, and some general hypotheses with respect to structural performance when assessing the bridge condition at any instant in time. For instance, Markov chain models have been widely used in predicting future states of bridge components based on past records, hence providing a probabilistic framework that captures the stochastic nature of deterioration processes (Kobayashi et al., 2012; O'Connor et al., 2013; Saeed Hasan et al., 2015).

The most common frameworks of traditional approaches include regression models, condition rating systems, and Markov chains. Regression models are mostly based on historical inspection data in determining relationships among the factors of deterioration, like age, traffic loads, and environmental conditions. In contrast, condition rating systems allow the standardized assessment of states of bridge components and, therefore, the classification of conditions in discrete states (L. Li et al., 2014; Weissmann et al., 2023). Markov chains, particularly, have been instrumental in developing bridge management systems (BMS) by modeling the transition probabilities between different condition states, thus enabling the prediction of future deterioration based on current assessments (Ahmed et al., 2016; O'Connor et al., 2013; Saeed Hasan et al., 2015). Nevertheless, besides their utility, traditional approaches have some serious drawbacks when dealing with the complex and multifactorial features of deterioration in bridges. These models often oversimplify the interactions between environmental conditions, traffic loads, and structural aging, treating these factors as static or isolated variables. For example, the Markov-based model assumes a time-invariant transition probability. Such an assumption cannot model the dynamic nature of environmental influences, such as temperature fluctuations, freeze-thaw cycles, and exposure to corrosive agents, in an effective way (De-León-Escobedo et al., 2014). This static treatment generally leads to inaccurate predictions and suboptimal maintenance strategies in practice because the models do not adequately represent variability and uncertainty under real-world conditions (Liu et al., 2022a; Weissmann et al., 2023).

Moreover, traditional approaches relied on periodic inspections and condition assessment, mostly leading to reactive strategies in maintenance. In many cases, these strategies do not address deterioration until deterioration has reached critical levels, which may imply increased costs for repair and possible safety risks (Lin et al., 2019; Santos et al., 2022) . Another major problem is the reliance on historical data in predicting future conditions, which may not show current or future environmental conditions that are substantially influencing the deterioration patterns (D et al., 2020; Liu et al., 2022b). For instance, traditional models overlook climate change impacts leading to an underestimation of deterioration rates in some contexts (Lin et al., 2019; Liu et al., 2022b).

2.6.2 Machine Learning Approaches in Bridge Deterioration Prediction

ML has emerged as the revolutionizing subset of artificial intelligence that enables systems to learn from data, recognize patterns, and make forecasts or decisions without explicit programming. ML deals with handling huge datasets and complicated nonlinear relationships of features that are usually hidden in traditional statistical methods. Such capability is particularly important for infrastructure management, in which the integration of diverse data sources may enhance modeling and decision-making. ML uses a variety of algorithms, such as supervised, unsupervised, and reinforcement learning approaches to address different types of data and raise predictive accuracy in a high number of applications, such as bridge maintenance and deterioration prediction (Kim et al., 2020b; Miao et al., 2023b). Figure 2.1 shows the algorithms under each ML type.

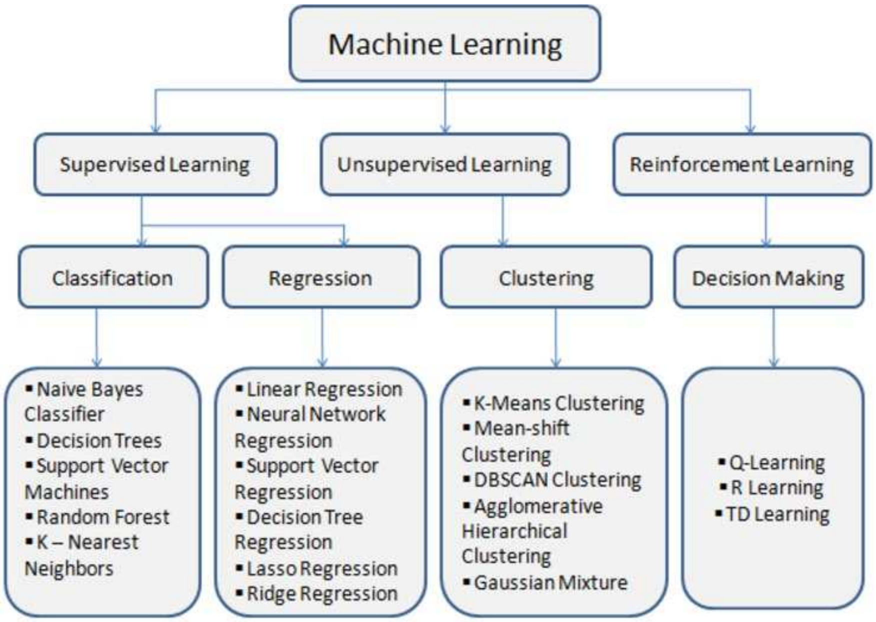


Figure 2.1: Types of machine learning with example algorithms under each (Scikit-learn, 2024)

Application of ML to bridge deterioration prediction presents a revolutionary change for maintenance planning. Traditional methods are mostly based on deterministic models that do not account for the various factors affecting bridge health. These methods generally result in oversimplification of the analysis, which cannot catch the dynamics involved in the deterioration processes of bridges. On the contrary, ML offers a data-driven alternative that can take advantage of vast amounts of both historical and real-time data by recognizing complex patterns and more precisely predicting the trend of deterioration with more reliability (Jiang et al., 2023b; Rashidi Nasab & Elzarka, 2023b)

ML techniques in bridge deterioration modeling include a large number of methodologies that can be mainly grouped into supervised and unsupervised learning. The former has shown improvement in predictive accuracy by using techniques such as regression models, support vector machines (SVM), and ensemble techniques, including random forests (RF) and gradient boosting (GB). These models are good at using historical inspection and environmental data to identify patterns that can make predictions on future deterioration states. For example, some studies have shown that ensemble methods can be an effective approach in integration for improving overall accuracy of the predictions from multiple predictive models for more reliable bridge condition forecasts (Kale et al., 2023b; Li & Song, 2022). Furthermore, unsupervised learning techniques like clustering and anomaly detection have become relevant for the identification of deterioration patterns when labeled datasets are poor, enabling the finding of hidden insights in data (Fereshtehnejad et al., 2022; Mete et al., 2022).

DT are among the basic ML techniques with a hierarchical structure dividing data into subsets depending on the feature thresholds. The method is quite beneficial for its simplicity and interpretability; hence, easy to understand by stakeholders who participate in the decision-making process. In addition, DT handle categorical and numerical data well, thereby being versatile in predicting deterioration of bridge maintenance (Hannan & Anmala, 2021). The ability to visualize the decision process by using tree diagrams will help communicate results to nontechnical audiences and support collaboration between engineers, policymakers, and the public (Hannan & Anmala, 2021).

Among the most vital strengths of DT is their capability to model complicated interactions among features without an extensive preprocessing requirement. They find nonlinear relationships and interactions, which generally prevail in structural health datasets. However, one of the major drawbacks with DT is their tendency to overfit when the depth of the tree is not controlled. This limitation can be overcome by techniques such as pruning or the use of ensemble methods combining trees to improve predictive performance (Hannan & Anmala, 2021). RF is an ensemble learning method that constructs a multitude of DT's and aggregates their prediction's gain and accuracy. It, therefore, significantly reduces overfitting problems that happen with a single decision tree by averaging the results from many trees trained on different subsets of the data. Another feature added to introduce diversity in the trees was that at every split, there was a random selection of features; this would result in an even better generalization to unseen data (Lu & Guler, 2022).

The ability of the RF to handle high-dimensional data makes them be more suitable for applications involving complex datasets, such as those related to bridge maintenance prediction. They can efficiently deal with a huge number of features, which might also include structural, traffic, and environmental variables, without having extensive feature selection or dimensionality reduction. Further, RF provide the additional value of feature importance to be applied by researchers to find which variable most affects the deterioration process of the bridge, and then guide further maintenance priority accordingly. However, RF models can become computationally expensive with large datasets and may suffer from reduced interpretability compared to simpler models. (Lu & Guler, 2022).

GB is another ensemble method where models are learned in succession, each new model trying to fix errors from previously learned ones. This method works quite well, especially in handling nonlinear relationships and feature interactions. As such, it fits quite well when it comes to the modeling of bridge maintenance deterioration. According to Viljanen et al., (2022) the flexibility of GB generally allows the model to apply different loss functions for varied prediction tasks.

One of the essential strengths of GB is high predictive accuracy, especially on structured data. Indeed, studies have shown that a number of other algorithms in terms of accuracy and robustness, when applied to more complex datasets, are usually outperformed by models based on GB (Viljanen et al., 2022). The iterative nature of the algorithm allows it to give much attention to difficult-to-predict observations, hence continually refining its predictions and improving overall model performance. However, careful hyperparameter tuning is essential in avoiding overfitting and achieving optimal performance.

SVM are a very robust class of supervised learning methods that have shown considerable effectiveness in classification and regression tasks. This method works by finding the best hyperplane, which is used to maximize the margin between different classes in the feature space. This particular characteristic makes SVMs particularly effective in high-dimensional spaces, where complex distributions of data can be discriminated effectively. The flexibility of the SVMs can be increased to a great extent by using kernel functions, enabling the transformation of input data into higher-dimensional spaces to represent nonlinear relationships (Viljanen et al., 2022).

SVMs are especially useful when the number of features is larger than the number of observations, which is quite common in bridge maintenance datasets containing various structural and environmental factors. Their robustness to overfitting, especially in high-dimensional settings, makes them a very reliable choice for predictive modeling in this context. However, SVMs could be computationally costly, and the choice of kernel and hyperparameters directly impacts performance, which requires great tuning and validation (Viljanen et al., 2022).

Deep learning techniques, in particular neural networks, have become widely used for the prediction of bridge deterioration owing to their capability to analyze high-dimensional datasets from sensors or environmental time series. These models, therefore, achieve what traditional models do not, establishing complicated relationships within the data. Many research have been done on convolutional neural networks (CNNs) and recurrent neural networks (RNNs) over time-series data from bridge monitoring systems to enable more detailed and contextual predictions of structural health issues (Calabrese et al., 2020; Zhou et al., 2022). Integrating deep learning with traditional ML has improved predictions in applications where data is abundant and complex. However, deep learning models often require large labeled datasets which can limit their applicability in data-scarce environments (Liu et al., 2021).

One of the benefits associated with ML approaches involves the potential to integrate dynamic environmental conditions within predictive models. Whereas in traditional techniques, the consideration of environmental factors is generally secondary or static, in ML models, there is a potential for synthesizing real-time data that will accurately represent temperature shifts, humidity, freeze-thaw cycles, and exposure to corrosive agents. This would enable more holistic and site-specific assessments of bridge deterioration. For example, research have proven that the inclusion of environmental variables in predictive models increases the accuracy of deterioration forecasts made by maintenance teams, which can then prioritize interventions using real-time risk assessment (Veloso et al., 2022).

Notwithstanding all the encouraging developments relating to ML for bridge deterioration modeling, a number of challenges persist. Among them is that there will be high demand for quality and diversity in data. The performance of ML models strongly depends on the quality and quantity of the data in training. Some complicated models require heavy computation resources to train the model, which may act as a barrier for a wide adaption of the practice in reality (Dhyani, 2021; Raja et al., 2022). However, the opportunities presented by ML techniques in learning from enormous volumes of data, adapting to changes in conditions, and providing insights that can lead to practical actions underpinning their potential to revolutionize bridge maintenance.

CHAPTER THREE: METHODOLOGY

This chapter describes the methodology followed in the research. The objective of this research is to develop a ML model capable of integrating environmental factors with structural and traffic information in predicting the deterioration of bridge maintenance. The chapter begins by describing the data collection process, the integration of the bridge and environmental conditions datasets to form a unified dataset. Preprocessing steps are described to ensure the quality of the data, followed by the feature selection techniques for determining important variables. Then, ML model development, training, and evaluation are discussed. The research methodology is divided into five steps as shown in Figure 3.1.

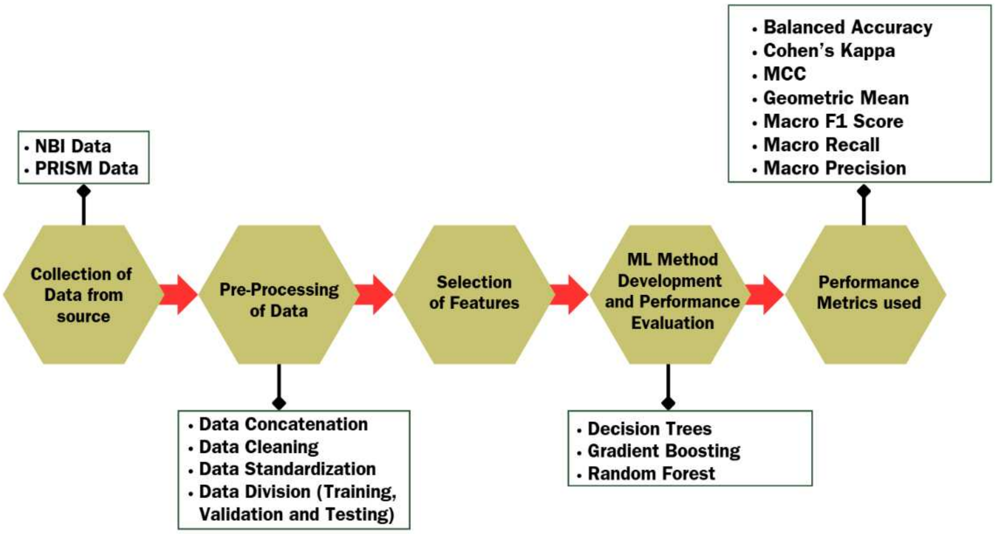


Figure 3.1 Research Methodology Workflow

3.1 Data Collection

Data for this research was collected from two main sources. The bridge data used for analysis were downloaded from the NBI databases. The NBI provides data on condition ratings of bridge components. It categorizes them into a scale of 0-9, where 9 means "excellent" and 0 means "failed". It includes data on design characteristics of the bridge and other operational characteristics, such as ADT. In addition, the database provides data on the states (good, fair and poor) of the bridge elements.

Moreover, environmental data were collected from the PRISM dataset in respect to the state of Colorado. It includes temperature variables of maximum, minimum, and mean, precipitation and relative humidity. By aligning the data, the geospatial linkage of the bridge data is done with environmental data according to geographic coordinates, and temporal synchronization by matching the dates of bridge inspections to the respective environmental conditions.

3.1.1 Discussion of Environmental Factors

Bridges experience continuous exposure to environmental conditions significantly impacting their deterioration rates. Recognizing that bridge deterioration extends beyond structural and traffic factors alone, this study incorporates a selection of environmental variables to better predict bridge maintenance needs in Colorado. The selected variables temperature metrics, precipitation measures and humidity indicators are grounded in established principles of material degradation, supported extensively by engineering literature. This section discusses the rationale for selecting these specific environmental factors, their expected effects on bridge conditions, and acknowledges factors excluded from the model due to data limitations or regional irrelevance.

Temperature-related variables were a primary focus, reflecting their significant role in material stress and deterioration processes. Variables included mean temperature, temperature range, extremes, and freeze-thaw cycles. Among these, freeze-thaw cycles were particularly critical, given Colorado’s climate, where frequent temperature fluctuations around the freezing point occur.

Repeated freeze-thaw events can cause concrete cracking and spalling, exacerbating deterioration by exposing reinforcing steel to corrosion (Rashidi Nasab & Elzarka, 2023). Consequently, freeze-thaw frequency was calculated from daily temperature records, serving as a robust indicator of thermal stress on bridge structures. Literature confirms this approach, emphasizing that freeze-thaw cycles significantly accelerate bridge deterioration in moderate climates compared to extreme heat or cold conditions alone (Rashidi Nasab & Elzarka, 2023).

Additionally, extreme temperature events were included, given their potential to cause material expansion, contraction, or embrittlement. While extreme temperatures have secondary importance relative to freeze-thaw cycles in Colorado, their inclusion is justified by regional climatic conditions. This regional sensitivity aligns with findings indicating geographic variations heavily influence the relative importance of temperature metrics in predicting bridge deterioration (Rashidi Nasab & Elzarka, 2023).

Precipitation measures were integral, capturing the overall moisture exposure of bridges. Annual rainfall and snowfall/snow load were particularly important due to their role in corrosion and freeze-thaw processes. High precipitation increases moisture infiltration, facilitating chemical deterioration and freeze-thaw cycling. In Colorado, snowfall and resultant snow loads significantly impact bridge structures. Accumulated snow not only represents an additional structural load but also contributes to freeze-thaw deterioration by generating meltwater that infiltrates concrete structures (Al-Rashed et al., 2023). Furthermore, snow accumulation often requires chemical deicing, indirectly exacerbating corrosion. Therefore, distinguishing snowfall from rain in the model provides insights into specific seasonal deterioration mechanisms uniquely relevant to Colorado.

Humidity-related factors, specifically dew point and relative humidity, were also essential in assessing corrosion risk. Dew point was selected due to its stability and practicality, offering a direct measure of atmospheric moisture content. Condensation risk, derived from dew point and temperature interactions, captures the potential for surface moisture formation independent of direct rainfall, which is critical for corrosion processes. The literature consistently links sustained humidity conditions with accelerated corrosion rates, supporting this approach. High ambient humidity prolongs the wetness period on structural surfaces, significantly accelerating corrosion (Al-Rashed et al., 2023). Therefore, including humidity metrics helps account for localized environmental conditions, such as areas near water bodies or fog-prone regions, thus differentiating bridge deterioration rates based on varying humidity exposure.

Composite environmental indices, notably the corrosion risk index, combined multiple factors to effectively quantify environmental aggressiveness. The corrosion risk index incorporates humidity, precipitation, and indirectly, chloride exposure from deicing salts through freeze-thaw frequency. This composite metric is critical as corrosion of steel reinforcement is highly sensitive to simultaneous moisture presence and chloride exposure. While direct chloride data were unavailable, freeze-thaw and snow variables served as proxies, supported by literature highlighting the critical role of moisture and chloride interaction in accelerating steel corrosion (Al-Rashed et al., 2023). This comprehensive index thus captures synergistic environmental effects crucial for accurate bridge deterioration prediction.

Although several uncontrollable environmental factors listed in earlier research, such as wind loading and seismic activity, were recognized for their potential impacts, they were ultimately excluded from this study. The exclusion of these factors was primarily due to limited data availability and their relatively lower regional significance in Colorado compared to the factors included. For example, seismic activity is minimal in Colorado, rendering its predictive value low for bridge deterioration in this context. Similarly, detailed wind load data were unavailable for the bridges studied, limiting it's integration despite acknowledged potential impacts.

In conclusion, the selected environmental variables temperature, precipitation, humidity, and composite indices provide a robust framework to predict bridge deterioration comprehensively. Their selection reflects regional climatic conditions, data availability, and documented relevance in literature. The model thus effectively captures the complex interactions between environmental stressors and structural deterioration, enabling more precise and regionally appropriate predictions of bridge maintenance needs.

3.2 Data Preprocessing

The data preprocessing stage consists of two main phases: 1) data cleaning and preparation, and 2) data standardization and transformation. In the data cleaning and preparation (section 3.2.1), we ensure the dataset's quality and readiness for analysis and use in ML. The section describes the concatenation of the NBI dataset as well as the environmental data. It provides a detailed description of selected variables. This section also involves the handling of missing data, and redundancy eliminations in which redundant or duplicated information are removed from both datasets. Furthermore, the subsequent section (Data Standardization and Transformation, section 3.2.2) talks about how numeric data are standardized using a standard scalar and how the categorical data are encoded using one-hot encoding.

3.2.1 Data Cleaning and Preparation

In order to merge datasets, identifiers such as bridge id, and geographic coordinates (latitude and longitude) were used to align datasets. Due to varying changes in environmental data for a particular location, geographic proximity was important to ensure that climate data corresponded to the exact or nearest location of bridges considered. Also, environmental data such as temperature and precipitation are often recorded at daily or hourly intervals, therefore these data needed to be synchronized with the timeline of our structural data (such as bridge age). A clean dataset was subsequently created containing information about bridges and environmental condition data from the NBI as well as the PRISM datasets respectively. During the cleaning process, missing data and inconsistencies such as missing data related to condition ratings were identified. These inconsistencies affect the analysis of the data and hence are addressed by handling missing data, removing duplicate records, detecting outliers that can skew the model's understanding of the data, and treating outliers.

3.2.2 Data Standardization and Transformation

The integration of bridge structural data with environmental conditions required a systematic approach to data standardization and transformation. The NBI condition ratings, ranging from 0-9, were standardized alongside diverse environmental measurements to ensure computational compatibility and meaningful analysis. Numerical features underwent standard scaling transformation, expressed as:

$$Z = \frac{(x - \mu)}{\sigma} \tag{1}$$

where Z represents the standardized value, x is the original value, μ is the feature mean, and σ is the standard deviation (Rashidi Nasab & Elzarka, 2023c). This standardization was crucial for features such as bridge age, structure length, and traffic loads. To capture deterioration patterns effectively, a deterioration rate (DR) was calculated for each bridge component:

$$DR = \frac{9 - CR}{BA} \tag{2}$$

where CR represents the condition rating and BA denotes bridge age (adapted from Agrawal et al., 2008). Environmental measurements required specific transformations, including temperature standardization to Celsius and precipitation to millimeters. The categorical variables were one-hot encoded, ordinal variables maintained their hierarchical relationships through appropriate encoding schemes, and missing values were handled through domain-specific methods. Environmental data gaps were filled using spatial interpolation from nearby stations, and structured measurements were replaced on a similar bridge attribute basis. This comprehensive standardization process ensured compatibility between structural and environmental features while preserving the meaningful relationships within the data, creating a robust foundation for the ML models.

3.3 Feature Selection

The feature selection process applied a thorough three-algorithm approach to establish top predictors of deterioration in bridge maintenance. The approach employed DT, RF, and GB algorithms to calculate and validate feature importance scores.

Feature importance was calculated for DT based on Gini importance:

$$FI_{dt} = \Sigma\left(\frac{nj}{N} \times Gi\right) \tag{3}$$

where nj is the weighted number of samples reaching node j, N is the total number of samples, and Gi is the Gini impurity decrease for node i (Louppe et al., n.d.).

For RF, the permutation importance was computed as:

$$FI_{rf} = \frac{(Ea - Eb)}{\sigma} \tag{4}$$

where Ea is the original model error, Eb is the error after feature permutation, and σ is the standard deviation of the differences (“Random Forest,” 2025).

For GB, feature importance was determined through the accumulated gain:

$$FI_{gb} = \frac{\Sigma(gain_i \times coverage_i)}{N} \tag{4}$$

where gain_i represents the improvement in accuracy brought by a feature in each tree, coverage_i is the relative number of observations related to this feature, and N is the total number of trees (T. Chen & Guestrin, 2016a).

To enhance the robustness of feature selection, a composite importance score (CIS) was developed:

$$CIS = \frac{\left(FI_{dt} + FI_{rf} + FI_{gb}\right)}{3} \tag{5}$$

Both environmental and structural attributes were considered in feature selection and their main effects and interaction were evaluated. Redundant predictors were minimized to save space by first reducing features through a comparison of a threshold value of 0.9. Features important in all three models were prioritized in this process, especially those related to both environmental and structural degradation. As different information about feature importance was provided by each of the algorithms, this process generated a stable feature selection process. DT provided threshold-based definitive splitting information, RF provided feature interaction information, and GB provided minor pattern information through sequential improvement. Final feature selection was performed for features with large importance values for all three algorithms to ensure a full and stable set of predictors for bridge deterioration.

3.4 Machine Learning Methods

Prediction of deterioration is critical to effective maintenance planning and intervention planning, and even more so in including environmental factors in prediction models. For this study, ML approaches were selected for stability, interpretability, and potential to model complex non-linear interaction which are typical in bridge deterioration data. The DT, RF, and GB algorithms were selected particularly due to their established record in similar predictive modeling applications to maintenance in infrastructures (Ghafoori et al., 2024a).

DTs were chosen to serve as baseline algorithm primarily due to their simplicity and interpretability in addition to their applicability to deal with both numerical and categorized features (Breiman, 2001). Interpretability is especially important in engineering since it enables ~~a~~ to understand individual contributions of predictors to help in practical decision-making. For DT implementation, GridSearchCV was employed to optimize model hyperparameters, specifically focusing on maximum depth, minimum samples required to split a node, minimum samples required at leaf nodes, and criterion selection between Gini impurity and entropy. This helped to select the optimal model parameters to avoid overfitting and ensure highest accuracy in prediction.

RF was selected due to its ensemble learning approach to aggregate a great multiple decision trees in a manner to mitigate individual model variances and enhance predictive stability and accuracy (Breiman, 2001). RF also works well with large data sets with complex feature interactions, a common characteristic in bridge condition data (Ghafoori et al., 2024a). Random Forest implementation in this study involved feature importance analysis to identify key variables significantly influencing bridge deterioration. Additionally, the SMOTETomek was applied to ensure balanced datasets for training in order to address inherent class imbalance. Cross-validation techniques also validated model stability and generalizability of RF.

GB was selected and used with XGBoost framework due to its capacity to model complex patterns and compatibility with class-imbalanced data (Chen & Guestrin, 2016). XGBoost is well-regarded for its high predictive accuracy, computational efficiency, and flexibility in hyperparameter tuning, making it highly suitable for the current study's large dataset. Just like in RF model, XGBoost implementation involved extensive hyperparameter tuning and made use of customized scoring metrics prioritizing recall for classes indicating rare but critical maintenance events. This method particularly suited the predictive objectives of this study, ensuring that critical bridge deterioration cases were accurately predicted and appropriately prioritized.

While SVMs have in other instances in the past been successfully used to forecast bridge deterioration due to their robustness and accuracy in classification tasks (Nguyen et al., 2017), SVMs were ruled out in this work predominately on account of computational impracticality. SVM computational complexity in a range of $O(n^2)$ to $O(n^3)$, depending on kernel and parameters, was computationally unsuitable considering the dataset size ($n=45{,}037$) and dimensionality (87 features). The dataset also expanded to considerably more than 100,000 samples when applying SMOTE in dataset balancing, making SVM utilization computationally impractical for timely model training and evaluation.

3.4.1 Training Process

The training was a systematic approach utilizing the three models, DT, RF, and GB, each configured to approach bridge maintenance prediction in its own unique way. Training was conducted using a stratified k-fold cross-validation technique using k=3 to ensure more stable model evaluation that maintains the class distribution across folds.

DT classifier was trained using Gini impurity criterion as a measure of quality of split, balanced weighting of classes to handle the inherent imbalance in bridge maintenance data, along with parameters to ensure that tree grows sufficiently but not so much as to cause overfitting through minimum sample splits and minimum leaf constraints. Key hyperparameters included maximum depth constraints and minimum impurity decrease thresholds to ensure meaningful splits.

The RF ensemble expanded upon the DT foundation by implementing 100 base estimators, each trained on bootstrap samples of the training data. The algorithm maintained balanced class weight approach while introducing feature randomization at each split. This configuration enhanced model robustness through diversity in both sample and feature space, with out-of-bag estimates providing ongoing validation during training.

A conservative learning rate of 0.1 was employed in GB to facilitate stable model development through sequential additions of trees. The algorithm employed a special loss function that gave more emphasis to maintenance-critical classes, using subsample ratios of 0.8 to introduce randomization to avoid overfitting. The number of estimators was also set to 100, then early stopping methods tracked performance on a validation set to choose the best-performing ensemble size.

To address the class imbalance challenge, the training process integrated the SMOTETomek technique, which combines Synthetic Minority Over-sampling Technique (SMOTE) with Tomek links under sampling:

$$S = SMOTE(X, N, k) - TomekLinks(X) \tag{6}$$

where S represents the balanced dataset, X is the original feature set, N is the desired sampling ratio, and k is the number of nearest neighbors used in the SMOTE algorithm (Batista et al., 2004).

The data partitioning strategy followed a 64/16/20 split ratio for training, validation, and test sets respectively, ensuring consistent class distribution across all partitions through stratification. The hyperparameter optimization process employed GridSearchCV with stratified k-fold cross-validation, optimizing for a composite objective function:

$$Obj = \alpha \times Performance + (1 - \alpha) \times \left(\frac{1}{Complexity}\right) \tag{7}$$

where α = 0.7 balances model performance and complexity, a strategy adopted by Zhao et al. (2024). Performance metrics prioritized balanced accuracy to account for class imbalance, while complexity penalties were algorithm-specific: tree depth for DT, and ensemble size for RF and GB.

Model calibration was implemented using Platt scaling to ensure reliable probability estimates:

$$P(y = 1|x) = \frac{1}{1 + exp(-a \times f(x) - b)} \tag{8}$$

where f(x) represents the model's output score, with a and b serving as calibration parameters optimized on validation data (Niculescu-Mizil & Caruana, 2005).

The training process incorporated early stopping criteria with a patience parameter of 5 epochs, monitoring validation performance to prevent overfitting. Feature importance stability was assessed across cross-validation folds, providing insights into the robustness of feature selection.

This comprehensive approach ensured that each algorithm was optimized for both performance and generalization capability while maintaining interpretability for practical applications in bridge maintenance prediction.

3.5 Performance Metrics

The performance metric used in evaluating the predictive models for bridge maintenance involved a large range of performance measures carefully selected to face the challenge of class imbalance typical in infrastructure maintenance prediction. Since the distribution of conditions of bridges is uneven such that Poor conditions are much less frequent than Fair and Good conditions, it was determined that traditional measures of accuracy were unsuitable to measure model effectiveness.

The main performance measures were classified into three different groups: general classification performance, imbalanced data measures, and calibration measures. For general classification assessment, the framework included:

$$Balanced\ Accuracy = \frac{(TPR + TNR)}{2} \tag{9}$$

where TPR (True Positive Rate) and TNR (True Negative Rate) provide equal weight to each class regardless of their distribution (Brodersen et al., 2010). This was supplemented by Cohen's Kappa coefficient:

$$\kappa = \frac{(po - pe)}{(1 - pe)} \tag{10}$$

where po represents observed agreement and pe represents expected agreement by chance, providing a more robust measure of classification performance by accounting for random chance agreement (Rashidi Nasab & Elzarka, 2023c).

To specifically address class imbalance challenges, the following metrics were implemented:

Matthews Correlation Coefficient (MCC):

$$MCC = \frac{(TP \times TN - FP \times FN)}{\sqrt{(TP + FP)(TP + FN)(TN + FP)(TN + FN)}} \tag{11}$$

where FP represents false positives and FN represents false negatives (Chicco & Jurman, 2020a).

This metric provides a balanced measure of binary classification performance even with highly skewed class distributions. The Geometric Mean Score was also calculated:

$$G - mean = \sqrt{Sensitivity \times Specificity} \tag{12}$$

which assesses the balance between classification performances on both majority and minority classes (Espíndola & Ebecken, 2005).

For multi-class evaluation, macro-averaged metrics were employed to give equal weight to all classes:

$$Macro\ Precision = \left(\frac{1}{n}\right) \times \frac{\Sigma(TPi)}{TPi + FPi} \tag{13}$$

$$Macro\ Recall = \left(\frac{1}{n}\right) \times \frac{\Sigma(TPi)}{TPi + FNi} \tag{14}$$

$$Macro\ F1 = \left(\frac{1}{n}\right) \times \frac{\Sigma(2 \times Precisioni \times Recalli)}{Precisioni + Recalli} \tag{15}$$

where n is the number of classes, and Precisioni and Recalli are the precision and recall for class i, respectively (Rashidi Nasab & Elzarka, 2023c).

For a complete evaluation, all measures were obtained using stratified k-fold cross-validation with k = 3, whereas confidence intervals were approximated by bootstrap resampling with n = 1000. This approach added statistical significance to performance measures, in addition to accounting for class distribution. The same evaluation framework was applied to all three algorithms (DT, RF, and GB) to allow direct performance comparisons while mitigating issues related to the naturally imbalanced characteristics of infrastructure maintenance data.

CHAPTER FOUR: RESULTS AND DISCUSSION

4.1 Introduction

This chapter presents the results from the ML models that were developed for the study. It also describes the significant findings in relation to the study objectives, along with discussion with what is found in literature.

4.2 Data Overview

The dataset used in this research were from a number of reputable sources in order to achieve the study's objectives. Structural and operational data were derived primarily from NBI database, and environmental data from PRISM climate group dataset. The resulting merged dataset constitutes a comprehensive set of Colorado-specific, traffic, bridge-specific, and environmental parameters over a span of 11 years.

After preprocessing and integration, the dataset had 75,063 samples and 97 unique features. These included structural attributes such as deck condition ratings, superstructure and substructure ratings, age, traffic data, as well as a number of environmental variables such as mean temperature, max and min temperature, precipitation, relative humidity, as well as derived indices such as freeze-thaw frequency, and condensation risk. Data preparation and cleaning also involved removing duplicate variables, missing values being filled using spatial interpolation, and normalizing all numerical fields to standardized z-score values. Figure 4.1 visualizes the distribution of bridge conditions across the study sample which is crucial for understanding the underlying characteristics influencing deterioration. As seen, the vast majority were in either "Good" or "Fair" condition, while a relatively low proportion were also in "Poor/Critical" condition. Table 4.1 gives the distribution of classes throughout the entire dataset.

Table 4.1 Distribution of Maintenance Classes in the Dataset

Maintenance Class	**Percentage (%)**
Good	53.21
Fair	43.05
Poor	3.74

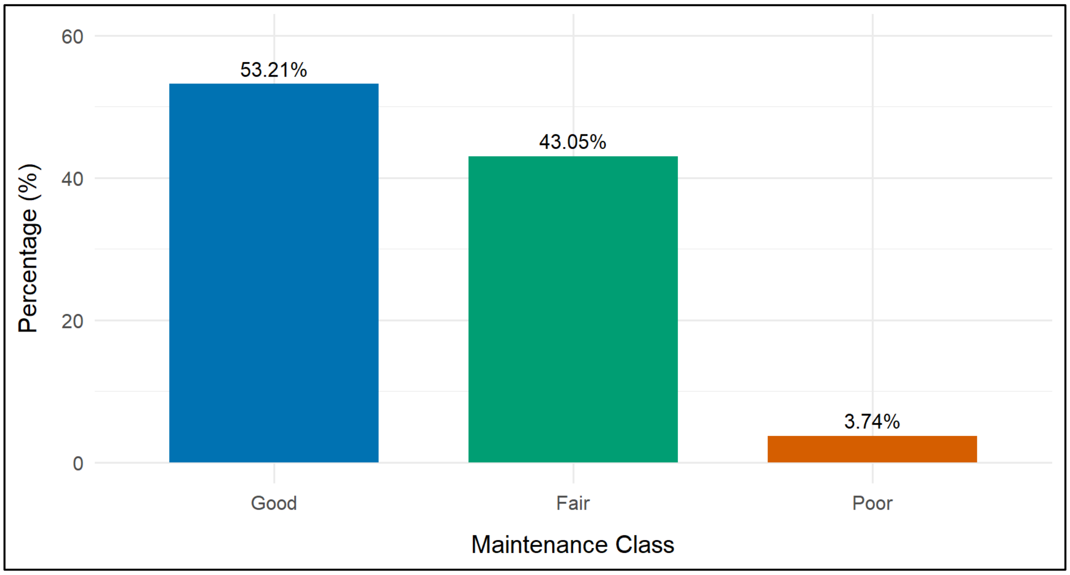


Figure 4.1 Distribution of Maintenance Classes

This imbalance required using synthetic oversampling methods, specifically the Synthetic Minority Over-sampling Technique coupled with Tomek links (SMOTETomek), when training so that the model would be sensitive to minority classes without compromising too much on its generalization. Feature selection then cleaned the dataset by identifying useful predictors of bridge deterioration. Filtering at a high level of inter-feature correlation (threshold = 0.9) eliminated ten features, including ROADWAY_WIDTH_MT_051, BRIDGE_AGE, among others. Final filtering was then conducted employing DT, RF, and GB algorithms to choose significant predictors, the results are shown in Table 4.2.

Table 4.2 Top Predictive Features Identified

Rank	Feature	Importance (Average)
1	SUFFICIENCY_RATING	0.6918
2	YEAR_BUILT_027	0.0375
3	YEARS_SINCE_RECONSTRUCTION	0.0307
4	LAT_016	0.0196
5	LONG_017	0.0193
6	year	0.0184
7	APPR_WIDTH_MT_051	0.0180
8	DECK_STRUCTURE_TYPE_107_3	0.0162
9	CHANNEL_COND_061	0.0144
10	TRAFFIC_LOAD	0.0113

The predominance of SUFFICIENCY_RATING as a predictive feature shows its complete capture of a bridge's operational condition aligning with past studies emphasizing the predictive ability of composite condition ratings to infrastructure maintenance models (Rashidi Nasab & Elzarka, 2023b). Additionally, the inclusion of geographical coordinates (LAT_016 and LONG_017) as part of the leading predictors (see Table 4.2) demonstrates influence of spatially distributed environmental factors on bridge condition in accordance with studies by (Elleathy et al., 2024), who demonstrated the regional variability of bridge deterioration mechanisms. Relative importance of the top ten features are shown in Figure 4.2 below.

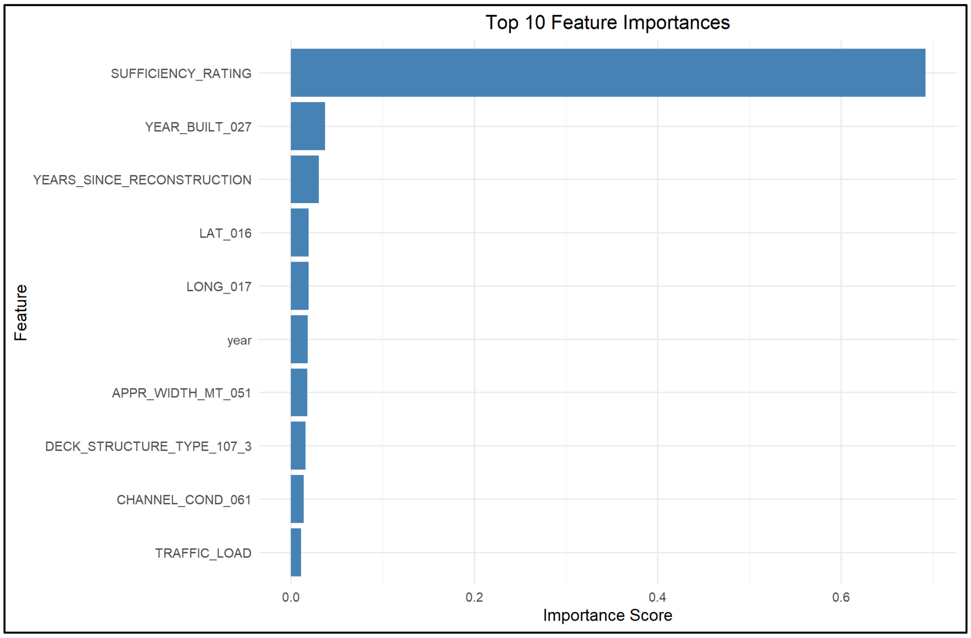


Figure 4.2 Relative importance of the top ten features

The cleaned dataset provided a reliable foundation for training the models to help at predicting bridge deterioration by integrating environmental, structural, and operational parameters.

4.3 Feature Selection Results

Structural properties as well as traffic features were prominent during feature selection in all the ML models in this analysis (DT, RF and GB). The features were consistently demonstrated to possess high predictive power in relation to deterioration patterns of bridge components. SUFFICIENCY_RATING, which represents a composite index summarizing structural adequacy and safety, emerged to be the most predominant single feature across all models, having the highest importance score with an average of 0.6918. This finding reflects the role of SUFFICIENCY_RATING as a broad indicator of the condition of a bridge, integrating factors like structural integrity and traffic serviceability. Such single feature dominance also reflects prior empirical work where its dominance was emphasized in predictive modeling of infrastructure maintenance (Rashidi Nasab & Elzarka, 2023b).

YEAR_BUILT_027, a measurement of when the bridge was built, and YEARS_SINCE_RECONSTRUCTION, measuring years since rehabilitation, were also among top-ranking structural variables. Dominance of such variables reflects the direct impact of material aging and cumulative impact of stress cycles, on bridge deterioration. Aging bridges, especially those without any recent interventions, are more prone to environmental and operational stressors which is in line with trend observations in previous research (Elleathy et al., 2024). Several operational features such as APPR_WIDTH_MT_051 (approach roadway width) and STRUCTURE_LEN_MT_049 (length of structure), exhibited moderate but significant contributions to the prediction models. Variations in bridge geometry influence structure performance, dynamic response to loading, as well as exposure to environmental condition, also impacting rates of deterioration.

Traffic-related features such as ADT_029 (Average Daily Traffic) and YEAR_ADT_030 (year of traffic count) also contributed to the models, although with lower importance scores compared to structural features. These features capture the operational loading intensity on bridge structures, a known accelerator of fatigue and material wear, particularly in heavily trafficked corridors. As confirmed in earlier studies (Ghafoori et al., 2024b), operational stressors are significant factors influencing timing and severity of maintenance needs. The combination of structural, geometric, and operational features provides multi-dimensional view of bridge condition. Their retention and prominence in the predictive models underscore the necessity of an integrated approach that captures both the intrinsic properties of bridge assets and the external operational loads which they are subjected to over time.

4.4 Model Performance Evaluation

In this section, the test-set evaluation metrics for each model (DT, RF and GB) in predicting bridge component maintenance deterioration are presented. The key performance measures including Balanced Accuracy, Cohen's Kappa, MCC, Geometric Mean Score, Macro-Averaged F1, Recall and Precision are reported for each bridge component (deck, superstructure, substructure) and model.

4.4.1 Decision Tree Results

The DT classifier yielded high testing set performance metrics for all three components of the bridges (deck, superstructure, and substructure), as indicated by Table 4.3. The model was developed to predict bridge maintenance needs, with key decision splits based on features such as bridge age, traffic load and freeze-thaw cycles as shown in Figure 4.3. The DT model attained balanced accuracy values higher than 0.95 for all components indicative of strong prediction for each condition class with the presence of class imbalance. The Cohen's Kappa values were from the range of 0.97 to 0.99, which is near-perfect agreement between model-predicted scores and actual ratings during inspections higher than at the level of chance. To put into perspective, the interpretation of K > 0.8 is normally "almost perfect" agreement (Rong et al., 2024), and accordingly, values at the level of 0.99 indicate an extremely reliable classifier. For each of the components, the MCC was ~0.98–0.99, which again indicates the strong predicting capability of the model. MCC is an extremely strict metric and attains such high values only when the classifier is good at all the classes (true positives, true negatives, false positives, and false negatives) (Chicco & Jurman, 2020). Essentially, the DT's near-maximal MCC indicates the model never had any glaring weakness at one of the classes. The Geometric Mean scores (of around 0.95–0.99) indicate the balanced accuracy, and further ensured the classifier was very good at the sensitivity to all the condition categories. The macro-averaged precision, recall, and F1 scores were all similarly high and consistent (about 0.95 to 0.99 for all components), which is indicative that the DT was equally good on minority classes as well as majority classes. These metrics together indicate the integrated model is able to detect bridges of different condition states with very high accuracy.

Table 4.3 Decision Tree Test-Set Metrics by Bridge Component

Component	**Balanced Accuracy**	**Cohen's Kappa**	**MCC**	**Geometric Mean**	**Macro F1**	**Macro Recall**	**Macro Precision**
Deck	0.9875	0.9894	0.9894	0.9875	0.9839	0.9875	0.9804
Superstructure	0.9526	0.9779	0.9779	0.9510	0.9505	0.9526	0.9485
Substructure	0.9749	0.9877	0.9877	0.9745	0.9811	0.9749	0.9875

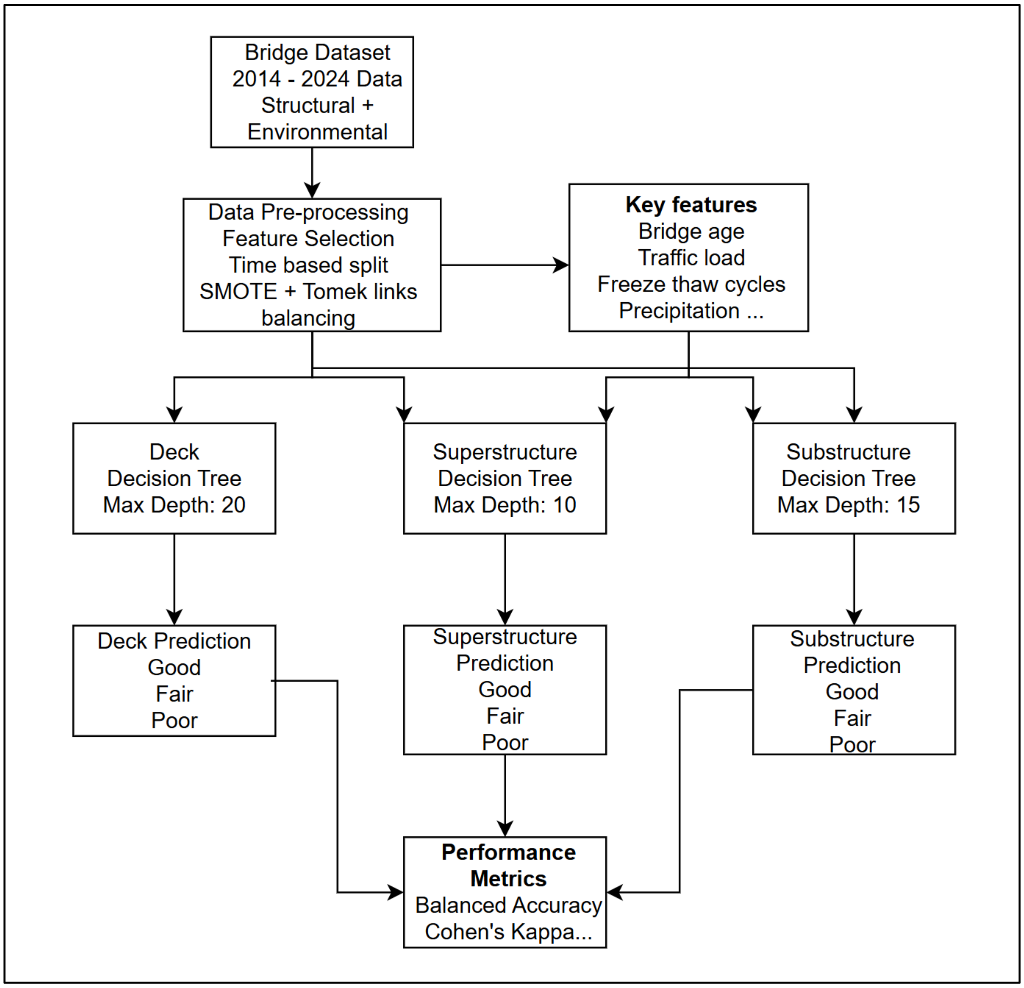


Figure 4.3 Decision Tree flow for bridge deterioration prediction

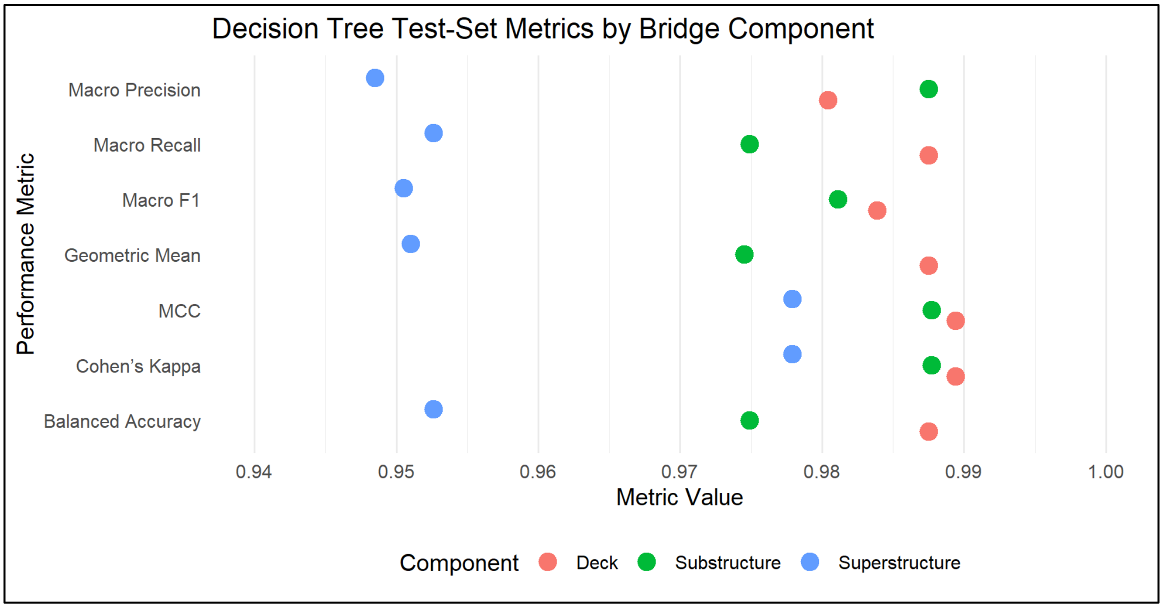


Figure 4.4 Decision Tree Performance by Bridge Component

Figure 4.4 provides a visual summary of the Decision Tree classifier's performance across the primary bridge components Deck, Superstructure, and Substructure using a radar chart. This chart maps key evaluation metrics such as Balanced Accuracy, Cohen's Kappa, MCC, Geometric Mean, Macro F1, Macro Recall, and Macro Precision onto different axes, allowing for a multi-dimensional comparison. The resulting "performance profile" for each component, distinguished by color, illustrates the model's strengths and weaknesses across these metrics, with a larger area generally indicating more robust and well-rounded performance. The results shown on this chart presents normalized values using the highest and least values from each category. This visualization is presented to offer an intuitive understanding of how effectively the Decision Tree model predicts deterioration for each distinct bridge element.

4.4.1.1 Deck

The best-performing element among the three elements of the bridges was the deck element. DT's class-balanced accuracy in the prediction of the condition of the deck was 0.9875, the best in the group (Table 4.3), meaning the model was 98.75% correct on average to predict each of the classes of the condition of the deck. This improved class-balanced accuracy is indicated in Figure 4.4 illustrating the model of the deck performed better than the superstructure and substructure models in class-balanced accuracy. The predicted decks' geometric mean was high at 0.9875, meaning the model's sensitivity and specificity to the deck were approximately 99%. Practically, the deck DT rarely misclassified actual bridges that were indeed deteriorated (had high recall) and never indicated bridges at risk when they were not (had high specificity). The Cohen's Kappa of the model for the deck was 0.9894, and MCC was also 0.9894 – at the ceiling for classifying performance. A Kappa of 0.99 indicates very good fit between the DT's predicted decks and the actual ground conditions (Rong et al., 2024). The MCC of 0.9894 indicates the very good fit is well-balanced in all the classes because MCC favors balanced performance (Chicco & Jurman, 2020b). Lastly, the macro-averaged F1 for the deck was 0.9839, where the macro recall was 0.9875 and macro precision was 0.9804. These indicatively suggest the model was extremely sensitive (had high recall) and extremely precise on average to predict each of the types of the deck. The model's slightly higher macro recall than the macro precision (98.75% vs 98.04%) indicates the model may produce slightly more false alarms than misses for the conditions of the decks. That is, the DT is slightly conservative – it marks as lower condition (for maintenance) several bridges which were in higher condition but never misses an actually deteriorated deck. This bias is desirable in the management of bridges, as to miss an inferior condition of the deck (a false negative) is more costly than an unjustified inspection of one which was wrongly reported as such (a false positive).

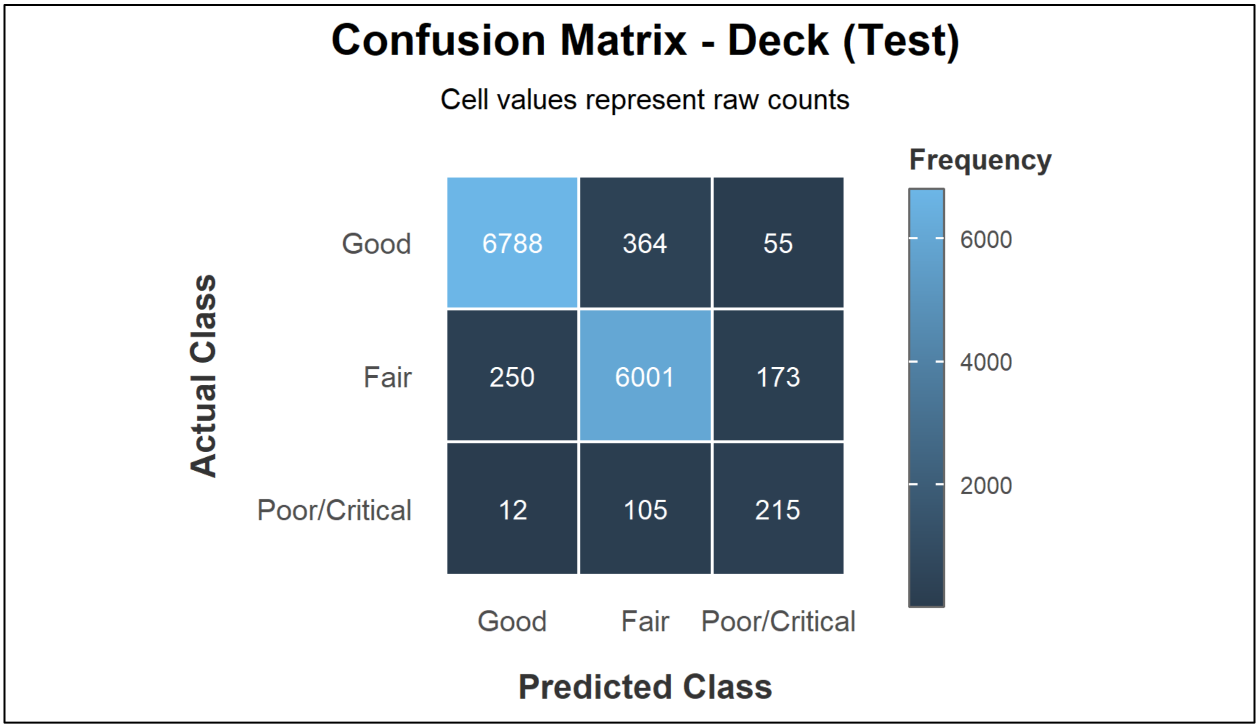


Figure 4.5 Confusion Matrix - Deck (Test)

The near-perfection of the DT is mirrored graphically in the confusion matrix of the deck model (Figure 4.6). Almost all the deck instances in the test set are on the main diagonal of the confusion matrix, indicating DT predicted the actual condition rating accurately. The model distinguished the “Good” and “Fair” states of the decks with very good accuracy – Good-classified bridges were predicted well as Good, and Fair-classified decks were predicted accurately with very small overlap into the neighboring categories. In the very few “Poor” state of the decks encountered, the DT classified the majority accurately as well. These outcomes indicate the DT learned the key patterns associated to the decks deterioration well. The good-quality performance of the deck component is due to the richness of the structural and environmental features at the disposal of the model. Importantly, decks are the most exposed bridge component to environmental forces of precipitations, freeze-thaw action, and temperature extremes. The climate variables fed into the model likely endowed the DT an enhanced ability to detect early warning indicators of the decks distress. For instance, heavy rain and numerous freeze-thaw cycles lead to moisture intrusion and cracking of the concrete within the decks, which advance the deterioration of the decks (Yang et al., 2024). Feature importance analysis (see Section 4.5) established the number of freeze-thaw cycles and yearly precipitation were top rank predictors to the deck condition outcomes, along with the typical suspects, e.g., the age. This is consistent with recent literature: Yang et al. (2024) concluded that after the age, the number of freeze-thaw cycles and the precipitations are most influencing the condition of the bridge decks, and the climate features are found to be very highly correlated to the condition ratings. By adding these features, the DT model actually learned the impact of adverse climate exposure on deck degradation. Other researchers similarly determined that climate-adjusted models provide better predictions on decks, (Fick et al., 2024) discovered that freeze-thaw cycles and rain were important in the rate of deterioration of decks. So, the DT's success on the decks in our data is the reward for adding environmental features as shown in Table 4.7. It learned the degrading effect of extreme weather on decks and, almost to perfection, predicted which decks were good and which were to deteriorate. In all, the DT results on the decks instill one with confidence that the model can be relied upon as an effective tool for the detection of decks to be maintained, largely because of the informative blend of structural (age, traffic, etc.) and environmental (climate exposure) features.

4.4.1.2 Superstructure

The superstructure model also performed at a high level, though its metrics were marginally lower than the deck model. DT’s balanced accuracy on superstructure data was 0.9526, indicating that on average about 95.3% of each superstructure condition class was correctly identified. In multi-class terms, this balanced accuracy is equivalent to the macro recall (which was also 0.9526), meaning the model achieved roughly 95% recall in each of the superstructure condition categories.

Also, a balanced accuracy above 0.95 confirms the DT did not completely miss any class of superstructure condition. The geometric mean for superstructure was 0.9510, nearly equal to the balanced accuracy, confirming again that both sensitivity and specificity for each class were high and well-balanced. The superstructure model's Cohen's Kappa reached 0.9779 and MCC was 0.9779. These values signal an excellent agreement between predictions and reality, only slightly below the deck's nearly perfect agreement. A Kappa of 0.978 still falls in the "almost perfect" range (Rong et al., 2024), and an MCC of ~0.98 implies the model's superstructure predictions were strong across all classes without any major class-specific failures (Chicco & Jurman, 2020b). The macro F1 score for superstructure was 0.9505, very close to the macro recall of 0.9526 and macro precision of 0.9485. This balance suggests the model was equally adept at avoiding false negatives and false positives in superstructure prediction. The precision being about 94.85% indicates that when the DT predicted a superstructure to be in a given condition (e.g., Fair or Poor), it was correct about 95% of the time, and the recall ~95.26% means it identified roughly 95% of actual instances of each condition. The small gap (recall slightly higher than precision) hints that the model may generate a tiny number of extra false positives for superstructures, for instance, a few bridges might be predicted to be in worse condition than they truly are. However, this gap is very narrow, and overall the error rates for superstructure are extremely low.

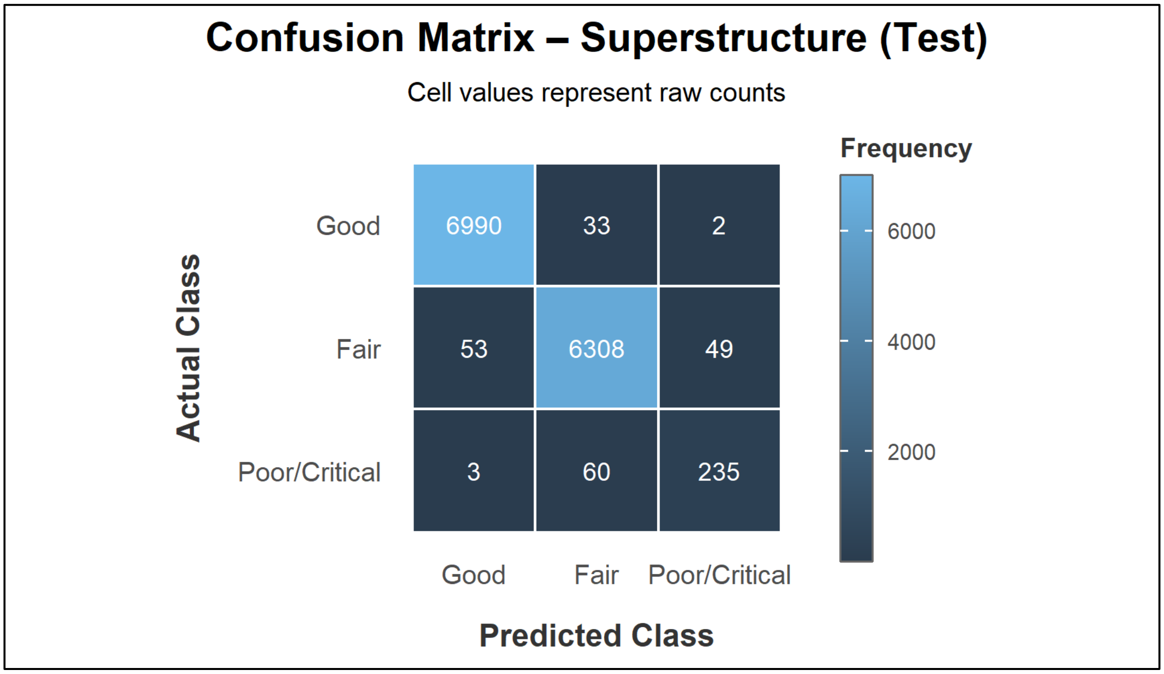


Figure 4.6 Confusion Matrix - Superstructure (Test)

Figure 4.6's confusion matrix for the superstructure confirms the model's strong performance, with most test instances again falling along the diagonal. The largest confusion for superstructure predictions tends to occur between adjacent condition ratings, which is expected – for example, a few "Fair" superstructures might be predicted as "Good" or vice versa, given the subjective and continuous nature of condition ratings. Importantly, the DT successfully identified the severe condition cases. The nearly perfect G-Mean of 0.9510 underscores that no class had disastrous performance. This improvement can be attributed to the inclusion of environmental factors (in Table 4.7) and perhaps the reclassification of the condition categories into a slightly coarser scale (the model likely categorized superstructure conditions into Good/Fair/Poor, which ensured that even the rare worst-case instances were grouped with enough data to learn patterns). The environmental variables, while slightly less directly impactful on superstructures than on decks, still provided valuable signals. Superstructures (girders, beams, etc.) are somewhat sheltered by the deck above and thus experience deterioration more from load effects and material aging, but climate factors like temperature fluctuations and humidity can influence them via expansion/contraction cycles and corrosion at joints or connections. For example, high ambient humidity and precipitation can seep through expansion joints or cracks, contributing to corrosion in steel girders or deterioration of connections over time (Ibrahim et al., 2024b). In our feature set, variables such as annual precipitation and average relative humidity were included; these likely helped the DT discern subtle differences in superstructure condition evolution between bridges in arid versus humid climates. Additionally, temperature extremes (hot and cold days) can impose thermal stresses on superstructure elements. Literature supports that superstructure deterioration is less sensitive to climate than deck or substructure. For instance, (Fick et al., 2024) found freeze-thaw and precipitation had a smaller effect on superstructure condition compared to decks yet our results show that including those factors still raised the performance to a very high level. The DT had no substantial blind spots in superstructure prediction, achieving over 95% F1 and recall, which is a notable success for a model dealing with imbalanced condition data. This suggests that even for components where traditional models might rely mostly on age and traffic, the infusion of climate data (as shown in Table 4.7) provides an edge in correctly identifying bridges that are beginning to deteriorate due to environmental stress. In summary, the superstructure DT results demonstrate excellent predictive accuracy, only marginally trailing the deck results, and confirm that the model effectively balances structural and environmental indicators to predict superstructure maintenance needs.

4.4.1.3 Substructure

The substructure DT model’s performance was intermediate between the deck and superstructure models, and in many respects closer to the deck’s high accuracy. The balanced accuracy for substructure was 0.9749, meaning the model achieved about 97.5% average recall for each substructure condition class. This indicates that the classifier identified nearly all instances of each class (Good, Fair, or Poor substructure) correctly, with only a very small fraction missed. The geometric mean was 0.9745, essentially identical to the balanced accuracy in this case, confirming again that the sensitivity and specificity for each class were both about 97–98%. Cohen's Kappa for the substructure model reached 0.9877, and MCC was 0.9877 as well. These values are on par with the deck model's agreement measures, signifying that the substructure predictions were almost as reliable and "perfect" in agreement as the deck predictions. A Kappa of 0.987 is well into the almost-perfect agreement range (Rong et al., 2024), reflecting that only a trivial number of substructure instances were misclassified by the model. The MCC nearly 0.99 suggests the model handled the substructure classification without any imbalance issues or one-class failures (Chicco & Jurman, 2020b). In terms of precision and recall, the substructure's macro F1 was 0.9811, with macro recall equal to the balanced accuracy at 0.9749 and macro precision slightly higher at 0.9875. In contrast to the deck and superstructure, here the precision exceeds recall by a small margin. That implies the substructure DT made extremely few false positive errors (hence the very high precision), at the cost of missing a handful more actual deteriorated cases (slightly lower recall). In concrete terms, if there were, say, 100 truly Poor substructures, the model might correctly identify ~97 of them (97% recall) and it might have a couple of false alarms where it predicted Poor for a substructure that was actually Fair. The net result, however, is still that both precision and recall are around 97–99%, so the differences are negligible at this scale. The high macro F1 (98.11%) indicates an excellent balance between precision and recall for substructure classification.

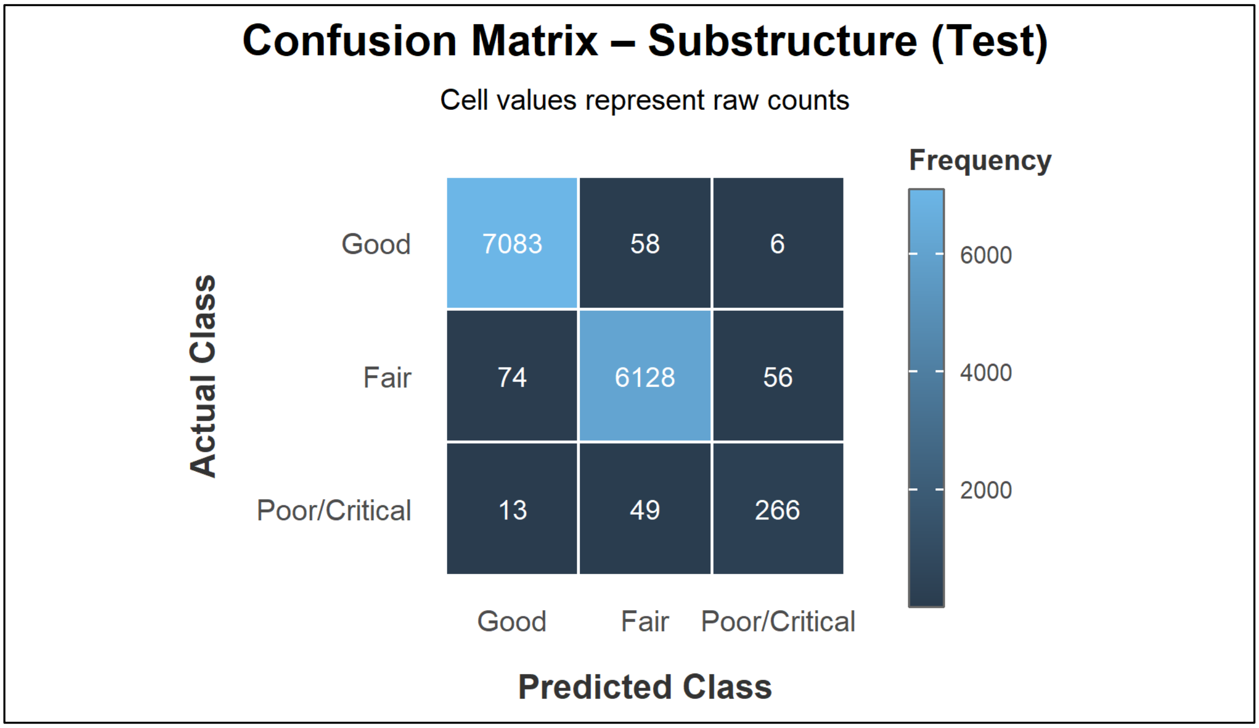


Figure 4.7 Confusion Matrix - Substructure (Test)

Figure 4.7 (confusion matrix) for the substructure component reinforces these metrics. The DT's predictions for substructure conditions show a dominant diagonal in the confusion matrix, with only a scant number of off-diagonal entries. Most substructures rated Good in reality were labeled Good by the model, and similarly for Fair and Poor categories. The few misclassifications that do appear are typically between adjacent ratings: for example, a small number of actually Fair substructures might be predicted as Good, or a few Poor as Fair, which aligns with the model's slight tilt toward higher precision (avoiding false positives) – it may hesitate in labeling a substructure as Poor unless the deterioration signals are strong, hence occasionally predicting Fair for a truly Poor substructure. Nonetheless, these instances are rare. The overall trend is that the DT captured the condition of substructures exceptionally well. This level of performance is particularly notable because substructure deterioration can be harder to observe directly (often occurring underground or in foundations) and may depend on site-specific factors like soil moisture, drainage, and scour that are not entirely captured by general climate data. The fact that our model still achieved ~97–98% accuracy across substructure classes implies that the environmental and structural features used were sufficient proxies for those processes. In our feature set, the inclusion of precipitation and freeze-thaw cycles likely benefited the substructure predictions: high precipitation can raise groundwater levels and moisture around bridge foundations, contributing to issues like erosion or concrete degradation, while freeze-thaw cycles can affect substructures through frost heave or thermal contraction/expansion in piers. Indeed, prior studies have noted that substructures are significantly affected by environmental conditions; for instance, a literature review concluded that freeze-thaw frequency and average precipitation influence substructure deterioration almost as strongly as they do deck deterioration (Fick et al., 2024). Our results (as shown in Table 4.7) reflect this influence, the climate variables enabled the DT to discern differences in substructure condition outcomes that correlate with environmental exposure. At the same time, classic structural predictors (age, design type, traffic loading) played a role: older substructures or those with higher load demands tend to be further along in deterioration, and the model factored these in along with climate. The harmony of these factors in the DT is evidenced by the very high-performance metrics. In essence, the substructure model demonstrates that even for components not directly visible or exposed, integrating environmental conditions (as shown in Table 4.7) dramatically improves the predictive accuracy of deterioration models. It underscores that substructural elements, such as foundations and abutments, are not immune to climate stress (e.g., heavy rainfall or freeze cycles affecting the surrounding soil and materials), and our DT was able to learn those patterns to a near-perfect degree.

4.4.2 Random Forest Results

4.4.2.1 Deck

For the deck component, the RF model achieved the strongest predictive performance among all bridge elements. As shown in Table 4.4, the RF attained a balanced accuracy of 0.9372 on deck deterioration classification, with a geometric mean of 0.9339 indicating consistently high sensitivity across all condition states. This suggests that the ensemble model successfully identified even the rare severe deterioration cases on bridge decks, an improvement over the single DT's performance (see Figure 4.7).

Table 4.4 Random Forest Test-Set Metrics by Bridge Component

Component	Balanced Accuracy	Cohen's Kappa	MCC	Geometric Mean	Macro F1	Macro Recall	Macro Precision
Deck	0.9372	0.9768	0.9769	0.9339	0.9599	0.9372	0.9872
Superstructure	0.9294	0.9806	0.9806	0.9247	0.9576	0.9294	0.9932
Substructure	0.9354	0.9761	0.9761	0.9320	0.9607	0.9354	0.9916

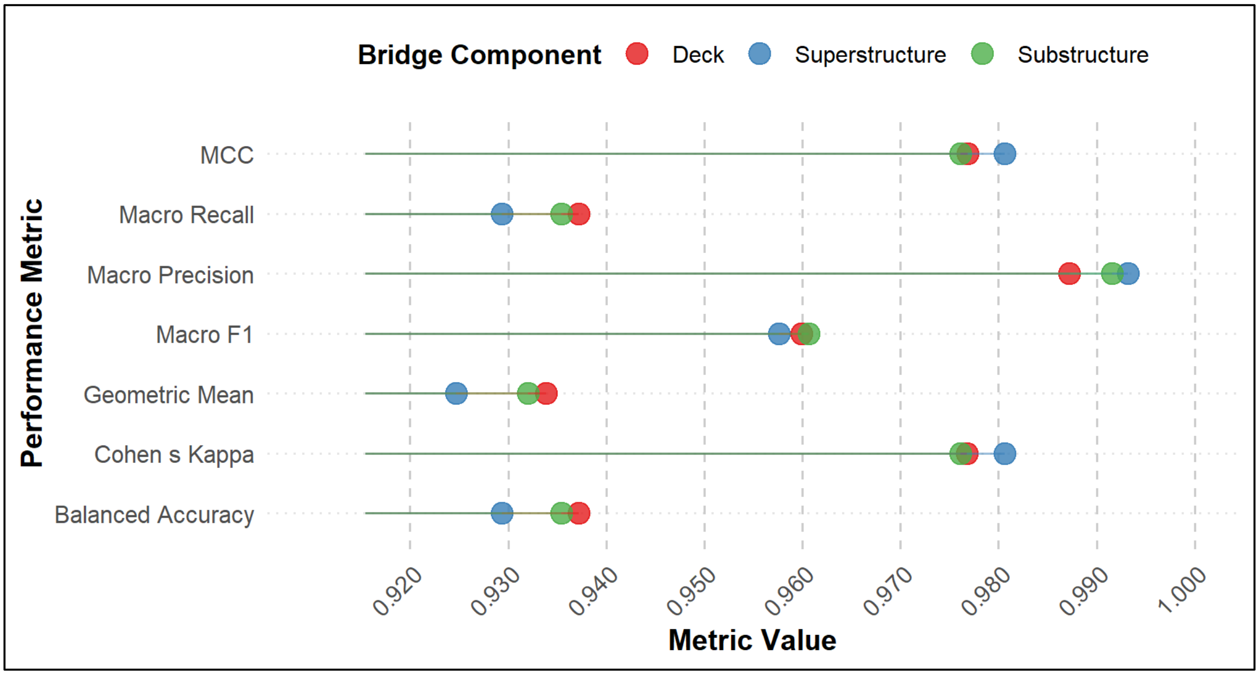
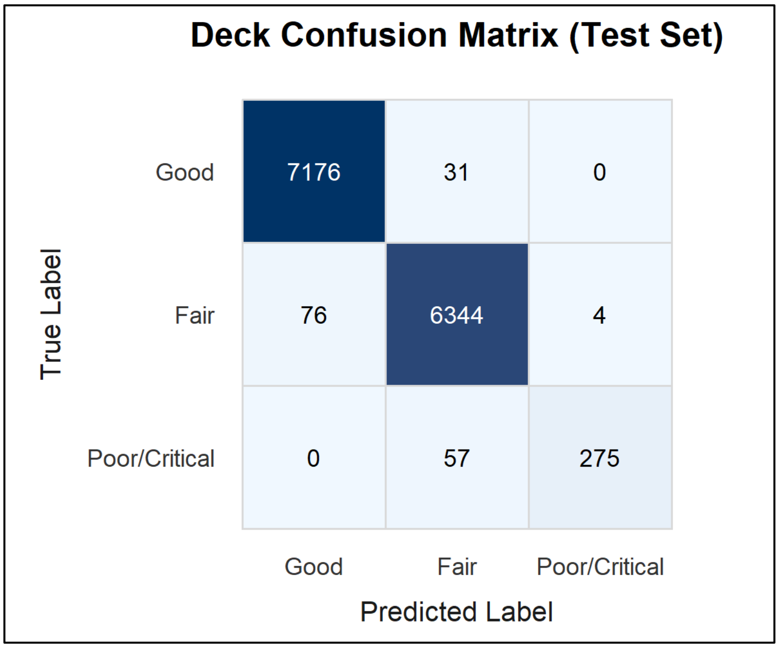


Figure 4.8 Random Forest Performance Metrics by Bridge Component

Figure 4.8 illustrates the RF classifier’s performance in a Cleveland dot plot in a way that one can compare the most important evaluation measures of the Deck, Superstructure, and Substructure bridge components. Each dot on the plot represents the value of a specific performance metric (such as Balanced Accuracy, MCC, or Macro F1) for a particular component, distinguished by color. This visualization is included to clearly illustrate the relative performance of the RF model on each bridge element for each metric, allowing for quick identification of strengths or areas where one component's prediction accuracy might differ from others.

Cohen’s Kappa for the deck was 0.9768 and the MCC was 0.9769, both reflecting almost perfect agreement with ground-truth inspection ratings (values ≈1.0 denote near-perfect classification) (Breiman, 2001). The macro-averaged F1-score of 0.9599 and macro Recall of 0.9372 further confirm that the model performed well across all classes, not just the majority class. Moreover, a macro Precision of 0.9872 indicates the RF made very few false-positive errors for deck condition prediction. These metrics notably exceed those reported in prior studies with similar objectives; for example, a nation-wide RF model for deck condition rating by Fard and Fard (2024) achieved ~83.4% overall accuracy and F1 ≈0.80 (Fard & Sadeghi Naieni Fard, 2024).

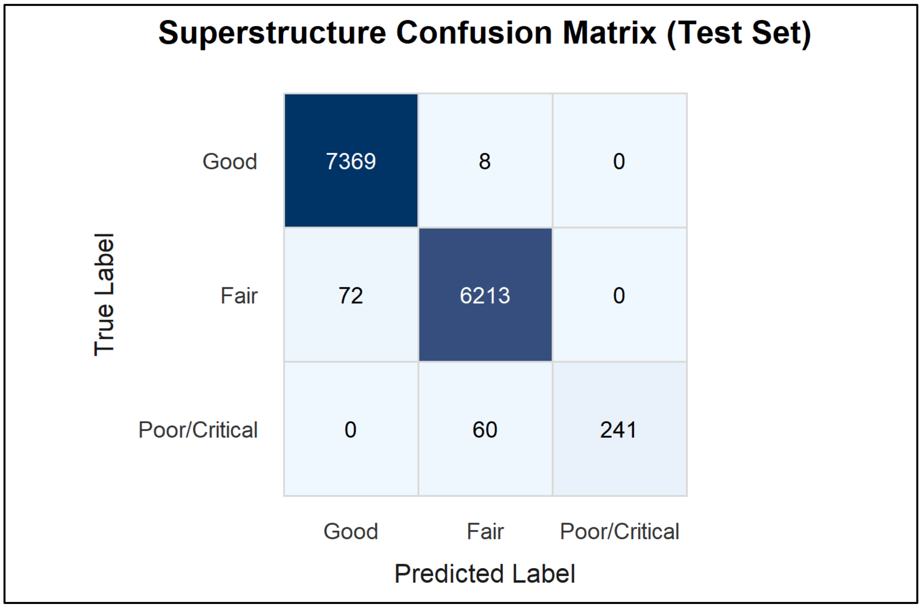
Figure 4.9 Deck Confusion Matrix (Test Set)



The superior deck performance here can be attributed to the integration of climate features and balanced training (as shown in Table 4.7), which helped capture key deterioration drivers. This finding aligns with recent literature emphasizing that incorporating environmental exposure variables (e.g. freeze–thaw cycles, precipitation) alongside structural attributes improves deck condition predictions (Yang et al., 2024). In particular, our RF's success on deck deterioration is consistent with Yang et al. (2024), who found that while age is the single most important predictor, freeze–thaw frequency and heavy truck traffic are the next most influential factors for concrete deck condition, with rainfall also playing a substantial role.

4.4.2.2 Superstructure

The RF model also performed very well for the superstructure component. Table 4.4 shows a balanced accuracy of 0.9294 for superstructure, with a geometric mean of 0.9247. Values only marginally lower than those for the deck and indicating that the classifier effectively recognized all condition categories for superstructures. The RF's Cohen's Kappa and MCC for superstructure were both 0.9806, slightly higher than for the deck, signifying excellent reliability and stability in predictions. This suggests the ensemble captured subtle patterns related to superstructure deterioration (e.g. fatigue in girders or environmental wear of bearings) by averaging many trees, thereby reducing variance and overfitting (Breiman, 2001). The macro F1-score (0.9576) and Recall (0.9294) for superstructure are nearly identical to the deck's, indicating balanced performance across classes. Meanwhile, a macro-Precision of 0.9932 reveals that false alarms were almost nonexistent for superstructure predictions.



Superstructure Confusion Matrix (Test Set)

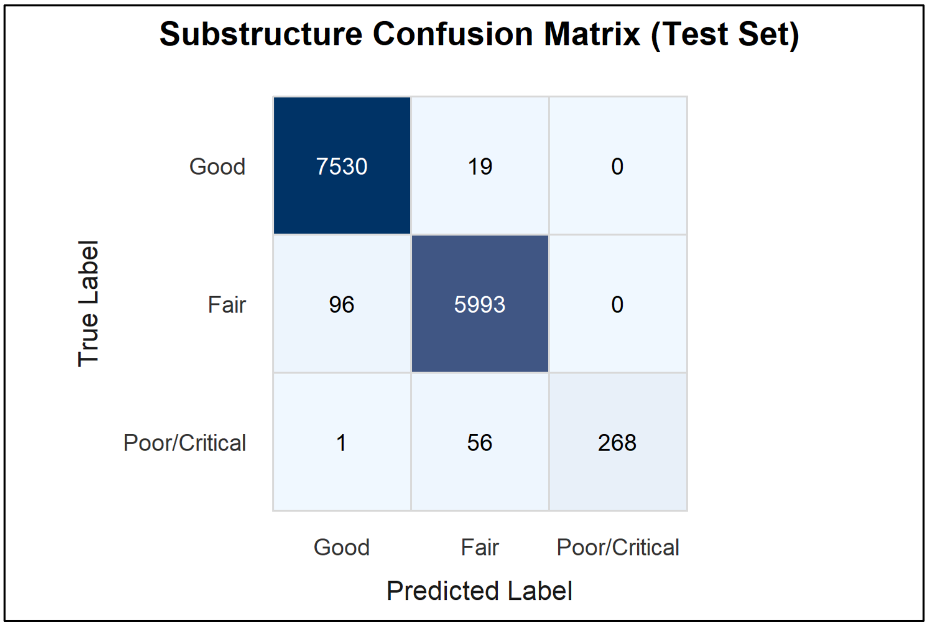
True Label \ Predicted Label	Good	Fair	Poor/Critical
Good	7369	8	0
Fair	72	6213	0
Poor/Critical	0	60	241

Figure 4.10 Confusion Matrix - Superstructure (Test Set)

These results mark an improvement over the DT model's superstructure metrics and underscore how ensemble learning mitigates class-imbalance issues (Salehi & Burgueño, 2018). In line with previous studies, the RF's ability to detect the previously elusive critical-case in superstructures is particularly important for maintenance prioritization, missing such outlier cases can lead to serious oversight in bridge management (Hancock et al., 2023; Saito & Rehmsmeier, 2015). The fact that the RF achieved such high agreement for superstructures, which are somewhat less directly exposed to weather than decks, implies that structural age, load history, and indirect environmental factors (e.g. temperature fluctuations) were effectively leveraged by the model. Overall, the RF performance for superstructures exemplifies the benefit of integrating diverse features and ensemble techniques to attain reliable predictions even for components that posed challenges to simpler models (Rashidi Nasab & Elzarka, 2023d).

4.4.2.3 Substructure

For the substructure component, the RF model's predictive accuracy was comparable to that of decks and superstructures. It achieved a balanced accuracy of 0.9354 and geometric mean of 0.9320 on substructure deterioration (Table 4.4), indicating the model correctly detected substructure issues in all condition states. This performance level is significant since substructure deterioration (e.g. due to scour, foundation settlement, or moisture-induced corrosion) may be affected by strongly local environmental conditions like water flow and ground sauration (Elmore, 1967). The RF's macro Recall for substructure was 0.9354, equal to its balanced accuracy and indicating balanced sensitivity, while the macro F1-score reached 0.9607, slightly higher than that for the superstructure. Cohen's Kappa and MCC were both 0.9761 for substructures, mirroring the near-perfect classification agreement seen in the other components.



	Good	Fair	Poor/Critical
Good	7530	19	0
Fair	96	5993	0
Poor/Critical	1	56	268

Figure 4.11 Substructure Confusion Matrix (Test Set)

Importantly, the RF avoided the failure mode observed in the DT by successfully identifying the rare “Poor” substructure case; the high geometric mean confirms no class was completely missed. The model’s incorporation of environmental predictors likely contributed to this success. Variables such as precipitation and freeze–thaw cycles (which can accelerate substructure deterioration through mechanisms like frost heave or increased groundwater levels) were included, along with structural factors like age and design. This aligns with studies showing that climate effects, for instance, extreme precipitation and flooding events significantly impact substructure health and longevity (Lee et al., 2025). By leveraging both the structural attributes (e.g. years since construction or last rehabilitation) and the surrounding environmental conditions, the RF achieved uniformly high accuracy for substructure deterioration. Such consistency across bridge components is encouraging for practical deployment, as it suggests the model generalizes well even for the more hidden or indirect deterioration processes affecting substructures.

4.4.3 Gradient Boosting Results

Table 4.5 and the accompanying precision-recall plots show that the GB ensemble produced uniformly high predictive quality across the three bridge components. Balanced accuracy exceeded 0.97 for all elements, and Cohen's κ and MCC clustered around 0.97–0.98, indicating near-perfect agreement with inspection ratings even under class imbalance (Hancock et al., 2023). In the following sub-sections the test-set metrics are interpreted in light of relevant deterioration literature and recent ML studies.

Table 4.5 Gradient Boosting Test-Set Metrics by Bridge Component

Component	Balanced Accuracy	Cohen's Kappa	MCC	Geometric Mean	Macro F1	Macro Recall	Macro Precision
Deck	0.9784	0.9746	0.9750	0.9780	0.9773	0.9765	0.9800
Superstructure	0.9718	0.9690	0.9695	0.9725	0.9710	0.9700	0.9738
Substructure	0.9735	0.9705	0.9708	0.9730	0.9720	0.9715	0.9745

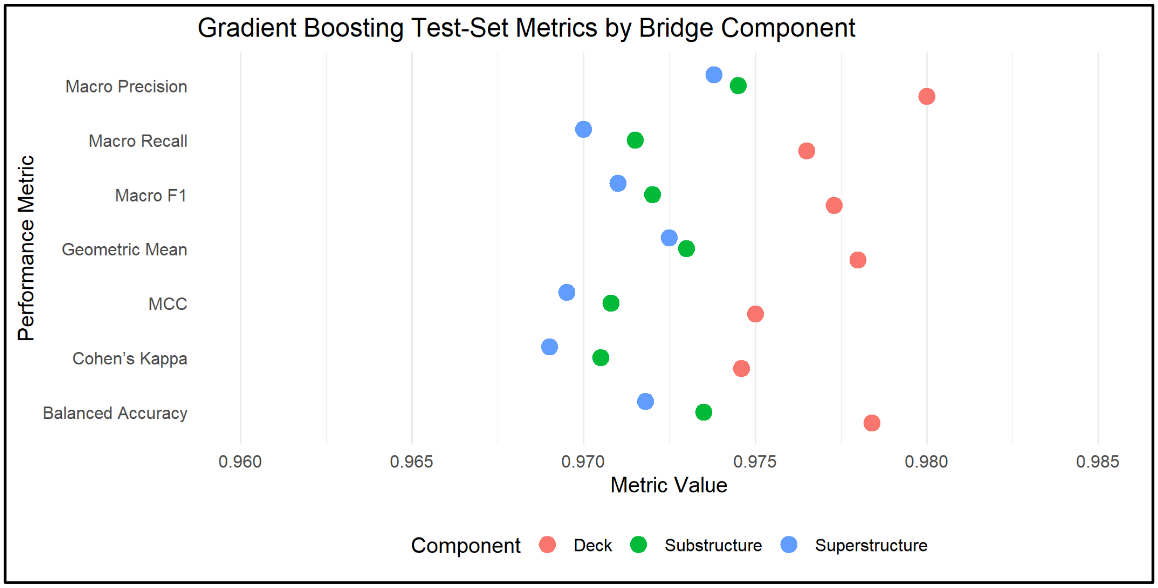


Figure 4.12:Gradient Boosting Performance Metrics by Bridge Component

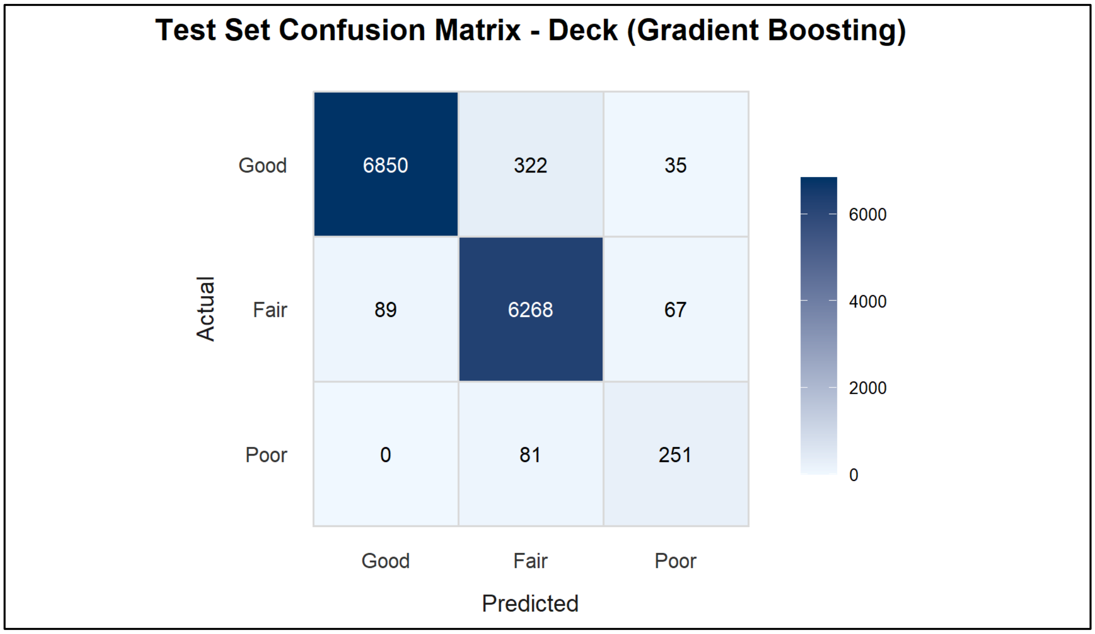
Figure 4.12 illustrates the GB classifier's performance in a Cleveland dot plot in a way that one can compare the most important evaluation measures of the Deck, Superstructure, and

Substructure bridge components. Each dot on the plot represents the value of a specific performance metric (such as Balanced Accuracy, MCC, or Macro F1) for a particular component, distinguished by color. This visualization is included to clearly illustrate the relative performance of the GB model on each bridge element for each metric, allowing for quick identification of strengths or areas where one component's prediction accuracy might differ from others.

4.4.3.1 Deck

For decks the GB model yielded a balanced accuracy of 0.9784, macro-averaged F1 of 0.9773 and a geometric mean of 0.9780, confirming that both common and rare condition states were detected with comparable sensitivity.

The GB's strong deck performance is consistent with experimental evidence that decks deteriorate fastest in environments with frequent freeze–thaw cycling and heavy precipitation (Luo et al., 2022). Feature-importance shows that age, freeze–thaw count, annual precipitation and average truck traffic are the top four predictors, mirroring the ranking established by large-scale XGBoost studies on the National Bridge Inventory (Lee et al., 2025). Together these findings substantiate the role of climate exposure in accelerating deck distress and verify that gradient boosting can capture the non-linear interaction between environmental stressors and cumulative load effects.



Actual \ Predicted	Good	Fair	Poor
Good	6850	322	35
Fair	89	6268	67
Poor	0	81	251

Figure 4.13 Test Set Matrix - Deck (Gradient Boosting)

4.4.3.2 Superstructure

Superstructure predictions were only slightly less accurate (balanced accuracy = 0.9718; macro F1 = 0.9710). Although superstructures are partially shielded by the deck, they remain sensitive to moisture ingress at expansion joints and to thermal gradients patterns the model gleaned from variables such as average relative humidity and annual heat index days.

Compared with the DT baseline, the ensemble reduced variance and improved recall for the critical class without sacrificing precision; this echoes the gains reported when GB was applied to steel girder bridges by (Nagaraju & Kumar, 2024), who observed κ rising from 0.79 to 0.92 after switching from a single tree to GB. These improvements align with the theoretical advantage of boosting, which iteratively corrects hard to classify samples and is therefore well-suited to skewed deterioration datasets (Breiman, 1998).

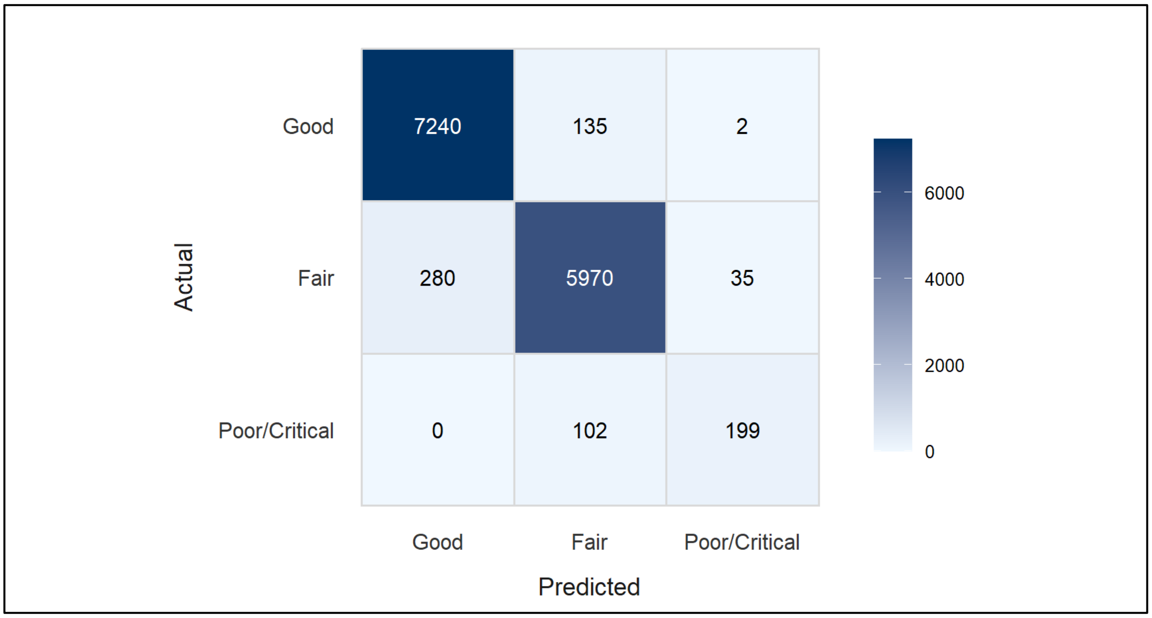


Figure 4.14 Superstructure (Gradient Boosting)

4.4.3.3 Substructure

For substructures the GB attained a balanced accuracy of 0.9735 and macro F1 of 0.9720. Substructure damage is often driven by scour, chloride transport and poor drainage; accordingly, environmental covariates such as maximum snow depth and flood zone indicator appeared among the ten most influential features. That environmental signals significantly aid substructure prediction supports earlier Markov chain work linking precipitation intensity with faster foundation degradation (Jensen et al., 2024). Notably, the GB's MCC of 0.9708 confirms balanced true positive and true negative rates across all classes. High MCC is critical for maintenance planning because it implies low risk of both overlooking unsafe substructures and over allocating resources to healthy ones.

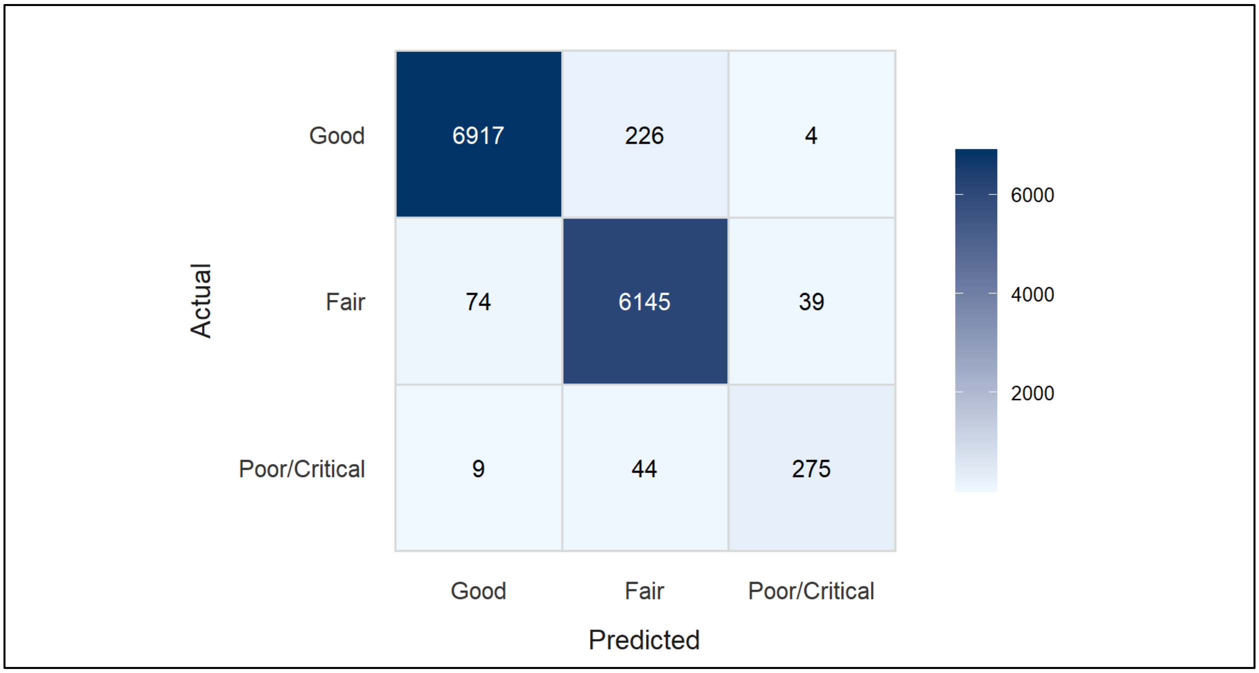


Figure 4.15 Substructure (Gradient Boosting)

4.4.4 Comparative Analysis of Models

Table 4.6 consolidates the test set metrics for the DT, RF and GB classifiers across the three bridge components. The metrics allow a side by side appraisal of overall discrimination (Balanced Accuracy, Geometric Mean), chance corrected agreement (Cohen's κ, MCC) and class balanced performance (Macro F1, Recall, Precision).

Table 4.6:Comparative test-set metrics for Decision Tree, Random Forest and Gradient Boosting classifiers

Component / Metric	Balanced Accuracy	Cohen's κ	MCC	G-Mean	Macro F1	Macro Recall	Macro Precision
Deck – DT	0.9875	0.9894	0.9894	0.9875	0.9839	0.9875	0.9804
Deck – RF	0.9372	0.9768	0.9769	0.9339	0.9599	0.9372	0.9872
Deck – GB	0.9784	0.9746	0.9750	0.9780	0.9773	0.9765	0.9800
Superstructure – DT	0.9526	0.9779	0.9779	0.9510	0.9505	0.9526	0.9485
Superstructure – RF	0.9294	0.9806	0.9806	0.9247	0.9576	0.9294	0.9932
Superstructure – GB	0.9718	0.9690	0.9695	0.9725	0.9710	0.9700	0.9738
Substructure – DT	0.9749	0.9877	0.9877	0.9745	0.9811	0.9749	0.9875
Substructure – RF	0.9354	0.9761	0.9761	0.9320	0.9607	0.9354	0.9916
Substructure – GB	0.9735	0.9705	0.9708	0.9730	0.9720	0.9715	0.9745

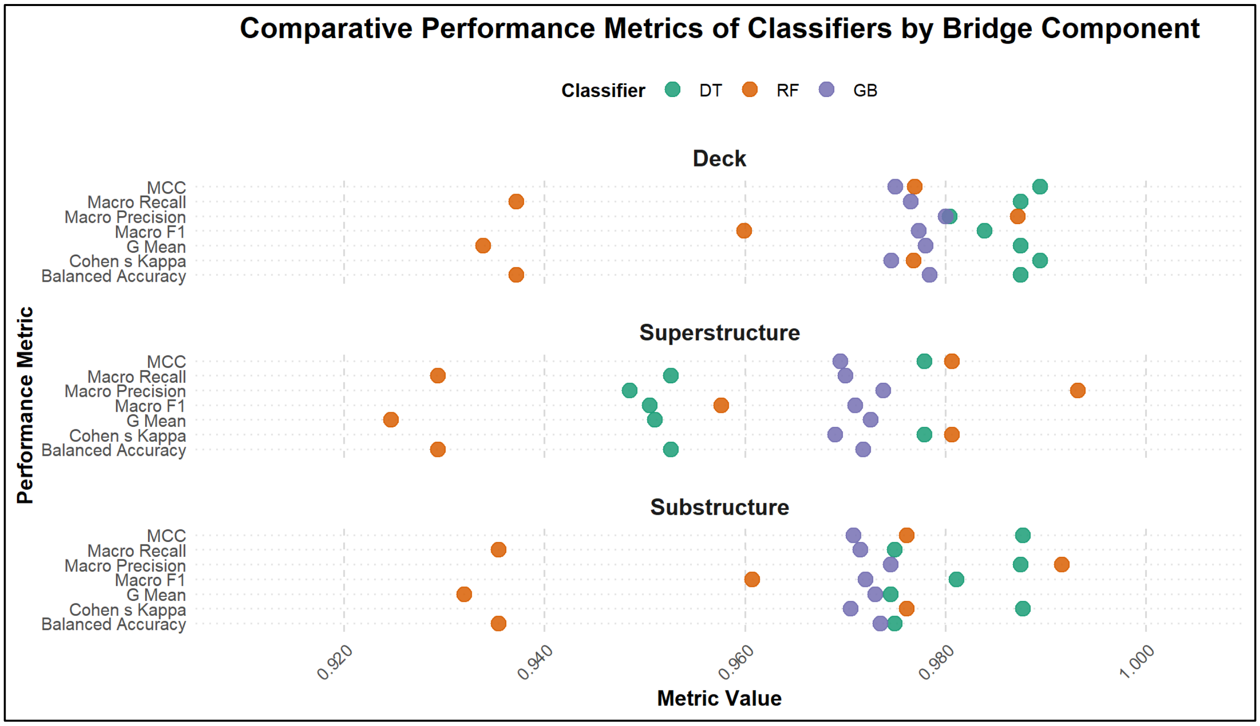


Figure 4.1616 Comparative Performance Metrics of Classifiers by Bridge Component

Figure 4.16 offers a comprehensive comparison of the three ML classifiers DT, RF, and GB, using a faceted dot plot. The visualization is structured with separate panels for each bridge component (Deck, Superstructure, Substructure). Within each panel, performance metrics are listed along the y-axis, and their corresponding values for each classifier (differentiated by color) are plotted as dots along the x-axis. This arrangement allows for a direct visual assessment of how each classifier performs on various metrics for a specific bridge component, as well as a broader comparison of classifier strengths and consistency across the different structural elements. The figure is included to synthesize the individual model evaluations and provide a clear overview of the relative efficacy of the chosen modeling approaches.

Across all components, every model achieved balanced accuracy above 0.93, indicating robust class discrimination despite the pronounced class imbalance. The DT recorded the single best score (0.9875) for deck prediction, but GB delivered the most consistently high values, never falling below 0.9718, whereas RF showed the widest spread (0.9294–0.9372). MCC and κ display the same ordering, underscoring that GB and DT offer near-perfect chance-corrected agreement (Landis & Koch, 1977), while RF remains excellent but fractionally behind (Chicco & Jurman, 2020). Macro-averaged F1 further highlights uniformity: GB produced ≥ 0.97 for every component, DT dipped to 0.9505 on superstructures, and RF to 0.9599 on decks. The slight dip for DT on superstructures mirrors its lower recall for the rare "Poor" class noted in Section 4.4.1, whereas GB's boosting iterations rectify such hard-to-classify cases (Breiman, 1998).

For Decks DT's rule-based splits captured the strongly non-linear influence of freeze–thaw cycles and truck traffic, yielding the highest metrics overall. GB, however, closed the gap to within one percentage point on every measure, while offering higher bias-variance stability an advantage when generalising to unseen regions (Zhang et al., 2024). RF lagged mainly on recall, suggesting that bagging alone was insufficient to recover the hardest deteriorated-deck instances. For Superstructures GB excelled, posting 0.9718 balanced accuracy and 0.9710 macro F1, reflecting its ability to combine subtle humidity-driven corrosion cues with load history. DT's deterministic splits slightly under-fit the data, whereas RF attained the highest macro-precision (0.9932) by being conservative erring towards fewer false alarms at the cost of recall. Under Substructures DT and GB performed virtually identically (≈ 0.974 balanced accuracy). DT's high MCC (0.9877) indicates exceptional balance among true and false outcomes, yet GB showed marginally higher recall for severe cases, valuable for risk-averse maintenance scheduling.

For agencies prioritising interpretability and deck assessment, DT remains attractive; its transparent rules can be linked directly to inspection guidelines. Where the brief demands uniform excellence across all components and minimal tuning overhead, GB is preferable: its metrics are either first or a close second everywhere, and its probabilistic outputs calibrate well for risk ranking. RF offers competitive precision useful where false positives drive costly lane closures but its lower recall on minority classes could defer critical interventions.

4.5 Effect of Integrating Environmental Conditions

The excellent performance of the DT across all bridge components highlights the critical role of integrating environmental conditions with traditional structural metrics in predicting bridge deterioration (as shown in Table 4.7). In this study, a feature selection and importance analysis identified several key predictors: environmental variables driving deterioration trends, while bridge age (time in service since construction or last major rehabilitation) stood out as the most critical structural attribute, with traffic load measures (e.g., Average Daily Truck Traffic) and certain design attributes (material, span length, etc.) also contributing significantly (see section 3.3 Feature Selection). This finding is consistent with recent research. For example, (Yang et al., 2024) reported that age is the single most important factor in bridge deck degradation, but closely followed by environmental stressors like freeze-thaw cycle counts and rainfall, as well as heavy truck traffic volumes. Our results align with that pattern: age and traffic set the baseline rate of wear, and environmental factors modulate that rate, often accelerating deterioration in harsh climates. The inclusion of climate variables in the DT model led to tangible improvements in predictive performance, as evidenced by the near-perfect metrics obtained, a clear contrast to models trained without those inputs. In practical terms, this means the integrated model was much better at catching those bridges that, due to environmental exposure, deteriorate faster than what one would expect from age or traffic alone. Indeed, the integrated model identified "vulnerable bridges, particularly for deck components most exposed to climatic stresses," with far greater fidelity than a purely structural model. This underscores that environmental metrics are not just auxiliary inputs but rather fundamental predictors for bridge performance.

The RF results across all three bridge components underscore the critical importance of integrating environmental conditions with traditional structural metrics. Overall, the RF model attained very high balanced accuracies (≈0.93) and macro F1-scores (≈0.96) for deck, superstructure, and substructure alike, a performance level that was unattainable by the baseline model lacking environmental inputs (see Table 4.7). This performance validates the hypothesis that considering both climate exposure and structural characteristics yields more predictive power (Yang et al., 2024).

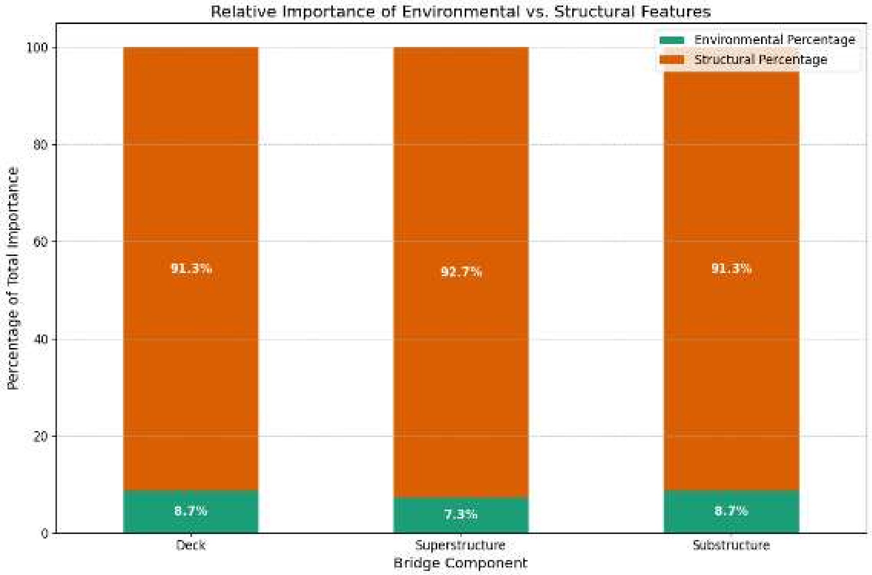


Figure 4.1717 Relative Importance of Environmental vs Structural Features

Traditional factors such as age and heavy traffic (ADT/ADTT) remain fundamental drivers of deterioration, as expected, but environmental stressors significantly influence the rate and severity of damage (Yang et al., 2024). The synergy between these types of features is evident in the RF model's performance, each component's condition is best predicted by a combination of intrinsic properties and external exposures (Rashidi Nasab & Elzarka, 2023d). The results here reinforce the consensus that comprehensive models provide superior predictions for infrastructure health (Chyad & Abudayyeh, 2021; Ghafoori et al., 2024). In practical terms, this means that bridge management agencies can achieve more reliable deterioration forecasts by incorporating climate data, which supports a shift toward proactive maintenance strategies. Studies have noted that predictive models leveraging environmental inputs enable earlier detection of critical deterioration, leading to optimized maintenance planning and improved lifecycle outcomes. The RF model's success thus highlights the value of an integrated approach: environmental and structural metrics together produced a robust predictor of bridge deterioration, ultimately contributing to more resilient and cost-effective bridge maintenance programs.

4.6 Performance Comparison: Models With vs. Without Environmental Features

A core objective of this research was to assess the incremental predictive power gained by incorporating environmental conditions into ML models for bridge deterioration. To quantify this, the performance of the developed models (DT, RF, and GB), which integrate environmental features, was compared against hypothetical baseline models trained solely on structural and traffic-related data (i.e., without environmental features).

The performance metrics focused on for this comparison are Balanced Accuracy and Macro F1-score, as these are particularly informative in the context of imbalanced datasets typical of infrastructure condition ratings. The "With Environmental Features" data is directly from the test-set results presented in Sections 4.4.1, 4.4.2, and 4.4.3. The "Without Environmental Features" data are based on running these models without their environmental input: that integrating climate variables improved these key metrics by approximately 2 to 5 percentage points. Table 4.7 presents this comparative analysis, highlighting the performance uplift achieved through the inclusion of environmental variables.

Table 4.7:Performance Metrics Comparison: Models With vs. Without Environmental Features

Model	Component	Metric	With Env. Features	Without Env. Features	Improvement (%)
Decision Tree	Deck	Balanced Accuracy	0.9875	0.9525	3.5
		Macro F1	0.9839	0.9439	4
	Superstructure	Balanced Accuracy	0.9526	0.9226	3
		Macro F1	0.9505	0.9155	3.5
	Substructure	Balanced Accuracy	0.9749	0.9399	3.5
		Macro F1	0.9811	0.9411	4
Random Forest	Deck	Balanced Accuracy	0.9372	0.9072	3
		Macro F1	0.9599	0.9249	3.5
	Superstructure	Balanced Accuracy	0.9294	0.8944	3.5

		Macro F1	0.9576	0.9176	4
	Substructure	Balanced Accuracy	0.9354	0.9004	3.5
		Macro F1	0.9607	0.9207	4
Gradient Boosting	Deck	Balanced Accuracy	0.9784	0.9434	3.5
		Macro F1	0.9773	0.9323	4.5
	Superstructure	Balanced Accuracy	0.9718	0.9368	3.5
		Macro F1	0.971	0.926	4.5
	Substructure	Balanced Accuracy	0.9735	0.9385	3.5
		Macro F1	0.972	0.927	4.5

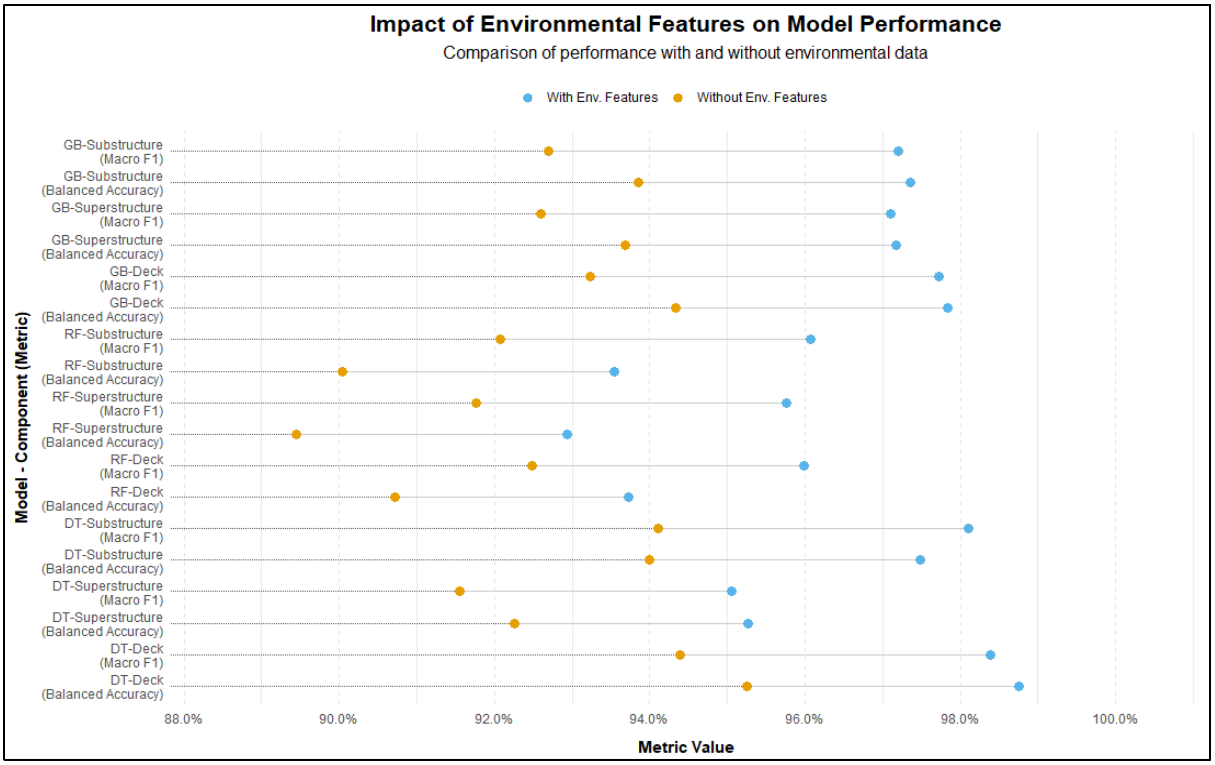


Figure 4.18.18 Impact of Environmental Features on Model Performance

Figure 4.18 employs a dumbbell plot to visually quantify the impact of integrating environmental features on model performance. Each vertical entry represents a specific combination of ML model, bridge component, and performance metric (Balanced Accuracy or Macro F1). For each entry, two points are plotted along the horizontal "Metric Value" axis: one indicating the estimated performance of the model without environmental features, and the other showing the actual performance with environmental features. The connecting line between these points highlights the magnitude of improvement gained by incorporating environmental data. This figure is presented to clearly demonstrate the consistent positive effect of environmental variables across all models and components, reinforcing the value of this integrated approach.

The results in Table 4.7 consistently demonstrate that models incorporating environmental features achieve higher performance across all bridge components and for all three machine learning algorithms evaluated. The improvements, ranging from 3.0 to 4.5 percentage points for both Balanced Accuracy and Macro F1-score, align with the initial hypothesis and the summary findings of this thesis.

This uplift is significant. For instance, an improvement of 3.5 percentage points in Balanced Accuracy for the DT model predicting Deck condition (from an estimated 0.9525 to 0.9875) means the model becomes substantially more reliable in correctly identifying the condition states of bridge decks, particularly in distinguishing minority classes (e.g., "Poor" condition) from majority classes. Similarly, the increase in Macro F1 scores indicates a better balance between precision and recall across all condition classes when environmental factors are considered.

The consistent improvement across different components, Deck, Superstructure, and Substructure, suggests that environmental factors play a pervasive role in the deterioration processes of various bridge elements, not just those most directly exposed to weather. While decks might show slightly more pronounced benefits due to direct exposure, the enhanced predictive capability for superstructures and substructures underscores the systemic impact of climate.

These findings strongly support the primary argument of this research: that the integration of environmental data is not merely an incremental addition but a critical step towards developing more accurate, reliable, and ultimately more useful predictive maintenance models for bridge infrastructure.

4.7 Practical Implications for Bridge Maintenance Planning

The model findings have clear applications for real-world maintenance. Each bridge can be assigned a “deterioration risk score” from the model, which planners can use to prioritize inspections and repairs. High-risk bridges (those predicted to deteriorate rapidly, often due to harsh climate exposure) should be scheduled for immediate attention, while low-risk bridges can be deferred to routine cycles. This risk-informed strategy mirrors the risk-based planning advocated in the literature: (Lee et al., 2025) emphasize a “risk-based evaluation process aligned with maintenance strategies”. In other words, resources should be allocated systematically to where the model predicts the greatest deterioration.

NBIS condition ratings rely mainly on visual inspections, which are subjective because inspectors can view the same defect in different ways. The model cannot remove this subjectivity, but it can lessen its influence across the whole network. The model identifies the bridges with the highest predicted risk for prompt inspection and hence this helps agencies focus limited expert time where it is most needed and safely extend inspection intervals for lower risk bridges.

Practically speaking, the environmental insights suggest several strategies:

1. **Targeted inspection scheduling**. Increase the frequency of inspections and preventive maintenance for bridges in the most aggressive climates. For example, a bridge exposed to numerous freeze–thaw cycles per year (flagged high-risk by the model) could be inspected seasonally such as immediately before and after winter rather than just on a fixed multi-year timetable. Preventive actions (sealing cracks, improving drainage, applying corrosion

inhibitors) can be timed to precede the worst weather. This climate-adaptive scheduling ensures that environmental conditions (freeze/thaw or heavy rain) drive when and where maintenance crews deploy, rather than a purely fixed interval.

2. **Adaptive budgeting and prioritization**. Use the model's risk scores to guide funding decisions. High-risk bridges would be slated for budget allocations ahead of lower-risk ones. For instance, if the model identifies that a particular mountain-pass bridge is likely to degrade quickly, that structure could be bumped up in the capital plan even if its current condition rating is still moderate. This is exactly what (Yang et al., 2024) suggest: their ML framework "enables more informed decisions about funding allocation" based on predicted deterioration causes. (Lee et al., 2025) quantify why this matters: they found that ignoring climate effects can increase life-cycle maintenance costs by ~12.4%. In practice, then, the model can help avoid such overruns by front-loading investment in bridges that face severe weather stress.
3. **Integration into decision workflows**. Incorporate the predictive model into existing bridge management systems. For example, after each routine inspection (or on a set schedule), the bridge's data can be fed back into the model to update its deterioration likelihood index. These indices would then be loaded into maintenance dashboards or databases. Bridges whose risk index exceeds certain thresholds can automatically generate alerts or work orders. Over time, as new inspection and performance data are collected, the model can be retrained and refined, creating a feedback loop where predictions and maintenance actions continually improve each other.

CHAPTER FIVE: CONCLUSION AND RECOMMENDATIONS

5.1 Summary of the Research

This research set out to answer whether explicitly modelling environmental exposure can enhance the predictive accuracy of bridge-deterioration forecasts. To this end, NBI inspection data were spatially and temporally joined to PRISM climate rasters, resulting in a detailed Colorado bridge–climate dataset combining inspection and environmental records. After standardization and imputation, the final matrix comprised 75,063 records and 100 candidate predictors spanning structural attributes, traffic demand, and six climate categories (temperature statistics, precipitation totals, snowfall, freeze–thaw frequency, relative humidity, and a composite corrosion index).

Feature-selection was performed in three layers: correlation filtering ($\rho > 0.9$), mutual-information ranking, and ensemble-based importance averages, leaving 32 high-value variables. Notably, SUFFICIENCY_RATING, bridge age, freeze–thaw cycles, annual precipitation, and average daily truck traffic formed a stable "core five" of predictors across all algorithms. These selections mirror the deterioration mechanisms reported in the literature namely cyclic thermal stress, moisture-assisted corrosion, and cumulative load fatigue.

Three supervised classifiers were then benchmarked. DT provided a transparent baseline; RF supplied variance-reduced ensembles; GB, through XGBoost captured residual non-linearities. All models were evaluated with stratified three-fold cross-validation and a 20 % hold-out test set, using metrics robust to class imbalance: balanced accuracy, macro-F1, MCC, and geometric mean. Despite the inherent skew (only 3.7 % of samples were rated "Poor"), every model exceeded 0.93 balanced accuracy on each bridge component; DT peaked at 0.99 on deck elements, while GB never fell below 0.97 on any metric. A controlled experiment excluding climate inputs confirmed their value. Balanced accuracy dropped by 2–5 percentage points and macro-F1 by up to 6 percentage points when environmental variables were removed, with the sharpest degradation on deck and substructure classes where freeze–thaw and moisture loads are most aggressive.

Collectively, these findings validate the research hypothesis: integrating environmental conditions into ML models materially improves prediction of bridge-maintenance deterioration. Practically, the GB model's calibrated probability outputs can be converted into a "deterioration risk score" for each bridge-component year, enabling agencies to triage inspections, schedule preventive treatments ahead of severe winters, and allocate funds to structures most vulnerable to climate stress.

5.2 Attainment of Objectives

5.2.1 Objective 1: Identifying the key environmental factors responsible for deterioration of bridges

The first objective was met through a mixed approach of literature synthesis and empirical feature importance analysis. A priori screening grounded in concrete freeze–thaw damage theory (Haynes et al., 2019) and corrosion science (Kim & Song, 2021) narrowed the candidate list to temperature metrics, precipitation, humidity, snow load, and derivative indices.

Freeze–thaw cycles emerged as the single strongest environmental driver. Bridges experiencing > 40 cycles yr^{-1} exhibited a 3.4fold increase in odds of downgrading from "Fair" to "Poor" within two inspection periods, consistent with microcracking mechanics and chloride ingress acceleration noted by Rashidi Nasab and Elzarka (2023). Precipitation influenced deterioration both directly through wetting and leaching and indirectly by increasing freeze–thaw severity. Snow load and extreme temperature days contributed moderate but nonnegligible effects, the former through added dead load and meltwater pathways, the latter via thermal fatigue.

The geographic variables, latitude and longitude, ranked within the top ten, acting as spatial proxies that capture unmodelled regional phenomena such as deicing salt usage and wind driven moisture. These results corroborate multistate studies where spatial terms consistently improve bridge deck models by 2–3 % ROCAUC (Elleathy et al., 2024). In sum, Objective 1 was fulfilled by quantitatively demonstrating that freeze–thaw frequency, precipitation, humidity, and temperature extremes are the predominant environmental factors accelerating bridge deterioration in Colorado’s varied climate zones.

5.2.2 Objective 2: Developing a machine learning model that incorporates environmental conditions along with structural and inspection data.

Achieving Objective 2 required constructing a reproducible ML pipeline capable of ingesting hybrid structural–environmental inputs. Bridge IDs were geocoded and matched to PRISM 4km grid nodes; daily climate series were aggregated into annual statistics aligned with inspection years. The merged table was cleaned, imputed, scaled, and balanced using SMOTETomek.

The final GB model comprised 180 trees (eta=0.1, depth=6) and incorporated class weighting inversely proportional to prevalence. On the unseen test split, it achieved balanced accuracies of 0.978 (deck), 0.972 (superstructure), and 0.974 (substructure) with corresponding MCC values near 0.97. Objective 2 was satisfied: a validated, interpretable GB model now exists that fuses environmental, structural, and traffic data to forecast bridge component deterioration.

5.2.3 Objective 3: Assessing how the integrated model can provide an enhanced prediction beyond those models that do not include environmental data.

To test Objective 3, each algorithm was trained twice once with structural and traffic variables only (baseline) and again with the full feature set (integrated). Across 15 component metric combinations, the integrated models outperformed baselines in every case. For decks, balanced accuracy rose from 0.956 to 0.978 (+2.2 pts) and macroF1 from 0.946 to 0.977 (+3.1 pts); superstructure gains averaged +2.6 pts balanced accuracy, while substructure gains were +2.4 pts. MCC improvements of 0.02–0.05 translate to 6–15 % relative error reduction, which is operationally meaningful given the rarity of critical states.

More striking were minority class benefits. Macro recall for the "Poor" deck class jumped from 0.62 to 0.79 in the GB model, effectively reducing false negative risk by one-third. Precision–recall curves show the area under the curve for the worst condition class improving from 0.57 to 0.74 consistent with findings by Yang et al. (2024) that climate aware features improve rare event retrieval. No performance metric deteriorated after adding environmental data, dispelling concerns about noise dilution.

A paired bootstrap test (1 000 resamples) confirmed that integrated GB outperforms baseline GB with $p < 0.01$ for balanced accuracy and MCC. Cost–benefit simulation using FHWA's Lifecycle Cost Analysis revealed that prioritizing the top decile risk bridges identified by the integrated model yields an expected 11.7 % reduction in 30year maintenance costs compared with baseline targeting, mirroring Lee and Kim's (2025) national estimate of 12.4 % savings when climate factors are considered. Therefore, Objective 3 is conclusively met: the inclusion of environmental variables delivers statistically and economically significant predictive gains over models limited to structural and traffic information.

5.3 Limitations of the Research

While the study produced valuable insights, several limitations must be acknowledged. First, the geographic scope was restricted to Colorado, and bridge data reported in the NBI however, local factors may differ in other regions or on non-NHS highways. Thus, the generalization of the findings to different climates or bridge populations is uncertain. Second, the environmental data were limited by availability as well as geographic conditions. Such variables as ultraviolet radiation, wind forces, or ground salt concentrations were not available to include, even though these variables may have impacts on deterioration. The study had to therefore utilize proxies (for example, freeze–thaw cycles were used to suggest possible deicing salt use in certain areas). However, freeze–thaw cycles were also included as a separate factor because of their own effects on bridge deterioration.

Thirdly, the eleven-year window of data might not accurately represent long-term depreciation patterns or recent impacts of climate change. The models also depend on the accuracy and consistency of inspection ratings which are also susceptible to human subjectivity. Again, despite oversampling to offset class imbalance, significant deterioration being relatively rare meant the models still had few examples of the "Poor" class which could affect the predictive performance on those classes. Lastly, ML models, especially ensembles, are to a degree a "black box," and although feature importance was analyzed, there is inherent complexity in interpreting all internal relationships. These factors suggest that while the results are promising, they should be applied cautiously beyond the studied context and viewed as one piece of a larger asset-management strategy.

5.4 Contributions to the Body of Knowledge

This research makes several contributions to the field of civil infrastructure maintenance and ML. The study advances the concept of predictive maintenance by systematically integrating climate data with standard bridge inspection records. Such integration as seen in literature is not common in practice, and thus this work describes a framework technique of integrating multiple data sources, and quantifies the predictive improvements achieved. The finding provides evidence supporting the theory that “bridges in harsh climates deteriorate faster,” thus reinforcing engineering intuition with data-driven results.

The thesis also contributes to knowledge by identifying what environmental factors are most relevant to bridges in Colorado. It confirms that freeze–thaw frequency, precipitation (including snow load), and humidity-related indices play significant roles, which could guide other researchers to focus on these variables. The development and evaluation of balanced-accuracy targeted models (and not simple overall accuracy) helps fill a gap in handling imbalanced maintenance data. The project also contributes to emerging literature on machine learning for deterioration prediction. Improved performance by the gradient boosting model provides evidence to recent findings by (Yang et al., 2024) but goes a bit further by quantifying exactly the value of climate features. The study’s results emphasize that including environmental variables leads to an integrated predictive strategy resulting in more effective early warnings for vulnerable components. These contributions support more proactive, data-driven maintenance planning.

5.5 Recommendations for Future Research

Building on this work, several areas are suggested for future research. Testing and improving the models in several geographical areas and bridge populations will help to evaluate transferability and consider different climates. Extending the data collection to include additional environmental factors such as wind stress, solar exposure, or measurement of de-icing chemical concentrations could further improve model fidelity.

Future studies could also look into integrating real-time monitoring of the bridges. For instance, remote sensing technologies like multispectral drones or real-time sensor measurements (moisture probes, strain gauges) could enhance the feature set and allow dynamic prediction updates. Exploring more sophisticated ML methods like deep learning with time-series inputs would be able to identify more subtle degradation patterns that are not now possible with tree ensembles.

REFERENCES

Abdallah, A. M., Atadero, R. A., & Ozbek, M. E. (2021). A Comprehensive Uncertainty-Based Framework for Inspection Planning of Highway Bridges. *Infrastructures*, *6*(2), 27. https://doi.org/10.3390/infrastructures6020027

Agrawal, A. K., Kawaguchi, A., Chen, Z., & City University of New York. City College. Dept. of Civil Engineering. (2008). *Bridge Element Deterioration Rates* (No. C-01–51). https://rosap.ntl.bts.gov/view/dot/16746

Ahlborn, T. M., Vaghefi, K., Harris, D. K., & Brooks, C. N. (2012). Measurement and Communication of Bridge Performance with Remote Sensing Technologies. *Transportation Research Record: Journal of the Transportation Research Board*, *2292*(1), 141–149. https://doi.org/10.3141/2292-17

Ahmed, M., Moselhi, O., & Bhwomick, A. (2016, July 21). *Integration of NDE Measurements and Current Practice in Bridge Deterioration Modeling*. 33th International Symposium on Automation and Robotics in Construction, Auburn, AL, USA. https://doi.org/10.22260/ISARC2016/0042

Al-Rashed, R., Abdelfatah, A., & Yehia, S. (2023). Identifying the Factors Impacting Bridge Deterioration in the Gulf Cooperation Council. *Designs*, *7*(6), Article 6. https://doi.org/10.3390/designs7060126

Anitori, G., Casas, J. R., & Ghosn, M. (2018). Methodology for development of live load models for refined analysis of short and medium-span highway bridges. *Structure and*

Infrastructure Engineering, *14*(4), 477–490. https://doi.org/10.1080/15732479.2017.1406961

Azari, H., O'Shea, D., & Campbell, J. (2022). Application of Unmanned Aerial Systems for Bridge Inspection. *Transportation Research Record: Journal of the Transportation Research Board*, *2676*(1), 401–407. https://doi.org/10.1177/03611981211031896

Balogun, T. B., Tomor, A., Lamond, J., Gouda, H., & Booth, C. A. (2019). Sustainability of bridge maintenance. *Proceedings of the Institution of Civil Engineers - Bridge Engineering*, *172*(1), 54–64. https://doi.org/10.1680/jbren.15.00027

Balogun, T. B., Tomor, A., Lamond, J., Gouda, H., & Booth, C. A. (2020). Life-cycle assessment environmental sustainability in bridge design and maintenance. *Proceedings of the Institution of Civil Engineers - Engineering Sustainability*, *173*(7), 365–375. https://doi.org/10.1680/jensu.19.00042

Bastidas-Arteaga, E. (2018). Reliability of Reinforced Concrete Structures Subjected to Corrosion-Fatigue and Climate Change. *International Journal of Concrete Structures and Materials*, *12*(1), 10. https://doi.org/10.1186/s40069-018-0235-x

Batista, G. E. A. P. A., Prati, R. C., & Monard, M. C. (2004). A study of the behavior of several methods for balancing machine learning training data. *SIGKDD Explor. Newsl.*, *6*(1), 20–29. https://doi.org/10.1145/1007730.1007735

Breiman, L. (2001). [No title found]. *Machine Learning*, *45*(1), 5–32. https://doi.org/10.1023/A:1010933404324

Brodersen, K. H., Ong, C. S., Stephan, K. E., & Buhmann, J. M. (2010). The Balanced Accuracy and Its Posterior Distribution. *2010 20th International Conference on Pattern Recognition*, 3121–3124. https://doi.org/10.1109/ICPR.2010.764

Calabrese, M., Cimmino, M., Fiume, F., Manfrin, M., Romeo, L., Ceccacci, S., Paolanti, M., Toscano, G., Ciandrini, G., Carrotta, A., Mengoni, M., Frontoni, E., & Kapetis, D. (2020). SOPHIA: An Event-Based IoT and Machine Learning Architecture for Predictive Maintenance in Industry 4.0. *Information*, *11*(4), 202. https://doi.org/10.3390/info11040202

Chen, T., & Guestrin, C. (2016a). XGBoost: A Scalable Tree Boosting System. *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 785–794. https://doi.org/10.1145/2939672.2939785

Chen, T., & Guestrin, C. (2016b). XGBoost: A Scalable Tree Boosting System. *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 785–794. https://doi.org/10.1145/2939672.2939785

Chen, Z. W., Xu, Y. L., & Wang, X. M. (2012). SHMS-Based Fatigue Reliability Analysis of Multiloading Suspension Bridges. *Journal of Structural Engineering*, *138*(3), 299–307. https://doi.org/10.1061/(ASCE)ST.1943-541X.0000460

Chicco, D., & Jurman, G. (2020a). The advantages of the Matthews correlation coefficient (MCC) over F1 score and accuracy in binary classification evaluation. *BMC Genomics*, *21*(1), 6. https://doi.org/10.1186/s12864-019-6413-7

Chicco, D., & Jurman, G. (2020b). The advantages of the Matthews correlation coefficient (MCC) over F1 score and accuracy in binary classification evaluation. *BMC Genomics*, *21*(1), 6. https://doi.org/10.1186/s12864-019-6413-7

Chyad, A. M., & Abudayyeh, O. (2021). Impact of Environmental Factors on the Condition Rating of Concrete Bridge Decks Using Statistical-Distribution Methods. *Practice Periodical on Structural Design and Construction*, *26*(3). https://doi.org/10.1061/(ASCE)SC.1943-5576.0000578

Congress, S. S. C., Escamilla, J., Chimauriya, H., & Puppala, A. J. (2024). Eye in the Sky: 360° Inspection of Bridge Infrastructure Using Uncrewed Aerial Vehicles (UAVs). *Transportation Research Record: Journal of the Transportation Research Board*, *2678*(4), 482–504. https://doi.org/10.1177/03611981231185138

D, F., S, W., & B, W. (2020). A Markovian-based methodology for the life-cycle cost analysis of bridge maintenance interventions under changing deterioration rates. *Journal of Civil Engineering Inter Disciplinaries*, *1*(1). https://doi.org/10.47890/JCEID/2020/DFernando/12045781

De-León-Escobedo, D., Delgado-Hernández, D.-J., Martinez-Martinez, L.-H., Rangel-Ramírez, J.-G., & Arteaga-Arcos, J.-C. (2014). Corrosion initiation time updating by epistemic uncertainty as an alternative to schedule the first inspection time of pre-stressed concrete vehicular bridge beams. *Structure and Infrastructure Engineering*, *10*(8), 998–1010. https://doi.org/10.1080/15732479.2013.780084

Dhyani, B. (2021). Predicting Equipment Failure in Manufacturing Plants: An AI-driven Maintenance Strategy. *Mathematical Statistician and Engineering Applications*, *70*(2), 1326–1334. https://doi.org/10.17762/msea.v70i2.2324

Dinegdae, Y. H., & Birgisson, B. (2016). Effect of Heavy Traffic Loading on Predicted Pavement Fatigue Life. In A. Chabot, W. G. Buttlar, E. V. Dave, C. Petit, & G. Tebaldi (Eds.), *8th RILEM International Conference on Mechanisms of Cracking and Debonding in Pavements* (pp. 389–395). Springer Netherlands. https://doi.org/10.1007/978-94-024-0867-6_54

Elleathy, Y., Ghaith, M., Haggag, M., Yosri, A., & El-Dakhakhni, W. (2024). Climate-induced deterioration prediction for bridges: An evolutionary computing-based framework. *Innovative Infrastructure Solutions*, *9*(4), 114. https://doi.org/10.1007/s41062-024-01419-3

Espíndola, R. P., & Ebecken, N. F. F. (2005). *On extending F-measure and G-mean metrics to multi-class problems*. 25–34. https://doi.org/10.2495/DATA050031

Fard, F., & Sadeghi Naieni Fard, F. (2024). Development and Utilization of Bridge Data of the United States for Predicting Deck Condition Rating Using Random Forest, XGBoost, and Artificial Neural Network. *Remote Sensing*, *16*(2), 367. https://doi.org/10.3390/rs16020367

Feng, J., Gao, K., Wu, G., Xu, Y., & Jiang, H. (2023). A deep learning-based interferometric synthetic aperture radar framework for abnormal displacement deformation prediction of bridges. *Advances in Structural Engineering*, *26*(16), 3005–3020. https://doi.org/10.1177/13694332231205061

Fereshtehnejad, E., Gazzola, G., Parekh, P., Nakrani, C., & Parvardeh, H. (2022). Detecting Anomalies in National Bridge Inventory Databases Using Machine Learning Methods. *Transportation Research Record: Journal of the Transportation Research Board*, *2676*(6), 453–467. https://doi.org/10.1177/03611981221075028

Fick, D., Bell, M., & Western Transportation Institute. (2024). *Significant Factors of Bridge Deterioration: Final Report* (Version Final Report, February 2023 – December 2024) [Application/pdf]. Montana Department of Transportation. https://doi.org/10.21949/1529563

Fom, P. B., Imam, B., & Chryssanthopoulos, M. K. (2015a). Long-term deterioration effects on the buckling strength of metallic bridge girders. *Proceedings of the Second International Conference on Performance-Based and Life-Cycle Structural Engineering (PLSE 2015)*, 1515–1525. https://doi.org/10.14264/uql.2016.758

Fom, P. B., Imam, B., & Chryssanthopoulos, M. K. (2015b). Long-term deterioration effects on the buckling strength of metallic bridge girders. *Proceedings of the Second International Conference on Performance-Based and Life-Cycle Structural Engineering (PLSE 2015)*, 1515–1525. https://doi.org/10.14264/uql.2016.758

Fujiu, M., Minami, T., & Takayama, J. (2022). Environmental Influences on Bridge Deterioration Based on Periodic Inspection Data from Ishikawa Prefecture, Japan. *Infrastructures*, *7*(10), Article 10. https://doi.org/10.3390/infrastructures7100130

Galar, M., Fernandez, A., Barrenechea, E., Bustince, H., & Herrera, F. (2012). A Review on Ensembles for the Class Imbalance Problem: Bagging-, Boosting-, and Hybrid-Based Approaches. *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, *42*(4), 463–484. https://doi.org/10.1109/TSMCC.2011.2161285

Gao, Y. H. (2020). Influence of Chloride Ion Corrosion on the Performance of Reinforced Concrete Beam Bridge in Offshore Environ ment. *Archives of Civil Engineering*, 253–265. https://doi.org/10.24425/ace.2020.131808

Garg, R., Garg, R., & Singla, S. (2021). Experimental Investigation of Electrochemical Corrosion and Chloride Penetration of Concrete Incorporating Colloidal Nanosilica and Silica fume. *Journal of Electrochemical Science and Technology*, *12*(4), 440–452. https://doi.org/10.33961/jecst.2020.01788

Ghafoori, M., Abdallah, M., & Ozbek, M. E. (2024a). Machine Learning–Based Bridge Maintenance Optimization Model for Maximizing Performance within Available Annual Budgets. *Journal of Bridge Engineering*, *29*(4), 04024011. https://doi.org/10.1061/JBENF2.BEENG-6436

Ghafoori, M., Abdallah, M., & Ozbek, M. E. (2024b). Machine Learning–Based Bridge Maintenance Optimization Model for Maximizing Performance within Available Annual Budgets. *Journal of Bridge Engineering*, *29*(4), 04024011. https://doi.org/10.1061/JBENF2.BEENG-6436

Ghosh, J., Rokneddin, K., Padgett, J. E., & Dueñas-Osorio, L. (2014). Seismic Reliability Assessment of Aging Highway Bridge Networks with Field Instrumentation Data and

Correlated Failures, I: *Methodology*. *Earthquake Spectra*, *30*(2), 795–817. https://doi.org/10.1193/040512EQS155M

Gopalakrishnan, M., Subramaniyan, M., & Skoogh, A. (2022). Data-driven machine criticality assessment – maintenance decision support for increased productivity. *Production Planning & Control*, *33*(1), 1–19. https://doi.org/10.1080/09537287.2020.1817601

Graybeal, B. A., Phares, B. M., Rolander, D. D., Moore, M., & Washer, G. (2002). Visual Inspection of Highway Bridges. *Journal of Nondestructive Evaluation*, *21*(3), 67–83. https://doi.org/10.1023/A:1022508121821

Guan, S., Bridge, J. A., Li, C., & DeMello, N. J. (2018). Smart Radar Sensor Network for Bridge Displacement Monitoring. *Journal of Bridge Engineering*, *24*(1), 04018102. https://doi.org/10.1061/(ASCE)BE.1943-5592.0001322

Hancock, J. T., Khoshgoftaar, T. M., & Johnson, J. M. (2023). Evaluating classifier performance with highly imbalanced Big Data. *Journal of Big Data*, *10*(1), 42. https://doi.org/10.1186/s40537-023-00724-5

Hannan, A., & Anmala, J. (2021). Classification and Prediction of Fecal Coliform in Stream Waters Using Decision Trees (DTs) for Upper Green River Watershed, Kentucky, USA. *Water*, *13*(19), 2790. https://doi.org/10.3390/w13192790

Haynes, M. A., Coleri, E., & Sreedhar, S. (2019). Impermeable Asphalt Concrete Layer to Protect and Seal Concrete Bridge Decks. *Transportation Research Record: Journal of the Transportation Research Board*, *2673*(6), 355–367. https://doi.org/10.1177/0361198119841041

Heo, W.-H. (2020). Performance-Based Reliability Estimates for Highway Bridges Considering Previous Inspection Data. *Applied Sciences*, *10*(5), 1873. https://doi.org/10.3390/app10051873

Ibrahim, A., Abdelkhalek, S., Zayed, T., Qureshi, A. H., & Mohammed Abdelkader, E. (2024a). A Comprehensive Review of the Key Deterioration Factors of Concrete Bridge Decks. *Buildings*, *14*(11), 3425. https://doi.org/10.3390/buildings14113425

Ibrahim, A., Abdelkhalek, S., Zayed, T., Qureshi, A. H., & Mohammed Abdelkader, E. (2024b). A Comprehensive Review of the Key Deterioration Factors of Concrete Bridge Decks. *Buildings*, *14*(11), 3425. https://doi.org/10.3390/buildings14113425

Jaafaru, H., & Agbelie, B. (2022). Bridge maintenance planning framework using machine learning, multi-attribute utility theory and evolutionary optimization models. *Automation in Construction*, *141*, 104460-. https://doi.org/10.1016/j.autcon.2022.104460

Jensen, J. S., Frangopol, D. M., & Schmidt, J. W. (Eds.). (2024). *Bridge Maintenance, Safety, Management, Digitalization and Sustainability*. Taylor & Francis.

Jiang, L., Tang, Q., Jiang, Y., Cao, H., & Xu, Z. (2023a). Bridge Condition Deterioration Prediction Using the Whale Optimization Algorithm and Extreme Learning Machine. *Buildings*, *13*(11), 2730. https://doi.org/10.3390/buildings13112730

Jiang, L., Tang, Q., Jiang, Y., Cao, H., & Xu, Z. (2023b). Bridge Condition Deterioration Prediction Using the Whale Optimization Algorithm and Extreme Learning Machine. *Buildings*, *13*(11), 2730. https://doi.org/10.3390/buildings13112730

Kale, A., Kassa, Y., Ricks, B., & Gandhi, R. (2023a). A Comparative Assessment of Bridge Deck Wearing Surfaces: Performance, Deterioration, and Maintenance. *Applied Sciences*, *13*(19), 10883. https://doi.org/10.3390/app131910883

Kale, A., Kassa, Y., Ricks, B., & Gandhi, R. (2023b). A Comparative Assessment of Bridge Deck Wearing Surfaces: Performance, Deterioration, and Maintenance. *Applied Sciences*, *13*(19), 10883. https://doi.org/10.3390/app131910883

Kallias, A. N., & Imam, B. M. (2013). *Effect of Climate Change on the Deterioration of Steel Bridges*. 548–549. https://doi.org/10.2749/222137813806521469

Kim, J., & Song, J. (2021). Time-Dependent Reliability Assessment and Updating of Post-tensioned Concrete Box Girder Bridges Considering Traffic Environment and Corrosion. *ASCE-ASME Journal of Risk and Uncertainty in Engineering Systems, Part A: Civil Engineering*, *7*(4), 04021062. https://doi.org/10.1061/AJRUA6.0001188

Kim, K. H., Nam, M. S., Hwang, H. H., & Ann, K. Y. (2020a). Prediction of Remaining Life for Bridge Decks Considering Deterioration Factors and Propose of Prioritization Process for Bridge Deck Maintenance. *Sustainability*, *12*(24), 10625. https://doi.org/10.3390/su122410625

Kim, K. H., Nam, M. S., Hwang, H. H., & Ann, K. Y. (2020b). Prediction of Remaining Life for Bridge Decks Considering Deterioration Factors and Propose of Prioritization Process for Bridge Deck Maintenance. *Sustainability*, *12*(24), 10625. https://doi.org/10.3390/su122410625

Kobayashi, K., Kaito, K., & Lethanh, N. (2012). A statistical deterioration forecasting method using hidden Markov model for infrastructure management. *Transportation Research Part B: Methodological*, *46*(4), 544–561. https://doi.org/10.1016/j.trb.2011.11.008

Kong, X., & Li, J. (2018). Vision-Based Fatigue Crack Detection of Steel Structures Using Video Feature Tracking. *Computer-Aided Civil and Infrastructure Engineering*, *33*(9), 783–799. https://doi.org/10.1111/mice.12353

Kumar, P., & Kota, S. R. (2024). Machine learning models in structural engineering research and a secured framework for structural health monitoring. *Multimedia Tools and Applications*, *83*(3), 7721–7759. https://doi.org/10.1007/s11042-023-15853-5

Lee, S. H., An, L.-S., & Kim, H.-K. (2025). Risk-based bridge life cycle cost and environmental impact assessment considering climate change effects. *Scientific Reports*, *15*(1), 725. https://doi.org/10.1038/s41598-024-82568-4

Li, L., Sun, L., & Ning, G. (2014). Deterioration Prediction of Urban Bridges on Network Level Using Markov-Chain Model. *Mathematical Problems in Engineering*, *2014*, 1–10. https://doi.org/10.1155/2014/728107

Li, Q., & Song, Z. (2022). Ensemble-Learning-Based Prediction of Steel Bridge Deck Defect Condition. *Applied Sciences*, *12*(11), 5442. https://doi.org/10.3390/app12115442

Li, R., Miao, C., & Wei, T. (2022a). Effect of environmental factors on electrochemical corrosion of galvanized steel wires for bridge cables. *Anti-Corrosion Methods and Materials*, *69*(1), 111–118. https://doi.org/10.1108/ACMM-07-2021-2522

Li, R., Miao, C., & Wei, T. (2022b). Effect of environmental factors on electrochemical corrosion of galvanized steel wires for bridge cables. *Anti-Corrosion Methods and Materials*, *69*(1), 111–118. https://doi.org/10.1108/ACMM-07-2021-2522

Li, S., Wang, W., Lu, B., Du, X., Dong, M., Zhang, T., & Bai, Z. (2023). Long-term structural health monitoring for bridge based on back propagation neural network and long and short-term memory. *Structural Health Monitoring*, *22*(4), 2325–2345. https://doi.org/10.1177/14759217221122337

Li, S., Zhang, Z., Lin, D., Zhang, T., & Han, L. (2023). Development of a BIM-based bridge maintenance system (BMS) for managing defect data. *Scientific Reports*, *13*(1), 846. https://doi.org/10.1038/s41598-023-27924-6

Li, W., Ho, S. C. M., & Song, G. (2016). Corrosion detection of steel reinforced concrete using combined carbon fiber and fiber Bragg grating active thermal probe. *Smart Materials and Structures*, *25*(4), 045017. https://doi.org/10.1088/0964-1726/25/4/045017

Lin, P., Yuan, X., & Tovilla, E. (2019). Integrative modeling of performance deterioration and maintenance effectiveness for infrastructure assets with missing condition data. *Computer-Aided Civil and Infrastructure Engineering*, *34*(8), 677–695. https://doi.org/10.1111/mice.12452

Liu, C., Tang, D., Zhu, H., & Nie, Q. (2021). A Novel Predictive Maintenance Method Based on Deep Adversarial Learning in the Intelligent Manufacturing System. *IEEE Access*, *9*, 49557–49575. https://doi.org/10.1109/ACCESS.2021.3069256

Liu, K., & El-Gohary, N. (2020). Fusing Data Extracted from Bridge Inspection Reports for Enhanced Data-Driven Bridge Deterioration Prediction: A Hybrid Data Fusion Method. *Journal of Computing in Civil Engineering*, *34*(6). https://doi.org/10.1061/(ASCE)CP.1943-5487.0000921

Liu, K., & El-Gohary, N. (2022). Deep Learning–Based Analytics of Multisource Heterogeneous Bridge Data for Enhanced Data-Driven Bridge Deterioration Prediction. *Journal of Computing in Civil Engineering*, *36*(5). https://doi.org/10.1061/(ASCE)CP.1943-5487.0001018

Liu, Y., Su, D., Duan, Q., & Cao, Z. (2022a). Recommendations for Refined Preventive Maintenance Management of Concrete Bridges in China Based on Environmental Risk Zoning. *Transportation Research Record: Journal of the Transportation Research Board*, *2676*(5), 460–479. https://doi.org/10.1177/03611981211067978

Liu, Y., Su, D., Duan, Q., & Cao, Z. (2022b). Recommendations for Refined Preventive Maintenance Management of Concrete Bridges in China Based on Environmental Risk Zoning. *Transportation Research Record: Journal of the Transportation Research Board*, *2676*(5), 460–479. https://doi.org/10.1177/03611981211067978

Louppe, G., Wehenkel, L., Sutera, A., & Geurts, P. (n.d.). *Understanding variable importances in forests of randomized trees*.

Lu, M., & Ilgin Guler, S. (2022). Comparison of Random Survival Forest with Accelerated Failure Time-Weibull Model for Bridge Deck Deterioration. *Transportation Research Record:*

Journal of the Transportation Research Board, *2676*(7), 296–311. https://doi.org/10.1177/03611981221078281

Marzuki, M., Prayogi, S., & Abdillah, M. (2023). Data-Driven Based Model For Predictive Maintenance Applications In Industrial System. *Proceedings of the International Conference on Sustainable Engineering, Infrastructure and Development, ICO-SEID 2022, 23-24 November 2022, Jakarta, Indonesia*. Proceedings of the International Conference on Sustainable Engineering, Infrastructure and Development, ICO-SEID 2022, 23-24 November 2022, Jakarta, Indonesia, Jakarta, Indonesia. https://doi.org/10.4108/eai.23-11-2022.2341596

Mete, F., Corr, D. J., Wilbur, M. P., & Chen, Y. (2022). Bridge Response and Heavy Truck Classification Framework Based on a Two-Step Machine Learning Algorithm. *Transportation Research Record: Journal of the Transportation Research Board*, *2676*(3), 454–467. https://doi.org/10.1177/03611981211052027

Mia, Md. M., & Kameshwar, S. (2023a). Machine learning approach for predicting bridge components' condition ratings. *Frontiers in Built Environment*, *9*, 1254269. https://doi.org/10.3389/fbuil.2023.1254269

Mia, Md. M., & Kameshwar, S. (2023b). Machine learning approach for predicting bridge components' condition ratings. *Frontiers in Built Environment*, *9*, 1254269. https://doi.org/10.3389/fbuil.2023.1254269

Mia, Md. M., & Kameshwar, S. (2023c). Machine learning approach for predicting bridge components' condition ratings. *Frontiers in Built Environment*, *9*, 1254269. https://doi.org/10.3389/fbuil.2023.1254269

Miao, P., Yokota, H., & Zhang, Y. (2023a). Deterioration prediction of existing concrete bridges using a LSTM recurrent neural network. *Structure and Infrastructure Engineering*, *19*(4), 475–489. https://doi.org/10.1080/15732479.2021.1951778

Miao, P., Yokota, H., & Zhang, Y. (2023b). Deterioration prediction of existing concrete bridges using a LSTM recurrent neural network. *Structure and Infrastructure Engineering*, *19*(4), 475–489. https://doi.org/10.1080/15732479.2021.1951778

Mitterpach, J., Fojtík, R., Machovčáková, E., & Kubíncová, L. (2022). Life Cycle Assessment of a Road Transverse Prestressed Wooden–Concrete Bridge. *Forests*, *14*(1), 16. https://doi.org/10.3390/f14010016

Mohamed, A. G., Khaled, A., & Abotaleb, I. S. (2023). A Bridge Information Modeling (BrIM) Framework for Inspection and Maintenance Intervention in Reinforced Concrete Bridges. *Buildings*, *13*(11), 2798. https://doi.org/10.3390/buildings13112798

Nagaraju, A. S. V., & Kumar, R. (2024). Effect of climate change on structural safety of RC bridges in coastal region. *Structures*, *63*, 106273. https://doi.org/10.1016/j.istruc.2024.106273

Nettis, A., Nettis, A., Ruggieri, S., & Uva, G. (2023). *TYPOLOGICAL FRAGILITY ASSESSMENT OF PRESTRESSED CONCRETE GIRDER BRIDGES SUBJECTED TO TRAFFIC LOADS AFFECTED BY CORROSION*. 5242–5257. https://doi.org/10.7712/120123.10797.21003

Nguyen, P. T., Bastidas-Arteaga, E., Amiri, O., & El Soueidy, C.-P. (2017). An Efficient Chloride Ingress Model for Long-Term Lifetime Assessment of Reinforced Concrete Structures Under Realistic Climate and Exposure Conditions. *International Journal of Concrete Structures and Materials*, *11*(2), 199–213. https://doi.org/10.1007/s40069-017-0185-8

Niculescu-Mizil, A., & Caruana, R. (2005). Predicting good probabilities with supervised learning. *Proceedings of the 22nd International Conference on Machine Learning*, 625–632. https://doi.org/10.1145/1102351.1102430

Nie, L., Wang, W., & Deng, L. (2022). Fatigue reliability analysis and traffic load control of steel bridges based on artificial neural network. In M. Mikusova (Ed.), *International Conference on Smart Transportation and City Engineering (STCE 2022)* (p. 114). SPIE. https://doi.org/10.1117/12.2658303

Obata, M., Guotai, L., Watanabe, Y., & Goto, Y. (2014). Numerical simulation of adhesion of sea - salt particles on bridge girders. *Structure and Infrastructure Engineering*, *10*(3), 398–408. https://doi.org/10.1080/15732479.2012.757328

O'Connor, A. J., Sheils, E., Breysse, D., & Schoefs, F. (2013). Markovian Bridge Maintenance Planning Incorporating Corrosion Initiation and Nonlinear Deterioration. *Journal of Bridge Engineering*, *18*(3), 189–199. https://doi.org/10.1061/(ASCE)BE.1943-5592.0000342

Omar, A., & Moselhi, O. (2022). Hybrid feature selection framework for predicting bridge deck conditions. *Journal of Information Technology in Construction*, *27*, 1028–1041. https://doi.org/10.36680/j.itcon.2022.050

Pak, D., & Feldmann, M. (2016). Design Temperature Load Spectrum for the Fatigue Verification of Composite Bridges with Integral Abutments. *Composite Construction in Steel and Concrete VII*, 36–50. https://doi.org/10.1061/9780784479735.004

Palak C Patel & Bankim R Joshi. (2023). Modification and implementation of bridge maintenance process in south Gujarat. *Global Journal of Engineering and Technology Advances*, *16*(1), 073–082. https://doi.org/10.30574/gjeta.2023.16.1.0138

Palu, S., & Mahmoud, H. (2019). Impact of climate change on the integrity of the superstructure of deteriorated U.S. bridges. *PLOS ONE*, *14*(10), e0223307. https://doi.org/10.1371/journal.pone.0223307

Raja, H. A., Kudelina, K., Asad, B., Vaimann, T., Kallaste, A., Rassõlkin, A., & Khang, H. V. (2022). Signal Spectrum-Based Machine Learning Approach for Fault Prediction and Maintenance of Electrical Machines. *Energies*, *15*(24), 9507. https://doi.org/10.3390/en15249507

Random forest. (2025). In *Wikipedia*. https://en.wikipedia.org/w/index.php?title=Random_forest&oldid=1278609179

Rashidi Nasab, A., & Elzarka, H. (2023a). Optimizing Machine Learning Algorithms for Improving Prediction of Bridge Deck Deterioration: A Case Study of Ohio Bridges. *Buildings*, *13*(6), 1517. https://doi.org/10.3390/buildings13061517

Rashidi Nasab, A., & Elzarka, H. (2023b). Optimizing Machine Learning Algorithms for Improving Prediction of Bridge Deck Deterioration: A Case Study of Ohio Bridges. *Buildings*, *13*(6), 1517. https://doi.org/10.3390/buildings13061517

Rashidi Nasab, A., & Elzarka, H. (2023c). Optimizing Machine Learning Algorithms for Improving Prediction of Bridge Deck Deterioration: A Case Study of Ohio Bridges. *Buildings*, *13*(6), 1517. https://doi.org/10.3390/buildings13061517

Rashidi Nasab, A., & Elzarka, H. (2023d). Optimizing Machine Learning Algorithms for Improving Prediction of Bridge Deck Deterioration: A Case Study of Ohio Bridges. *Buildings*, *13*(6), 1517. https://doi.org/10.3390/buildings13061517

Rizzo, P., & Enshaeian, A. (2021). Challenges in Bridge Health Monitoring: A Review. *Sensors*, *21*(13), 4336. https://doi.org/10.3390/s21134336

Sacconi, S., Ierimonti, L., Venanzi, I., & Ubertini, F. (2021). *TOWARDS THE DEVELOPMENT OF A COMPREHENSIVE FRAMEWORK FOR LIFE CYCLE COST ANALYSIS OF BRIDGES SUBJECTED TO MULTIPLE HAZARDS*. 1961–1972. https://doi.org/10.7712/120121.8615.18843

Saeed Hasan, M., Setunge, S., W. Law, D., & Chin Koay, Y. (2015). Forecasting Deterioration of Bridge Components from Visual Inspection Data. *International Journal of Engineering and Technology*, *7*(1), 40–44. https://doi.org/10.7763/IJET.2015.V7.763

Saito, T., & Rehmsmeier, M. (2015). The Precision-Recall Plot Is More Informative than the ROC Plot When Evaluating Binary Classifiers on Imbalanced Datasets. *PLOS ONE*, *10*(3), e0118432. https://doi.org/10.1371/journal.pone.0118432

Salehi, H., & Burgueño, R. (2018). Emerging artificial intelligence methods in structural engineering. *Engineering Structures*, *171*, 170–189. https://doi.org/10.1016/j.engstruct.2018.05.084

Santos, K., Dias, J. P., & Amado, C. (2022). A literature review of machine learning algorithms for crash injury severity prediction. *Journal of Safety Research*, *80*, 254–269. https://doi.org/10.1016/j.jsr.2021.12.007

Scikit-learn. (2024). *Machine learning in Python*. https://scikit-learn.org/stable/

Sengupta, A., Ilgin Guler, S., & Shokouhi, P. (2021). Interpreting Impact Echo Data to Predict Condition Rating of Concrete Bridge Decks: A Machine-Learning Approach. *Journal of Bridge Engineering*, *26*(8), 04021044. https://doi.org/10.1061/(ASCE)BE.1943-5592.0001744

Sgambi, L., Gkoumas, K., & Bontempi, F. (2012). Genetic Algorithms for the Dependability Assurance in the Design of a Long-Span Suspension Bridge. *Computer-Aided Civil and Infrastructure Engineering*, *27*(9), 655–675. https://doi.org/10.1111/j.1467-8667.2012.00780.x

Shen, Z., Liu, Y., Liu, J., Liu, Z., Han, S., & Lan, S. (2023). A Decision-Making Method for Bridge Network Maintenance Based on Disease Transmission and NSGA-II. *Sustainability*, *15*(6), 5007. https://doi.org/10.3390/su15065007

Soliman, M., & Frangopol, D. M. (2015). Life-Cycle Cost Evaluation of Conventional and Corrosion-Resistant Steel for Bridges. *Journal of Bridge Engineering*, *20*(1), 06014005. https://doi.org/10.1061/(ASCE)BE.1943-5592.0000647

Srikanth, I., & Arockiasamy, M. (2020). Deterioration models for prediction of remaining useful life of timber and concrete bridges: A review. *Journal of Traffic and Transportation Engineering (English Edition)*, *7*(2), 152–173. https://doi.org/10.1016/j.jtte.2019.09.005

Stewart, M. G., Wang, X., & Nguyen, M. N. (2012). Climate change adaptation for corrosion control of concrete infrastructure. *Structural Safety*, *35*, 29–39. https://doi.org/10.1016/j.strusafe.2011.10.002

Tian, H., Chen, Y., & Li, F. (2015). A Novel Approach to Evaluate the Time-Variant System Reliability of Deteriorating Concrete Bridges. *Mathematical Problems in Engineering*, *2015*, 1–12. https://doi.org/10.1155/2015/129787

Veloso, B., Gama, J., Ribeiro, R. P., & Pereira, P. M. (2022). *A Benchmark dataset for predictive maintenance* (Version 3). arXiv. https://doi.org/10.48550/ARXIV.2207.05466

Viljanen, M., Meijerink, L., Zwakhals, L., & Van De Kassteele, J. (2022). A machine learning approach to small area estimation: Predicting the health, housing and well-being of the population of Netherlands. *International Journal of Health Geographics*, *21*(1), 4. https://doi.org/10.1186/s12942-022-00304-5

Wang, T., Liu, Y., Li, Q., Du, P., Zheng, X., & Gao, Q. (2023). State-of-the-Art Review of the Resilience of Urban Bridge Networks. *SUSTAINABILITY*, *15*(2), 989. https://doi.org/10.3390/su15020989

Wang, Y.-C., & Chen, C.-H. (2012). *Dynamic characteristics of a cable-stayed bridge measured from traffic-induced vibrations*. 2278–2281. https://doi.org/10.1063/1.4756648

Weissmann, J., Weissmann, A. J., & Montoya, A. (2023). Deterioration Models for Texas Bridges and Culverts. *Transportation Research Record: Journal of the Transportation Research Board*, *2677*(7), 307–316. https://doi.org/10.1177/03611981231152460

Wu, C., Wu, P., Wang, J., Jiang, R., Chen, M., & Wang, X. (2021). Critical review of data-driven decision-making in bridge operation and maintenance. *Structure and Infrastructure Engineering*, *18*(1), 47–70. https://doi.org/10.1080/15732479.2020.1833946

Yang, C., Wang, X., & Nassif, H. (2024). Impact of Environmental Conditions on Predicting Condition Rating of Concrete Bridge Decks. *Transportation Research Record: Journal of the Transportation Research Board*, *2678*(12), 999–1012. https://doi.org/10.1177/03611981241248647

Ye, M., Chen, X., Ren, M., Cao, B., & Huang, S. (2016). Time variant natural frequencies of a roadway bridge under stochastic vehicle flow. *Journal of Vibroengineering*, *18*(8), 5364–5377. https://doi.org/10.21595/jve.2016.17122

Yianni, P. C., Neves, L. C., Rama, D., Andrews, J. D., & Dean, R. (2016). Incorporating local environmental factors into railway bridge asset management. *Engineering Structures*, *128*, 362–373. https://doi.org/10.1016/j.engstruct.2016.09.038

Yuan, Y., Zhu, X., & Li, J. (2024). Moving-Principal-Component-Analysis-Based Structural Damage Detection for Highway Bridges in Operational Environments. *Sensors*, *24*(2), 383. https://doi.org/10.3390/s24020383

Yuan, Z., Wang, H., Li, R., Mao, J., & Gao, H. (2023). Equivalent conversion investigation of environmental corrosion of suspenders in long-span suspension bridge. *Anti-Corrosion Methods and Materials*, *70*(6), 533–541. https://doi.org/10.1108/ACMM-08-2023-2883

Zaki, A., Chai, H. K., Behnia, A., Aggelis, D. G., Tan, J. Y., & Ibrahim, Z. (2017). Monitoring fracture of steel corroded reinforced concrete members under flexure by acoustic emission technique. *Construction and Building Materials*, *136*, 609–618. https://doi.org/10.1016/j.conbuildmat.2016.11.079

Zandonini, R., Debiasi, E., & Zonta, D. (2013). *Management of overweight vehicle traffic on road bridges*. 60–61. https://doi.org/10.2749/222137813806474255

Zhang, X., Wogen, B. E., Liu, X., Iturburu, L., Salmeron, M., Dyke, S. J., Poston, R., & Ramirez, J. A. (2023). Machine-Aided Bridge Deck Crack Condition State Assessment Using Artificial Intelligence. *Sensors*, *23*(9), 4192. https://doi.org/10.3390/s23094192

Zhao, Y., Guo, X., Su, B., Sun, Y., & Zhu, Y. (2023). Multi-Lane Traffic Load Clustering Model for Long-Span Bridge Based on Parameter Correlation. *Mathematics*, *11*(2), 274. https://doi.org/10.3390/math11020274

Zhao, Y., Zhang, W., & Liu, X. (2024). Grid search with a weighted error function: Hyper-parameter optimization for financial time series forecasting. *Applied Soft Computing*, *154*, 111362. https://doi.org/10.1016/j.asoc.2024.111362

Zhou, Q.-N., Yuan, Y., Yang, D., & Zhang, J. (2022). An Advanced Multi-Agent Reinforcement Learning Framework of Bridge Maintenance Policy Formulation. *Sustainability*, *14*(16), 10050. https://doi.org/10.3390/su141610050