Multiscale Machine Learning and Numerical Investigation of Ageing in Infrastructures

By
Keyvan Aghabalaei Baghaei

A thesis submitted in fulfilment of the requirements for the degree of
Doctor of Philosophy

School of Civil Engineering
Faculty of Engineering
The University of Sydney

September 2023

Statement of originality

This is to certify that to the best of my knowledge, the content of this thesis is my own work. This thesis has not been submitted for any degree or other purposes. I certify that the intellectual content of this thesis is the product of my own work and that all the assistance received in preparing this thesis and sources have been acknowledged.

Keyvan Aghabalaei Baghaei

Authorship attribution statement and list of publications

Chapters 3 and 4 in this thesis contains materials that are published in:

- Aghabalaei Baghaei, K. and Hadigheh, S.A. Durability assessment of FRP-to-concrete bonded connections under moisture condition using data-driven machine learning-based approaches, Composite Structures, 2021, 114576,
  DOI: https://doi.org/10.1016/j.compstruct.2021.114576

- Aghabalaei Baghaei, K. and Hadigheh, S.A. A machine learning approach to modelling the bond strength of adhesively bonded joints under water immersion condition, 10th International Conference on FRP Composites in Civil Engineering (CICE 2021), 2022, Istanbul, Turkey.

  This includes Sections 3.4, 3.5 in Chapter 3 and Sections 4.1, 4.2.1, 4.3, 4.4.1, and 4.5 in Chapter 4.

  K. Aghabalaei Baghaei: Data curation, Formal analysis, Investigation, Software, Validation, Visualization, Writing - original draft. **S.A. Hadigheh**: Conceptualization, Data curation, Funding acquisition, Investigation, Methodology, Project administration, Resources, Supervision, Validation, Writing - review & editing.

Chapter 5 in this thesis contains materials that are published in:

- Aghabalaei Baghaei, K. and Hadigheh, S.A. Artificial neural network-based characterisation of the bond between FRP bar and concrete under environmental conditions, The 20th European Conference on Composite Materials (ECCM 20), 2022, Lausanne, Switzerland.

  This includes Section 5.9.

  K. Aghabalaei Baghaei: Data curation, Formal analysis, Investigation, Software, Validation, Visualization, Writing - original draft. **S.A. Hadigheh**: Conceptualization, Data curation, Funding acquisition, Investigation, Methodology, Project administration, Resources, Supervision, Validation, Writing - review & editing.

In addition to the statement above, in cases where I am not the corresponding author of the published item, permission to include the published material has been granted by the corresponding author.

Keyvan Aghabalaei Baghaei, 30 September 2023

As supervisor for the candidature upon which this thesis is based, I can confirm that the authorship attribution statements above are correct.

Ali Hadigheh

Acknowledgements

I would like to express my sincere gratitude to my supervisor, Dr Ali Hadigheh, for his support, guidance, and encouragement throughout my doctoral studies. His expertise and advice were invaluable to me, and I am grateful for the opportunity to have worked with him.

I would also like to thank the members of the examining committee for their time and for reviewing my dissertation.

I also extend my thanks to Dr Slaven Marusic. His experience and knowledge were a great source of help for me to gain invaluable industrial experience.

I deeply appreciate the University of Sydney for providing financial support through the Engineering and IT Research Scholarship (EITRS) and the Postgraduate Research Support Scheme (PRSS) during my candidature, that helped me advance my research. I also appreciate the Australian Postgraduate Research Intern (APR.Intern) Program Scholarship that supported my research during my internship at Aurecon.

I would also like to thank my dear colleagues and friends, Dr Sina Kazemzadeh Azad, Dr Bahman Lahoorpoor, Ziyuan Wang, Haowei Huang, and Bev Stoch. I have always enjoyed talking with them and sharing ideas about research and life.

Last but not least, my deepest gratitude goes to my family. I am always grateful to have the endless support of my parents, Shahin and Abdollah. I appreciate their support, love, and encouragement during this long and challenging journey. I also thank my dear brother, Peyman, who has been a great source of mentorship throughout the important stages of my life. Talking to him is always a great pleasure, which gives me the confidence to pursue my goals.

Dedicated to my Mom & Dad

Abstract

Infrastructure is a critical components of a country's economic growth. It is important to maintain their functionality and ensure their reliable service to society. Ageing is an inevitable phenomenon that impacts serviceability over time. Interaction with extreme service environments can adversely affect the long-term performance of infrastructure and accelerate ageing. Rehabilitation strategies are helpful in preserving and extending the service life of infrastructure. However, their efficiency is compromised since the impact of ageing at smaller scales is often overlooked in the design of rehabilitation techniques. As a result, higher safety factors are introduced, which lead to costly and less efficient solutions.

Machine learning has been introduced as a powerful tool to solve engineering problems with uncertainties since it can handle complex and noisy data. This research focuses on using machine learning to improve the efficiency of analysing the ageing impact on infrastructure at multiple scales. It is done by integrating a series of empirical datasets and numerical modelling with machine learning models. The results demonstrate that the developed frameworks make an effective contribution to modelling the long-term performance of infrastructure.

First, a data-driven campaign is developed to analyse the condition of an ageing infrastructure. Historical data on the maintenance of a railway system is explored to identify the factors contributing to the failure of the infrastructure assets. A framework is developed based on data analytics and machine learning to predict the state of various assets and identify any faults. The influence of various asset characteristics on the infrastructure defects is modelled. The findings indicate the potential of the proposed framework for planning efficient predictive maintenance strategies, which can optimise costs and improve the serviceability of the railway network.

The ageing of fibre-reinforced polymer (FRP)-strengthened concrete structures due to interaction with environmental conditions is investigated using machine learning. A series of comprehensive databases are established on the durability study of connections in the FRP-reinforced members. Different machine learning methods are explored and used to characterise the long-term performance of the bond by modelling its strength and failure mode. The outcome of the study shows that the machine learning-based frameworks can accurately analyse the ageing of composite structures under a variety of environmental conditions.

The environmental ageing of glass fibre/epoxy composite materials is investigated by a micromechanics-based machine learning model. A mathematical framework is developed to automatically generate unidirectional (UD) composite representative volume elements (RVEs) with periodic microstructures while accounting for the interphase areas around the fibres. The microstructures are analysed by the finite element (FE) method to obtain their effective transverse properties. In order to provide a computationally fast and accurate framework, a machine learning model is developed based on FE simulations of artificially created 1,200 microstructures. The generated data is processed using the two-point correlation function (TPC) and principal component analysis (PCA) techniques to statistically characterise and extract features of the microstructure images. The processed data is used to develop an artificial neural network (ANN) model to map the microstructure geometric features to the effective mechanical properties of composites under humid conditions. The results show that machine learning can accurately analyse the impact of ageing at the material scale.

Finally, a multiscale FE and machine learning framework is developed to further expand the understanding of the composite material ageing process. A moisture diffusion analysis is performed by Fick’s law to simulate water uptake of resin and UD composite samples under water immersion condition. The knowledge is used to obtain the moisture concentration values in the composite by solving a macro-model diffusion problem. The analysis results are then downscaled to solve a micro-stress problem and obtain micromodel stress fields. A numerical homogenisation is finally used to upscale the results and obtain the composite transverse behaviour during immersion in water. Given the computational effort required by the numerical approach, a machine learning model is developed to replace the micromodel FE simulations. An ANN model is developed based on the multiscale simulation results of the composite under moisture immersion. The efficiency of the developed model is tested by comparing its predictions to the numerical simulation of composite samples under 1200 hours of water uptake. A feature importance analysis is also performed to investigate the contribution of geometrical, mechanical, and environmental factors to the composite material degradation.

In this thesis, different machine learning frameworks are developed based on experimental studies and numerical modelling. These frameworks demonstrate how machine learning improves the analysis of ageing across multiple scales of infrastructure. The resulting

understanding can help develop more efficient strategies for the rehabilitation of ageing infrastructure.

Table of Contents

4 Assessment of Ageing Impact on FRP-Strengthened Concrete Structures Under Environmental Conditions Using Machine Learning Methods 87

5 Modelling Transverse Performance of FRP Materials Under Humid Environment Using Micromechanical Analysis and Machine Learning Method 123

List of Figures

List of Tables

CHAPTER 1

Introduction

1.1 Background

Infrastructures are the key components of every society. They are crucial for economic and social growth. Preserving the functionality of the infrastructures during their lifetimes is important to ensure their safe and reliable service to society. On the other hand, ageing is an inevitable phenomenon that affects the functionality of the infrastructures over time. Interactions with the service environment, such as extreme weather conditions and increased usage, can accelerate the ageing process [1].

Rehabilitation of existing infrastructures can protect them from extreme environmental conditions and improve their efficiency, thereby extending their service lives. Maintenance and strengthening are the common practices used to improve the performance of infrastructures. Maintenance refers to a series of regular inspections and repairs to preserve the serviceability of the infrastructure assets. Maintenance procedures are normally carried out in response to a failure or in accordance with a predefined schedule. Although infrastructure performance is preserved, the current strategies are often inefficient because unnecessary or poorly scheduled maintenance activities can lead to increased costs and infrastructure downtime [2].

Besides maintenance strategies, strengthening of structures using fibre-reinforced polymer (FRP) composite materials is an alternative that has been used in the construction industry for more than two decades [3]. The prominent properties of FRPs, such as their high strength-to-weight ratio and resistance to aggressive environments, have extended their application for strengthening purposes, particularly in concrete structures. The efficiency of the FRP-reinforced structures is determined by the overall composite system's capacity to transfer load, with the bond between the FRP and the structure serving as a determining component. Although FRP strengthening has proven effective in improving the load-bearing capacity of structures, its long-term performance is compromised due to the interaction with extreme environmental conditions. This necessitates exploring the long-term performance of the composite structures. Durability is commonly studied through accelerated ageing experiments, in which the strength of a composite structure is tested after being exposed to different conditions such as moisture, varying temperature, and chemical environment.

Although durability experiments provide valuable information on ageing at the structural level, they often miss the underlying interactions between the material and its surrounding environment at lower scales. This creates uncertainty around the long-term performance of the structures, which leads to introducing high safety factors in the design of strengthening solutions. Therefore, it is important to study the impact of environmental ageing on the composites at the material level. Computational micromechanics is a numerical method to analyse composite materials, including FRPs, at microscopic scale. This method benefits from incorporating composite constitutive material behaviours in modelling. As a result, it is possible to investigate the distinct interaction of the FRP components, namely, fibre, matrix, and interphase, with the surrounding environment. Therefore, employing this method allows for a better understanding of the material degradation caused by ageing and its impact on the overall performance of infrastructure at larger scales.

Despite its precision and high fidelity, numerical modelling requires considerable computational effort to predict the evolution of composite material degradation over time. In recent years, machine learning has been introduced as an efficient tool to provide solutions to engineering problems with uncertainties [4]. Its ability to learn from data of different types and sizes, has made it an attractive option for various engineering applications. As a subset artificial intelligent, machine learning is a main contributor to data-driven approaches. It comprises algorithms that can learn pattens and relationships in the data, identify important features, and make predictions based on that learning. Machine learning methods are mainly divided into supervised and unsupervised learning. In supervised learning, a model is developed through learning from the available data in order to predict the unknown future examples. By splitting data into train and tests sections, the ML uses the training part to discover the existing relations between the parameters. The test data is utilized to assess the performance of the developed ML model. Supervised learning methods are divided into regression and classification. In a regression problem, it is intended to make predictions by mapping the data to a continuous function. Methods such as linear regression, ridge and lasso regression, support vector regression, and regression decision trees are instances of famous techniques that are used to develop prediction models in this thesis. In classification, the purpose of the model is to map the data into discrete categories. This method is used to classify data into the intended groups. Methods such as logistic regression, classification decision trees, support vector machines, naïve Bayes, and K-nearest neighbours are the examples of popular algorithms used in different chapters of the thesis for classification purposes.

The Integration of experimental and numerical studies with machine learning has significant implications for addressing the impacts of ageing across multiple scales of infrastructure. By harnessing the power of these combined approaches, new frameworks can be developed that offer various practical benefits, as follows:

- Traditional methods of analysing the effects of ageing on infrastructure often involve time-consuming and costly experimental testing or complex numerical simulations. However, by incorporating machine learning, multiscale ageing studies can be processed rapidly with vast amounts of data and make accurate predictions. This results in significantly reducing the computational time and resources required and accelerates the decision-making process, allowing for efficient rehabilitation strategies.
- The existing prediction models are highly empirical, and their predictive capabilities are limited to a small number of experimental results, which undermines their application in practice. By employing data-driven machine learning-based techniques, the efficiency of the predictions is enhanced through the integration of various data sources.
- The capacity of estimating the impacts of ageing across multiple scales allows infrastructure stakeholders to identify potential defects, optimise resource allocation, and prepare for timely maintenance and strengthening plans. This approach can prevent costly failures, improve infrastructure safety, and extend its service life, resulting in significant economic and societal improvements.

1.2 Scope of research

The goal of the current research is to demonstrate how machine learning can improve the analysis of the ageing impact on infrastructures at multiple scales. We want to use machine learning in particular to predict the state of structures and materials that are ageing as a result of interaction with their environment. We believe that this understanding can help improve the analysis of infrastructure ageing across multiple scales. This can lead to the development of targeted and effective strategies for maintaining the infrastructures. The research is mainly divided into four stages. The following describes each of these phases.

1. To perform a literature review on the methods for analysing the impacts of ageing on infrastructures at multiple scales.
2. To develop a machine learning framework based on historical maintenance information for analysing the condition of the assets within an infrastructure system.
3. To develop machine learning frameworks based on experimental studies to analyse the impact of environmental ageing on the long-term performance of composite structures.
    3.1. To investigate the durability of the adhesive bond under moisture condition in externally strengthened concrete structures with FRP composites.
    3.2. To investigate the durability of the bond under various environmental conditions in concrete structures with FRP bars.
4. To develop a robust machine learning framework based on micromechanical numerical simulations for analysing the impact of natural humid ageing on the FRP composite materials.
    4.1. Simulating the FRP’s transverse behaviour using the finite element (FE) method.
    4.2. Predicting the FRP mechanical properties using a machine learning method.
5. To develop a multiscale FE and machine learning framework to model the composite material ageing process at macro- and microscales under water immersion conditions.
    5.1. To simulate the kinetics of composite water uptake by solving a macroscale diffusion problem.
    5.2. To perform a microscale stress analysis to reflect the impact of material moisture content on its transverse properties.

1.3 Research questions

With the aforementioned research scopes, the following questions are tended to be addressed throughout the thesis.

1. Since infrastructure ageing causes faults and failures in assets, it is important to anticipate these defects before they occur. In this regard, how can we develop practical solutions to efficiently for the potential failures of assets for predictive maintenance purposes?

2. Although FRP composites are regarded as effective solutions for strengthening structures, their contributions can be compromised due to interactions with surrounding environments. How can we address the current limitations in durability study of composite structures, e.g., lack of universal agreement on the long-term performance of these structures and predictive capabilities of existing models, by machine learning?
3. Numerical modelling based on computational micromechanics has been largely used as an effective tool for studying FRP composites at the material scale. How can we address the high computational effort in numerical simulations by the aid of machine learning?

1.4 Outline of the thesis

The rest of the current chapter provides a brief overview of the topics covered in the thesis. The remaining chapters are organised as follows:

- Chapter 2 explores the impacts of ageing on infrastructures and the rehabilitation methods to address the resulting deterioration. It discusses the different maintenance strategies including the FRP strengthening technique. It also emphasises the challenges associated with these methods and approaches to assessing the impact of ageing in infrastructures at different scales.
- Chapter 3 presents a framework based on data analytics and machine learning to analyse the state of a railway infrastructure system. Using the provided data on the history of railway maintenance, the framework is developed to predict the assets' conditions in order to identify potential defects. The influence of different asset characteristics on the predictions is studied, and equations are proposed to assess the condition of the infrastructure in practice. The study proposes a predictive approach that can be used for maintenance purposes and is efficient and cost-effective.
- Chapter 4 presents a machine learning-based framework to address the ageing issues in composite structures. In doing so, two comprehensive databases are established on the durability experiments of reinforced concrete members with FRP composites. These datasets were then used by the developed machine learning framework for two purposes: (i) modelling the long-term performance of FRP-to-concrete adhesively bonded connections under moisture conditions, (ii) modelling the FRP bar-to-concrete bond behaviour subjected to various conditions, including water immersion, chemical exposure, and freeze-thaw environments. The framework is able to model the strength and failure mode of the bond with respect to information on the geometrical, mechanical, and environmental properties

of the composite structures. The developed approach is able to address the lack of universal agreement in the literature regarding the long-term performance of the bond in composite structures.

- Chapter 5 presents a combination of micromechanical numerical modelling and machine learning to investigate the ageing issue in the composite materials. It is intended to develop a machine learning prediction model to predict the transverse properties of the glass fibre/epoxy composite subjected to humidity ageing. First, a framework for automatically generating microstructures for micromechanical analysis is proposed. A FE approach is then established to obtain the influence of various parameters on the composite's transverse behaviour. A machine learning model is finally developed using the microstructure images and the outputs of the finite element analysis in order to predict the transverse modulus of the FRP under humid conditions. The methodology developed in this chapter demonstrates the efficiency of machine learning in capturing the impact of material constituents on the macroscopic properties of FRP composites as they age. Therefore, an understanding of the underlying degradation mechanism of the material is obtained. This can aid in predicting the long-term performance of composite structures and designing effective strengthening solutions.
- Chapter 6 presents a multiscale FE framework integrated with machine learning to model the ageing process due to the water immersion in the composite material. In the proposed approach, a diffusion problem is solved by the FE simulation of moisture uptake in a composite macro-model. The resulting water concentration values are used to modify the constitutive material properties at the microscale. A micro-stress analysis is then performed to obtain the composite traverse properties. The proposed framework is validated by comparing its responses to the experimental observations. A machine learning model is trained based on the outputs of the multiscale numerical framework to model the material degradation with time. A feature importance analysis is also performed to study the contribution of different factors to the material's ageing. The findings of this chapter can extend our understanding of the ageing process in composite materials.
- Chapter 7 concludes the research findings and provides suggestions for the future study.

Finally, Appendices A, B, and C present the information on the copyright agreement for using figures in Chapter 2, as well as the details of the databases and the prediction equations used in Chapter 4 to model the durability of FRP-concrete connections.

The outline of the thesis is also illustrated in Figure 1-1.

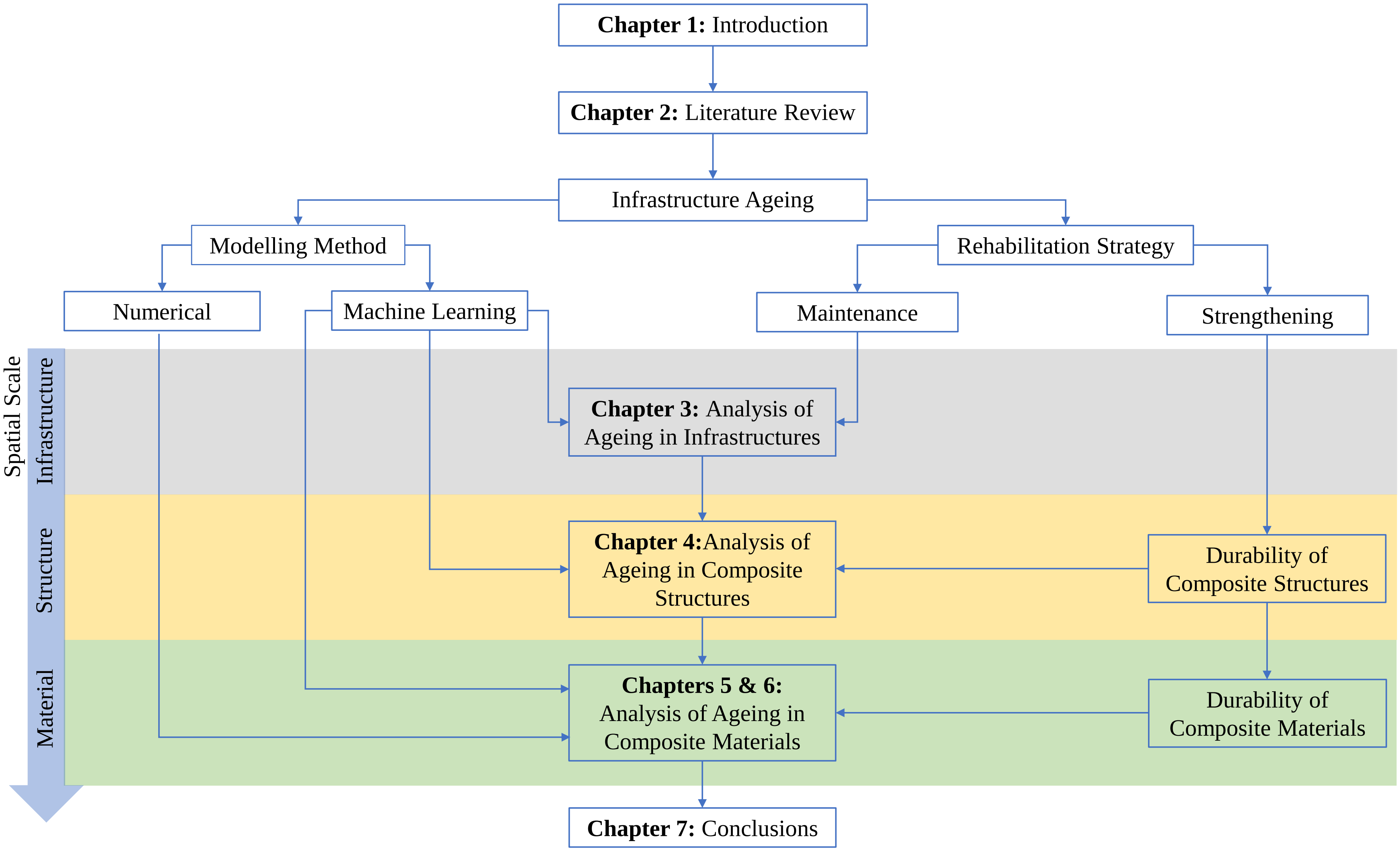


Figure 1-1. Outline of the thesis

CHAPTER 2

Literature Review

2.1 Introduction

Infrastructures provide the foundation for economic growth, facilitate the operation of society, and improve the quality of life for residents. Maintaining the service life of infrastructures is important to ensure their continued functionality in society. Changing environmental condition, increased levels of usage, and natural disasters, can accelerate infrastructure ageing and reduce its service life [1]. These conditions can cause the materials and components of infrastructures to degrade and eventually fail over time.

Environmental ageing can affect the operating assets, compromising the reliability and safety of the infrastructure system. At the structural scale, ageing reduces the long-term durability and causes degradation of key components, such as beams and columns. Furthermore, the properties of the materials can change during the life of the structures. These factors can reduce the service life and increase the maintenance cost. Therefore, it is necessary to understand the impacts of the ageing process across different levels of infrastructure.

Rehabilitation by maintenance and strengthening techniques can help protect existing infrastructures from deterioration. They can be used to ensure the integrity and performance of infrastructures throughout their service life. Regular maintenance practices can help identify and address potential issues before they become major problems. Strengthening with new materials can also help protect the existing structural elements from aggressive environments and increase their load-bearing capacity. This results in extending the service life of the infrastructure system.

Machine learning, as one of the major divisions in the artificial intelligence field, has been used for more than a decade to automate the solutions to structural engineering problems [4]. Its capability for learning from data and capturing the input-output relationships of mathematically complex systems has made it as an efficient prediction tool. The application of machine learning in the rehabilitation of infrastructures can help optimise the service life of the assets through modelling their degradation and forecasting their failure [5]. This can help identify the areas that are most at risk and prioritise the rehabilitation efforts accordingly.

2.2 Ageing of infrastructures

Infrastructures consist of networks of facilities in different sectors such as transportation, energy, water, and communications. Their reliability and serviceability are crucial for a society’s functionality. Besides, a country’s economic growth is dependent on its investments in infrastructures. As an example, a 2022–2023 federal budget indicated that the Australian government has allocated $9.6 billion for infrastructure projects, with the majority of investment going to the transport sector [6].

Deterioration is an unavoidable issue that affects infrastructure’s performance over time. A report by the Australian Infrastructure Audit indicated that 60% of Australia’s assets are considered long-term assets [7]. Ageing can lead to decreased serviceability, increased maintenance costs, and eventually, the need for replacement. The ageing process can be influenced by factors such as building materials, environmental exposure, infrastructure function, and loadings. Figure 2-1 depicts the impacts of deteriorating infrastructures in different sectors. It is seen that ageing can affect a wide range of sectors and have economic, environmental, social, and safety effects.

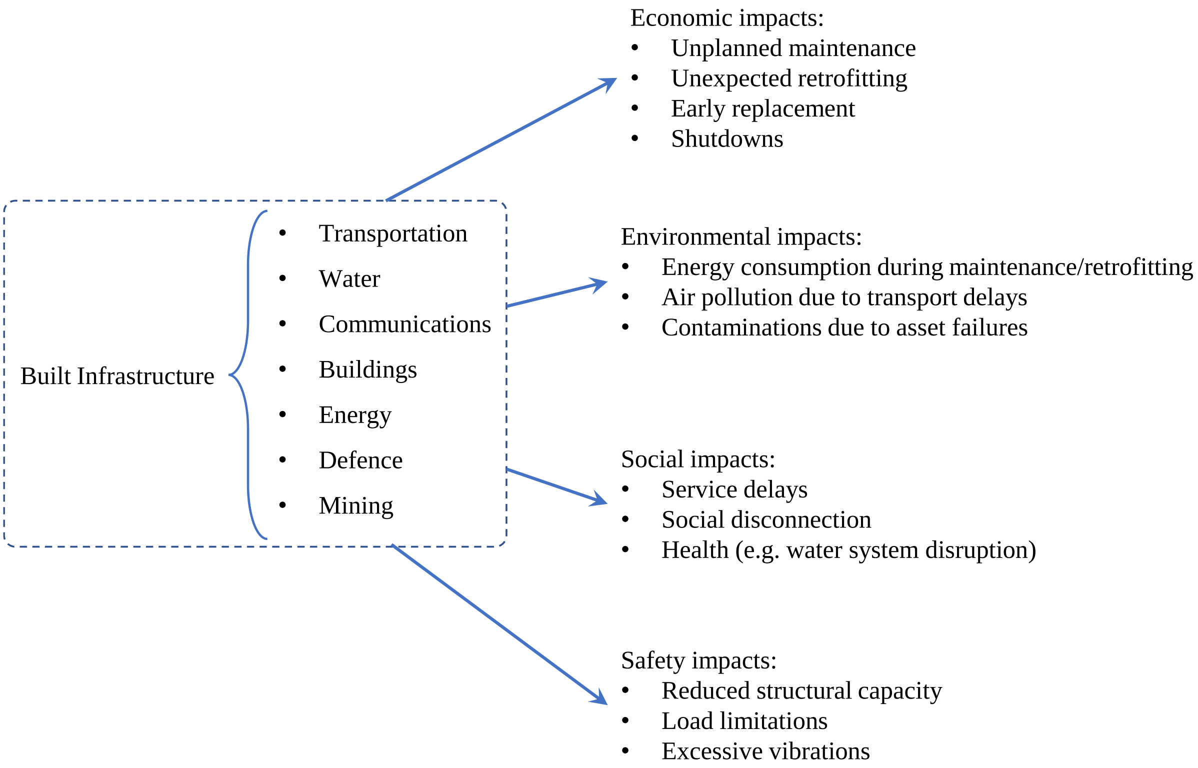


Figure 2-1. Impact of ageing infrastructures [8]

To characterise ageing of an infrastructure, it is necessary to understand the built components of it at multiple scales. The onset of deterioration starts at the microscopic level, where the building blocks of the infrastructure system start to degrade due to interaction with the service environment. Figure 2-2 depicts the impact of extreme environment on a bridge system across multiple scales. The bridge comprises several components, with reinforced concrete as its key material. Due to the porous nature of concrete, aggressive agents can penetrate it and migrate to the steel, which results in corrosion at the microscale. This indicates that a deterioration process is initiated by the interaction of the service environment with infrastructure at the microscopic scale.

Besides the steel reinforcement corrosion, other conditions such as exposure to chemical environments such as alkaline and acidic cause expansion followed by stress inside the concrete, which results in cracks and deterioration. In addition, cycles of repetitive freeze and thaw can cause spalling in the concrete, which leads to structural failure. The deterioration mechanisms can compromise the durability of the structures and limit their serviceability. This results in increased rehabilitation costs, reduces efficiency, and raises safety risks in the infrastructure. Understanding the deterioration mechanisms can help develop maintenance and strengthening strategies in order to extend the service life of assets and improve the efficiency of the built infrastructure.

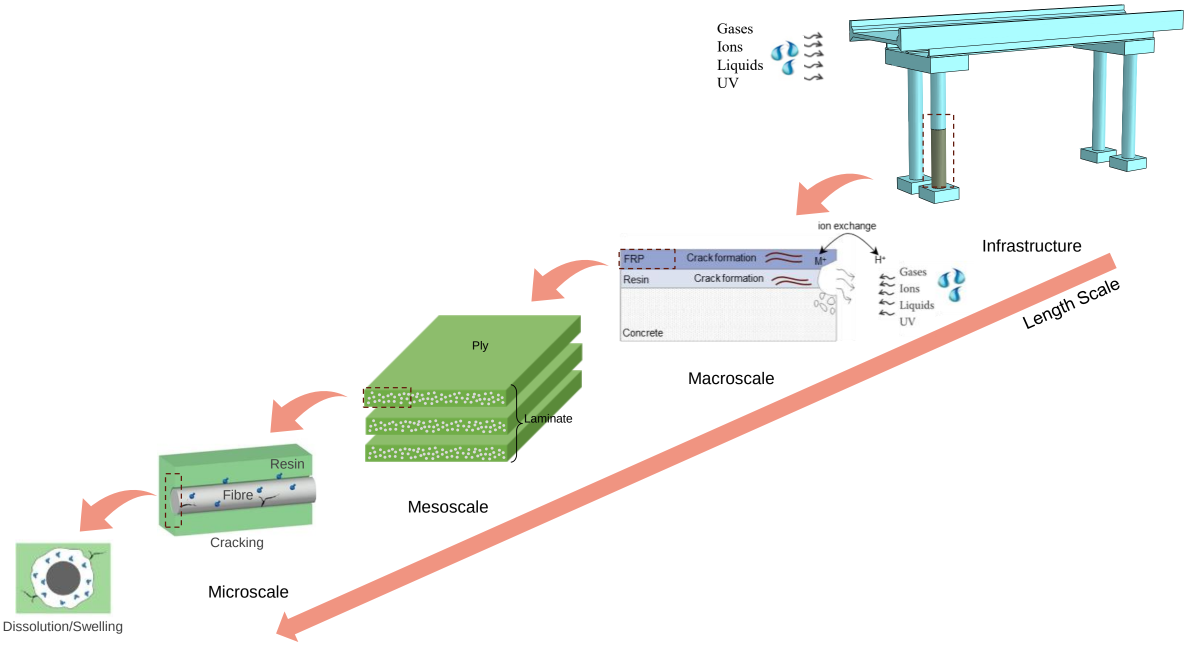


Figure 2-2. Mechanism of ageing across multiple scales [9, 10]

2.3 Rehabilitation of existing infrastructures

In the recent decade, a considerable number of structures have shown obvious indications of degradation [11]. This has increased the need to rehabilitate the ageing infrastructures. The performance of structure is indicated either with qualitative parameters describing its overall condition, such as a damage index, or by quantitative values, such as a reliability index, which explains the physical properties including strength and displacement [12]. As the performance of structures degrades over time, rehabilitation strategies are implemented to preserve, repair, and improve the condition of the assets. This can extend the service life of structures, support society's economic growth, and reduce the environmental impact of constructing new infrastructure. Maintenance and strengthening are the common rehabilitation approaches to ensure the reliability and durability of the structures. Maintenance involves regular inspections and repairs to keep the infrastructure in good working condition [13]. On the other hand, strengthening refers to the techniques that involve modifying the structures to retain or improve their performance [14]. Despite the differences, these rehabilitation techniques pursue a common goal, which is enhancing the reliability and extending the service life of the infrastructure. A maintenance practice can help identify the region in the structure that requires repair or strengthening. On the other hand, upgrading the structural capacity by strengthening techniques can reduce the need for frequent maintenance attempts, which can reduce the rehabilitation costs of the infrastructure. In Sections 2.3.1 and 2.3.2 these rehabilitation strategies will be addressed in more detail.

2.3.1 Maintenance of infrastructures

Maintenance refers to a series of technical and managerial policies taken into action during the service life of an infrastructure in order to restore it to the state that it can serve beyond an acceptable performance threshold [15]. It involves a series of regular inspections and repairs to maintain the serviceability of the infrastructure assets. The maintenance strategies are broadly divided into reactive and proactive approaches. They are different based on their intervention time and the extent of the maintenance task. Figure 2-3 displays different maintenance strategies. In the reactive method, maintenance is performed after a failure occurs in the asset due to a lack of scheduled downtime. It includes repairing the defects so that the asset can resume its service. Although reactive maintenance is cost-effective at the early stages of the asset's service life, it becomes less efficient in the long term when degradation leads to frequent failures [16]. Moreover, accidental failures can create serious safety issues and impact the environment.

Proactive maintenance is applied before the asset reaches its minimum performance threshold. This practice involves analysing the degradation of assets and applying necessary corrections before the cost of maintenance increases. A proactive approach is divided into preventive and predictive methods. The preventive strategy is applied based on a time schedule or asset condition. The former is known as the time-based method, where the maintenance is performed following the predefined schedule regardless of the asset's current condition [17]. Although simple, this can result in unnecessary maintenance tasks. In the condition-based approach, regular inspections are performed to check the condition of the assets for any degradation and take necessary maintenance action. Besides improving reliability and efficiency, the preventive method can reduce the asset's useful life [17]. Also, frequent shutdowns for maintenance impacts the serviceability of the assets.

Besides preventive and reactive, which are conventional maintenance approaches, predictive method is the most cost-effective type, which can increase the asset efficiency by 90% [5]. Predictive approach extends asset's service life and minimises the costs by providing maintenance before the asset failure. This can avoid the unexpended downtimes and maximise the serviceability. In this method, the defects are predicted before the occurrence and the resulting risk from the failure is avoided. The predictive maintenance (PdM) can be divided into three approaches, namely, physical, data-driven, and hybrid models [18]. In physical method, the degradation of the asset is modelled to assess its performance for maintenance. Data-driven models predict the system performance by learning it from the historical data. Common tools for data driven PdM are statistical, and machine learning methods. Hybrid methods are the combination of the mentioned two methods where the data generated by the physical model is used to predict the asset's future performance [18]. The challenges concerning the PdM are the implementation of these methods in complex systems and the data availability.

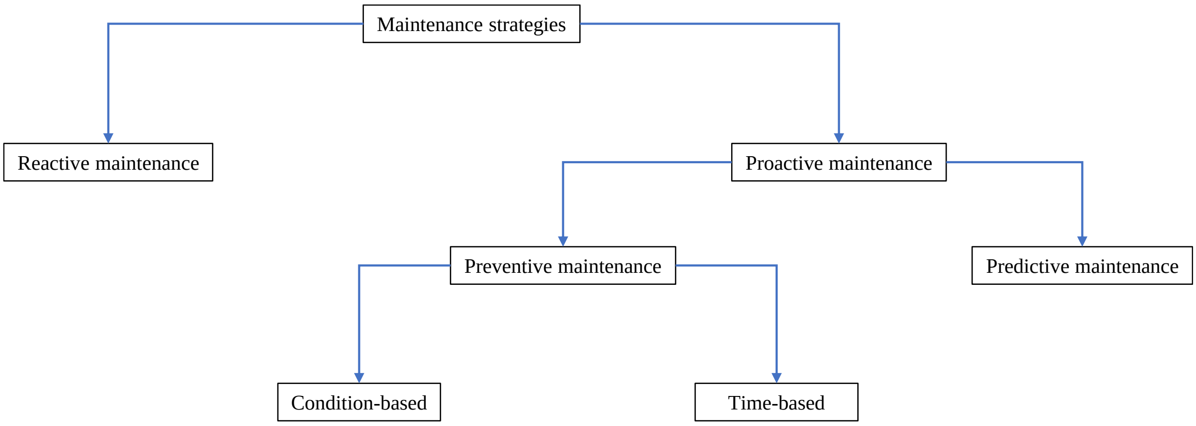


Figure 2-3. Division of maintenance strategies

Figure 2-4 shows the condition history of an asset over time when its service starts at $t = 0$. It is seen that, at the beginning, the performance is indicated as v_0. However, the structural condition degrades with time. Preventive maintenance is applied at t_1 to prevent the system from failure and improve its condition. However, in a reactive approach, the asset is operated until it fails, and then maintenance is applied at t_2 in order to return it to service. The predictive practice is applied at t_3, when the asset has reached the end of its useful life. This reduces the need for expensive replacements and repairs.

Although a preventive approach can enhance infrastructure reliability, the overall cost due to frequent maintenance is high. Moreover, for the reactive approach, the repair cost due to the lack of maintenance is high as well. However, in predictive method, the optimal time for maintenance allows for the asset to use its potential useful life by avoiding unnecessary maintenance tasks. Therefore, devising a strategy with optimal maintenance cost is essential to preserve the serviceability of the infrastructure and extend its service life.

Although maintenance techniques are beneficial, they often overlook the underlying mechanisms that lead to the failure of the asset. As a result, further strategies can be proposed that involve how the structures interact with the service environment in their design. Strengthening techniques are suitable options for mitigating infrastructure ageing by increasing the structural capacity in response to the environmental conditions. This can help structures better withstand the stresses due to exposure to various loads and environments.

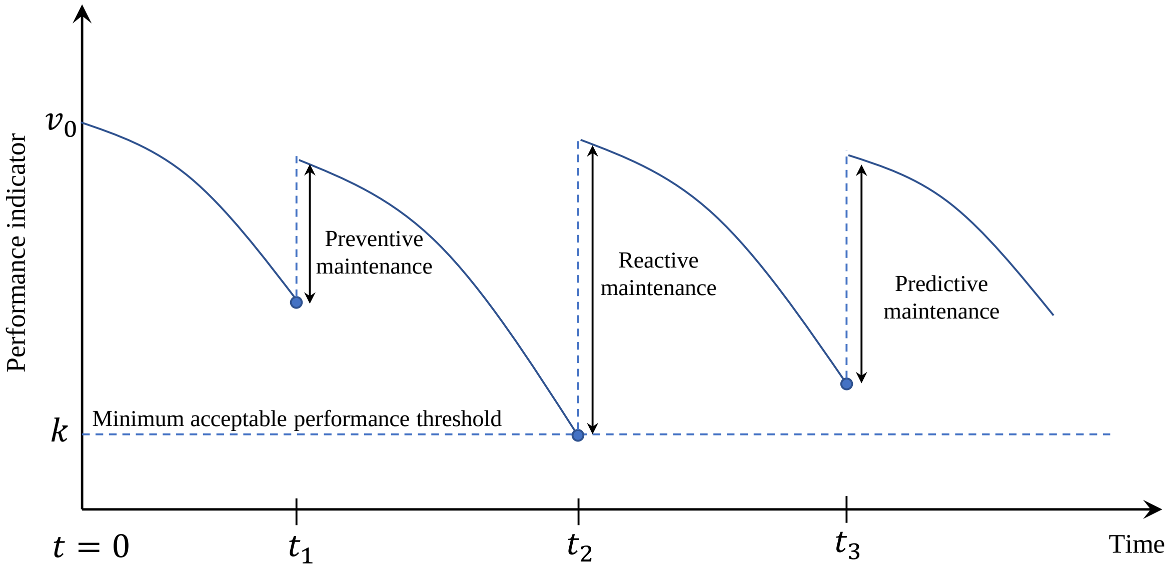


Figure 2-4. Preventive, reactive, and predictive maintenance methods

2.3.2 Strengthening of infrastructures

Apart from maintenance strategies, structural strengthening to protect the existing assets from extreme service environments is an alternative to extend the service life of the infrastructures. Strengthening refers to the technique used to increase the load-carrying capacity and enhance the performance of structures. This can increase the serviceability and enhance the durability of structures by reducing the risk of failure during their long-term interaction with the service environment.

A selection of a proper strengthening technique depends on various factors, such as the condition and extent of damage in the existing structure, the geometrical configuration of structural members, and the cost of the strengthening solution [19]. Some of the common forms of strengthening are jacketing, external plate bonding, and post-tensioning. Jacketing involves enlarging the section of the reinforced concrete members by adding a new layer of material around the existing member. Common forms of jacketing are reinforced concrete and steel plate. In the former, the deteriorated concrete and corrosion are first removed from the member. Then, the surface is cleaned and prepared for the addition of new reinforcement. Finally, the new layer of concrete is applied. However, when the enlargement of the section by additional reinforcement is limited, steel jackets are fabricated and welded around the existing member. Although jacketing techniques improve the load-bearing capacity of the members, they are labour-intensive and prone to corrosion in the case of steel jackets [19]. External plate bonding is a widely used method where a new layer of steel is attached to the external surface of the beam or slab members using adhesives. This method results in improved flexural and shear strength and stiffness of the members. However, the durability of this method is compromised due to the steel corrosion and the stress concentration at the ends of the plates [20]. External post-tensioning is a cost-effective solution that is applied to the prestressed reinforced concrete members. With the main applications in the bridge infrastructure, this method involves installing prestressed tendons to the outside of a structural member by anchoring them to its ends. This can help improve the flexural and shear capacities of the structure and reduce deflections [21]. Although this method can improve structural capacity, its application is complicated and requires skilled labour.

Apart from the aforementioned methods, using advanced materials has proven to be a promising approach for strengthening existing structures. Fibre-reinforced polymer (FRP) composites have been extensively used to repair and strengthen concrete structures in the last two decades [3]. The prominent properties of these materials, such as their high strength-to-weight ratio and resistance to aggressive environments have extended their application for strengthening purposes. Figure 2-5 presents the Ashby plot of the relationship between the density and Young's modulus for different materials. It is seen that composites represent a higher strength-to-weight ratio compared to most classes of materials. FRPs have been successfully used for strengthening of concrete structures, either through external bonding or as internal reinforcement [9, 22]. These reinforcing methods have been recognised as promising techniques for improving the long-term performance of existing structures. FRP applications in structures can improve the strength capacity and serviceability of the concrete members by providing additional confinement.

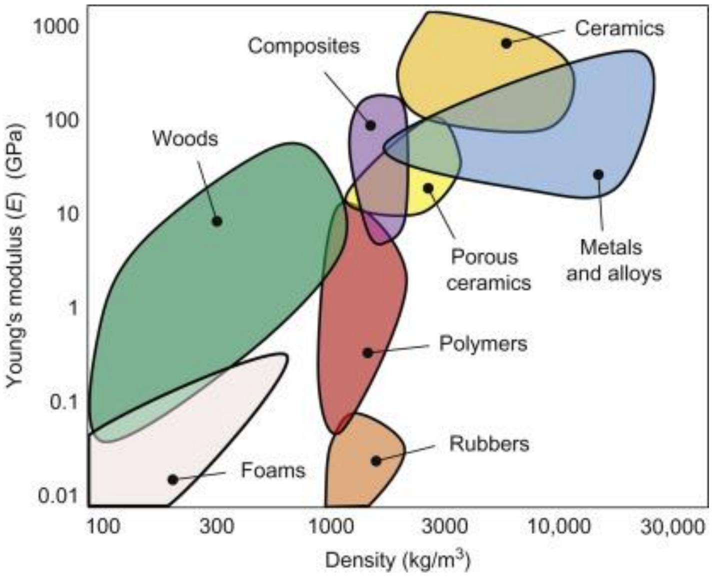


Figure 2-5. Young’s modulus of materials according to their density [23]

The strengthening of concrete element by FRP composites is commonly performed by hand lay-up and pultruded laminate systems. Hand lay-up is a cost-effective and flexible method that can be applied to a wide variety of components [24]. In this method, the concrete surface is first prepared and cleaned. Then unidirectional or woven fibre sheets are cut into the desired dimensions. Finally, resins are used to impregnate the sheets with rollers or brushes. Pultrusion is an automatic technique to mould FRP laminates. Impregnating involves either injecting the resins or running the fibres through a resin solution. A heated die is used in this method to cure the resins and mould the impregnated fibres. Finally, the cured laminates are cut into the desired shapes. The pultruded laminates are bonded to the concrete members with adhesive resins, the type of which is determined by the laminate’s application.

Besides the aforementioned techniques, resin infusion and resin transfer moulding techniques have been introduced recently for structural strengthening. These methods have been used in various industry sectors and are able to provide high-quality composites at low cost and on large scales [25]. To process the composite, the FRP sheet is placed on a rigid mould, and a vacuum is applied to create a low-pressure environment to draw the resin into the fabric. Applying vacuum pressure can reduce the formation of voids in the composite and result in components with better mechanical properties [26].

FRP materials have been successfully used in the construction industry for strengthening purposes. However, the interaction of the structures with extreme environments during their service life can undermine the effectiveness of the FRPs, leading to degradation of the strengthened structures. This highlights the importance of considering the durability of composite structures when designing FRP strengthening solutions for structures.

2.4 Durability study of FRP composites under environmental ageing

The efficiency of the structural members reinforced with FRPs is dependent on the combined performance of the entire composite structure in bearing the load. The bond between the FRP and structural member is a critical component in transferring the load, which determines its load-bearing capacity. The composite structure is expected to retain a large proportion of its strength during its service life. However, exposure to a wide range of environmental conditions can undermine the serviceability of composite structures, especially by compromising the integrity of the bond. This necessitates studying the durability of the FRP-reinforced structural elements.

The durability of the composite structure is commonly studied by performing accelerated ageing tests. In these experiments, the effect of environmental conditions on long-term performance is evaluated by measuring the bond strength in the FRP-strengthened member. Among various service conditions, exposure to moisture, varying temperatures, and chemical environments can greatly influence the connection of the composite structure [27]. In the following, the influence of each environmental condition is explored.

2.4.1 Impact of moisture diffusion

Moisture is one of the most prevalent environments that can degrade the long-term performance of structures strengthened with FRPs, particularly concrete members. It can diffuse into the interface of the composite connection and degrade the bond between FRP and concrete [27]. The epoxy used in the FRP-concrete adhesive connections and the FRP composite is prone to moisture absorption. The interaction of water and epoxy can lead to reversible mechanisms, such as plasticisation and swelling, or irreversible degradation by hydrolysis [28]. As a result of plasticisation, the mobility of the polymer chain increases due to the weak hydrogen or Van der Waals bonding between the water and polymer molecules [28]. This also creates swelling due to disruption in the molecular arrangement. During hydrolysis, the breakage of the polymer chain leads to a decrease in the glass transition temperature (T_g), which irreversibly compromises the stiffness of the epoxy [29]. Moreover, chemical reaction of water with fibres, particularly the glass-type fibres, can induce stress corrosion leading to microcrack growth [30]. Moisture can also weaken the bond by diffusing into the interface of fibre and matrix, which can result in microcracks in the matrix [31]. Figure 2-6 explains the degradation mechanism of the FRP material subjected to moisture condition. It is seen that water diffusion into the fibre-matrix interface causes the resin to swell and dissolve, which leads to microcracks in the composite.

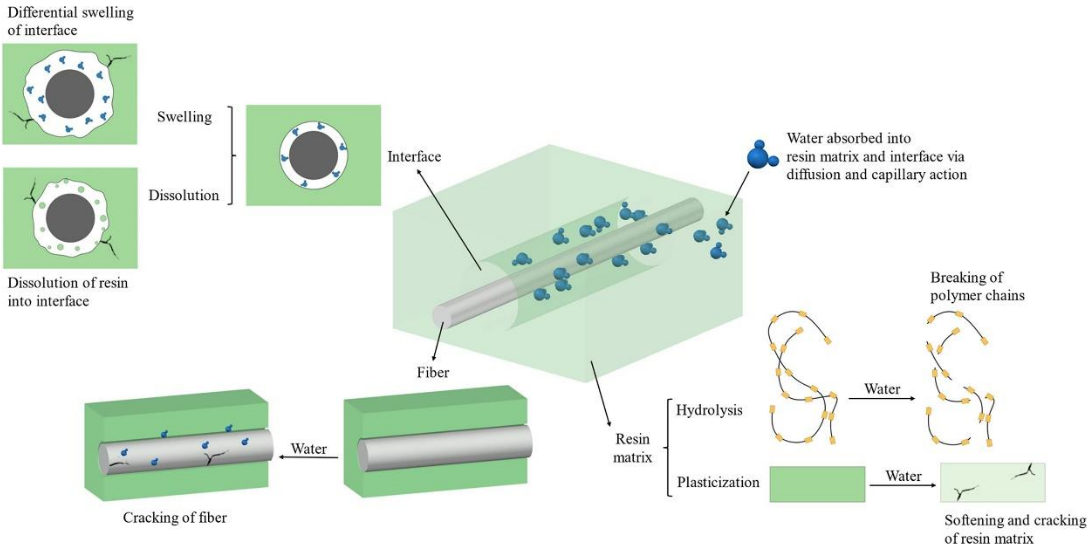


Figure 2-6. Degradation of FRP material under moisture condition [10]

Several studies evaluated the bond's durability in FRP-concrete connections under moisture conditions. In these works, the effect of environmental conditions was investigated by measuring the bond strength in the FRP-strengthened concrete members. Most of the studies revealed the detrimental impact of moisture on the bond, where the exposure duration was highly correlated with the strength degradation rate [32]. The pull-out test results also indicated a shift in failure from the concrete layer to the interfacial region [33], implying that moisture can degrade the bond with time through weakening the interface in the adhesively bonded connections. In other studies, it was revealed that bond strength was either negligibly changed or increased after exposure [22, 34]. The reason for such behaviour was related to the improvement in the concrete's compressive strength with time [34].

2.4.2 Impact of temperature

Thermal exposure can adversely impact the mechanical properties of the composite structures. Several studies investigated the impact of cyclic temperature changes by performing freeze-thaw experiments on FRP-concrete connections. The difference in the thermal expansion of the FRP and concrete was the reason for premature debonding [35]. Moreover, at temperatures below freezing, resin hardening and microcracking lead to degradation of the interface between matrix and fibre in the FPR material [36]. On the other hand, other studies argued that the FRP-concrete bond was minimally impacted by the cycles of freeze and thaw [37].

The effect of elevated temperature on the adhesively bonded connections is normally controlled by the resin matrix or adhesive. When exposed to temperatures close to or above T_g, molecular relaxation occurs in the resin, due to which the stiffness of the material decreases [38]. In addition, a chemical reaction between polymer and oxygen molecules causes oxidation, which decreases T_g and accelerates ageing in the elevated temperatures [39]. The softening of the polymer at elevated temperatures also increases its moisture absorption, which accelerates the degradation of the bond in a humid environment [33].

The results on the durability of the composite structures under thermal exposure are contradictory in some studies because the experiments were undertaken with different adhesive material types, temperature ranges, and heating protocols. For the cases with improved bonding properties, it was seen that the elevated temperature was beneficial by providing a post-cure to the resin [40]. However, it is widely accepted that the FRP-concrete connection is adversely impacted by temperature, where the bond strength is compromised and the failure mode is changed to interfacial debonding [27].

2.4.3 Impact of chemical environment

The long-term performance of a composite member in chemical environment is largely determined by the durability of its components in reaction with the chemical solutions [27]. Figure 2-7 shows the interaction of a composite structure with the corrosive environment. It is seen that due to the chemical interaction with the aggressive agents, the bond is compromised, which results in the separation of the FRP layer from the concrete substrate.

Research into the effects of saline environments has shown that lower concentrations of salt are more detrimental to the bond, while concentrations between 5% and 15% indicate no strength reduction [41]. It was also shown that the rate of degradation in saline and alkaline solutions is proportional to temperature since higher temperatures accelerate the corresponding chemical process [41]. Glass FRP (GFRP) was also shown to be more vulnerable to salty and alkaline environment than carbon FRP (CFRP) [42]. Several studies found that the degradation of the adhesive layer due to interactions with the mentioned solutions can also reduce the durability of composite connections [43]. Hadigheh et al. [9] conducted a study of the influence of acidic environments and discovered that acid ingress has a negative impact on the durability of the composite structures owing to the chemical interactions between acid and the structural components.

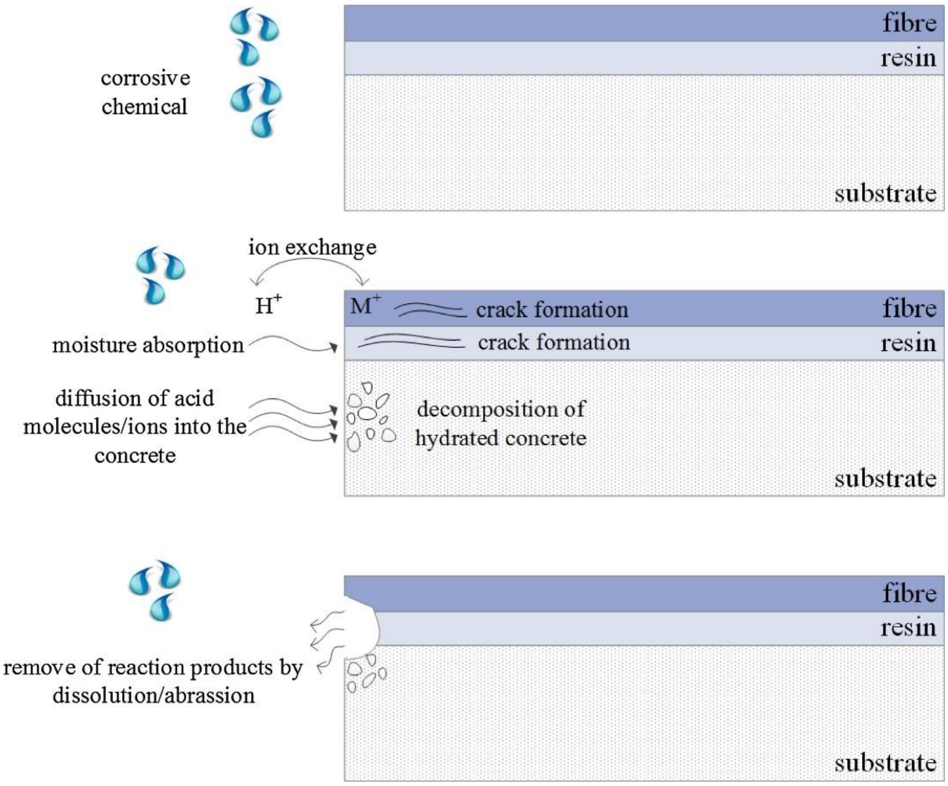


Figure 2-7. Reaction of composite structural element with chemical environment [9]

The review of the studies on the durability of the composites indicates that there is a lack of universal agreement on the long-term performance of composite structures. One reason for this is the lack of a standard procedure for performing the durability tests, which results in inconsistency between the experimental protocols in different studies. This demonstrates the need for modelling the degradation of the composite by accurate procedures. The following section investigates the numerical approach for analysing composite structure ageing at multiple scales.

2.5 Numerical modelling of environmental ageing in FRP composites

Although studying ageing at the macroscopic scale provides valuable information on the durability of structures, it often misses the material degradation mechanisms at the underlying scales. FRP composites are two-phase materials composed of fibres that are heterogeneously dispersed in a continuous matrix. This results in a complex mechanical behaviour that is poorly understood [44]. Owing to the size of the fibre, which is in the range of a few micrometres, the processes influencing the material’s durability occur at a scale much smaller than the macroscopic design level. The lack of understanding of material behaviour leads to introducing safety factors in the design phase. Besides, the experimental studies at the mesoscale also fail to provide knowledge on the microscopic drivers of the composite behaviours.

With the advent of high computational powers, the application of high-fidelity numerical approaches for developing accurate models has increased recently. Computational micromechanics has been regarded as a powerful tool to study the impact of material degradation on composite mechanical behaviour at the microscale. This method allows for incorporating the constitutive models of the material to predict its response using the finite element method.

Two strategies have been used in the studies to simulate the behaviour of the microstructures. The embedded cell method, which is suitable to model local damage mechanisms in the microstructure, and the periodic representative volume element (RVE), which is used to model the effective mechanical properties of heterogeneous material. RVE is a geometrical representation of the composite microstructure which is used to obtain its mechanical behaviour at the macroscale. The determination of the RVE with precise dimensions is critical. Identifying the size of RVE for a specific composite is attained by performing a sensitivity analysis. A few dozen of inclusions were found to be sufficient to represent the composite effective behaviour [45].

The generation of RVE that is close to reality can be attained by image-based techniques [46, 47]. Although these methods are straightforward and able to capture defects and real fibre distribution, they are costly, time-consuming, and require expertise in the field. An alternative to this is generating ideal RVEs using mathematical algorithms [48, 49]. Despite being faster compared to imaging methods, the mathematical generation of RVEs requires computational effort. One of the simple, yet mostly used, algorithms to create microstructures is random sequential adsorption (RSA) [50]. In this method, the centres of the fibres are generated sequentially in arbitrary positions, however, checked for overlapping. The inclusions are generated until the desired volume fraction is reached. This method has been modified by several studies to generate RVEs with higher fibre volume fractions [49, 51]. Besides RSA, different techniques such as nearest neighbour, random sequential expansion, and hexagonal packing algorithms were also investigated to overcome the limitation of RSA [48, 52, 53]. In the case of unidirectional (UD) composites, the fibres in the microstructure are considered to be infinitely long and straight and are represented by circles in the cross section of the RVE.

Figure 2-8 depicts the UD composite ply, which is produced by iteratively generating periodic RVEs using the RSA algorithm.

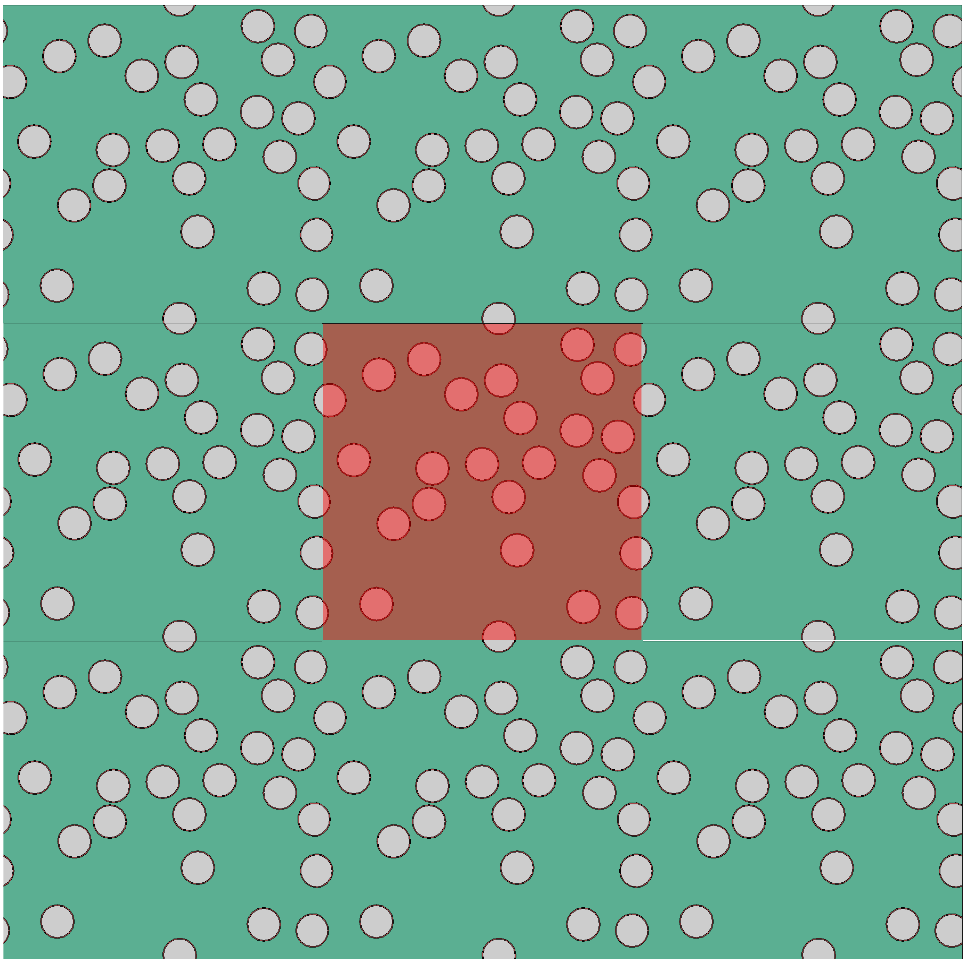


Figure 2-8. Schematic view of a periodic RVE (red) within the region of a UD composite ply

Figure 2-9 shows three scales of interest in modelling the FRP composites. Studying materials in the microscale offers high-fidelity analysis as the material components are distinctly modelled and the degradation mechanism can be included. The major purpose of multiscale analysis is to use the benefits of smaller scale models to provide predictions of macroscopic scale within a computationally efficient approach. Numerical homogenisation is one of the most popular multiscale techniques. It involves averaging the microscale properties across RVE domain to describe the macroscopic behaviour. The idea is to translate the material's heterogenous behaviour into homogenous equivalent medium, which can be analysed easily. This is a bottom-up approach in which the composite's response at the material component scale is passed to the larger scales to obtain the structural behaviour [54]. Homogenisation is a computationally efficient method and does not require experimental calibration to obtain macroscopic properties.

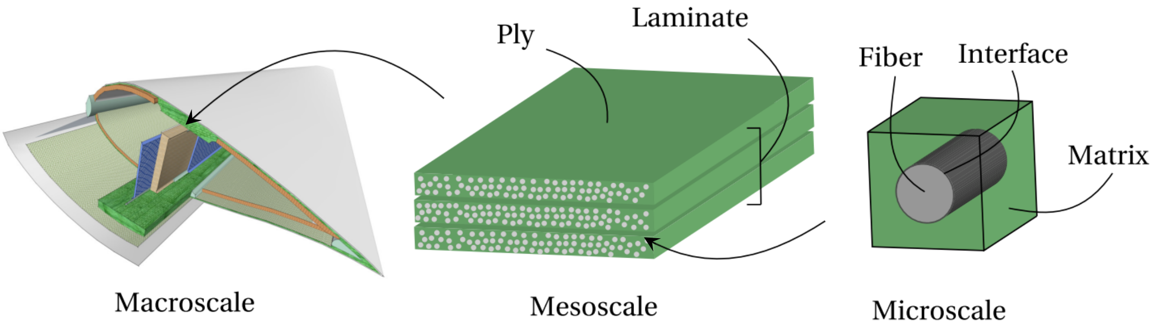


Figure 2-9. Schematic representation of composite structure in multiple scales [55]

Moisture is one of the common environments that can impact the durability of composite materials. The numerical modelling of ageing due to moisture conditions is critical to understanding composite degradation. Since material ageing by moisture starts with water ingress, the diffusion process has been investigated at multiple scales. Joliff et al. [56], [57, 58] studied water diffusion in the unidirectional fibre composite while including the interphase region in the modelling. The study showed a gap between the bulk resin and matrix diffusivity, which was explained by introducing an interphase region around the fibres. They identified that the diffusivity in the interphase area was five times higher than that in the matrix. It was also seen that fibre distribution and interphase properties influenced the water ingress behaviour in the composite. Rocha et al. [59] used a combined experimental and numerical study to explore the anisotropic water diffusion in the unidirectional glass fibre/epoxy composites. The results showed that the fibres acting as barriers against water diffusion were partly responsible for the anisotropy. Despite the findings of Joliff et al. [57] on the diffusion coefficient, it was shown that an assumption of linear variation in the diffusivity within interphase gives a realistic representation of this region.

The FRP material deterioration is resulted from the damages at the resin matrix, fibres, and the fibre-matrix interface. Rocha et al. [44], [60] proposed a numerical framework based on computational homogenisation for multiscale modelling of hygrothermal ageing in the glass-epoxy composites. In order to incorporate the degradation of the constitutive materials, the matrix damage and the interface debonding models were included in the micromechanical analyses. Riaño et al. [61] investigated the effect of interphase on the degradation of the elastic properties of glass fibre/epoxy composites subjected to natural ageing using a micromechanical approach. Quino et al. [62] also used finite element simulations to model the hygroscopic stress from water diffusion into a unit cell of GFRP composite exposed to hydrothermal conditions.

Comparing numerical models to experimental results revealed that the multiscale approach is a promising technique for capturing the deterioration of FRP materials under environmental ageing.

Although numerical methods are accurate tools to model ageing in infrastructures, they require considerable computational effort to predict the evolution of degradation over time. Specifically, in the finite element (FE) modelling of composites, the RVE of the composite needs to be large enough to model the mechanical response of the heterogenous material. Capturing the microstructural complexity requires high-resolution meshes with fine elements, leading to increased computational effort [56]. Furthermore, composite materials exhibit a hierarchical structure throughout their length scale from microscopic details to macroscopic behaviour. Accurately modelling the entire hierarchy can be computationally demanding as it requires resolving various length scales concurrently [60]. These factors can undermine the computational efficiency of FE, particularly for simulating nonlinear responses or crack propagation in composites with complex microstructural geometries [63, 64]. Machine learning algorithms, on the other hand, can propose accurate prediction models with less computational cost, making them an alternative to numerical modelling. In the next section, machine learning methods and their application in analysing ageing infrastructure across multiple scales are reviewed.

2.6 Machine learning approach to analysing ageing infrastructures

In recent years, data-driven approaches have been introduced as alternatives to computationally expensive and time-consuming simulations. Data-driven methods can provide accurate and low computational cost solutions to structural engineering problems associated with uncertainties [4]. Machine learning, as a subset of artificial intelligence, is the main contributor to data-driven approaches. It focuses on developing algorithms that can learn patterns and relationships in the data and make predictions based on that learning. The ability to handle different types of data and make accurate predictions with efficient computation is an important aspect of machine learning.

Machine learning is generally divided into supervised and unsupervised methods. Supervised learning techniques are a set of algorithms that can learn from data and make predictions by capturing the relationships among the system variables. Based on the type of data, machine learning methods can be used as regression or classification tools to predict continuous or categorical values. Unsupervised learning is a set of algorithms that are trained on unlabelled data. This method aims to identify patterns and structures in the data, and group similar data points together into clusters. Unlike supervised learning, there is no specific target output to predict in this method. Instead, the algorithm is performed by finding the underlying relationships within the data. Some of the main applications of unsupervised learning are clustering, dimensionality reduction, anomaly detection, and autoencoders. Figure 2-10 illustrates the categories of machine learning with some of the popular examples used in engineering.

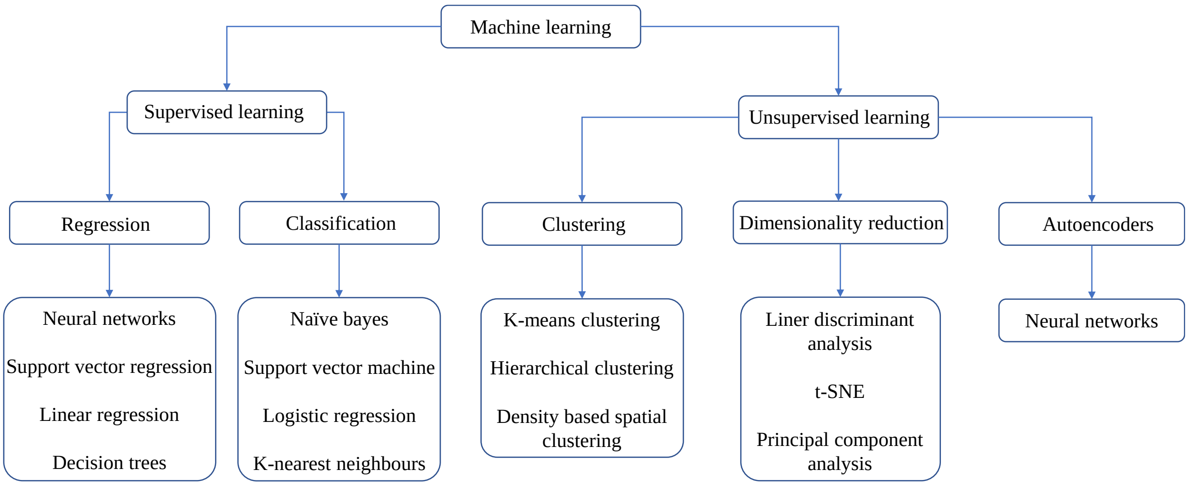


Figure 2-10. Machine learning techniques

The flexibility in handling large and various data types, generalisability to new data, and accurate forecasting of the future states of systems have made machine learning a popular tool in addressing engineering challenges. As seen in Section 2.3, maintenance and strengthening techniques can help extend the infrastructure's service life. In addition, understanding the degradation mechanisms at multiple scales can aid in the development of targeted and effective rehabilitation practices. By combining the insights gained from experimental and numerical studies with machine learning algorithms, new frameworks can be developed that are not only accurate and cost-effective but also lead to more informed decisions regarding the maintenance and strengthening of the infrastructure.

The important stages of developing machine learning models are database establishment, data pre-processing, model training, and model deployment. Considering the case of study, the collected databases can be historical or synthetic. Data pre-processing refers to preparing the data by techniques such as data cleaning and transformation and feature engineering. Model development and validation involves training a machine learning model by the processed data and test its efficiency with new data. Finally, the developed prediction model can be deployed for specific purposes. In the following, the applications of machine learning in analysing ageing of infrastructures in multiple scales will be discussed.

2.6.1 Machine learning assessment of ageing infrastructures for predictive maintenance

It was seen in Section 2.3 that identifying potential faults and failures can help devise efficient maintenance solutions, which can ensure the reliability and safety of the infrastructures. Data analytics and machine learning are powerful tools that can help develop accurate predictive models. The review of the current applications of machine learning shows that this strategy is used in vast areas of industry, with the focus on forecasting the potential failures of the assets. The application of machine learning in infrastructure maintenance can help develop more efficient strategies that are scheduled at optimal costs. Among different machine learning techniques, decision tree-based methods have been the most widely used, accounting for 33% of the applications. This is followed by neural networks, support vector machines (SVM), and k-means clustering methods with 27%, 25%, and 13% applications in the maintenance practices, respectively [65].

The railway network is one of the fundamental infrastructures, as it has a considerable impact on the economic and social growth of a society. The significant increase in the usage of railways, as one of the oldest built infrastructure systems, can lead to degradation of these systems over time. This can compromise their serviceability and create safety issues. Monitoring the condition of the railway system and identifying potential faults can help develop effective strategies for the network's maintenance.

Railway tracks are critical components of the transportation system, and failure in these members can cause safety issues and service delays, compromising the system's efficiency. The improvements in computer vision have made it possible for computers to interpret visual information. By using deep learning algorithms, it is possible to train models that are able to detect defects in railway tracks. Gibert et al. [66] proposed a multi-task learning approach to combine a series of neural network-based architectures to detect the faults in different components of the railway track from its images. Faghih-Roohi et al. [67] proposed a convolutional neural network model for the automatic detection of faults from rail surface video recordings. They were also able to demonstrate the efficiency of the trained model in identifying the types of defects. In similar studies, the defects in different track components, including rails, fasteners, sleepers, and turnouts, were automatically identified from imagery resources using different architectures of neural networks [68-70].

Classical machine learning techniques have also been a source of interest for researchers to evaluate the conditions of infrastructures. They are useful to understand the relationships between different infrastructure characteristics and asset faults. Li et al. [71] developed a predictive model based on the historical data collected from different railway maintenance resources. A SVM model was developed to reduce the false alarm rate on the rail network, thereby removing extra maintenance schedules and costs. Allah Bukhsh et al. [72] used decision-tree-based classifiers to identify the maintenance need, activity type, and status of the triggered maintenance for the railway point machine assets. The data was collected from different resources describing the asset's details, condition history, and maintenance work orders. Ghofrani et al. [73] proposed an integration of physics-based and data-driven processes to predict the defect frequency in the railway. Using the information from historical data, they used finite element modelling to identify the propagation of the crack in the rail. The generated data was then used to train an ensemble model for predicting the defects. In other studies, various techniques such as logistic regression, multivariate regression, genetic algorithms, and support vector regression were used to model the degradation of railway track systems [74-76].

The successful application of machine learning to assessing the state of infrastructures using big data is a promising development, indicating their ability to identify potential failures. This can help plan PdM strategies to improve the serviceability of the infrastructure. However, the degradation of the infrastructure components greatly impacts its overall performance. Since they cannot be fully captured by inspection practices, it is important to understand the underlying ageing process of assets at structural and material levels. This can help to gain a complete understanding of the long-term performance of infrastructure. In the next subsection, the possibility of applying machine learning to capture the degradation process of infrastructure at both structural and material levels will be discussed.

2.6.2 Machine learning analysis of composite structures at multiple scales

In Section 2.3, it was shown that strengthening with FRP composites is an alternative to maintenance strategies for extending the service life of infrastructure. On the other hand, the durability studies showed that interaction with the service environment impacts the long-term performance of the composite structures. In order to better understand the degradation mechanism of composites, high-fidelity numerical approaches were seen as promising solutions to capture the ageing of these materials. Machine learning methods can offer accurate and computationally efficient solutions to study the long-term performance of composites, while avoiding the limitations of experimental and numerical practices. While machine learning has been effectively used to model the response of composites to a variety of loading scenarios at various scales, its application for studying the durability of these materials is missing from the literature. In this subsection, we investigate how machine learning can be used to predict composite properties at different scales.

Machine learning can help us understand the complex relationships between composites and their service environment by making use of large datasets and advanced algorithms. Subsection 2.3.2 demonstrated that FRPs can be used to strengthen existing structures, but their performance is greatly dependant on the integrity of the connection. With the increased use of FRPs, notably in concrete structures, it is possible to explore the complicated behaviour of the FRP-concrete connections using the current databases. Machine learning studies on bond strength has primarily been conducted using neural network architectures. Mashrei et al. [77] investigated the performance of backpropagation neural networks in identifying the bond strength of CFRP-reinforced concrete connections with only 150 test results. Zhou et al. [78] surveyed the last 20 years of data on the strength of FRP-concrete connections to develop neural network models. Comparing its prediction with the analytical models indicated the superior performance of machine learning in modelling the bond. Haddad and Haddad [79] used neural networks to study the change in bond strength with different geometrical and mechanical characteristics of the connection.

An integration of Fuzzy logic with neural networks creates a hybrid model that benefits from the advantages of both techniques. The new model is called the adaptive neuro-fuzzy inference system (ANFIS) and has been used to model the different characteristics of the connections in the FRP composite structures. Comparing its efficiency to other machine learning methods, Mansouri et al. [80] showed that ANFIS was able to outperform other models in predicting bond behaviour. Apart from bond strength, other properties of composite structures, such as beam deflections, debonding strain and stress, and compressive strength of the columns were also investigated by ANN-based models [81, 82].

These studies show that machine learning can be used successfully to model the bond behaviour of composite structures. This can avoid the time and cost associated with experimental testing and provide an effective method for identifying the critical factors affecting the FRP-concrete connections. However, macroscale experiments are often limited in capturing the underlying interactions of the composite structures with their service environments at smaller scales. Although numerical methods are an effective choice to model the constitutive behaviour of FRPs at the microscale, the computational power associated with these methods increases for complex geometries and loading scenarios. Machine learning is proposed as an efficient solution to model the multiscale behaviour of composite materials. In the following, the successful integration of this method with numerical approaches to predict composite properties is addressed.

The structure-property relationship in composites has encouraged researchers to investigate the impact of microstructure geometry on the material response. Making a mapping from geometrical information to the effective properties of composites has been addressed by deep learning methods in several studies. Since training an efficient prediction model requires a considerable amount of data, using image-based techniques to train machine learning models is costly and cumbersome. An alternative to this is to artificially generate microstructure images using mathematical tools.

Yang et al. [83] employed a convolutional neural network (CNN) together with principal component analysis (PCA) to identify the properties of binary composites using images of their microstructure. In doing so, CNN was trained by generating 100,000 artificial 2D microstructures and obtaining stress-strain curves using FE modelling. Comparing the CNN predictions with FE results showed less than 10% error. Yang et al. [84] used different architectures of CNN to automatically extract important geometric features from 3D RVEs. They artificially generated 8,550 microstructures to estimate stiffness in composite materials, with the hard phase volume fraction ranging from 25% to 75%. In this study, the rule of mixtures and linear regression methods were also used to predict stiffness. The comparison of the prediction accuracy indicated that CNN was a superior solution to extract features and map them to material properties.

In a similar study, Cecen et al. [85] generated a dataset of 3D microstructure with 5,900 voxels. They combined the features extracted by the CNN and the correlation functions to train a deep learning model and improve the elastic modulus predictions of two-phase composite materials. The result of this study also indicated the importance of computationally efficient predictions by machine learning compared to numerical approaches. Similar methodologies have also been used to address the structure-property links such as fluid permeability and ionic conductivity in heterogeneous materials including sandstone and ceramics [86, 87].

Besides deep learning, classical machine learning methods have been used to predict macroscale mechanical properties from composite microstructures in several studies. In these works, the microscale geometric features were extracted by statistical functions. Pathan et al. [88] and Ford et al. [89] trained various machine learning models such as gradient boosting, SVM, and linear regression to predict the effective elastic properties of composites with different shapes and volume fractions of inclusions. They used two-point correlation functions to quantitatively characterise the microstructure images. The obtained information was then used by a dimension reduction method to extract the most important features. The selected features were then used by the machine learning models to make predictions. In a similar study, neural networks were used with correlation functions to investigate the impact of fibre arrangements on the matrix damage initiation of 2D ceramic composites [90]. The comparison of the predictions in the mentioned studies indicates that classical machine learning can draw an accurate mapping from microstructure to effective properties as good as FE models while requiring less computing time and cost than deep learning methods.

These studies show that machine learning methods can successfully capture the behaviour of composite materials at various scales. The combination of machine learning and physics-based modelling can aid in the identification of patterns and correlations that numerical models cannot easily identify. This can lead to generalised predictive simulations that can be used to design composite materials with tailored properties [91].

2.7 Conclusions

It was seen that maintenance and strengthening are techniques that can be used to rehabilitate degraded infrastructures. Although helpful, their efficiency is compromised due to the lack of knowledge on the impact of ageing at smaller scales. As a result, understanding ageing mechanisms at multiple levels is required in order to develop targeted maintenance and strengthening strategies. Despite the abundance of research on the topic, there is not yet an efficient approach for capturing the interaction between infrastructures and service environments. Studies often focus on experimental and numerical approaches. Although the macroscale experiments provide information on the performance of infrastructures subjected to the surrounding environment at the structural level, they often miss the underlying interactions of the materials. Numerical approaches can be an alternative to accounting for these interactions at smaller scales. However, this comes at the cost of computational efficiency.

Machine learning can be introduced as an effective tool to address the current challenges in studying infrastructure ageing. By integrating the knowledge across different scales, machine learning can detect and predict the signs of degradation in infrastructure components. This can enable targeted maintenance and efficient strengthening design solutions. It is intended to show how machine learning can help analyse the impact of ageing on infrastructure at multiple scales throughout the chapters of this study. A number of limitations and gaps to overcome include:

- Analysis of ageing in infrastructures for maintenance purposes: Since infrastructure ageing causes faults and failures in assets, it is important to anticipate these defects before they occur. However, the current maintenance strategies are costly, either due to unnecessary maintenance in preventive practices or expensive repairs as a result of infrequent maintenance in reactive approaches. This can undermine optimal resource allocation and waste the useful life of assets in an infrastructure system. In this regard, while machine learning is known to efficiently forecast the potential failures of assets for predictive maintenance purposes [65], the resulting prediction models are often too complex to be used in practice. Moreover, the available solutions are not general, making them inapplicable to different types of infrastructure assets.
- Analysis of ageing in FRP strengthened structures: Although FRP composites are regarded as effective solutions for strengthening structures, their contributions can be compromised due to interactions with surrounding environments. Despite the fact that the ageing of composite structures has been widely addressed in experimental studies, there is a lack of universal agreement on the long-term performance of these structures. In addition, the existing models are highly empirical, and their predictive capabilities are limited to small databases, which undermines their practicality in proposing strengthening solutions for concrete structures. The proposed prediction models in the literature often miss the underlying interaction of the materials with the surrounding environment. A lack of understanding of performance at the material scale can create additional uncertainty about the long-term performance of composite structures.
- Analysis of ageing in composite materials: Numerical modelling based on computational micromechanics has been identified as an effective tool for studying FRP composite ageing at the material scale. Despite their high fidelity and precision, they require considerable computation effort and time to predict the evolution of degradation with time, which compromises their efficiency in the decision-making process for quick assessment and immediate action in real-time scenarios. Therefore, an efficient alternative is needed to

accelerate the solution of multiscale modelling of composite ageing due to interaction with the service environment.

CHAPTER 3

Assessment of An Ageing Railway Infrastructure Using Data Analytics and Machine Learning for Predictive Maintenance

3.1 Introduction

This chapter presents a machine learning method to analyse the impact of ageing on an infrastructure system by predicting its condition. It was seen in the literature that, although reactive methods are effective at the early stages, as the infrastructure gets older, they become less efficient. Also, in preventative approaches, unnecessary maintenance schedules can undermine the serviceability of the asset. On the other hand, it was seen that predictive methods increase the assets' useful life by providing maintenance before failure. In this chapter, we intend to develop a data-driven model that can predict upcoming defects in the railway system. A historical database on the maintenance work orders of a railway system is used to train machine learning classification and regression models. A comparative analysis is performed to nominate the most accurate model for analysing the state of different assets within the network and modelling the failure of the infrastructure. By establishing the likelihood of a defect before it occurs, maintenance can be scheduled at an optimal time, which can improve the serviceability of the railway network.

Among the various infrastructures, railways are of the highest importance as they are the key components for the efficient transportation of freight and passengers. Besides, the influence of the railway infrastructure on the economy is significant. For instance, a study of the Inland Rail business case in 2015 indicated that the cross-country railway system is expected to add up to $16 billion to Australia's gross domestic product (GDP) within the first 50 years of operation

[92]. Application of railways under severe environmental conditions with extreme traffic loads can accelerate the degradation of these structures over time. Regular asset maintenance therefore helps preserve the functionality of a railway system. Although regular maintenance tasks can increase the asset's lifetime, the resulting costs are normally significant. A report published by the Transport Asset Holding Entity (TAHE) [1] showed that more than $295 million was invested for regular maintenance of the rail network in 2021–2022. Besides being costly, regular asset maintenance and repair is correlated with infrastructure downtime, which can create delays and increase traffic in a transportation system.

The advent of machine learning and data-driven methodologies has made it feasible to employ these ideas in asset management to optimise the maintenance cost and the useful life of the assets [93]. By integrating these methods with maintenance activities, it is possible to predict the asset condition based on the information learned from its historical data.

In the current study, a data-driven model for predictive maintenance is established to predict the conditions of the assets in a railway infrastructure using machine learning techniques. The work in this chapter is part of a research internship project conducted a global consulting company and in collaboration with a rail network operator. A database on the historical maintenance work orders was provided by the industry partner. It is intended to develop suitable data pipelines using statistical and machine learning tools to predict and identify defects in the assets. First, data is described, and pre-processing tasks are employed to clean the data. Feature engineering techniques are then used to explore the database and identify important parameters. Finally, a series of machine learning models are developed to analyse defects and evaluate the influence of important characteristics on asset failures. The practicality of defect predictions is also demonstrated by machine learning-based equations.

3.2 Infrastructure maintenance overview

The creation of maintenance work orders for railway systems is normally completed by the plant maintenance module in the SAP Enterprise Resource Planning (ERP) software [94]. This module manages the maintenance activities by creating statements on the condition of the inspected asset. Depending on the condition, work orders are generated with response time frames based on the maintenance priority level. Figure 3-1 illustrates the framework for creating the maintenance work order. First, a notification on the asset condition is created by the inspection engineer. Then, depending on its status, the required work is ordered using asset details, such as location, type, and detected failure with priority level. The corresponding response time to the defect is set for the maintenance action, and the notification is updated in the final.

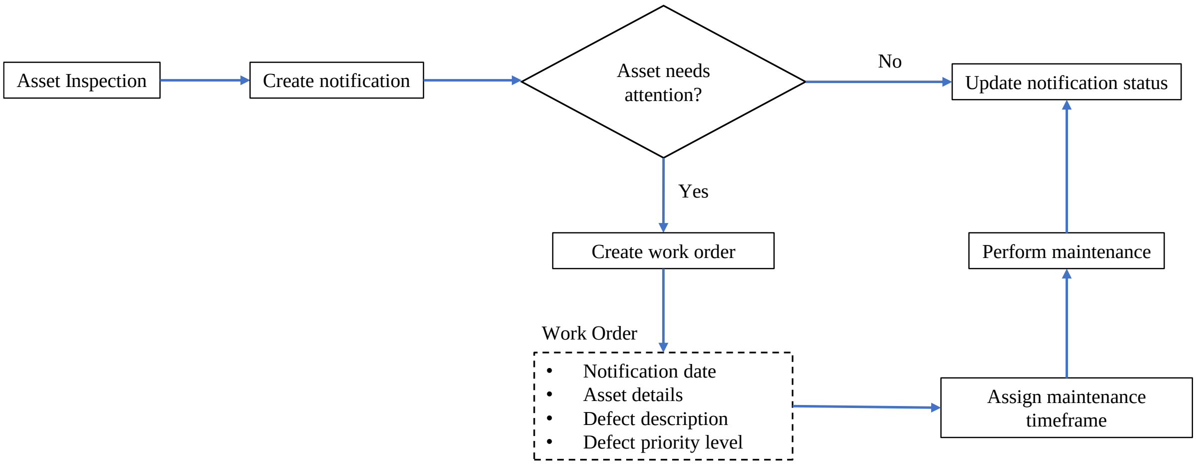


Figure 3-1. Framework for maintenance work order creation and action

The historical maintenance data covered the work orders from 2017 to 2021. A review was undertaken by subject matter experts in five phases to evaluate the data in detail and propose recommendations to the technical maintenance plan (TMP). Four different data sources were explored, as described in the following:

- Asset register data, which provided information on the asset type and functional location details. Track, turnout, and point machine (PM) are the assets present in the data. While the track is a continuous asset, the turnout and the PM are examples of discrete assets.
- Network utilisation data, consisting of information on the asset usage in terms of loadings and car counts.
- TMP service schedules, providing details on the turnout and PM assets in each functional location.
- Maintenance historical data, which includes details on the detected asset failures.

Figure 3-2 illustrates the schematic details of a turnout system with PM and track assets. Turnouts make it possible for a train system to switch between the tracks. Turnout is mainly composed of switch and closure rails as well as the crossing component with wing rails. It is seen that track A is divided into two tracks (B and C), and the PM is operated to switch the turnout position so that the train wheels are positioned towards either track B or C. The check rails are also used to ensure the safety of the train passing the crossing component.

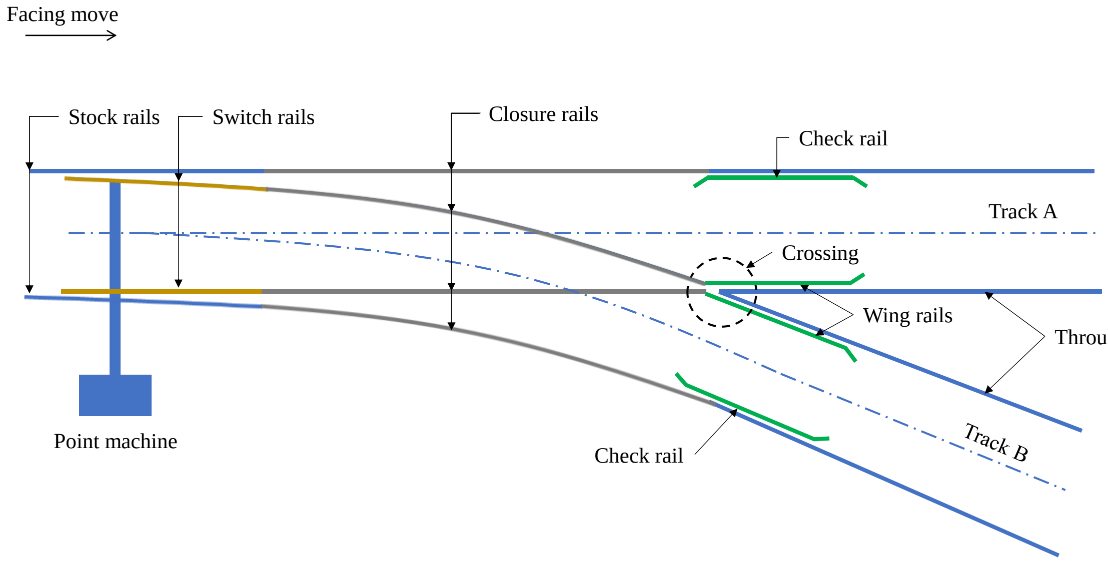


Figure 3-2. Schematic details of turnout system with PM and track assets

The review of the technical maintenance plans in five phases provided details on the characteristics of the historical work orders. The extent of failure rates in various assets suggests that current practices for maintaining the infrastructure assets are inefficient. Furthermore, the quality of the data indicates the necessity for performing an in-depth data analysis to understand the potential patterns among different fields. In this study, the capability of a predictive approach in evaluating railway condition is investigated. It is accomplished by pre-processing the historical maintenance data and developing prediction models using machine learning techniques.

3.3 Data description and pre-processing

A comprehensive database of available work orders for the asset maintenance was utilised for the research. A total of 2531 historical data points with the information on the asset location, usage condition, and defects were assembled. An initial data cleaning was required before proceeding to analysis. There were multiple fields with blank, missing, or duplicated values that were removed from the database. In some cases, features were renamed to improve understanding. Table 3-1 presents the details of the features available in the database. It is seen that 23 numerical parameters are presented with various units. For the three asset types in the data, these parameters quantify their location, quantity, and defect count. Also, the number and load of the yearly passing trains and carriages are available in the database.

Table 3-1. Details of numerical features in the database

Feature type	Feature name	Feature description	Unit
Numerical	Base code	Number to specify a railway track	–
	Start point	Inspection start point	km
	End point	Inspection end point	km
	Mid-point	Inspection middle point	km
	Car weight	Total weight of passing trains per year	Megaton
	Service count	Number of trains	–
	Car count	Number of carriages	–
	Turnout count	Number of inspected turnout assets	–
	PM count	Number of inspected PM assets	–
	Track length	Length of the inspected track	km
	Turnout defect count	Number of defects in inspected turnouts	–
	PM defect count	Number of defects in inspected PMs	–
	Track defect count	Number of defects in a length of inspected track	–
	Total defects	Total asset defect counts in inspection	–
	R_1 defect count	Asset defect count in Region 1	–
	R_2 defect count	Asset defect count in Region 2	–
	R_3 defect count	Asset defect count in Region 3	–
	R_4 defect count	Asset defect count in Region 4	–
	R_5 defect count	Asset defect count in Region 5	–
	R_6 defect count	Asset defect count in Region 6	–
	R_7 defect count	Asset defect count in Region 7	–
	R_8 defect count	Asset defect count in Region 8	–
	R_9 defect count	Asset defect count in Region 9	–
Categorical	Description	Railway line class	–
	Area code	Asset suburban location	–
	Territory	Asset territorial location	–

Table 3-1 shows the list of the categorical features as well. It is seen that three parameters in this table describe the railway class and the location of the assets. Since the machine learning-based prediction models in this study work with numerical variables, it is required to convert the labels of the categories to numbers. Therefore, the one-hot encoding method was used to transform each label to binary values [95]. It works by adding additional features that either have a 0 or 1 value for each category, where 1 indicates the presence of the categorical label Considering that description and area code are large in the number of categories, encoding these features will add a large number of unnecessary parameters to the data, increasing the computational cost of the analyses. Therefore, the one-hot method was only used to encode the territory variable, while the other categorical features were removed from the data.

For the territory field, which represents the asset locations in the database, it was found that many data points were not assigned any names. Also, a large number of assets were located outside of operator's jurisdiction. Figure 3-3a shows the distribution of assets across different territories. It is seen that 17% of the data lacks proper territory names. As a step in the data cleaning, the assets that do not belong to the operator were excluded from the database. For the assets with unknown locations, the jurisdiction map was used to identify the missing territories. The start point and end point features in the database for each work order reflect the inspection's beginning and ending distances from a nominal origin. As a result, for each asset without a location name, the values of these features were used to determine the suburb and, consequently, the territory of the work order on the map. Due to the confidentiality of the data, we were not able to share the map in this section. Figure 3-3b shows the revised territorial distributions after identifying the unknown locations of the assets. It is noted that R_8 and R_9 include 36% of all assets, while R_1 accounts for only 1% implying that assets are not evenly distributed amongst given regions.

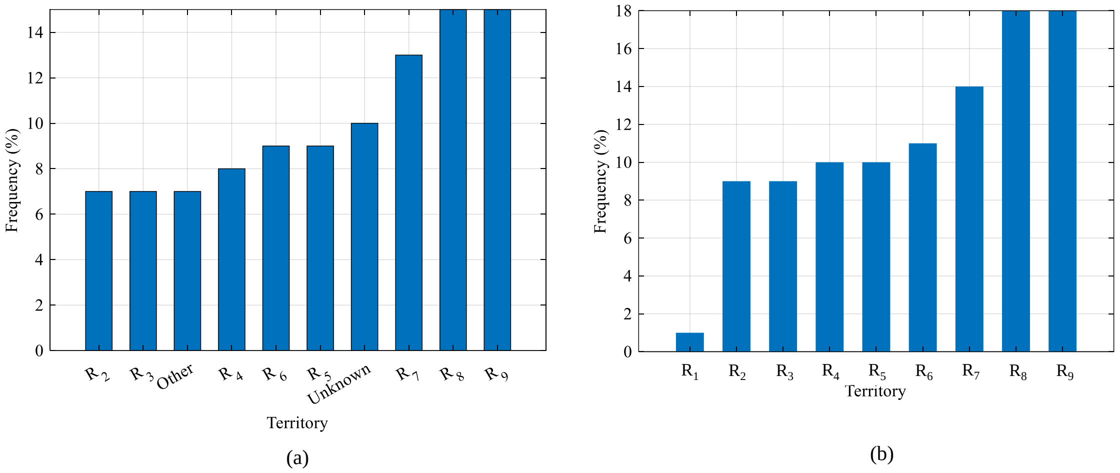


Figure 3-3. Distribution of assets across territories: (a) with missing information (b) after processing data

Since large scales variables can mask the effect of the small scale ones on the prediction model, it is important to rescale the features to a unit range [96]. In this work, the parameters in Table 3-1 were standardised to have a zero mean and unit variance values by:

$$X' = \frac{X - \mu}{\sigma} \tag{3-1}$$

where X' is the scaled variable; $X, \mu,$ and σ denote the feature with the mean and standard deviation values, respectively.

The data processing provided 2304 data points with 23 numerical and 1 categorical feature. The cumulative values of asset characteristic indicated the considerable asset utilisation as well as the great number of defects in the assets. Due to the sensitivity of the information, these values were not presented in this chapter.

Besides data size, the efficiency of a prediction model is also impacted by the feature space in the data [97]. In order to better understand the pattern among the data and select suitable variables for machine learning applications, it is important to investigate the inter-feature relationships as well. It was done by performing a feature engineering procedure on the data.

3.4 Machine learning techniques for modelling ageing infrastructure

In order to study the ageing of the assets by the processed data, a series of supervised machine learning algorithms were explored and used to model the asset defects. Two types of data-driven analyses were performed in this study: regression and classification. Regression technique was used to predict the continuous outputs, while classification was used for categorical outputs. Depending on the complexity of the relation between input and output variables, machine learning methods were divided into parametric and non-parametric approaches. Parametric methods make simplification assumptions during the learning process; however, non-parametric techniques do not assign any assumptions during prediction. Since the complexity of the relationship among the input and output features were unknown in this study, a series of regression and classification approaches were evaluated to realise the models with the highest prediction performance.

3.4.1 Linear regression and classification methods

Linear methods are a commonly used machine learning technique that assumes a linear relationship between the input and the output variables. Linear classification and regression models create the following output from a training data point $(\boldsymbol{x}_i, y_i)$, where $\boldsymbol{x}_i$ is the input parameter vector and y_i is the output label:

$$f(\boldsymbol{x}) = b + \boldsymbol{\omega}^T \phi(\boldsymbol{x}) \tag{3-2}$$

In Eq. (3-2), b is bias and $\boldsymbol{\omega}$ is the weight vector. In a linear model, $\phi(\boldsymbol{x}) = \boldsymbol{x}$, and the algorithm involves minimising the error between labels and classified values by:

$$\underset{\boldsymbol{\omega}}{argmin} \sum_{i=1}^{N} \xi(\boldsymbol{\omega}; \boldsymbol{x}_i, y_i) \tag{3-3}$$

where ξ is the loss function. The loss functions for linear classification (LC) and regression (LR) models are commonly written as logistic regression and least squares errors, respectively, as in Eqs. (3-4)–(3-5).

$$\xi(\boldsymbol{\omega}; \boldsymbol{x}, y)_{LC} = \log\left(1 + e^{-yf(\boldsymbol{x})}\right) \tag{3-4}$$

$$\xi(\boldsymbol{\omega}; \boldsymbol{x}, y)_{LR} = \frac{1}{2}(y - f(\boldsymbol{x}))^2 \tag{3-5}$$

3.4.2 Ridge and lasso regression

The regression coefficients estimated by Eq. (3-3) are unbiased and suffer from high variance. However, shrinking these coefficients towards zero through applying penalty on the cost function can reduce their variance [98]. Ridge and lasso are two popular machine learning techniques for shrinking the coefficient estimates. The ridge coefficients are derived by minimising a penalised cost function as:

$$\beta_{Ridge} = \underset{\beta}{argmin} \sum_{j=1}^{N} \left(y_j - \beta_0 - \sum_{i=1}^{M} X_{ji}\beta_i \right)^2 + \lambda \sum_{i=1}^{M} {\beta_i}^2 \qquad (3\text{-}6)$$

where λ is a hyperparameter that controls the shrinkage of the regression coefficients. As the value of λ increases, the ridge regression coefficients approach to zero.

Unlike ridge regression, where the model includes all input variables, lasso performs variable selection by imposing some of the coefficients to be equal to zero. The feature selection in lasso enhances the interpretability of the model in datasets with numerous variables. The lasso coefficients are derived as:

$$\beta_{Lasso} = \underset{\beta}{argmin} \sum_{j=1}^{N} \left(y_j - \beta_0 - \sum_{i=1}^{M} X_{ji}\beta_i \right)^2 + \lambda \sum_{i=1}^{M} |\beta_i| \qquad (3\text{-}7)$$

3.4.3 Decision trees

Classification and regression trees (CART), referred to as decision trees, are a set of machine learning methods that develop prediction models by dividing the feature space into high-dimensional rectangular regions, known as leaf nodes. By partitioning the variable space into D regions, $R_1, \dots, R_D$, the regression model can be written as:

$$f(x) = \sum_{d=1}^{D} c_d I\,(x \in R_d) \qquad (3\text{-}8)$$

where c_d is the mean value for the observations in the region R_d. The prediction performance of a decision tree can be tuned by the hyperparameters related to the tree size, such as the minimum number of observations per leaf node (Min_Leaf_Size) and the maximum number of decision nodes in the decision tree (Max_Num_Splits).

The linear model trees offer additional learning systems to the decision trees by fitting linear regression models to the observations in the leaf nodes. M5P, as an example of a linear model tree, uses a more computationally efficient approach to build a tree compared to CART. This method uses the standard deviation of the data reached to each node as a criterion for splitting the feature space. After building the tree, a line is fitted to predict the observations of each leaf node. Finally, a pruning technique is used to reduce the complexity and overfitting in the built tree.

3.4.4 Support vector machines

The mathematical representation of $\boldsymbol{\omega}^T\phi(\boldsymbol{x})$ in Eq. (3-2) can be cumbersome in nonlinear classification and regression tasks. Thus, kernel methods are proposed to map data into high-dimensional space. Support vector machine (SVM) is a popular machine learning method extended from the maximal margin classifier. The goal of support vector is to find a hyperplane equation, which either separates the binary classes or predicts the continuous outputs with minimum error. It is addressed by solving the L^1-norm problem as Eq. (3-9) for classification and Eq. (3-10) for regression SVMs, respectively.

$$\underset{\boldsymbol{\omega}}{argmin}\ \frac{1}{2}\boldsymbol{\omega}'\boldsymbol{\omega} + C\sum_{j=1}^{N}(\xi_j)\,;\ \begin{cases} y_j f(\boldsymbol{x}_j) \geq 1 - \xi_j \\ \xi_j > 0 \end{cases} \tag{3-9}$$

$$\underset{\boldsymbol{\omega}}{argmin}\ \frac{1}{2}\boldsymbol{\omega}'\boldsymbol{\omega} + C\sum_{j=1}^{N}(\xi_j + {\xi_j}^*)\,;\ \begin{cases} \forall j: y_j - f(\boldsymbol{x}_j) \leq \varepsilon + \xi_j \\ \forall j: f(\boldsymbol{x}_j) - y_j \leq \varepsilon + {\xi_j}^* \\ \forall j: {\xi_j}^* \geq 0 \\ \forall j: \xi_j \geq 0 \end{cases} \tag{3-10}$$

where ξ_j and ${\xi^*}_j$ are slack variables, and C is the box constraint that controls the penalty imposed to predictions with deviations from the observations larger than ε. By solving Eq. (3-9) in a Lagrangian dual formulation, the SVM output can be obtained for binary classes as:

$$\boldsymbol{x} \rightarrow sgn\left(\left[\sum_{j=1}^{N} \alpha_j y_j K(\boldsymbol{x}_j, \boldsymbol{x})\right]\right) \tag{3-11}$$

where sgn is a sign function indicating the belonging of input $\boldsymbol{x}$ to classes ± 1.

Similarly, the solution of Eq. (3-10) obtains:

$$f(\boldsymbol{x}) = \sum_{j=1}^{N} (\alpha_j - {\alpha_j}^*) K(\boldsymbol{x}_j, \boldsymbol{x}) + b \tag{3-12}$$

where α_i and $\alpha_i{}^*$ are Lagrangian multipliers, and $K(x_i, x)$ is called the kernel function. The above formulations show that the SVM prediction can be tuned by the values of C, ε, and the type of kernel function.

3.4.5 Ensemble methods

In classification and regression tasks, ensemble learning stands for approaches that integrate a number of models to make predictions. The learning process starts with generating a number of models, followed by a pruning step to remove redundant learners. Finally, it ends up with combining the finalised models with superior prediction performance. In this research, different boosting and bagging methods were used to combine decision tree learners to classify and predict the study outputs. Bagging involves sampling the training data via the bootstrap method where the data points are uniformly sampled to create subsamples of the predictors for training the learning algorithms [99]. This method is generally used to reduce the variance in decision tree algorithms. The ensemble prediction is equal to the average response of all the tree models as:

$$P_b(x) = \frac{1}{B}\sum_{b=1}^{B} f_b(x) \tag{3-13}$$

where B denotes the number of classifiers, and $f_b(x)$ is the prediction of a single decision tree.

The idea behind boosting method is to transform a weak learning algorithm into a high-accuracy prediction model. Initially, the data is equally weighted and trained with the first model. In each successive step, the incorrectly predicted observations gain more weight, and the correctly predicted samples decrease in weight. The next prediction model is then trained on the data with updated weights.

To train an ensemble of classification learners, bagging method along with a series of boosting techniques such as adaptive boosting (AdaBoost), gentle adaptive boosting (GentleBoost), adaptive logistic regression (LogitBoost), and random under-sampling boosting (RUSBoost) were used in this research [100-102]. For the regression problem, an ensemble of tree learners was aggregated by bagging and the least-square boosting (LSBoost) methods.

3.4.6 Artificial neural networks

Apart from classical machine learning models, artificial neural networks (ANNs) hold a multilayer structure which is loosely inspired by the interconnection of the neurons in the human brain. It consists of input, hidden, and output layers and performs two stages of feedforward and backpropagation to make predictions. During the learning process, the information passes through the layers to compute the output of the network in a feedforward manner. For a fully connected network with L layers, the output of the network is calculated in forward propagation. For a neuron k in the lth layer, the output ${a_k}^{(l)}$ is written as:

$$ {a_k}^{(l)} = \sigma\left(\sum_{m=0}^{N_{l-1}} {\theta_{km}}^{(l-1)} {a_m}^{(l-1)}\right) \tag{3-14} $$

where $\sigma(\bullet)$ is the activation function; N_{l-1} denotes the number of neurons in layer $l-1$; θ_{km} is the weight between the mth node in layer $l-1$ and the kth node in layer l. To measure the fitting error between the predicted and the target outputs, the cost function is written as:

$$ J(\theta) = \sum_{j=1}^{N} (y_j - a_1^{(L)})^2 \tag{3-15} $$

where $a_1^{(L)}$ denotes the output of the neuron in the output layer. To minimise $J(\theta)$, a backpropagation approach is utilised, and the gradient of the cost function with respect to its weights is computed. During the ANN training, the obtained gradients are used by different optimisation algorithms to update the weights until the error between the predicted and the target values reaches its minimum value. A successful application of ANN depends on the suitable choice of hyperparameters. The number of hidden layers and neurons together with the dropout value, which is the rate of randomly ignored neurons during the training process to avoid overfitting, describe the ANN architecture. Besides, the learning rate controls the step size in updating the neuron weights during the training process.

3.4.7 Naïve Bayes

Naïve Bayes is a probabilistic classifier based on the Bayes theorem which assumes that the input parameters perform independently in predicting a class. For a given class c and the input variables $x_1, \dots, x_M$, the Bayes theorem is written as:

$$ P(y = c|x_1, \cdots, x_M) = \frac{P(y = c)P(x_1, \dots, x_M|y = c)}{P(x_1, \dots, x_M)} \tag{3-16} $$

Using Bayes theorem assumption, the conditional probability of input variables for a given class c, $P(x_1, \dots, x_M | y = c)$ is written separately for each variable. Thus, Eq. (3-16) is formulated as:

$$P(y = c | x_1, \cdots, x_M) = \frac{P(y = c) \prod_{i=1}^{M} P(x_i | y = c)}{P(x_1, \dots, x_M)} \tag{3-17}$$

Since $P(x_1, \dots, x_M)$ is constant for all classes, the predicted output belongs to the class that maximises the nominator of Eq. (3-17) as:

$$f(x) = \underset{c}{argmax} P(y = c) \prod_{i=1}^{M} P(x_i | y = c) \tag{3-18}$$

The accuracy of naïve Bayes can be tuned by the assumption made regarding the distribution of the variables x_i over the class c.

3.4.8 K-nearest neighbours

K-nearest neighbours (KNN) is an instance-based machine learning method that identifies the K number of data points nearest to the observation x_j and predicts the conditional probability of x_j belonging to a class c as:

$$P(y = c | x = x_i) = \frac{1}{K} \sum_{j \in N_K} I(y_j = c) \tag{3-19}$$

where y_j and N_K are the class and the number of points closest to x_i, respectively.

3.5 Establishment of machine learning models

To apply the aforementioned machine learning methods, the dataset first needed to be split into training and test parts. The training subset was used to create prediction models, and the efficiency of the models was evaluated by the test set. The predictive performance of the machine learning models highly depends on the hyperparameter values. Therefore, it is necessary to explore the best combinations of hyperparameters to minimise the prediction error. In this study, Bayesian optimisation was utilised to automatically find the hyperparameters and avoid the time-consuming process of tuning with trial-and-error. Bayesian optimisation works based on Bayes theorem and searches for the optimal values of an objective function, $e(\theta)$. In this technique, optimisation initiates by randomly selecting points in the search space as $D_{1:t} =$

$\{(\theta_1, e(\theta_1)), \dots, (\theta_t, e(\theta_t))\}$. With the assumption of Gaussian process (GP) for the prior probability distribution over e, the posterior predictive distribution can be written as [103]:

$$P(e|D_{1:t}, \theta_t) = \mathcal{N}(\mu(\theta_t), \sigma^2(\theta_t)) \tag{3-20}$$

where $P(e|D_{1:t}, \theta_t)$ is called a surrogate function; $\mathcal{N}(\bullet)$ denotes the Gaussian distribution, and $\mu(\theta_t)$ and $\sigma^2(\theta_t)$ are the predictive mean and variance functions. A critical step of Bayesian optimisation is to find the next point during the search for the optimum of the objective function. The next point is obtained by maximising an acquisition function as in Eq. (3-21).

$$\theta_{t+i+1} = \underset{\theta}{argmax} a(\theta|D_{1:t+i}) \tag{3-21}$$

In the above equation $a(\bullet)$ is the acquisition function, and i denotes the number of the points which have already been found by maximising the acquisition function. An acquisition function is defined so that the point corresponding to its maximum is expected to provide the best-observed value for the objective function among the points found so far by Eq. (3-21). θ_{t+i+1} is used to augment the data set as $D_{1:t+i+1} = \{D_{1:t+i}, (\theta_{i+1}, e(\theta_{i+1}))\}$ and update the surrogate function for the next point exploration. Sampling continues until the predetermined threshold of maximum iterations is reached. Figure 3-4 illustrates the Bayesian optimisation flow chart used for asset defect prediction.

80% of the data was selected randomly and used for model training. A 10-fold CV technique was used to resample the training data for hyperparameter tuning. During model tuning, Bayesian optimisation iteratively searched for the hyperparameters. At each iteration, the prediction models were tuned by the explored hyperparameters, and their CV errors were obtained. Finally, the model with the minimum error was selected, and its efficiency was evaluated by the remaining 20% of data as the test set. For the failure mode identification, a stratified 10-fold CV method was used to avoid imbalanced classification. During the model tuning, the hyperparameters were explored by Bayesian optimisation, and the optimised machine learning model with minimum CV error was selected. At the end of optimisation, the efficiency of the selected model was evaluated by the out-of-fold predictors. In this study, MATLAB software was used as the programming language throughout the prediction model development steps including data pre-processing stage, machine learning and Bayesian optimisation algorithms development, model training, and the performance evaluation.

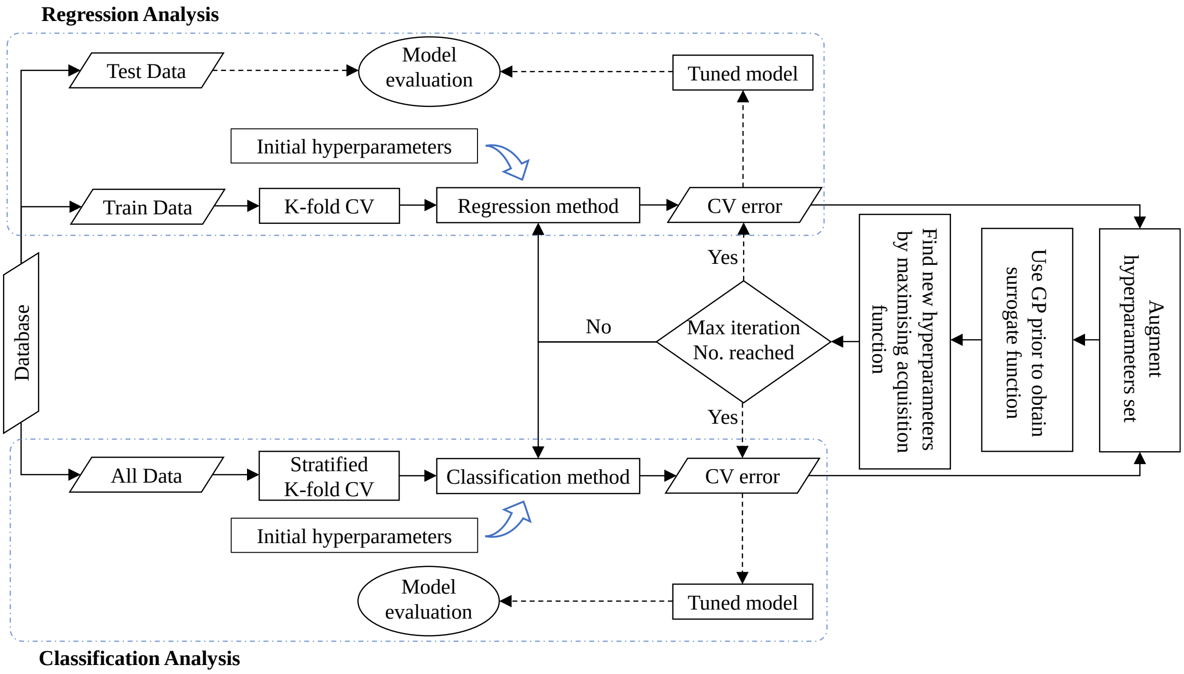


Figure 3-4. Bayesian optimisation flow chart for asset failure prediction

3.6 Feature engineering

In this section, the data is explored for any potential patterns among the features. A series of statistical and machine learning tools were developed to investigate data at different levels of detail. This assists in finding the correlation between the input features and the output variable. Furthermore, a feature selection procedure is employed to nominate a subset of parameters for use in the machine learning model.

3.6.1 Data exploration

The infrastructure maintenance overview in Section 3.2 addressed the details of the faults, indicating the necessity for studying the defects in the assets. In order to evaluate the relationships between the features and the asset defects, the scatter plots of various parameters with the total defects feature were created to explore potential correlations. However, due to the sensitivity of data, these plots were not presented in this work.

Scatter plots showed that no correlation can be drawn among the data points, implying that analysing data at the present level of detail does not map meaningful patterns between the features. Analysing the distribution of defect counts in three asset types, namely, track, turnout, and PM, in relation to the total defect numbers revealed that the total defects value grows as the number of failures in each asset type increases. Evaluating how the total number of faults changes in relation to the defect counts across territories, indicated correlations between the features. The good relationship between the features representing the defect count is expected, as for each data point, the sums of the fault numbers across the assets as well as the territories are equal to the total defects count.

In order to investigate the feature relationships in detail, the work orders were filtered based on the territories so that the asset characteristics were quantified across different regions in the data. The distribution of different features across the territories showed that the assets in R_8 experienced the highest numbers of passing cars and loadings. This resulted in the highest number of defects in this territory. It was also obtained that R_1 owns the least number of assets, making it the least used territory with the fewest defects. A graphical visualisation on the distribution of different features across territories was not added to the chapter due to the sensitivity of the data.

3.6.2 Feature selection

As a part of feature engineering, the selection of a subset of parameters with high relevance to the target output is important in order to improve model accuracy and remove overfitting [104]. Feature selection methods can help identify and remove the non-informative parameters in the data that do not contribute to the model's efficiency. This is also desirable for reducing the complexity of the prediction model and increasing the training speed. Among various methods, intrinsic techniques select features as a part of model training. In these cases, the importance of the features after model training is used as a metric to score the relevance of input parameters to the target variable. Besides, feature importance can provide insight into the data by exploring the relationship between the predictors and the output parameter. Estimation of feature importance by permutation is a technique to demonstrate the influence of a model parameter in predicting the output variable. The drop in model accuracy caused by permuting the values of an input feature is a metric to measure its importance [105]. For an influential feature, permuting its values greatly decreases the model's accuracy.

Analysing data by the scatterplots showed us that no meaningful patterns were apparent among the datapoints. Considering the ability of ensembles to handle noisy data, a series of ensemble models were trained using the data scaled by Eq. (3-1) to rank the importance of various asset properties in relation to the defects.

The steps for ranking features by permutation are explained in the following.

1. For each tree model t in the ensemble model:
    a. The model error ε_t from out-of-bag observations was estimated.
    b. For each input parameter X_i, the data points were permuted.
        i. The model error from the permuted values of the out-of-bag observations ε_{ti} was estimated.
        ii. The difference between out-of-bag error before and after permutation $\mathrm{E}_{ti} = \varepsilon_{ti} - \varepsilon_t$ was obtained.
2. The mean μ_i and standard deviation σ_i of the difference over all trees was obtained for each input parameter.
3. The out-of-bag permutation importance for each feature was obtained as μ_i/σ_i.

Figure 3-5 illustrates the relative importance of various asset properties to the defect types. For each parameter on the horizontal axis, four bars of different colours show the importance of the feature for different asset defects. The blue, red, orange, and purple bars identify the importance of the turnout, track, PM, and total defect numbers, respectively.

It is seen in Figure 3-5 that for track length, the highest importance is assigned to the track defects. For the turnout counts, it is obvious that the turnout defect received the highest influence from this feature. Similarly, the PM count has a greater impact on PM defects than other types of asset defects. This implies that as the number of assets in the network increases, the probability of asset failure grows as well. Furthermore, it is shown in Figure 3-5 that the weight and number of passing cars create the most defects in the track among other asset types. This is due to the fact that railway tracks are subjected to a variety of vertical loadings as well as lateral and creep forces. As the wheel load grows, the contact stress between the rail and wheel increases, resulting in more defects in these assets compared to turnouts and PMs.

The features were chosen by ranking their importance to the total defects using the purple bars in Figure 3-5. It is seen that the total asset defects are greatly governed by the number of PMs and track length in the network. It is also seen that local properties, including the inspection locations and the asset territories, have the lowest influence on the defects, with the mid-point and end point being the least important features in the dataset.

The correlation coefficients among the numerical features are presented by a heatmap in Figure 3-6. It is clear that there is a strong correlation between asset local properties and the asset utilisation characteristics. Specifically, the correlation value among the start point, end point, and mid-point variables is 1. Also, car count is correlated to car weight and service count parameters at 0.99 and 0.95, respectively. High correlation values imply multicollinearity between the features, which can cause problems in determining the unique effect of each input parameter on the target variable. Therefore, these features were removed from the data.

Furthermore, the features on the defect numbers across different asset types and territories in the database are considered as redundant because they are non-informative and provide the same information as the total defects feature. Considering that redundant features can cause overfitting and increase the model's complexity, they were also removed from the database [106]. The nominated parameters for the further analysis of the asset defects are listed in Table 3-2.

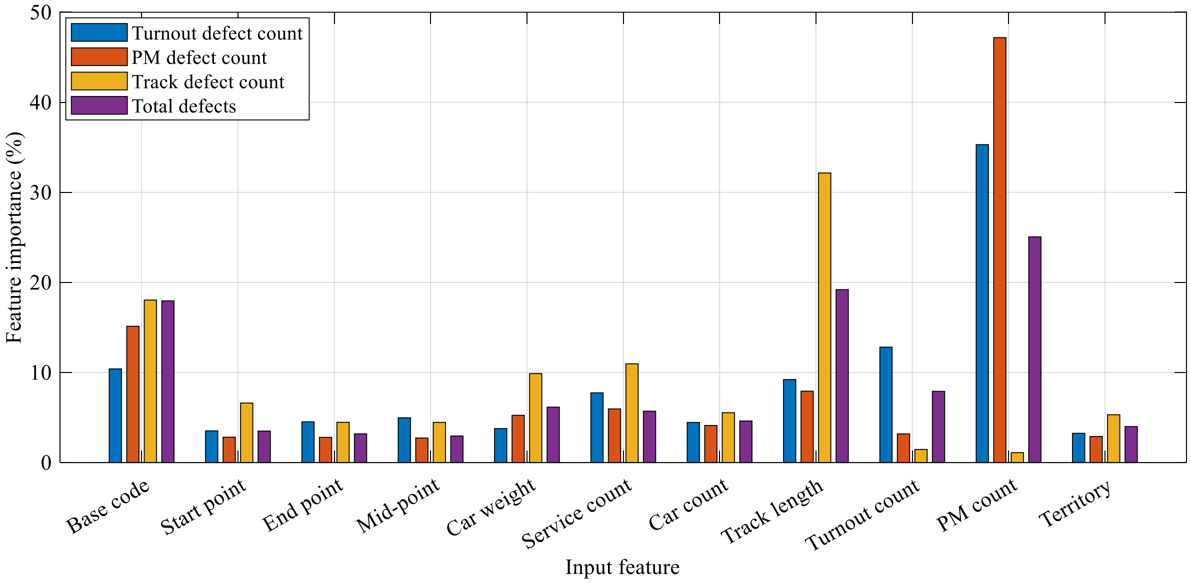


Figure 3-5. Feature importance in relation to defect count in various asset types

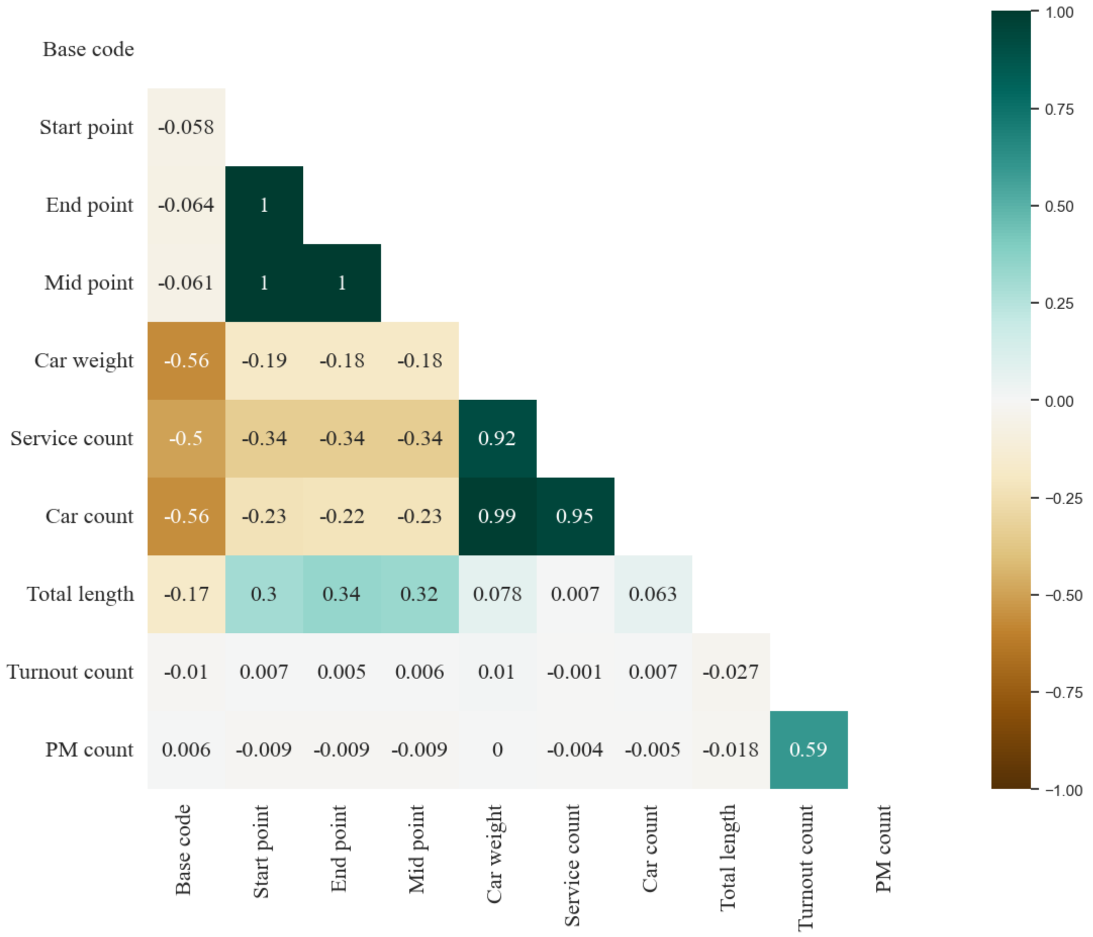


Figure 3-6. Triangle correlation heatmap between numerical features

Table 3-2. The selected features for defect prediction

Feature	Feature type	Feature name
Input parameter	Numerical	Base code
		Start point
		Car weight
		Service count
		Track length
		Turnout count
		PM count
	Categorical	Territory
Output parameter	Numerical	Total defects

3.7 Asset defect prediction

A defect in the assets can refer to a number of different factors, such as ageing, extreme environmental conditions, and lack of maintenance. They can result in service disruptions, delays, and cancellations, impacting the overall efficiency of the transport system. The ageing of assets can cause a series of defects across the network. A railway track comprises several different components working together to support the movement of the trains. Some of the main parts of track are rails, sleepers, ballast, fasteners, and subgrade. Figure 3-7 shows a cross-section view of a railway track with its components. Defects in these parts can undermine the serviceability of the asset. Some of the failures that occur are wheel burn, squat, and fatigue in the rail, poor ballast condition, fastening wear and tear, and sleeper fractures.

Common failures observed in turnouts are wear, fatigue, and plastic deformation occurring in the crossing and switch rails [107]. Point machines are electric devices used to control the switching between the rails. Failure in these mechanical systems is generally caused by ageing, environmental conditions, misalignment, and obstruction [108].

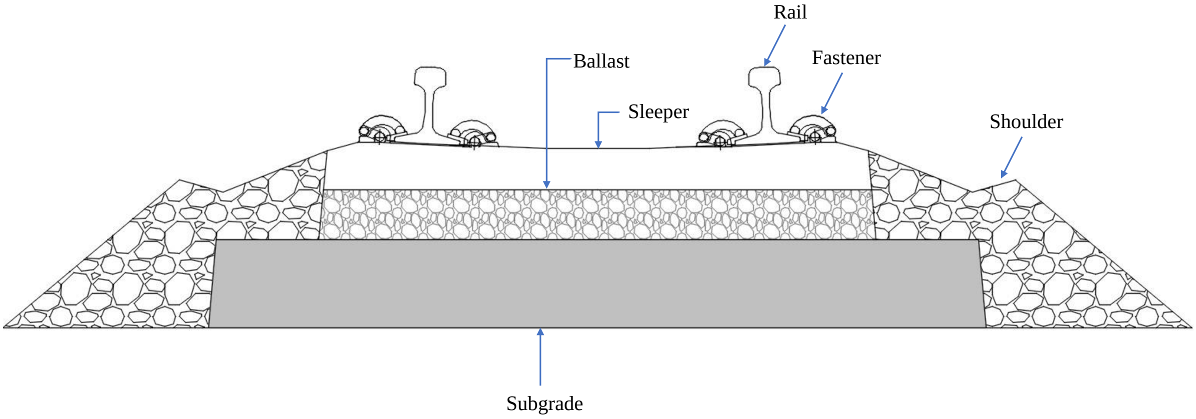


Figure 3-7. Cross-section view of a railway track

Predicting the defects in the assets is critical for ensuring the safety, reliability, and efficiency of the railway operations, as well as reducing the cost and environmental impact of asset faults. In order to identify defects using the information from the historical maintenance work orders, a two-stage classification technique was used in this study. In the first step, a binary classifier filtered the data by separating it into two positive and negative classes. The positive data was then passed to the second stage, where a multi-class classifier was trained to identify the class of the filtered data. A two-stage classifier can improve the efficiency of the prediction model by reducing the number of data points that the multi-class classifier needs to process [109].

Figure 3-8 shows the flowchart for applying the proposed classifier to the processed dataset using the selected features in Table 3-2. The first model was trained to identify the presence or absence of defects in the asset. The portion of the data with positive defects was then used in the second classifier to identify the type of faulty assets.

The classification performance of machine learning models was measured by precision, recall, and accuracy metrices. Precision measures the ratio of the correctly predicted positive instances out of all instances predicted as positive. It is calculated by the following equation.

$$Percision = \frac{True\ positives}{True\ positives + False\ positives} \tag{3-22}$$

Recall measures the ratio of correctly predicted positive instances out of all actual positive instances in the dataset. It is calculated as:

$$Percision = \frac{True\ positives}{True\ positives + False\ negatives} \tag{3-23}$$

Accuracy measures the overall performance of a classification model as a percentage of correctly classified instances. It is calculated as:

$$Percision = \frac{True\ positives + True\ negatives}{True\ positives + True\ negatives + False\ positives + False\ negatives} \tag{3-24}$$

In Eqs. (3-22)–(3-24), true positives and true negatives indicate the number of instances predicted as positive and negative, respectively. Similarly, false positives and false negatives denote the number of observations incorrectly predicted as positive and negative, respectively.

In order to identify the defective assets, a classification analysis was performed by training five machine learning models, decision tree, naïve Bayes, KNN, SVM, and ensemble. The data were sampled by 10-fold CV to train and tune the models' hyperparameters. The list of hyperparameters with their optimal values presented in Table 3-3, which were obtained after 30 steps of tuning by Bayesian optimisation for each model. Figure 3-9 shows the out-of-fold prediction results of the first-stage classifiers. Out-of-fold prediction refers to the predictions made by the model based on the observations from the validation set during each iteration of the CV process. It means that the model, which was trained on the training folds, is used to predict the type of faulty asset for the instances in the validation fold that were excluded during training. The values in the green and red cells of the confusion matrices in Figure 3-9 indicate the numbers of correctly classified and misclassified observations, respectively. Comparing the classification performance of the models shows that ensemble outperforms the remaining classifiers in identifying the assets' conditions with the accuracy of 90%. Also, a precision value of 91.6% in Figure 3-9e indicates the precision of the ensemble in detecting the faulty assets.

Considering the high prediction accuracy of ensemble model, it was selected for the second stage of classification. At this stage, a portion of the data identified as faulty asset by the first ensemble model were selected to training at the second stage. It was seen in the selected data that 11%, 4%, and 3% of the defective assets were turnout, track, and PM, respectively. Also, in 82% of inspections, more than one type of asset was found faulty. The selected data was imbalanced since the number of defective assets for track and PM was small in the dataset. A 10-fold stratify CV method was employed to ensure that the observations for the minor classes were present in all folds [110]. Furthermore, in order to balance the data, the majority classes were undersampled using a method called Tomek Links. This method is based on the idea that removing examples from majority classes that are closest to the minority category can improve the prediction performance of the classifier [111].

The resulting dataset comprised 1669 maintenance records. The out-of-fold prediction details of the second-stage classifier is shown in Figure 3-10. It is apparent in this figure that the model can classify the faulty asset types with 94.9% accuracy. The model's precision in identifying failure in more than one type of asset is 95.6%. For the PM, turnout, and track assets, these numbers are 100%, 86.2%, and 100%, respectively. This implies that the model predictions of the defective system are in good agreement with the observations. In the confusion matrix, the recall metric for a specific class describes the number of observations from that class that are predicted correctly. The values of this metric in Figure 3-10 also indicate the model was able to correctly identify most types of defective assets. However, the low recall values for PM and track are related to the fact that these asset types only comprise 7% of the data, making it difficult for the classifier to identify them.

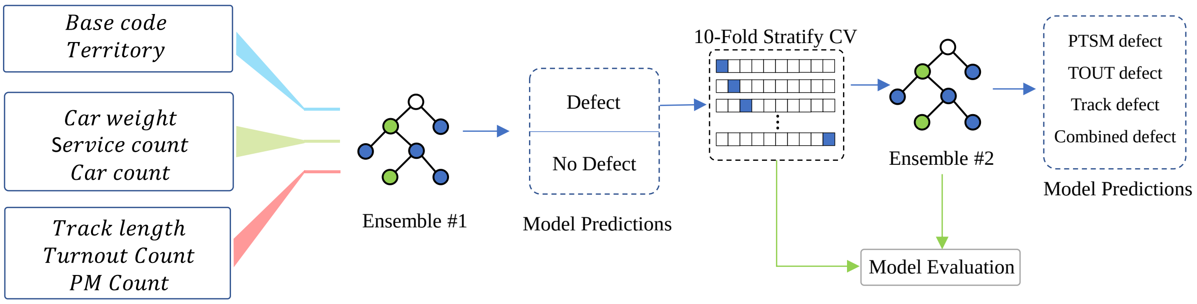
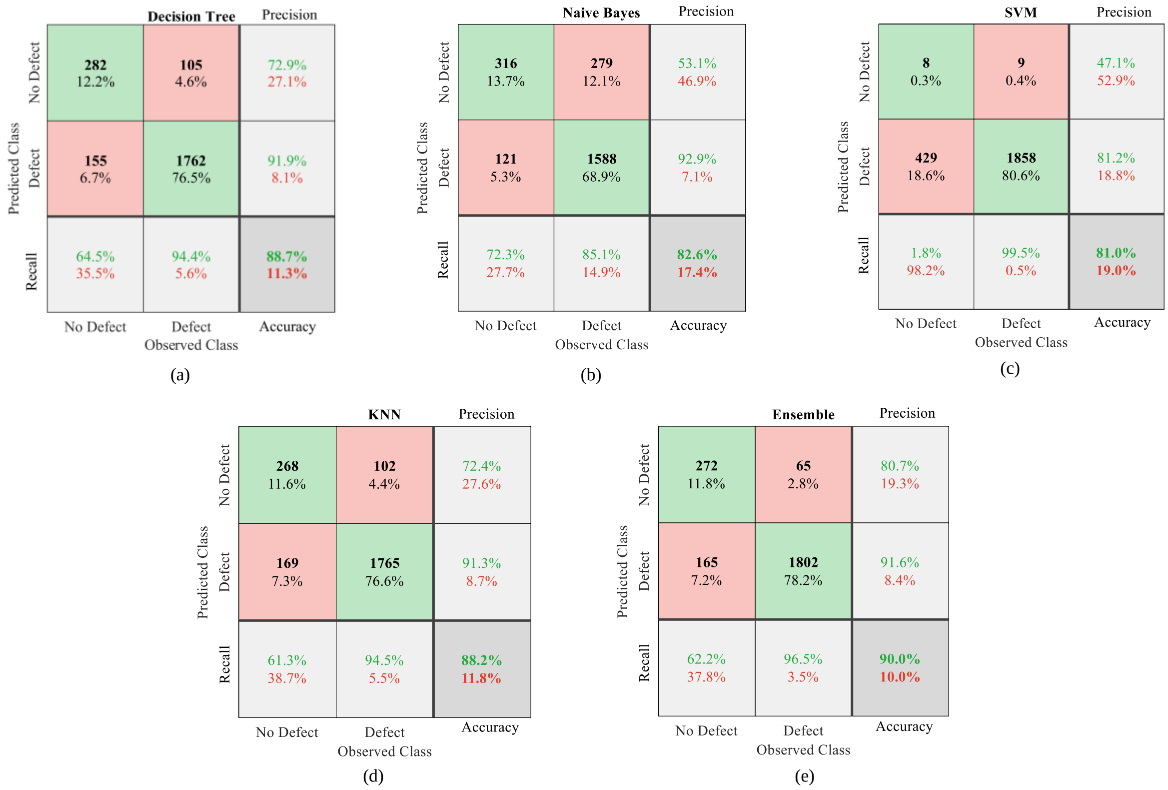


Figure 3-8. Two-stage classification flowchart

Table 3-3. List of selected hyperparameters with optimal values for first-stage classifiers

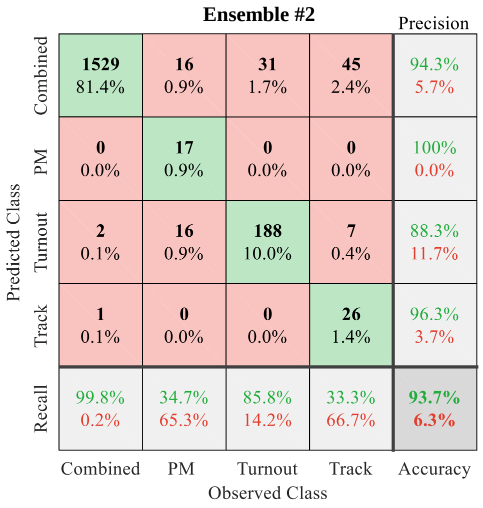
Model	Hyperparameter	Search range	Optimal value
Decision Tree	Min_Leaf_Size	1-1152	25
	Max_Num_Splits	1-2303	1341
Naïve Bayes	Feature distribution	Normal, Kernel	Kernel
SVM	C	0.001-1000	109.28
	Kernel function	Gaussian, Linear, Polynomial	Gaussian
KNN	K	1-215	9
Ensemble	Method	Bagging, AdaBoost, LogitBoost	LogitBoost
	Num_Learning_Cycles	10-500	498



(a) (b) (c)

(d) (e)

Figure 3-9. Confusion matrices of the first-stage classifiers using out-of-fold predictions



Predicted Class \ Observed Class	Combined	PM	Turnout	Track	Precision
Combined	**1529** 81.4%	**16** 0.9%	**31** 1.7%	**45** 2.4%	94.3% 5.7%
PM	**0** 0.0%	**17** 0.9%	**0** 0.0%	**0** 0.0%	100% 0.0%
Turnout	**2** 0.1%	**16** 0.9%	**188** 10.0%	**7** 0.4%	88.3% 11.7%
Track	**1** 0.1%	**0** 0.0%	**0** 0.0%	**26** 1.4%	96.3% 3.7%
Recall	99.8% 0.2%	34.7% 65.3%	85.8% 14.2%	33.3% 66.7%	**93.7%** **6.3%**

Accuracy

Figure 3-10. Confusion matrix of second-stage ensemble model

3.8 Influence of asset characteristics on the infrastructure defects

A parametric study of the asset defects provides valuable information that can help make informed decisions about the performance of the assets, resulting in increased efficiency, safety, and reliability of the railway system. In addition, understanding the change in the defects can be used to predict the performance of the assets under different conditions, which can help the railway system plan for future maintenance work orders. To study the influence of the asset properties, a regression analysis was performed by developing machine learning models, namely, linear, decision tree, SVM, and ANN. A 10-fold CV technique was used to sample the training data for tuning the models to optimally predict the total asset defects. Table 3-4 presents the list of hyperparameters and their optimal values obtained by Bayesian optimisation for each model. It is seen that a ridge regression with 0.00025 regularisation value was obtained as the linear model. For the decision tree, a model with a minimum number of 21 observations per leaf node and 34 decision nodes was selected. The optimal SVM model uses a Gaussian kernel function with a box constraint (C) of 999.5 and a deviation (ε) of 0.0012 derived during hyperparameter optimisation. An ANN model architecture with three hidden layers and a Relu activation function for the neurons was found during optimisation.

The prediction performance of the trained models was checked using root-mean-square error (RMSE) and goodness of fit (R^2) metrices. RMSE and R^2 are the most widely used metrices as they provide insights on how close the predictions are to the regression line. RMSE measures the standard deviation of prediction errors. RMSE emphasises greater errors by squaring the difference between the predicted and true values, penalising models that make considerable errors. In addition, RMSE is the most commonly used cost function when fitting the machine learning hyperparameters to the training data. R^2 represents the proportion of the variation in the target variable that can be explained by the input features. It is also a function of RMSE, as an increase in RMSE results in a decrease in the R^2 value. The equations for R^2 and RMSE are presented as:

$$RMSE = \sqrt{\frac{1}{N}\sum_{j=1}^{N}\left((y_{pred})_j - (y)_j\right)^2} \tag{3-25}$$

$$R^2 = 1 - \frac{\sum_{j=1}^{N}\left((y_{pred})_j - (y)_j\right)^2}{\sum_{j=1}^{N}\left(\bar{y}_{pred} - (y)_j\right)^2} \tag{3-26}$$

where y and y_{pred} denote the real and predicted values, and $\bar{y}_{pred}$ shows the average of predicted values. Table 3-5 presents the metrices of various prediction models for the train and the test data. It is seen that ANN accurately predicts the number of the faulty assets across the network. The close agreement between training and test data metrices for ANN implies its reliability in making predictions from new data.

To visualise the predictive capability of regression models, the correlations between real and predicted measurements of faulty assets are depicted in Figure 3-11 for all data. The correlation plots in this figure show that the predictions are located mostly within ±50% of the real data. It is also seen in Figure 3-11c and Figure 3-11d that data points are least skewed for SVM and ANN models. Based on the prediction results provided in Table and Figure, it is concluded that ANN assigns low prediction errors and high correlation values to the real data points. Hence, it was selected to model the influence of the input parameters on the defects of the assets in the network.

For each numerical parameter in Table 3-2, a combination of features with sequential values for the parameter of study and fixed values for the remaining variables was generated and used by the ANN model. In order to model the infrastructure defects, the change in the defect count with each input parameter was depicted. Figure 3-12a and Figure 3-12b shows the change in the total defect numbers with the base code and start point parameters. It is seen that as the inspection location gets further from the start point, the defect numbers decrease. This indicates that the maintenance schedule is more frequent near the start point. It may also indicate that assets closer to the start point are used more frequently, causing more faults than assets further away. Moreover, it is possible that the assets near this point are older and therefore more likely to have defects than those in further locations.

Figure 3-12c and Figure 3-12d depict the change in the defect numbers with the asset usage parameters. It is seen that there is a decrease in defect count with the annual increase in weight to 10 megatons. Similarly, for the first 30,000 passing services, the defect number decreases. It is possible that the assets used at lower weight and service numbers are older than the assets that are used at higher weight and service counts for the mentioned ranges. Older assets are more likely to have defects than the newer ones, which may explain the decrease in defect count. Furthermore, the applied load and service ranges may be within the design load of the assets. As a result, an increase in asset usage of up to 10 megatons per year and 30,000 services does not always lead to an increase in fault numbers. Figure 3-12c and Figure 3-12d show that the defect numbers increase after initially decreasing. This indicates that the assets have a certain service life and are designed to withstand a specified level of stress. As the load and service counts grow, the assets may be stressed beyond their load-bearing capacity, resulting in more defects.

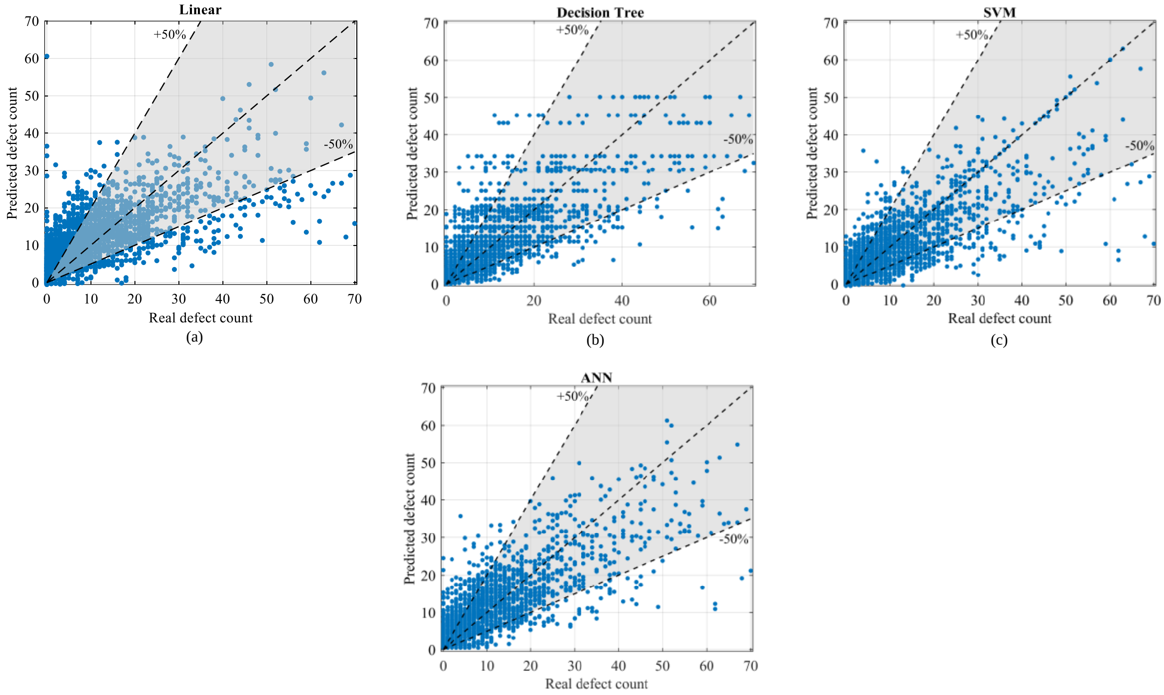
Figure 3-12e to Figure 3-12g shows the change in the defect counts in relation to the asset numbers. It is indicated that the increase in the number of assets leads to an increase in total defect counts. One possible reason is that the maintenance schedule may not be able to keep up with the increasing number of assets, resulting in an overall increase in the number of defects.

Table 3-4. List of hyperparameters with optimal values for various regression techniques

Model	Hyperparameter	Search range	Optimal value
Linear	Regularisation	Ridge, Lasso	Ridge
	λ	0-0.5	0.00025
Decision Tree	Min_Leaf_Size	1-200	21
	Max_Num_Splits	1-315	34
SVM	C	0.001-1000	995.5
	ε	0.001-1000	0.0012
	Kernel function	Gaussian, Linear, Polynomial	Gaussian
ANN	Number of hidden layers	1-4	3
	Number of hidden neurons	1-33	256, 21, 299
	Activation function	Relu, Tanh, Sigmoid	Relu

Table 3-5. Statistical metrices of machine learning models

Model	Training Data		Test Data	
	RMSE	R^2	RMSE	R^2
Linear	9.621	0.682	9.645	0.679
Decision Tree	7.906	0.799	10.760	0.618
SVM	7.968	0.801	8.511	0.770
ANN	7.890	0.806	8.340	0.772



(a) (b) (c) (d)

Figure 3-11. Real versus predicted defect count by various machine learning models

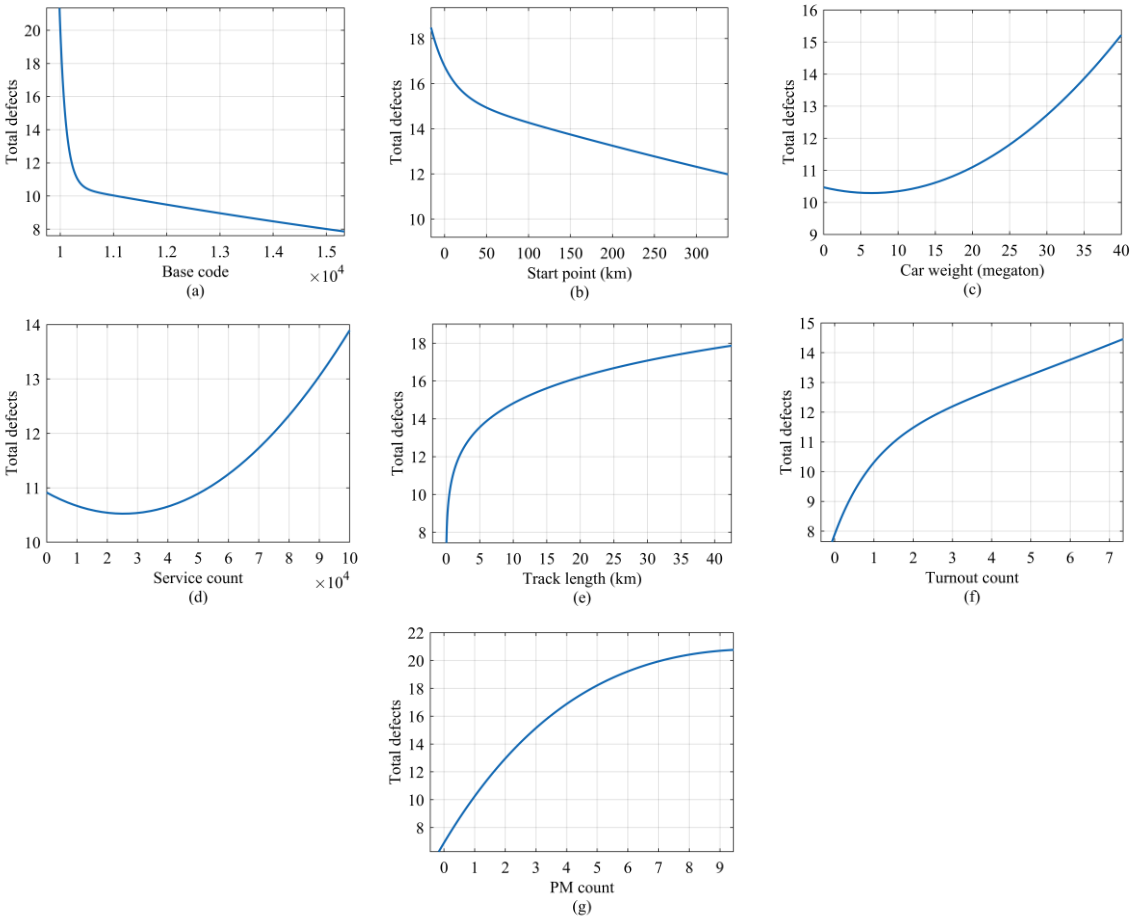


Figure 3-12. Parametric study of total asset defects

3.9 Practical equations for asset defect estimation

Developing practical equations by machine learning models is helpful in forecasting asset defects from historical data. This can help minimise the impact of defects on the railway system and improve its efficiency. Linear model equations are widely used as they are computationally efficient, easy to interpret, and can handle multiple input features. The detection of defects using historical maintenance information is a binary classification problem. Therefore, a logistic regression model was trained and used to detect the fault as a function of asset characteristics. Eq. (3-27) presents the logistic regression equation.

$$\mathrm{Ln}\left(\frac{f(\boldsymbol{x})}{1-f(\boldsymbol{x})}\right) = 8.7194 + \sum_{i=1}^{8} (a_1)_i X_i \tag{3-27}$$

where $f(\boldsymbol{x})$ is the probability of defect occurrence, X_i denotes the input variable, and a_1 is the logistic regression coefficient. Table 3-6 lists the input variables with the corresponding coefficients. The failure mode was identified by using the a_1 coefficient values in Eq. (3-27). A positive output from this equation indicates the occurrence of defects, while a negative value implies that the asset does not fail. The accuracy of Eq. (3-27) in identifying the defects is shown by a confusion matrix in Figure 3-13. It is seen that the equation is able to classify defects with 91% accuracy. In addition, it can identify 96.4% of the defective assets with 92.7% precision, implying its high accuracy in finding the faulty assets.

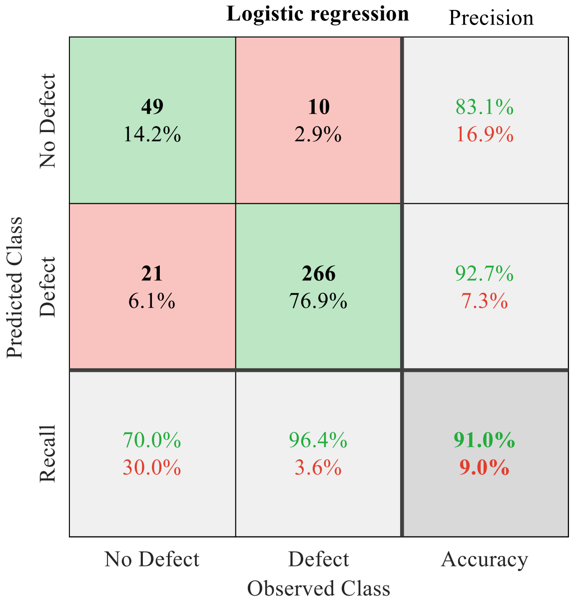


Figure 3-13. Confusion matrix of logistic regression equation

In order to estimate the number of defects from asset characteristics, a linear regression model was developed using the processed data. Eq. (3-28) shows the relationship between the input features and the defect count. The coefficient b_1 values of this equation were derived during the model training to minimise the error between the real and predicted values.

$$Y(\boldsymbol{x}) = 17.40 + \sum_{i=1}^{8} (b_1)_i X_i \tag{3-28}$$

In Eq. (3-28), $Y(\boldsymbol{x})$ estimates the defect counts, and b_1 is the linear regression coefficient. The value of b_1 for each parameter is also listed in Table 3-6. In order to check the efficiency of Eq. (3-28) in estimating the defect count, the correlation between the predictions and the real values from the test data is depicted in Figure 3-14. It is seen that most of the predictions are within ±70% of the real defect numbers. Also, a correlation value of 0.72 means that there is a strong relationship between the two outputs, implying that Eq. (3-28) is able to predict the total number of asset defects with a good degree of accuracy.

Table 3-6. Input variables with coefficients for logistic and linear regression equations

i	X_i		a_1	b_1
1	Base code		-0.0007	-0.001
2	Start point		-0.0039	-0.048
3	Car weight		0.0984	0.365
4	Service count		-7.7138	-3.57
5	Track length		-0.0266	1.18
6	Turnout count		0.1177	2.57
7	PM count		2.7433	5.13
8	Territory	R_1	-1.7591	-10.11
		R_2	0.4744	7.56
		R_3	-0.0638	0.21
		R_4	-0.0008	-1.38
		R_5	1.0381	0.35
		R_6	-0.2763	-0.80
		R_7	-0.3477	3.51
		R_8	0.3138	-3.20
		R_9	-0.5169	-2.59

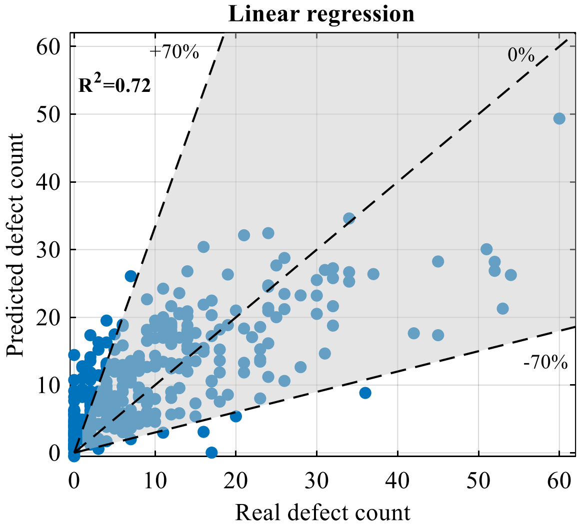


Figure 3-14. Correlation between real and predicted defect counts by linear regression equation

3.10 Conclusions

In this chapter, a data-driven machine learning-based maintenance approach was proposed to predict potential failures in the rail assets. In collaboration with the industry partner, historical data on the preventive maintenance work orders was used to develop prediction models and study the railway asset defects across the network. Data exploration techniques were used to statistically analyse the database's features. In order to select the parameters with a high influence on the infrastructure defects, an intrinsic feature selection method with an ensemble model was used. The importance of the features was ranked by the permutation technique, where the influence of various railway system properties on different defect types was investigated. It was seen that the failure was impacted by the asset quantity and usage properties within the network. Also, the number of assets greatly impacted the failure rate, while local properties slightly influenced the defects.

The prediction of the defects was performed by a two-stage classification technique. The first classifier was trained to identify the defects present in the network based on the processed historical data. The prediction results were used to filter the data for defective assets. A multi-class classification was performed by the second model while taking into account the undersampling of the majority class and using stratified sampling. The prediction results demonstrated the accuracy of the two-stage classifier in identifying the defects in the assets. A regression task was performed to further study the influence of the infrastructure characteristic on the defects. The parametric study results indicated the change in total defect counts, with the features representing the location, usage, and quantity of the assets. It was seen that the assets in further locations from the inspection start point had fewer defects. An initial decrease followed by an increase in the failures caused by the load and service count values was related to the asset's age, service life, and load-bearing capacity. It was also discovered that the total number of defects increased with the asset count in the network.

In order to use the prediction models in practice, linear machine learning equations were extracted and used to estimate defects using the information on the history of the asset's maintenance. A logistic regression equation was used to identify the probability of an occurring defect in the asset with 91% accuracy. Furthermore, a linear regression model was trained to estimate the number of asset failures using a closed-form equation. The good correlation between the prediction results and test data indicated the accuracy of the proposed equation in estimating asset defects.

The ability of the developed models to identify defects and predict the condition of the infrastructure based on historical maintenance activities allows for planning predictive strategies. As a result, the maintenance schedules can be optimised so that the occurrence of unexpected defects is reduced, and the service life of the assets are increased. Therefore, the reliability and serviceability of the ageing railway network can be improved.

This chapter demonstrated how we used machine learning to address the issue of ageing by modelling the failures of an infrastructure system, which can lead to improved maintenance strategy. Although the proposed framework makes predictions efficiently, it is solely focused on detecting defective assets and their numbers across the network, limiting its practicality in various contexts, including severity assessment of damage to the assets, estimating the remaining useful life of the assets, and analysing the cost associated with the required predictive maintenance actions. Additionally, the developed methodology in this chapter overlooks the underlying mechanisms that causes asset failures. Predicting the type of defect in the railway can also allow for more targeted maintenance strategies. Also, including the impact of extreme environments in the analyses can help develop prediction models for the durability study of the assets. However, due to the limited information in the database, it was not possible to develop prediction models for these purposes. Therefore, a future direction of the research can be to incorporate more comprehensive data to train machine learning models for the purpose of addressing the aforementioned challenges. It was seen in the literature that strengthening solutions with composite materials are a suitable option to improve the functionality and extend the service life of structures. In the next chapter, we will discuss the ageing issue in strengthened structures and how machine learning can help analyse this phenomenon at the structural scale.

CHAPTER 4

Assessment of Ageing Impact on FRP-Strengthened Concrete Structures Under Environmental Conditions Using Machine Learning Methods

4.1 Introduction

The impact of ageing on the efficiency of an infrastructure system was analysed by modelling the defects in the previous chapter. It was seen that forecasting the condition of the infrastructure can help in planning for more efficient maintenance strategy. Strengthening of structures is also another technique to improve the serviceability of infrastructure. Despite being helpful, ageing due to extreme environmental conditions can also impact the long-term performance of the strengthened structures. This necessitates analysing ageing in the structures that benefit from strengthening treatments. In the current chapter, we tend to identify degradation at the structural scale. This can aid in gaining a better understanding of how ageing affects the structural performance, allowing for more efficient strengthening design.

Structural elements are normally exposed to extreme environmental conditions during their service life, which can accelerate their ageing process. Therefore, it is important to protect structures from surrounding environmental conditions. Fibre-reinforced polymer (FRP) materials have been successfully used for strengthening and reinforcing concrete structures [9, 22]. They have been recognised as promising techniques for improving the durability of existing structures. The load transfer efficiency of the composite structures depends on the integrity of the bond between FRP and concrete [112]. However, exposure to a harsh environment can deteriorate the connection by diffusing the aggressive agents to the FRP-to-concrete interface and compromising its durability.

Several studies have investigated the effect of environmental ageing into the interface of FRP and concrete by performing durability tests [113-115]. Also, reduction factors have been proposed to estimate degradation in the connections under various exposure conditions [116]. What has not been addressed in the durability studies is the inconsistency among the test results, which creates uncertainty regarding the long-term performance of connections. Moreover, the existing models are highly empirical, and their predictive capabilities are limited to a small number of experimental results, which undermine their application in practice.

To address these limitations, the current study attempts to assess deterioration of the bond between FRP and concrete due to interactions with environmental factors using machine learning techniques. The externally and internally bonded FRP-concrete structural members are investigated for their connection durability. Frist, the durability experiments under different environmental conditions are reviewed to establish comprehensive databases for the bond analysis. A series of machine learning techniques for durability assessment are explored, and a framework for developing prediction models is proposed. Figure 4-1 illustrates the graphical abstract for environmental durability modelling of composite structures using machine learning models. The durability of the adhesively bonded connections and FRP bar-to-concrete joints is evaluated by identification of the long-term bond strength and the failure modes.

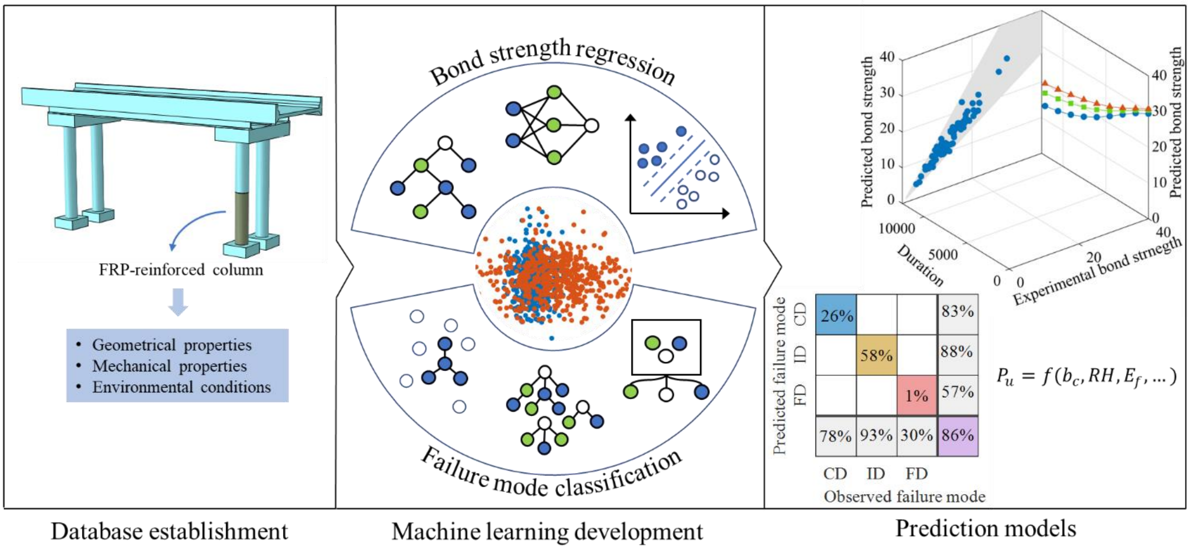


Figure 4-1. Environmental durability modelling at structural level using machine learning

4.2 Experimental data exploration

Durability of the bond between FPR and concrete is commonly studied by performing accelerated ageing tests. In these experiments, the effect of environmental factors on the connection is studied by determining the bond strength and failure mode using various types of pull-out tests. The test is conducted by measuring the tensile force required to pull the FRP plate or bar off the concrete. The bond strength is then calculated by dividing the force by the bond area. In the following, the experimental data on the durability study of the bond in the strengthened and reinforced concrete connections by FRP materials is explored.

4.2.1 Data exploration of adhesively bonded connections

The external bonding of FRP composites to concrete elements has been recognised as a promising technique in improving the long-term performance of existing structures. Durability study of the adhesively bonded FRP-to-concrete connections is commonly done by pull-out lap shear test. These tests are simple and widely used procedures to observe the bond performance through measuring the strength and failure mechanism of the connections. These tests are critical in terms of investigating the long-term performance of FRP strengthened concrete structures under various environments.

In order to prepare samples for FRP-to-concrete bonded connections, concrete prisms are cast and cured normally for 28 days. A surface preparation technique is then employed by removing the mortar layer and exposing the aggregates on the top surface, of the concrete blocks. Treating the surface before FRP attachment is crucial to removing dust and moisture from the surface as they can reduce the load capacity of the bond. The FRP composite sheets or laminates are then cut into desirable sizes and attached to the prepared concrete surface using different techniques such as wet lay-up, pultrusion, and resin infusion techniques. To study the durability of the connections to moisture conditions, the prepared samples are exposed to constant relative humidity and temperature by placing them inside a humidity chamber. After the required ageing period, the samples are taken out immediately and placed in the test rig to perform a single lap shear test. To ensure that pure shear is applied to the interface, the specimen is fixed in the rig with its bond line aligned with the loading grip. The tensile load is applied to the end of the plate until a fracture occurs in the sample. Figure 4-2a shows the front and side views of the FRP-to-concrete bonded sample. The moisture conditioning of the specimen using a humidity chamber is schematically illustrated in Figure 4-2b. Figure 4-2c shows the single lap shear test setup where the FRP plate is bonded to the concrete substrate and loaded in one direction.

Moisture is one of the most recurrent environmental conditions. It can deteriorate the connection through diffusing into the interface of FRP and concrete [117, 118]. Since moisture degrades the material properties of components in the connection, failure can occur in FRP, concrete, or the interfacial region between these two materials. Typically, five types of failure are observed for a connection, namely, (i) concrete debonding (CD), (ii) concrete/adhesive split, (iii) adhesive decohesion, (iv) FRP/adhesive split, and (v) FRP delamination (FD). Failure modes (ii) and (iv) can be grouped as interfacial debonding (ID). These modes of failure are depicted in Figure 4-2d.

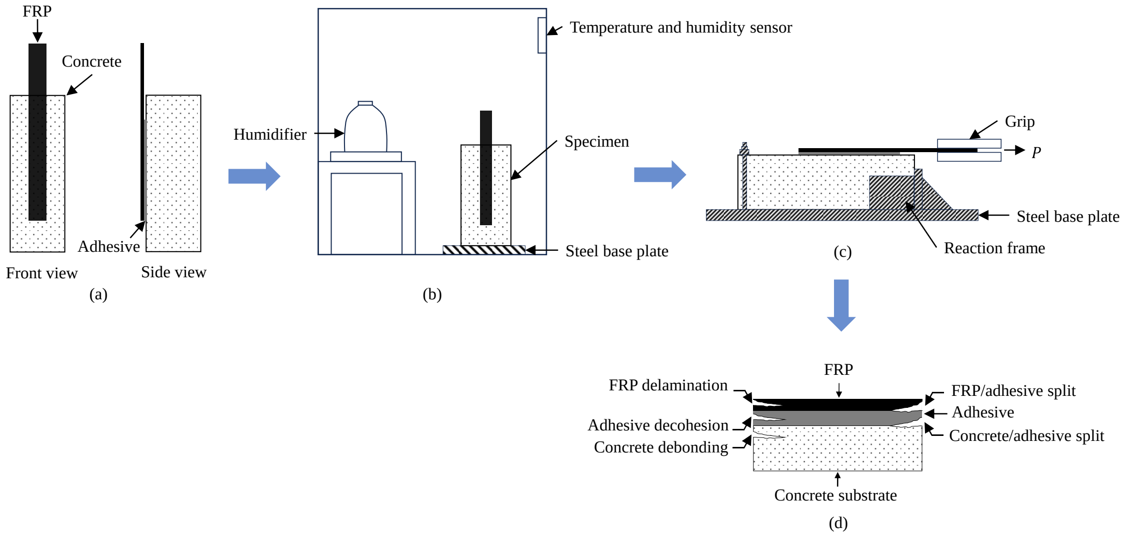


Figure 4-2. (a) FRP-to-concrete bonded connection, (b) Moisture conditioning in humidity chamber, (c) single lap shear test setup, (d) failure modes in the connection

This study establishes a comprehensive database of available experimental results on the durability of adhesively bonded connections collected from 21 research articles and dissertations. A total of 567 shear test results with information on the geometrical and mechanical properties of FRP-to-concrete connection elements and the moisture conditions were assembled. The following assumptions have been made during the database establishment to make data useful for machine learning applications:

(i) The compressive strength of a concrete cube was converted into the compressive strength of a cylinder according to Eurocode 2 [119].

(ii) The duration of tests was converted into hours, and one-month time was equalled to 672 hours.

(iii) The equivalent duration of a cyclic test was obtained by summing the durations of each cycle.

(iv) FRP/adhesive and concrete/adhesive separations were considered as interfacial failure modes.
(v) The unconditioned samples were assigned with zero duration.

Table B-1 in Appendix B shows the list of references used to collect data. The availability of data for various parameters is also presented in this table. It is seen that most of the studies have not reported the mechanical properties of adhesives and the ultimate strain and the tensile strength of the FRP. A systematic approach to deal with missing data is to remove the corresponding features from the database [120], as deleting the references considerably reduces the quantity of data. After cleaning the dataset, 429 data remained out of which 85 data points belonged to the test results of unconditioned FRP-to-concrete samples; 251 test results related to the continuous moisture exposure conditions, and 93 data points were the tests performed under cyclic moisture environments. The database also included 146 concrete debonding (CD), 270 interfacial debonding (ID), and 13 FRP delamination (FD) failure modes.

The selected features for the application of machine learning after data cleaning are summarised in Table 4-1. In this table, b_c is the concrete width; L_f, t_f, b_f, and E_f are bond length, thickness, width, and elastic modulus of FRP, respectively; f'_c is the concrete compressive strength; RH denotes the relative humidity; P_u is the experimental bond strength.

Table 4-1. Selected features for machine learning modelling

Feature type	Category	Description	Unit	Range
Input feature	Geometrical property	b_c	mm	50.8-210
		L_f	mm	48.8-400
		t_f	mm	0.11-1.5
		b_f	mm	18.7-100
	Mechanical property	E_f	GPa	26.5-409.3
		f'_c	MPa	19-67.3
	Environmental condition	Exposure type		Continuous, Cyclic
		Temperature	°C	(-20)-60
		RH	%	30-100
		Duration	hour	0-131,04
		Sustained load	kN	0-12.5
Output feature	Test result	P_u	kN	3.4-56
		Failure mode		CD, ID, FD

The distribution of data for input features is also depicted in Figure 4-3. This figure shows that the data are skewed and have different scales for various parameters. Input features with different scales can make it difficult for machine learning methods to model the output variable since the effect of parameters of small scales is undermined by the variables of greater scale [121]. Besides in a skewed distribution, the underlying machine learning calculations can be influenced by the tail of distribution [120]. Therefore, scaling the range and distribution of input variables can help to increase the prediction performance of machine learning models.

In this study, Yeo-Johnson power transformation was used as in Eq. (4-1) to normalise data and reduce skewness by creating a Gaussian-like distribution [122]:

$$\psi(x,\lambda)=\begin{cases}(x+1)^{\lambda}/\lambda & if\ x\geq 0,\lambda\neq 0,\\ \ln{(x+1)} & if\ x\geq 0,\lambda=0,\\ -[(-x+1)^{2-\lambda}-1]/(2-\lambda) & if\ x<0,\lambda\neq 2,\\ -\ln{(-x+1)} & if\ x<0,\lambda=2.\end{cases} \tag{4-1}$$

where x is the input variable; λ determines the shape of the transformed distribution; ψ is the Yeo-Johnson transformation function. The probability distribution of input variables after transformation is depicted in Figure 4-4. It is seen in this figure that the variables are translated to the same range and have normal distributions.

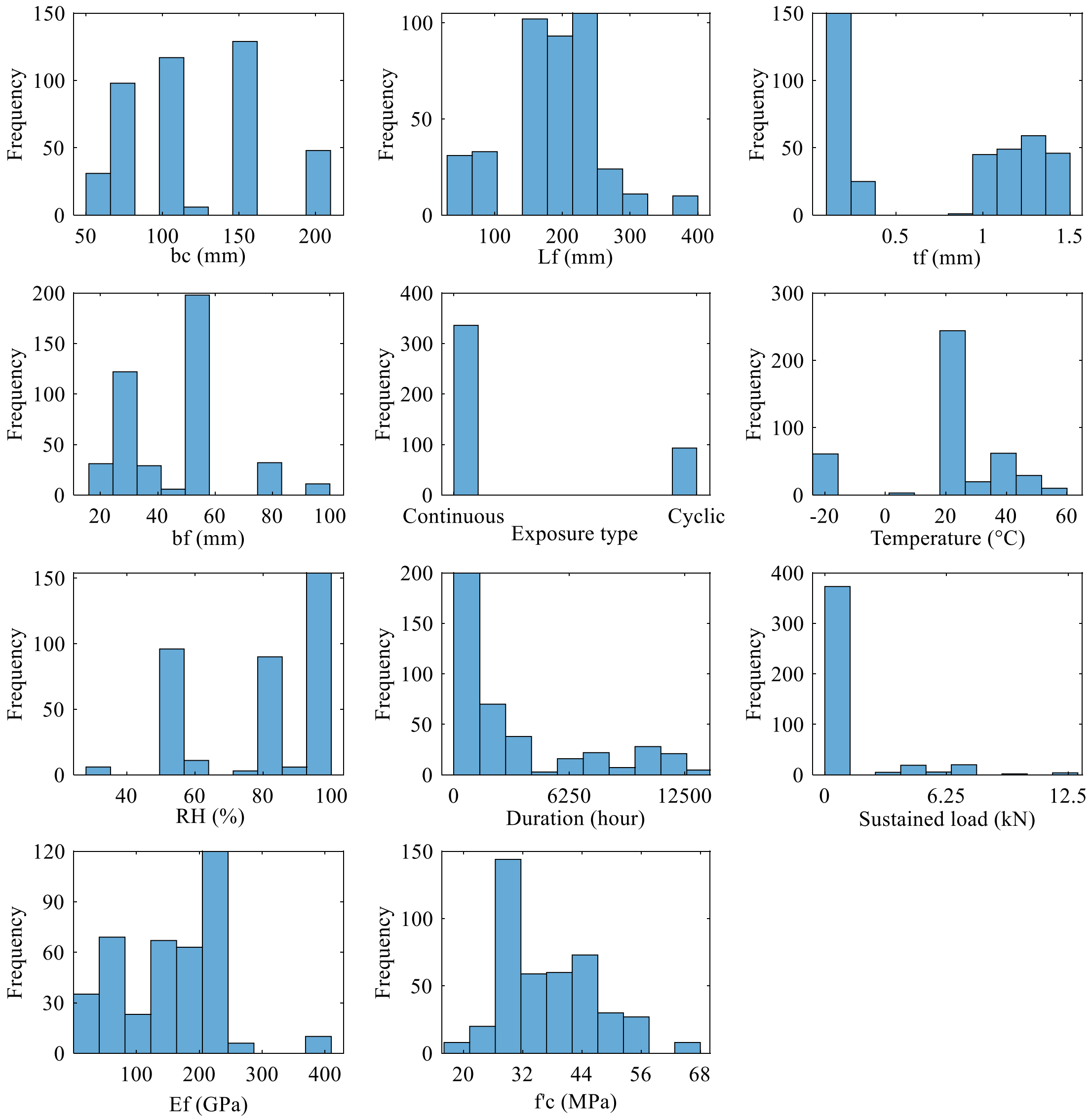


Figure 4-3. Distribution of input features

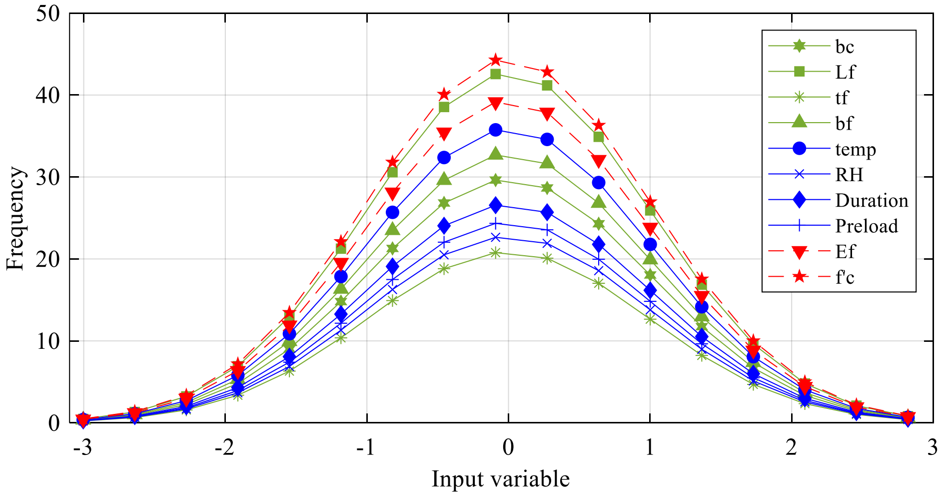


Figure 4-4. Probability distribution of transformed input variables

4.2.2 Data exploration of FRP bar-to-concrete connections

Reinforcing bars made of FRP have become increasingly popular as a substitute for steel in concrete structures, most notably because of their resistance to corrosion [123]. It has been established that under environmental degradation, the resin layer of FRP may deteriorate, which undermines the integrity of the bond between bar and concrete, resulting in the bond's failure [117, 124, 125].

The study of the bond integrity between FRP bar and concrete is performed by pullout test. For the experimental study of the bond, cubic or cylindrical concrete specimens with FRP bars embedded with a length of five time of the bar diameter are cast. The pullout samples are then cured for 28 days and subjected to environmental conditions for the required ageing period. After the conditioning period, the specimen is immediately placed in the test rig, and pullout load is applied at the free end of the FRP bar. To prevent the gripping problems, a steel tube is attached to the end of the FRP bar by injecting epoxy resin. The configuration of the pullout test for the FRP bar-concrete sample is shown in Figure 4-5.

Typically, three major failure modes can occur for a FRP bar-to-concrete connection: pull-out, concrete splitting, and FRP rupture. The failure types are depicted in Figure 4-6. In a pull-out, for a suitable embedment length, the pull-out force can surpass the concrete tension strength, resulting in shear failure. However, when the tensile stress around the bar exceeds the concrete's tensile strength, the concrete starts splitting as cracks appear along the bar and on the surface of the concrete. FRP rupture occurs due to the reduction in the tensile strength of the bar after exposure to harsh conditions.

In the current study, a database of 772 durability experimental results of FRP bars under ambient (UC), water immersion (W), saline (S), alkaline (AL), acidic (AC), and freeze-thaw (FT) conditions were compiled from 22 references. Table B-2 in Appendix A summarises the availability of the geometrical and mechanical properties of concrete, FRP bar, and resin together with the environmental parameters for the observations of bond strength in different references. In this table, $cover/d_b$ is the ratio of concrete cover to bar diameter; f'_c is the concrete compressive strength; d_b, L_b/d_b and E_f denote the diameter, embedment length to diameter ratio, and elastic modulus of the FRP bar. While compiling the data, the concrete compressive strength values were converted into cylinder corresponding values according to Eurocode 2 [126]. The pH values for samples under water immersion conditions were assigned as 7 [127]. If not specified, the ambient temperature for unconditioned samples was assumed to be 20 °C. In cyclic tests, the experiment time was equal to the sum of the duration of each cycle. Also, for unconditioned samples, the zero value was assigned for the duration parameter.

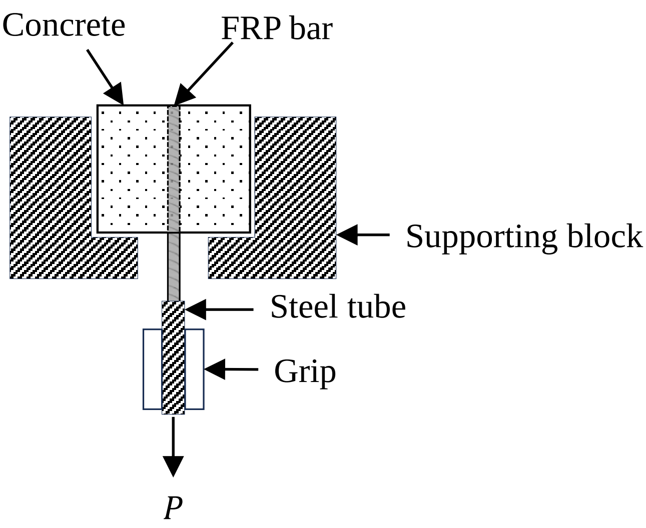


Figure 4-5. Pullout test configuration

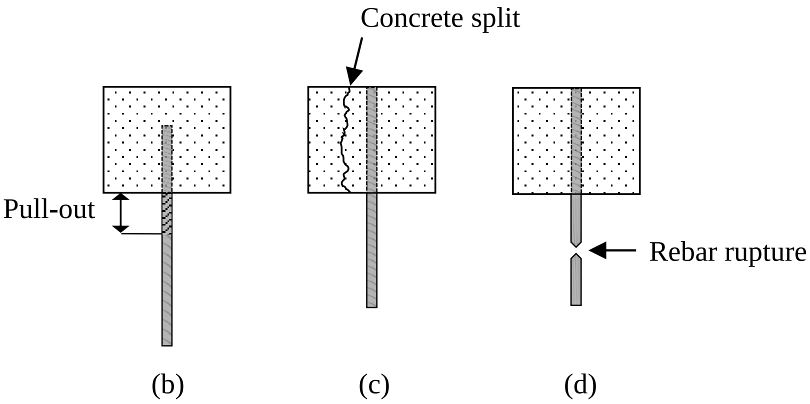


Figure 4-6. FRP bar-to-concrete failure modes: (a) Pull-out; (b) concrete split; (c) Rebar rupture

It is seen in Table B-2 that five references miss the data on the geometrical and mechanical properties as well as the test duration. While removing the variables with missing information can save the size of the dataset, it can also reduce the accuracy of the machine learning model. Therefore, the data for the five references was removed from the database. After performing data cleaning, 502 pull-out test results were collated to investigate the durability of the bond. In this dataset, 88 specimens were tested under UC setting and 48 samples were examined in W condition; 71, 218, and 31 specimens were tested in AL, S, and AC solutions, respectively; and 46 specimens were tested in FT cyclic environment. The dataset also includes 454 pull-out (PO) and 48 concrete-splitting (SP) failure modes.

The next step after cleaning the data is to check the scale of the input parameters. Table 4-2 describes the selected features for machine learning applications. It is seen that two types of numerical and categorical variables are present in the dataset. Since machine learning algorithms operate on numerical data, it is necessary to convert the labels of the categorical data into numbers. In this study, the one-hot encoding method was used to transform the categorical features into binary values [95]. It works by adding additional features with 0 or 1 values for each category with 1 denoting the presence of the categorical label in the data. It is also seen from Table 4-2 that the variables have different ranges. Therefore, rescaling the parameters to a unit range is done by Eq. (4-1) to avoid the large-scale variables masking the effect of small-scale features on the prediction model. Figure 4-7 shows the distribution of scaled data for the input features. It is seen that the data is normally distributed, while the mean and standard deviation values for all parameters are 0 and 1, respectively. Therefore, following encoding the categorial variables and applying power transformation, a total of 28 features were prepared for in developing machine learning models.

Table 4-2. Details of the features selected for machine learning application

Feature type	Feature description	Unit	Range
Numerical property	d_b	mm	8–19
	$cover/d_b$	–	1.84–8.88
	L_b/d_b	–	3.84–7
	f'_c	mm	20–58.5
	E_f	Gpa	33–152.1
	Temperature	°C	(-20)–90
	Duration	hour	48–5760
	pH	–	2–13
	Salt concentration	%	3–5
Categorical property	Bar surface texture	–	Sand coated, Ribbed, Smooth
	Bar material	–	Carbon, Glass, Basalt, Steel
	Resin material	–	Polyester, Vinyl ester, Epoxy
	Test type	–	Concentric, Eccentric
	Exposure type	–	Continuous, Cyclic
	Conditioning type	–	Water immersion, Alkaline, Saline, Acidic, Freeze-thaw

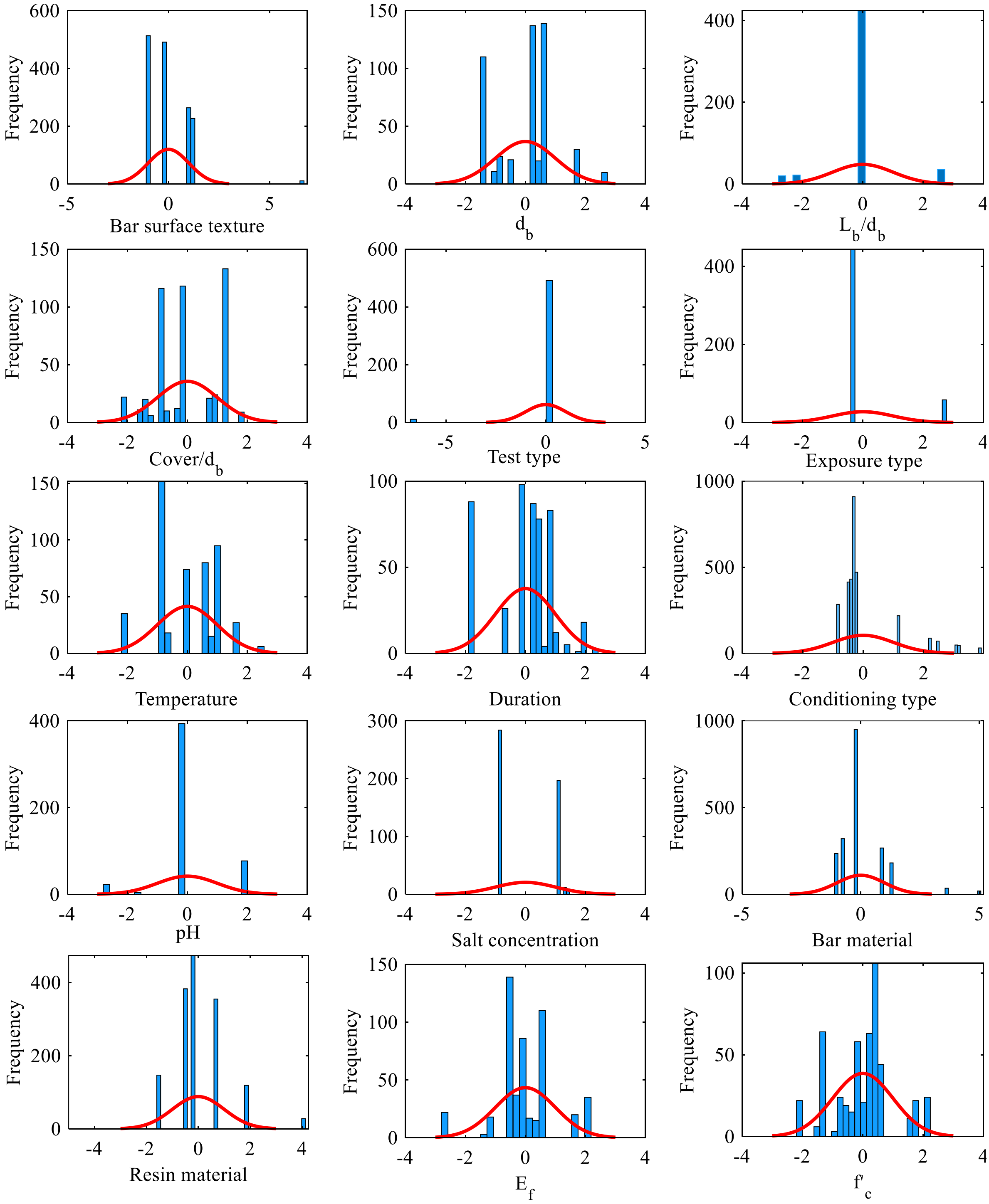


Figure 4-7. Distribution of input features after Yeo-Johnson transformation

4.3 Machine learning model development

The proposed regression and classification machine learning algorithms were used to predict the durability of the bond in the adhesively bonded and FRP bar-to-concrete connections. In order to develop prediction models and evaluate their performance, the data was randomly divided into training and test sets. 80% of data were used to train machine learning, and the remaining data points were used to test the accuracy of the model. The training process included Bayesian optimisation technique where 10-fold CV technique was used to sample the data for tuning the machine learning hyperparameters. The model with minimum CV error was nominated as the optimal model and used to model the strength and failure mode of the bond between the FRP and concrete connections. The accuracy and generalisation of the model to new data was checked by the test set. Figure illustrates the framework developed to model bond degradation in this Chapter.

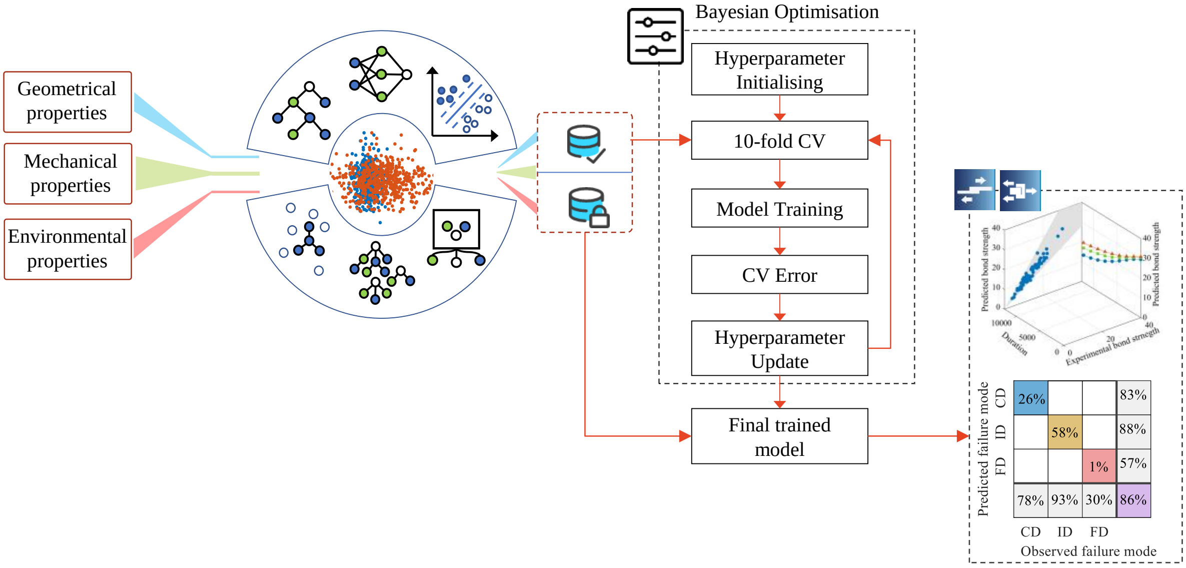


Figure 4-8. Workflow for modelling bond degradation

4.4 Results and discussions

The established databases were used to assess the performances of machine learning models in modelling the effect of environmental agents on the bond in adhesively bonded joints and FRP bar-to-concrete connections. The regression and classification models with tuned hyperparameters were developed and compared in performance to select accurate models for the parametric analysis of environmental factors. Moreover, closed-form equations were proposed to estimate the long-term performance of the connections for practical purposes.

4.4.1 Durability assessment of FRP-to-concrete adhesively bonded connections under moisture condition

The long-term performance of the adhesively bonded connections under moisture conditions was modelled by analysing the processed dataset from Section 4.2.1 using the developed machine learning framework. This provides an efficient way to predict the bond strength and failure mode of the connection and reduces the cost and time required for durability experiments.

4.4.1.1 Strength prediction of adhesively bonded connections

The bond strength prediction of FRP-to-concrete connections under moisture condition is crucial to understand the performance of FRP-strengthened reinforced concrete members during exposure. A machine learning approach can help accurately determine bond strength in relation to degrading factors. A regression analysis was performed in this section to identify the best model for the prediction of bond strength. Since hyperparameter values are critical in developing accurate prediction models, a combination of Bayesian optimisation with 10-fold CV was adopted to tune these values. Table 4-3 presents the results of Bayesian optimisation for various prediction models. The obtained values in this table were used to train machine learning models and test their efficiency.

Table 4-3. List of hyperparameters with optimal values for various regression techniques

Model	Hyperparameter	Search range	Optimal value
Ridge	λ	0-0.5	0.0394
Lasso	λ	0-0.5	0.0218
Decision Tree	Min_Leaf_Size	1-200	1
	Max_Num_Splits	1-315	85
SVM	C	0.001-1000	301.2
	ε	0.001-1000	0.0031
	Kernel function	Gaussian, Linear, Polynomial	Gaussian
ANN	Number of hidden layers	1-4	2
	Number of hidden neurons	1-33	30, 5
	Learning rate	0-0.1	0.0302
	Dropout	0-0.499	0.0272

To measure the predictive performance of the models tuned by the values in Table 4-3, mean RMSE and R^2 metrices were used. Table 4-4 presents the metrices of various prediction models for the train and the test data. It is seen that decision tree, SVM, and ANN provide better prediction performance than linear models. Moreover, the metric values of the test set indicate the highest efficiency of ANN in predicting the unseen bond strength values amongst other models. The close measurements between training and test data for ANN imply that the developed model is not overfitted to the training dataset, which indicates the generalisation ability of ANN in fitting to the new data.

Table 4-4. Statistical metrices of machine learning models

Model	Training data		Test data	
	RMSE	R^2	RMSE	R^2
Linear	4.553	0.647	3.893	0.762
Ridge	4.574	0.646	3.979	0.761
Lasso	4.554	0.647	3.895	0.763
Decision Tree	1.683	0.951	2.132	0.931
SVM	1.693	0.952	2.210	0.923
ANN	1.687	0.952	1.738	0.953

To visualise the predictive capability of machine learning models, the correlations of experimental measurements with predicted values of the test set are depicted in Figure 4-9. Figure 4-9a to Figure 4-9c show that for the linear, lasso, and ridge regression, the points are distributed within ±50% of the equality line. For the remaining models (Figure 4-9d to Figure 4-9f), the predicted values are approximately within ±30% of the experimental values. Based on the performance metrices provided in Table 4-4 and Figure 4-9, it is seen that ANN is able to assign low prediction errors and high correlation values to the train and test data, which demonstrates the predictive capability of this method amongst other regression techniques. Hence, ANN is selected as the machine learning model for the bond strength prediction of FRP-to-concrete connections under moisture conditions.

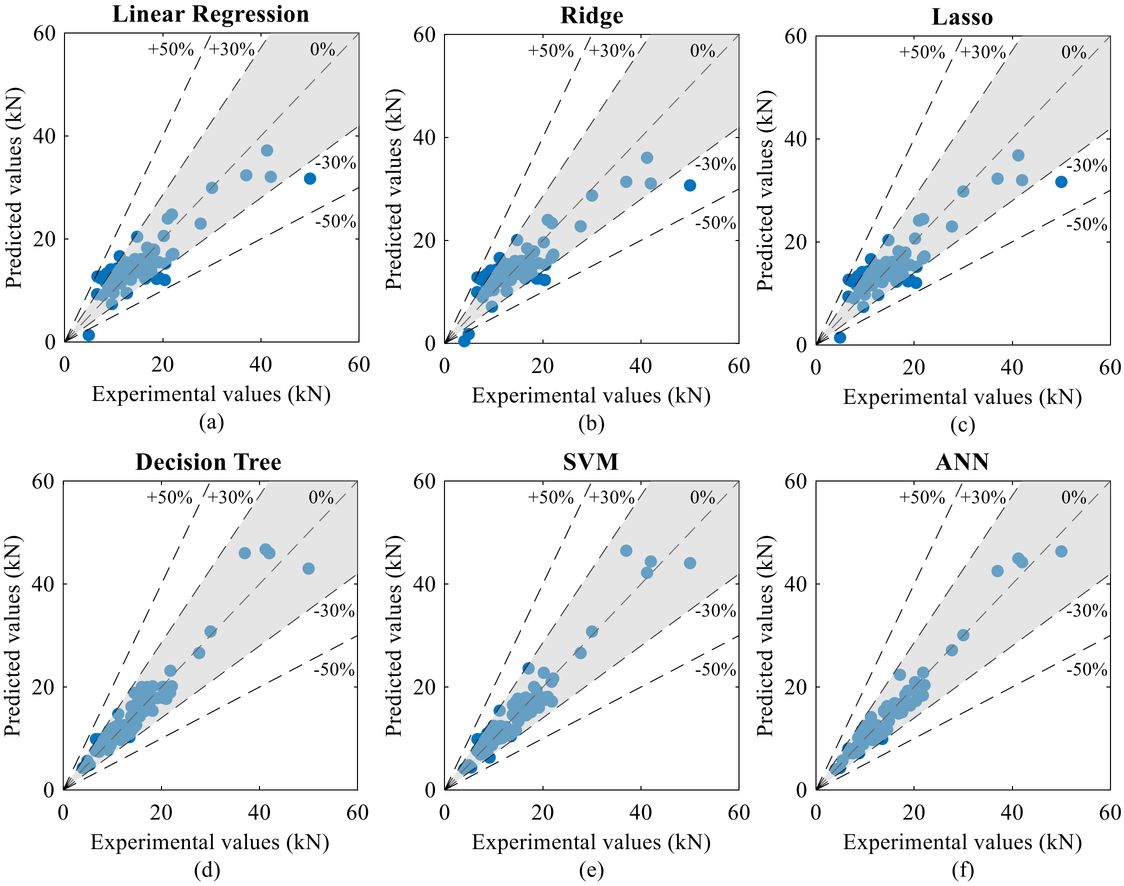


Figure 4-9. Experimental versus predicted bond strength values of test set by various machine learning models

4.4.1.2 Influence of environmental conditions on the strength of adhesively bonded connections

An evaluation of bond strength under various moisture conditions was performed to demonstrate the impact of environmental agents on the FRP-concrete interface. Based on the outcomes of Section 4.4.1.1, ANN was used to study the variation of bond strength under moisture conditions. For each environmental factor, a dataset including a range of sequential values of the environmental variable and fixed values for the remaining parameters was generated and loaded in the ANN model. Figure 4-10 represents the fully connected ANN model with input, hidden and output layers as well as the input and output variables used to perform the parametric analysis.

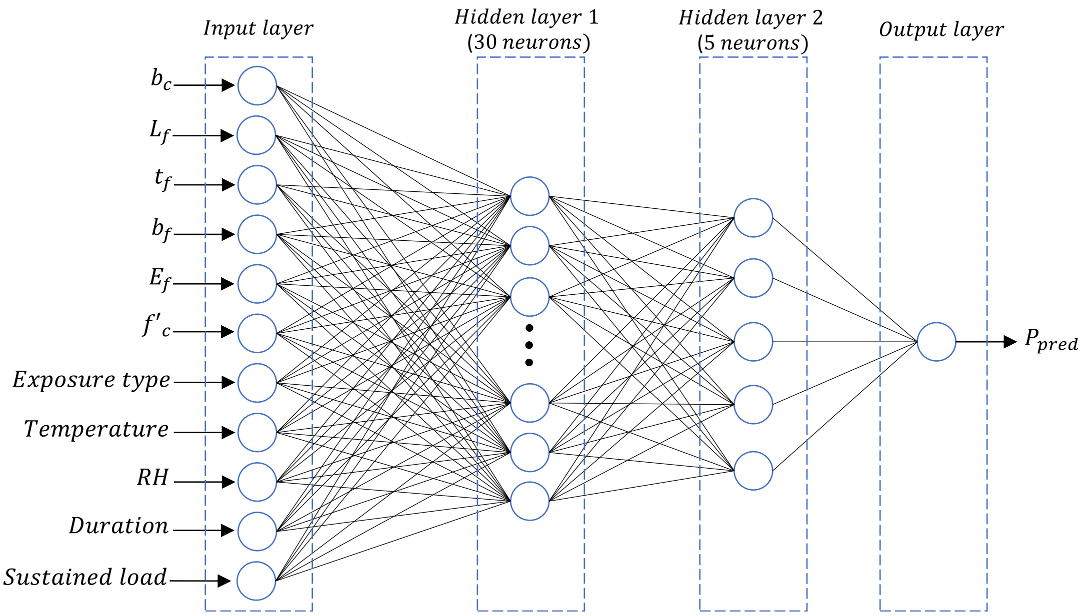
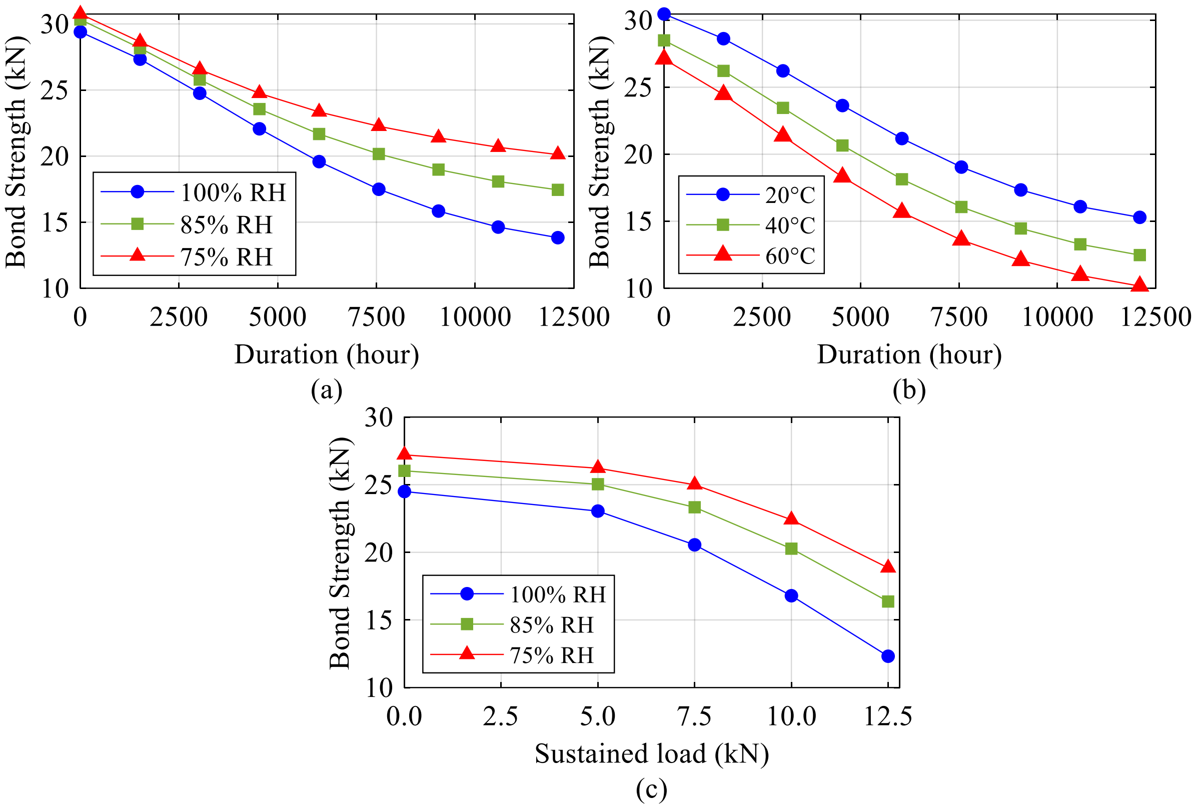


Figure 4-10. ANN layout for the prediction of bond strength

Figure 4-11 shows the loss in the bond strength during 18 months of continuous exposure to 30 °C moisture with different levels of relative humidity (RH). This figure shows that, for a higher percentage of RH, the bond strength values are less than those with lower RH levels. It is also seen in Figure 4-11a that the rate of bond loss increases with RH, implying that the increase in the moisture content adversely impacts the integrity of the bond.

Figure 4-11b illustrates the effect of temperature variation and time on bond strength under 100% RH. It is seen that by the progression of exposure, the bond strength value decreases. Moreover, a lower value of bond strength is associated with the higher temperature, indicating that the combination of moisture and elevated temperature escalates the bond degradation. Figure 4-11c shows the combined effects of sustained loading and moisture on the bond strength after six months of exposure to continuous moisture in a 30 °C solution. In addition, the rate of bond loss agrees with the level of RH, implying that moisture content intensifies the impact of sustained loading on the strength capacity of the bond.

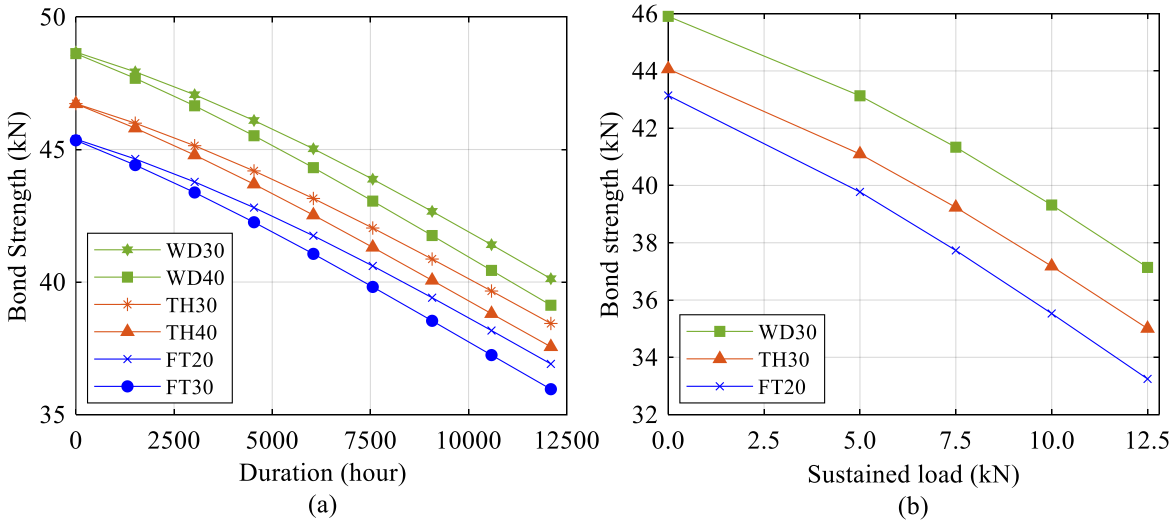


(a)

(b)

(c)

Figure 4-11. Strength loss under continuous moisture: (a) impact of RH at 30 °C, (b) impact of temperature variation at 100% RH, and (c) combined impact of sustained load and RH variation at 30 °C.



(a)

(b)

Figure 4-12. Strength loss of FRP-to-concrete bond under cyclic moisture

The impact of various cycling conditions on bond strength is illustrated in Figure 4-12. The losses in the bond strength during exposure to wet-dry cycles in +30 °C (WD30) and +40 °C (WD40) temperatures, thermal cycles from +20 °C to +50 °C (TH30) and +60 °C (TH40), and freeze-thaw cycles with 20 °C (FT20) and 30 °C (FT30) temperature variations are displayed in Figure 4-12a. It is seen that freeze-thaw cycles have the highest contribution to the bond loss, with FT30 as the most severe condition. Besides, in each conditioning regime, lower bond strength values are related to higher temperature variations. These demonstrate that FRP-to-concrete connections are susceptible to the variation in temperature and moisture content. Figure 4-12b shows the bond strength loss after six months of exposure to different wet-dry, thermal, and freeze-thaw cycles and sustained loading conditions. It is seen that as the value of sustained load increases, the bond strength decreases for all cyclic conditions. Moreover, the freeze-thaw cycle has a higher contribution to bond degradation compared to wet-dry and thermal cycle conditions.

4.4.1.3 Strength prediction equation of adhesively bonded connections

For the practical application of the developed ANN model, one possible approach is to mathematically represent the ANN algorithm. Although, the established ANN model accurately predicts the FRP-to-concrete bond strength, it generates complicated mathematical relations between the input and the output variables, thereby undermining its practical application. To address this issue, the M5P algorithm is used to develop an equation for bond strength, and its prediction accuracy is compared with respect to the developed ANN model. In the M5P formulation, predicted bond strength (P_{pred}) can be written as a function of the input features, presented in Table 4-1, in the general form of:

$$P_{pred} = f(b_c, L_f, t_f, b_f, C, T, RH, D, Pr, E_f, f'_c) \tag{4-2}$$

where C denotes exposure type; T is temperature; D is duration, and Pr expresses sustained load values. Since M5P provides linear regression equations for the observations in each leaf node, Eq. (4-2) can be written as the linear combination of input parameters as:

$$\begin{aligned} P_u = {} & a + b \times b_c + c \times L_f + d \times t_f + e \times b_f + f \times C + g \times T + h \times RH + i \times D \\ & + j \times Pr + k \times E_f + l \times f'_c \end{aligned} \tag{4-3}$$

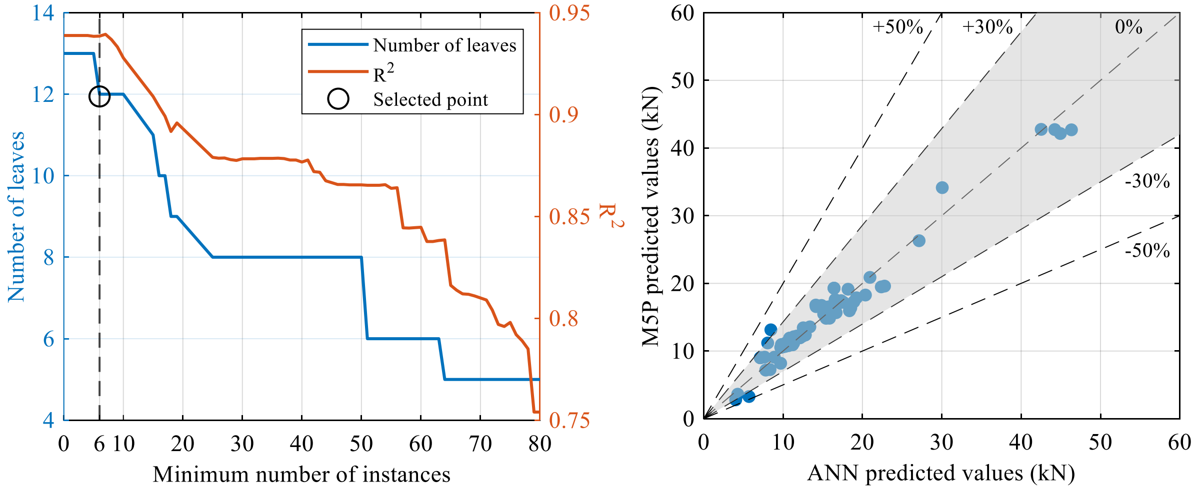
where a is the constant, and $b, c, d, e, f, g, h, i, j, k,$ and l are the coefficients of the input variables derived from the M5P algorithm. Since Yeo-Johnson transformation was applied to the data, it is necessary to rewrite Eq. (4-3) in a transformed form as:

$$
\begin{aligned}
P_u = a &+ b \times [6.553(b_c + 1)^{0.175} - 14.957] + c \times \left[0.019\left(L_f + 1\right)^{0.961} - 2.927\right] \\
&+ d \times \left[5.511 - 7.104\left(t_f + 1\right)^{-0.582}\right] + e \times \left[63.621\left(b_f + 1\right)^{0.034} - 72.071\right] \\
&+ f \times [1.932 - 2.459(C + 1)^{-6.277}] \\
&\qquad + h \times [9.53 \times 10^{-7}(RH + 1)^{3.227} - 1.828] \\
&+ g \times T' + i \times [0.406(D + 1)^{0.23} - 2.033] \\
&\qquad + j \times [\,2.583 - 2.971(Pr + 1)^{-3.882}] \\
&+ k \times \left[1.53 \times 10^{-4}\left(E_f + 1\right)^{0.809} - 2.526\right] \\
&\qquad + l \times [10.208 - 42.44(f'_c + 1)^{0.3971}]
\end{aligned}
\tag{4-4}
$$

In the above equation, C takes 0 or 1 value for continuous or cyclic moisture conditions. Since Eq. (4-1) is constrained to the sign of input variable, the transformed temperature variable, T', in Eq. (4-4) is written separately for positive and negative values as:

$$
T' = \begin{cases} 0.034(T + 1)^{1.148} - 1.271 & if\ T \geq 0 \\ 0.0097(1 - T)^{0.852} - 1.2464 & if\ T < 0 \end{cases}
\tag{4-5}
$$

The performance of M5P depends on the minimum number of observations in each leaf node. To find an M5P model with acceptable performance, R^2 values and the number of leaves are plotted as functions of the minimum observation numbers per leaf in Figure 4-13a. This figure indicates the inverse relationship of the minimum number of instances in each leaf node with R^2 and the number of leaves. It is seen that, for a minimum of 6 instances, M5P builds a tree with 12 leaves, which predicts the training data with $R^2 = 0.938$. Figure 4-13b shows the ANN predictions versus predicted bond strength values by M5P equations using the test data. The data points are mostly scattered between ±30% of the equality line, indicating that Eqs. (4-4)–(4-5) are able to predict bond strength as accurate as the developed ANN model. Figure 4-14 shows the constructed M5P model tree for predicting bond strength. In this figure, each leaf node indicates a fitted linear model with its coefficient values given in Table C-1 in Appendix C. For each linear model, P_{pred} is calculated using the values of this table and Eqs. (4-4)–(4-5), which facilitates the accurate and rapid identification of FRP-to-concrete bond strength under moisture condition.



(a)

(b)

Figure 4-13. M5P model tree: (a) learning process, (b) ANN versus M5P prediction of bond strength values of the test set

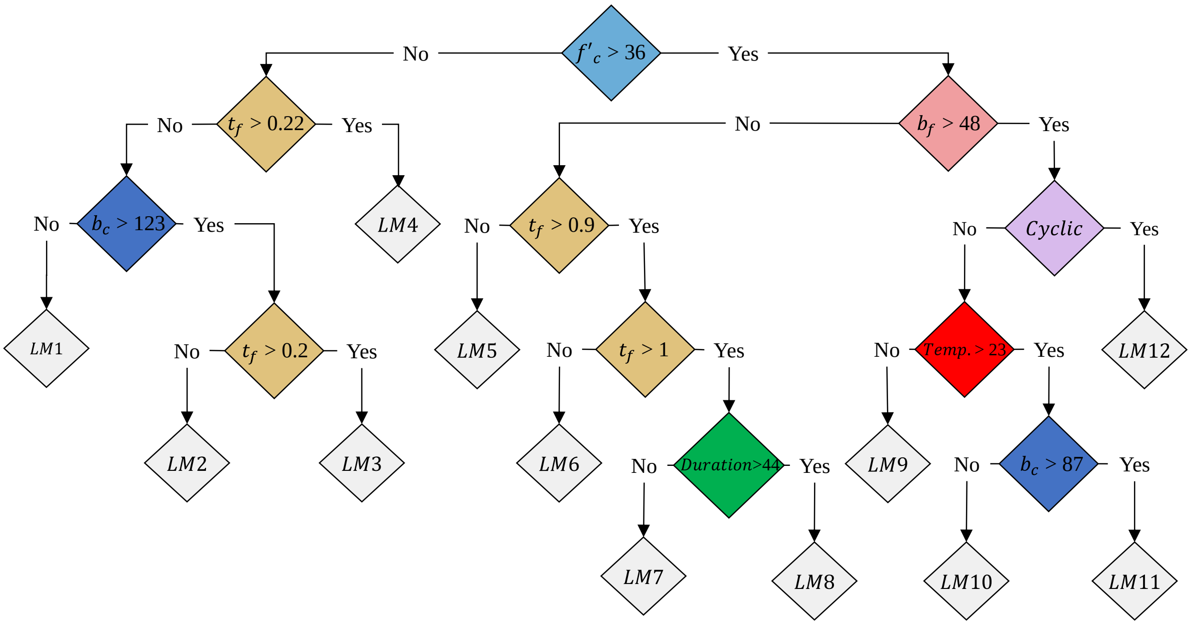


Figure 4-14. M5P model tree for the prediction of bond strength

4.4.1.4 Failure mode identification of adhesively bonded connections

The identification of failure mode in adhesively bonded connections is crucial to understand the debonding resistance of the FRP strengthened concrete elements. The lack of mechanics-based approaches prevents the rapid identification of failure mode in the connections exposed to moisture conditions.

A classification-based data-driven approach was developed in this study and used to predict the failure modes. To realise the effect of moisture, the distribution of experimental bond strength values with environmental parameters across the failure mode classes are depicted in Figure 4-15. Figure 4-15a shows the scatterplot of bond strength with time for different failure modes. It is seen that, at the early stages of exposure, the failure mode of samples is dominantly CD, which demonstrates the load transfer capacity of the FRP-to-concrete bond. However, by the progression of exposure, general trends of the shift in failure from the concrete layer to the interfacial region and the decrease in bond strength values are evident. These imply that moisture can degrade the bond with time through weakening the FRP-concrete interface. The low values of bond strength for the data points marked with FD in Figure 4-15a are related to the low tensile strength or the small cross-sectional area of FRP composites.

Figure 4-15b shows the change in failure modes from CD to ID with the increase in temperature, which indicates that elevated temperature with moisture can deteriorate the FRP-concrete bond. The combined effects of sustained loading with moisture on the failure mode are illustrated in Figure 4-15c. It is seen that the application of sustained load on the specimens during the moisture exposure largely shifts the failure mode from CD to ID.

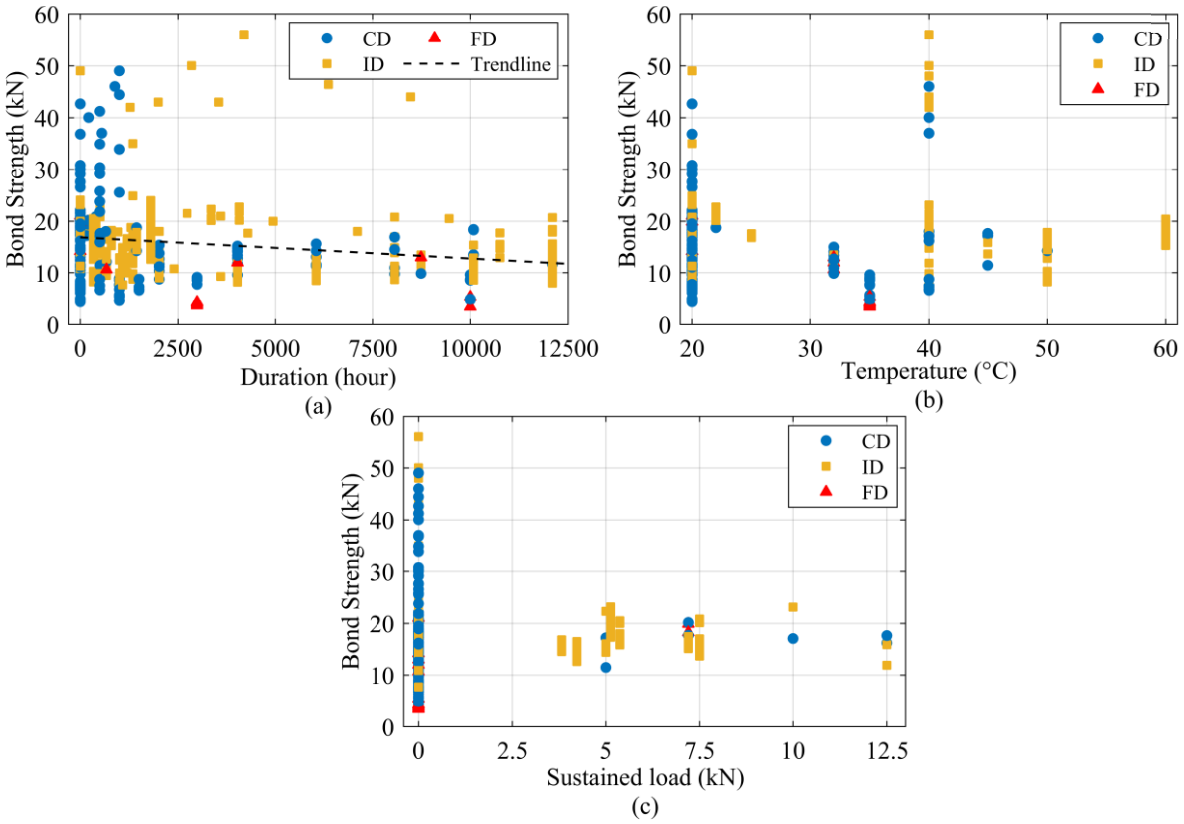


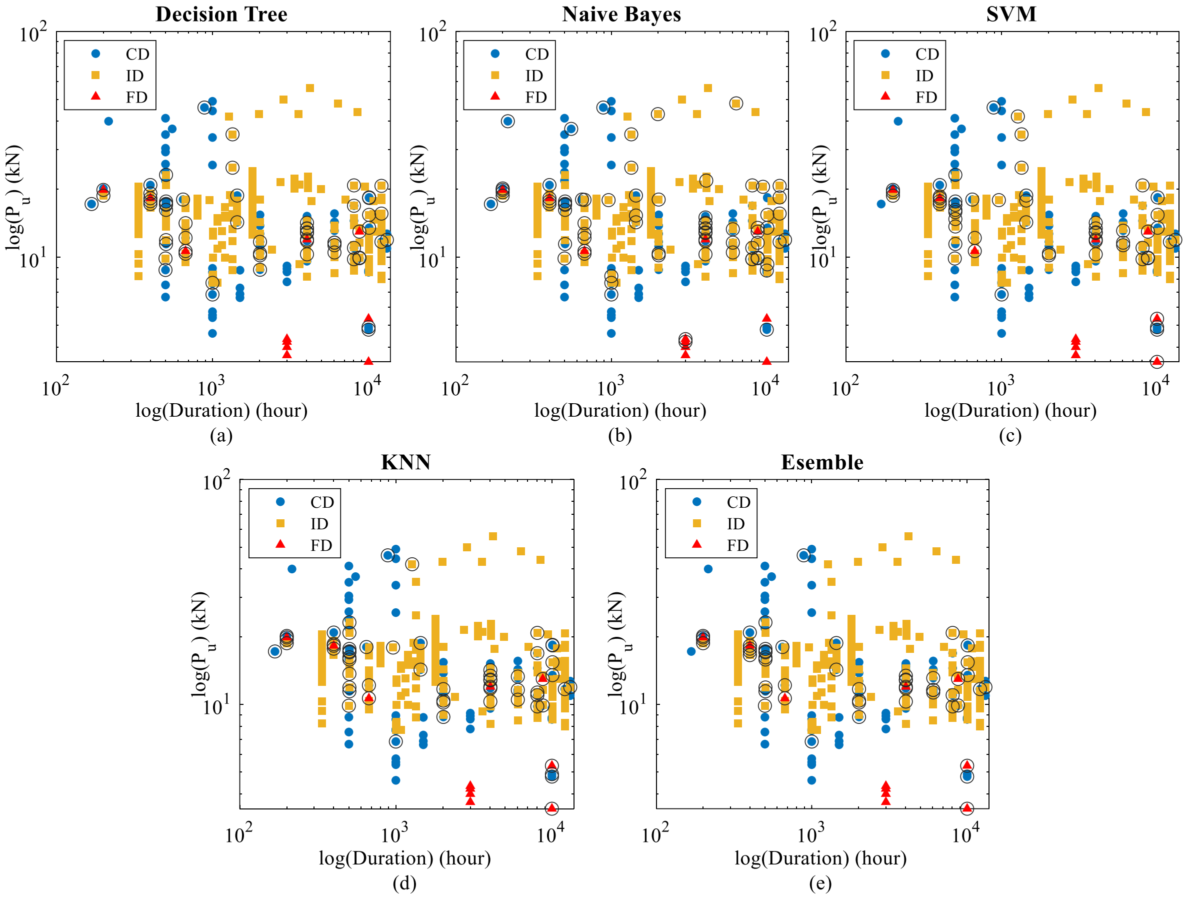
Figure 4-15. Scatterplot of bond strength versus environmental parameters: (a) duration, (b) temperature, and (c) sustained load

To identify the failure modes of FRP-to-concrete connections, a classification analysis was performed. Since there was no adhesive decohesion failure in the database, CD, ID, and FD failure modes were considered in this study for classification. It is also seen from Figure 4-15 that the data points of the FD class are fewer than those of CD and ID classes. Since the data distribution is imbalanced among the classes, the test set is likely to have a few or no FD class instances. Therefore, a stratified 10-fold CV technique was used to resample the data to ensure that the FD class samples are present in the out-of-fold predictions. To build high-performance prediction models, Bayesian optimisation was adopted, and the hyperparameters with minimum prediction errors were determined as presented in Table 4-5.

Table 4-5. List of hyperparameters with optimal values for various classification techniques

Model	Hyperparameter	Search range	Optimal value
Decision Tree	Min_Leaf_Size	1-214	4
	Max_Num_Splits	1-428	363
Naïve Bayes	Feature distribution	Normal, Kernel	Kernel
SVM	C	0.001-1000	961.82
	Kernel function	Gaussian, Linear, Polynomial	Gaussian
KNN	K	1-215	3
Ensemble	Method	Bagging, AdaBoost, RUSBoost	Bagging
	B	10-500	39

Figure 4-16 shows the classification performance of machine learning models tuned by the values in Table 4-5. The failure modes in this figure are depicted as functions of bond strength and duration parameters. The blue, orange, and red markers, indicate CD, ID, and FD failure modes. The wrongly classified (misclassified) values were also marked by black circles in this figure. It is seen in Figure 4-16 that the number of misclassified instances is different among various machine learning techniques. Generally, ensemble and SVM have the lowest prediction error, while naive Bayes presents the highest numbers of misclassification.



(a) (b) (c)

(d) (e)

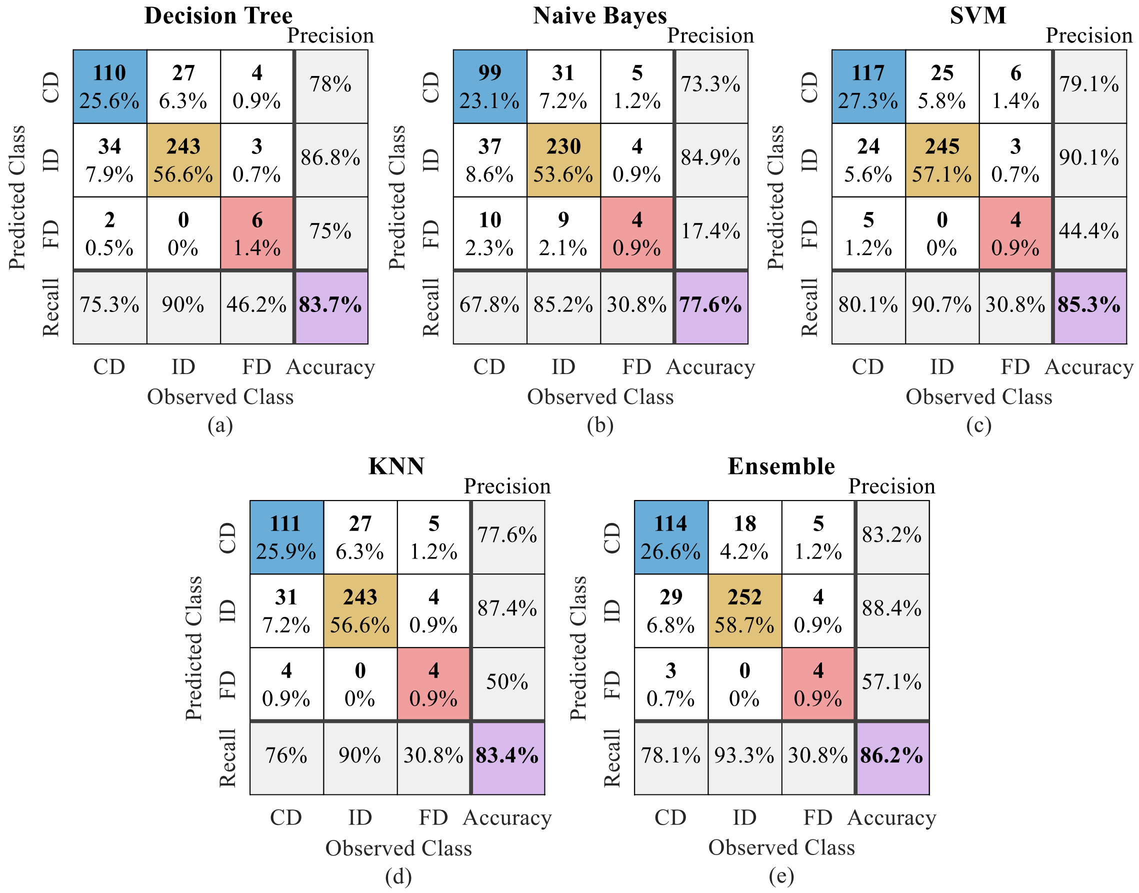
Figure 4-16. Failure mode scatter plots for various classification models using out-of-fold predictions

The out-of-fold predictive performance of machine learning methods was also evaluated in detail by confusion matrices in Figure 4-17. The diagonal cells in the confusion matrix indicate the number of observations classified correctly, and the off-diagonal numbers correspond to the misclassified instances. Three metrics, namely, precision, recall, and accuracy, were used to evaluate the classification performance and are given in the fourth columns and rows of the confusion matrices. For a specific class like ID, recall is the percentage of failure mode observations correctly predicted as ID; precision describes the percentage of predicted failure modes observed as ID. For each machine learning model, accuracy indicates the percentage of correctly predicted failure modes out of all observations. A classifier with high recall, precision, and accuracy values can identify the failure modes with a greater correction rate. It is seen from Figure 4-17 that:

- Ensemble method has the highest accuracy of 86.2% followed by SVM (85.3%), decision tree (83.7%), KNN (83.4%), and naïve Bayes (77.6%).

- The lower precision and recall rates of FD, compared to other classes, indicates that the classifiers have difficulty in identifying this failure mode.
- The off-diagonal numbers in the third columns of confusion matrices indicate that the FD class is mostly predicted as the CD failure mode. Since the failure types of both CD and FD are material decohesion, these classes can be grouped as a single failure mode for a naive classification task [114].
- The poor predictive performance of naïve Bayes indicates that the assumption of independent input features by this technique is not accurate in classifying the failure modes.
- The high accuracy rates of tree-based models and the SVM method imply that decision boundaries between the classes are highly nonlinear. Decision trees separate data by orthogonal boundaries, while SVM projects data into higher dimensional space by Gaussian kernel function to make separations.

Based on its high accuracy, recall, and precision rates, the bagging ensemble method is suggested as a machine learning model to classify the failure modes of FRP-to-concrete connections exposed to moisture conditions.



Decision Tree

Predicted Class \ Observed Class	CD	ID	FD	Precision
CD	**110** 25.6%	**27** 6.3%	**4** 0.9%	78%
ID	**34** 7.9%	**243** 56.6%	**3** 0.7%	86.8%
FD	**2** 0.5%	**0** 0%	**6** 1.4%	75%
Recall	75.3%	90%	46.2%	**83.7%**
	CD	ID	FD	Accuracy

(a)

Naive Bayes

Predicted Class \ Observed Class	CD	ID	FD	Precision
CD	**99** 23.1%	**31** 7.2%	**5** 1.2%	73.3%
ID	**37** 8.6%	**230** 53.6%	**4** 0.9%	84.9%
FD	**10** 2.3%	**9** 2.1%	**4** 0.9%	17.4%
Recall	67.8%	85.2%	30.8%	**77.6%**
	CD	ID	FD	Accuracy

(b)

SVM

Predicted Class \ Observed Class	CD	ID	FD	Precision
CD	**117** 27.3%	**25** 5.8%	**6** 1.4%	79.1%
ID	**24** 5.6%	**245** 57.1%	**3** 0.7%	90.1%
FD	**5** 1.2%	**0** 0%	**4** 0.9%	44.4%
Recall	80.1%	90.7%	30.8%	**85.3%**
	CD	ID	FD	Accuracy

(c)

KNN

Predicted Class \ Observed Class	CD	ID	FD	Precision
CD	**111** 25.9%	**27** 6.3%	**5** 1.2%	77.6%
ID	**31** 7.2%	**243** 56.6%	**4** 0.9%	87.4%
FD	**4** 0.9%	**0** 0%	**4** 0.9%	50%
Recall	76%	90%	30.8%	**83.4%**
	CD	ID	FD	Accuracy

(d)

Ensemble

Predicted Class \ Observed Class	CD	ID	FD	Precision
CD	**114** 26.6%	**18** 4.2%	**5** 1.2%	83.2%
ID	**29** 6.8%	**252** 58.7%	**4** 0.9%	88.4%
FD	**3** 0.7%	**0** 0%	**4** 0.9%	57.1%
Recall	78.1%	93.3%	30.8%	**86.2%**
	CD	ID	FD	Accuracy

(e)

Figure 4-17. Confusion matrices of various classification models using out-of-fold predictions

4.4.1.5 Influence of environmental factors on the failure of the adhesively bonded connections

Feature importance is commonly used to explain the importance rank of each input feature to the target variable. In this study, the impact of influencing factors (i.e. temperature, RH, duration, and sustained loading) on the failure modes was measured by the out-of-bag permutation importance. In this method, the out-of-bag observations of the features were first randomly permuted and used to estimate the bagging ensemble model error. The importance of parameters was then ranked based on the error raised by permuting their values. Figure 4-18 shows the contributions of environmental factors in identifying the failure modes by the ensemble model. It is seen that duration has the most influence on the failure mode of FRP-to-concrete connections, followed by RH, sustained load, and temperature, which have lower and relatively close impacts on the type of failure.

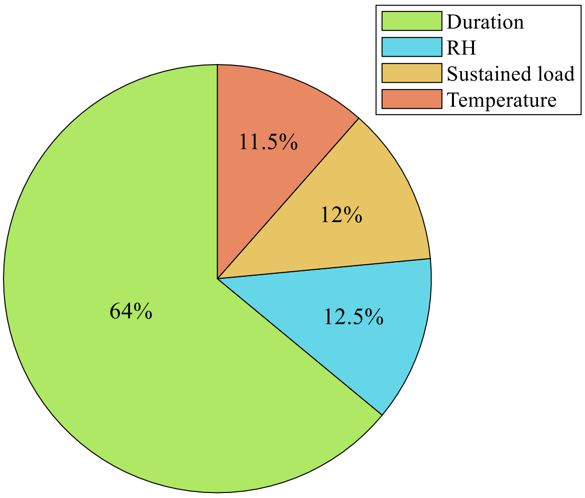


Figure 4-18. Relative importance of environmental parameters in failure mode identification

4.4.2 Durability assessment of FRP bar-to-concrete connections under different environmental conditions

In this section, a series of prediction models were established to compare their performance in predicting the bond strength and identifying the failure mode of the bond between FRP bar and concrete. The processed dataset in Section 4.2.2 used together with the selected features in Table 4-2 to identify bond degradation.

4.4.2.1 Failure mode identification of FRP bar-to-concrete bond

In order to accurately realise the bond degradation under conditioning regimes, it is important to study the failure mechanism of the FRP bar in concrete. The data shows that there are two major failure modes of pull-out (PO) and concrete splitting (SP). A majority (90%) of the samples failed at the interface of the bar and concrete due to pull-out, while only 10% failed in an SP failure mode.

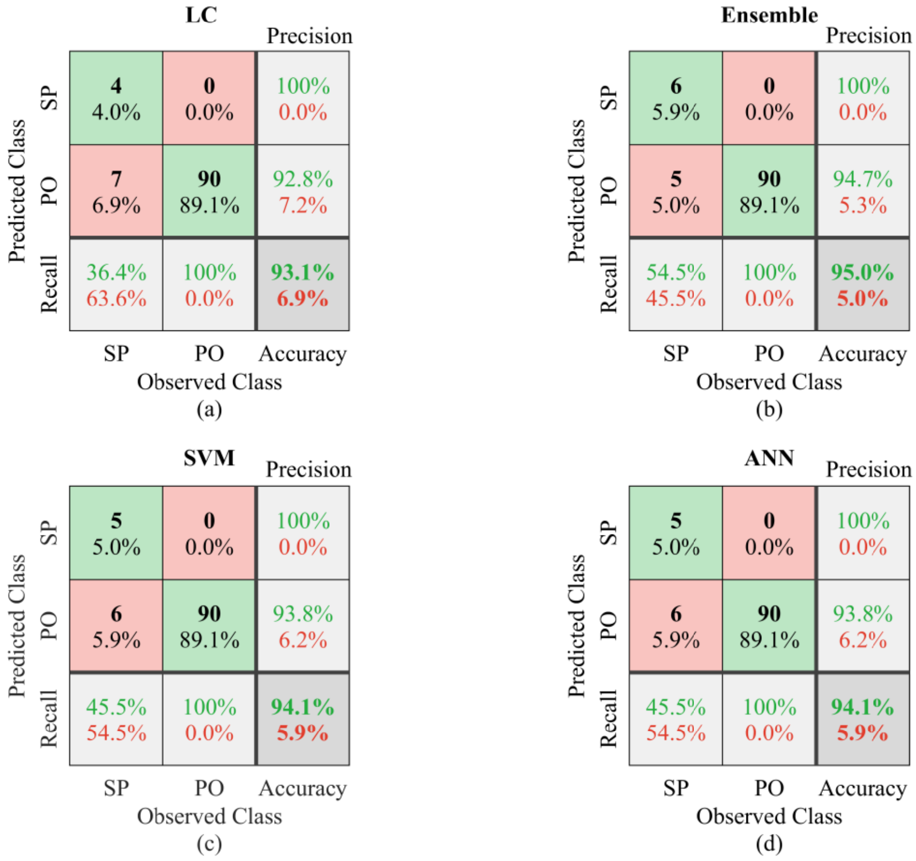
To demonstrate the classification capability of machine learning techniques and select the best prediction model, four methods were used, and their accuracies were evaluated. Since samples with SP failure modes constitute only 10% of the dataset, it is possible that during the random data split, no SP failure modes would be assigned to the test data. To solve the imbalanced classification problem, in this paper, the dataset was cross validated in a 10-fold stratified approach to guarantee that SP samples were available in all the validation folds.

To avoid manual tuning of hypermeters, in this study, Bayesian optimisation was implemented, and the best hyperparameters were selected by cross validating the training data into 10 folds. Table 4-6 shows the list of hyperparameters for each classifier along with their ranges and the optimised values. The selected values were then used to retrain and test the machine learning models.

Table 4-6. List of selected hyperparameters for optimised classification models

Model	Hyperparameter	Search range	Optimised value
LC	λ	0 – 249.37	2.86e-5
	Regularisation	Ridge, Lasso	Ridge
Ensemble	Method	Bagging, GentleBoost, LogitBoost, AdaBoost, RUSboost	AdaBoost
	Number of Learners	10 – 500	497
	Learning rate	0.001 – 1	0.984
	MinLeafSize	1 – 200	5
SVM	C	0.001 – 1000	999.95
	Kernel function	Gaussian, Linear, Polynomial	Linear
ANN	Number of layers	1 – 5	2
	Number of neurons	1 – 33	13, 1
	Learning rate	0 – 0.1	0.019
	Dropout	0 – 0.499	0.046

In order to compare the classification performance of the models, the precision, recall, and accuracy criteria were used in predicting the test dataset. The number of samples with PO and SP failure modes in the test subset was 90 and 11, respectively. Figure 4-19 shows the confusion matrices of the models. The numbers of the correctly classified observations in each method are located in the diagonal cells together with off-diagonal cells quantifying the misclassified failure modes. It is evident that all the trained classifiers provide accurate predictions on the unseen data, while the ensemble outperforms the other methods with 95% accuracy. This is followed by ANN and SVM (94.1%) and LC (93.1%). It is also seen that the SP recall values obtained by LC, SVM, and ANN are lower than the one by ensemble, which indicates that this failure mode is mostly predicted as PO. It is also evident in Figure 4-19 that the highest precision value for PO is attained by the ensemble model.



LC

Predicted Class	SP	PO	Precision
SP	**4** 4.0%	**0** 0.0%	100% 0.0%
PO	**7** 6.9%	**90** 89.1%	92.8% 7.2%
Recall	36.4% 63.6%	100% 0.0%	**93.1%** **6.9%**
	SP	PO	Accuracy

(a)

Ensemble

Predicted Class	SP	PO	Precision
SP	**6** 5.9%	**0** 0.0%	100% 0.0%
PO	**5** 5.0%	**90** 89.1%	94.7% 5.3%
Recall	54.5% 45.5%	100% 0.0%	**95.0%** **5.0%**
	SP	PO	Accuracy

(b)

SVM

Predicted Class	SP	PO	Precision
SP	**5** 5.0%	**0** 0.0%	100% 0.0%
PO	**6** 5.9%	**90** 89.1%	93.8% 6.2%
Recall	45.5% 54.5%	100% 0.0%	**94.1%** **5.9%**
	SP	PO	Accuracy

(c)

ANN

Predicted Class	SP	PO	Precision
SP	**5** 5.0%	**0** 0.0%	100% 0.0%
PO	**6** 5.9%	**90** 89.1%	93.8% 6.2%
Recall	45.5% 54.5%	100% 0.0%	**94.1%** **5.9%**
	SP	PO	Accuracy

(d)

Figure 4-19. Confusion matrices of machine learning models for the test data

4.4.2.2 Influence of environmental factors on FRP bar-to-concrete bond failure

In order to rank the contribution of environmental agents to the bond failure by the ensemble method, the importance of parameters for each tree learner in the model was evaluated, and the weighted average was computed for each variable. In each tree, the feature importance is derived by the sum of the changes in the node risk divided by the total number of branch nodes.

The risk in node i is defined as the node error multiplied by the node probability according to Eq. (4-6).

$$R_i = P_i E_i \tag{4-6}$$

where P_i is the node probability, and E_i is the node error. Figure 4-20 depicts the contribution of the environmental factors to the failure of the FRP bar-to-concrete bond. It is evident that the failure of the bond is most affected by the duration of exposure, followed by the magnitude of temperature, pH level, cyclic/continuous exposure type, salt concentration, and the concentric/eccentric test type.

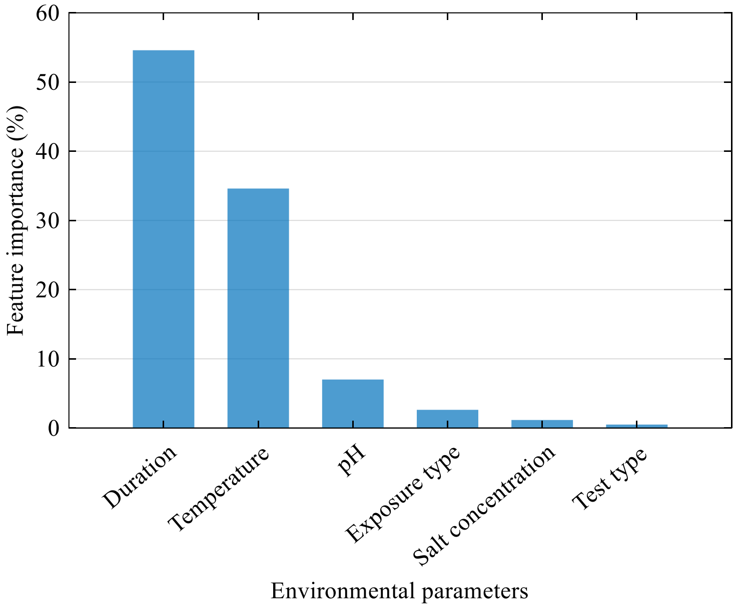


Figure 4-20. Contribution of environmental parameters to the bond failure

4.4.2.3 Failure mode prediction equation of FRP bar-to-concrete bond

In order to make practical use of the developed prediction models in engineering problems, it is important to select an interpretable machine learning model. Among the machine learning methods, linear models can provide a simple formulation. Therefore, the identified failure mode can be expressed as a linear combination of input parameters as in Eq. (4-7).

$$\mathrm{Ln}\ (\frac{f(\boldsymbol{x})}{1-f(\boldsymbol{x})}) = 3.063 + \sum_{i=1}^{14} (a_1)_i \mathrm{Ln}\ [(a_2)_i (x_i + 1)^{(\lambda)_i} + (a_3)_i] \tag{4-7}$$

where x_i denotes the input variable; i is the index of the variable; $a_1, a_2, a_3,$ and λ are the coefficients derived by the Yeo-Johnson transformation in Eq. (4-1). Table C-2 in Appendix C shows the list of the input variables with their corresponding coefficients. The failure mode is identified by substituting the values of this table in Eq. (4-7). For the PO failure mode, the output of Eq. (4-7) is positive, while for a sample failed with concrete splitting, it is negative. Therefore, there is an increase in the chance of PO failure mode when the values of input variables with positive a_1 grow, whereas the likelihood of SP failure mode increases when the values of the parameters associated with negative a_1 increase in value. Table 4-1 lists the ranges of input variables that were used to train the LC model, so it is critical that the values of the input parameters fall within these ranges in order for Eq. (4-7) to make valid predictions.

4.4.2.4 Degradation modelling of FRP bar-to-concrete bond

Bond degradation modelling is a crucial step in understanding the impact of environmental conditions on the durability of the FRP bars in concrete. The properties of the bond can be studied by performing a pull-out test as shown in Figure 4-21. For developing the prediction model of bond strength, four machine learning regression techniques were trained with 80% of the data, and their accuracies were checked by the test set.

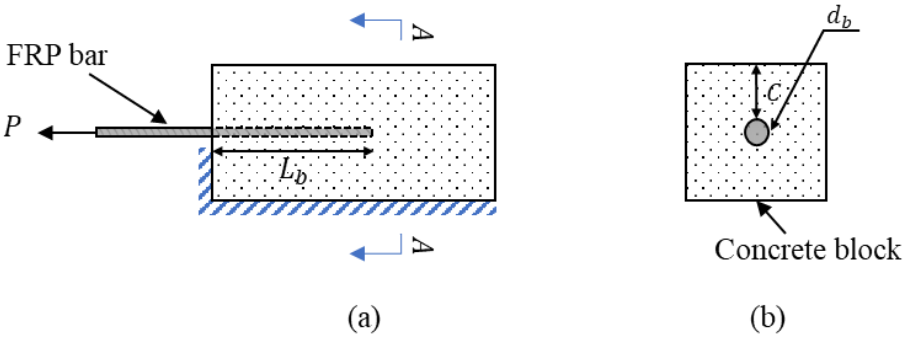


Figure 4-21. Pull-out test setup: (a) side view; (b) front view

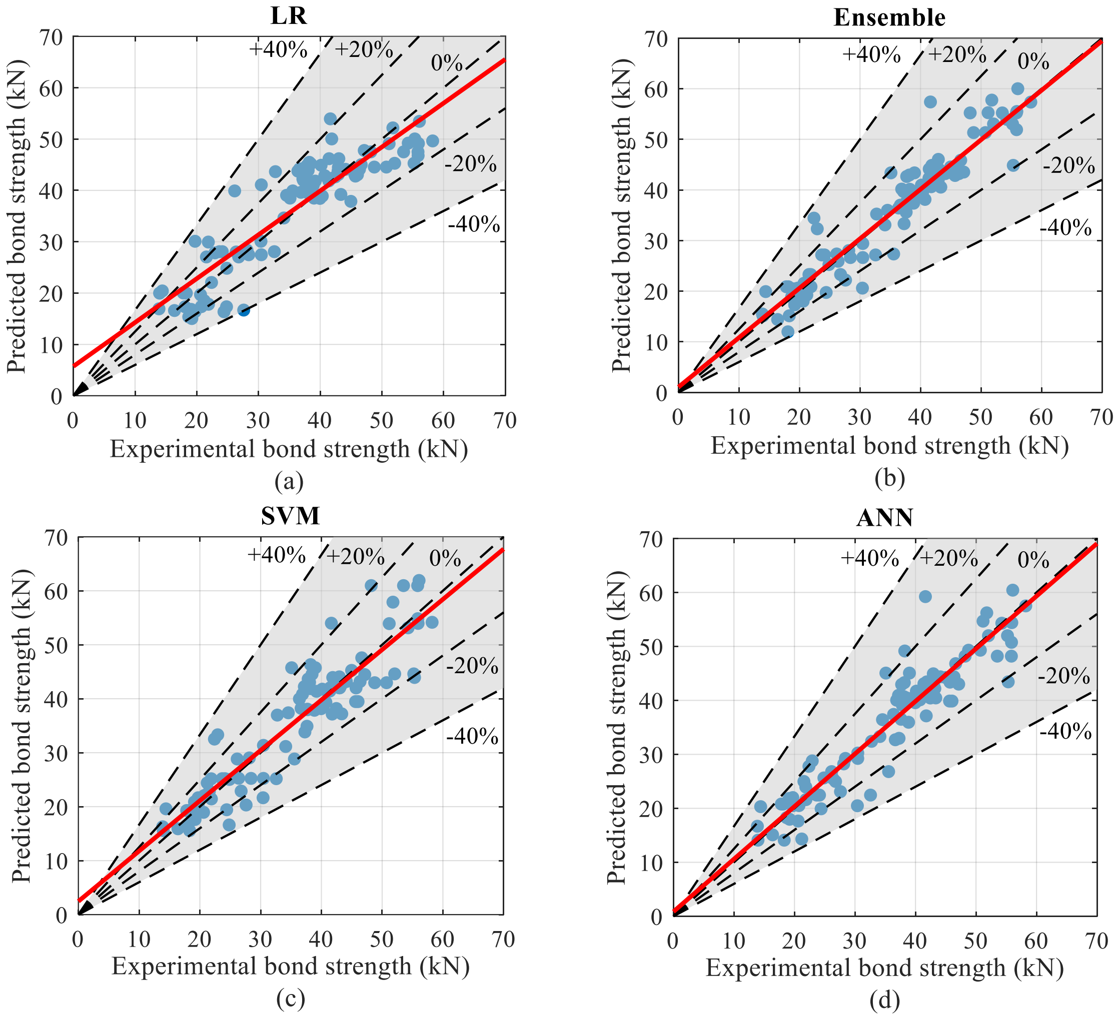
To predict the bond strength by regression analysis, the training data was cross validated through 10 folds during the Bayesian optimisation in order to confirm the generalisation ability of the models. Table 4-7 lists the hyperparameters with their optimal values for each model. The obtained values in this table were used to retrain and test the models' accuracies with the new data. The performance check of the models was attained by using RMSE and R^2 evaluation criteria. The values of these metrics are presented in Table 4-8 for the training, validation, and test sets. The prediction error and correlation values in this table imply that all the developed models are capable of predicting bond strength with an acceptable accuracy. However, the ensemble outperforms other models in making accurate predictions. The correlations between the predicated and real values in the test set are also depicted in Figure 4-22. It is seen that the correlations are distributed within ±40% of the equality line with the least skewness for ensemble and ANN models. For a detailed visual comparison, the correlation points in this figure were fitted with a red line for each model. It is evident that for ensemble and ANN, this line overlaps the equality line, indicating the higher prediction accuracy of these models. However, a gap between the fit and equality lines in the LR and SVM denotes a degree of error in the predictions. This figure implies that an ensemble of weak decision tree learners can accurately capture the relationship between the FRP bar, concrete, and environmental parameters with the bond strength.

Table 4-7. List of selected hyperparameters for optimised regression models

Model	Hyperparameter	Selected value
LR	λ	2.79e(-3)
	Regularisation	Ridge
Ensemble	Method	LSBoost
	Number of Learners	496
	Learning rate	0.0831
	MinLeafSize	17
SVM	C	0.0107
	Kernel function	Polynomial
ANN	Number of layers	4
	Number of neurons	29, 24, 21, 17
	Learning rate	0.1
	Dropout	0.0

Table 4-8. Values of evaluation metrics for the regression models

Model	Training set		Validation set		Test set	
	RMSE	R^2	RMSE	R^2	RMSE	R^2
LR	5.78	0.84	6.03	0.82	5.39	0.86
Ensemble	2.72	0.96	3.91	0.92	4.05	0.92
SVM	3.87	0.93	5.07	0.87	4.89	0.88
ANN	3.72	0.93	4.61	0.89	4.37	0.9



(a) (b) (c) (d)

Figure 4-22. Correlations between real and predicted values: (a) LR; (b) ensemble; (c) SVM; (d) ANN

4.4.2.5 Prediction of FRP bar-to-concrete bond strength by code formulations

Environmental reduction factors allow to design tensile strength of FRP bars based on composite material type and exposure characteristics. Figure 4-23 depicts the equilibrium of forces in an FRP bar embedded in a concrete prism. The applied force to the end of bar creates bond stress which acts on the bar's surface. In Figure 4-23, A_b is the cross-section area of the rebar, and f_{fu} is the tensile stress in the rebar due to the pull-out force. The proposed reduction factors in standards are applied to the ultimate tensile stress to imply the degradation of the bar. Eq. (4-8) determines the reduced bond strength (P^*) based on the reduction factors proposed in the standards.

$$P^* = C_E f_{fu} A_b \tag{4-8}$$

In this paper, the reduction factors (C_E) were obtained from AS 5100.8 [128], ACI 440.1R-15 [123], and fib Bulletin No.40 [129]. The C_E values in ACI 440 for the samples exposed to environmental conditions with carbon and glass fibres are 0.9 and 0.7, respectively, whereas AS 5100 assigns reduction factors based on the exposure classification. Therefore, for immersion in water, C_E values are 0.85 and 0.65 for carbon FRP (CFRP) and glass FRP (GFRP) composite bars, respectively; for aggressive exposures, C_E is 0.85 and 0.5 for CFRP and GFRP, respectively, which implies the vulnerability of glass fibres in harsh environments.

The long-term bond strength in *fib* is obtained by applying the reduction factor as in Eq. (4-9). In this equation, R_{10} is the standard reduction in tensile strength factor per decade; n denotes the influences of moisture condition, environmental temperature, and the desired service life of composite; and γ_f is called material factor.

$$C_E = \frac{[(100 - R_{10})/100]^{n+2}}{\gamma_f} \tag{4-9}$$

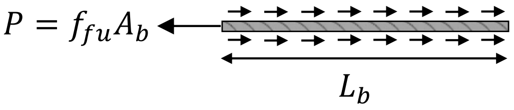
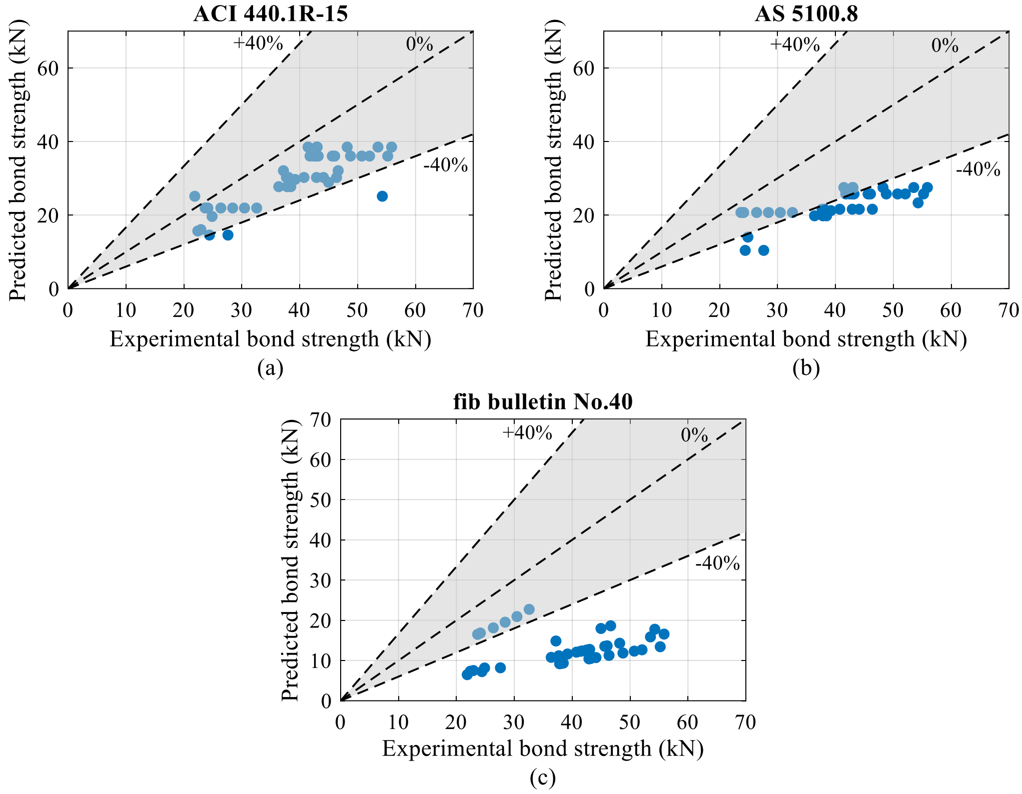


Figure 4-23. Force transfer from bar to concrete through the bond

To obtain the accuracy of code predictions, the long-term bond strength values in the test set were estimated by Eqs. (4-8)–(4-9) and depicted in relation to the experimental values in Figure 4-24. It is seen that the predicted values by standards are mostly located under the equality line, implying that the environmental reduction factors propose conservative estimations of bond strength. The lower R^2 values for AS 5100 and *fib* compared to the machine learning results in Table 4-8 indicate the low accuracy of these formulations in predicting bond strength. However, the estimations made by ACI 440 agrees well with the experimental results. There is also evident in Figure 4-24 that *fib* assigns a larger safety factor for bond strength than other guidelines, whereas the reduction factors by ACI 440 result in the least conservative estimations.



(a)

(b)

(c)

Figure 4-24. Correlations between real values and code estimations: (a) ACI 440.1R-15 (b) AS 5100.8; (c) *fib* Bulletin No. 40

4.4.2.6 Influence of environmental conditions on FPR bar-to-concrete bond strength

Due to the fact that exposure to harsh environments plays an important role in the degradation of the bond, it is critical to study the change in the bond integrity with environmental parameters. In doing so, ensemble models were loaded by a series of datasets with sequential values of the environmental variables and constant values for the remaining parameters. Figure 4-25a depicts the bond strength variation under 5000 hours of exposure to continuous W, S, AL, and AC as well as cyclic FT environments. It is seen that the exposure duration correlates with the bond degradation. The lower bond strength values of AC in relation to the other continuous conditioning types validate the detrimental impact of acid on the bond integrity. In addition, the increase in the rate of bond loss with time implies the inverse effect of the continuous exposure to chemical contents in S, AL, and AC solutions. Whitin the first 2000 hours, it is seen that the bond strength decreases the most by the FT cycles. However, the greatest loss in bond strength occurs after 5000 hours as a result of a high degradation rate in an acidic environment.

Figure 4-25b shows the bond strength variation with different pH values in red. It is seen that the bond strength values increase with the pH level for $pH < 7$, which implies that the scale of acidity contributes to the bond degradation. The inverse impact of higher pH levels on the bond is also obtained from Figure 4-25b. It is evident that for $pH > 7$, a higher alkalinity environment inversely affects the bond strength. The contribution of the saline environment to the bond loss is also seen in Figure 4-25b in blue. As the salt concentration increases, it is clear that the bond strength decreases. This means that the FRP bar-to-concrete connection is susceptible to the change in the salt concentration.

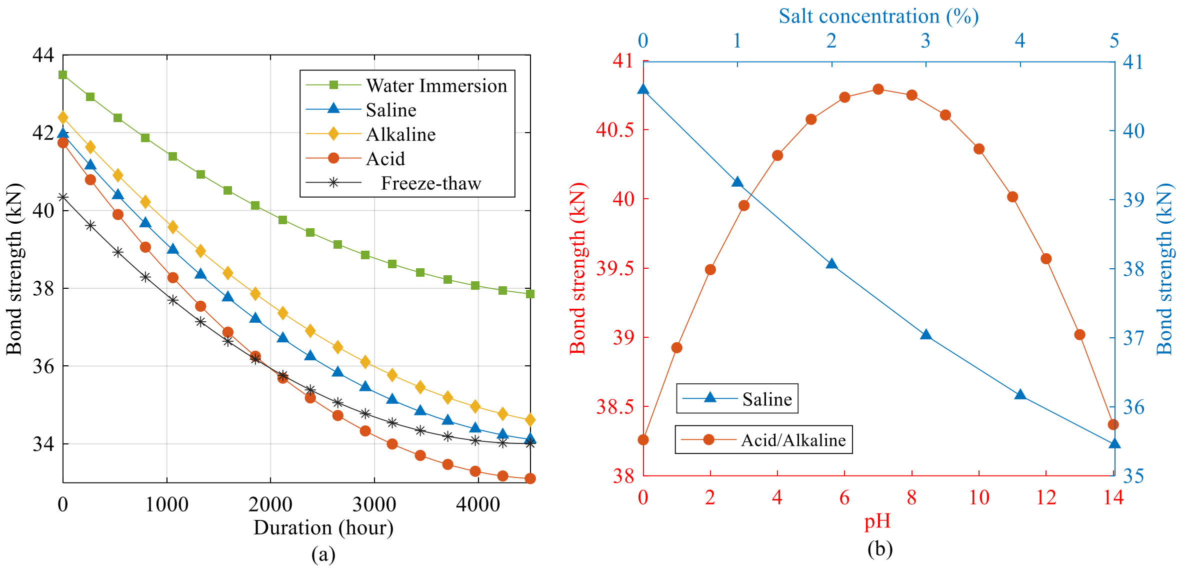


Figure 4-25. Bond strength loss: (a) impact of various continuous exposures with time; (b) impact of pH level on salt concentration variations

4.4.2.7 Strength prediction equation of FRP bar-to-concrete bond

Due to its simplicity, the LR model in Figure 4-22a was used to generate an equation for the relationship between the long-term bond strength and the geometric, mechanical, and environmental factors. Since the input variables were transformed by Eq. (4-1) to train the model, the transformed linear equation can be written as Eq. (4-10):

$$\hat{P} = 37.86 + \sum_{i=1}^{14} (b_1)_i \left[(b_2)_i (x_i + 1)^{(\gamma)_i} + (b_3)_i\right] \tag{4-10}$$

where b_1 is the coefficient of the input variable x_i derived from the linear regressor; b_2, b_3 and, γ are the coefficients obtained after applying the Yeo-Johnson transformation. The values of these coefficients for input variables are presented in Table C-3 in Appendix C. Therefore, the long-term bond strength is calculated using Eq. (4-10) and the coefficients from Table C-3 with respect to the ranges of the input variables as denoted in Table 4-2.

4.5 Conclusions

Given the impact of the service environment on the efficiency of the infrastructure system, it is important to study durability at the structural level. This can provide more detailed information on how structural components are impacted by environmental ageing. Strengthening and reinforcing concrete structures with FRP materials is a successful means of extending their service life. However, the durability of the bond between FRP and concrete is compromised due to interaction with the surrounding environment. The literature showed that there is a lack of accurate procedures for studying the durability of FRP-reinforced concrete elements. In this chapter, we developed a machine learning framework to study the long-term performance of the bond between FRP and concrete subjected to aggressive environments.

Two comprehensive databases were established by exploring the experimental data on the durability of the connections in the externally and internally bonded FRP-concrete connections. A series of machine learning techniques were explored, and a framework based on Bayesian optimisation was established to predict the failure mode and bond strength of the connections with high accuracy.

The assessment on the durability of the FRP-to-concrete adhesively bonded connection was performed based on a comprehensive database of 429 durability test results of samples under moisture condition. The detailed evaluation of prediction results showed that ANN has the best performance in predicting bond strength. The study on the effects of conditioning regimes indicated the adverse impact of various continuous and cyclic moisture environment on bond integrity. A practical equation based on the M5P algorithm was also developed to calculate the bond strength as a function of geometrical, mechanical, and environmental variables of the connections under moisture conditions. The failure mode was also identified by a series of classifiers. The prediction results showed that the ensemble model has the best performance in predicting failure modes. The study on the relative importance of environmental factors showed that exposure duration is the most critical factor, while RH, sustained load, and temperature have relatively similar impacts on the failure mode.

The durability of the bond between FRP bar and concrete was also evaluated under different environmental conditions. A total of 502 pull-out test results under ambient, water immersion, saline, alkaline, acidic, and freeze-thaw conditions were compiled and used by the developed framework to model bond degradation. A detailed evaluation of the prediction performance indicated that the ensemble outperforms the linear, SVM, and ANN models in the accurate identification of failure mode and bond strength. The importance analysis of environmental factors showed that the FPR bar-to-concrete bond failure was mostly affected by the duration of exposure. A comparison between the machine learning results and code formulations indicated the conservative estimation of bond strength in various standards. This implies that higher safety factors are used in the durability design of the FRP-reinforced concrete structures. Furthermore, the influence of different conditioning regimes showed the detrimental effect of the acidic environment on the bond integrity. Also, bond loss was linked to extreme pH levels and higher salt concentrations in environments with chemical content. Two practical equations based on linear models were formulated to mathematically capture the relationships between the failure mode and bond strength with the geometrical, mechanical, and environmental characteristics of the FRP bar and concrete.

The proposed framework in this chapter demonstrated that using machine learning can lead to reliable and practical approaches to studying the structural effects of environmental ageing. However, it is critical to recognise the study's limitations, as well as suggest potential areas for future research and improvement in the durability assessment of composite structure connections. Firstly, the study focused on specific harsh environments, including moisture, water immersion, alkaline, acidic, and freeze-thaw conditions. Future research can explore a wider range of environmental conditions, such as exposure to UV radiation, impact or cyclic loading scenarios, and combinations of different aggressive environments. Furthermore, this study primarily focused on bond strength and failure mode prediction. While these are important factors, future research can consider the time-dependant behaviour of reinforced concrete members strengthened with FRP composites, such as creep and stress relaxations. In addition to this, it is important to characterise the stress versus the relative movement between the FRP and concrete. Particularly, modelling the bond-slip relationship of the connections under extreme environmental conditions can help to understand how stress distributes within the interface over time. This can lead to more durable designs for the connections.

In the following chapter, we will look into the degradation of FPRs at the material level to see how ageing affects the performance of composite structures at smaller scales.

CHAPTER 5

Modelling Transverse Performance of FRP Materials Under Humid Environment Using Micromechanical Analysis and Machine Learning method

5.1 Introduction

Machine learning was used in the previous chapter to investigate how environment ageing affects the durability of FRP-reinforced concrete elements. It helped gain an understanding of the long-term performance of composites at the structural level. Despite providing valuable information on ageing, the developed prediction models missed the underlying interaction of the materials with the surrounding environment. The lack of understanding of performance at the material scale can create uncertainty about the long-term performance of composite structures. In this chapter, the environmental degradation of composites at microscale using computational micromechanics is investigated. The numerical modelling is used to predict the long-term properties of the FRP materials. Although numerical modelling is high fidelity and accurate, it is not as cost-effective as machine learning in computation. As a result, a machine learning-based framework is created to facilitate analysing the impact of ageing on FRPs at material scale.

In the numerical part of the study, a two-dimensional representative volume element (2D RVE) generation framework is first developed to automatically create periodic microstructures with random fibre distributions. The transverse properties of the unidirectional (UD) glass fibre-reinforced polymer composite are then determined using micromechanical finite element (FE) simulations. Numerical, analytical, and experimental results are also compared to confirm the validity of the FE modelling. Finally, a parametric study is performed to analyse the influence of different characteristics of numerical modelling on the composite transverse behaviour.

In the machine learning part of this chapter, 1,200 artificial microstructures are generated by the mentioned RVE generation framework. The FE modelling is then used to obtain the mechanical response of the FRP composites impacted by humid conditions. A data pre-processing procedure is taken to characterise the generated microstructure images using the two-point correlation (TPC) functions. Since the obtained features from TPC are large in size, the principal component analysis (PCA) technique is used to reduce the dimension and extract important features. The processed data is then used to develop an artificial neural network (ANN) model. In order to find an optimal network architecture for the prediction model, the Bayesian optimisation is used to tune the ANN hyperparameters. Finally, the contribution of the features extracted by PCA is studied by an ensemble machine learning model.

Figure 5-1 shows the framework used throughout this chapter to analyse the impact of ageing on FRP composites at microscale using machine learning.

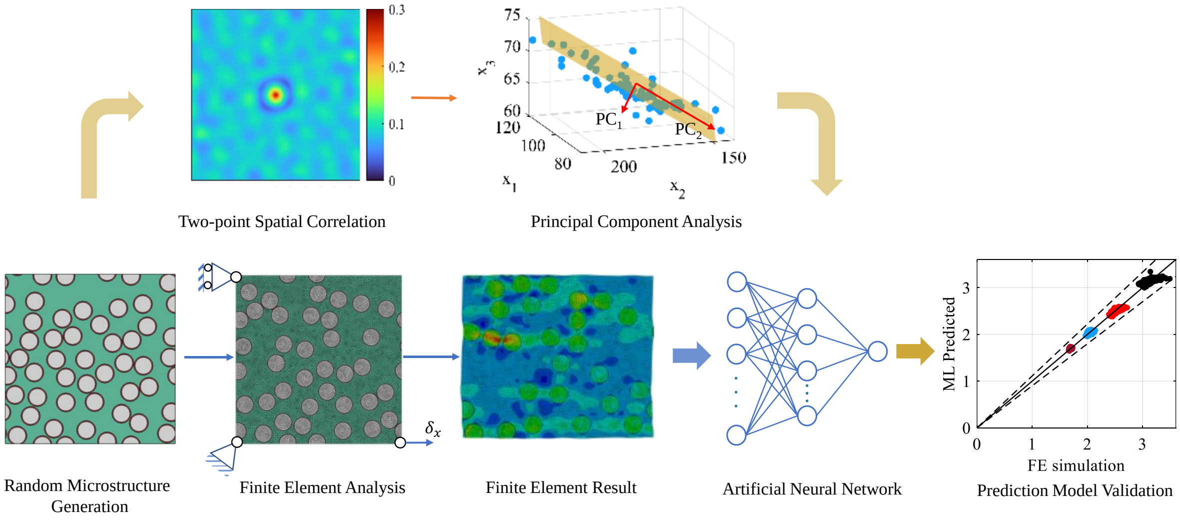


Figure 5-1. Flowchart of micromechanics-based machine learning approach to analysing ageing in FRP composites

5.2 Generation of representative volume element (RVE)

The concept of RVE is the basis for the study of composites at the microscale. A RVE of a heterogenous material is identified as the volume that is large enough to represent the mechanical response of the composite [45]. An RVE can be characterised by analysing the images of the composite microstructure. However, this method, besides being time-consuming, does not capture the periodicity of the microstructure, which is essential for applying periodic boundary conditions in the numerical analysis. Therefore, the synthetic generation of fibre distributions with random arrangements in a matrix is required. Despite the existence of numerical tools in the literature to generate RVEs [51, 130], an automatic generation of a periodic UD micromodel while including an interphase region between fibre and matrix is still lacking. Despite its negligible thickness compared to the RVE size, the influence of the interphase on the mechanical properties of the composite is acknowledged in the literature [131].

In order to generate a spatially random inclusions in a 2D matrix, the random sequential adsorption (RSA) algorithm is employed. In RSA, the fibre locations are randomly and sequentially generated within a predefined space. For each fibre, its position is checked with the position of the existing particles in the domain to avoid any overlapping between the fibres. The fibre inclusion in the RVE is continued until the desired volume fraction is achieved [50]. Figure 5-2 depicts the flowchart for the numerical procedure of the RVE generation. The algorithm inputs the fibre radius, interphase thickness, RVE dimensions, and fibre volume fraction. These values are used to locate the first coordinate of the fibre centre. In the following, the overlapping criterion is checked to ensure that no intersection occurs between the inclusions. In addition, for the fibres located on the edges of the RVE, periodicity conditions are required to apply in order to create a periodic RVE.

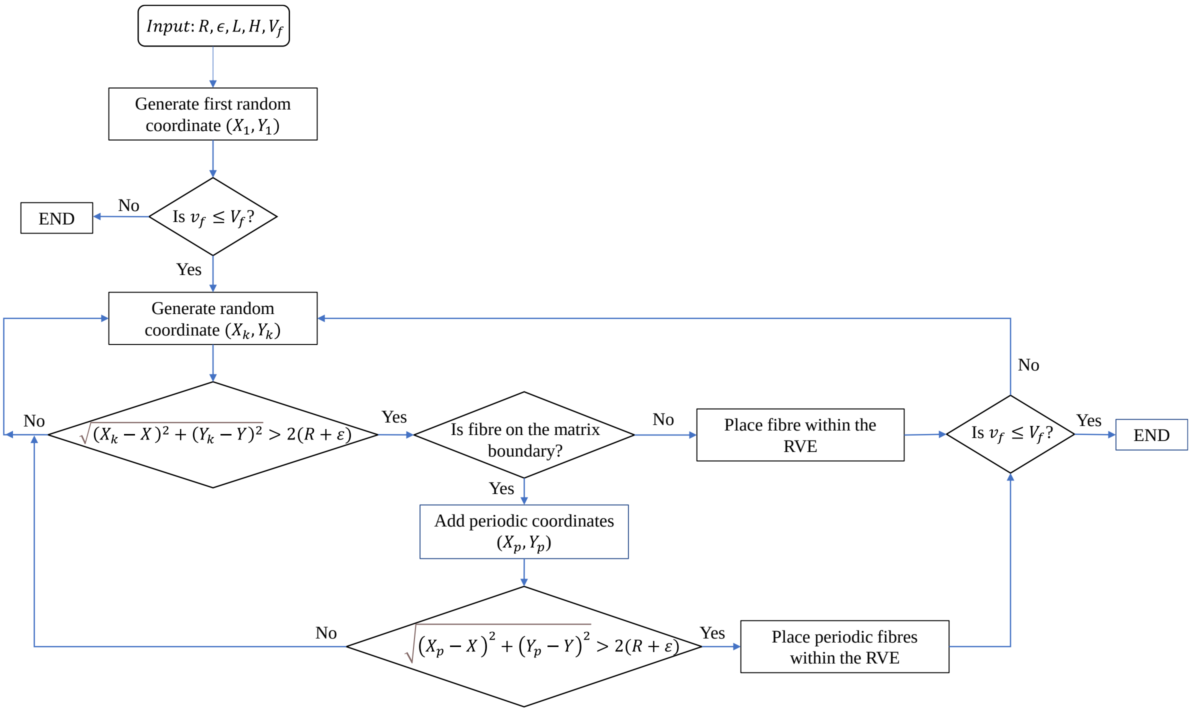


Figure 5-2. Flowchart for RVE automatic generation

5.2.1 Fibre overlap check

Since the coordinates of the fibre centres are generated randomly, it is possible that the fibres intersect each other. Therefore, it is necessary to check for overlap before placing them in the modelling space. Figure 5-3 shows a typical structure of an RVE, which includes interphase and circular fibres. The width and the height of the area are denoted by L and H. The distance between a k-th circle with coordinate (X_k, Y_k) and the other circles (d_i) is calculated as:

$$d_i = \sqrt{(X_k - X_i)^2 + (Y_k - Y_i)^2} \tag{5-1}$$

where $i = 1,2, \dots, k - 1$. With R and ε being the fibre and interphase radius and thickness, respectively, the k-th fibre is only added if $d_i > 2(R + \varepsilon)$ for all the existing fibres.

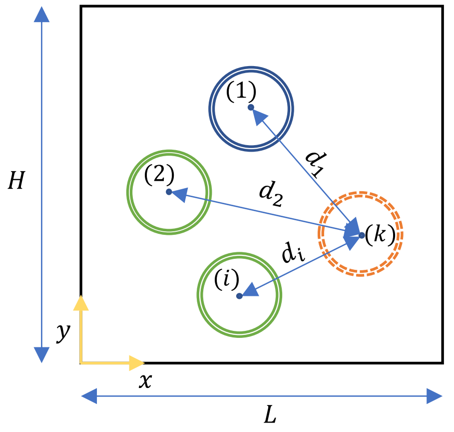


Figure 5-3. Random coordinate generation without overlaps

5.2.2 Periodicity of RVE

For the fibres that are located at the boundaries, a section of their area is placed outside the RVE. In order to apply periodicity, it is required to replicate the out-of-bounds sections on the direct opposite sides of the RVE. This also avoids the wall-effect of the edges, so that the inclusions can penetrate the boundaries and RVE is viewed as a sample of a larger area [132]. Figure 5-4 shows the method for applying periodicity to the RVE. For a fibre with the coordinate (X_k, Y_k) included in the area $L \times H$, three types of positions are possible on the boundaries, which are:

1. Fibre on the left/right edges where $|X_k - X| < (R + \varepsilon)$ or $|X_k - (X + L)| < (R + \varepsilon)$
2. Fibre on the top/bottom edges where $|Y_k - Y| < (R + \varepsilon)$ or $|Y_k - (Y + H)| < (R + \varepsilon)$
3. Fibre on the corners by one of the following conditions:
    - Fibre on the left-bottom corner, where $\sqrt{(X_k - X)^2 + (Y_k - Y)^2} < (R + \varepsilon)$
    - Fibre on the right-bottom corner, where $\sqrt{(X_k - (X + W))^2 + (Y_k - Y)^2} < (R + \varepsilon)$
    - Fibre on the left-up corner, where $\sqrt{(X_k - X)^2 + (Y_k - (Y - H))^2} < (R + \varepsilon)$
    - Fibre on the right-up corner, where $\sqrt{(X_k - (X + W)^2 + (Y_k - (Y - H))^2} < (R + \varepsilon)$

It is seen in Figure 5-4 that *Circle-1* is located on the bottom edge. The periodicity of the material requires to replicate the circle on the direct opposite edge. Therefore, $Circle - 1^Y$ is created on the top edge by translating a copy of the *Circle-1* along the y-axis by L. A similar transition is applied to the fibres on the left/right boundaries.

However, *Circle-2* is located on the right-up corner of the area, and the outside fraction of the fibre is shared with three adjacent RVEs. It is necessary to duplicate the circle along the *x*- and

y-axes and the diagonal of the area. As a result, $Circle - 2^x$ is formed along the x-axis by translating a copy of the *Circle-2* by L. $Circle - 2^y$ is a duplication of *Circle-2* along the y-axis by H, and $Circle - 2^{XY}$ is created by moving the centre of a copy of *Circle-2* by L and H along the x- and y-axes, respectively. The created circles on the edges and corners are depicted by the dashed lines in Figure 5-4.

5.2.3 Numerical generation of 2D RVEs

2D RVEs were automatically generated by developing a Python script based on the flowchart depicted in Figure 5-2 and passing it to the ABAQUS CAE software [133]. Figure 5-5 illustrates the examples of the generated RVEs with different fibre volume fraction (ν_f), diameter, and RVE size values. This demonstrates that the suggested algorithm can generalise over a wide variety of input value ranges. Figure 5-5a shows a 40×40 μm^2 RVE with a 25% fibre volume fraction and 0.5 μm interphase thickness. In Figure 5-5b, fibres with 35% ν_f are dispersed in a 150×150 μm^2 area, and a higher volume fraction of 40% for an RVE with 200×200 μm^2 area and 1.5 μm interphase thickness is depicted in Figure 5-5c.

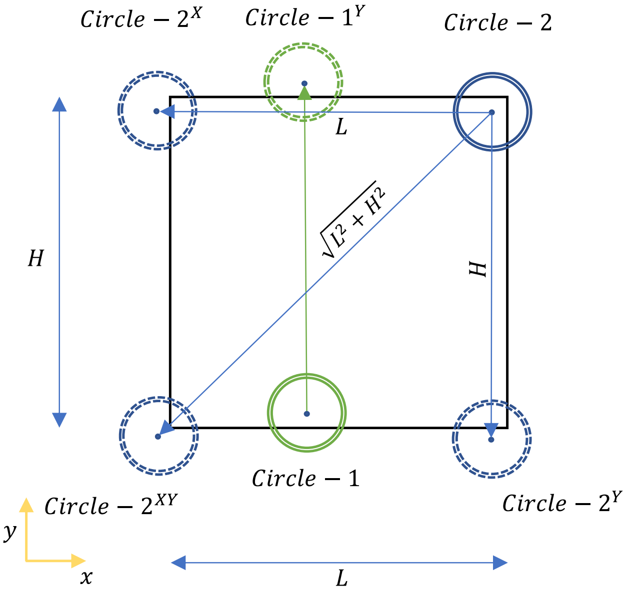


Figure 5-4. Applying periodicity of material to the RVE

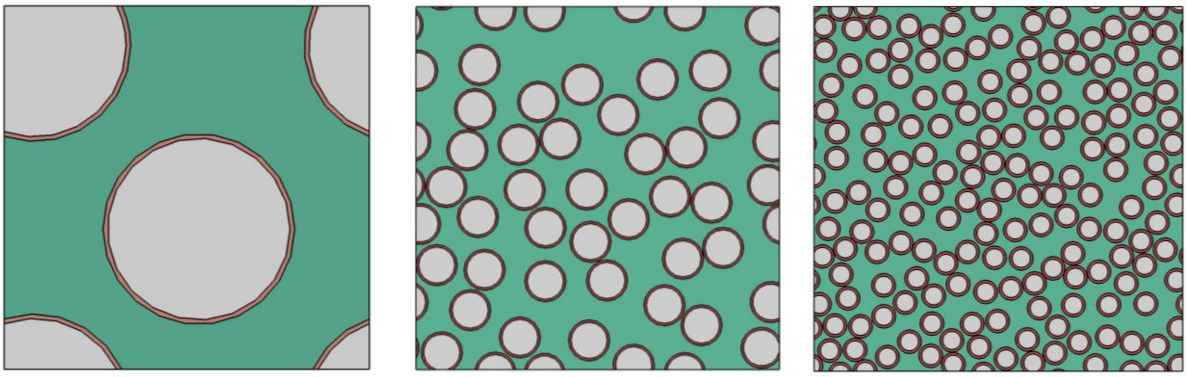

(a) (b) (c)

Figure 5-5. Typical 2D RVEs: (a) 40×40 μm^2, R = 10 μm, v_f = 25%, ε = 0.5 μm; (b) 150×150 μm^2, R = 7.5 μm, v_f = 35%, ε = 1 μm; (c) 200×200 μm^2, R = 5 μm, v_f = 40%, ε = 1.5 μm

5.3 Periodic boundary condition

Periodic boundary conditions (PBCs) are a set of constraints applied to the opposing edges of a 2D RVE. PBC implies that the shape of the opposite edges remains identical under a particular loading history [132]. Also, to maintain stress continuity across the borders, the stress vectors on the opposite edges must be in the opposite directions. The use of PBC has demonstrated its capacity to provide accurate estimations of effective mechanical properties [134, 135].

The PBC on a RVE domain Γ_{RVE} is enforced when the opposing edges are kinematically constrained. Figure 5-6a shows a 2D RVE of circular fibres where the left, right, top and bottom edges are labelled as Γ_L, Γ_R, Γ_T, and Γ_B. The periodicity implies that the edges Γ_L and Γ_T are kinematically tied to the edges Γ_R and Γ_B and the vertices N_1, N_2, and N_4. Considering N_L, N_R, N_T, and N_B as the representative nodes on left, right, top, and bottom edges, respectively, the periodicity of displacement for these nodes is written as:

$$U_{(x,y)}^{N_R} = U_{(x,y)}^{N_L} + U_{(x,y)}^{N_4} - U_{(x,y)}^{N_1} = 0 \tag{5-2}$$

$$U_{(x,y)}^{N_T} = U_{(x,y)}^{N_D} + U_{(x,y)}^{N_2} - U_{(x,y)}^{N_1} = 0 \tag{5-3}$$

where $U_{(x,y)}^{N_i}$ explains the displacement of the node N_i in x- and y-axes. Eq. (5-2) constraints the displacement of the node N_R to the degrees of freedom in nodes N_L, N_4, and N_1. Similarly, $U_{(x,y)}^{N_T}$ is related to the displacements in nodes N_D, N_2, and N_1 in Eq. (5-3). These equations are the basis for the mathematical application of PBCs to Γ_{RVE}.

The PBC requirement for applying stress equilibrium on the opposite edges of the RVE for the domain Γ_{RVE} subjected to the tensor stress $\boldsymbol{\sigma}$ is written:

$$\sigma . n_R(\Gamma_R) = -\sigma . n_L(\Gamma_L) \tag{5-4}$$

$$\sigma . n_T(\Gamma_T) = -\sigma . n_B(\Gamma_B) \tag{5-5}$$

where n_j is the unit outward normal to the edge Γ_j.

5.3.1 Definition of uniaxial load case

Commercially produced composites often have a random microstructure due to the lack of control over the fibre placement. This results in transversely isotropic properties of the composite at the mesoscale. Due to the weaker load-bearing capacity of composites in the transverse direction than the longitudinal direction, it is important to analyse the microstructure under transverse loading.

Eqs. (5-2)–(5-3) indicate that a sequence of equations can be written between all nodes on the opposite edges of the RVE to apply periodicity. In order to induce transverse tension condition, it is necessary to place certain constrains on the nodes N_1, N_2, and N_4. Figure 5-6b indicates a 2D RVE where $U^{N_1}_{(x,y)} = 0$ in order to prevent rigid body motion. A tensile uniaxial loading parallel to *x*-axis is applied to the node N_4 by defining a positive displacement $U^{N_4}_{(x)} = \delta_x$, while the node N_2 is constraint along the *x*-axis as $U^{N_2}_{(x)} = 0$. Table 5-1 shows the details of the constraints for the nodes on the vertices of the RVE for prescribing the uniaxial load.

Table 5-1. Nodal constrains to impose uniaxial deformation along *x*-axis

Direction	Node			
	N_1	N_2	$N3$	$N4$
x-axis	$U^{N_1}_{(x)} = 0$	$U^{N_2}_{(x)} = 0$	$U^{N_3}_{(x)} \neq 0$	$U^{N_4}_{(x)} = \delta_x$
y-axis	$U^{N_1}_{(y)} = 0$	$U^{N_2}_{(y)} \neq 0$	$U^{N_3}_{(y)} \neq 0$	$U^{N_4}_{(y)} = 0$

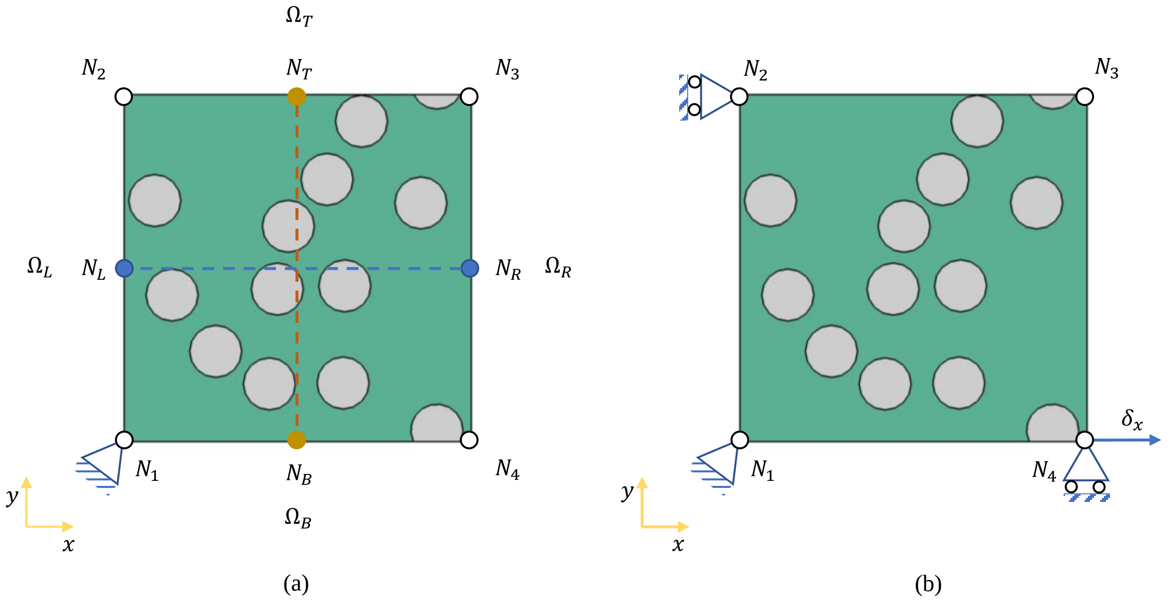


Figure 5-6. A typical 2D RVE: (a) Edge and corner nodes for applying PBC, (b) defining uniaxial tensile displacement

5.4 Numerical homogenisation

In order to address the effective mechanical properties of heterogenous materials, homogenisation methods are used. There have been various studies to characterise the effective performance of composites. One of the most commonly used homogenisation techniques for composites with periodic microstructures is the Eshelby-based mean-field homogenisation model [136]. This method provides a cost-effective, yet accurate, solution for predicting the effective mechanical properties [137]. Figure 5-7 shows a macro- to micro-scale transition of a heterogenous material via computational homogenisation. It is assumed that the difference between the RVE and the micromodel length scales is sufficiently large so that the principle of the separation scales can be applied. [138].

Considering the first-order homogenisation and zero boundary fluctuation model [139], the macro strain ε_{macro} is applied to the boundary of the RVE and the displacement vector is written as:

$$U_{micro} = \varepsilon_{macro}\Delta x \tag{5-6}$$

where Δx is the relative position vector. The macroscale stress tensor is also computed following the Hill-Mandel principle as a volume-averaged stress at the microscale, which gives:

$$\sigma_{macro} = \frac{1}{\Gamma} \int_{\Gamma_{RVE}} \sigma_{micro} \, d\Gamma \quad (5\text{-}7)$$

The average stress and strain are related with the stiffness tensor *C* in the form of:

$$\sigma_{macro} = C \varepsilon_{macro} \quad (5\text{-}8)$$

By choosing a unit strain value, the components of the *C* matrix are obtained by averaging the stress tensor over the elements. Commercial software such as ABAQUS is able to compute the stress and volume for each element.

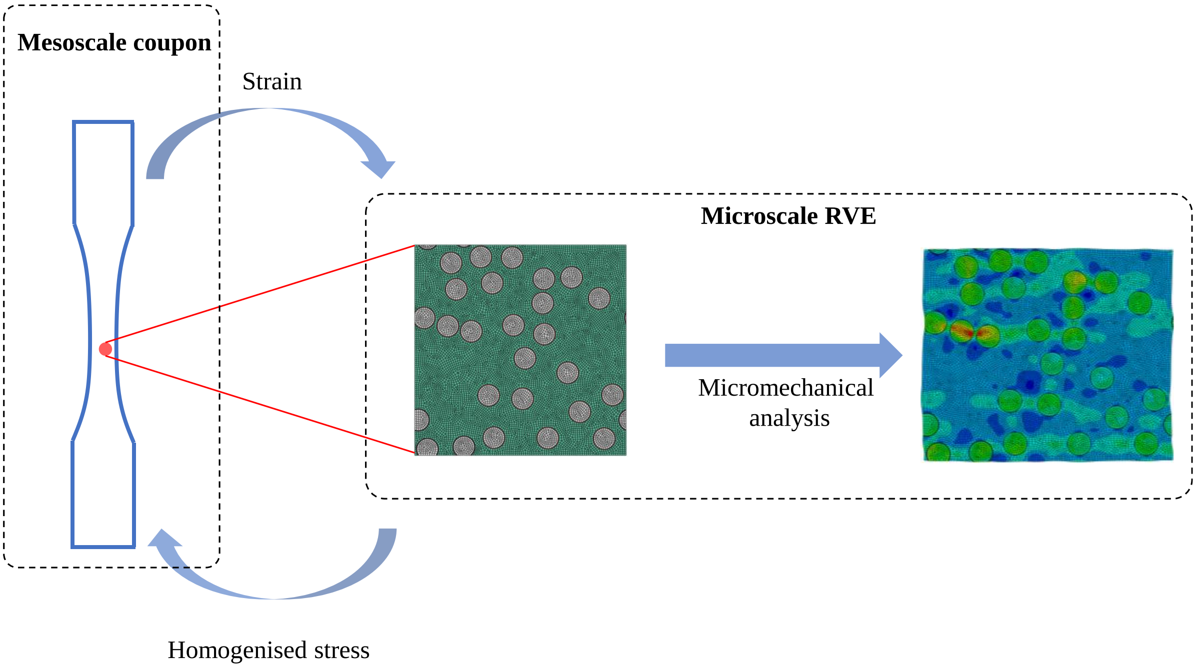


Figure 5-7. A representation of computational homogenisation with the scale transitions

5.5 Constitutive material properties

In this section, the properties of the materials used to numerically model the mechanical response of the composite are elaborated. The unidirectional composite laminate used for the current study is based on a thermoset epoxy type resin reinforced with 20% volume fraction of Type-E glass fibres. The elastic properties of the resin were obtained by three-point bending tests [140, 141]. The elastic modulus of the fibre was directly taken from the information provided by the manufacturer [142]. For the interphase region, atomic force microscopy (AFM) was used between the 50 nm to 500 nm distance from the fibre to identify its elastic modulus [143]. The Poisson’s ratio for the fibre and matrix were obtained from the literature [144]. Riaño et al. [143] performed a sensitivity analysis of the Poisson's ratio for the interphase and assumed a value equal to the matrix's Poisson's ratio.

In order to study the impact of humid condition, the above materials were subjected to a tropical climate for 76 weeks, where the average humidity was 82% ± 4%, and the average temperature was between 20.3 °C and 33.4 °C. AFM identified a degradation of the initial matrix modulus for about 16%. Furthermore, the interphase area expanded by a factor of two compared to its original size [61, 141]. The details of the mechanical properties introduced to the numerical modelling are presented in Table 5-2. In this table, t_0 and t_{76} indicate the mechanical properties at the initial state and after 76 weeks of exposure to the humidity.

Table 5-2. Properties of the UD composite constituents used in numerical modelling

	Matrix		Fibre		Interphase area		
	E_m (GPa)	ν_m	E_f (GPa)	ν_f	E_i (GPa)	ν_i	e_i (nm)
t_0	1.43	0.4	72	0.25	1.9	0.4	500
t_{76}	1.2	0.4	72	0.25	1.1	0.4	1000

5.6 Numerical approach

Within the numerical framework of this chapter, the UD composite components, namely, fibre, matrix, and interphase are considered homogenous and linear materials with isotropic mechanical behaviour and plain strains. In addition, the interface is assumed to be perfectly tied to the fibre and matrix [61, 143].

For the micromechanical analysis, fibre arrangement of the UD composites is commonly assumed to be regular with square or hexagonal layouts [145]. Although this assumption simplifies the analysis, it misses the understanding of the material behaviour influenced by the randomness of the fibre distributions. Therefore, in this study, numerical analyses are performed on UD composites with random fibre arrangement and periodicity on the boundary edges.

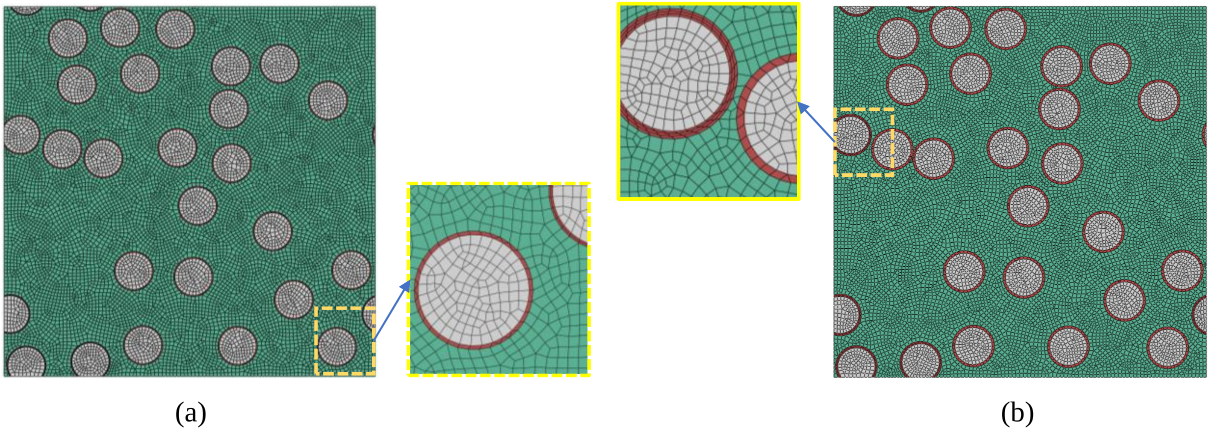
The RVE generation framework proposed in Section 5.2 was used together with the details of the material mechanical properties in Table 5-2 to numerically model the microstructure of the UD composites. The influence of the RVE characteristics on the transverse response of the composite is numerically analysed using FE modelling, and the numerical results are validated by comparing them to existing experimental and analytical studies.

5.6.1 Validation of finite element model

A comparison between the numerical, experimental, and analytical results was conducted to confirm the validity of the proposed FE models. Table 5-3 presents the details of the various parameters used to simulate the transverse tension in ABAQUS. The mechanical properties of the matrix, fibre, and interface for modelling were also obtained from Table 5-2. The ageing effect was considered in this table by introducing degraded mechanical properties to the composite elements and modifying the thickness of the interphase region. Two numerical models for initial (t_0) and aged (t_{76}) states were developed using the corresponding values from Table 5-2. Figure 5-8 illustrates the FE models with periodic random fibre distributions used to simulate the tensile test for the unaged and aged composite samples.

Table 5-3. Parametric values for numerical simulation

Parameter	Value
RVE size	150×150 μm^2
Fibre diameter	15 µm
Fibre volume fraction	20 %
Element type	CPE4
Element size	1.5 µm
Boundary condition	PBC


(a)

(b)

Figure 5-8. RVE models for numerical validation: (a) unaged sample (t_0), (b) sample after 76 weeks of exposure to humid conditions

The results obtained from numerical modelling in this study are presented in Table 5-4. They are compared with the experimental [141] and analytical studies as well as the numerical findings by Riaño et al. [61]. In the experimental study, 2 mm thick UD composite plates with 20% volume fraction of E-glass fibres and a two-component epoxy resin were moulded and cured in a cycle of 2 hours at 60 °C followed by 2 hours at 120 °C and 2 hours at 140 °C. The plates were exposed to natural weathering condition for 76 weeks in Danang, Vietnam with an average relative humidity of was 82% ± 4%, and the average temperature in a range of 22.6 °C (± 2.3 °C) – 29.8 °C (± 3.6 °C). To characterise the mechanical properties, the composite samples of 40×15×2 mm^2 size were tested within a three-point bending configuration in the transverse direction using a MTS tensile machine. Figure 5-9. Three-point bend test setup

shows the schematic view of the test setup. A constant loading nose speed of 1 mm/min was applied to the plate according to NF EN 2746 [140].

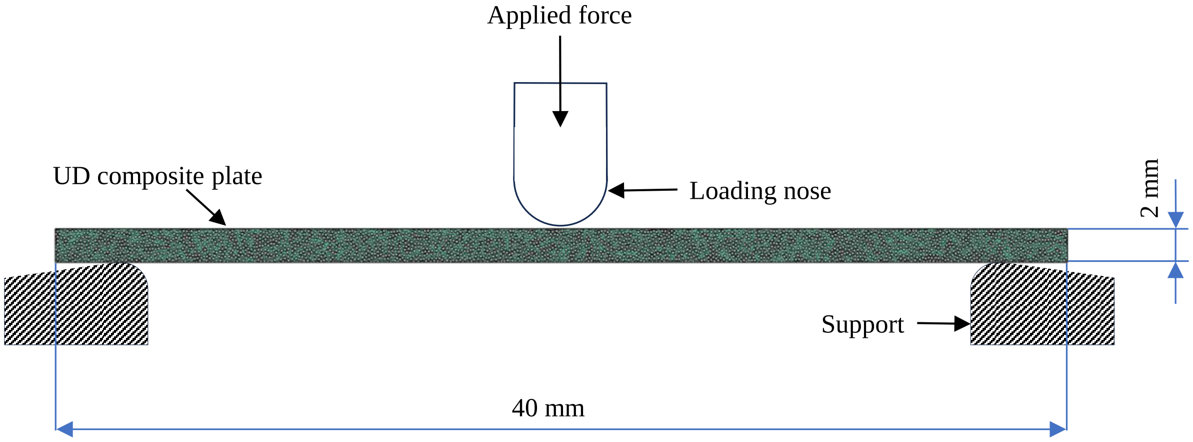


Figure 5-9. Three-point bend test setup

In this study, two analytical models called Ruess [146] and Hashin and Shtrikman [147] were used to obtain the effective elastic modulus. In the Reuss model, transverse loading is applied to the RVE, and it is assumed that the stress values for fibre, matrix, and composite are the same. Instead, for the Hashin and Shtrikman, the lower bound for elastic modulus (E^{HS-}) is calculated as:

$$E^{HS-} = \frac{9K^{-}G^{-}}{3K^{-} + G^{-}} \tag{5-9}$$

where

$$\begin{gathered} K^{-} = K_m + \cfrac{v_f}{\cfrac{1}{K_f - K_m} + \cfrac{3v_m}{3K_m + 4G_m}} \\ G^{-} = G_m + \cfrac{v_f}{\cfrac{1}{G_f - G_m} + \cfrac{6(K_m + 2G_m)v_m}{5G_m(3K_m + 4G_m)}} \end{gathered} \tag{5-10}$$

In Eq. (5-10), K_m and G_m denote the matrix bulk and shear moduli, respectively. Similarly, K_f and G_f are the corresponding moduli values for the fibre.

The values of Young's modulus derived from different methods are presented in Table 5-4. It is seen that the values presented by the numerical and analytical models are consistent with the experimental outcomes. Figure 5-10 provides a detailed comparison of the results. In this figure, the difference between the experimental results and the numerical and analytical estimations is calculated as:

$$Difference\ (\%) = \frac{E_{Experimental} - E_{Numerical/Analytical}}{E_{Experimental}} \times 100 \tag{5-11}$$

where $E_{Experimental}$ denote the modulus from the experiments, and $E_{Numerical/Analytical}$ indicates the results obtained by either numerical or analytical models. It is seen in Figure 5-10 that the Reuss model estimates the modulus value with a minimum error of 12.7% for the initial condition. However, for the aged composite, the FE simulation provides results with a 1.3% difference from the experimental findings. One reason for the overestimation of modulus in the initial stage by the FE model is that the defects present in the composite were not taken into account in the numerical modelling. An experimental study by three-point-bending test, on the other hand, can capture these defects influencing the composite global behaviour. A good agreement between the numerical and experimental results after 76 weeks of exposure can be related to the fact that the effect of matrix and interphase degradation is greater than the impact of the composite defects on the elastic modulus.

The von Mises stress fields in the RVEs are also depicted for initial and aged states in Figure 5-11. It is seen that due to the periodicity on the boundaries, the opposite edges have the same deformations. Figure 5-11 displays the stress contours for the unaged and aged models, respectively. In order to highlight the impact of ageing, the stress intensities in both figures were measured by the normalised stress scale of the unaged model. It is implied from this figure that higher values of stress are assigned to the RVE in Figure 5-11a compared to the aged model in Figure 5-11b. This is related to the higher stiffness of the matrix in the initial stage than in the aged model. In addition, the increase in size and drop in the mechanical properties of the interphase region contribute to the decrease in stress level of the aged model. The values are normalised in Figure 5-11 in order to avoid displaying stress values that are too high. This is due to the fact that no damage initiation criteria were considered in the modelling. As a result, the stress values passes the glass fibre composite ultimate stress in transverse direction, which is 30–40 MPa [148].

Table 5-4. Young’s modulus values for unaged and aged composites obtained from experimental, numerical, and analytical studies

Young’s modulus (GPa)					
Sample	Experimental result	Numerical result	Analytical result		Riaño et al. [61]
			Reuss	HS^{-}	
t_0	2.060	2.508	1.779	2.52	2.52
t_{76}	2.033	2.061	1.494	2.12	2.080

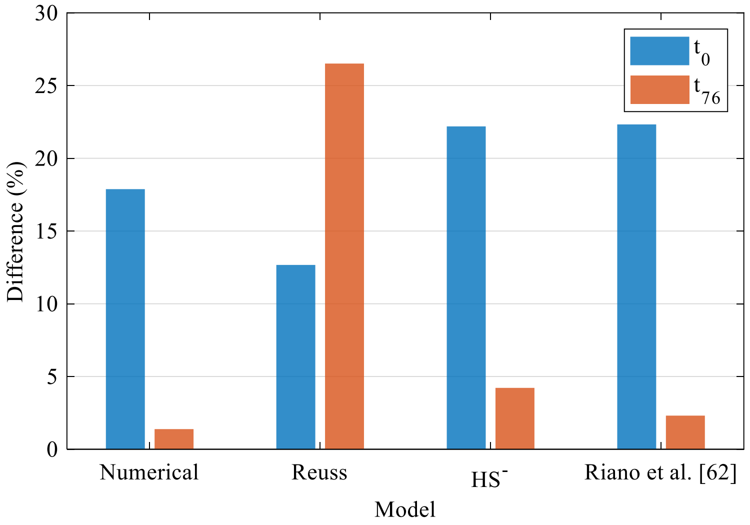
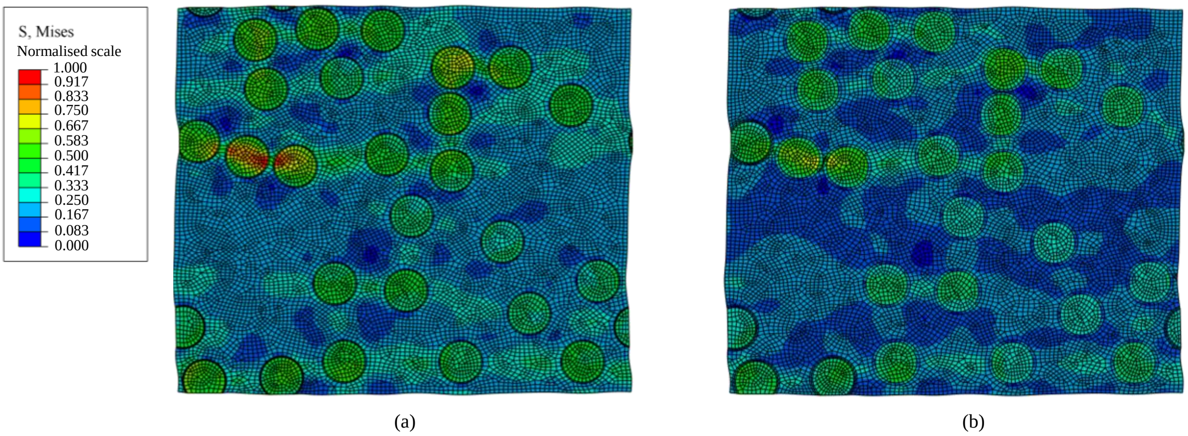


Figure 5-10. Difference between the numerical and analytical estimations and the experimental results



(a)

(b)

Figure 5-11. Von mises stress contour: (a) unaged stage, (b) 76-week aged stage

5.6.2 Parametric study of the influencing factors on the composite transverse behaviour

UD fibre-reinforced composites provide high strength in the longitudinal direction. However, their transverse strength is limited by the fibres, as they provide low strength in this direction. This results in stress concentration at the fibre and matrix interface, leading to micro-cracking and eventually fracture of the material. Therefore, the presence of fibres compromises the transverse strength of the matrix. This necessitates investigating the performance of the composites in the transverse direction.

A parametric study was performed in this section to characterise the influence of a series of parameters on the transverse properties of the UD fibre-reinforced composites. For all the RVEs analysed in this section, the mechanical properties were obtained from Table 5-2 for the composite materials at the initial stage (t_0).

5.6.2.1 Influence of the mesh

In order to demonstrate the influence of mesh, a microstructure with random fibre distribution was generated. It is a 2D RVE with a 150×150 μm^2 section area and a 20% fibre volume fraction subjected to periodic boundary conditions. Four different elements (CPE3, CPEM6, CPE4, CPE8) were chosen, each with different combination of triangular or quadrangular geometric orders. The geometric order determines the type of interpolation between the nodes, which is used to describe the shape of the element. ABAQUS offers triangular, quadrangular, and quadrangular-dominated shape elements. As a result, quadrangularly dominant geometries were selected for CPE4 and CPE8. For the RVEs with CPE3 and CPEM6 type elements, triangular shapes were chosen.

In order to check the influence of the element type and size, the simulations were performed with a mesh size ranging from 0.5–6 μm. Figure 5-12 shows the change in the calculated elastic modulus for various element types. The dashed line describes the reference modulus, which is calculated by averaging the convergent moduli of various element types. It is seen that the elements with triangular shapes diverge from the reference line, implying that global mesh size greatly affects this type of elements. For the quadrangular-dominated shape elements, the maximum convergence error of 4.7% is assigned to CPE8, while it is 1.0% for the CPE4 element. Therefore, the CPE4 elements with a size of 1.5×1.5 μm^2 were selected to avoid convergence problems. Figure 5-13 depicts an example of a 2D RVE with ν_f = 20% meshed by the CPE4 elements.

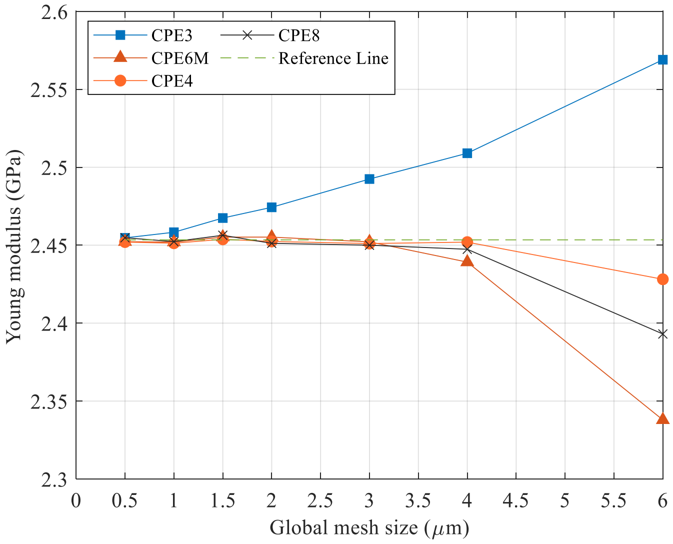


Figure 5-12. Influence of element size and type on the Young's modulus

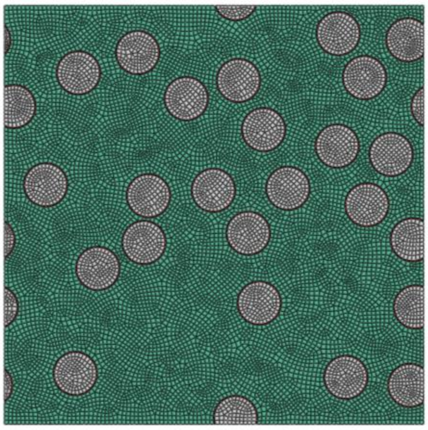


Figure 5-13. A 2D RVE composite with 20% fibre volume fraction meshed by CPE4 elements

5.6.2.2 Influence of the boundary conditions

The influence of the boundary conditions on the elastic modulus can be associated with the volume fraction of the matrix inclusions in the composite [149]. Therefore, the impact of boundary conditions on the elastic behaviour of the RVE was analysed for a range of fibre volume fraction values from 10% to 48%. Two different RVE models were analysed for each volume fraction, using either kinematic uniform boundary conditions (KUBCs) or PBCs. KUBCs are a set of kinematic coupling constraints where the group of nodes on the boundaries are constrained to the degrees of freedom of a reference node. As a result, the applied displacement field on the boundary creates uniform strains inside the RVE.

The KUBC was applied to the RVE using an ABAQUS plugin called Homtools [150]. By this method, the macroscopic strain is introduced as the average strain in the RVE, which is equal to the degree of freedom of the reference point. The macroscopic stress is processed as a ratio of the reaction forces at the reference point to the RVE area.

Figure 5-14 compares the results of the effective elastic modulus of the models with various fibre volume fractions, which were analysed by applying the aforementioned boundary conditions. It is seen that the modulus value increases with the volume fraction in both scenarios. Given that the fibre is stiffer than the matrix, increasing the volume fraction of fibres contributes to the stiffness of the microstructure. Figure 5-14 clearly shows that, as the fraction of the fibre volume grows, the gap between the elastic modulus values for the two boundary condition types increases. The errors between the results are 2.95% and 21.97% for 10% and 48% volume fractions, respectively. The trend lines also indicate that elasticity increases with a higher ratio for the models with KUBCs than the ones with PBCs.

The application of KUBC by Homtools in ABAQUS requires the manual selection of all edge objects in the RVE. This is cumbersome for periodic microstructures with high fibre volume fractions. In addition, this method does not provide a real stress state at the edges of the RVE. For this study, a Python script was developed to automatically apply PBC to the 2D RVEs and perform the post-processing to estimate the homogenised elastic modulus. Therefore, PBC was selected for the further analyses in this chapter.

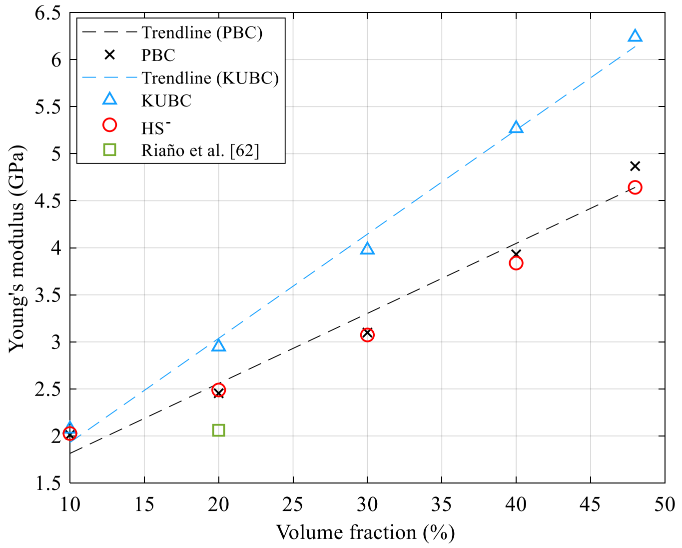


Figure 5-14. Influence of BC with fibre volume fraction on Young’s modulus

5.6.2.3 Influence of RVE size

The parametric study of the mesh in Section 5.6.2.1 calibrated the element size to 1.5×1.5 μm^2 in order to obtain converged finite element results. It is also important to select an RVE size large enough so that the effective mechanical properties do not change when the RVE dimensions are increased. A meaningful parameter for investigating the impact of RVE size on the homogenised elasticity is the ratio of the RVE length to the fibre diameter [151]. Although a ratio greater than 8 has been suggested in the literature [152], a parametric study was conducted to determine the ideal size. The plotted values in Figure 5-15 indicate the change in the Young's modulus for the models with 20% fibre volume fraction and different RVE length to fibre diameter (L_{RVE}/D_{fibre}) ratios. It is seen that consistent results are obtained for $L_{RVE}/D_{fibre} \geq 10$. Therefore, a ratio of 10 was selected for the further numerical analyses.

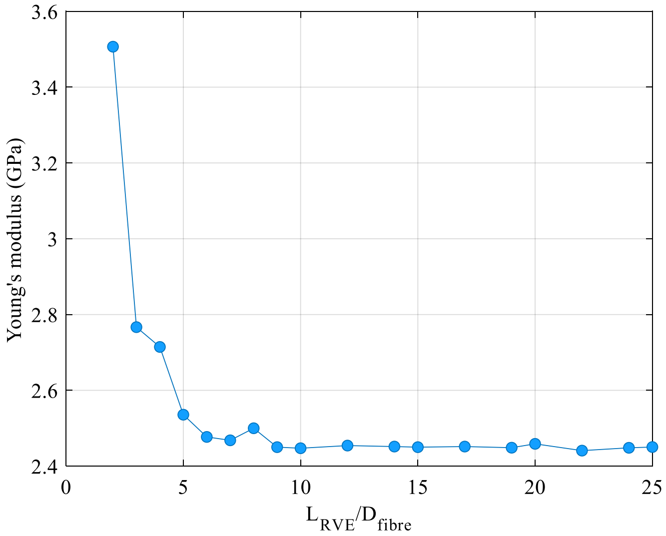


Figure 5-15. Influence of RVE length /fibre diameter on Young's modulus

5.6.2.4 Influence of the random fibre distribution

In order to study the impact of the fibre arrangement randomness on the homogenised response of the microstructure, 60 different micromodels with a variety of fibre volume fractions were analysed. Figure 5-16 shows examples of four RVEs with 20%, 30%, 40%, and 48% fibre volume ratios. An RVE size of 150×150 μm^2 with a 15 μm fibre diameter was chosen for all simulations. The change in the elastic modulus with ν_f for the micromodels is presented in Figure 5-17a. For each fibre volume ratio, 15 models were analysed. In the case of FE modelling, it is seen that the response in models with higher volume fractions is scattered compared to models with lower numbers of fibres inside the RVE area. A maximum volume fraction of 48% was attained in generating the RVEs using the framework in Section 5.2 This is due to the limitations of the RSA algorithm in generating fibre locations while avoiding fibre overlaps. Figure 5-16a also shows that the values of Young's modulus found by the HS$^-$ model are the same for each ν_f. This implies that the analytical model is not able to capture the influence of RVE geometric features on the transverse behaviour.

Figure 5-17b shows the standard deviations (SDs) of the micromodel responses to the random fibre distributions. It is seen that the SD increases with the fibre volume fraction. The SD values are 3.3%, 3.5%, 7.6%, and 9.1% for the composites with 20%, 30%, 40%, and 48% ν_f values, respectively. This implies a greater influence of the random distribution on composites with higher fibre volume ratios.

The influence of fibre distributions on the stress concentration is analysed by the local stress fields in the RVE. The normalised stress contours for two models with the same volume fraction ($\nu_f = 48\%$) and different distributions of fibres are displayed in Figure 5-18. It is seen that stress concentrations are present in regions where the fibres are close together. Comparing the stress plots shows that the local stress values generated by the fibre arrangements in Figure 5-18a are greater than those in Figure 5-18b. This indicates that the values and locations of the stress concentrations, which are sources of damage, depend on the distribution of fibres in the matrix.

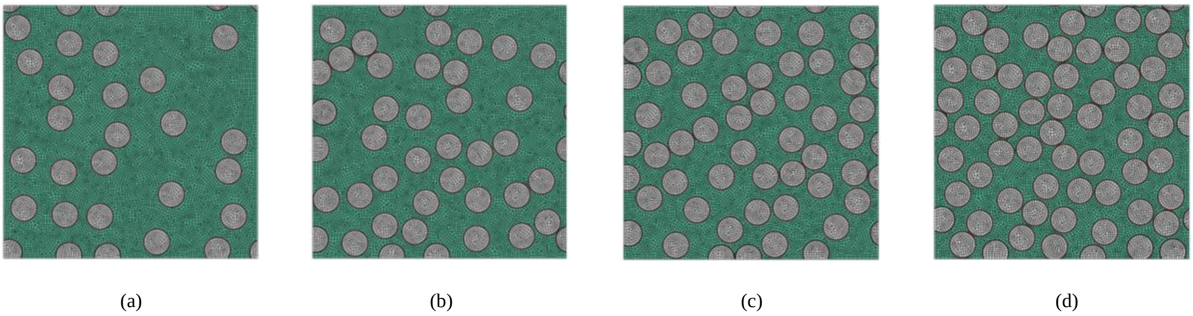

(a) (b) (c) (d)

Figure 5-16. 2D microstructures with random distributions: (a) ν_f =20%, (b) ν_f =30%, (c) ν_f =40%, (d) ν_f =48%

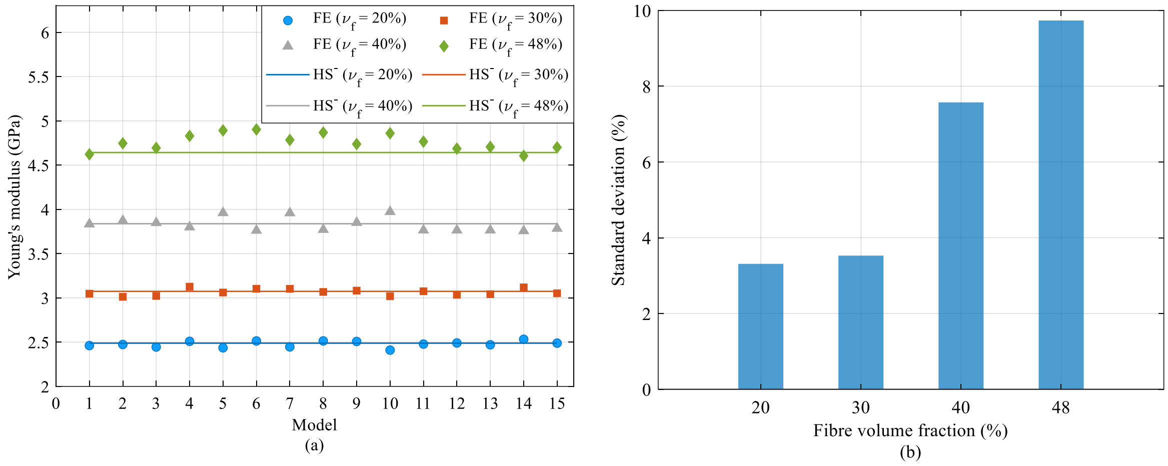


Figure 5-17. (a) Young's modulus values for random fibre distributions, (b) Standard deviations of Young's modulus for different fibre volume fractions

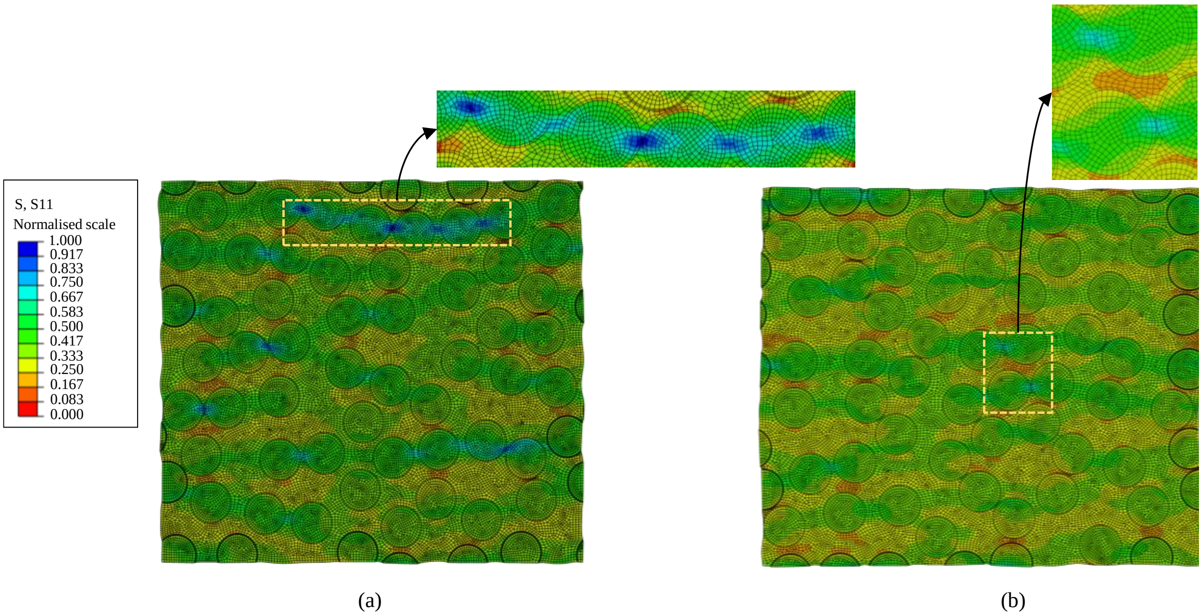


Figure 5-18. Stress concentrations in RVEs with ν_f =48%: (a) high local values, (b) low local values

5.6.2.5 Influence of the interphase characteristics on the composite transverse behaviour

The influence of the interphase region's properties on the transverse behaviour of UD composites is studied to demonstrate its impact on the elastic modulus. Interphase is a key element in determining the crack propagation and, subsequently, the rupture mechanism of the composite [153]. Considering the role of the interphase in transferring the stress, it can protect the fibres from environmental degradation and deviate the cracks [154]. Therefore, the overall performance of the composite is affected by the quality of the interphase area.

The impact of the interphase zone on the elastic behaviour of the composite was studied by modelling the RVEs with various interphase thickness and fibre volume fraction values. Figure 5-19 depicts the Young's modulus of the models for different fibre volume fractions with the interphase thickness values varying from 0 to 1000 nm. The simulations were performed with consistent fibre distribution and L_{RVE}/D_{fibre} values for each volume ratio. It is seen that the elastic modulus is directly correlated to the interphase thickness, indicating that composites become more rigid as the interphase thickness increases. It is because the interphase used for the simulations is tougher than the matrix. Therefore, as the thickness increases, the soft matrix is gradually replaced by the hard interphase, resulting in a tougher composite. It is also observed that the influence of thickness on elastic modulus increases with volume fraction.

The difference in numerical responses of the models with and without the interface are also presented in Figure 5-20, which is calculated in percentage as:

$$Difference\ (\%) = \frac{E_{model\ with\ interpahse} - E_{model\ without\ interpashe}}{E_{model\ with\ interphase}} \times 100 \qquad (5\text{-}12)$$

It is seen in Figure 5-20 that the gap between the numerical responses increases with ν_f. Specifically, as fibre volume rate grows, the proportion of the interphase region in the matrix increases, and its impact on the overall elastic behaviour of the composite becomes significant.

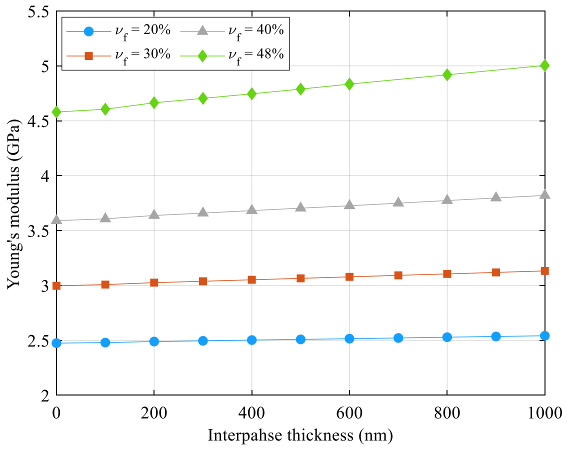


Figure 5-19. Influence of interphase thickness and fibre volume fraction on Young's modulus

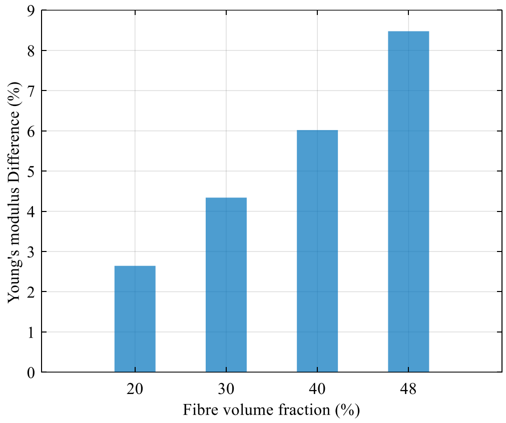


Figure 5-20. Young's modulus difference between models with and without interphase area for various fibre volume fractions

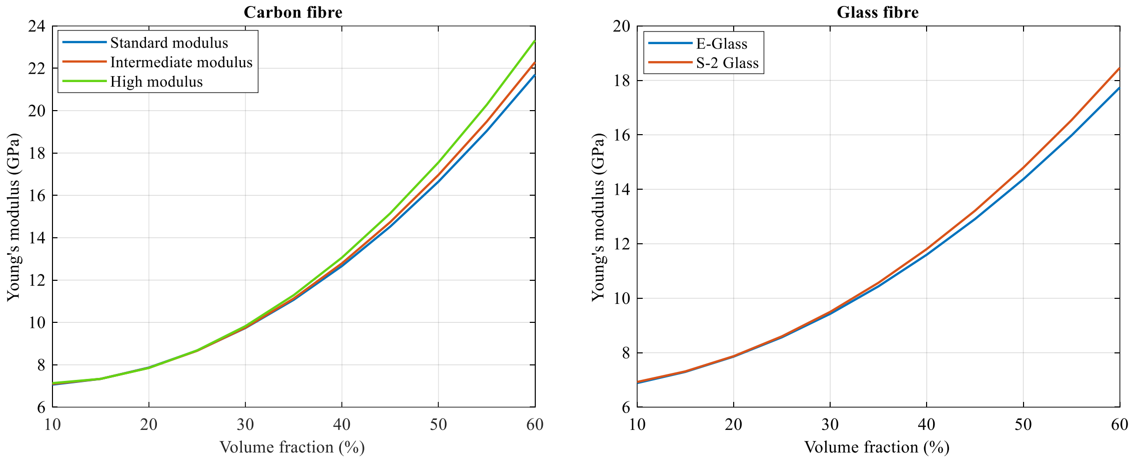
5.6.2.6 Influence of fibre characteristics on the composite transverse behaviour

In Sections 5.6.2.1–5.6.2.5, we were able to study the influence of various parameters on the composite transverse behaviour. In these analyses, the mechanical properties of the fibre and matrix were set using the values provided in Table 5-2 in the parametric study. In this section, the influence of different fibre types on the transverse performance of the composite is investigated. Glass and carbon are the common types of fibres used in composite manufacturing. 30 RVEs with varying volume fraction and modulus values for the glass and carbon fibres were modelled and analysed to obtain the composite transverse properties. In order to investigate the influence of fibre modulus on composite performance, the information of glass and carbon fibres with varying stiffness were obtained from the manufacture catalogues [155-157]. The E-glass and S-2 glass fibres, together with standard, intermediate, and high modulus carbon fibres, were used in the FE modelling. Table 5-5 presents the details of different fibre types.

Figure 5-21 depicts the change in the effective transverse moduli for the glass and carbon fibre-reinforced epoxy matrix composites. Figure 5-21a shows the results of the composites modelled by carbon fibres with different transverse moduli. It is seen that higher transverse properties are assigned to composite with stiffer fibres. Figure 5-21b shows the RVEs modelled by glass fibres. It is seen that S-2 glass material contributes to a higher Young's modulus in composites compared to E-glass fibres. Both figures show that the fibre volume fraction is correlated with the composite transverse properties. Also, as the volume fraction increases, the difference between the composite modulus values grows as well. This demonstrates the role of fibres in modifying the transverse properties of composites at higher volume ratios. From these figures, it is implied that the properties of the reinforcement material affect the composite's transverse behaviour.

Table 5-5. Material properties used to simulate the influence fibre on the composite transverse performance

Fibre Material	E_f (GPa)	ν_f
E-glass	72	0.25
S-2 glass	87	0.25
Standard modulus carbon	230	0.2
Intermediate modulus carbon	300	0.2
High modulus carbon	500	0.2



(a) (b)

Figure 5-21. Influence of fibre type with different volume fractions on composite transverse modulus

5.7 Microstructure generation

In order to develop a micromechanics-based machine learning model, it is first necessary to generate microstructural data. Using the proposed framework in Section 5.2, square microstructures with $L_{RVE}/D_{fibre} = 10$ were generated. The parametric study in Section 5.6.2 demonstrated the influence of fibre volume fraction in combination with the fibre random distribution and the interphase on the elastic behaviour of the composite. Therefore, a total of 1,200 microstructures with varying v_f and interphase thickness were created. Table 5-6 shows the details of the generated data. For each combination of fibre volume fraction ranging from 10% to 40% and three interphase thickness values of 0, 500, and 1000 nm, 100 RVE models with random fibre distributions were generated.

Each RVE was saved as an image with 570×570 pixels of resolution, and its corresponding effective elastic modulus was obtained using the finite element simulations. The fibre, matrix, and interphase region were all modelled as homogenous and linear materials with isotropic mechanical behaviour and plain strains [57, 61, 143]. The material properties of the constituents were obtained from Table 5-2 for the aged composite (t_{76}). The geometry of the RVEs were meshed with CPE4 elements of 1.5×1.5 μm^2 size to avoid convergence problems during the modelling. The PBCs were applied to the edges of the RVE. Finally, a tensile uniaxial displacement loading parallel to x-axis was imposed by introducing a positive displacement δ_x to the node N_2 of the model as shown in Figure 5-6b.

Table 5-6. Details of the generated microstructural images

Fibre volume fraction (%)	Interphase thickness (nm)	Number of the generated microstructures
10	0, 500, 1000	300
20	0, 500, 1000	300
30	0, 500, 1000	300
40	0, 500, 1000	300

5.8 Data pre-processing

Pre-processing the generated data into a format that is suitable for machine learning is an initial step in developing efficient prediction models. In this section, the relevant microstructure geometric features are extracted and processed from the generated data.

5.8.1 Digital representation of composite material

A composite microstructure usually contains a series of geometric descriptors such as size, volume fraction, arrangement, and relative location of phases. The Statistical distribution of these features characterises the composite internal structure. The geometry of the composites with random inclusions has been defined with various statistical measures such as Voronoi polygon areas, neighbouring fibre distances and orientations, and Ripley's K function in the literature [49]. However, these methods are not able to provide a unique description of a composite microstructure.

Among the various statistical characterisation methods for materials, n-point spatial correlations have been shown to precisely quantify the geometry of the composite with random inclusions [158, 159]. The 1-point correlation, as the simplest form of the spatial correlation, identifies the probability of finding the given material phase in the microstructure, which determines the volume fraction of the material. Two-point correlation (TPC) functions have been successfully used in the accurate property prediction and reconstruction of composite materials [88, 160]. The TPC function f_r^{np} quantifies the probability of finding the local states of *n* and *p* at the start and end of a vector of length *r*, randomly placed in a microstructure RVE image with *S* pixels. It is written as:

$$f_r^{np} = \frac{1}{S}\sum_{s=1}^{S} m_s^n m_{s+r}^p \tag{5-13}$$

where *s* is the spatial position of the pixel. For a vector of length *r* with its tail located at pixel *s*, m_s^n is equal to 1 if the material phase *n* is present at *s* and 0 if any other phase is located in this spatial position. With the assumption of periodic RVE, *S* represents the total number of placements for the vector of length *r* in the microstructure. For *n*=*p*, Eq. (5-13) is described as auto-correlation, which states the probability of finding a specific phase *p* at the material. However, $n \neq p$ denotes a cross-correlation function for Eq. (5-13).

For a two-phase composite material with interphase region, nine TPC functions are defined as:

$$f_r^{np} = \begin{bmatrix} f_r^{FF} & f_r^{FM} & f_r^{FI} \\ f_r^{MF} & f_r^{MM} & f_r^{MI} \\ f_r^{IF} & f_r^{IM} & f_r^{II} \end{bmatrix} \tag{5-14}$$

where *F*, *M*, and *I* denote the fibre, matrix, and interphase phases, respectively. Considering that the resulting f_r^{np} in Eq. (5-14) comprises a very large dataset, various techniques in the literature have investigated the interdependencies between the TPCs. It has been shown that for a two-phase material, only one of the four TPC functions is independent [161]. In the case of a microstructure with N distinct local states, $\frac{1}{2}(N(N-1))$ functions are independent [162]. A study by Niezgoda et al. [158] also showed that for a periodic microstructure of *N* phases, *N-1* correlations are independent. Therefore, following this research, calculating only two of the nine TPC functions in Eq. (5-14) are sufficient to obtain the remaining functions in a three-phase material. Considering the convolution nature of Eq. (5-13), discrete Fourier transforms (DFTs) were used to accelerate the calculations of the TPC functions [163].

Figure 5-22 shows an example of a microstructure with a 30% fibre volume fraction and its corresponding TPC plots for f_r^{FF}, f_r^{FM}, and f_r^{FI}, respectively. The centre of the contour in Figure 5-22b corresponds to $r = 0$ and has an intensity value of 0.3 which is equal to the fibre volume fraction. Figure 5-22c illustrates the cross-correlation function between fibre and matrix. This contour implies the probability of finding a vector of length r, starting at the fibre's centre and ending at the matrix phase. Similarly, Figure 5-22d depicts the intensity values for the vectors with the head and tail in interphase and fibre phases, respectively. These contours provide useful information regarding the volume fraction and the ratio between the fibre diameter and the RVE length [88].

A three-phase composite material used in this study comprises a total of 324,900 pixels and 649,800 two-point spatial correlations. Hence, developing a machine learning prediction model from all the TPCs of the 1,200 images is a computationally expensive task. One possible approach is to reduce the dimension of the data for the machine learning application. In this study, the high-dimensional data is projected onto lower-dimensional space using the principal component analysis (PCA) method.

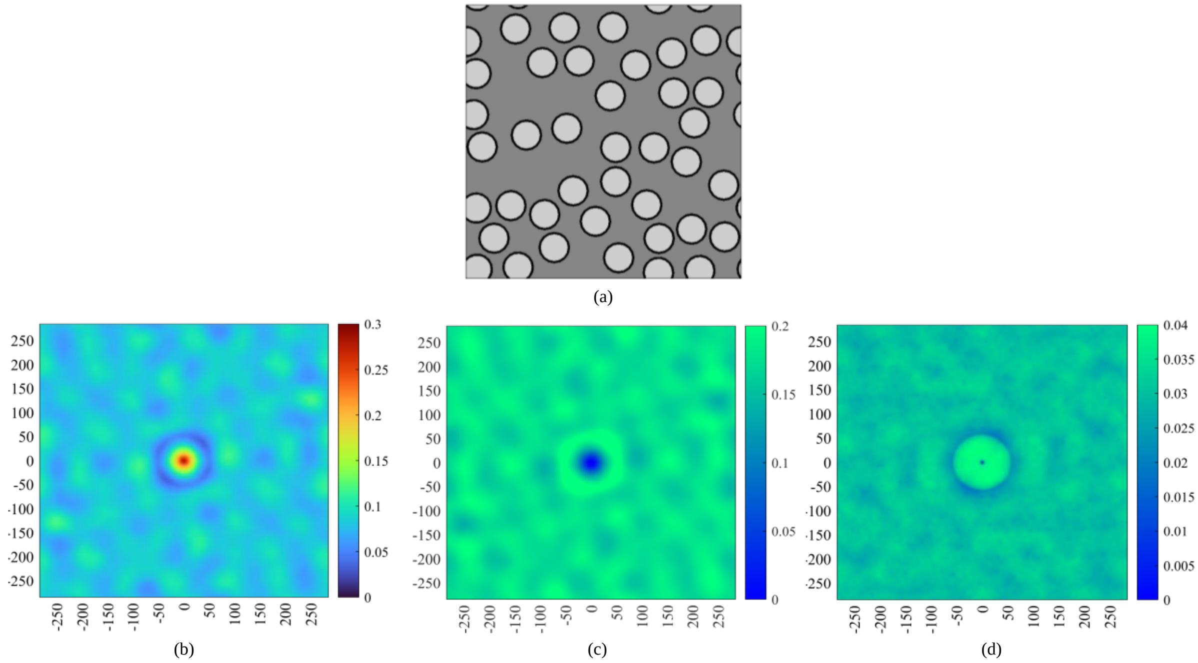


Figure 5-22. Top: 2D UD composite microstructure with 30% fibre volume fraction; Bottom: TPC function contour plots: (b) f_r^{FF}, (c) f_r^{FM}, (d) f_r^{FI}

5.8.2 Principal component analysis

Principal component analysis (PCA) is a data compression technique that has been vastly used in machine learning applications [88]. It is done by projecting the n-dimensional data onto a space with k orthogonal principal axes, where for each axis, the variance of the projected data is maximum. These axes are numbered in descending variance order. In this case, the first principal component (PC_1) is derived by decomposing the original data along the eigenvector with a direction that results in the highest variance in the projected data. Subsequently, the following principal components (PCs) are defined as the eigenvectors that are orthogonal to the preceding ones and the variances of the data along them are in a descending order [164]. Once the eigenvectors are derived, the data are projected into the space with new principal axes.

Figure 5-23 illustrates an example of projecting data with three variables onto a new space with two principal components. PC_1 and PC_2 are two vectors that define the highest variances in the data. Figure 5-23b reveals that the data variance along the PC_1 axis is larger than that along the PC_2 axis.

The mathematical principle of the PCA is based on the singular value decomposition of the data matrix A with the size $n \times p$, which contains the two-point correlation functions as:

$$A = \begin{bmatrix} fr_{n=1,p=1}^{FF} & \cdots & fr_{n=1,p=649,800}^{MM} \\ \vdots & \ddots & \vdots \\ fr_{n=649,800,p=1}^{FF} & \cdots & fr_{n=649,800,p=649,800}^{MM} \end{bmatrix} = USV^T \tag{5-15}$$

In Eq. (5-15), S is a $p \times p$ diagonal matrix with the eigenvalues of the principal components. U is a $n \times p$ orthonormal matrix with the columns representing the principal axes of A. Also, V is a $p \times p$ matrix which contains the orthogonal eigenvectors of the matrix A^T. Once the matrix A is decomposed into matrices of eignen values and eigenvectors, dimensionality reduction is performed by reconstructing the matrix A by the first k principal components, where $k <$ max (n, p). The new matrix A' is written in a reduced dimension as:

$$A'_{(n \times k)} = U_{(n \times k)} S_{(k \times k)} V^T{}_{(k \times k)} \tag{5-16}$$

The principal components are ranked based on their corresponding eigenvalues. Therefore, the first component has the greatest variance. In order to quantify the amount of variance captured by each component, the cumulative variance of the PCs is calculated. Figure 5-24 shows the change in the variance of the data with the PC number. It is seen that the first two PCs capture about 98.5% of the variance in the data, indicating that they contain the most prominent geometrical features of the microstructure. The presence of redundant features in the data can compromise the prediction model's performance as they cause overfitting and increase the complexity of the model [106]. Reducing the number of parameters by selecting the most important features using PCA is a useful approach to improving efficiency of the machine learning model [165].

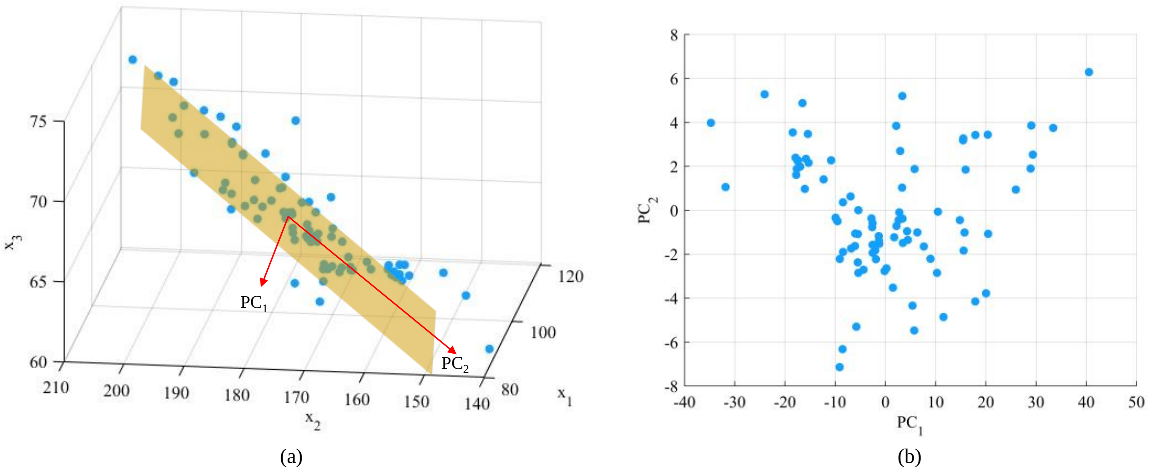


Figure 5-23. (a) PCA projections: (a) Data of three variables with the selected PCs, (b) projected data onto the new coordinate system with principal axes

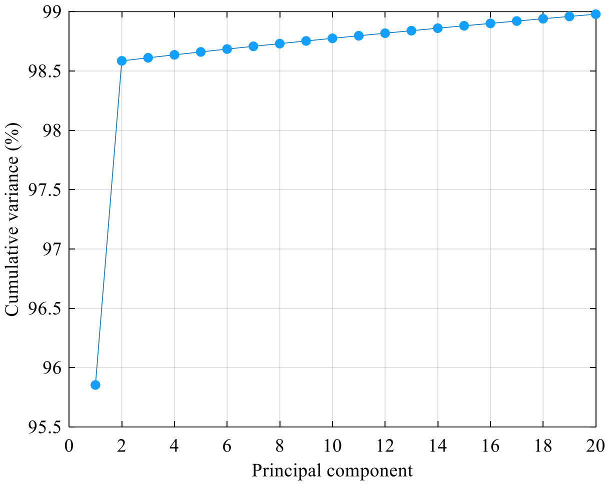


Figure 5-24. Cumulative variance of input data for different PCs

5.9 Machine learning approach

In this study, artificial neural networks (ANNs) method was used to perform a regression task in predicting the homogenised elastic modulus from the extracted features of the composite microstructures. Despite classical machine learning, ANN has become a most widely used model in theory and practice owing to its capability to manage big-data analytics [166]. Because of this, it can represent both complex and simple models, and its calculation cost can be adjusted to meet the problem requirements [167]. In order to develop and evaluate an ANN prediction model, 1,200 square RVE images with varying ranges of physical, mechanical, and environmental properties were created. Table 5-7 presents the list of the features in the dataset with their corresponding ranges. Environmental properties include information on the relative humidity (R_H), temperature, and exposure duration of the humid condition. In addition, each RVE was saved as an image with 570×570 pixels of resolution, and its corresponding transverse modulus output (E_{22}) was obtained using by the FE simulations using Abaqus. For all of the FE models, the fibre, matrix, and interphase region were considered as homogenous and linear materials with isotropic mechanical behaviour and plain strains [57, 61, 143]. The geometry of the RVEs were meshed with CPE4 elements of 1.5×1.5 μm^2 size to avoid convergence problems, and the PBCs were applied to the edges of the RVE. The data with reduced dimensions were then randomly divided into 80% training and 20% test sets. Moreover, to confirm the model's generalisation ability and avoid overfitting, the training data was cross validated by 10 different folds. Considering that the efficiency of a machine learning model depends on the hyperparameter values, it is necessary to explore the best combinations of hyperparameters to minimise the prediction error. In this study, Bayesian optimisation was used to automatically tune the ANN hyperparameters. Figure 5-25 shows the framework for the development and evaluation of ANN. It is seen that 10-fold cross validation (CV) is integrated with Bayesian optimisation to search for an architecture with minimum prediction error. In the model evaluation, the accuracy of the selected ANN is validated by the test data.

Table 5-7. Details of the features in the generated database

Feature type	Category	Description	Unit	Range
Input parameter	Physical descriptor	RVE image	pixel	570×570
	Mechanical property	E_f	GPa	12.97–72
		E_m	GPa	1.2–4.28
		E_i	GPa	1.1–4.28
	Environmental property	RH	%	50–100
		Temperature	°C	20–70
		Duration	week	0–76
Output parameter	Mechanical property	E_{22}	GPa	1.65–8.28

5.9.1 Machine learning model development and validation

The ANN model with an efficient architecture was obtained by searching for the best hyperparameters using Bayesian optimisation. Table 5-8 presents the list of hyperparameters, search ranges, and the optimal values. In this table, activation function indicates the shape of the function used to transform the weighted sum of the input to an output during FF; λ denotes the generalisation term used in the loss function equation; and weight initializer describes the functions used to sample the weights from either a uniform or normal distribution by Glorot or He functions, respectively [168, 169].

The tuning process of the hyperparameters is also shown in Figure 5-26a. Each point in this figure shows the minimum of the estimated cost function for the ANN model after it has been trained using the hyperparameters discovered in each iteration. The best point in this figure implies the set of hyperparameters from which the minimum loss is derived. Based on this figure, a minimum error of 0.0014 was obtained at iteration 10 during the hyperparameter tuning. Figure 5-26b depicts the learning curve of ANN model during training with the selected hyperparameters. This figure shows a steep decrease in the mean squared error (MSE) within the first 10 epochs. It is also seen that the MSE gradually decreases and converges to a minimum error value after 100 epochs.

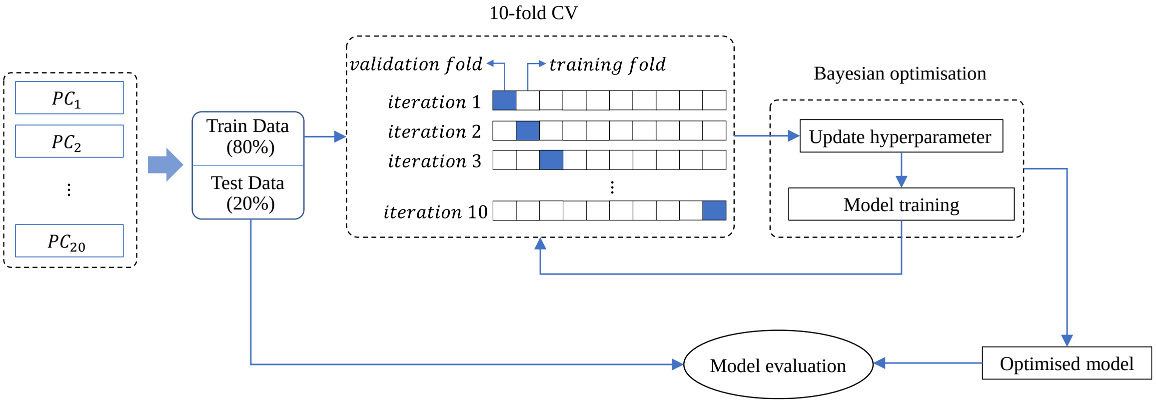
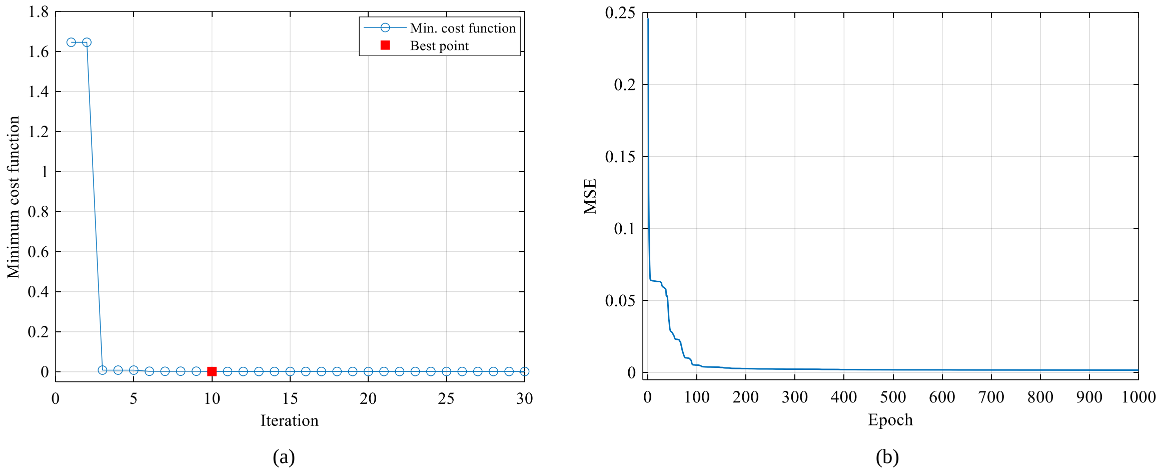


Figure 5-25. Flowchart of ANN optimisation and validation

Table 5-8. List of hyperparameters with selected values for ANN

Hyperparameter	Search range	Optimal value
Number of layers	1 – 3	1
Layer size	1 – 300	5
Activation function	Relu, Tanh, Sigmoid	Sigmoid
λ	1.04e(-8) – 104.16	1.67e(-8)
Weight initialiser	Glorot, He	Glorot



(a) (b)

Figure 5-26. (a) Hyperparameter tuning process, (b) learning curve of ANN model

The prediction performance of the model was checked by introducing evaluation criteria. Two metrices, namely, $RMSE$ and R^2 were used to show the efficiency of the ANN. Table 5-9 shows the values of the metrices for the training, validation, and test sets. These values imply that the developed ANN is capable of making predictions with low error and high correlation. Also, the high accuracy of the model in predicting the test data indicates that it has overcome overfitting to the training data.

In addition, to visualise the accuracy of the model, the correlation between the finite element simulations and the ANN predictions is illustrated in Figure 5-27. This figure compares the elastic modulus values for various ranges of fibre volume fractions between the numerical and machine learning approaches. It is seen that for all sets of fibre volume fractions, the predictions are located within $\pm 10\%$ of the FE simulations. Although ANN is able to map the microstructure geometrical features to the elastic modulus with high accuracy, the efficiency of the prediction model can be improved by increasing the training data size with RVE models of more variable volume fractions. The statistical measures of ANN for different folds are also presented in Figure 5-28. The RMSE and R^2 metrices for the training and validation folds, together with their corresponding mean values, are depicted in this figure. The agreement of the error and correlation values between the training and validation folds implies that the model was able to accurately perform out-of-fold-prediction without overfitting to the training data.

A comparison of the computation cost required for estimating the transverse modulus was made to demonstrate the computational efficiency of ANN over the FE model. Table 5-10 shows the computation time of modelling RVEs with fibre volume fractions ranging from 10% to 50%. It is seen that as the fibre volume ratio increases, the estimations of the FE model become increasingly time-consuming. The computation time starts at 24.8 seconds for $\nu_f = 10\%$ and reaches 31.3 seconds for the micromodel with 50% ν_f. However, it is seen that for the ANN model, regardless of the fibre volume fraction, the required time to predict the transverse modulus remains at 0.001 seconds, which is insignificant compared to the computation time of the FE models.

Table 5-9. Values of evaluation metrics for ANN

Model	Training set		Validation set		Test set	
	RMSE	R^2	RMSE	R^2	RMSE	R^2
ANN	0.0857	0.987	0.0854	0.987	0.0952	0.985

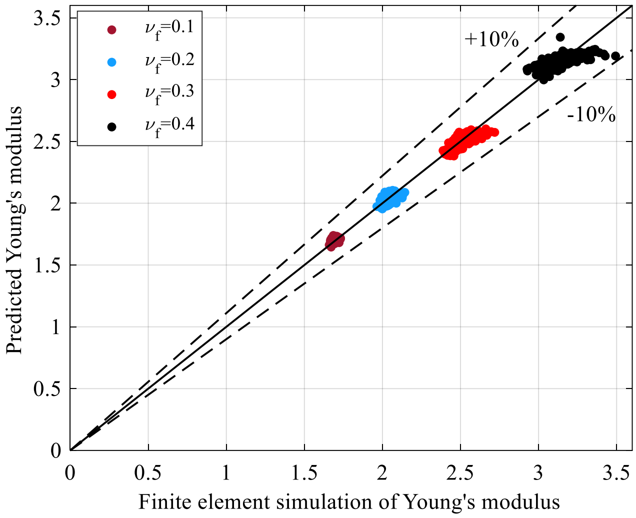


Figure 5-27. Correlation between the finite element simulations and the ANN predictions

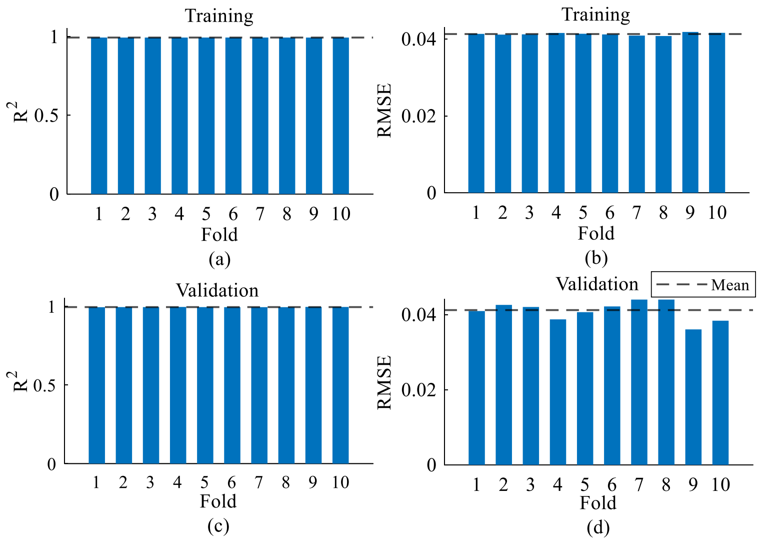


Figure 5-28. 10-fold CV measurements: (a, c) R^2, (b, d) RMSE

Table 5-10. FE and ANN computation times in seconds

Model	Fibre volume fraction (%)				
	10	20	30	40	50
FE	24.8	27.6	28.9	29.2	31.3
ANN	0.001	0.001	0.001	0.001	0.001

5.9.2 Influence of geometric features on composite transverse behaviour

In this study, a prediction model was developed based on the features extracted from the TPC functions by PCA. In order to demonstrate the impact of each PC on the Young's modulus, a feature importance analysis was performed by training an ensemble regression model with all the data. A hyperparameter tuning by Bayesian optimisation was performed to select the model with the highest accuracy. The details of the optimal hyperparameter values are listed in Table 5-11. Therefore, the selected model is an ensemble of 179 decision trees with a minimum number of one leaf node observations, which were combined with the least-squares boosting (LSBoost) method.

In order to rank the importance of extracted features by the ensemble model, the weighted average of feature importance for the tree learners was obtained. With each decision tree model, the feature contributions were calculated by multiplying the node errors by the probabilities of the nodes and averaging over the branch nodes. Figure 5-29 shows the importance of extracted features by PCA. It is observed that PC_1 and PC_2 account for 42% and 23% of influence on the elastic modulus, respectively, whereas the remaining PCs each contribute less than 4%. Moreover, the contour plots of the important components are provided in Figure 5-30. Figure 5-30a shows the plot of the first PC. It is seen that the highest intensity is located at the centre of the image. The PC_1 provides information on the fibre volume fraction. The second PC plot in Figure 5-30b accounts for an important geometrical feature, which is fibre size [89]. Considering the high importance of the geometric features extracted by PC_1 and PC_2, it is implied that the volume fraction and size of fibres have a high influence on the transverse elastic modulus of the UD composites.

The PC patterns of higher orders are generally more complex and encode further geometric features. It is seen in Figure 5-30c and Figure 5-30d that the 3rd and 4th PCs stand for the arrangement of fibres in the horizontal and vertical directions [88]. PC_5 in Figure 5-30e is also important for predictions of the elastic modulus. Although the interpretation of the pattern is not straightforward, it can encode information on the inter-fibre distance.

Table 5-11. Selected hyperparameter values for ensemble model

Model	Hyperparameter	Selected value
Ensemble	Method	LSBoost
	Number of Learners	179
	Learning rate	0.1165
	MinLeafSize	1

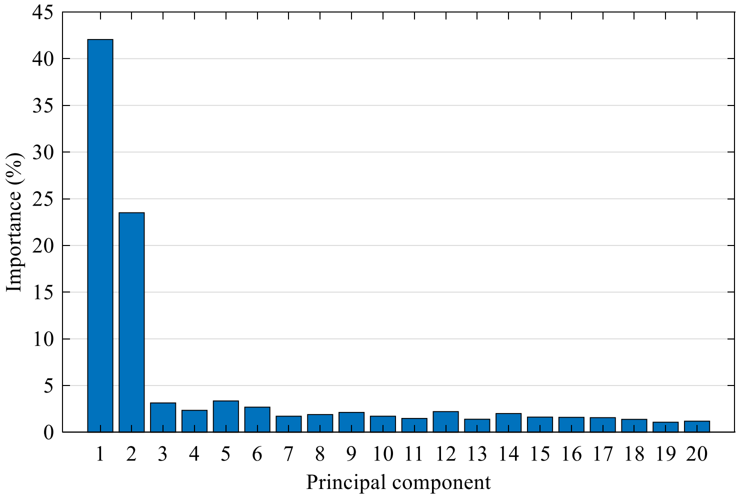


Figure 5-29. Influence of PCs on Young’s modulus prediction

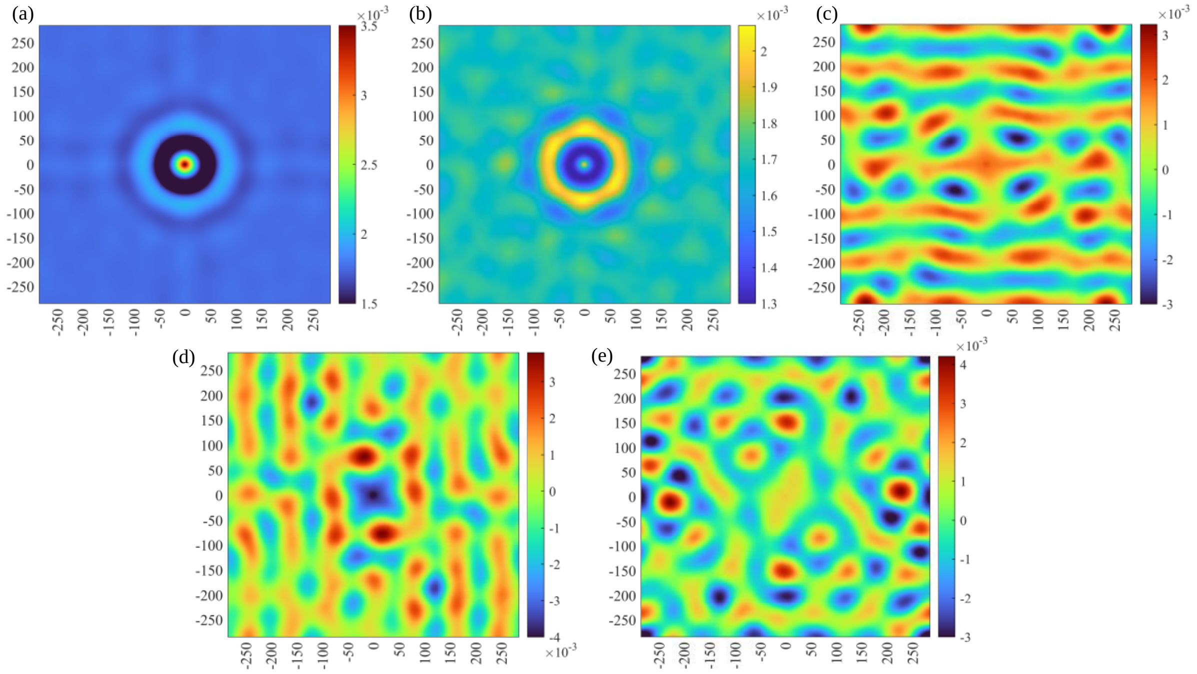


Figure 5-30. Contour plots of different PCs: (a) PC_1, (b) PC_2, (c) PC_3, (d) PC_4, (e) PC_5

5.10 Conclusions

Given the importance of studying the ageing impact on FRP composites, micromechanics and machine learning methods were used in this chapter to analyse the environmental degradation at the material scale. A numerical study was performed to verify the impact of the humid condition on the transverse behaviour of the UD glass fibre-reinforced polymer composites under uniaxial tensions. A framework was developed to automatically create RVEs with random fibre distributions that would also account for the interphase region. The fibre overlaps and the periodicity of the area were checked during the microstructure generation. The periodic boundary constraints were defined along opposite edges of the RVE under uniaxial loading. A representation of obtaining homogenised mechanical properties through calculating the average strain and stress values at the macroscale was provided. A comparison between the numerical, experimental, and analytical results was conducted to confirm the validity of the finite element modelling.

A parametric study was performed to assess the influence of a series of variables on the transverse behaviour of the composite. A comparison of different mesh characteristics revealed that an element type of CPE4 with a length of 1.5 μm avoided convergence problem in the numerical modelling. The feasibility of applying PBC and KUBC with different fibre volume fractions was compared, and PBC was selected to automatically create constraints on the boundaries of the RVE. For a microstructure with a 20% fibre volume ratio, it was observed that $L_{RVE}/D_{fibre} \geq 10$ obtained consistent results in terms of the Young's modulus. The study on the fibre arrangement in the RVE revealed that higher volume ratios of fibres contribute to the impact of a random distribution on the homogenised response of the microstructure. It was also seen that for higher values of fibre volume fraction, the increase in the interphase thickness is correlated to the increase in the elastic modulus values. This provided evidence of the interphase region's involvement in the transverse behaviour of the composite.

In order to demonstrate the feasibility of machine learning in analysing aged composite materials, it was necessary to generate the microstructural data first. Therefore, a total of 1,200 microstructure images with varying fibre volume fraction and interphase thickness values were created at a resolution of 570×570 pixels. The effective elastic modulus under the influence of humid condition was calculated for each RVE using FE simulations. A pre-processing procedure was organised to prepare data for the machine learning application. TPC functions were defined for each pixel of the three-phase composite material to statistically characterise the geometry of the microstructure. Considering that the resulting TPCs comprised a very large feature space, PCA technique was used to reduce the dimensions and extract important geometric features. Among the 20 PCs used for feature extraction, it was seen that the first two components captured about 98.5% of the variance in the data, indicating that they contained the prominent geometrical features of the microstructure.

The resulting PCs were then used as inputs to develop an ANN model and evaluate its efficiency. The PCA data was randomly divided into 80% training and 20% test sets. A Bayesian optimisation technique was used in conjunction with a 10-fold CV method to tune the hyperparameters of the ANN and confirm the generalisation ability of the prediction model to the new data. It was discovered that the prediction model can accurately map the microstructure geometric features to the elastic modulus of aged composites. The error and correlation values of the folds in the training and validation sets also confirmed the good performance of ANN in performing out-of-fold predictions without overfitting to the training data. An importance analysis of the geometric features was performed by an ensemble model. It was seen that volume fraction and size of fibres greatly impacted the effective elastic modulus of the glass fibre/epoxy composite. The contour plots of PCs of high importance also revealed the key geometric information extracted from the TPC functions.

The developed framework in this chapter demonstrated an efficient approach to analysing ageing impact at the material scale. Although numerical modelling is a helpful tool to study the influence of ageing in composite materials, its integration with machine learning proposes a framework that is computationally fast and accurate. It should be noted that developing machine learning models using the FE results requires significant computational effort. However, once trained, machine learning can perform further predictions much faster and with less expensive computational effort. In addition, the ability to generalise to new data makes it possible to perform predictions outside the training range. This expands the application of machine learning models in predicting without a further need to perform FE simulations. Therefore, the long-term behaviour of composite materials can be quickly simulated over a wide range of modelling parameters.

The proposed framework can also support the development of efficient maintenance strategies for composite structures. By predicting material degradation under different ageing conditions, it is possible to identify potential structural damages and plan timely maintenance interventions. This can help prevent costly failures and extend the service life of composite structures. The framework can also provide valuable insights for material design. Furthermore, by understanding the composite structure-property relationship, the framework can guide the selection of suitable geometrical, mechanical, and environmental parameters to achieve the desired material properties for composite manufacturing. In addition, the framework can reduce the need for costly experimental testing by providing estimations of composite material durability. By leveraging the computational efficiency of the proposed framework, informed decisions can be made regarding the use of composite materials for the strengthening of structures.

CHAPTER 6

Multiscale Finite Element and Machine Learning Modelling of Ageing in FRP Composites Under Water Immersion Condition

6.1 Introduction

In the previous chapter, we developed a machine learning framework based on the micromechanical analysis to investigate the impact of humid ageing on the transverse behaviour of unidirectional (UD) composites. The results indicated the promising performance of the proposed approach in mapping the geometric features at the microscale to the effective mechanical properties. In order to further understand how ageing impacts the long-term performance of composite materials, it is required to study the ageing process during interaction with the surrounding environment. The literature review in Chapter 2 showed that moisture diffusion is one of the most prevalent phenomena that inversely influences the durability of composite structures and materials. The interaction of water with the polymer molecules can lead to a chain of chemical reactions that eventually degrade the mechanical properties. Therefore, studying the diffusion of moisture can help understand the degradation of the composite material over time.

In this chapter, analytical and numerical approaches are first employed to simulate the kinetics of moisture transport through the composite material. Among various methods in the literature, Fick's law is regarded as a simple yet robust solution to model the diffusion of water in composite materials [56, 170]. The analytical and finite element methods are used to simulate moisture ingress in a composite containing unidirectional glass fibres. A comparison with an experimental study is made to validate the analytical and numerical models and obtain the diffusion coefficient values.

The results from the moisture diffusion analysis are then used to build a multiscale framework to model non-uniform material degradation without making any prior assumptions about the constitutive models at the macroscale. Particularly, the constitutive behaviour of each point in the macro-model is obtained by the finite element (FE) analysis at the microscale. It is accomplished by homogenising the mechanical behaviour of the volume element representing each point in the macro-model. The developed framework is able to include the effects of microstructure features on the macro-model. Therefore, it is used to model material degradation during exposure to water immersion conditions. This is accomplished by combining a macroscale diffusion analysis with a stress analysis at the microscale. At each time step, the responses from the water diffusion kinetics at the macroscale are passed to the representative volume elements (RVEs) at the microscale to obtain the transverse properties. The framework is validated by comparing its estimations with the durability experiment results of a UD composite plate immersed in water.

In the proposed multiscale approach, FE analyses are performed at each step of diffusion for all the RVEs embedded in the macro-model in order to update its mechanical properties. This results in a significant computational effort that can undermine the efficiency of the framework. To overcome this issue, a data driven machine learning-based model is developed to surrogate the simulations at the microscale. Machine learning can map the geometrical, mechanical, and environmental properties of the composite material to its transverse properties. Once trained, the prediction model can predict the transverse performance of the composite material during the water immersion without the need to perform further FE analysis.

Figure 6-1 shows the multiscale framework for modelling the composite material degradation under the water immersion condition. At each time step, the diffusion is modelled by Fick’s law at the macroscale. The water content values at every integration point are then used to update the material constituent properties and solve the stress problem at the microscale. The analysis results are finally upscaled by the numerical homogenisation to obtain the effective elastic properties of the composite in macroscale. In order to develop a machine learning-based framework, the information from the diffusion and stress problems at the macro- and microscales, respectively, is used to train a prediction model. The estimations of the effective transverse properties during the exposure to water immersion conditions are compared with the experimental observations to validate the efficiency of the proposed framework in modelling the multiscale aging process of composite materials.

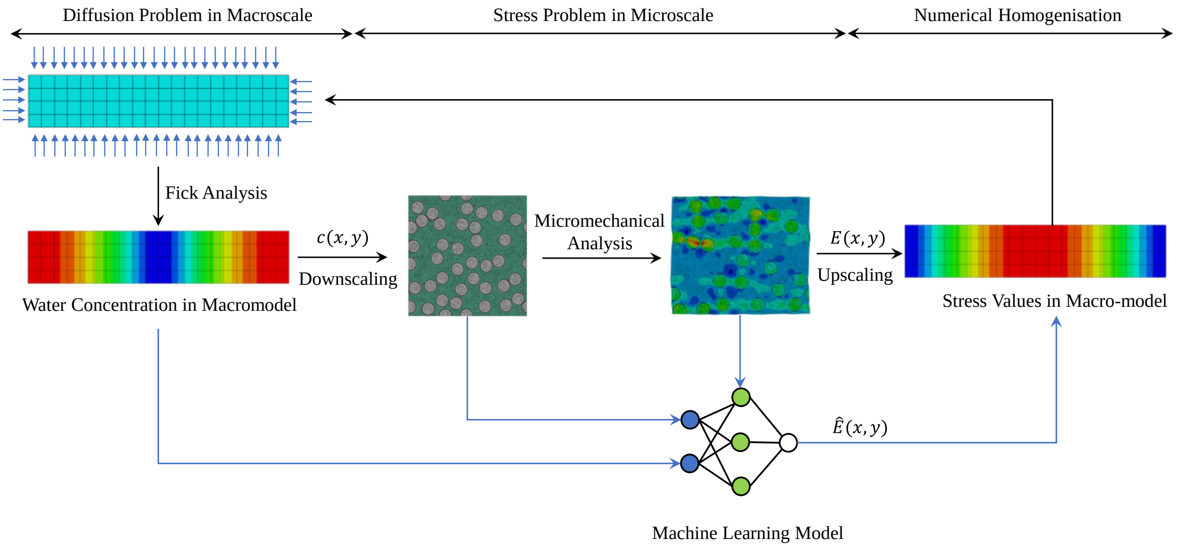


Figure 6-1. Flowchart for multiscale modelling of the ageing mechanism for UD composite under water immersion condition

6.2 Moisture diffusion study of UD composite material

The diffusion of moisture into the composite can be explained by three stages. At first, water diffuses into the material through the existing voids and defects facilitated by the capillary effect, where the moisture content increases linearly with time. In the following, the chemical reaction of water with the composite matrix leads to breakage of the polymer chain, due to which further water molecules ingress into the composite [29]. As a consequence, the water content of the material rises slowly in the second phase. The diffusion continues until the material reaches saturation at the third stage. The phases of moisture absorption are depicted by the gravimetric curve in Figure 6-2. The moisture content (M_t) is normally obtained by measuring the weight gained by the composite samples during exposure to moisture conditions as in Eq. (6-1).

$$M_t = \frac{W_t - W_0}{W_0} \times 100 \tag{6-1}$$

where W_t and W_0 denote the weight of the samples at the aged and unaged states.

The diffusion of moisture into the FRP composite material is normally addressed by the Fickian model in the literature [57]. The Fick's law describes the kinetics of diffusion as a function of the concentration gradient, where the diffusive agent moves from high to low concentration points in the medium. In this model, the diffusion flux is related to the material diffusivity and the concentration gradient. Fick's law is mathematically represented as:

$$\frac{\partial_c}{\partial_t} = D\frac{\partial^2{}_c}{\partial x^2} \tag{6-2}$$

where $\frac{\partial_c}{\partial_t}$ denotes the change in concentration with time; D represents the material diffusivity, and $\frac{\partial^2{}_c}{\partial^2{}_x}$ explains the change in concentration with the location of the diffusive agent. The analytical solution of Fick's equation describes the concentration of the diffusive agent as a function of time and location in a sample with the thickness of h [171].

$$M_t = M_s\left[1 - \frac{8}{\pi^2}\sum_{m=0}^{\infty}\frac{1}{(2k+1)^2}exp\left(-\frac{D(2k+1)^2\pi^2 t}{h^2}\right)\right] \tag{6-3}$$

where M_s is the moisture content in the saturated sample. Eq. (6-3) can be used to approximate the content in the short- and long-term duration of exposure as:

$$M_t = M_s\left[\frac{4}{\pi^2}\sqrt{\frac{Dt}{h^2}}\right] \quad \text{for} \quad \frac{Dt}{h^2} < 0.04 \tag{6-4}$$

$$M_t = M_s\left[1 - \frac{8}{\pi^2}exp\left(\frac{-Dt}{h^2}\pi^2\right)\right] \quad \text{for} \quad \frac{Dt}{h^2} > 0.04 \tag{6-5}$$

In these equations, D is obtained from the gradient of the linear part (θ) in the gravimetric curve as:

$$D = \pi\left(\frac{h}{4M_s}\right)^2\theta^2 \tag{6-6}$$

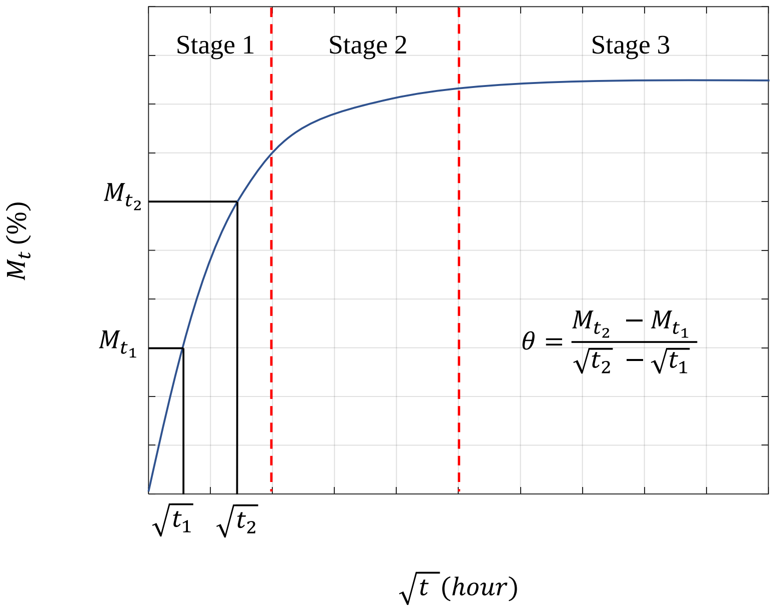


Figure 6-2. Gravimetric curve of diffusion by Fick’s law

6.2.1 Moisture absorption modelling

The diffusion problem at the macroscale, as depicted in Figure 6-1, explains finding the water concentration values in the macro-model using Fick’s analysis. In this section, analytical and numerical models were developed to obtain the moisture uptakes for the epoxy resin and the UD glass fibre composite samples with 50% volume fractions. The simulation results were validated by the experiment observations of the samples with a size of 75×75×2 mm^2 that were immersed in 70 °C deionised water for 16 weeks [56]. Figure 6-3 shows the experimental values of the moisture content for the resin specimen. It is seen that the saturation content is 2.3%. A Fickian model was developed and fitted to the experimental data. As a result, a diffusion coefficient of 15×10^{-7} mm^2/s was derived for resin.

In order to obtain the moisture absorption curve for the UD composite, a 2D FE model of the sample was developed in the Abaqus CAE software [133]. It was first assumed that the inclusion of fibres did not change the diffusivity of the resin matrix. Therefore, a coefficient value of 15×10^{-7} mm^2/s was assigned to the matrix. Since the fibres are not impacted by the moisture, an arbitrary value of 1×10^{-24} mm^2/s was assigned to them [172]. Due to the symmetry, only a quarter of the sample was modelled, and a constant water concentration was applied to the model’s boundaries. The mathematical framework described in Chapter 5 was used to generate the model with random inclusions. Figure 6-4 shows the schematic FE model of the sample with boundary conditions and random fibre distributions. It is obvious that moisture diffuses from the left and bottom edges into the sample. The model was meshed by triangular-shaped elements with linear geometric order (DC2D3), and a size of 5 µm was assigned to the elements. Figure 6-5 depicts a part of the model meshed by DC2D3 elements.

The analogy between Fickian mass diffusion and Fourier's law of heat transfer was used to model the water uptake [118]. This is because heat transfer is also driven by the gradient of temperature. Therefore, similar to Eq. (6-2), the change in the temperature is proportional to the thermal diffusivity as:

$$\frac{\partial T}{\partial t} = \alpha \frac{\partial^2 T}{\partial x^2} \tag{6-7}$$

where T is temperature, and α denotes thermal diffusivity which is written in the form of:

$$\alpha = \frac{k}{\rho C_p} \tag{6-8}$$

with k, ρ and C_p denoting the conductivity, density, and special heat parameters. Based on the analogy, the values of these parameters were adjusted to replicate the moisture diffusion in the FE model.

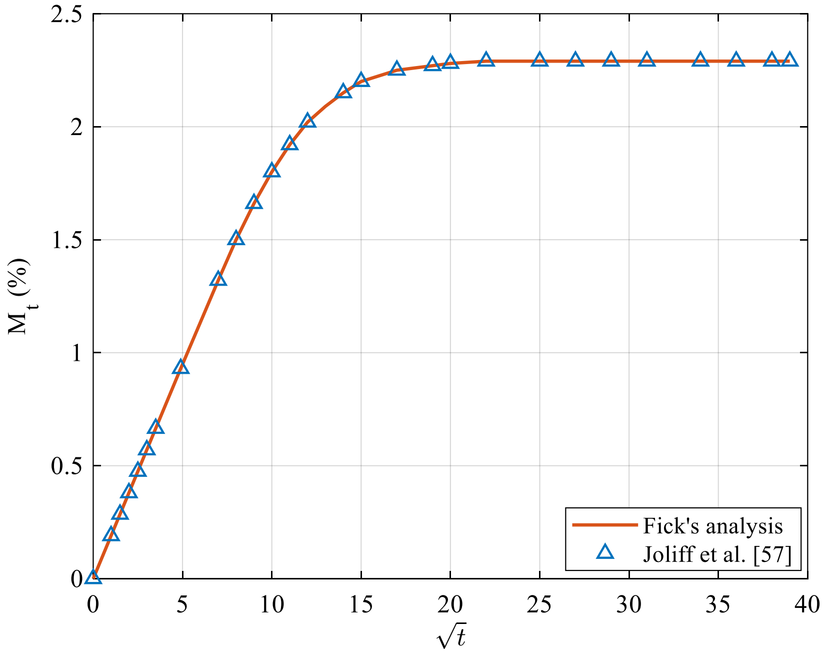


Figure 6-3. Moisture uptake of the epoxy resin obtained by the Fick's equation

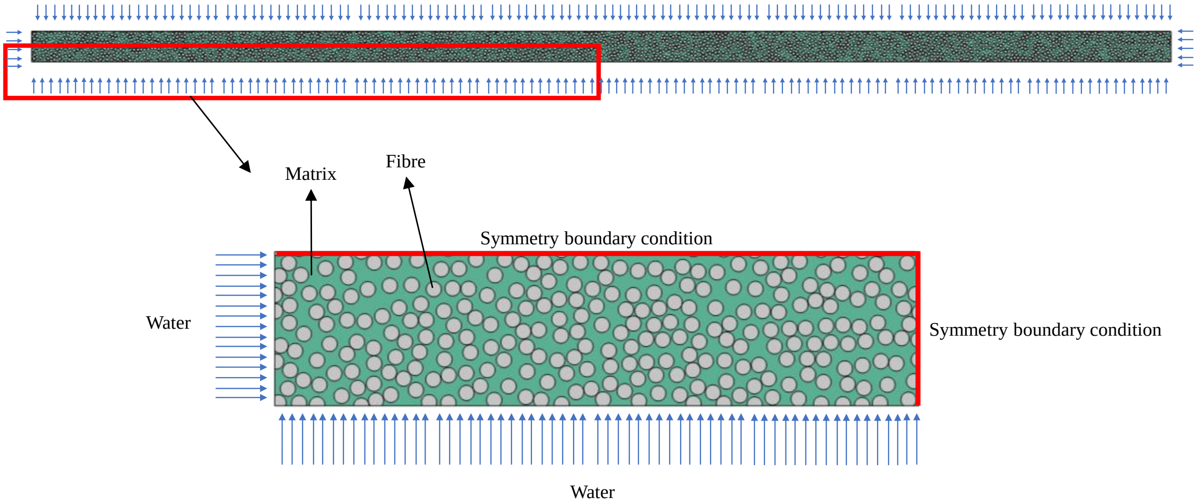


Figure 6-4. FE model of water diffusion into the UD composite sample

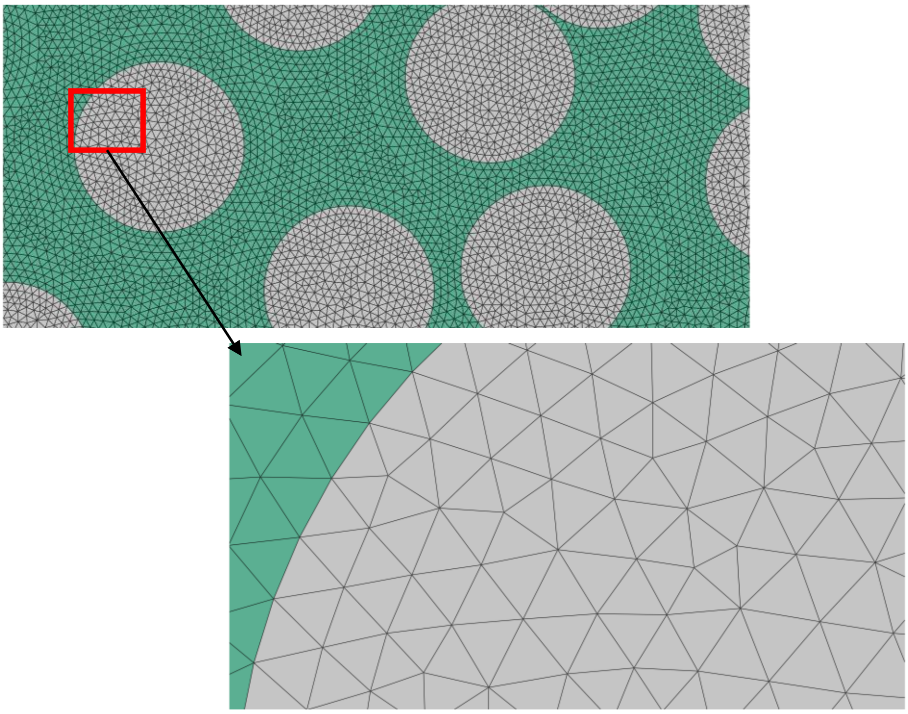


Figure 6-5. A 2D model of the composite meshed by DC2D3 elements

The kinetics of diffusion in the composite were identified by the moisture uptake values from the experiment and were fitted by Fick’s law and the FE simulation results. Figure 6-6 shows the composite absorption curves. It is seen that there is a good agreement between the values from the experiment and Fick’s equation. A saturation level of 0.6% was obtained for the composite material, which is lower than the ones in the neat resin sample. This implies the barrier effect of fibres, which limits the amount of water that can penetrate the composite. The fitting of the experimental data to Fick's equation resulted in a diffusion coefficient of $7.75\times10^{-}$

^{7} mm^2/s. In the FE model, the water uptake at each time step was obtained by averaging the concentration values of all integration points. It is seen in Figure 6-6 that assigning the diffusion coefficient of 15×10^{-7} mm^2/s to the matrix leads to slightly lower water uptake. Therefore, the diffusivity was increased to 20×10^{-7} mm^2/s in order to fit the FE model with $R^2 = 0.99$ to the experiment results. This indicates that the diffusivity of the resin with the fibre inclusions is higher than that of the neat resin, which implies that the matrix can be modified in the presence of fibres in the composite. This is related to the fact that the mobility of the polymer chain is higher at the interphase regions around the fibres, which highlights the necessity for including the interphase in the numerical modelling [56].

In order to illustrate the water uptake, the isolines of the moisture at different time steps are illustrated in Figure 6-7. The evolution of immersion from the left and bottom edges of the sample shows that water constantly diffuses into the composite matrix until it saturates the sample. In this figure, NT11 represents the water concentration values at different points on the specimen. Figure 6-8 depicts the change in the moisture content along the sample at different stages of water immersion. It is seen that due to the constant concentration at the sample's boundaries, the water uptake is greatest at the left and right edges, and lowest in the middle. In addition, with the progression of the exposure, the water content increases along the sample length until it reaches saturation after 800 hours.

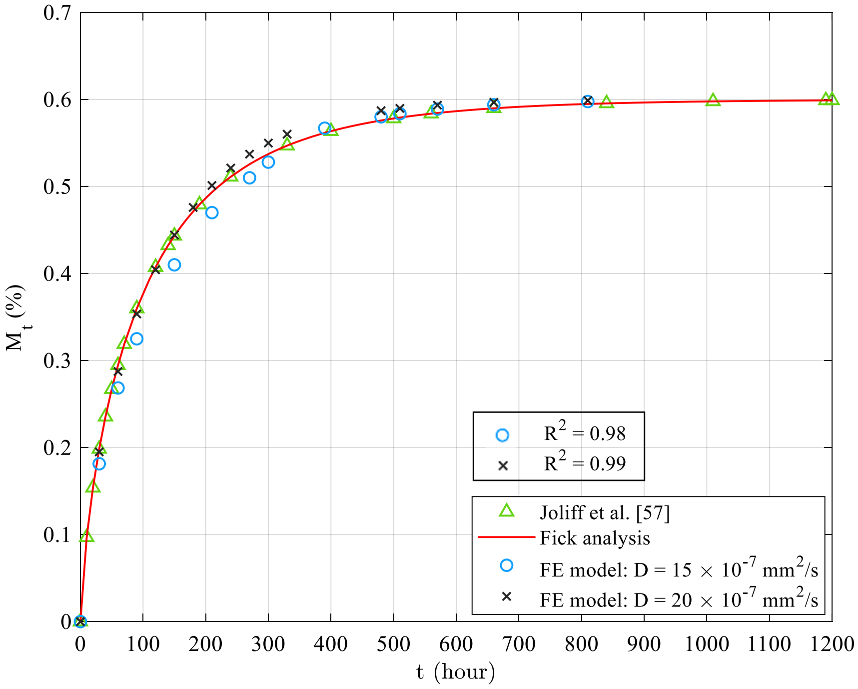
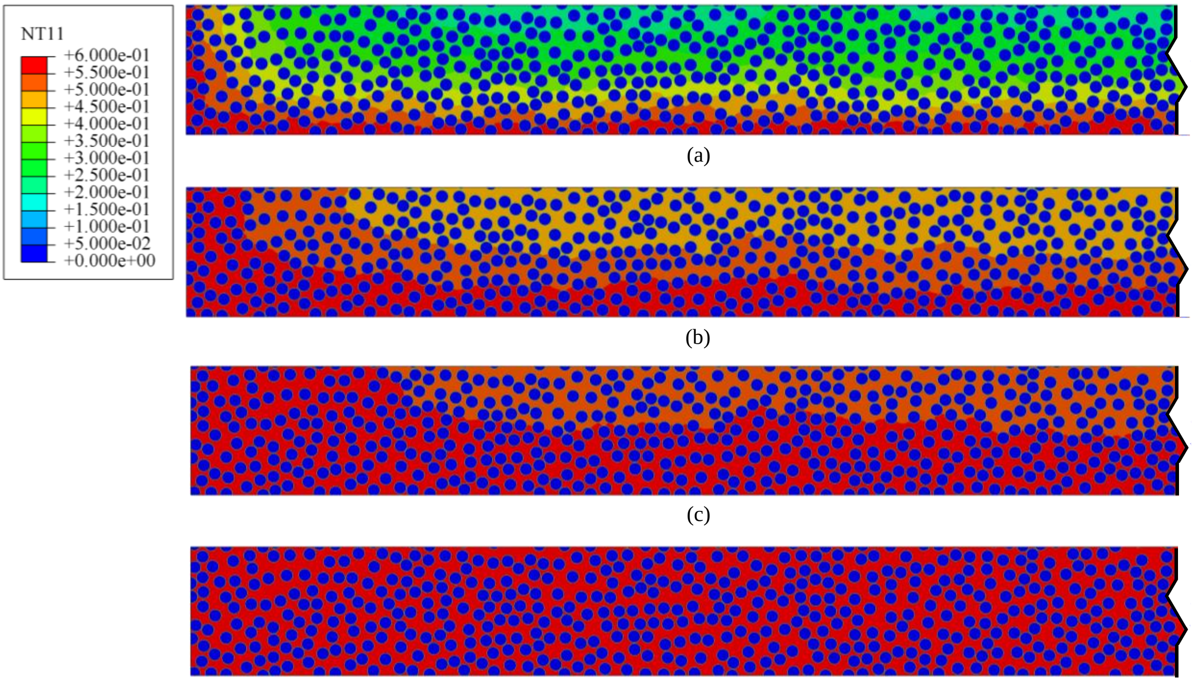


Figure 6-6. Moisture uptake of the composite modelled by the Fickian and FE methods



(a)

(b)

(c)

(d)

Figure 6-7. Moisture diffusion into the composite at different time steps (in hours): (a) 90, (b) 210, (c) 300, (d) 500

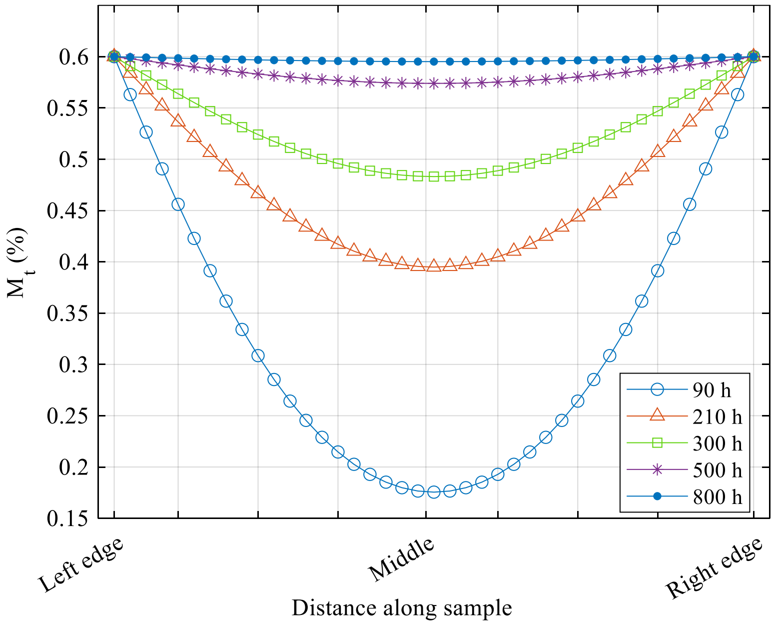


Figure 6-8. Moisture concentration along the length of the composite sample at different time steps

6.3 Numerical multiscale modelling of ageing in UD composite

In the previous section, the mechanism of water immersion into the composite was explained by the kinetics of diffusion. It is also important to numerically elaborate the process of material ageing due to moisture absorption. Figure 6-7 showed that the moisture content was different between the fibre and matrix in the composite. It was also obvious that the concentration of water was not homogenous during immersion. This can result in a non-uniform degradation of the composite material. Therefore, a multiscale approach is required to concurrently study material ageing at the macro- and microscale. The results in Chapter 5 showed that the proposed framework was able to obtain macroscopic transverse behaviour from a single micromechanical simulation. However, the distribution of water with varying concentrations in the macro-model changes the material constituent properties at different locations. This requires performing micromechanical analysis for all the points in the macro-model at each step of moisture diffusion.

Based on the proposed framework in Figure 6-1, a macro-model diffusion problem was solved by the FE model using Fick’s law. A benefit of this approach is that no prior assumption was not made about the constitutive model at the macroscale. However, for each point in the sample, the water concentration values were passed to the micromodel for a stress analysis. Due to the separation of scales between the micro- and macro-models, it was assumed that the water content was constant within the microstructure [138, 173].

At the microscale, two-dimensional (2D) RVEs were created as the UD fibres randomly placed in the matrix with the interphase regions around them using the algorithm described in Chapter 5. A micromechanical stress analysis was performed by FE to obtain the transverse mechanical properties. The RVE constituent models were considered homogenous and linear materials with plain strains, and the interphase was assumed to be perfectly tied to the fibre and matrix [57, 61].

In order to include the aging process caused by moisture diffusion, the degradation levels of the matrix and interphase were set to be proportional to the water concentration in the micromodel [44]. A degradation factor (d_t) is defined as:

$$d_t = \frac{d_s}{M_s} M_t \tag{6-9}$$

where d_s denotes the degradation level of the saturated material. By using d_t, the matrix and interphase properties during ageing are modified as:

$$E_t = (1 - d_t)E \tag{6-10}$$

where E_t and E represent the aged and unaged transverse elastic properties. These parameters were used to update the mechanical properties of the RVE constituent models. The microscale stress problem was then solved, and the effective elastic modulus was derived by numerically homogenising the response over the area of the RVE.

6.3.1 Macroscale diffusion problem

It was seen that comparing the moisture concentration values obtained from the Fickian model to the experimental data in Section 6.2.1 resulted in a diffusion coefficient of 7.75×10^{-7} mm^2/s for the UD composite in the transverse direction. This information was used to solve the macro-model diffusion problem. A plane FE model of 75×2 mm^2 was created and meshed by the quadrangular geometries with linear interpolation between the nodes (DC2D4) and a size of 0.05 mm. Figure 6-9 shows an example of the elements used to mesh the macro-model. The final model comprises 4125 nodes. A constant concentration of 0.6% was applied to the boundaries of the model to simulate the immersion condition.

In order to obtain the water uptake at all integration points, the FE diffusion problem was solved at each time step. Figure 6-10 shows the averaged moisture concentrations during water immersion. It is seen that the estimations are in good agreement with the experimental observations. A snapshot of the moisture concentration after 500 hours of exposure was also depicted in Figure 6-11 to show the distribution of moisture content at various points of the macro-model. These values were then passed to the micromodel in order to perform a micro-stress analysis in the next section.

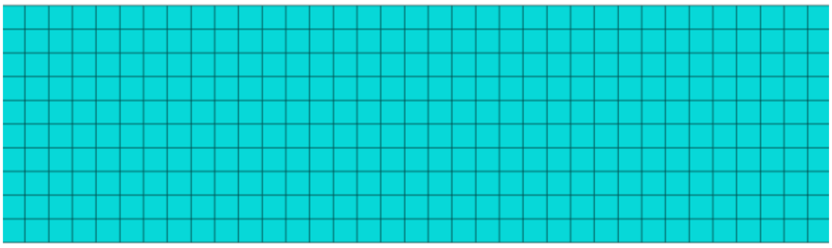


Figure 6-9. Macro-model meshed by DC2D4 elements

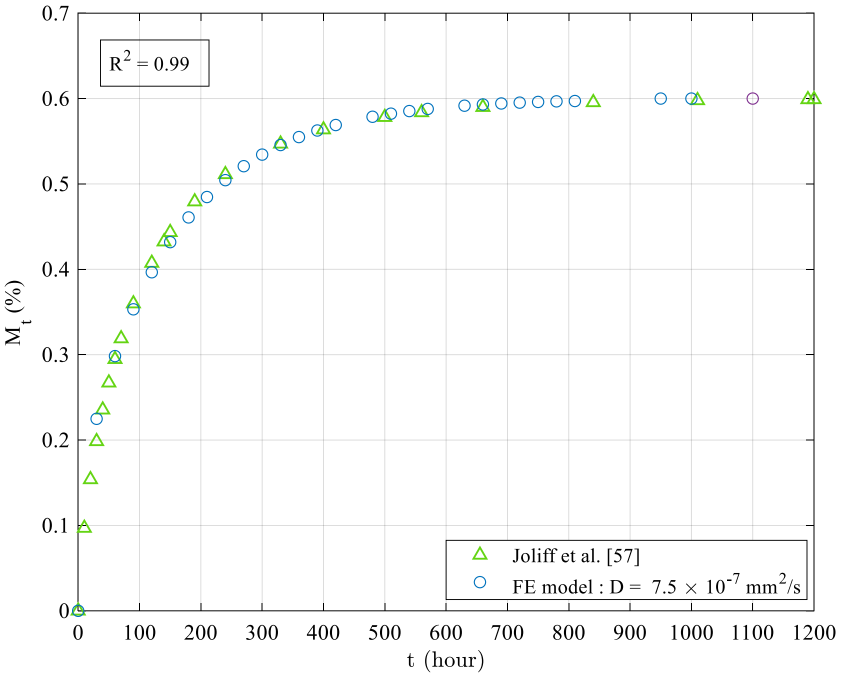


Figure 6-10. Moisture uptake of the composite during immersion by the experimental data and FE model

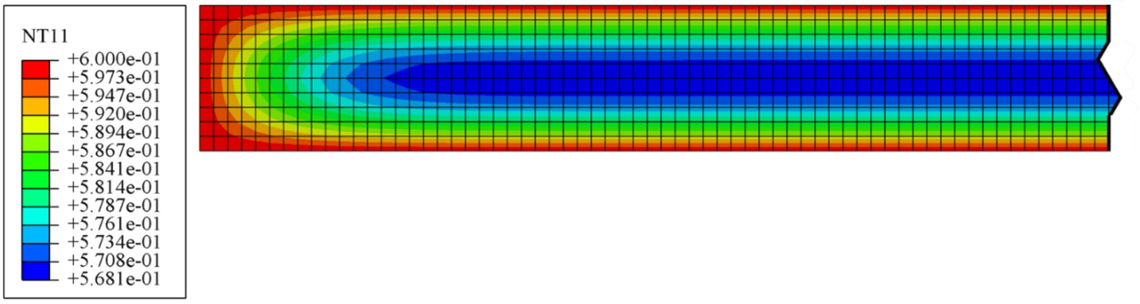


Figure 6-11. Water concentration field at the FE macro-model

6.3.2 Microscale stress problem

In this section, the water uptake values from the macro-model were passed to the microscale stress problem to study the impact of moisture on the composite material. In order to obtain the microscale stress values, an RVE with 50% fibre volume fraction, including interphase regions, was first automatically generated in Abaqus. Figure 6-12a depicts the created micromodel. However, it is seen that the interphase regions overlap each other, which creates issues in meshing and assigning constitutive material properties. This requires removing the intersected regions by editing the geometry of the RVE. The resulting micromodel after erasing the redundant lines is depicted in Figure 6-12b. The generated RVE has 150×150 μm^2 section area, including circular fibres with an interphase of 1 µm thickness around them. In order to perform FE analysis, the model was meshed with the quadrangular shape elements with linear interpolation between the nodes (CPE4). The length of the elements was set to 1.5 µm to avoid convergence problems, as shown in Chapter 5. The nodes of the micromodel were constrained to capture periodicity and have the same stress values at the opposing edges. Material ageing at the microscale was addressed by performing stress analysis, A tensile uniaxial displacement load was applied in the horizontal direction to the RVE vertex, with the other vertices constrained in both directions within the plane.

In order to perform a FE micro-stress analysis, the material properties of the epoxy matrix and fibre were obtained from an experimental study on the durability of the UD composite plate under water immersion condition [57]. These values are provided in Table 6-1. In this table, t_0 and t_S represent the initial and saturation stages of the sample during immersion. Since the details on the interphase were not provided in the experimental study, we assumed that the interphase shared the same mechanical properties as the epoxy matrix.

Considering the fact that moisture adversely impacts the epoxy matrix with time, it is important to include material degradation in the FE simulation of RVE [55]. Particularly, the moisture concentration values were used to update the mechanical properties of the composite. Based on the values in Table 6-1, the degradation at the saturation level (d_s) for the matrix and interphase is calculated as 0.172. In addition, Figure 6-10 showed that the composite saturation level was 0.6%. A degradation factor of 0.287 was derived for epoxy matrix by Eq. (6-9). Therefore, using Eq. (6-10), the modified properties at time t were obtained as a function of moisture content and in the form of:

$$(E_t)_m = (1 - 0.287M_t)E_m \tag{6-11}$$

where $(E_t)_m$ denotes the transverse properties of the matrix in time *t*. A flowchart on modifying the micromodel constituent properties under the impact of water diffusion is depicted in Figure 6-13. For each integration point in the macro-model, the water concentration values at the immersion time *t* are passed to modify the RVE constituent properties and perform the stress analysis.

Table 6-1. Epoxy matrix and glass fibre mechanical properties used for numerical modelling

	Matrix		Fibre	
	E_m (GPa)	ν_m	E_f (Gpa)	ν_f
t_0	2.7 ± 0.2	0.35	70	0.35
t_S	2.5 ± 0.1	0.35	70	0.35

The stress analysis results of the RVE FE model are presented in Figure 6-14. This figure shows the resulting normalised von Mises stress fields at two different integration points of the macro-model after 500 hours of exposure to water immersion. The RVEs in this figure represent the points in the macro-model with high and low moisture concentrations, respectively. Comparing the stress contour plots shows that the normalised values at Figure 6-14c are slightly greater than those at Figure 6-14b, which is due to the lower water content and therefore higher stiffness of the material at the inner section of the macro-model. This suggests that the presence of moisture in the macro-model results in non-uniform degradation at various regions of the material. Therefore, a stand-alone microscale stress analysis is not sufficient to accurately represent the macroscale transverse behaviour under water immersion conditions. These necessitate using the multiscale framework, as provided in Figure 6-1 to couple the macro- and microscale analyses.

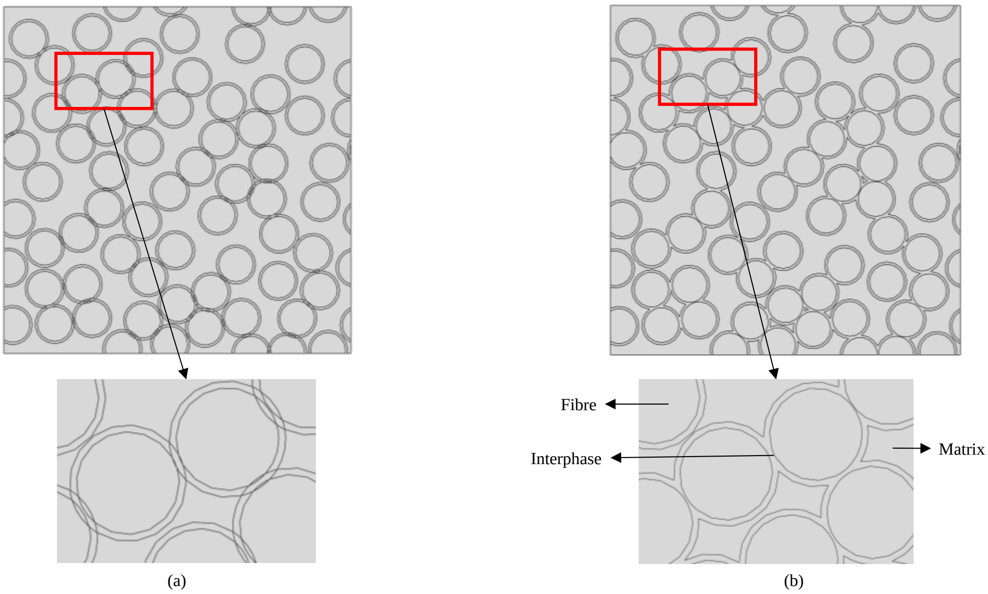


Figure 6-12. Interphase modelling in the RVE: (a) intersecting lines, (b) Removed redundant lines

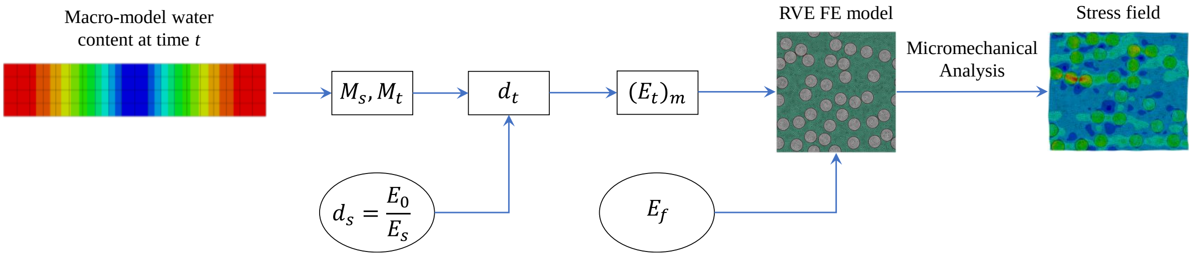


Figure 6-13. Flowchart for performing micro-stress analysis

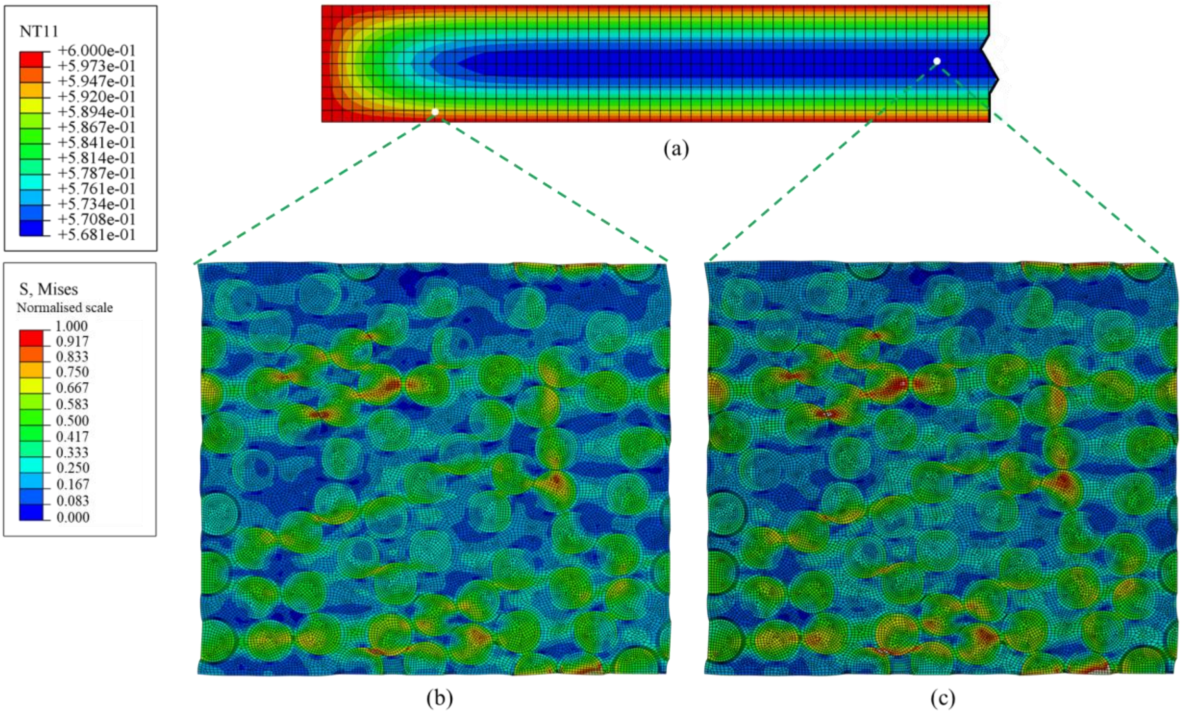


Figure 6-14. Normalised von Mises stress values at different points of the macro-model

6.3.3 Validation of the multiscale numerical framework

Sections 6.3.1 and 6.3.2 addressed the diffusion and stress problems and the transition of knowledge across multiple scales. These outputs were used to perform computational homogenisation and calculate the material properties impacted by the water immersion condition in this section. In order to validate the multiscale framework, the simulation results were compared with the observations from the three-point-bending test experiments of UD composite plates of the thermoset epoxy type resin reinforced with 50% E-glass fibres [57]. In the experiment study, the ageing effect was captured by submerging the composite plates in 70 °C water for 16 weeks. The mechanical properties of the fibre and matrix were also provided in Table 6-1.

For the numerical validation, two plate samples were modelled to simulate the macro-model water diffusion at the initial (t_0) and saturation (t_s) states. Figure 6-10 showed that the water content at t_0 was 0, and the sample attained the saturation level of 0.6% after 800 hours of immersion. Therefore, the resulting moisture contents at all integration points were downscaled to the micromodel to obtain the stress values and calculate the effective transverse properties. The details of the parameters used in the multiscale framework simulation are provided in Table 6-2. In this table, the macro- and micromodel parameters were used for diffusion and stress problems, respectively.

Table 6-2. Parametric values for the multiscale numerical simulation

Scale	Parameter	Value
Macro-model	Plate size	75×2 mm^2
	D	7.75×10^{-7} mm^2/s
	M_s	0.6%
	Element type	DC2D4
	Element size	0.05 mm
Micromodel	d_s	0.28
	RVE size	150×150 μm^2
	Fibre diameter	15 μm
	v_f	50 %
	Element type	CPE4
	Element size	1.5 μm
	Boundary condition	PBC

The simulation results, together with the experimental data, are provided in Table 6-3. It is seen that the values estimated by the proposed multiscale framework are consistent with the experimental outcomes. It is also worth mentioning that the required time to complete the numerical simulations for the initial and saturated stages was 2400 seconds. The difference between the experimental data and the numerical and analytical estimations is also provided in

Table 6-3. It is seen that the numerical model is able to make predictions at the initial stage with a maximum error of 1.8%. For the saturation stage, it is obvious that both analytical (HS^-) and numerical models can estimate the transverse property of the degraded composite with a good degree of accuracy.

Table 6-3. Results of experimental, numerical, and analytical studies obtained for the composite plate at the initial and saturation states

Parameter	Sample	Experiment [57]	Numerical method		Analytical method (HS^-)	
			Result	Difference	Result	Difference
Young's modulus (GPa)	t_0	8.8 ± 0.4	8.64	1.8%	8.44	1.2%
	t_S	7.7 ± 0.2	7.78	1.0%	7.87	4.1%

6.4 Machine learning modelling of material degradation

In Section 6.3, we were able to validate the proposed multiscale framework by comparing its results with the experimental data at the initial and saturated stages. The results in Table 6-3 showed that the moisture condition could ultimately decrease the transverse properties of the composite material. In this section, we will model the composite material's degradation over the immersion time. However, it was mentioned above that using the multiscale framework to only model the two unconditioned and aged stages required considerable analysis time. This indicates that modelling the transverse behaviour for each time step until saturation requires tremendous computational effort. Therefore, in the current section, we will integrate machine learning with the multiscale framework to model the composite material degradation during immersion in water. Particularly, we want to develop a surrogate model for the micro-stress problem shown in Figure 6-1 to replace it with FE simulations at the microscale.

In order to develop a machine learning model, the data first needed to be generated. Figure 6-10 showed the moisture content of the macro-model during immersion. A macroscale diffusion problem was solved for the first 500 hours of immersion, and the results were passed to the microscale stress problem. As shown in the Figure 6-13 flowchart, at each time step, the water concentration values of the integration points from the FE macro-model were passed on to the RVEs to perform the micromechanical analysis. This was achieved by solving the 533 micro-stress problems for different moisture concentration values in Abaqus and obtaining the composite transverse modulus values. Information from the macro-model, namely, immersion duration and moisture content values, as well as the micromodels' geometric and mechanical characteristics, were compiled with the Abaqus analysis results to generate the database. Table 6-4 describes the features used to create a machine learning model and simulate material degradation by moisture diffusion. The information on the geometrical, mechanical, and environmental properties of the composite at multiple scales was used to train a prediction model.

Table 6-4. Parameters of the first 500 hours of immersion for training machine learning

Feature	Feature type	Feature name	Feature values
Input parameter	Geometrical properties	Fibre size (μm)	7.5
		Fibre volume fraction (%)	50
	Mechanical properties	Fibre transverse modulus (GPa)	70
		Fibre Poisson ratio	0.35
		Matrix transverse modulus (GPa)	2.4–2.681
		Matrix Poisson ratio	0.35
		Interphase transverse modulus (GPa)	2.4–2.681
		Interphase Poisson ratio	0.35
	Environmental properties	Immersion duration (hours)	0–500
		Moisture content (%)	0–0.6
Output parameter		Composite transverse modulus (GPa)	7.788–8.854

In order to surrogate the microscale FE simulations with a machine learning model, an ANN was trained using the information from Table 6-4 in order to map the input features to the composite transverse modulus. Since the model was trained by the data from the first 500 hours of immersion, its accuracy in predicting material degradation after this time was tested with new data. It was carried out by checking the ANN model predictions using the information outside the range of training parameters. In order to avoid overfitting to the training data, the 10-fold cross validation (CV) method was used to check the model's accuracy with the validation folds. A Bayesian optimisation approach, as mentioned in Chapter 4, was also used to optimise the hyperparameters of the prediction model. The selected values for different hyperparameters of the ANN model are listed in Table 6-5. The optimised model comprises 3 hidden layers with 3, 5, and 4 neurons in each layer activated by Tanh functions. $\lambda = 0.157$ and Glorot function were selected as generalisation term and weight initialiser during the hyperparameter tuning. The optimal architecture led to a minimum CV error of 0.349 at the 17th iteration of optimisation.

The flowchart for developing an optimised ensemble model based on the knowledge from the multiple scale simulations is presented in Figure 6-15.

Table 6-5. List of selected hyperparameters for optimised ANN model

Hyperparameter	Search range	Optimal value
Number of layers	1 – 3	3
Layer size	1 – 300	3, 5, 4
Activation function	Relu, Tanh, Sigmoid	Tanh
λ	1.04e(-8) – 104.16	0.157
Weight initialiser	Glorot, He	Glorot

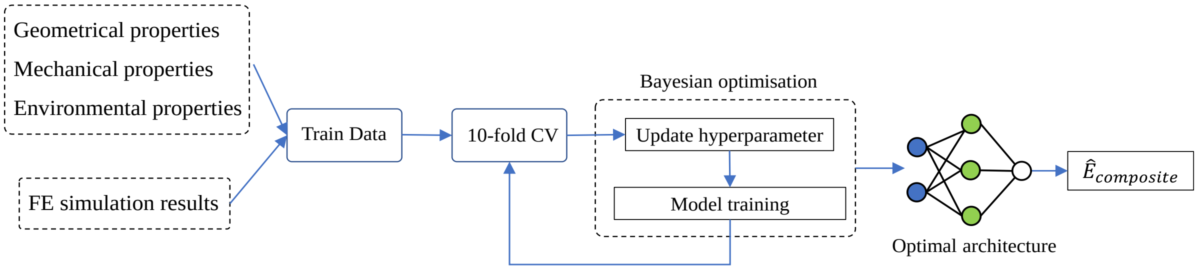


Figure 6-15. Flowchart of ANN model optimisation and development

6.4.1 Machine learning model validation

The prediction performance of the proposed framework was validated by the $RMSE$ and R^2 metrices. In order to see if the ANN was able to model the material degradation after 500 hours, its predictions were also tested with the numerical simulation results. In doing so, the multiscale framework was used to further solve the water diffusion and stress problems for 600–1200 hours of immersion with 328 simulations. The resulting transverse modulus values were then checked with the predictions of ANN model that was trained by the simulations of the first 500 hours of conditioning.

Table 6-6 shows the values of the metrices for the training, validation, and test datasets. These values imply that the developed model is able to predict material degradation with high accuracy. The performance of ANN in predicting the new data indicates that it can model the ageing after 500 hours with a minimum error from the numerical simulation results. This also implies that the model can generalise to new data without sacrificing its accuracy.

Table 6-6. Values of evaluation metrices for ensemble

Model	10-fold CV					
	Training folds		Validation folds		Test data	
	RMSE	R^2	RMSE	R^2	RMSE	R^2
ANN	0.0018	0.999	0.0018	0.999	0.0009	0.981

In addition, to visualise the accuracy of the machine learning model, the correlation between the FE simulations and the ANN predictions is illustrated in Figure 6-16. This figure compares the transverse modulus values between the numerical and machine learning approaches for the composite model immersed in water for 1200 hours. It is seen that the predictions are located within ± 1 % of the numerical simulations, indicating the strong correlation between the two methods. This implies that the developed prediction model is able to accurately map the geometrical, mechanical, and environmental properties of the composite material to its transverse properties during ageing under water.

A comparison of the computation cost required for modelling material degradation was made to demonstrate the computational efficiency of ANN over the multiscale FE model. Table 6-7 shows the computation time required for the multiscale modelling of composite ageing under 500 hours of water immersion conditions. It is seen that it took 647.77 seconds for the FE approach to model the degradation in the transverse properties of the composite plate with a 50% fibre volume fraction. However, it is seen that the developed ANN model computes the material degradation under 500 hours of exposure in 0.012 seconds, implying its superior computational efficiency when compared with the multiscale FE approach.

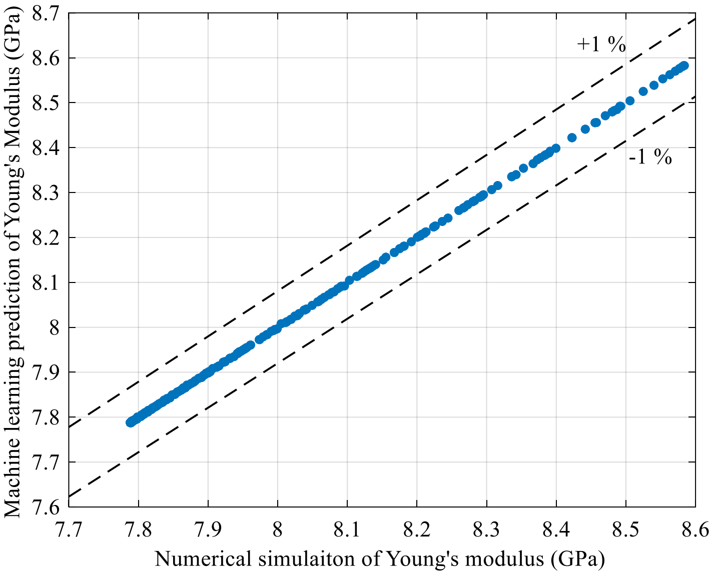


Figure 6-16. Correlation between the multiscale FE simulations and the ANN model predictions

Table 6-7. Computation time required for 500 hours of material degradation modelling using multiscale FE and ANN approaches

Model	Computation time (seconds)
Multiscale FE	647.77
ANN	0.012

6.4.2 Influence of input parameters on composite material degradation

In order to rank the contribution of geometrical, mechanical, and environmental factors on the degradation of composite, the permutation analysis was performed, as described in Chapter 3. Thereby, the out-of-bag observations of the features were first randomly permuted and used to estimate the ANN model error. The importance of parameters was then ranked based on the error raised by permuting their values. Figure 6-17 shows the contributions of input parameters in the composite transverse properties. It is seen that environmental properties have the most influence on the material degradation, followed by the mechanical and geometrical features of the composite.

Considering the high contribution of environmental and mechanical properties in determining the long-term performance of the composite material, a sensitivity analysis was performed using the ANN approach. Specifically, the robustness of ANN predictions to changes in the input features was investigated in a parametric study. Figure 6-18a depicts the change in the values of the transverse modulus with time for a composite plate with different fibre volume fractions. It is seen that with the progression of the water immersion condition, the material loses its Young's modulus with time. This implies the adverse impact of water on the composite material, particularly the epoxy matrix, during exposure, which is a result of the irreversible change and breakage of the polymer chain due to hydrolysis [29]. It is also seen that the degradation rate drops with time, which is proportional to the water content levels inside the material. At the early stages of immersion, due to the larger gradients between the surface and interior of the material, the moisture content increases, which results in a lower diffusivity and a slower degradation rate of the composite.

Figure 6-18b depicts the adverse impact of the moisture content on the material properties of composite. It is seen that the transverse properties decrease with the increase in moisture content, which is due to the degradation in the matrix and the interphase region proportionate to the water concentration value in the material.

Figure 6-18c shows the positive relationship between the transverse properties of the composite material constituents, matrix and interphase, and the material's Young's modulus. This figure shows that as the transverse properties of these constituents improve, the Young's modulus of the composite also increases, underscoring the critical role of the matrix and interphase region in determining the overall mechanical performance of the composite plate.

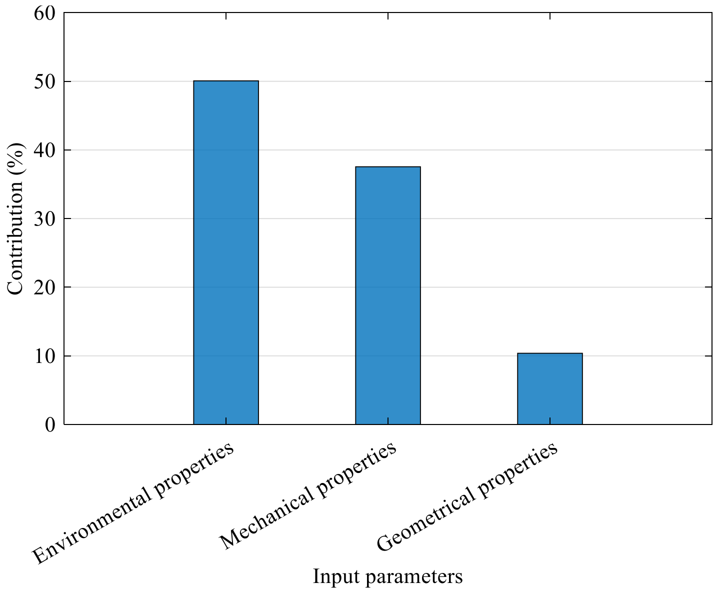
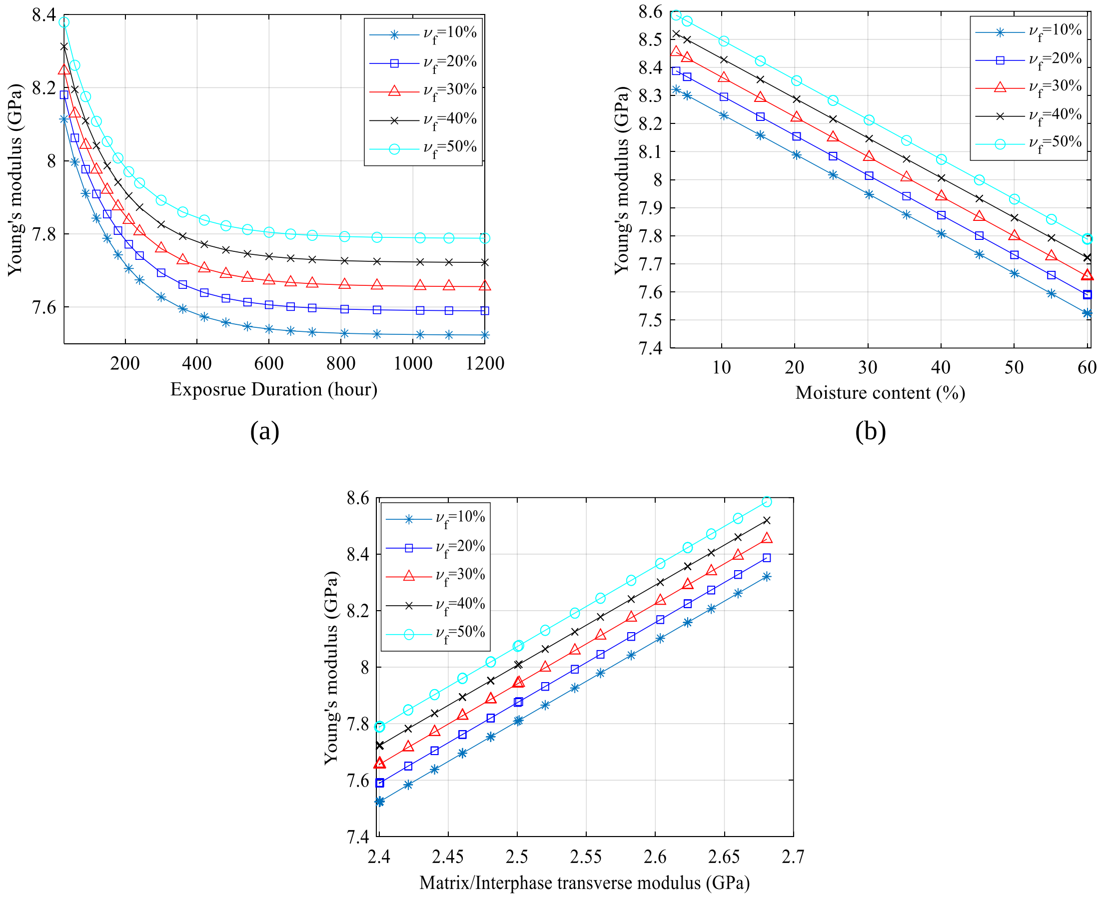


Figure 6-17. Contribution of input parametrs on composite material degradation



(a) (b)

(c)

Figure 6-18. Variations in composite material's Young's modulus under water immersion conditions: (a) exposure duration, (b) moisture content, (c) matrix/interphase transverse modulus

6.5 Conclusions

In this chapter, a framework was proposed in which a numerical scheme was integrated with the machine learning approach to model the process of material ageing on the FRP composites. The degradation of the composite was studied under the water immersion condition. Given the knowledge that the moisture ingress into the material is a non-uniform process, a multiscale framework based on computational homogenisation was proposed to model the composite ageing process. This required solving the moisture diffusion and stress problems at multiple scales and obtaining the homogenised response of the composite.

Analytical and numerical studies were performed to understand the kinetics of moisture diffusion into the composite material immersed in water. In doing so, first, Fick's law was used to simulate the stages of water uptake in the epoxy resin and UD glass fibre composite samples. The fitting of gravimetric curves to the experimental observations obtained the diffusivity of these materials. The information from the resin water uptake was then used in the numerical modelling of the composite moisture absorption. An analogy between Fickian mass diffusion and Fourier's law of heat transfer was established to model moisture diffusion into the composite. Comparing the FE simulation results with the experimental data revealed that the diffusivity of the epoxy matrix was modified by the presence of fibres in the composite. The evolution of moisture concentration along the length of the composite was also numerically modelled, and it was seen that the sample reached saturation after 800 hours of immersion in water.

The knowledge from water diffusion kinetics was then used to employ the multiscale framework in order to model the composite material's degradation with time. This required solving a macroscale diffusion problem. The composite diffusivity obtained from fitting Fick's law to the experimental observations was first used to simulate moisture uptake by a FE macro-model. The resulting water content values were then used to model material degradation at the microscale. An RVE with a 50% fibre volume fraction was created for each integration point, and its constitutive material properties were modified based on the water content value of the corresponding point in the macro-model. A FE simulation was performed to model each RVE stress response and obtain the effective traverse properties by homogenisation. In order to validate the multiscale framework, the response of the unaged and aged composite samples under water immersion conditions was modelled. It was seen that the transverse modulus estimations at the initial and saturated stages had a very small errors (i.e., 1.8% and 1.0%) from the three-point-bending experiment results, respectively.

Given the computational effort required in the numerical method, a machine learning ANN model was developed to surrogate the FE simulations at the microscale. First, the multiscale response of a composite sample immersed in water for 500 hours was obtained using the multiscale framework. This involved solving 533 micro-stress problems for different moisture concentration values. A database was then generated based on the simulation results, which included information about the composite material at multiple scales. A Bayesian optimisation technique was used together with a 10-fold CV method to tune the hyperparameters and avoid overfitting to the training data. The optimised model was then used to predict the transverse behaviour during the 1200 hours of immersion. The high correlation between the predictions and numerical results implied the ability of the ANN model to accurately map the geometrical, mechanical, and environmental properties of the composite material to its transverse properties during ageing. It was discovered that, although the model was trained by the FE simulations of the first 500 hours of moisture uptake, it was able to make accurate predictions beyond its training range. The ageing of the composite material during exposure was also modeled by the ANN model. It was seen that the transverse properties degraded with time, which was related to the chemical reaction of the epoxy matrix with the water molecules. In addition, the drop in the degradation level with time was due to the change in moisture concentration gradients with time. Particularly, the increase in exposure time and moisture content resulted in lower diffusivity and a slower degradation rate for the composite. A study on the influence of the input parameters on the material degradation was carried out by permuting the out-of-bag observations and calculating the ANN prediction errors. It was seen that water immersion conditions had the most influence, followed by the mechanical and geometrical properties of the composite material.

The current study helped us understand how immersing the UD FRP composite in water impacts its mechanical properties with time. The integration of the proposed multiscale framework with the machine learning was a promising approach to accurately model the material degradation with time. However, it should be noted that although machine learning is an effective technique to accelerate multiscale modelling of ageing, it requires considerable computational effort for development. The findings of the current study can guide the development of maintenance strategies for composite structures exposed to water-immersion conditions. By estimating the degradation of composite material, maintenance intervals can be scheduled accordingly in order to prevent unexpected failures in the composite structures and optimise the maintenance resources. Furthermore, understanding the degradation enables a more accurate estimation of the service life of composite structures. By taking into account the anticipated degradation rate, the engineers can determine how long it will take for the composite material to reach the critical performance threshold. This information is useful for decision-making and planning maintenance strategies necessary for the composite structures. In addition, understanding material failure can guide risk mitigation strategies for composite structures in water-immersed environments.

CHAPTER 7

Conclusions

In this research, the use of machine learning to improve the computational cost and accuracy in the analysis of ageing infrastructures at multiple scales was investigated. The thesis demonstrated the innovative application of machine learning models and their efficiency in accurately predicting the ageing mechanisms in infrastructure, structural elements, and composite materials. Four novel frameworks were developed to study the influence of ageing in structures from macro- to microscales. In Chapter 3, a data-driven machine learning-based maintenance approach was proposed to predict potential failures in a railway system. It was concluded that anticipating defects by learning from the history of the assets can aid in scheduling effective predictive strategies for infrastructure maintenance. In Chapter 4, the ageing in concrete structures strengthened/reinforced with FRP composites was discussed. The proposed framework was found to be capable of capturing the ageing effects in externally and internally reinforced concrete elements under various environmental conditions. In Chapters 5 and 6, the influence of ageing process on the performance of FRP composite materials was analysed. It was seen that the integration of numerical and machine learning approaches leads to an efficient solution to modelling the interaction of composite with the surrounding environment at the macro- and microscales.

This research extended the current knowledge on the long-term performance of structures by leveraging machine learning techniques and showcasing their successful integration within the developed framework. By conducting comprehensive evaluations and comparisons of different machine learning models and assessing their performance, the most suitable approaches were identified for the multiscale ageing problem. Additionally, by demonstrating the successful applications of machine learning-based prediction frameworks, the thesis contributed to the field by providing practical insights for structure maintenance and strengthening.

Based on the findings of each chapter, the following conclusions about the long-term performance of ageing infrastructures at various scales can be drawn.

7.1 Conclusions

7.1.1 Analysis of ageing in infrastructures

Sydney Trains railway network is an old infrastructure that suffers from frequent shutdowns for regular maintenance activities [174]. Besides being costly, this issue impacts the serviceability of the infrastructure and leads to environmental concerns such as noise and air pollution. This demonstrates the lack of an effective strategy for maintaining the railway system. In our study, we established a comprehensive dataset based on 4.5 years of maintenance records and used it to develop a framework for modelling defects in the different assets of the railway network. The data was processed, and features were explored to select the ones that were important for analysing the railway assets. It was accomplished by ranking the features using the permutation technique. Considering the noisy nature of the data, an ensemble model of decision trees was trained for feature importance analysis. It was seen that the failure was impacted by the asset quantity and usage properties within the network. Also, the number of assets greatly impacted the failure rate, while local properties slightly influenced the defects.

A two-stage classification model was developed to predict the condition of the assets within the network. Five classification algorithms, namely, decision tree, naïve Bayes, support vector machine (SVM), k-nearest neighbour (KNN), and ensemble, were first trained for identifying the presence of defects in the network. The machine learning models were tuned for best prediction performance using Bayesian optimisation and the 10-fold cross validation (CV) technique. Each classifier was trained 30 times in search of optimal hyperparameters, resulting in a minimum CV error. A comparison between the out-of-fold predictions showed that an ensemble model of 498 decision trees was able to outperform the remaining models in detecting defective assets across the network with 90% accuracy. Nominating ensemble as an accurate predictor of faulty assets, it was used in the second stage of defect type identification. The prediction results from the first stage were undersampled using the Tomek Links method to balance the data with the majority classes. The resulting dataset, comprising 1669 maintenance records, was passed to the ensemble model for a multiclass classification approach. It was observed that the second model was able to identify the type of defective asset with 93.7% accuracy, implying good agreement between the predictions and observations of the historical data.

A regression analysis was performed to model the influence of different infrastructure characteristics on the defects. It was done by training four regression models: linear regression (LR), decision tree, support vector regression (SVR), and artificial neural network (ANN). A hyperparameter tuning attempt was conducted by training the model with 80% of the data using Bayesian optimisation. Through sampling the training data with 10 folds during optimisation, it was obtained that the ANN model of three hidden layers with 256, 21, and 299 neurons in each layer and the Relu activation function were able to predict the number of defects from the selected input features. A comparison between the prediction results of different models demonstrated the superior performance of ANN, with 0.806 and 0.772 agreement between the prediction and real values, respectively. Therefore, ANN was used to perform a parametric study of the asset defects with respect to the location, usage, and number of the assets in the network. It was seen that the assets in further locations from the inspection start point had fewer defects. An initial decrease followed by an increase in the failures caused by the load and service count values was related to the asset's age, service life, and load-bearing capacity. It was also discovered that the total number of defects increased with the asset count in the network.

Practical equations were also developed based on the maintenance history of infrastructure in order to predict the condition of the railway system. Closed-form equation of the logistic and linear regression models were used to identify the faulty assets and estimate the number of asset failures, respectively. The proposed equation was able to accurately model the asset defects across the infrastructure network.

7.1.2 Analysis of ageing in FRP-strengthened concrete structures

In addition to maintenance practices, strengthening techniques are options for extending the service life of ageing infrastructure. Strengthening concrete structures by the external application of fibre-reinforced polymer (FRP) composites or reinforcing them with FRP bars is effective in improving their durability in service. Despite increasing the load-bearing capacity, the long-term performance of the composite structure is compromised due to exposure to extreme environments. Particularly, the bond between FRP and concrete is susceptible to deterioration with time. Therefore, a machine learning framework was developed to model the durability of the bond in response to different environmental conditions.

Two comprehensive databases were established on the durability of the bond in the externally and internally FRP-bonded concrete members subjected to various environmental conditions. For the FRP-to-concrete adhesively bonded connections, five regression models, including ridge, lasso, decision tree, SVM, and ANN, were developed to predict the bond strength under moisture conditions. A 10-fold VC-based Bayesian optimisation was first applied to 80% of the data as a training set to tune the machine learning hyperparameters. The detailed evaluation of prediction results showed that ANN with 30 and 5 neurons at two hidden layers had the best performance in predicting bond strength, with 0.952 and 0.953 values of R^2 on training and test sets, respectively. A visualisation of model predictions confirmed that ANN was able to assign low prediction errors and high correlation values to the train and test data. A study by ANN on the effects of conditioning regimes showed that different continuous and cyclic moisture conditioning regimes adversely impacted the integrity of the bond. In order to accurately predict the bond strength based on the geometrical, mechanical, and environmental properties of the connections, a practical equation using the M5P decision tree method was proposed. It was seen that for a minimum of 6 instances, M5P built a tree with 12 leaves, which predicted the training data with $R^2 = 0.938$.

To identify the failure mode of the adhesively bonded connections, a classification task was followed by training five models of decision trees, naïve Bayes, SVM, KNN, and ensemble. An optimal prediction performance for each model was obtained by applying Bayesian optimisation during the model training. Considering that the data distribution was imbalanced among classes, a stratified 10-fold CV was applied to ensure sampling of data from a minority class during model optimisation. The classification performance was checked by the confusion matrices. A test by the out-of-fold predictions assigned the highest accuracy value of 86.2% to the ensemble model in identifying the failure model of connection under moisture conditions compared to the remaining models. The parametric study of the environmental factors by ensemble demonstrated that the duration of exposure was the governing factor in determining the type of failure, followed by relative humidity (RH), sustained load, and temperature parameters.

The durability of the internally bonded FRP-concrete structures was studied by modelling the bond performance between FRP bars and concrete. A machine learning framework was developed to analyse the durability of the connection under exposure to moisture, chemicals, and varying temperature conditions. A comparison among the predictions of the linear classification (LC), ensemble, SVM, and ANN showed that an ensemble model tuned by 497 decision tree classifiers with a minimum number of 5 observations per leaf node identified the failure mode with 94.1% accuracy. The analysis of the environmental factors by ensemble indicated the significant influence of exposure duration on identifying the failure mode. A regression analysis was conducted to model the degradation of the FRP-to-concrete bond by LR, ensemble, SVR, and ANN. It was seen that an architecture of ANN with 4 hidden layers, including 29, 24, 21, and 17 neurons outperformed other regression models in predicting bond strength. A comparison between the experimental measurements and bond strength predictions revealed a good performance of ANN in training, validating, and testing with real observations. A comparison between the ANN results and the code formulations demonstrated that various standards use higher safety factors for the durability design of the FRP-reinforced concrete structures. A parametric study on the influence of different conditioning regimes on the durability of the bond by ANN showed the detrimental effect of chemical environments, especially at low pH levels and high salt concentrations. Two practical equations based on LC and LR models were also formulated to mathematically capture the relationships between the failure mode and bond strength, with the geometrical, mechanical, and environmental characteristics of the FRP bar and concrete, respectively.

7.1.3 Analysis of ageing in composite materials

Although analysis of ageing at the structural scale provides valuable information on the durability of composite structures, it often misses the underlying interaction of the materials with the service environments. A micromechanics-based machine learning framework was proposed to analyse the long-term performance of unidirectional (UD) glass fibre/epoxy composites under environmental ageing. A mathematical framework was developed to automatically create UD composite RVEs with periodic microstructures while accounting for the interphase region. The generated microstructures were used by a finite element (FE) numerical campaign to model the transverse behaviour of the FRP composites. It was seen that different characteristics, including mesh, boundary conditions, RVE size, fibre distribution, and interphase size, impact the effective mechanical properties of the composite. A parametric study revealed that the RVEs, which were meshed with the CPE4 element of 1.5 μm length and constrained by periodic boundary condition, provided converged responses. In addition, higher values of fibre volume fraction contributed to the impact of random fibre distributions and interphase thickness on the effective elastic properties of the composites. The study on the influence of the fibre characteristics showed that both carbon and glass-type fibres modified the transverse behaviour. It was seen that fibre type and volume fraction modified the effective modulus of the composite materials.

Although numerical modelling is helpful in analysing the ageing impact of composite materials, its integration with machine learning provides a computationally fast and accurate framework. A regression analysis of different machine learning models in Chapters 3 and 4 demonstrated that ANN was an accurate model in predicting continues output variables. Therefore, in Chapter 5, an ANN-based framework was developed by generating 1,200 microstructure RVEs and analysing them using the FE method. The generated data was processed by the two-point correlation (TPC) functions and the principal component analysis (PCA) method, and the important geometric features were extracted from the microstructure images. A database was established using the extracted physical descriptors and the mechanical and environmental properties of the composite microstructures. The data was then divided into 80% training and 20% test sets. An integration of 10-fold CV with hyperparameter tuning technique using Bayesian optimisation resulted in ANN model with 1 hidden layer with 5 neurons and the Sigmoid activation function with $\lambda = 1.67e(-8)$ and Glorot weight initialiser. The resultant architecture obtained a CV error of 0.0014 at the 10th step of hyperparameter optimisation. The prediction results demonstrated that the developed model was able to accurately predict the elastic modulus of microstructures with varying fibre volume fractions and distributions. Particularly, the trained ANN was able to map the important geometric features of microstructures from the training, validation, and test sets to their effective mechanical properties under humid conditions with a good accuracy. A parametric study was performed by the trained ANN model to analyse the influence of the geometric features on the composite transverse behaviour. It was seen that the volume fraction and size of fibres greatly impacted the effective elastic modulus of the UD composites.

In order to further extend our understanding of the ageing process, a multiscale FE and machine learning framework was developed to model the material degradation under water immersion conditions. The kinetics of moisture diffusion into the composite was first investigated by performing analytical and numerical studies. In doing so, Fick's law was used to simulate the water uptake in the epoxy resin and UD glass fibre composite samples. The results of the analytical study were then used to develop FE models of diffusion for the composite. It was discovered that the diffusivity of the epoxy matrix was modified by the presence of fibres in the composite.

The knowledge from the water diffusion kinetics was then used to model the composite material's degradation under water immersion by developing a multiscale numerical framework. In the proposed approach, a diffusion problem was solved by the FE simulation of moisture uptake in a macro-model. The resulting water concentration values were used to modify the constitutive material properties at the microscale. A micro-stress analysis was then performed to obtain the composite traverse properties. The proposed framework was validated by comparing its responses to the experimental observations of composite samples under water immersion conditions.

Given the computational effort required for FE simulations, a machine learning model was developed to replace the micro-stress analyses in the framework. First, A database using the simulation results from multiscale modelling of a composite sample exposed to 500 hours of immersion was generated. Similar to previous chapter, an ANN model was trained using the generated data. A Bayesian optimisation framework was proposed to develop a model with the best prediction performance. In order to avoid biased performance of ANN, a 10-fold CV was used for data sampling during the 30 epochs of hyperparameter tuning. It was demonstrated that the trained model with three hidden layers comprising 16, 30, and 14 neurons and the Relu activation function resulted in a minimum CV error of 1.94e(-7) at the 29th iteration. The developed ANN-based framework was able to accurately map the geometrical, mechanical, and environmental properties of the composite material to its transverse properties during ageing. We also discovered that, although the model was trained by the simulations of the first 500 hours of moisture uptake, it was able to make accurate predictions beyond its training range. The developed ANN method was used to model the ageing of the composite material with the immersion duration. It was seen that the material's transverse performance decreased with time, and the drop in the degradation level was proportional to the moisture concentration gradients during the diffusion. Finally, we performed a feature importance analysis to identify the contribution of different parameters to the degradation. It was seen that water immersion conditions had the most influence, followed by the mechanical and geometrical properties of the composite material.

In our studies, we saw that integrating numerical simulations with machine learning was an efficient way to accelerate material degradation modelling. It was accomplished by developing prediction models based on the FE analysis results. However, it should be noted that the accuracy of machine learning is dependent on the size and quality of the database. Generating data by FE simulations requires considerable computational effort, particularly for modelling complex physical phenomena. However, it is important to note that once trained, the machine learning model can perform predictions much faster, with less expensive computational effort, and without the need to perform further FE simulations.

7.2 Suggestions for future studies

The machine learning integrated frameworks in this thesis provided accurate and efficient models to analyse the infrastructural ageing at multiple scales. However, there are some limitations associated with the proposed frameworks, as listed in the following:

- The accuracy and generalisation ability of the machine learning models strongly depend on the quality of the training data. Biased and insufficient data can impact the model's ability to make accurate predictions.
- It was seen throughout the thesis that machine learning algorithms such as ANN and ensemble models were able to accurately map the input features to multiscale infrastructure ageing. However, the interpretation of how these models work is challenged due to their complexity. This can impact the reliability of these methods for making informed decisions for infrastructure maintenance and strengthening based on their results.
- Environmental parameters included in the datasets for machine learning training are limited to certain ranges. Extrapolating the machine learning prediction to environmental conditions outside of the parameters' training ranges can result in uncertain predictions.
- The proposed micromechanics-based ANN in Chapters 5 and 6 can make rapid and accurate predictions. However, it was required to simulate a large number of FE models to cover different material and environmental configurations for training machine learning models.

The following suggestions can be made for future studies:

- The proposed framework for railway infrastructure ageing was able to address the issue of ageing by modelling assets failure across the network. However, the model focused solely on detecting faulty assets and their numbers across the network. This limits the practicality of the proposed model in various applications such as assessing the severity of damages to the assets, predicting the remaining useful life, and analysing the cost associated with the required maintenance actions. Furthermore, the developed methodology overlooked the underlying mechanisms contributing to the asset failure. Predicting the type of defect in the railway can also allow for more targeted maintenance strategies. Also, including the impact of extreme environments in the analyses can help develop prediction models for the durability study of the assets. However, due to the limited information in the database, it was not possible to develop prediction models for these purposes. Therefore, a future direction of the research can be to incorporate more comprehensive data to train machine learning models for analysis of failure type, durability, and remaining service life of the infrastructure.

- The mentioned considerations can improve the efficiency of predictive maintenance planning, due to which unplanned shutdowns and disruptions to services can be reduced. Additionally, the capability to accurately predict the remaining useful life can aid in identifying critical assets for immediate action. An efficient predictive plan can also contribute to optimal resource and budget allocation for infrastructure maintenance, extending the service life of the assets.
- The frameworks developed to study the durability of the connections in the FRP-strengthened concrete structures were able to model the long-term strength and failure mode of the bond. However, the study focused on specific conditions, including moisture, water immersion, alkaline, acidic, and freeze-thaw environments. The durability prediction models for these conditions were developed based on a limited range of environmental parameters, which can create uncertainty around the performance estimation of composite structures under a broader range of environmental scenarios. Furthermore, this study primarily focused on bond strength and failure mode prediction. While these are important factors, it is essential to acknowledge that the performance of FRP-reinforced concrete structures is influenced by time-dependent behaviour of the concrete members such as creep, fatigue, and stress relaxation. Future research can consider a more comprehensive approach to address the limitations of the current study. This includes expanding the range of environmental parameters in the durability prediction models to capture a broader spectrum of exposure conditions. Additionally, integrating the time-dependent behaviour of the concrete and FRP materials during interaction with the environment will provide a deeper understanding of the ageing performance of composite structures under diverse environmental scenarios.
- Long-term performance prediction of the elements can aid in estimating the remaining useful life of FRP-strengthened concrete structures; thereby, maintenance activities can be planned in a targeted manner, which results in optimal resource allocation and minimised service downtime of the structures. Moreover, the predictive models can be integrated with

the structural health monitoring systems for continuous and real-time feedback on the structures' condition, leading to more efficient maintenance practices and improved overall performance of ageing structures.

- The developed micromechanics-based ANN framework was able to reproduce the experimental measurements of the transverse modulus of FRP composite materials. The application of a multiscale approach allowed for accurate analysis of the material ageing process under a limited range of humid and water immersion environmental parameters. However, immersion in water with different temperatures can greatly impact degradation behaviour and needs to be addressed in future studies. Considering the dependence of the material's chemical bond on the chemical composition of the immersion environment, multiscale ageing can be further studied under diverse immersion conditions, such as alkaline and acid. The micromechanical analysis of ageing included linear elastic models for the FRP material constituents. However, moisture can lead to plasticisation and swelling of the polymer. Therefore, a more accurate prediction of the material degradation can be obtained by incorporating damage models for the epoxy and the interface between fibre and matrix in the numerical simulation. This study investigated the material ageing of FRP composites at the microscopic scale. However, modelling the matrix and interphase ageing at this scale was not achievable without experimental calibrations. A further downscaling to the molecular level and modelling the interaction of FRP materials by advanced molecular dynamics simulations can improve the understanding of the ageing process of materials.

The ability to model the composite material's ageing within the proposed multiscale framework can aid in identifying potential damages in composite structures and planning timely maintenance interventions. This can help prevent costly failures and extend the service life of composite structures. Furthermore, by understanding the structure-property relationship, the framework can guide the selection of suitable geometrical, mechanical, and environmental parameters to design composite materials with desired properties. In addition, the ANN-based framework can reduce the need for costly experimental testing by providing estimations of composite material durability. By leveraging the computational efficiency of the proposed framework, informed decisions can be made regarding the use of composite materials for the strengthening of structures.

References

1. Cusson, D., Lounis, Z., and Daigle, L., Durability Monitoring for Improved Service Life Predictions of Concrete Bridge Decks in Corrosive Environments. Computer-Aided Civil and Infrastructure Engineering, 2011. 26(7): p. 524-541.
2. Sánchez-Silva, M. and Klutke, G.-A., Maintenance Concepts and Models, in Reliability and Life-Cycle Analysis of Deteriorating Systems, M. Sánchez-Silva and G.-A. Klutke, Editors. 2016, Springer International Publishing: Cham. p. 271-324.
3. Hollaway, L.C., A review of the present and future utilisation of FRP composites in the civil infrastructure with reference to their important in-service properties. Construction and Building Materials, 2010. 24(12): p. 2419-2445.
4. Salehi, H. and Burgueño, R., Emerging artificial intelligence methods in structural engineering. Engineering Structures, 2018. 171: p. 170-189.
5. Serradilla, O., Zugasti, E., Rodriguez, J., and Zurutuza, U., Deep learning models for predictive maintenance: a survey, comparison, challenges and prospects. Applied Intelligence, 2022: p. 1-31.
6. Program, I.I. Budget Announcements. [cited 2023; Available from: http://investment.infrastructure.gov.au/about/budget.aspx.
7. Australia, I. Australian Infrastructure Audit. 2019; Available from: https://www.infrastructureaustralia.gov.au/publications/australian-infrastructure-audit-2019.
8. Collins, F., & Blin, D, *Ageing of Infrastructure*, ed. I.A.L.-C. Approach. Vol. In A Life-Cycle Approach. 2018: CRC Press.
9. Hadigheh, S.A., Gravina, R.J., and Smith, S.T., Effect of acid attack on FRP-to-concrete bonded interfaces. Construction and Building Materials, 2017. 152: p. 285-303.
10. Liu, T., Liu, X., and Feng, P., A comprehensive review on mechanical properties of pultruded FRP composites subjected to long-term environmental effects. Composites Part B: Engineering, 2020. 191: p. 107958.
11. Neves, L.s.C., Frangopol, D.M., and Cruz, P.S., Cost of life extension of deteriorating structures under reliability-based maintenance. Computers & Structures, 2004. 82(13-14): p. 1077-1089.
12. Liu, M. and Frangopol, D.M., Optimal bridge maintenance planning based on probabilistic performance prediction. Engineering Structures, 2004. 26(7): p. 991-1002.
13. Frangopol, D.M. and Liu, M., Maintenance and management of civil infrastructure based on condition, safety, optimization, and life-cycle cost∗. Structure and Infrastructure Engineering, 2007. 3(1): p. 29-41.
14. Abdelrahman, A., Evaluation of Concrete Structures, in Strengthening of Concrete Structures: Unified Design Approach, Numerical Examples and Case Studies, A. Abdelrahman, Editor. 2023, Springer Nature Singapore: Singapore. p. 1-33.

15. UNE-EN 13306, *Maintenance - Maintenance terminology*. 2018, Génova, Madrid: Asociación Española de Normalización.

16. Deighton, M.G., Chapter 5 - Maintenance Management, in Facility Integrity Management, M.G. Deighton, Editor. 2016, Gulf Professional Publishing: Boston. p. 87-139.

17. Vaurio, J., On time-dependent availability and maintenance optimization of standby units under various maintenance policies. Reliability Engineering & System Safety, 1997. 56(1): p. 79-89.

18. Liao, L. and Köttig, F., A hybrid framework combining data-driven and model-based methods for system remaining useful life prediction. Applied Soft Computing, 2016. 44: p. 191-199.

19. Raza, S., Khan, M.K., Menegon, S.J., Tsang, H.-H., and Wilson, J.L., Strengthening and repair of reinforced concrete columns by jacketing: State-of-the-art review. Sustainability, 2019. 11(11): p. 3208.

20. Chen, J.F. and Teng, J., Anchorage strength models for FRP and steel plates bonded to concrete. Journal of structural engineering, 2001. 127(7): p. 784-791.

21. El-Zohairy, A., Salim, H., and Saucier, A., Steel–Concrete Composite Beams Strengthened with Externally Post-Tensioned Tendons under Fatigue. Journal of Bridge Engineering, 2019. 24(5): p. 04019027.

22. Nepomuceno, E., Sena-Cruz, J., Correia, L., and D'Antino, T., Review on the bond behavior and durability of FRP bars to concrete. Construction and Building Materials, 2021. 287: p. 123042.

23. Jahan, A., Edwards, K.L., and Bahraminasab, M., 3 - Screening of materials, in Multi-criteria Decision Analysis for Supporting the Selection of Engineering Materials in Product Design (Second Edition), A. Jahan, K.L. Edwards, and M. Bahraminasab, Editors. 2016, Butterworth-Heinemann. p. 41-61.

24. Daken, H.H., Ismail, A.A., and Rashed, F., Productivity improvement and resin-waste reduction in composites wet lay-up facilities. Journal of Materials Processing Technology, 1994. 41(1): p. 1-21.

25. Teoh, K.J. and Hsiao, K.-T., Improved dimensional infidelity of curve-shaped VARTM composite laminates using a multi-stage curing technique – Experiments and modeling. Composites Part A: Applied Science and Manufacturing, 2011. 42(7): p. 762-771.

26. Hayward, J.S. and Harris, B., The effect of vacuum assistance in resin transfer moulding. Composites Manufacturing, 1990. 1(3): p. 161-166.

27. Cabral-Fonseca, S., Correia, J.R., Custódio, J., Silva, H.M., Machado, A.M., and Sousa, J., Durability of FRP - concrete bonded joints in structural rehabilitation: A review. International Journal of Adhesion and Adhesives, 2018. 83: p. 153-167.

28. Chin, J.W., Aouadi, K., Haight, M.R., Hughes, W.L., and Nguyen, T., Effects of water, salt solution and simulated concrete pore solution on the properties of composite matrix resins used in civil engineering applications. Polymer composites, 2001. 22(2): p. 282-297.

29. Grammatikos, S., Zafari, B., Evernden, M., Mottram, J., and Mitchels, J., Moisture uptake characteristics of a pultruded fibre reinforced polymer flat sheet subjected to hot/wet aging. Polymer Degradation and Stability, 2015. 121: p. 407-419.

30. Khennane, A. and Melchers, R.E., Durability of Glass Polymer Composites Subject to Stress Corrosion. Journal of Composites for Construction, 2003. 7(2): p. 109-117.

31. Kootsookos, A. and Mouritz, A.P., Seawater durability of glass- and carbon-polymer composites. Composites Science and Technology, 2004. 64(10): p. 1503-1511.

32. Shrestha, J., Zhang, D., and Ueda, T., Durability Performances of Carbon Fiber–Reinforced Polymer and Concrete-Bonded Systems under Moisture Conditions. Journal of Composites for Construction, 2016. 20(5): p. 04016023.

33. Lai, W.L., Kou, S.C., Poon, C.S., Tsang, W.F., and Lai, C.C., Effects of elevated water temperatures on interfacial delaminations, failure modes and shear strength in externally-bonded CFRP-concrete beams using infrared thermography, gray-scale images and direct shear test. Construction and Building Materials, 2009. 23(10): p. 3152-3160.

34. Kabir, M.I., Shrestha, R., and Samali, B., Effects of applied environmental conditions on the pull-out strengths of CFRP-concrete bond. Construction and Building Materials, 2016. 114: p. 817-830.

35. Yun, Y. and Wu, Y.-F., Durability of CFRP–concrete joints under freeze–thaw cycling. Cold Regions Science and Technology, 2011. 65(3): p. 401-412.

36. Karbhari, V., Chin, J., Hunston, D., Benmokrane, B., Juska, T., Morgan, R., Lesko, J., Sorathia, U., and Reynaud, Durability gap analysis for fiber-reinforced polymer composites in civil infrastructure. Journal of composites for construction, 2003. 7(3): p. 238-247.

37. Colombi, P., Fava, G., and Poggi, C., Bond strength of CFRP–concrete elements under freeze–thaw cycles. Composite Structures, 2010. 92(4): p. 973-983.

38. Odegard, G.M. and Bandyopadhyay, A., Physical aging of epoxy polymers and their composites. Journal of Polymer Science Part B: Polymer Physics, 2011. 49(24): p. 1695-1716.

39. Polanský, R., Mentlík, V., Prosr, P., and Sušír, J., Influence of thermal treatment on the glass transition temperature of thermosetting epoxy laminate. Polymer Testing, 2009. 28(4): p. 428-436.

40. Moussa, O., Vassilopoulos, A.P., de Castro, J., and Keller, T., Time–temperature dependence of thermomechanical recovery of cold-curing structural adhesives. International Journal of Adhesion and Adhesives, 2012. 35: p. 94-101.

41. Aydin, H., Gravina, R.J., and Visintin, P., Durability of Adhesively Bonded FRP-to-Concrete Joints. Journal of Composites for Construction, 2016. 20(5): p. 04016016.

42. Woo, M.S. and Piggott, M.R., Water absorption of resins and composites: II. Diffusion in carbon and glass reinforced epoxies. Journal of Composites, Technology and Research, 1987. 9(4): p. 162-166.

43. Yang, Q., Xian, G., and Karbhari, V.M., Hygrothermal ageing of an epoxy adhesive used in FRP strengthening of concrete. Journal of Applied Polymer Science, 2008. 107(4): p. 2607-2617.

44. Rocha, I.B.C.M., van der Meer, F.P., Raijmaekers, S., Lahuerta, F., Nijssen, R.P.L., Mikkelsen, L.P., and Sluys, L.J., A combined experimental/numerical investigation on hygrothermal aging of fiber-reinforced composites. European Journal of Mechanics - A/Solids, 2019. 73: p. 407-419.

45. Kanit, T., Forest, S., Galliet, I., Mounoury, V., and Jeulin, D., Determination of the size of the representative volume element for random composites: statistical and numerical approach. International Journal of Solids and Structures, 2003. 40(13): p. 3647-3679.

46. Buryachenko, V.A., Pagano, N.J., Kim, R.Y., and Spowart, J.E., Quantitative description and numerical simulation of random microstructures of composites and their effective elastic moduli. International Journal of Solids and Structures, 2003. 40(1): p. 47-72.

47. Romanov, V., Lomov, S.V., Swolfs, Y., Orlova, S., Gorbatikh, L., and Verpoest, I., Statistical analysis of real and simulated fibre arrangements in unidirectional composites. Composites Science and Technology, 2013. 87: p. 126-134.

48. Vaughan, T. and McCarthy, C., A combined experimental–numerical approach for generating statistically equivalent fibre distributions for high strength laminated composite materials. Composites Science and Technology, 2010. 70(2): p. 291-297.

49. Melro, A.R., Camanho, P.P., and Pinho, S.T., Generation of random distribution of fibres in long-fibre reinforced composites. Composites Science and Technology, 2008. 68(9): p. 2092-2102.

50. Feder, J., Random sequential adsorption. Journal of Theoretical Biology, 1980. 87(2): p. 237-254.

51. Herráez, M., Segurado, J., González, C., and Lopes, C.S., A microstructures generation tool for virtual ply property screening of hybrid composites with high volume fractions of non-circular fibers – VIPER. Composites Part A: Applied Science and Manufacturing, 2020. 129: p. 105691.

52. Yang, L., Yan, Y., Ran, Z., and Liu, Y., A new method for generating random fibre distributions for fibre reinforced composites. Composites Science and Technology, 2013. 76: p. 14-20.

53. Gusev, A.A., Hine, P.J., and Ward, I.M., Fiber packing and elastic properties of a transversely random unidirectional glass/epoxy composite. Composites Science and Technology, 2000. 60(4): p. 535-541.

54. Greco, F., Leonetti, L., and Nevone Blasi, P., Adaptive multiscale modeling of fiber-reinforced composite materials subjected to transverse microcracking. Composite Structures, 2014. 113: p. 249-263.

55. Rocha, I.B.C.M. Numerical and experimental investigation of hygrothermal aging in laminated composites. PhD thesis, Delft University of Technology; 2019.

56. Joliff, Y., Belec, L., and Chailan, J.F., Modified water diffusion kinetics in an unidirectional glass/fibre composite due to the interphase area: Experimental, analytical and numerical approach. Composite Structures, 2013. 97: p. 296-303.

57. Joliff, Y., Rekik, W., Belec, L., and Chailan, J.F., Study of the moisture/stress effects on glass fibre/epoxy composite and the impact of the interphase area. Composite Structures, 2014. 108: p. 876-885.

58. Joliff, Y., Belec, L., Heman, M.B., and Chailan, J.F., Experimental, analytical and numerical study of water diffusion in unidirectional composite materials – Interphase impact. Computational Materials Science, 2012. 64: p. 141-145.

59. Rocha, I.B.C.M., Raijmaekers, S., Nijssen, R., Van Der Meer, F., and Sluys, L.J. Experimental/numerical study of anisotropic water diffusion in glass/epoxy composites. in *IOP Conference Series: Materials Science and Engineering*. 2016. IOP Publishing.

60. Rocha, I.B.C.M., van der Meer, F.P., Nijssen, R.P.L., and Sluys, L.J., A multiscale and multiphysics numerical framework for modelling of hygrothermal ageing in laminated composites. International Journal for Numerical Methods in Engineering, 2017. 112(4): p. 360-379.

61. Riaño, L., Chailan, J.-F., and Joliff, Y., Evolution of effective mechanical and interphase properties during natural ageing of glass-fibre/epoxy composites using micromechanical approach. Composite Structures, 2021. 258: p. 113399.

62. Quino, G., Tagarielli, V.L., and Petrinic, N., Effects of water absorption on the mechanical properties of GFRPs. Composites Science and Technology, 2020. 199: p. 108316.

63. Herráez, M., Mora, D., Naya, F., Lopes, C.S., González, C., and Llorca, J., Transverse cracking of cross-ply laminates: A computational micromechanics perspective. Composites Science and Technology, 2015. 110: p. 196-204.

64. Naya, F., González, C., Lopes, C.S., Van der Veen, S., and Pons, F., Computational micromechanics of the transverse and shear behavior of unidirectional fiber reinforced polymers including environmental effects. Composites Part A: Applied Science and Manufacturing, 2017. 92: p. 146-157.

65. Carvalho, T.P., Soares, F.A.A.M.N., Vita, R., Francisco, R.d.P., Basto, J.P., and Alcalá, S.G.S., A systematic literature review of machine learning methods applied to predictive maintenance. Computers & Industrial Engineering, 2019. 137: p. 106024.

66. Gibert, X., Patel, V.M., and Chellappa, R., Deep Multitask Learning for Railway Track Inspection. IEEE Transactions on Intelligent Transportation Systems, 2015. 18: p. 153-164.

67. Faghih-Roohi, S., Hajizadeh, S., Núñez, A., Babuska, R., and Schutter, B.D. Deep convolutional neural networks for detection of rail surface defects. in *2016 International Joint Conference on Neural Networks (IJCNN)*. 2016.

68. Yilboga, H., Eker, Ö.F., Güçlü, A., and Camci, F. Failure prediction on railway turnouts using time delay neural networks. in *2010 IEEE International Conference on Computational Intelligence for Measurement Systems and Applications*. 2010.

69. Chen, J., Liu, Z., Wang, H., Nunez, A., and Han, Z., Automatic defect detection of fasteners on the catenary support device using deep convolutional neural network. IEEE Transactions on Instrumentation and Measurement, 2017. 67(2): p. 257-269.

70. Jamshidi, A., Faghih-Roohi, S., Hajizadeh, S., Núñez, A., Babuska, R., Dollevoet, R., Li, Z., and De Schutter, B., A Big Data Analysis Approach for Rail Failure Risk Assessment. Risk Analysis, 2017. 37(8): p. 1495-1507.

71. Li, H., Parikh, D., He, Q., Qian, B., Li, Z., Fang, D., and Hampapur, A., Improving rail network velocity: A machine learning approach to predictive maintenance. Transportation Research Part C: Emerging Technologies, 2014. 45: p. 17-26.

72. Allah Bukhsh, Z., Saeed, A., Stipanovic, I., and Doree, A.G., Predictive maintenance using tree-based classification techniques: A case of railway switches. Transportation Research Part C: Emerging Technologies, 2019. 101: p. 35-54.

73. Ghofrani, F., Pathak, A., Mohammadi, R., Aref, A., and He, Q., Predicting rail defect frequency: An integrated approach using fatigue modeling and data analytics. Computer-Aided Civil and Infrastructure Engineering, 2020. 35(2): p. 101-115.

74. Zarembski, A.M., Einbinder, D., and Attoh-Okine, N., Using multiple adaptive regression to address the impact of track geometry on development of rail defects. Construction and Building Materials, 2016. 127: p. 546-555.

75. Schafer, D.H. and Barkan, C.P.L. A PREDICTION MODEL FOR BROKEN RAILS AND AN ANALYSIS OF THEIR ECONOMIC IMPACT. 2008.

76. Hu, C. and Liu, X. Modeling Track Geometry Degradation Using Support Vector Machine Technique. in *2016 Joint Rail Conference*. 2016.

77. Mashrei, M.A., Seracino, R., and Rahman, M.S., Application of artificial neural networks to predict the bond strength of FRP-to-concrete joints. Construction and Building Materials, 2013. 40: p. 812-821.

78. Zhou, Y., Zheng, S., Huang, Z., Sui, L., and Chen, Y., Explicit neural network model for predicting FRP-concrete interfacial bond strength based on a large database. Composite Structures, 2020. 240: p. 111998.

79. Haddad, R. and Haddad, M., Predicting fiber-reinforced polymer–concrete bond strength using artificial neural networks: A comparative analysis study. Structural Concrete, 2021. 22(1): p. 38-49.

80. Mansouri, I., Ozbakkaloglu, T., Kisi, O., and Xie, T., Predicting behavior of FRP-confined concrete using neuro fuzzy, neural network, multivariate adaptive regression splines and M5 model tree techniques. Materials and Structures, 2016. 49: p. 4319-4334.

81. Flood, I., Muszynski, L., and Nandy, S., Rapid analysis of externally reinforced concrete beams using neural networks. Computers & Structures, 2001. 79(17): p. 1553-1559.

82. Mansouri, I., Gholampour, A., Kisi, O., and Ozbakkaloglu, T., Evaluation of peak and residual conditions of actively confined concrete using neuro-fuzzy and neural computing techniques. Neural Computing and Applications, 2018. 29(3): p. 873-888.

83. Yang, C., Kim, Y., Ryu, S., and Gu, G.X., Prediction of composite microstructure stress-strain curves using convolutional neural networks. Materials & Design, 2020. 189: p. 108509.

84. Yang, Z., Yabansu, Y.C., Al-Bahrani, R., Liao, W.-k., Choudhary, A.N., Kalidindi, S.R., and Agrawal, A., Deep learning approaches for mining structure-property linkages in high contrast composites from simulation datasets. Computational Materials Science, 2018. 151: p. 278-287.

85. Cecen, A., Dai, H., Yabansu, Y.C., Kalidindi, S.R., and Song, L., Material structure-property linkages using three-dimensional convolutional neural networks. Acta Materialia, 2018. 146: p. 76-84.

86. Cang, R., Li, H., Yao, H., Jiao, Y., and Ren, Y., Improving direct physical properties prediction of heterogeneous materials from imaging data via convolutional neural network and a morphology-aware generative model. Computational Materials Science, 2018. 150: p. 212-221.

87. Kondo, R., Yamakawa, S., Masuoka, Y., Tajima, S., and Asahi, R., Microstructure recognition using convolutional neural networks for prediction of ionic conductivity in ceramics. Acta Materialia, 2017. 141: p. 29-38.

88. Pathan, M.V., Ponnusami, S.A., Pathan, J., Pitisongsawat, R., Erice, B., Petrinic, N., and Tagarielli, V.L., Predictions of the mechanical properties of unidirectional fibre composites by supervised machine learning. Scientific Reports, 2019. 9(1): p. 13964.

89. Ford, E., Maneparambil, K., Rajan, S., and Neithalath, N., Machine learning-based accelerated property prediction of two-phase materials using microstructural descriptors and finite element analysis. Computational Materials Science, 2021. 191: p. 110328.

90. Patel, D.K., Parthasarathy, T., and Przybyla, C., Predicting the effects of microstructure on matrix crack initiation in fiber reinforced ceramic matrix composites via machine learning. Composite Structures, 2020. 236: p. 111702.

91. Gu, G.X., Chen, C.-T., and Buehler, M.J., De novo composite design based on machine learning algorithm. Extreme Mechanics Letters, 2018. 18: p. 19-28.

92. Rail, I. Inland Rail Program Business Case. 2015; Available from: https://inlandrail.artc.com.au/inland-rail-program-business-case-2015/.

93. Baptista, M., Sankararaman, S., de Medeiros, I.P., Nascimento, C., Prendinger, H., and Henriques, E.M.P., Forecasting fault events for predictive maintenance using data-driven techniques and ARMA modeling. Computers & Industrial Engineering, 2018. 115: p. 41-53.

94. SAP. n.d.; Available from: https://www.sap.com/products/erp.html.

95. Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., and Dubourg, V., Scikit-learn: Machine learning in Python. the Journal of machine Learning research, 2011. 12: p. 2825-2830.

96. Witten, I.H. and Frank, E., *Data mining: practical machine learning tools and techniques*. 2017.

97. Guyon, I. and Elisseeff, A., An introduction to variable and feature selection. Journal of machine learning research, 2003. 3(Mar): p. 1157-1182.

98. James, G., Witten, D., Hastie, T., and Tibshirani, R., *An introduction to statistical learning*. Vol. 112. 2013, New York, NY: Springer.

99. Moreira, J., Soares, C., Jorge, A., and Sousa, J., Ensemble Approaches for Regression: A Survey. ACM Computing Surveys, 2012. 45: p. 10:1-10:40.

100. Freund, Y. and Schapire, R.E., A Decision-Theoretic Generalization of On-Line Learning and an Application to Boosting. Journal of Computer and System Sciences, 1997. 55(1): p. 119-139.

101. Friedman, J., Hastie, T., and Tibshirani, R., Additive logistic regression: a statistical view of boosting (with discussion and a rejoinder by the authors). The annals of statistics, 2000. 28(2): p. 337-407.

102. Seiffert, C., Khoshgoftaar, T.M., Hulse, J.V., and Napolitano, A. RUSBoost: Improving classification performance when training data is skewed. in *2008 19th International Conference on Pattern Recognition*. 2008.

103. Rasmussen, C. and Williams, C., *Gaussian Processes for Machine Learning*. 2006, Cambridge, Massachusetts: MIT Press.

104. Brownlee, J., *Machine learning mastery with Python: understand your data, create accurate models, and work projects end-to-end*. 2016: Machine Learning Mastery.

105. Breiman, L., Random Forests. Machine Learning, 2001. 45(1): p. 5-32.

106. Kuhn, M. and Johnson, K., *Applied predictive modeling*. Vol. 26. 2013: Springer.

107. Grossoni, I., Hughes, P., Bezin, Y., Bevan, A., and Jaiswal, J., Observed failures at railway turnouts: Failure analysis, possible causes and links to current and future research. Engineering Failure Analysis, 2021. 119: p. 104987.

108. Asada, T., Roberts, C., and Koseki, T., An algorithm for improved performance of railway condition monitoring equipment: Alternating-current point machine case study. Transportation Research Part C: Emerging Technologies, 2013. 30: p. 81-92.

109. Jabeen, H. and Baig, A.R., Two-stage learning for multi-class classification using genetic programming. Neurocomputing, 2013. 116: p. 311-316.

110. Aghabalaei Baghaei, K. and Hadigheh, S.A., Durability assessment of FRP-to-concrete bonded connections under moisture condition using data-driven machine learning-based approaches. Composite Structures, 2021: p. 114576.

111. TOMEK, I., Two Modifications of CNN. IEEE Transactions on Systems, Man, and Cybernetics, 1976. SMC-6(11): p. 769-772.

112. Hadigheh, S.A., Gravina, R.J., and Setunge, S., Influence of the Processing Techniques on the Bond Characteristics in Externally Bonded Joints: Experimental and Analytical Investigations. Journal of Composites for Construction, 2016. 20(3): p. 04015081.

113. Pan, Y., Shi, J., and Xian, G., Experimental and numerical study of the CFRP-to-concrete bonded joints after water immersion. Composite Structures, 2019. 218: p. 95-106.

114. Au, C. and Büyüköztürk, O., Peel and shear fracture characterization of debonding in FRP plated concrete affected by moisture. Journal of Composites for Construction, 2006. 10(1): p. 35-47.

115. Shrestha, J., Ueda, T., and Zhang, D., Durability of FRP concrete bonds and its constituent properties under the influence of moisture conditions. Journal of Materials in Civil Engineering, 2015. 27(2): p. A4014009.

116. Silva, M.A.G., Biscaia, H.C., and Marreiros, R., Bond–slip on CFRP/GFRP-to-concrete joints subjected to moisture, salt fog and temperature cycles. Composites Part B: Engineering, 2013. 55: p. 374-385.

117. Sarker, M., Hadigheh, S.A., and Dias-da-Costa, D., A performance-based characterisation of CFRP composite deterioration using active infrared thermography. Composite Structures, 2020. 241: p. 112134.

118. Hadigheh, S.A., Ke, F., and Kashi, S., 3D acid diffusion model for FRP-strengthened reinforced concrete structures: Long-term durability prediction. Construction and Building Materials, 2020. 261: p. 120548.

119. Eurocode 2, *Design of concrete structures - Part 1-1: General rules and rules for buildings*. 2004, Brussels, Belgium: British Standard. 230.

120. Kuhn, M. and Johnson, K., *Feature engineering and selection: A practical approach for predictive models*. 2019: CRC Press.

121. Witten, I.H., Frank, E., Hall, M.A., and Pal, C.J., *Data mining: practical machine learning tools and techniques*. 2017: Morgan Kaufman.

122. Yeo, I.K. and Johnson, R.A., A new family of power transformations to improve normality or symmetry. Biometrika, 2000. 87(4): p. 954-959.

123. ACI 440.1R-15, *Guide for the Design and Construction of Structural Concrete Reinforced with Fiber-Reinforced Polymer (FRP) Bars*. 2015, MI, USA: American Concrete Institute.

124. Altalmas, A., El Refai, A., and Abed, F., Bond degradation of basalt fiber-reinforced polymer (BFRP) bars exposed to accelerated aging conditions. Construction and Building Materials, 2015. 81: p. 162-171.

125. Li, J., Gravina, R.J., Smith, S.T., and Visintin, P., Bond strength and bond stress-slip analysis of FRP bar to concrete incorporating environmental durability. Construction and Building Materials, 2020. 261: p. 119860.

126. Eurocode 2, *Design of concrete structures - Part 1-1: General rules and rules for buildings*. 2004, Brussels, Belgium: British Standard.

127. Abbasi, A. and Hogg, P.J., Temperature and environmental effects on glass fibre rebar: modulus, strength and interfacial bond strength with concrete. Composites Part B: Engineering, 2005. 36(5): p. 394-404.

128. AS 5100.8, *Bridge design – Part 8: rehabilitation and strengthening of existing bridges*. 2017.

129. fib Bulletin No. 40, *FRP reinforcement in RC structures*. 2007.

130. Riaño, L. and Joliff, Y., An Abaqus™ plug-in for the geometry generation of Representative Volume Elements with randomly distributed fibers and interphases. Composite Structures, 2019. 209: p. 644-651.

131. Kari, S., Berger, H., Gabbert, U., Guinovart-Dıaz, R., Bravo-Castillero, J., and Rodrıguez-Ramos, R., Evaluation of influence of interphase material parameters on effective material properties of three phase composites. Composites Science and Technology, 2008. 68(3): p. 684-691.

132. Gitman, I.M., Askes, H., and Sluys, L.J., Representative volume: Existence and size determination. Engineering Fracture Mechanics, 2007. 74(16): p. 2518-2534.

133. Dassault Systèmes Simulia Corp., *Abaqus/CAE 2018*. 2017, RI, USA.

134. Gusev, A.A., Numerical Identification of the Potential of Whisker- and Platelet-Filled Polymers. Macromolecules, 2001. 34(9): p. 3081-3093.

135. Okereke, M.I. and Akpoyomare, A.I., A virtual framework for prediction of full-field elastic response of unidirectional composites. Computational Materials Science, 2013. 70: p. 82-99.

136. Eshelby, J.D., The determination of the elastic field of an ellipsoidal inclusion, and related problems. Proceedings of the royal society of London. Series A. Mathematical and physical sciences, 1957. 241(1226): p. 376-396.

137. Pierard, O., Friebel, C., and Doghri, I., Mean-field homogenization of multi-phase thermo-elastic composites: a general framework and its validation. Composites Science and Technology, 2004. 64(10): p. 1587-1603.

138. Geers, M.G.D., Kouznetsova, V.G., and Brekelmans, W.A.M., Multi-scale computational homogenization: Trends and challenges. Journal of Computational and Applied Mathematics, 2010. 234(7): p. 2175-2182.

139. Perić, D., de Souza Neto, E.A., Feijóo, R.A., Partovi, M., and Molina, A.J.C., On micro-to-macro transitions for multi-scale analysis of non-linear heterogeneous materials:

unified variational basis and finite element implementation. International Journal for Numerical Methods in Engineering, 2011. 87(1-5): p. 149-170.

140. Norme Française ET Européennes (NF EN 2746), *Série aérospatiale - Plastiques renforcés au verre textile - Essai de flexion. Méthode des trois pannes*. 1998: AFNOR.

141. Belec, L., Nguyen, T.H., Nguyen, D.L., and Chailan, J.F., Comparative effects of humid tropical weathering and artificial ageing on a model composite properties from nano- to macro-scale. Composites Part A: Applied Science and Manufacturing, 2015. 68: p. 235-241.

142. Bergeret, A. and Krawczak, P., Liaison renfort/matrice: Définition et caractérisation, in Techniques de l'Ingénieur. 2006, Editions T.I. p. AM5305.

143. Riaño, L., Belec, L., Chailan, J.-F., and Joliff, Y., Effect of interphase region on the elastic behavior of unidirectional glass-fiber/epoxy composites. Composite Structures, 2018. 198: p. 109-116.

144. Berthelot, J.M., *Matériaux composites*. TEC & DOC ed. Vol. 5. 2012, Paris.

145. Wongsto, A. and Li, S., Micromechanical FE analysis of UD fibre-reinforced composites with fibres distributed at random over the transverse cross-section. Composites Part A: Applied Science and Manufacturing, 2005. 36(9): p. 1246-1266.

146. Reuss, A., Berechnung der Fließgrenze von Mischkristallen auf Grund der Plastizitätsbedingung für Einkristalle. ZAMM - Journal of Applied Mathematics and Mechanics / Zeitschrift für Angewandte Mathematik und Mechanik, 1929. 9(1): p. 49-58.

147. Hashin, Z. and Shtrikman, S., On some variational principles in anisotropic and nonhomogeneous elasticity. Journal of the Mechanics and Physics of Solids, 1962. 10(4): p. 335-342.

148. Mantič, V., Interface crack onset at a circular cylindrical inclusion under a remote transverse tension. Application of a coupled stress and energy criterion. International Journal of Solids and Structures, 2009. 46(6): p. 1287-1304.

149. Martias, C., Joliff, Y., and Favotto, C., Effects of the addition of glass fibers, mica and vermiculite on the mechanical properties of a gypsum-based composite at room temperature and during a fire test. Composites Part B: Engineering, 2014. 62: p. 37-53.

150. Lejeunes, S. and Bourgeois, S., *Homtools: Homogeneization toolbox for Abaqus*. 2018.

151. Riaño, L., Belec, L., and Joliff, Y., Validation of a Representative Volume Element for unidirectional fiber-reinforced composites: Case of a monotonic traction in its cross section. Composite Structures, 2016. 154: p. 11-16.

152. Herráez, M., González, C., Lopes, C.S., de Villoria, R.G., Llorca, J., Varela, T., and Sánchez, J., Computational micromechanics evaluation of the effect of fibre shape on the transverse strength of unidirectional composites: An approach to virtual materials design. Composites Part A: Applied Science and Manufacturing, 2016. 91: p. 484-492.

153. Sih, G.C., Nait-Abdelaziz, M., and Vu-Khanh, T., *Particle and continuum aspects of mesomechanics: mesomechanics 2007*. Vol. 2. 2010: John Wiley & Sons.

154. Downing, T.D., Kumar, R., Cross, W.M., Kjerengtroen, L., and Kellar, J.J., Determining the interphase thickness and properties in polymer matrix composites using phase imaging atomic force microscopy and nanoindentation. Journal of Adhesion Science and Technology, 2000. 14(14): p. 1801-1812.

155. Owens Corning. 2023; Available from: https://www.owenscorning.com/en-us.

156. AGY Strength in Materials. 2022; Available from: https://www.agy.com/.

157. Toray Composite Materials America Inc. 2023; Available from: https://www.toraycma.com/.

158. Niezgoda, S.R., Fullwood, D.T., and Kalidindi, S.R., Delineation of the space of 2-point correlations in a composite material system. Acta Materialia, 2008. 56(18): p. 5285-5292.

159. Saheli, G., Garmestani, H., and Adams, B.L., Microstructure design of a two phase composite using two-point correlation functions. Journal of Computer-Aided Materials Design, 2004. 11(2): p. 103-115.

160. Swaminathan, S., Ghosh, S., and Pagano, N.J., Statistically Equivalent Representative Volume Elements for Unidirectional Composite Microstructures: Part I - Without Damage. Journal of Composite Materials, 2006. 40(7): p. 583-604.

161. Frisch, H.L. and Stillinger, F.H., Contribution to the Statistical Geometric Basis of Radiation Scattering. The Journal of Chemical Physics, 1963. 38(9): p. 2200-2207.

162. Gokhale, A.M., Tewari, A., and Garmestani, H., Constraints on microstructural two-point correlation functions. Scripta Materialia, 2005. 53(8): p. 989-993.

163. Brough, D.B., Wheeler, D., and Kalidindi, S.R., Materials Knowledge Systems in Python—a Data Science Framework for Accelerated Development of Hierarchical Materials. Integrating Materials and Manufacturing Innovation, 2017. 6(1): p. 36-53.

164. Kherif, F. and Latypova, A., Chapter 12 - Principal component analysis, in Machine Learning, A. Mechelli and S. Vieira, Editors. 2020, Academic Press. p. 209-225.

165. Brownlee, J., *Data preparation for machine learning: data cleaning, feature selection, and data transforms in Python*. 2020: Machine Learning Mastery.

166. Chiroma, H., Abdullahi, U.A., Abdulhamid, S.M., Alarood, A.A., Gabralla, L.A., Rana, N., Shuib, L., Hashem, I.A.T., Gbenga, D.E., Abubakar, A.I., Zeki, A.M., and Herawan, T., Progress on Artificial Neural Networks for Big Data Analytics: A Survey. IEEE Access, 2019. 7: p. 70535-70551.

167. Snoek, J., Larochelle, H., and Adams, R.P., Practical bayesian optimization of machine learning algorithms. Advances in neural information processing systems, 2012. 25.

168. Glorot, X. and Bengio, Y. Understanding the difficulty of training deep feedforward neural networks. in *Proceedings of the thirteenth international conference on artificial intelligence and statistics (AISTATS)*. 2010. Sardinia, Italy: JMLR Workshop and Conference Proceedings.

169. He, K., Zhang, X., Ren, S., and Sun, J. Delving deep into rectifiers: Surpassing human-level performance on imagenet classification. in *Proceedings of the IEEE international conference on computer vision*. 2015.

170. Luo, S., Leisen, J., and Wong, C., Study on mobility of water and polymer chain in epoxy and its influence on adhesion. Journal of applied polymer science, 2002. 85(1): p. 1-8.

171. Crank, J., *The mathematics of diffusion*. 1979: Oxford university press.

172. Bond, D.A., Moisture Diffusion in a Fiber-reinforced Composite: Part I – Non-Fickian Transport and the Effect of Fiber Spatial Distribution. Journal of Composite Materials, 2005. 39(23): p. 2113-2141.

173. Terada, K. and Kurumatani, M., Two-scale diffusion–deformation coupling model for material deterioration involving micro-crack propagation. International Journal for Numerical Methods in Engineering, 2010. 83(4): p. 426-451.

174. NSW, T.f. Maintenance and construction work. 2023; Available from: https://www.transport.nsw.gov.au/projects/community-engagement/sydney-trains-community/maintenance-and-construction-work.

175. Barger, J.D. Effects of aging on bond between FRP and concrete. Master's thesis, West Virginia University; 2000.

176. Pack, J.R. Environmental durability evaluation of externally bonded composites. Master's thesis, University of Cincinnati; 2003.

177. Tam, S.S.F. Durability of fibre reinforced polymer (FRP) and FRP bond subjected to freeze-thaw cycles and sustained load University of Toronto; 2007.

178. Subramaniam, K.V., Ali-Ahmad, M., and Ghosn, M., Freeze–thaw degradation of FRP–concrete interface: Impact on cohesive fracture response. Engineering Fracture Mechanics, 2008. 75(13): p. 3924-3940.

179. Gamage, J.C.P.H., Al-Mahaidi, R., and Wong, M.B., Durability of CFRP-strengthened concrete members under extreme temperature and humidity. Australian Journal of Structural Engineering, 2009. 9(2): p. 111-118.

180. Abbas, B.M.M. Durability of CFRP-concrete bond under sustained load in harsh environment. Ph.D. thesis, Monash University; 2010.

181. Benzarti, K., Chataigner, S., Quiertant, M., Marty, C., and Aubagnac, C., Accelerated ageing behaviour of the adhesive bond between concrete specimens and CFRP overlays. Construction and Building Materials, 2011. 25(2): p. 523-538.

182. Tuakta, C. and Büyüköztürk, O., Deterioration of FRP/concrete bond system under variable moisture conditions quantified by fracture mechanics. Composites Part B: Engineering, 2011. 42(2): p. 145-154.

183. Shi, J., Zhu, H., Wu, Z., Seracino, R., and Wu, G., Bond Behavior between Basalt Fiber–Reinforced Polymer Sheet and Concrete Substrate under the Coupled Effects of Freeze-Thaw Cycling and Sustained Load. Journal of Composites for Construction, 2013. 17(4): p. 530-542.

184. Kabir, M.I. Short and long term performance of concrete structures repaired/strengthened with FRP. Ph.D. thesis, University of Technology; 2014.

185. Pan, Y., Xian, G., and Silva, M.A.G., Effects of water immersion on the bond behavior between CFRP plates and concrete substrate. Construction and Building Materials, 2015. 101: p. 326-337.

186. Shrestha, J., Zhang, D., and Ueda, T., Durability performances of carbon fiber-reinforced polymer and concrete-bonded systems under moisture conditions. Journal of Composites for Construction, 2016. 20(5): p. 04016023.

187. Zheng, X.H., Huang, P.Y., Guo, X.Y., and Huang, J.L., Experimental study on bond behavior of FRP-concrete interface in hygrothermal environment. International Journal of Polymer Science, 2016. 2016: p. 5832130.

188. Gravina, R.J., Aydin, H., and Visintin, P., Extraction and analysis of bond-slip characteristics in deteriorated FRP-to-concrete joints using a mechanics-based approach. Journal of Materials in Civil Engineering, 2017. 29(6): p. 04017013.

189. Liu, S., Pan, Y., Li, H., and Xian, G., Durability of the bond between CFRP and concrete exposed to thermal cycles. Materials, 2019. 12(3): p. 515.

190. Al-Lami, K., Colombi, P., and D'Antino, T., Influence of hygrothermal ageing on the mechanical properties of CFRP-concrete joints and of their components. Composite Structures, 2020. 238: p. 111947.

191. Abedi, S. Evaluation of the bond and tensile strength of GFRP bars exposed to harsh environment. M.S. thesis, American Univ. of Sharjah; 2014.

192. Al-Tamimia, A., Abed Farid, H., and Al-Rahmani, A., Effects of harsh environmental exposures on the bond capacity between concrete and GFRP reinforcing bars. Advances in concrete construction, 2014. 2(1): p. 1-11.

193. Al-Dulaijan, S.U., Al-Zahrani, M.M., Nanni, A., Bakis, C.E., and Boothby, T.E., Effect of Environmental Pre-Conditioning on Bond of FRP Reinforcement to Concrete. Journal of Reinforced Plastics and Composites, 2001. 20(10): p. 881-900.

194. Alvarez, A., A. Zaidi, R. Mamsoudi. Bond-slip behaviour of FRP bars under low and high temperature - experimental and theoretical studies. in *3rd International Conference on Durability & Field Applications of Fiber Reinforced Polymer (FRP) Composites for Construction*. 2007. Quebec, Canada.

195. Al-Zahrani, M.M., Al-Dulaijan, S.U., Sharif, A., and Maslehuddin, M. Durability performance of glass fiber reinforced plastic reinforcement in harsh environments. in *6th Saudi engineering conference*. 2002.

196. Ammar, M.A. Bond durability of basalt fi bre-reinforced polymers (BFRP) bars under freeze-and-thaw conditions. M.S. thesis, Université Laval; 2014.

197. Calvet, V., Valcuende, M., Benlloch, J., and Cánoves, J., Influence of moderate temperatures on the bond between carbon fibre reinforced polymer bars (CFRP) and concrete. Construction and Building Materials, 2015. 94: p. 589-604.

198. Davalos, J.F., Chen, Y., and Ray, I., Effect of FRP bar degradation on interface bond with high strength concrete. Cement and Concrete Composites, 2008. 30(8): p. 722-730.

199. Dong, Z., Wu, G., Xu, B., Wang, X., and Taerwe, L., Bond durability of BFRP bars embedded in concrete under seawater conditions and the long-term bond strength prediction. Materials & Design, 2016. 92: p. 552-562.

200. Dong, Z.-Q., Wu, G., and Xu, Y.-Q., Bond and Flexural Behavior of Sea Sand Concrete Members Reinforced with Hybrid Steel-Composite Bars Presubjected to Wet–Dry Cycles. Journal of Composites for Construction, 2017. 21(2): p. 04016095.

201. Dong, Z., Wu, G., and Xu, Y., Experimental study on the bond durability between steel-FRP composite bars (SFCBs) and sea sand concrete in ocean environment. Construction and Building Materials, 2016. 115: p. 277-284.

202. Refai, A.E., Abed, F., and Altalmas, A., Bond Durability of Basalt Fiber–Reinforced Polymer Bars Embedded in Concrete under Direct Pullout Conditions. Journal of Composites for Construction, 2015. 19(5): p. 04014078.

203. Galati, N., Nanni, A., Dharani, L.R., Focacci, F., and Aiello, M.A., Thermal effects on bond between FRP rebars and concrete. Composites Part A: Applied Science and Manufacturing, 2006. 37(8): p. 1223-1230.

204. Hassan, M., Benmokrane, B., ElSafty, A., and Fam, A., Bond durability of basalt-fiber-reinforced-polymer (BFRP) bars embedded in concrete in aggressive environments. Composites Part B: Engineering, 2016. 106: p. 262-272.

205. Robert, M. and Benmokrane, B., Effect of aging on bond of GFRP bars embedded in concrete. Cement and Concrete Composites, 2010. 32(6): p. 461-467.

206. Yan, F., Lin, Z., Zhang, D., Gao, Z., and Li, M., Experimental study on bond durability of glass fiber reinforced polymer bars in concrete exposed to harsh environmental agents: Freeze-thaw cycles and alkaline-saline solution. Composites Part B: Engineering, 2017. 116: p. 406-421.

207. Yan, F. and Lin, Z., Bond durability assessment and long-term degradation prediction for GFRP bars to fiber-reinforced concrete under saline solutions. Composite Structures, 2017. 161: p. 393-406.

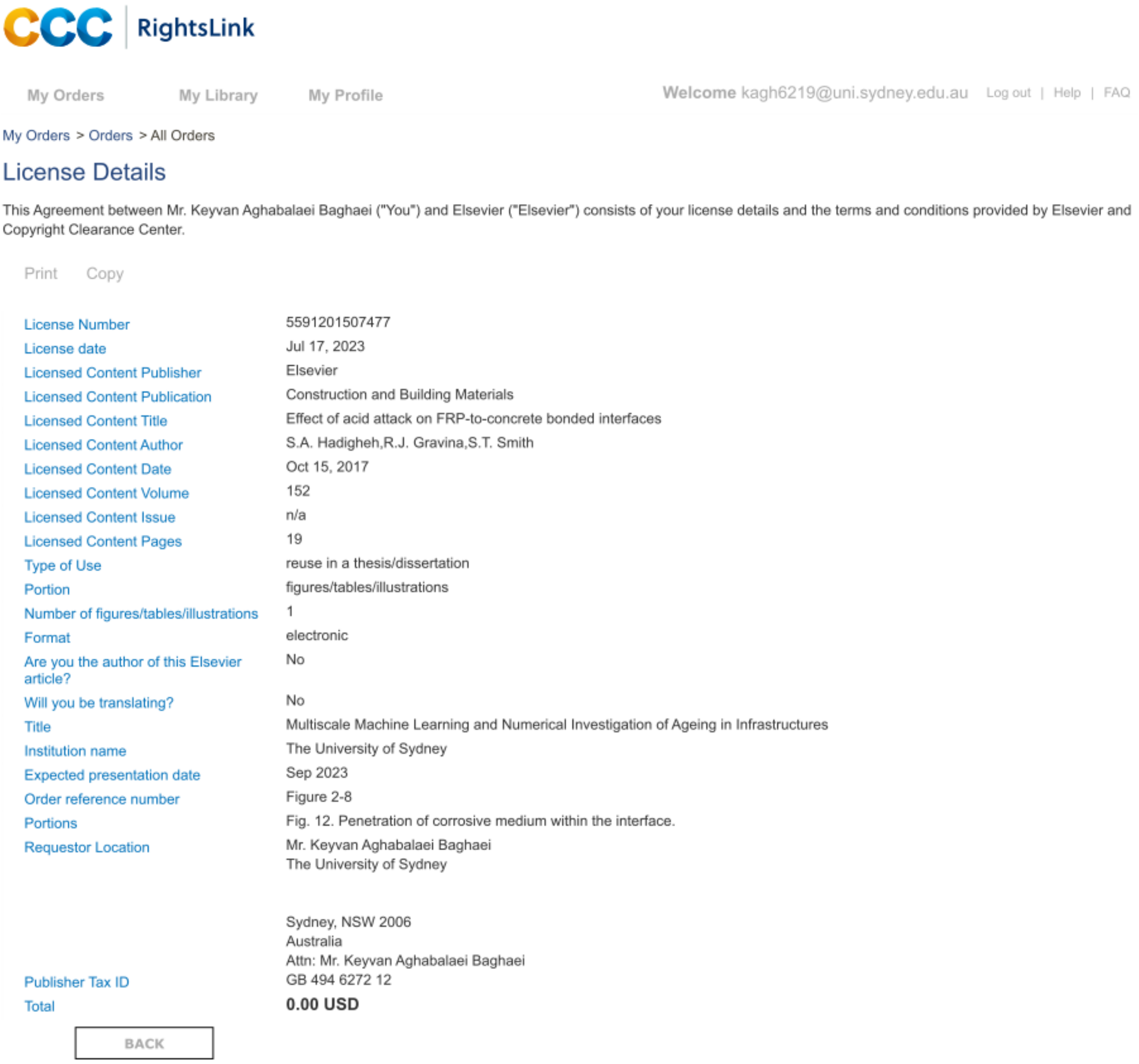
208. Zhou, J., Chen, X., and Chen, S., Durability and service life prediction of GFRP bars embedded in concrete under acid environment. Nuclear Engineering and Design, 2011. 241(10): p. 4095-4102.

209. Zhou, J., Chen, X., and Chen, S., Effect of different environments on bond strength of glass fiber-reinforced polymer and steel reinforcing bars. KSCE Journal of Civil Engineering, 2012. 16(6): p. 994-1002.

210. Mashima, M. and Iwamoto, K., Bond characteristics of FRP rod and concrete after freezing and thawing deterioration. Special Publication, 1993. 138: p. 51-70.

Appendix A

Copyright agreement

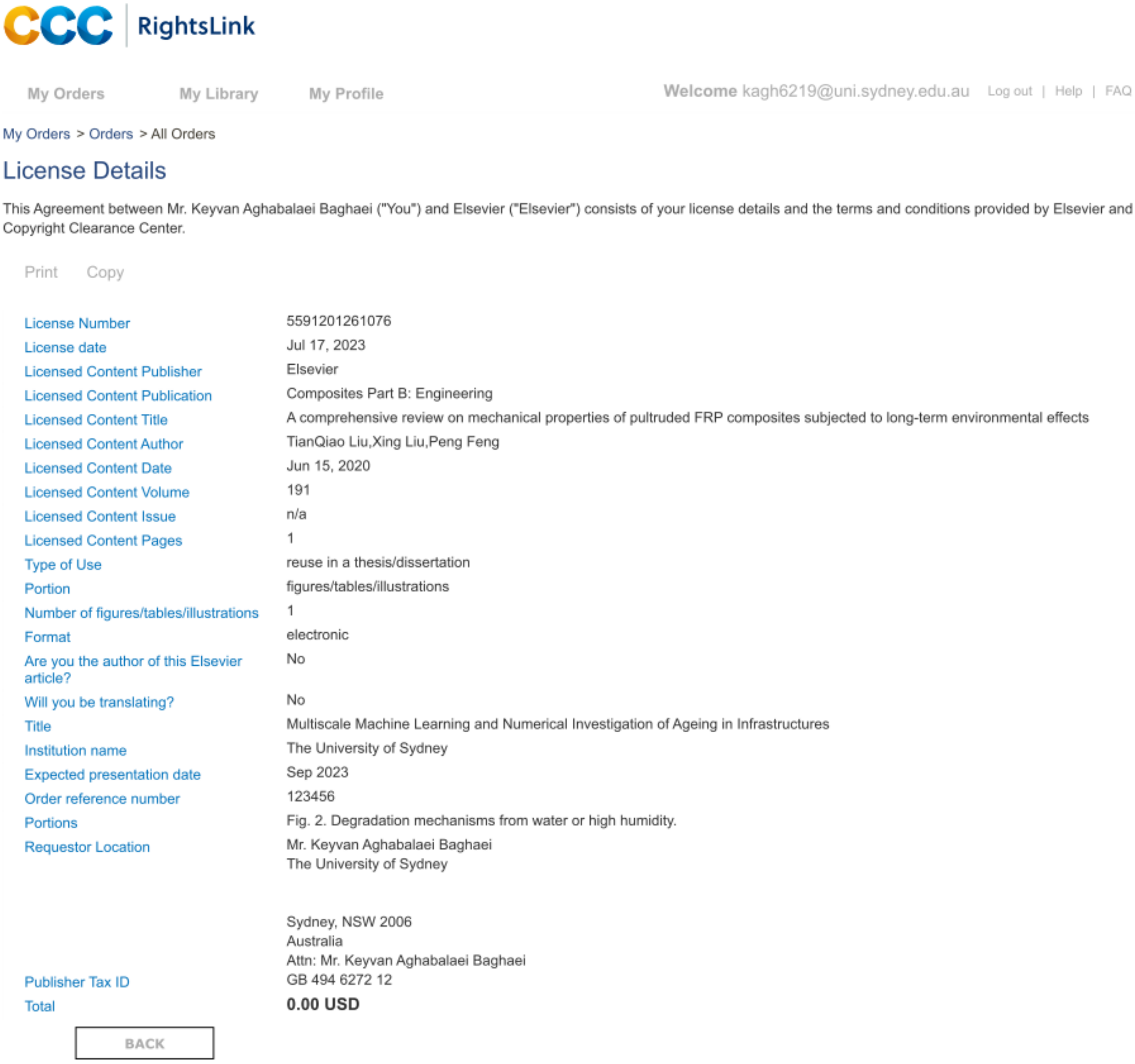


License Details

This Agreement between Mr. Keyvan Aghabalaei Baghaei ("You") and Elsevier ("Elsevier") consists of your license details and the terms and conditions provided by Elsevier and Copyright Clearance Center.

License Number	5591201507477
License date	Jul 17, 2023
Licensed Content Publisher	Elsevier
Licensed Content Publication	Construction and Building Materials
Licensed Content Title	Effect of acid attack on FRP-to-concrete bonded interfaces
Licensed Content Author	S.A. Hadigheh,R.J. Gravina,S.T. Smith
Licensed Content Date	Oct 15, 2017
Licensed Content Volume	152
Licensed Content Issue	n/a
Licensed Content Pages	19
Type of Use	reuse in a thesis/dissertation
Portion	figures/tables/illustrations
Number of figures/tables/illustrations	1
Format	electronic
Are you the author of this Elsevier article?	No
Will you be translating?	No
Title	Multiscale Machine Learning and Numerical Investigation of Ageing in Infrastructures
Institution name	The University of Sydney
Expected presentation date	Sep 2023
Order reference number	Figure 2-8
Portions	Fig. 12. Penetration of corrosive medium within the interface.
Requestor Location	Mr. Keyvan Aghabalaei Baghaei The University of Sydney Sydney, NSW 2006 Australia Attn: Mr. Keyvan Aghabalaei Baghaei
Publisher Tax ID	GB 494 6272 12
Total	**0.00 USD**

Figure A-1. Elsevier copyright agreement for using figures of Hadigheh et al. [9] in the thesis

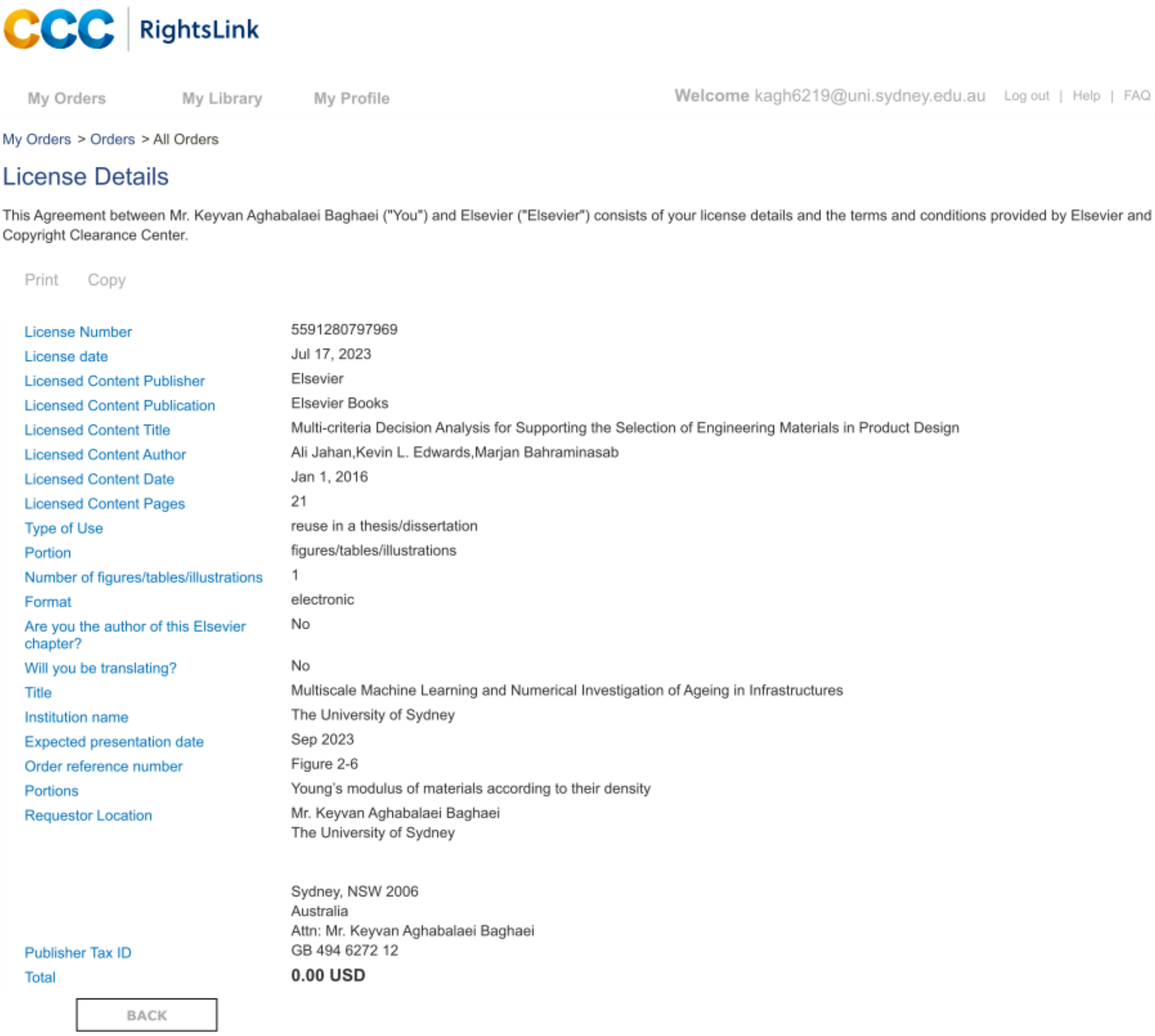


License Details

This Agreement between Mr. Keyvan Aghabalaei Baghaei ("You") and Elsevier ("Elsevier") consists of your license details and the terms and conditions provided by Elsevier and Copyright Clearance Center.

License Number	5591201261076
License date	Jul 17, 2023
Licensed Content Publisher	Elsevier
Licensed Content Publication	Composites Part B: Engineering
Licensed Content Title	A comprehensive review on mechanical properties of pultruded FRP composites subjected to long-term environmental effects
Licensed Content Author	TianQiao Liu,Xing Liu,Peng Feng
Licensed Content Date	Jun 15, 2020
Licensed Content Volume	191
Licensed Content Issue	n/a
Licensed Content Pages	1
Type of Use	reuse in a thesis/dissertation
Portion	figures/tables/illustrations
Number of figures/tables/illustrations	1
Format	electronic
Are you the author of this Elsevier article?	No
Will you be translating?	No
Title	Multiscale Machine Learning and Numerical Investigation of Ageing in Infrastructures
Institution name	The University of Sydney
Expected presentation date	Sep 2023
Order reference number	123456
Portions	Fig. 2. Degradation mechanisms from water or high humidity.
Requestor Location	Mr. Keyvan Aghabalaei Baghaei The University of Sydney Sydney, NSW 2006 Australia Attn: Mr. Keyvan Aghabalaei Baghaei
Publisher Tax ID	GB 494 6272 12
Total	**0.00 USD**

Figure A-2. Elsevier copyright agreement for using figures of Liu et al. [10] in thesis

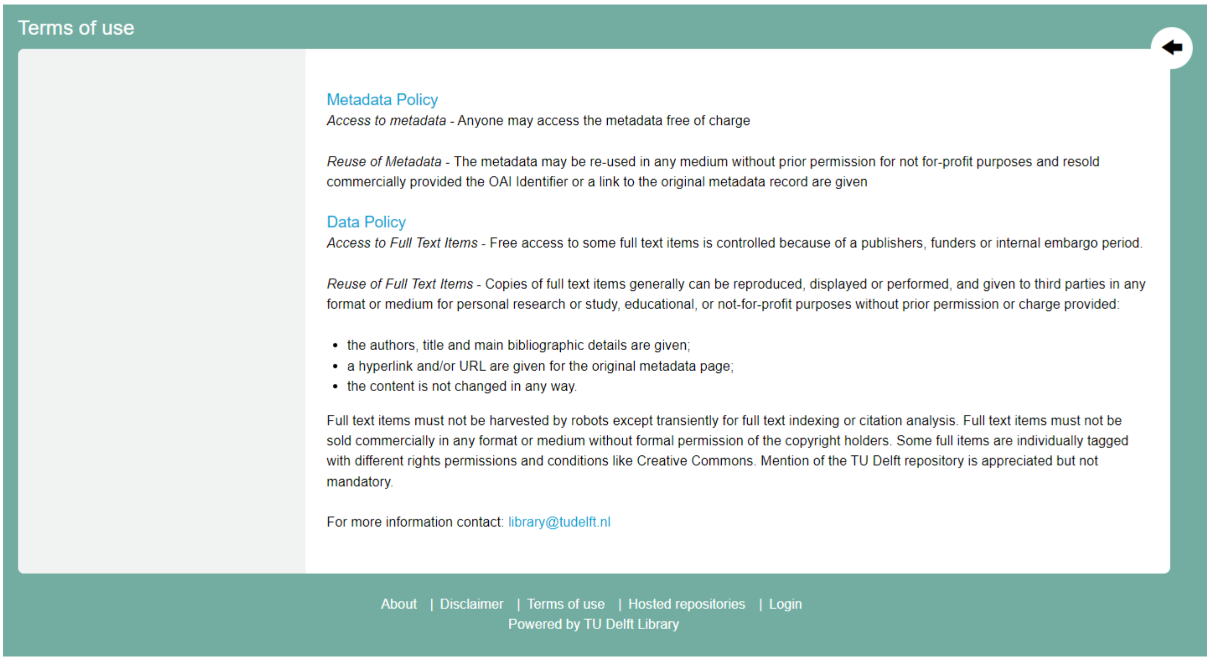


License Details

This Agreement between Mr. Keyvan Aghabalaei Baghaei ("You") and Elsevier ("Elsevier") consists of your license details and the terms and conditions provided by Elsevier and Copyright Clearance Center.

License Number	5591280797969
License date	Jul 17, 2023
Licensed Content Publisher	Elsevier
Licensed Content Publication	Elsevier Books
Licensed Content Title	Multi-criteria Decision Analysis for Supporting the Selection of Engineering Materials in Product Design
Licensed Content Author	Ali Jahan,Kevin L. Edwards,Marjan Bahraminasab
Licensed Content Date	Jan 1, 2016
Licensed Content Pages	21
Type of Use	reuse in a thesis/dissertation
Portion	figures/tables/illustrations
Number of figures/tables/illustrations	1
Format	electronic
Are you the author of this Elsevier chapter?	No
Will you be translating?	No
Title	Multiscale Machine Learning and Numerical Investigation of Ageing in Infrastructures
Institution name	The University of Sydney
Expected presentation date	Sep 2023
Order reference number	Figure 2-6
Portions	Young's modulus of materials according to their density
Requestor Location	Mr. Keyvan Aghabalaei Baghaei The University of Sydney Sydney, NSW 2006 Australia Attn: Mr. Keyvan Aghabalaei Baghaei
Publisher Tax ID	GB 494 6272 12
Total	**0.00 USD**

Figure A-3. Elsevier copyright agreement for using figures of Jahan et al. [23] in thesis



Metadata Policy

Access to metadata - Anyone may access the metadata free of charge

Reuse of Metadata - The metadata may be re-used in any medium without prior permission for not for-profit purposes and resold commercially provided the OAI Identifier or a link to the original metadata record are given

Data Policy

Access to Full Text Items - Free access to some full text items is controlled because of a publishers, funders or internal embargo period.

Reuse of Full Text Items - Copies of full text items generally can be reproduced, displayed or performed, and given to third parties in any format or medium for personal research or study, educational, or not-for-profit purposes without prior permission or charge provided:

- the authors, title and main bibliographic details are given;
- a hyperlink and/or URL are given for the original metadata page;
- the content is not changed in any way.

Full text items must not be harvested by robots except transiently for full text indexing or citation analysis. Full text items must not be sold commercially in any format or medium without formal permission of the copyright holders. Some full items are individually tagged with different rights permissions and conditions like Creative Commons. Mention of the TU Delft repository is appreciated but not mandatory.

For more information contact: library@tudelft.nl

Figure A-4. Delft University of Technology term of use policy for reusing metadata

Appendix B

Collected data for bond durability

Table B-1: References for data collection of adhesively bonded connections

References	Geometrical and mechanical properties											Exposure condition					Test results	
	Concrete		FRP						Adhesive									
	Width	Compressive strength	Bond length	Thickness	Width	Elastic modulus	Ultimate strain	Tensile strength	Thickness	Elastic modulus	Tensile strength	Exposure type	Temperature	Relative humidity	Duration	Sustained loading	Bond strength	Failure mode
Barger [175]	✓	✓	✓	✓	✓	✓	✓	×	×	×	×	✓	✓	✓	✓	✓	×	✓
Pack [176]	✓	✓	✓	✓	✓	✓	✓	✓	×	×	×	✓	✓	✓	✓	✓	✓	✓
Au and Büyüköztürk [114]	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Tam [177]	✓	✓	✓	✓	✓	✓	×	✓	×	×	×	✓	✓	✓	✓	✓	✓	✓
Subramaniam et al. [178]	✓	✓	✓	✓	✓	✓	✓	×	×	×	×	✓	✓	✓	✓	✓	✓	✓
Gamage et al. [179]	✓	✓	✓	✓	✓	×	×	×	✓	×	×	✓	✓	✓	✓	✓	✓	✓
Lai et al. [33]	✓	✓	✓	✓	✓	×	×	✓	✓	×	×	✓	✓	✓	✓	✓	✓	✓
Abbas [180]	✓	✓	✓	✓	✓	✓	×	×	✓	×	×	✓	✓	✓	✓	✓	✓	✓
Colombi et al. [37]	✓	✓	✓	✓	✓	✓	×	×	×	×	×	✓	✓	✓	✓	✓	✓	✓
Benzarti et al. [181]	✓	✓	✓	✓	✓	✓	✓	✓	×	×	×	✓	✓	✓	✓	✓	✓	✓
Tuakta and Büyüköztürk [182]	✓	✓	✓	✓	✓	✓	×	×	×	✓	×	✓	✓	✓	✓	✓	✓	✓
Shi et al. [183]	✓	✓	✓	✓	✓	✓	✓	✓	×	×	×	✓	✓	✓	✓	✓	✓	✓
Kabir [184]	✓	✓	✓	✓	✓	✓	×	✓	×	×	×	✓	✓	✓	✓	✓	✓	✓
Pan et al. [185]	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Shrestha et al. [115]	✓	✓	✓	✓	✓	✓	×	✓	×	✓	✓	✓	✓	✓	✓	✓	×	✓
Shrestha et al. [186]	✓	✓	✓	✓	✓	✓	×	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Zheng et al. [187]	✓	✓	✓	✓	✓	✓	×	×	×	×	×	✓	✓	✓	✓	✓	✓	✓
Gravina et al. [188]	✓	✓	✓	✓	✓	✓	✓	✓	×	×	×	✓	✓	✓	✓	✓	✓	✓
Liu et al. [189]	✓	✓	✓	✓	✓	✓	×	×	✓	×	×	✓	✓	✓	✓	✓	✓	✓
Pan et al. [113]	✓	✓	✓	✓	✓	✓	✓	×	✓	✓	×	✓	✓	✓	✓	✓	✓	✓
Al-Lami et al. [190]	✓	✓	✓	✓	✓	✓	✓	✓	×	✓	✓	✓	✓	✓	✓	✓	✓	✓

✓: Data available; ×: Data was not reported

Table B-2. Availability of variables in the collated references for FRP bar-to-concrete connection

	Material properties								Conditioning regime					
	Concrete		FRP					Resin						
Reference	$cover/d_b$	f'_c	Surface texture	Bar material	d_b	L_b/d_b	E_f	Resin material	Test type	Exposure type	Temperature	Duration	Conditioning type	pH / Salt concentration
Abbasi and Hogg [127]	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Abedi [191]	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Al-Tamimia et al. [192]	✓	✓	✓	✓	✓	✓	✓	✗	✓	✓	✓	✓	✓	✓
Al-Dulaijan et al. [193]	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✗	✓	✓
Altalmas et al. [124]	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Alvarez [194]	✓	✓	✓	✓	✓	✓	✓	✗	✓	✓	✓	✓	✓	✗
Al-Zahrani et al. [195]	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Ammar [196]	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✗
Calvet et al. [197]	✓	✗	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✗
Davalos et al. [198]	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✗
Dong et al. [199]	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Dong et al. [200]	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Dong et al. [201]	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Refai et al. [202]	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Galati et al. [203]	✓	✓	✓	✓	✓	✓	✓	✗	✓	✓	✓	✓	✓	✗
Hassan et al. [204]	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓

Robert and Benmokrane [205]	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	×
Yan et al. [206]	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Yan and Lin [207]	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Zhou et al. [208]	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Zhou et al. [209]	✓	×	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Mashima and Iwamoto [210]	×	×	✓	✓	×	×	×	×	✓	✓	✓	×	✓	×

✓: Data available; ×: Data was not reported

Appendix C

Details of the developed prediction models

Table C-1. Coefficients of the linear models developed by M5P

Linear Model	Intercept	Variable										
		b_c[a]	L_f[b]	t_f[c]	b_f[d]	C[e]	T[f]	RH[g]	D[h]	Pr[i]	E_f[j]	f'_c[k]
		Coefficient										
	a	b	c	d	e	f	g	h	i	j	k	l
LM1	9.6574	-0.2890	0.8135	2.2035	1.5568	0.0561	-0.431	0.2342	-0.2432	0	1.8075	0.1107
LM2	10.5329	-0.2890	0.2175	0.7224	1.5568	0.0561	-0.4113	0.1249	-0.4893	0	0.9640	-0.0842
LM3	13.0994	-0.2890	0.2175	2.8427	1.5568	0.0561	-0.4113	-0.7992	-0.3933	0	0.9640	-0.0842
LM4	15.8214	-0.3400	1.8165	7.9394	7.5542	0.0561	-0.1074	-0.0680	-0.0842	0	0	-0.0842
LM5	12.0914	0.8439	0.5488	0.8990	0.5986	0.2078	-0.1746	-0.1168	-0.1677	0.2489	-0.3688	-0.2423
LM6	22.4364	1.1677	1.7102	-4.2636	0.5986	0.1214	-0.4536	-0.1168	-0.3246	0.1153	-0.6893	0.2379
LM7	17.5093	3.3774	-0.0279	-1.3587	0.5986	0.1214	-0.4976	-0.1168	-0.4077	0.1153	-0.4110	1.4385
LM8	15.8919	0.6412	1.2294	-1.3587	0.5986	-0.1544	-0.7062	-0.1168	-0.3110	0.1153	-1.2269	0.2007
LM9	18.8392	-0.1271	0.3440	0.3160	0.6755	-0.1444	-0.6020	-0.9408	0.2997	0	-0.1146	-0.2972
LM10	19.9311	-0.1271	0.3440	0.3594	0.6755	-0.1444	-0.5359	-0.8789	0.7175	0	1.4197	-0.2972
LM11	18.1998	-0.1271	0.3440	0.3594	0.6755	-0.1444	-0.5359	-1.0151	0.9913	0	-0.1146	-0.2972
LM12	19.8379	-1.0211	0.3440	0.4825	0.6755	-0.2357	-0.4188	-0.4506	-2.8776	0	3.0173	-0.2972

[a] Concrete width; [b] FRP bond length; [c] FRP thickness; [d] FRP width; [e] Exposure type; [f] Temperature; [g] Relative humidity; [h] Duration; [i] Sustained load.
[j] FRP elastic modulus; [k] Concrete compressive strength

Table C-2. Coefficients of the bond failure equation

i	Input variable		Coefficients			
			a_1	a_2	a_3	γ
1		d_b	0.03	5.30	-13.48	0.37
2		$cover/d_b$	0.27	0.21	-2.99	1.39
3		L_b/d_b	-0.17	-111.66	-1.24	-1.71
4		f'_c	0.24	0.29	-0.32	0.82
5		E_f	0.24	-1569.64	-5.02	0.84
6	Temperature	$+$	0.29	0.03	-1.66	1.12
		$-$	0.29	-0.03	-1.60	0.88
7		Duration	0.19	0.10	-1.94	0.43
8		pH	0.25	1.75	-6.01	0.58
9		Salt concentration	0.06	-3.62	2.71	-0.58
10	Bar surface texture	Sand coated	0.32	4.89	-5.94	0.50
		Ribbed	-0.35	-4.28	3.38	-0.89
		Smooth	0.10	-6.87	6.72	-65.45
11	Bar material	Carbon	-0.13	-3.95	3.68	-20.57
		Glass	-0.04	3.90	-4.96	0.61
		Basalt	0.02	-2.56	1.81	-2.49
		Steel	0.21	-4.44	4.24	-44.99
12	Resin material	Polyester	0.07	-4.38	4.14	-5.54
		Vinyl ester	0.14	0.14	-1.70	3.99
		Epoxy	-0.41	-2.42	1.86	-25.71
13	Test type	Concentric	0.31	0.00	-6.68	65.45
		Eccentric	0.00	0.00	-6.68	65.45
14	Exposure type	Cyclic	-0.11	-3.09	2.73	-12.61
		Continuous	0.00	-3.09	2.73	-12.61
15	Conditioning type	AC	-0.38	-4.18	3.92	-23.22
		AL	0.27	-2.89	2.48	-10.07
		FT	-0.21	-3.49	3.17	-15.65
		W	0.08	-3.42	3.10	-14.99
		S	0.06	-3.71	2.83	-1.14
		UC	0.00	-2.59	2.13	-8.21

Table C-3. Coefficnets of the bond strength equaiton

i	Input variable		Coefficients			
			b_1	b_2	b_3	γ
1		d_b	15.50	5.50	-13.80	0.37
2		$cover/d_b$	0.92	0.20	-2.97	1.39
3		L_b/d_b	1.48	-140.44	6.53	-1.71
4		f'_c	4.99	0.27	-0.29	0.82
5		E_f	7.85	-1629.63	2.68	0.84
6	Temperature	$+$	-1.14	0.03	-1.66	1.12
		$-$	-1.14	-0.03	-1.60	0.88
7		Duration	-0.62	0.10	-1.94	0.43
8		pH	-1.55	1.74	-5.98	0.58
9		Salt concentration (%)	0.63	-3.61	2.73	-0.58
10	Bar surface texture	Sand coated	0.75	4.79	-5.85	0.50
		Ribbed	-0.11	-4.37	3.46	-0.89
		Smooth	-2.17	-6.81	6.66	-65.45
11	Bar material	Carbon	-2.10	-3.92	3.64	-20.57
		Glass	1.41	3.82	-4.89	0.61
		Basalt	0.36	-2.53	1.78	-2.49
		Steel	-1.80	-5.68	5.49	-44.99
12	Resin material	Polyester	-0.62	-4.35	4.10	-5.54
		Vinyl ester	-1.54	0.15	-1.70	3.99
		Epoxy	1.98	-2.40	1.84	-25.71
13	Test type	Concentric	-0.08	0.00	-6.66	65.45
		Eccentric	0.00	0.00	-6.66	65.45
14	Exposure type	Cyclic	0.81	-3.14	2.79	-12.61
		Continuous	0.00	-3.14	2.79	-12.61
15	Conditioning type	AC	-1.13	-4.14	3.89	-23.22
		AL	2.27	-2.87	3.14	-10.07
		FT	-1.17	-3.46	3.15	-15.65
		W	-0.60	-3.39	3.07	-14.99
		S	0.28	-3.69	2.82	-1.14
		UC	-0.38	-2.66	2.20	-8.21