UNIVERSITAT ROVIRA I VIRGILI

ENHANCING PROPERTY PREDICTION IN BUILDING MATERIALS THROUGH DATA-AUGMENTED NEURAL NETWORKS

Ana Carolina Souza Rosa

Enhancing property prediction in building materials through data-augmented neural networks

ANA CAROLINA SOUZA ROSA

DOCTORAL THESIS

2024

Ana Carolina Souza Rosa

Enhancing property prediction in building materials through data-augmented neural networks

Doctoral Thesis

Supervised by: Dr. Dieter Boer

Co-supervised by: Dr. Assed Haddad

DEPARTMENT OF MECHANICAL ENGINEERING

UNIVERSITY ROVIRA I VIRGILI

TARRAGONA

2024

UNIVERSITAT ROVIRA I VIRGILI

ESCOLA TÈCNICA SUPERIOR D'ENGINYERIA QUÍMICA
DEPARTAMENT D'ENGINYERIA QUÍMICA
Av. Països Catalans, 26
Campus Sescelades
43007 Tarragona (Spain)

We state that the present study, entitled "Enhancing property prediction in building materials through data-augmented neural networks", presented by Ana Carolina Souza Rosa for the award of the degree of Doctor, has been carried out under our supervision at the Department of Mechanical Engineering of this university.

Tarragona, 28th June 2024

Doctoral Thesis Supervisor.

DIETER THOMAS BOER - DNI X1992293X

Digitally signed by DIETER THOMAS BOER - DNI X1992293X
Date: 2024.06.28 11:44:24 +02'00'

Dr. Dieter Boer

Dr. Assed Haddad

Acknowledgements

....

Agradecimientos

…

Summary

Energy-efficient construction practices have become crucial in modern construction and urban planning. These practices aim to reduce energy consumption, improve thermal comfort, promote sustainability, and lower long-term costs. The rising demand for energy-efficient buildings has led to advancements in construction materials, techniques, and smart technologies. Consequently, the construction industry is evolving to create sustainable and cost-effective structures that benefit both the environment and the economy.

Researchers are focusing on a wide range of building materials for the building envelope, which is vital for maintaining stable internal temperatures despite external fluctuations. Careful selection of materials during the design phase can significantly reduce energy consumption, leading to more sustainable and efficient projects. The durability and cost of these materials are also critical factors, as durable materials reduce maintenance needs and costs, while cost-effective materials make sustainable construction more accessible.

Concrete remains one of the most versatile and widely used materials in construction, valued for its mechanical strength and adaptability. It can be tailored with different compositions, aggregates, and additives to meet specific project requirements, such as enhanced strength and improved thermal insulation. This adaptability supports energy efficiency and long-term resilience in construction projects.

The evolution of concrete mix design now includes considerations for sustainability and thermal properties. This process involves selecting ingredients to achieve desired characteristics for specific applications. Traditional methods of concrete mix design are often labor-intensive and time-consuming. To overcome these limitations, researchers are developing computational tools, including machine learning and deep learning techniques, to optimize concrete mix design efficiently.

Machine learning, a branch of artificial intelligence, creates models that learn from data to make predictions. In construction, machine learning models predict material properties, reducing the time and experiments needed to determine ideal compositions. Significant advances have been made using machine learning and deep learning to predict concrete properties. However, developing reliable models requires robust databases, which is a major challenge.

This study aims to develop a deep learning model to predict the thermal properties of concrete based on its composition and density. Following this, a secondary model

will generate synthetic data to improve the training of predictive models. The study's second stage will apply this methodology to other construction materials, demonstrating the versatility and potential of deep learning in advancing sustainable and efficient construction practices.

Resumen

Las prácticas de construcción energéticamente eficientes se han vuelto cruciales en la construcción moderna y la planificación urbana. Estas prácticas tienen como objetivo reducir el consumo de energía, mejorar el confort térmico, promover la sostenibilidad y disminuir los costos a largo plazo. La creciente demanda de edificios energéticamente eficientes ha llevado a avances en materiales de construcción, técnicas innovadoras y tecnologías inteligentes. En consecuencia, la industria de la construcción está evolucionando para crear estructuras sostenibles y rentables que beneficien tanto al medio ambiente como a la economía.

Los investigadores se están centrando en una amplia gama de materiales para la envolvente del edificio, que es vital para mantener temperaturas internas estables a pesar de las fluctuaciones externas. La cuidadosa selección de materiales durante la fase de diseño puede reducir significativamente el consumo de energía, lo que lleva a proyectos más sostenibles y eficientes. La durabilidad y el costo de estos materiales también son factores críticos, ya que los materiales duraderos reducen las necesidades y los costos de mantenimiento, mientras que los materiales rentables hacen que la construcción sostenible sea más accesible.

El hormigón sigue siendo uno de los materiales más versátiles y ampliamente utilizados en la construcción, valorado por su resistencia mecánica y adaptabilidad. Puede ajustarse con diferentes composiciones, agregados y aditivos para satisfacer requisitos específicos del proyecto, como mayor resistencia y mejor aislamiento térmico. Esta adaptabilidad apoya la eficiencia energética y la resiliencia a largo plazo en los proyectos de construcción.

La evolución del diseño de mezclas de concreto ahora incluye consideraciones de sostenibilidad y propiedades térmicas. Este proceso implica seleccionar ingredientes para lograr características deseadas para aplicaciones específicas. Los métodos tradicionales de diseño de mezclas de concreto a menudo son laboriosos y consumen mucho tiempo. Para superar estas limitaciones, los investigadores están desarrollando herramientas computacionales, incluyendo técnicas de aprendizaje automático y aprendizaje profundo, para optimizar el diseño de mezclas de concreto de manera eficiente.

El aprendizaje automático, una rama de la inteligencia artificial, crea modelos que aprenden de los datos para hacer predicciones. En la construcción, los modelos de aprendizaje automático predicen las propiedades de los materiales, reduciendo el tiempo y los experimentos necesarios para determinar las composiciones ideales. Se han logrado avances significativos utilizando el aprendizaje automático y el

aprendizaje profundo para predecir las propiedades del concreto. Sin embargo, desarrollar modelos confiables requiere bases de datos robustas, lo cual es un gran desafío.

Este estudio tiene como objetivo desarrollar un modelo de aprendizaje profundo para predecir las propiedades térmicas del concreto basado en su composición y densidad. Posteriormente, un modelo secundario generará datos sintéticos para mejorar el entrenamiento de los modelos predictivos. La segunda etapa del estudio aplicará esta metodología a otros materiales de construcción, demostrando la versatilidad y el potencial del aprendizaje profundo para avanzar en prácticas de construcción sostenibles y eficientes.

Table of contents

CHAPTER I.

Introduction

I. Introduction

1.1 Background and motivation

Energy-efficient construction practices are becoming essential components of modern construction and urban planning [REF]. These practices are based on several key principles, including reducing energy consumption, improving thermal comfort, promoting sustainability, and reducing long-term costs [REF]. The increasing demand for energy-efficient buildings has driven significant advances in new construction materials, innovative construction techniques, and smart technologies designed to meet these needs. As a result, the construction industry is evolving to incorporate these advancements, aiming to create more sustainable and cost-effective structures that benefit both the environment and the economy.

To address the growing demand for more sustainable buildings, researchers have been diligently investigating a wide range of building materials for the building envelope [REF]. The building envelope serves as the primary barrier to maintaining a stable internal temperature, regardless of fluctuations in external temperatures. This crucial function highlights the importance of carefully selecting materials during the initial design phase of a building [REF]. By choosing the adequate materials, energy consumption can be significantly reduced, leading to the development of projects that are not only more sustainable but also more efficient in the long term.

Additionally, the durability and cost of these materials play a significant role in their selection. Durable materials ensure the longevity of the building, reducing the need for frequent repairs and replacements, which in turn lowers maintenance costs and resource consumption [REF]. Cost-effective materials that do not compromise on quality or performance make sustainable construction more accessible and financially viable. This meticulous approach to material selection is essential in advancing the construction industry towards greener, more durable, and economically feasible practices.

Among building materials, concrete is widely recognized as one of the most versatile and extensively used in a variety of projects, including buildings, infrastructure, and other architectural constructions. Its exceptional mechanical strength allows it to withstand heavy loads and resist adverse environmental and usage conditions, making it a reliable choice for both structural and non-structural applications. Additionally, its malleability allows it to adapt to different shapes and sizes, which makes it a flexible and indispensable material for a wide range of engineering projects.

Concrete can be formulated with various compositions, aggregates, and additives to meet specific requirements, such as enhanced strength, improved thermal insulation, and other desired properties tailored to the needs of the project [REF]. This adaptability extends to incorporating different constituents that enhance its performance, sustainability, and durability. By adjusting the mix design, engineers can create concrete that not only meets the structural demands but also contributes to energy efficiency and long-term resilience [REF].

Since its invention, the composition of concrete mixtures has continuously evolved to meet the ever-changing demands of the construction sector. Traditionally, designers focused on achieving specific targets for properties like compressive strength and workability. However, with the growing emphasis on sustainability, thermal properties have also become essential considerations during the initial design phase [REF]. This shift necessitates the use of different constituents and proportions in the concrete mixture, tailored to the specific requirements of each application [REF].

This process, known as concrete mix design, involves the meticulous selection of the type and amount of each ingredient to produce concrete with desired characteristics for a given application [REF]. While traditional methods of concrete mix design are widely used and hold significant value, they are often labor-intensive and time-consuming. To address these limitations, researchers have been developing and applying computational tools to design concrete mixes with optimized properties [REF]. Among these innovative approaches, the use of machine learning and deep learning techniques has shown great promise in enhancing efficiency and precision in mix design.

Machine learning is a branch of artificial intelligence responsible to create models that learn from and make decisions based on a database. The algorithms read the database and employ some statistical techniques to recognize patterns and make predictions or decisions. The application of machine learning algorithms has expanded across various areas, including civil construction [REF]. In the realm of building materials, numerous studies have employed models with diverse techniques and architectures to predict material properties [REF]. This approach significantly reduces the time and number of experiments needed to determine the ideal composition that meets project requirements.

The literature highlights significant advances in using machine learning and deep learning models, which is a subset of machine learning, to predict concrete properties [REF]. However, developing reliable models requires access to comprehensive and robust databases. This limitation presents a major challenge, restricting the exploration and validation of new models across different datasets. To address this issue, the present work aims to develop a model capable of generating synthetic data to supplement the training of predictive models for concrete properties.

The initial stage of this study involves creating a deep learning model designed to predict the thermal properties of concrete based on its constituent composition and material density. Once this predictive model is established, the next step involves developing a secondary model to generate synthetic data that mirrors the characteristics of real data. By producing high-quality synthetic data, we aim to enhance the training process, thereby improving the accuracy and robustness of the predictive models.

Moreover, the second stage of this study seeks to extend this methodology to other construction materials, demonstrating the versatility and adaptability of the developed model. By applying these techniques across various materials, we hope to reveal the broader applicability and potential of deep learning in advancing sustainable and efficient construction practices.

1.2 General objectives

To develop and apply advanced deep learning models to predict and optimize the properties of construction materials, thereby enhancing the efficiency, sustainability, and performance of construction projects.

To achieve the overall goal, the following specific objectives must be met:

- Develop a robust deep learning model that can accurately predict concrete properties based on its composition and density.

- Create a data augmentation model to generate synthetic data that mirrors real data characteristics, thereby enhancing the training of predictive models.
- Extend the methodology to encompass other construction materials. By applying our predictive and data augmentation models to a broader range of materials, highlighting the versatility and effectiveness of these approach

1.3 Machine learning modeling approaches

Machine learning is an area that has been explored for years and has a wide range of applications for solving different problems. This thesis will emphasize the use of the model used to predict the properties of certain building materials.

The sequence of a machine learning model begins with the input of numerical or textual data to train the algorithm. During the training stage, the initial algorithm analyzes the imputed values and then adjusts its parameters to identify and learn certain patterns and relationships within the data. After training, new data is fed into the algorithm to evaluate its performance and ensure the effectiveness of the model developed.

Machine learning' models can be broken into three main categories: supervised learning, unsupervised learning, e reinforced learning. In supervised learning, the training step uses labeled data and the algorithm learns to map inputs to the corresponding outputs. On the opposite, unsupervised learning models are trained on unlabeled data and the algorithm must retrieve some patterns within the data on its own. In reinforced learning, the model learning process is based on rewards and penalties according to its actions, which aims to maximize the total reward over time.

Deep learning is a subset of machine learning and Artificial Neural Network (ANN) is its cornerstone. The ANN algorithm is inspired by the structure and the function of the human brain. ANNs are composed of a set of networks of interconnected nodes, which work together to learn complex relationships between a set of inputs and outputs. This makes it well-suited for predicting the output based on a variety of parameters. In this way, ANNs can be used to solve problems that conventional or other computational methods have difficulties [REF19]. ANNs provide an alternative method for predicting building material properties that is faster, cheaper, and more accurate than traditional methods. In this thesis, the prediction model algorithm is built using the multilayer perceptron network.

Another type of deep learning model that has be gaining prominence are Generative Adversarial Networks (GANs). The first GAN model was developed by Ian et al. in 2014 [20]. In this work, a framework was proposed for estimating generative models via the adversarial process, in which two models were trained simultaneously. Over the years, numerous algorithms have been developed and improved, often leveraging architectures like convolutional neural networks. In this thesis we will use CopulaGAN synthesizer.

1.3.1 Multilayer Perceptron Network

The Multilayer Perceptron (MLP) is a versatile artificial neural network that can be employed to solve many problems, including classification, regression, and pattern recognition. A typical MLP model comprises three blocks of interconnected neurons: the input layer in which each neuron represents a data feature; the hidden layers that could have one or more layers depending on the complexity of the problem; and the output layer representing the response output (Figure I-1). Each neuron in the network is bonded to other neurons through connection weights. Each neuron in a neural network receives an input (X_i) that is multiplied by a weight (W_i) and is summed with each other and added to the bias value (b). Then, the result is transferred to the activation function, which adjusts the final output (Figure I-2). Each layer of neurons gathers input from the previous layers, and the outputs of neurons within one layer become the inputs to neurons in the following layer. Finally, the last layer produces the predictions of the model.

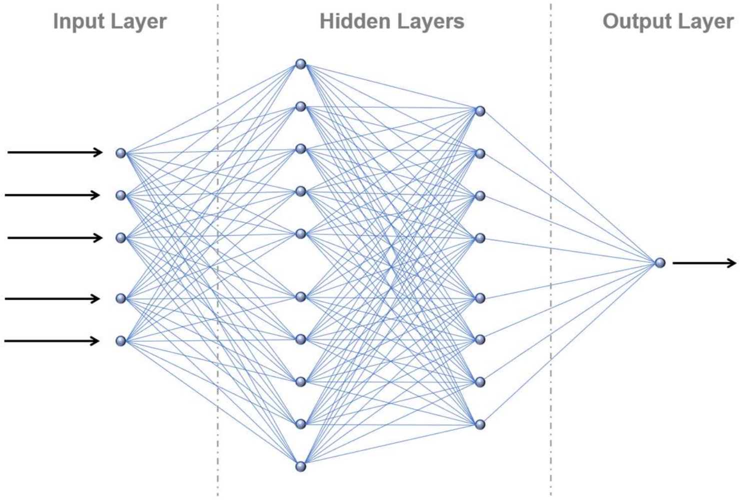

Figure I- 1. Example of a MLP network

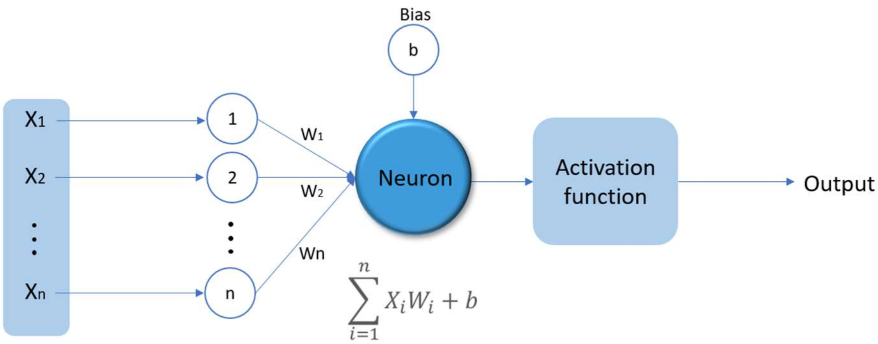

Figure I- 2. Graphical representation of a neuron

1.3.2 Generative Adversarial Network

Generative Adversarial Networks, or GANs, are a powerful type of neural networks that consists of two main components: a generator and a discriminator. The generator's objective is to learn the patterns within the input data so it can create new examples that closely resemble the original dataset. On the other hand, the discriminator's role is to spot the subtle differences between the real data and the synthetic data produced by the generator.

This whole process, known as adversarial training, involves these two networks working against each other that is represented in Figure I-3. The generator tries to fool the discriminator by making its synthetic data look as real as possible, while the discriminator aims to get better at distinguishing between the real and synthetic data. Essentially, the generator captures the distribution of the dataset, and the discriminator assesses the likelihood that a given data sample came from the real training data rather than being generated artificially.

GANs were introduced by Goodfellow et al. [20] and became a revolutionary development in the world of generative modeling. They have different applications, such as image generation, super-resolution imaging, style transfer, and data augmentation [REF. Deep Learning with Python, 2017]. Regarding tabular data, it can create synthetic data to expand real datasets and prevent over-fitting in such data-limited situations, helping to improve the training of an ML model. Although GAN started with image generation, some authors have already used this technique to create tabular data and obtain satisfactory results. Since the introduction of GAN, several algorithms to model tabular data have been used, such as Conditional Tabular GAN (CTGAN) [21], TabGAN, and CopulaGAN [22].

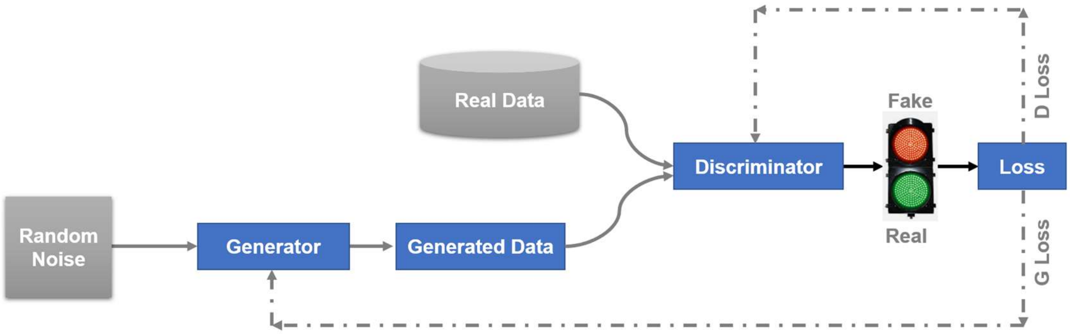

Figure I- 3. Schematic example of a Generative Adversarial Network

1.4 Outline: problems addressed

[Brief description]

1.4.1 Use of operational research techniques for concrete mix design: A systematic review (Article 1)

Concrete is one of the most widely used construction materials, thanks to its workability, durability, and the availability of raw materials. It's typically a mix of cement, water, and aggregates, but can include a variety of other materials to achieve different properties. This flexibility allows designers to tailor concrete mixtures to meet specific performance requirements like workability, durability, and strength.

Designing good concrete mixtures is crucial for meeting these requirements, and various methods exist to help find the best mix. Traditional methods, such as those by the American Concrete Institute (ACI) and British standards, often rely on empirical data and step-by-step procedures to determine proportions. However, these methods sometimes fall short, especially in accounting for the impact of supplementary materials like pozzolans and fly ash, and they may not always find the optimal mix.

In recent years, computational methods, including mathematical programming and machine learning, have been explored to improve concrete mix design. These methods can model the complex relationships between mix components and concrete properties more accurately, leading to better, more efficient solutions.

Many studies have applied machine learning to predict concrete properties or optimize mix compositions. Some researchers focus on making the process more accessible and cost-effective, while others aim for more sustainable and economical mixes. Reviews of the literature highlight various machine learning models and optimization algorithms used for these purposes.

This article aims to review the use of machine learning and mathematical programming in concrete mix design, providing a holistic understanding of these tools. It also proposes a framework to classify recent works in this area. The paper addresses three key questions: what machine learning algorithms are used in concrete mix design, what mathematical optimization algorithms are employed, and how these approaches are combined in the literature to design concrete mixes. Figure I-4 displays the graphical abstract of article 1, which represents the classification framework proposed and the techniques applied to predict the properties of concrete.

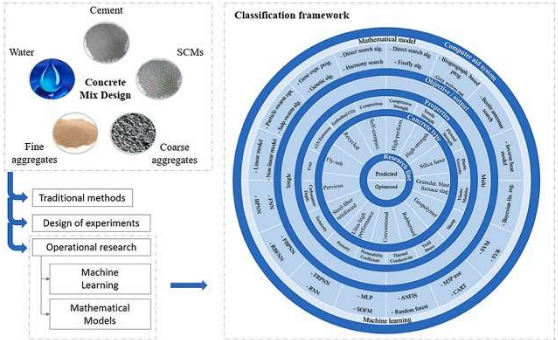

Figure I- 4. Graphical abstract of article 1: Indicating the different techniques applied to predict the properties of concrete and optimize the concrete mixture design

1.4.2 Methodology approach for prediction of the thermal conductivity of concrete by using neural networks (Article 2)

With rising awareness of sustainability, energy efficiency in buildings is becoming increasingly important. Researchers are focusing on developing materials that enhance thermal performance, leading to more eco-friendly and cost-effective construction practices. Materials that improve thermal efficiency can help maintain indoor comfort and reduce energy consumption.

In the building design phase, various methods and materials can be used to achieve energy efficiency. Active methods include technologies like heat pumps and solar panels, while passive methods involve using materials with low thermal conductivity to reduce heat transfer. Concrete, known for its durability and thermal mass, is commonly used for this purpose. Its ability to absorb, store, and gradually release heat makes it an effective material for stabilizing indoor temperatures.

Concrete's composition, including cement, water, fine and coarse aggregates, varies, affecting its thermal conductivity—a key property for energy efficiency. Accurate prediction of this property is crucial, as it can save time and resources in the design phase. Machine Learning (ML) models, especially Artificial Neural

Networks (ANNs), have shown promise in predicting various concrete properties but are less explored for thermal conductivity.

ANNs learn from data and can make accurate predictions, but they require large datasets for effective training. Generative Adversarial Networks (GANs) can address this by creating synthetic data to enhance limited datasets, improving ML model training. Despite GANs' success in other fields, few studies have applied them to predict concrete's thermal properties.

This work aims to bridge that gap by developing a methodology combining an ANN with a data augmentation model. Specifically, a Multilayer Perceptron (MLP) model will predict thermal conductivity, and a Copula GAN model will generate additional data. This approach will account for different types of concrete materials, providing a comprehensive solution for predicting thermal properties in diverse concrete compositions. Figure I-5 shows a graphical abstract of article 2 that highlights the motivation and the main outcome of the paper.

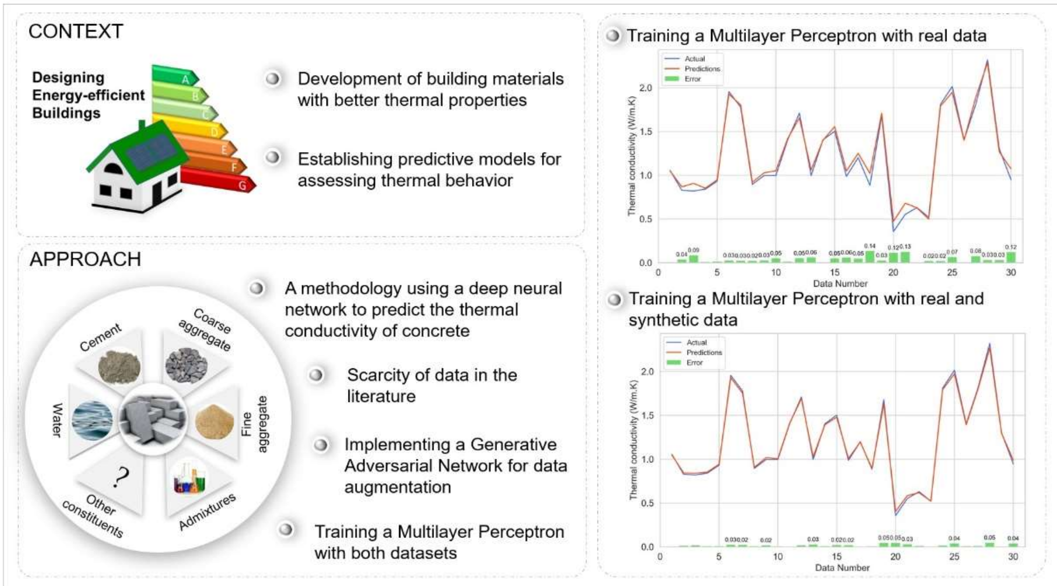

Figure I- 5. Graphical abstract of article 2: Combining a Multilayer Perceptron network and a Generative Adversarial Network to forecast the thermal conductivity of concrete

1.4.3 Data-Augmented Deep Learning Models for Assessing Thermal Performance in Sustainable Building Materials (Article 3).

Energy efficiency in buildings is driving the development of sustainable materials, with Phase Change Materials (PCMs) playing a significant role. PCMs, when added to concrete and cement, help regulate indoor temperatures by absorbing heat during the day and releasing it at night, thus enhancing energy efficiency. The accurate assessment of thermal properties in these materials is crucial but traditionally involves laborious and costly methods. Machine learning models offer an alternative but require extensive datasets, which are often unavailable.

Our study tackles the challenge of limited data by integrating deep neural networks, building on the methodology proposed in the second article. Given the small size of the original dataset, we use a Generative Adversarial Network (GAN) to augment it, followed by a Multilayer Perceptron (MLP) to predict the properties of cementitious composites enriched with PCM and nano-silica aerogel. This approach uses inputs such as mass composition and density to forecast thermal conductivity and compressive strength.

Our research aims to show that the proposed model, trained with augmented data, can accurately predict the thermal properties of concrete with different compositions and types of materials. Unlike the previous study, our dataset includes mixtures of cement, water, PCM, nano-silica aerogel, and superplasticizer. An improvement in our current model is its ability to predict two outputs simultaneously. By ensuring both thermal efficiency and structural resilience, this work significantly contributes to more sustainable construction practices. As illustrated in Figure I-6, the graphical abstract of Article 3 highlights the model's accuracy for both properties, demonstrating its applicability to specific case studies.

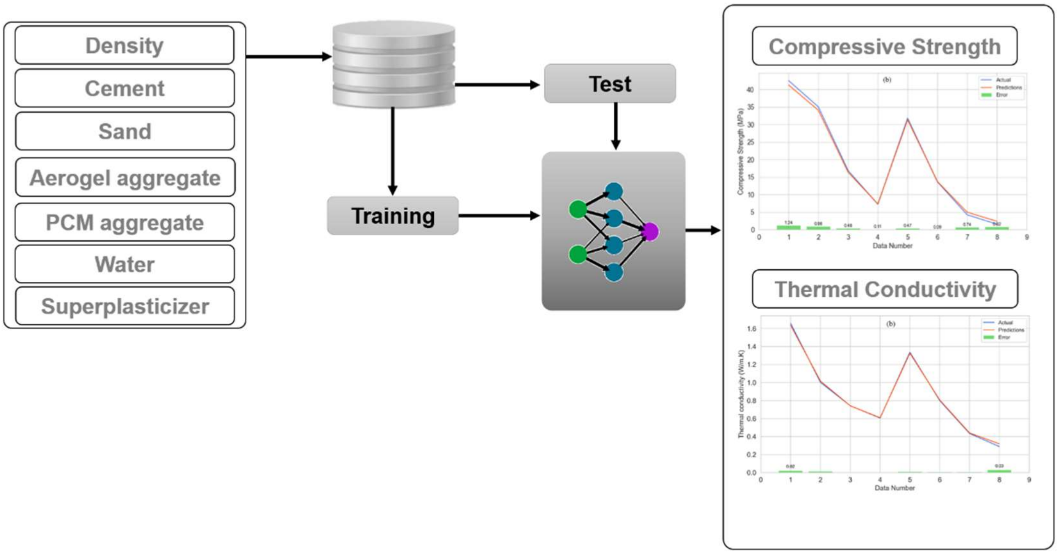

Figure I- 6. Schematic flowchart of article 3: expanding the methodology to a two-output case study

1.4.4 Enhanced Deep Learning Models with Data Augmentation for Ceramic Brick Property Assessment (Article 4).

Energy efficiency and sustainability are critical drivers in modern construction material research. While previous studies have demonstrated the potential of deep learning models in predicting the properties of cementitious composites, there is a pressing need to explore their applicability to other construction materials. This study aims to expand and validate the proposed deep learning model on ceramic bricks, a commonly used building material, to evaluate its performance across different material types.

Ceramic bricks are widely used due to their durability, thermal efficiency, and sustainability. Likely concrete, accurately predicting their properties is crucial for ensuring their optimal use in construction. To address the same challenges, we found on the second paper with concrete, as labor-expensive experiments and scarcity of dataset in the literature, our study employs an enhanced deep learning approach integrated with data augmentation techniques. We adopt expanded model from a previous work and applied to this case study. The integrated model with data augmentation and deep neural networks is designed to handle a variety of inputs, including alluvial soil, laterite soil, demolition and waste material, temperature, firing shrinkage, loss on ignition, and bulk density. The outputs of interest are water absorption and compressive strength.

Ceramic bricks are widely used due to their durability, thermal efficiency, and sustainability. Like concrete, accurately predicting their properties is crucial for ensuring optimal use in construction. To tackle the same challenges we encountered with concrete, such as labor-intensive experiments and the scarcity of datasets in the literature, our study employs an enhanced deep learning approach integrated with data augmentation techniques. We expand on the model from our previous work and apply it to this new case study. This integrated model, which combines data augmentation with deep neural networks, is designed to process a diverse range of inputs: alluvial soil, laterite soil, demolition and waste materials, temperature, firing shrinkage, loss on ignition, and bulk density. The key outputs it predicts are water absorption and compressive strength.

This research aims to demonstrate that the enhanced model, trained with augmented data, can accurately predict the properties of ceramic bricks with diverse compositions. By incorporating data augmentation, we aim to overcome the limitations posed by small datasets and improve the predictive accuracy of the model. Our study will also explore the robustness of the model in handling different input variables and its ability to generalize across another building material.

The expanded dataset, enriched by GAN-generated synthetic data, provides a comprehensive foundation for training the deep learning model. The use of MLP allows for effective mapping of the complex relationships between input variables and output properties. By validating the model on ceramic bricks, we aim to establish its versatility and effectiveness in predicting material properties beyond cementitious composites.

This work contributes significantly to sustainable construction practices by providing a reliable and efficient tool for assessing the properties of ceramic bricks. Ensuring both thermal efficiency and structural resilience, our approach has the potential to optimize material selection and enhance the overall performance of construction projects. The findings from this study could promote the development of more sustainable and efficient building materials.

1.5 General conclusions

This thesis is devoted to contributing in developing a general methodology to predict the properties of building materials based on its composition. To accomplish this, we have integrated a data augmentation model within artificial neural networks. We next provide a set of conclusions that we accomplished in this thesis:

Concrete Mix Design and Optimization:

Concrete is a fundamental construction material prized for its versatility, workability, and durability. Traditional methods for designing concrete mixtures, such as those established by the American Concrete Institute (ACI) and British standards, rely on empirical data and step-by-step procedures. However, these methods often fall short in accounting for the complex impacts of supplementary materials like pozzolans and fly ash, and they may not always yield the optimal mix. The integration of computational methods, including mathematical programming and machine learning, offers significant improvements. These methods can model the intricate relationships between mix components and the resulting concrete properties with greater accuracy, leading to more efficient and effective solutions. Studies highlight the potential of machine learning models in predicting concrete properties and optimizing mix compositions, paving the way for more sustainable and economical construction practices.

Advancements in Thermal Efficiency through ML and DL:

With the increasing emphasis on sustainability, enhancing the thermal efficiency of buildings has become crucial. Concrete's inherent thermal mass properties make it a prime candidate for improving energy efficiency. Machine learning (ML) and deep learning (DL) models, particularly Artificial Neural Networks (ANNs), show promise in predicting concrete properties, including thermal conductivity. The use of Generative Adversarial Networks (GANs) to augment datasets can significantly improve the accuracy of these models. By combining ANNs with GANs, researchers can address the challenges posed by limited datasets, leading to more precise predictions of thermal properties and better optimization of concrete mixtures for thermal performance.

Incorporating Phase Change Materials (PCMs):

Phase Change Materials (PCMs) integrated into concrete and cement composites enhance thermal regulation, absorbing heat during the day and releasing it at night. This integration improves energy efficiency in buildings. Traditional methods for assessing the thermal properties of these composites are labor-intensive and costly. By adopting deep learning techniques and data augmentation, researchers can overcome these challenges. Our study demonstrated that using a GAN to augment the dataset, followed by a Multilayer Perceptron (MLP) for prediction, can accurately forecast the thermal conductivity and compressive strength of PCM and nano-silica aerogel-enriched composites. This approach ensures both thermal efficiency and structural resilience, contributing significantly to sustainable construction practices.

Expanding the Model to Ceramic Bricks:

Building on previous work with concrete, our study expanded the deep learning model to assess ceramic bricks, another widely used building material known for its durability, thermal efficiency, and sustainability. By integrating data augmentation techniques and deep neural networks, the model was adapted to handle inputs specific to ceramic bricks, such as alluvial soil, laterite soil, demolition and waste material, temperature, firing shrinkage, loss on ignition, and bulk density. The model successfully predicted water absorption and compressive strength, demonstrating its versatility and effectiveness across different material types. This expanded application underscores the model's potential in optimizing material selection and enhancing the overall performance of various construction materials, thereby promoting more sustainable and efficient building practices.

Overall Impact:

The integration of advanced computational methods, particularly machine learning and deep learning models enhanced by data augmentation, represents a significant advancement in the field of construction materials. These methods provide more accurate predictions of material properties, facilitate the optimization of material compositions, and contribute to the development of more sustainable and efficient building practices. By addressing the limitations of traditional methods and expanding the applicability of these models to various construction materials, this research paves the way for future innovations in sustainable construction.

1.6 References

CHAPTER II

II. Identification of the prediction and optimization models

Use of operational research techniques for concrete mix design: A systematic review

Ana Carolina Rosa [a, c], Ahmed W. A. Hammad [b], Dieter Boer [c], Assed Haddad [d]

[a] Programa de Engenharia Ambiental, UFRJ, Universidade Federal do Rio de Janeiro, Rio de Janeiro, 21941-901, Brazil.

[b] School of Built Environment, UNSW Sydney (University of New South Wales), Sydney 2052, Australia

[c] Departament d'Enginyeria Quimica, Universitat Rovira i Virgili, Av. Països Catalans 26, 43007 Tarragona, Spain

Keywords: machine learning, mathematical programming, optimisation, concrete mix design.

2.1. Introduction

Concrete is one of the most commonly used construction materials in the world. Due to its unique feature combination of workability, mechanical properties, durability, and the wide availability of raw materials sources, concrete has become highly attractive for many applications in the civil industry [1]. Generally speaking, concrete can be defined as the mixture of a cement paste and water with fine and coarse aggregates. However, a concrete mix can be way more complex and have a wide range of constituents. It can be combined with materials, such as cement, pozzolans, fly ash, blast-furnace slag, silica, recycled concrete aggregate, polymers, fibres, and superplasticisers. Due to the vast possibility of constituents and the different variations in composition, it is possible to obtain a concrete mixture with different physical and chemical properties. Therefore, varying the components and their quantities allows the properties of both fresh and hardened states to be tailored to achieve the design specifications.

Each concrete application in the construction sector will require a specific mix composition to satisfy the performance requirements, such as workability, durability, and compressive strength. Therefore, designing good concrete is utterly vital to guarantee that the mixture will achieve the required specifications. In order to meet the best composition of concrete, designers have adopted concrete mix design to find an optimal concrete mixture. This procedure is defined as selecting suitable raw materials and proportioning the quantities of each constituent in order to produce a final product that meets the desired physical and chemical properties [2]. For this reason, promising development in the concrete mix design step can guarantee improvements in physical-mechanical performance, more outstanding durability of structures, and a reduction in production costs.

Most of the available mix design methods are based on empirical relationships, charts, and graphs developed from experimental investigations. They follow the same principles, and only minor variations exist in the mix design methods in selecting the mix proportions [3]. Some traditional methods for proportioning concrete are the American Concrete Institute (ACI) mix design method and the British mix design method [4]. According to DeRousseau [1], the conventional mix proportion classification can be broken down into two main methods: prescriptive specification and performance-based specification. The main difference between the two methods is that the first follows a step-by-step procedure to determine the quantity of each component. On the other hand, the second method does not require strict guidance. According to the design specification, a value of a required property is established; then, the designer can select any amount of cement, water, and aggregates and verify that the mixture achieves the value of the required property.

Although these traditional methods are way useful, the mathematical relationship between variables and the mixture composition can have some drawbacks. The first drawback of current concrete design methods is the lack of contemplation of Supplementary Cementitious Materials' influence (SCMs) on concrete properties. Since SCMs have been used lately to replace a portion of cement, they are prone to dramatically impact mechanical properties such as compressive strength [4]. Another drawback is that the mix design does not capture the true complexity of the relationship between the mixture design and the concrete's properties. The type and quantities of each material used are variable, and the estimation of these properties cannot be quantitatively and precisely evaluated. Lastly, the current concrete design methods do not present optimized values of concrete properties. The methods meet a satisfactory composition of the concrete constituents that reach the desired values of the properties. However, choosing a specific design does not guarantee that it is not an optimal concrete mix solution.

Besides the traditional methods, some research lines encompass the design of experiments to identify a good concrete mix composition. Compared to traditional methods, the application of the design of experiments can exhibit better solutions due to the numerous experiments. Although several studies have been investigating the problems related to concrete mix design, the optimisations of experimental design can undergo exponential growth if they increase the number of variables and their values [1]. Furthermore, due to the high degree of non-linearity between the dependent and independent variables, the regression models obtained to predict the properties of the concrete can hardly present a precise regression equation [5].

Therefore, both procedures have disadvantages and do not meet an optimised mix composition. For this reason, the computational aid system has been studied over the years in order to solve the problems regarding concrete mix design [6]. It formulates the problem by more accurately modelling the relationships between variables and the output response, leading to an optimal solution [7]. Thus, mathematical programming and machine learning have been prominent in the literature among the operational research techniques used to solve mixture proportioning issues. Both methods can be employed separately or as complementary approaches, frequently used to solve various engineering problems. Considering this fact, many researchers look to predict the properties of concrete in a more accessible, less costly, and less time-consuming way [8]. On the other hand, others seek to develop a more economical and sustainable concrete mix [9].

Several papers in the literature reviewed mathematical models and machine learning techniques in predicting concrete properties or the optimal composition of concrete. Chaabene et al. [10] examined machine learning (ML) models to forecast the mechanical properties of concrete. Another work presented a systematic review of the ML algorithms employed to predict compressive strength [11]. Besides that, Song et al. [12] reviewed the current literature on optimising mixing ratios using ML and metaheuristic algorithms, which offers insight into the continuous development of models in the field of hybrid optimisation. Based on these previous works, this article aims to examine relevant articles that have solved the concrete mix design problem using two types of operational research techniques: machine learning and mathematical programming. The novelty of this article lies in the lack of review articles focusing on these two techniquesapplied to the concrete mix design, which can bring a holistic understanding of the employment of these tools. Additionally, this article intends to propose a classification framework with the main features of the last works using mathematical programming or machine learning to solve problems with concrete mix design.

To place things into context, three major questions are addressed in the paper:

1. What are the main machine learning algorithms that have been used in order to come up with the concrete mix design?

2. What are the mathematical optimisation algorithms that have been used in order to help concrete mix design?

3. How is the literature linking both aspects, mathematical optimisation and machine learning, in order to design a concrete mix?

The rest of the paper is organised as follows: Section 2 briefly describes the concrete mix design research background. Next, section 3 describes the adopted study review methodology, presenting a systematic review and proposing a classification framework. The following section, section 4 evaluates the concrete mix design studies. Next, section 5 brings a discussion of the general findings. Finally, the final section contains concluding remarks on the general findings and research proposals.

2. Background

Before describing the proposed classification framework, some essential aspects of concrete mix design are summarised. Then, Section 2.1 and Section 2.2 outline mathematical optimisation models and machine learning methods adopted as solution strategies for the concrete mix design.

2.1. Overview of mathematical programming techniques

Mathematical models and computational optimisation methods are alternative solutions in concrete mix design to tackle the traditional methods and the time-consuming laboratory experiment optimisations. While the optimisation methods bring an optimal or near-optimal solution with low computational effort, the others only lead the designers to an ideal solution, which cannot be the most improved. These approaches initially establish one or more objectives (e.g., concrete properties) as functions of the decision variables (e.g., concrete constituents). Then, they search for the optimal concrete mixture by adopting optimisation algorithms. Mathematical methods can be classified into meta-heuristic, heuristic, and exact methods. Most researchers have frequently implemented them in the concrete mix design field to deal with mix design issues.

2.1.1. Exact methods

Exact methods have been used for a long time in mathematical optimisations and are very effective in modelling and optimising the properties of concrete. We can elucidate some model examples among the exact methods, such as linear [13], non-linear [14], and mixed programming [15]. Many authors have applied exact methods to estimate the optimal composition or predict other concrete properties. For instance, Jin et al. [15] proposed non-linear and mixed regression models to predict concrete strength based on mixture-design variables and curing age. The proposed model achieved higher accuracy than the linear method using the same variables and datasets. Another study that used non-linear programming was Habibi and Ghomashi's work [16]. Two mathematical models were developed based on sequential quadratic programming to optimise the self-compacting concrete mix design, minimising cost with compressive strength and slump as constraints. Miller et al. [17] developed a linear method for predicting Global Warming Potential (GWP) from concrete production and compressive strength for concretes containing replacement binder. The work proved the equations were suitable for predicting compressive strength and GWP for mixtures containing replacement binder. Fan and Miller [18] also took environmental problems into account. They developed a set of mathematical equations to find minimum Greenhouse Gas (GHG) emissions and to dictate the required concrete mix proportions, based on water-to-binder ratio and supplementary material, to achieve optimised GHG emissions. Li et al. [19] proposed a numerical simulation algorithm for calculating the overall elastic properties of ordinary concrete, which involves the finite element method in combination with Monte Carlo simulations. Comparisons with experimental test results of elastic modulus and Poisson's ratio indicated a good agreement with the proposed model.

2.1.2. Heuristic and Metaheuristic methods

Heuristic and metaheuristic methods are higher-level procedures aiming to find, develop, or select an algorithm that may provide a sufficiently good solution to an optimisation problem, especially with incomplete information [20]. While many heuristic algorithms are very specific and problem-dependent, a metaheuristic is a high-level problem-independent algorithm. Regarding concrete mix design issues, many studies have been using different heuristic metaheuristic methods to predict and optimise the properties of concrete. For example, Lee et al. [21] applied Harmony Search (HS) algorithm to find the optimum mix proportion satisfying the required strength and slump. The results showed that the HS algorithm could be a potent tool for high-performance concrete mix proportioning compared with other methods. Moayedi et al. [22] investigated the efficiency of the Ant Lion Optimisation (ALO) algorithm to predict the concrete slump. Besides, they compared it with Biogeography-Based Optimisation (BBO) and Grasshopper Optimisation Algorithm (GOA). The findings revealed that ALO outperformed both models. Another study[23] applied an Artificial Bee Colony (ABC) algorithm to predict chloride penetration in self-consolidating concretes. Comparisons among ABC and Genetic Algorithm (GA) and Particle Swarm Optimisation (PSO) were conducted and showed higher reliability [23].

Among the metaheuristic methods used to solve concrete mix design problems, many studies applying a genetic algorithm or some variation of swarm optimisation are notorious. For example, to improve the anti-cracking performance of high-strength concrete, Yue et al. [24] used GA to optimise concrete mix proportioning by minimising the cracking risk coefficient. Comparisons between a control concrete and the optimised one showed that crack resistance was considerably improved, and a reduction of 25% of the crack risk coefficient was achieved. Furthermore, Awoyera et al. [25] modelled the compressive strength, split tensile strength, and flexural strength of geopolymer concrete using gene expression programming, which showed a good prediction.

Regarding the swarm optimisations, Mashhadban et al. [7] investigated fibers' effects on self-compacting concrete properties using a hybrid model composed of PSO and a neural network. The model predicted splitting tensile strength, flexural strength, compressive strength, and fracture energy with a high level of accuracy. On the other hand, Kandiri et al. [26] estimated the compressive strength of the ground granulated blast furnace slag concrete by applying a multi-purpose Salp Swarm Algorithm (SSA).

Besides that, some authors have been employing other optimisation tools. For example, Knor et al. [27] applied a numerical solution based on the inverse heat transfer problem (IHTP) and direct search optimisation to investigate the thermal properties of hardening concrete, i.e., specific heat and thermal conductivity. Pazouki et al. [28] proposed a hybrid method using a neural network and firefly optimisation algorithm to predict and optimise the compressive strength of self-compacting concrete containing fly ash.

2.2. Overview of machine learning techniques

Machine learning is considered a branch of artificial intelligence that uses several algorithms to learn from the data. The learning step allows its behaviour to be modified based on its own experience, i.e., the system can predict or classify data based on previous data [29]. This behavioural modification consists of establishing some logical rules according to the input data to explore the potential of hidden patterns in the data used and improve the performance of a specific task. The use of these techniques allows the designers or the decision-makers to predict some desired outputs and choose a cleverer and more assertive decision regarding the properties of concrete [30].

The techniques have an outstanding performance in data processing and a remarkable ability to model the properties of concrete without explicit knowledge of the relationships between the properties of concrete and the variables [31]. For those reasons, they have been extensively used to model various properties of different types of concrete and solve concrete mix design issues. Another factor encouraging machine learning is the development of new types of concrete and the lack of mixing design guidelines, which can be noted in the Portland concrete guidelines that do not consider the new types of concrete developed.

With technological advancement and the search for more sustainable building materials, researchers have been looking for alternatives to replace part of the cement and aggregates, aiming to reduce CO_2 emissions and improve the quality of concrete. However, although the works regarding these materials are increasing, guidelines and standards do not cover concrete mixtures with these substitutions for common concrete mixtures. In addition, some types of concrete, such as fly ash-based geopolymer concrete, rubberised concrete, and recycled aggregate concrete, do not have specific guidelines to guarantee a good concrete mixture.

Machine learning has recently gained prominence in several areas, including the construction sector. Besides that, due to its accuracy in property predictions, some researchers have been efficiently applying these algorithms to solve concrete mix issues. The most used machine learning techniques in the literature to solve the concrete mix design issues can be broken down into five broad categories: Artificial Neural Networks (ANN); Support Vector Machine (SVM); Decision Trees (DT); Linear Regression; and Logistic Regression (Figure 1). The first three categories are the most employed in concrete mix design predictions and optimisations.

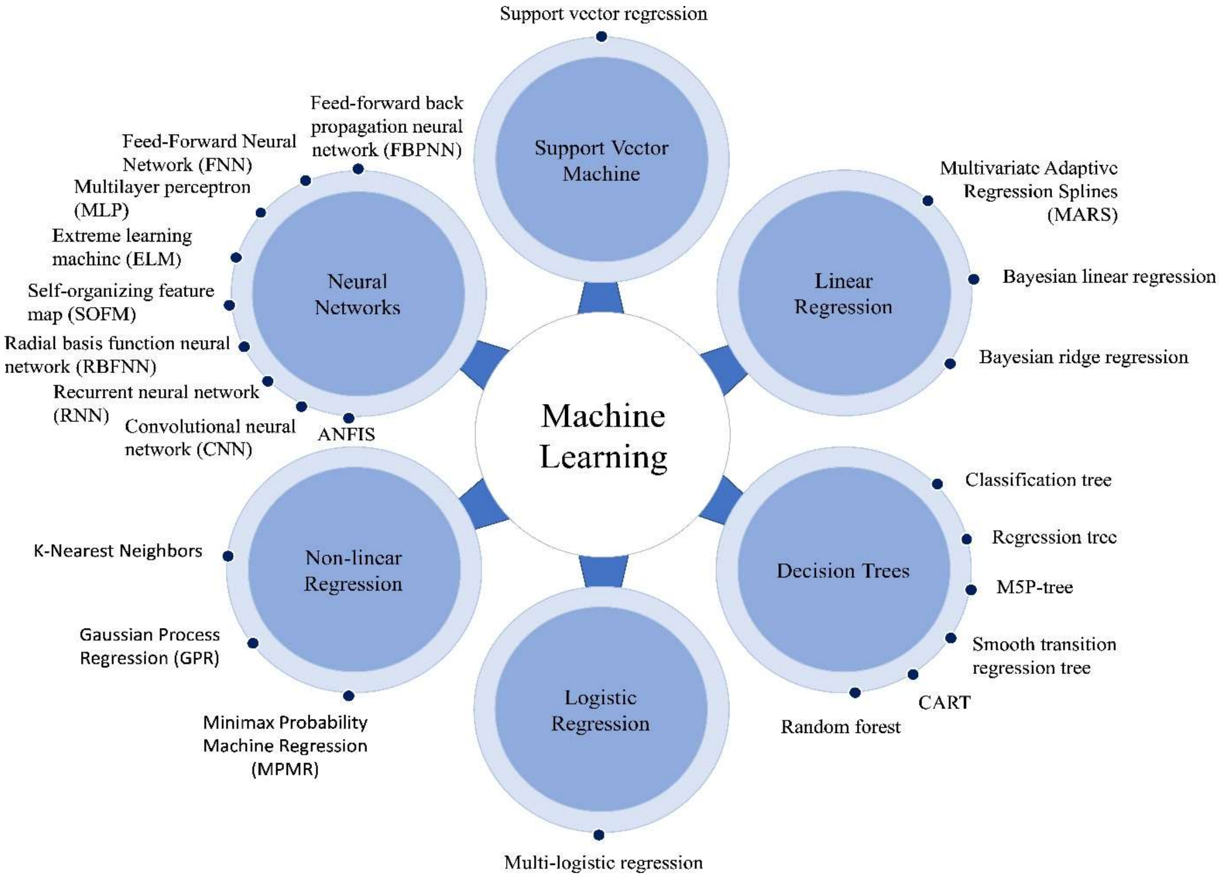

Figure II- 7. Main machine learning techniques used to solve concrete mix design problems.

2.2.1. Artificial neural networks

Artificial neural networks (ANN) are one of the most widespread machine learning techniques [32]. This model is inspired by the human neural system and the learning ability of the human brain. Its schematic representation consists of several layers of interconnected neurons responsible for letting the information flow from one neuron to another. Although the number of layers and neurons is directly related to the type of problem to be studied, ANN is generally composed of an input layer with the influencing variables, an output layer with the desired predicted variables, and at least one layer hidden between them. Yeh, a pioneer in this field, published some distinct works that influenced the development of the new architecture of ANN [33]. For instance, one of his works showed a higher accuracy of the strength model based on ANN when compared to a regression analysis model [34]. Furthermore, among his recent works, some models are analysing the strength of concrete using the design of experiments [35] and the accuracy of a neural network to predict the slump flow [36].

Different types of ANN are applied to solve concrete mix design problems in the literature. The ANN most used in modelling the prediction of concrete mixture properties are Feed-forward Neural Network (FNN) and the Backpropagation Neural Network (BPNN) approaches. To elucidate some examples, Yaprak et al. [37] applied a Feed-forward Backpropagation Neural Network (FBPNN) algorithm to predict the concrete compressive strength using different cement types, ages, and cure conditions as input parameters. Shahmansouri et al. [38] also predicted the compressive strength of geopolymer concrete using FNN, varying the specimen's age, NaOH concentration, and contents of neozolite, silica fume, and ground granulated blast-furnace slag. Yue et al. [24] applied FNN to predict concrete slump, compressive strength, tensile strength, and elastic modulus. Regarding BPNN, while Verm et al. [39] predicted the compressive strength by implementing this algorithm for fly-ash blended cement concrete mixes, Kellouche et al. [40] used BPNN to analyse the carbonation depth of fly-ash concrete. Both works indicated good agreement with the experiment results.

Other types of ANN have also been used frequently to predict concrete properties. One model widely applied to problems with concrete mix design is Multilayer Perceptron (MLP). Abbellán-García [41] applied MLP to predict the 28-day compressive strength of ultra-high-performance concrete using different combinations of supplementary cementitious materials:silica fume, fly ash, slag, glass powder, rice husk ash, fluid catalytic cracking residue, metakaolin, and limestone powder. Sargam et al. [30] compared nine machine learning algorithms, and the MLP model indicated a higher accuracy.

Besides these models, a few works employed other techniques. For example, Zhang et al. [42] established a strength prediction model of rubber fibre concrete using the Extreme Learning Machine (ELM), which showed that the model has high accuracy compared with other models as a conventional neural network. Nunez et al. [29] deployed a novel Recurrent Neural Network (RNN) to predict the compressive strength of recycled aggregate concrete. The model effectively captured the underlying features contributing to the compressive strength of the concrete. Czarnecki et al. [43] compared two ANN, Self-Organizing Feature Map (SOFM) and MLP, for predicting the compressive strength of cementitious composite containing ground granulated blast slag. Al-Mughanam et al. [44] implemented an Adaptive Neuro-Fuzzy Inference System (ANFIS) model to predict the compressive strength of self-compacting concrete containing treated palm oil fuel ash as a partial cement substitution. All of these models indicated high-efficient performance inpredicting the properties of concrete.

2.2.2. Support vector machine

The Support Vector Machine (SVM) is another machine learning technique highly chosen in concrete properties prediction. SVM is associated with learning algorithms that analyse regression and classification analysis data. According to Suykens and Vandewalle [45], SVM is a powerful and versatile method to deal with linear/non-linear classification, regression, and even outlier detection. The main important objective of the SVM classifier is to separate different classes with the most significant possible margin between two of them. In contrast, Support vector regression (SVR), a subset of SVM, tries to fit as many instances as possible on the street while limiting margin violations [46].

To name a few works applying SVM or SVR, Salimbahrami and Shakeri [47] adopted the SVM method to estimate recycled concrete's compressive strength. ANN and SVM techniques were compared, and although the results indicated SVM as an effective tool, its predictions are far closer to ANN. Besides, Thilakarathna et al. [48] carried out embodied carbon analysis using a set of machine learning algorithms to minimise the carbon footprint of high-strength and ultra-high-strength concrete without jeopardising the mechanical properties of the concrete. The outcomes found in this work were in good agreement with the ones discovered by Salimbahrami and Shakeri. On the other hand, Liu et al. [49] used SVM to predict the autogenous shrinkage of conventional concrete mixtures, and SVM was as accurate as ANN prediction. Other work introduced an advanced version of SVM, least-square SVM, to predict pervious concrete's permeability coefficient and compressive strength [50].

2.2.3. Decision trees

The Decision Trees (DT) methods also belong to the supervised learning family and can be applied to solve classification and regression problems. This predictive modelling is represented in a tree structure, where the leaves represent class labels and branches represent the characteristics set that lead to those class labels [48]. Due to its simplicity, some authors use this technique to estimate the properties of concrete mixtures. Random Forest (RF) and M5P trees stand out among the DT algorithms in this research area. Asthe other two methods of ANN and SVM presented previously, DT is also used to predict various properties of concrete.

For example, Gomaa et al. [51] modelled an RF model to predict slump flow and compressive strength as functions of influential parameters of the alkali-activated concrete. The model showed high efficiency in predicting both properties. Han et al. [52] presented an ensemble of ML models formulated by combining RF and SVM to predict the modulus of elasticity of recycled aggregate concrete, which produced more accurate results than those yielded by either standalone ML models. Behnood and Golafshani [53] studied the efficiency of the M5P algorithm in developing predictive models for the mechanical properties of concretes containing waste foundry sand as a partial or total replacement for fine aggregate. This research indicated a successful use of the M5P tree in generating predictive models for compressive strength, modulus of elasticity, flexural strength, and splitting tensile strength.

In addition to the tree-based techniques mentioned above, researchers like Anyaoha et al. [14] and Feng et al. [54] also used other DT techniques to predict the concrete's properties. The former used a boosting ensemble of Smooth Transition Regression Trees (STR-Tree) and a boosting algorithm to create a more robust model. The developed model predicted the compressive strength of high-performance concrete using its constituents and mixture proportions as input variables, which showed its dominance in prediction accuracy over the other methods. The latter employed an adaptive boosting algorithm integrated into weak models, such as Classification and Regression Tree (CART), ANN, and SVM. The outcomes indicated CART as the higher performance for the compressive strength prediction of concrete containing slag and fly ash.

3. Review Methodology

The systematic method adopted in this review aimed to address the research questions elucidated in the introduction section. Therefore, the review methodology can be broken down as follows:

Step 1: The first step is related to the search process for articles involving the concrete mix design topic and selecting the most relevant articles, which fit more into the scope of the questions raised.

Step 2: The second step involves analysing the previously selected articles to specify the general characteristics of the mathematical optimisation modelling and machine learning techniques used in this subject.

Step 3: The third step encompasses the development of a classification framework linking the predicted properties and the methods applied for the prediction and optimisation of concrete mixtures.

Step 4: Finally, the last step involves the evaluation of the material based on the proposed framework.

3.1 Systematic Review Search

The review search was conducted to find and select the most relevant articles. In order to ensure this, a set of keywords was well-defined. Two main groups of keywords have been used to find the literature related to concrete mix design problems:

Group 1: Words related to concrete mix design: "Concrete Mix Design"; "Concrete Mixture"; "Concrete Proportioning".

Group 2: Words related to mathematical modelling or machine learning: "Mathematical Programming"; "Mathematical Modelling"; "Optimisation"; "Algorithm"; "Metaheuristic"; "Machine Learning".

Combining the two groups, a general search query was achieved with the mais essential keywords and then employed on academic databases, including Scopus and Science Direct. The following query was applied in the databases for the extraction: "concrete mix design" OR "concrete mixture" OR "concrete mix" OR "concrete proportioning" AND ("mathematical programming" OR "mathematical modelling" OR "mathematical modeling" OR "mathematical model" OR "non-linear model" OR "nonlinear model" OR “particle swarm optimization” OR “particle swarm optimisation” OR “metaheuristic optimization” OR “metaheuristic optimisation” OR "machine learning" OR "algorithm" OR "genetic algorithm" OR "neural network" OR "support vector machine" OR "regression tree" OR "random forest" OR "linear regression" OR "polynomial regression" OR "support vector regression" OR "gaussian process regression" OR "boosted tree").

Initially, the search based on the abovementioned databases generated a total of 2040 papers. After the first screening, some constraints, such as conference papers, books, language, and a specific timeline period, were applied to the total of registers, which were reduced to 1087 papers after specifying these constraints After evaluating the title and removing the duplicates, the following screening checked the relevance of the remaining documents, which filtered 310 papers. This step mainly focused on selecting the documents linked to solving concrete mix design problems. Next, further examination was performed to guarantee that the scope of the selected papers was aligned with this review by scrutinising their abstracts, which resulted in a set of 148 remaining documents. Then, they were evaluated by thoroughly reading each paper's introduction and discussion sections. The remaining papers were chosen according to the application of mathematical methods and machine learning techniques to find the best composition of the concrete or any optimised property of this composition. Any article that only focused on experiment outcomes and did not solve an optimisation for concrete mix design or predict any concrete property based on its constituents was removed. After the last screening, a total of 100 papers were selected to be evaluated in this review. To summarise the systematic review procedure, Figure 2 describes each step adopted in this review.

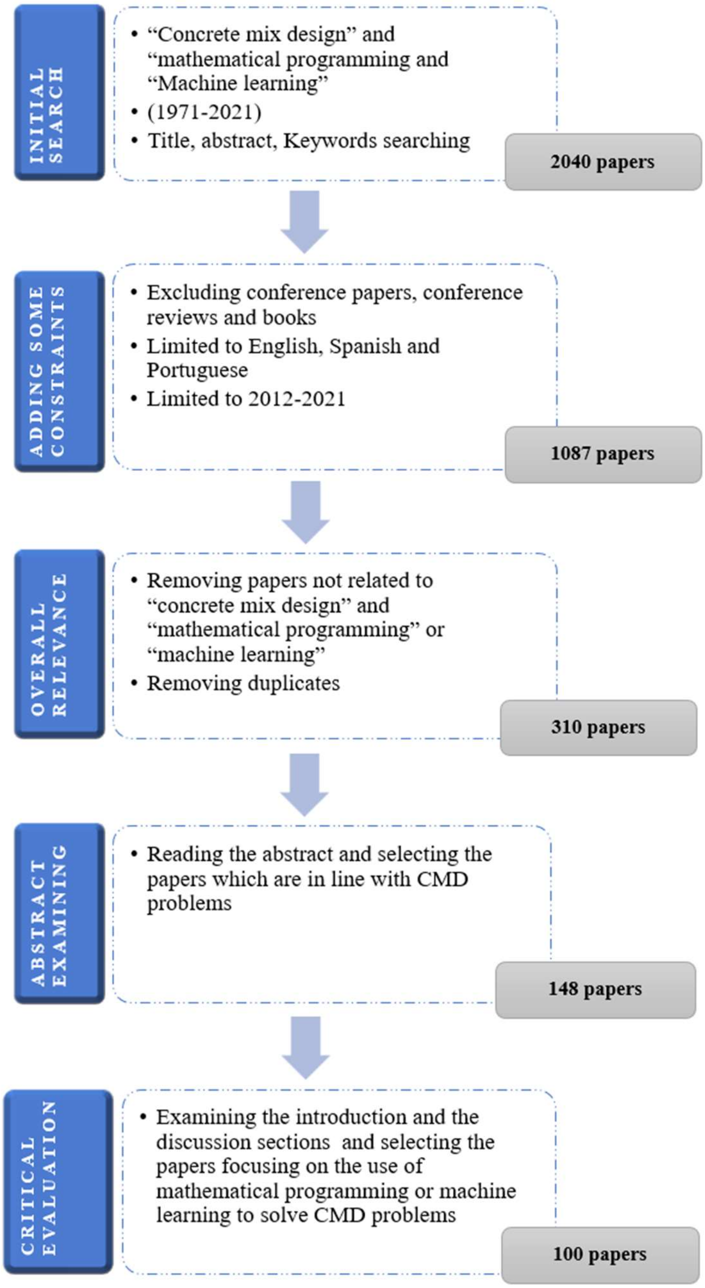

Figure II- 8. The procedure of systematic review search

3.2 Descriptive analysis

The selected articles were first evaluated according to the year of publication. Then, figure 3 was elaborated to reveal the publication rate trend considering the ten years considered in this analysis. It is noticeable that the number of papers related to concrete mix design has continuously risen over the years.

Despite being a common topic, this growth can be explained by the different types of concrete studied and several possibilities of the concrete mixture composition aligned with the development of computational techniques. In addition, the search for more sustainable building materials that promote less environmental impact and greater thermal comfort has also stimulated this growth.

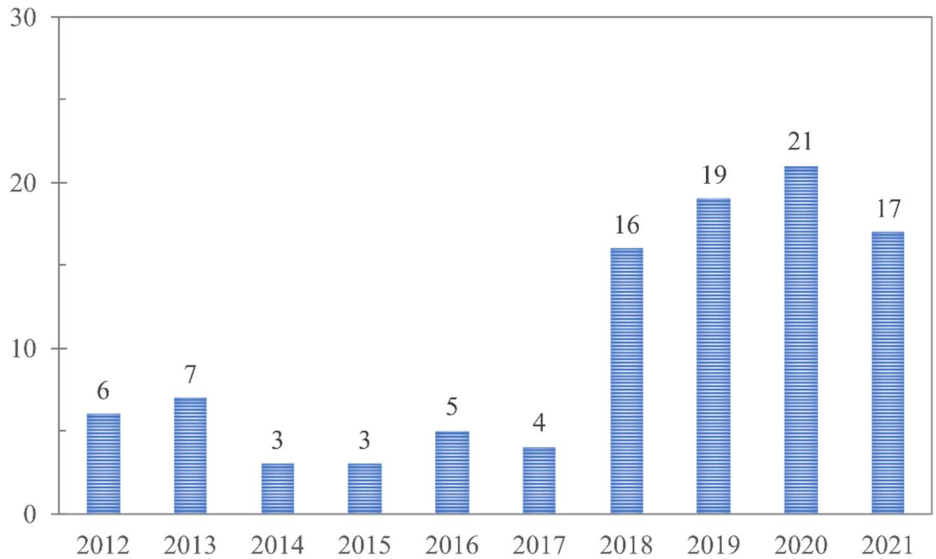

Figure II- 9. Distribution of reference papers by year

Following, the publishing journals most adopted to submit papers regarding concrete mix design were evaluated. Table 1 displays the spread of papers analysed across the publishing journals, bringing 100 documents from 40 different journals. Although there is a wide variety of publishing journals, the Construction and building materials journal contains the most significant publications involving concrete mix design problems, followed by the Journal of Building Engineering and the Journal of Cleaner Production.

Table 1. Distribution of reference papers by publication

Journal	Number of studies
Construction and Building Materials	35
Journal of Building Engineering	7
Journal of Cleaner Production	6
Cement and Concrete Research	3
Cement and Concrete Composites	3

Sustainability	3
ACI Materials Journal	3
Applied Sciences	2
Computers and Concrete	2
Frontiers in Materials	2
Frontiers of Structural and Civil Engineering	2
Advances in Civil Engineering	2
European Journal of Environmental and Civil Engineering	2
Materials	2
International Journal of Concrete Structures and Materials	1
Automation in Construction	1
Journal of the International Measurement Confederation	1
Neural Computing and Applications	1
Advanced Engineering Informatics	1
Advances in Engineering Software	1
Advances in Materials Science and Engineering	1
Alexandria Engineering Journal	1
Arabian Journal for Science and Engineering	1
Engineering with Computers	1
Expert Systems with Applications	1
Heat and Mass Transfer	1
Heliyon	1
International Journal of Industrial Engineering: Theory Applications and Practice	1
International Journal of Pavement Research and Technology	1
Journal of Computing in Civil Engineering	1
Journal of Construction Engineering and Management	1
Journal of Materials Research and Technology	1
KSCE Journal of Civil Engineering	1
Measurement	1
Polymers	1
Renewable and Sustainable Energy Reviews	1
Resources, Conservation and Recycling	1
Science of The Total Environment	1
Soft Computing	1
Structural Concrete	1

Depending on the uses of concrete, a huge variety of concrete can be designed for different objectives, such as high-rise buildings, sustainable buildings, economic buildings, residential buildings, and pavements. As previously described, concrete comprises cement, water, and solid aggregates. However, different materials can be added to this mixture to impart a specific property. Currently, there is a growing demand for building materials with higher performance, more sustainable, and less harmful to the environment. As a result, there is a great deal of adding other materials or partially replacing the cement, which is responsible for emitting significant portions of CO_2 into the atmosphere.

Therefore, the design of concrete mixes is a powerful assistant in specifying the composition of the constituent materials to achieve a specific characteristic. Each concrete mixture design will depend exclusively on the objectives to be achieved in the fresh and hardened states of concrete specified by the designers. Among the papers selected for this review, it was found that there are two well-defined research lines. The first research line is characterised by predicting a specific property of the concrete that one wishes to obtain from the composition of the constituents or a combination of the composition with other properties. While the second research line brings optimisation of the concrete mix design problem, where one or more objectives are selected to be minimised or maximised and thus obtain the optimum concrete mix for each case studied.

Figure 4 breaks down these two lines of research into three technique classes most applied in the retrieved papers. The first encompasses machine learning (ML) techniques, the second includes mathematical models (MM), and the third presents a hybrid system (HyS) containing both approaches.Machine learning is the most used method in papers where the properties prediction is sought, appearing in 45% of documents. Then comes hybrid systems with 24% and mathematical models with 14%. However, when the main objective is optimisation, there is an inversion of the techniques used. Mathematical models appear as the most applied tool with 12%, followed by hybrid systems with 5%. It is worth mentioning that there is no document in optimisation research presenting machine learning as the unique technique used.

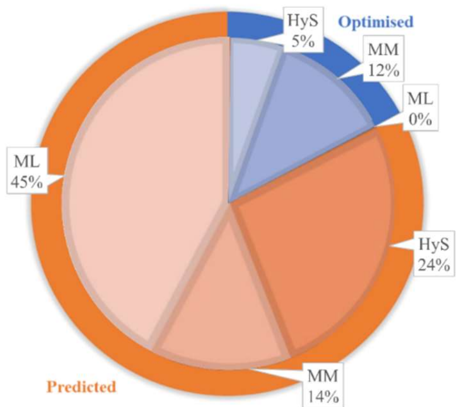

Figure II- 10. Papers distribution according to the research line

3.3. Literature review classification framework

In this section, a classification framework of the literature review is proposed. The framework classification identifies some categories which characterise the research area examined in the concrete mix design. Besides, each category is mapped to several suitable classes. The main purpose of this framework is to detail the works published in the literature in a consistent and organised way, thus enabling the comprehension of the study and comparisons for future works. Considering the two lines of research found in the literature, the classification of the framework can start from the initial objective of the papers: if it is for the prediction of some property or if it is the optimisation of mixtures with certain minimised or maximised properties. Figure 5 describes the classification framework broken down into five categories: research line, concrete type, concrete property evaluated, output/objective type of prediction and optimisation cases, and the type of computational aid system used.

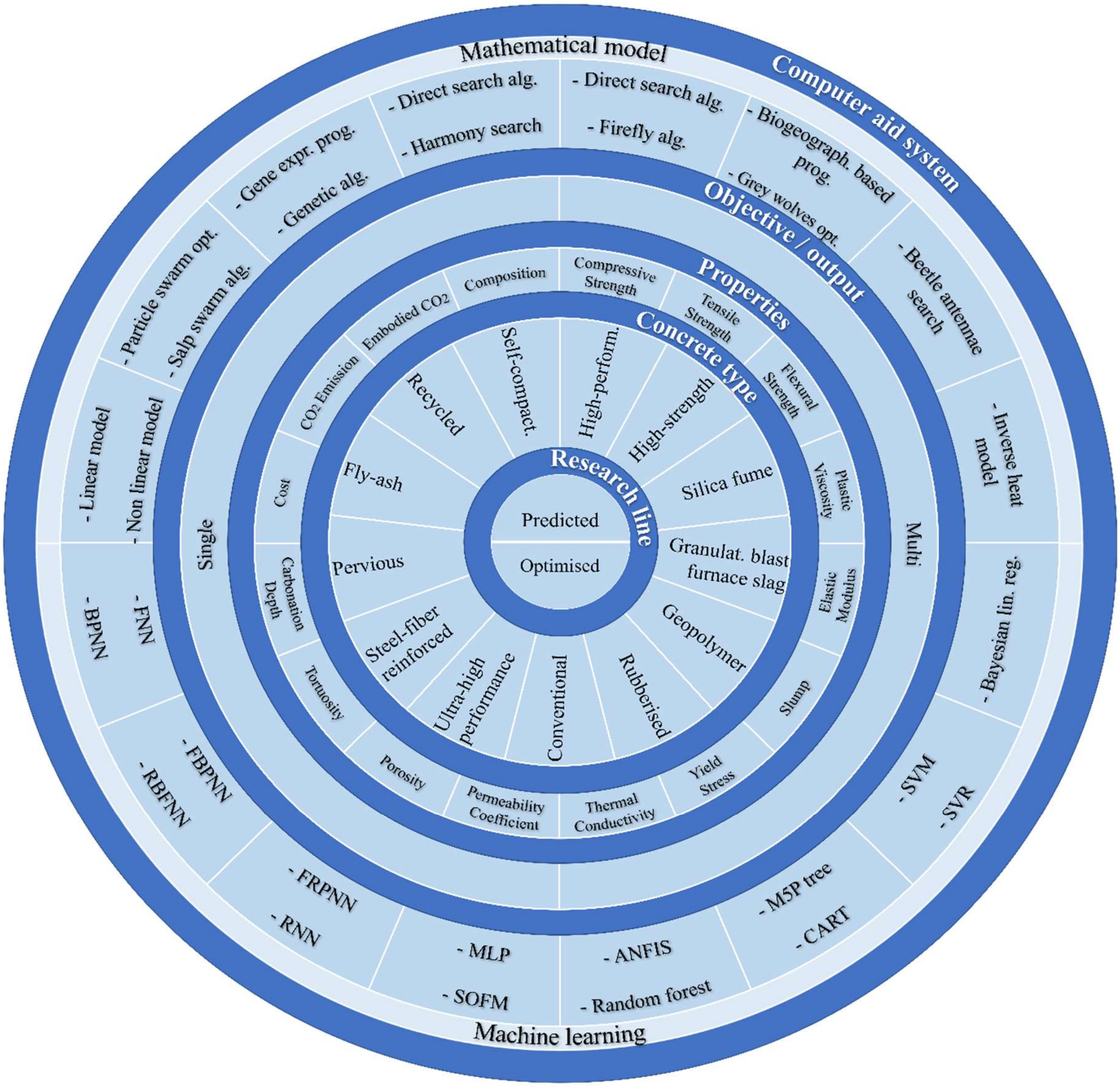

Figure II- 11. The proposed classification framework

The first category of the proposed classification framework is based on the research line adopted in each study. After analysing the selected papers, two specific lines were outlined. One considers the prediction of concrete properties, while the other focuses on optimizing the design of concrete mixes. After this indication, one can follow the next step and get into the second category to continue the work classification.

The second category examines the main types of concrete investigated in the literature review. It includes the types most used in the construction sector as well as new, more sustainable concretes. Among those evaluated in this review are: conventional concrete, rubberised concrete, recycled aggregate concrete, geopolymer concrete, fly ash concrete, ground granulated blast furnace slag concrete, silica fume concrete, self-compacting concrete, reinforced concrete, high-performance concrete, ultra-high-performance concrete, high-strength concrete, plastic concrete, pervious concrete, oil-palm shell concrete, self-consolidating concrete, pozzolanic concrete, and alkali-activated concrete.

According to the research line adopted, there is a wide variety of properties that researchers seek to predict or optimise. The next category allows the specification of which property is evaluated. Among the properties investigated, several properties are related to the concrete's fresh and hardened state. For example, some works consider slump and slump flow regarding the workability's concrete. Furthermore, many predictions and optimisations of compressive strength, flexural strength, tensile strength, and modulus of elasticity regarding mechanical properties are found in the literature. Besides that, some works evaluate other properties such as cost and sustainability-related properties of the material such as embodied carbon, CO_2 emission, carbonation depth, and thermal conductivity.

The following classification category characterises the output or objective type from each study. Here, it is evaluated if the prediction model output or the optimisation objective considers one or more properties, i.e., single or multi-output/objective. Both research lines present some works predicting or optimising just one property and others considering multiple properties to be predicted or optimised.

The last category considers the applied method to predict the properties or optimise the objective functions. As previously discussed in the introduction section, mathematical programming and machine learning are widely adopted in concrete mix design problems. The main techniques used are artificial neural networks, decision trees, random forest, support vector machine, ANFIS, linear model, non-linear model, particle swarm optimisation, salp swarm algorithm, firefly algorithm, harmony search, and beetle antennae search. Table 2 summarises the technique acronym used in this review.

Table 2. List of the techniques acronyms used in this paper

ABC	Artificial Bee Colony	HS	Harmony Search
ALO	Ant Lion Optimisation	ICA	Imperialist competitive algorithm
ANFIS	Adaptive neuro-fuzzy inference systems	KNN	K-nearest neighbour
ANN	Artificial Neural Network	LCA	Life cycle assessment

BAS	Beetle Antennae Search	LR	Linear regression
BBO	Biogeography-Based Optimisation	LSSVM	Least square support vector machine
BPNN	Back propagation neural network	MARS	Multivariate adaptive regression splines
CART	Classification and regression tree	MEP	Multi-expression programming
CNN	Convolutional neural network	MLP	Multilayer perceptron
DT	Decision Tree	MOLSSVR	Multi-output least squares support vector regression
EA	Evolutionary algorithm	PSO	Particle swarm optimization
ELM	Extreme Learning Machine	RBFNN	Radial basis function neural network
ESIM	Evolutionary support vector machine inference model	RF	Random forest
FA	Firefly algorithm	RNN	Recurrent neural network
FBPNN	Feed forward back propagation neural network	SLCA	Soccer league competition algorithm
FNN	Feed forward neural network	SOFM	Self-organizing feature map
FRPNN	Feedforward resilient propagation neural network	SQP	Sequential quadratic programming
GA	Genetic algorithm	SSA	Salp swarm algorithm
GEP	Gene expression programming	STR	Smooth Transition Regression
GMDH	Group method of data handling	SVM	Support vector machine
GOA	Grasshopper Optimisation Algorithm	SVR	Support vector regression
GPR	Gaussian process regression	WCA	Water cycle algorithm
GWO	Grey wolves optimization	wSVM	Weighted support vector machine

4. Evaluation of studies

In this section, studies are evaluated based on the proposed classification framework. In the following subsections, the papers selected in this review are organised according to the categories indicated in Figure 6.

4.1. Review according to the main objective of the work

The first category relies on the primary purpose of the paper. As previously diagnosed, papers addressing the topic of concrete mix design can be primarily classified into two lines of research: prediction and optimisation. According to Figure 5, the prediction of concrete properties represents most of the studies found in the literature. It reveals that in the analysis carried out on the papers selected, around 83% of the papers take into account the prediction of properties, and only 17% of the documents aim at optimising the concrete mix design problem, minimising or maximising the objective functions. However, some works adopt both research lines, i.e., the same study considers a property prediction and then performs a concrete mix optimisation.

Both research lines have employed machine learning techniques or mathematical models to characterize and analyze concrete mixtures. For instance, Huang et al. [55] followed the prediction line and developed an ANN model to determine the multiple mechanical properties of rubberised concrete based on concrete mix composites. On the other hand, Wang et al. [56] aimed to optimise a concrete mixture of slag-blended concrete with minimum CO2 emission. Zhang et al. [20] employed different ML models to predict concrete objectives first. As a second step, the best prediction model was selected as the objective function for the optimisation procedure.

4.2. Studies based on the concrete type

This section focuses on the types of concrete analysed in the retrieved papers. The documents bring up different combinations of concrete constituents that engender a wide variety of concrete types. Therefore, selecting materials and their quantities allows designing concrete with diverse compositions according to the project requirements. Due to this variety, it is possible to notice that studies have evaluated different types of concrete from 2012 to 2021.

Table 3 presents all types of concrete studied in prediction and optimisation research lines. It is worth mentioning that there is a considerable diversity of concrete in both research lines. However, there is a predominance of conventional concrete, concrete containing recycled aggregates, fly ash, silica fume, and slag. These five types of concrete are analysed by 55% and 67% of prediction and optimisation research studies, respectively. Both research lines reveal fly ash concrete as the most studied concrete type. It achieves first place in prediction research and second place in optimisation research.

Although fly ash-blended concrete appears as the most relevant on this list, other types of concrete are also noteworthy, such as plastic concrete and limestone concrete. Furthermore, with the ongoing search for more sustainable materials and mechanical properties similar to conventional concrete, many authors have been investigating new constituents and, thus, new types of concrete. To name a few, Iqbal et al. [57] modelled the mechanical properties of concrete containing waste foundry sand. Besides that, Chung and Tia [58] investigated the effects of cementitious materials on the compressive strength and surface resistivity of limestone concrete mixes. Zhang et al. [59] inspected the replacement of natural coarse aggregate with oil palm shell and elucidated the pros and cons of oil palm shell concrete.

Table 3. Concrete types evaluated in the literature review

Predicted

Concrete type	**Qt.**	**Concrete type**	**Qt.**
Concrete containing fly ash	17	Plastic concrete	3
Recycled aggregate concrete	15	High-strength concrete	2
Conventional concrete	12	Ultra-high-performance concrete	2
High-performance concrete	12	Pozzolanic concrete	2
Concrete containing slag	10	Alkali-activated concrete	1
Concrete containing silica fume	8	Waste foundry sand concrete	1
Reinforced concrete	7	Ultra-high-strength concrete	1
Rubberised concrete	5	Pervious concrete	1
Self-compacting concrete	4	Oil palm shell concrete	1
Self-consolidating concrete	4	Hydraulic concrete	1
Geopolymer concrete	3	Lime-stone concrete	1

Optimised

Concrete type	**Qt.**	**Concrete type**	**Qt.**
Concrete containing slag	6	High-performance concrete	2
Concrete containing fly ash	5	Reinforced concrete	1
Conventional concrete	3	Plastic concrete	1
Concrete containing silica fume	2	Self-compacting concrete	1
Recycled aggregate concrete	2	Pozzolanic concrete	1
High-strength concrete	2	Lime-stone concrete	1

4.3. Studies based on concrete property

Designing a concrete mixture relies on specifying the properties of the final product and guaranteeing it will meet the property requirements of the project. For this reason, many researchers and designers focus on experiments and mathematical models that better represent the relationship between concrete composition and its properties. Therefore, the third classification category proposed is the concrete property. Table 4 depicts the distribution of the papers according to the main properties predicted and the objectives optimised. Most of the predictions' outputs are related to the mechanical and chemical properties.

Concerning the predicted properties, several papers evaluate the best technique of machine learning and mathematical models to predict concrete properties. Compressive strength is the most investigated property in all studies, representing 43% of the papers evaluated. The predominance in most studies is justified by its importance in the durability design and predicting the service life of concrete structures. Still, regarding the compressive strength, there are some varieties related to the curing time; however, most works try to predict this property after 28 days. Two other highly evaluated properties are elastic modulus and slump, which appear in 9% and 7% of the works. Elastic modulus is another mechanical property that reflects the ability of concrete to deflect elastically. And the slump, related to the fresh concrete state, is a factor that directly contributes to the workability of the mixture. Forecasting the concrete properties, especially the mechanical properties, helps designers in the early characterization of the concrete, thus saving design time.

On the other hand, the papers consideringthe optimisation process adopt fewer properties as objective functions. Three appear most frequently in single or multi-optimisations among the main properties. 37% of the works look for the optimal concrete mix minimising the cost, 17% look over the compressive strength, and 14% use CO_2 emission as an objective function to minimise. Optimizations, despite being extremely important for the specification of a more suitable concrete mixture for a project, there is still a lack of studies focusing on different objective functions. Most studies still focus exclusively on compressive strength and cost.

Table 4. Main predicted and optimised properties found in concrete mix design problems

Predicted

Properties	**Qt.**	**Properties**	**Qt.**
Compressive strength	66	Permeability coefficient	1
Tensile strength	14	Coefficient of thermal expansion	1
Elastic modulus	13	Modulus of rupture	1
Flexural strength	11	Poisson's ratio	1
Slump	10	Dry shrinkage rate	1
Carbonation depth	4	Shrinkage	1
Mix proportion	4	Porosity	1
Plastic viscosity	3	Transport tortuosity	1
Ultra-pulse velocity	3	Specific heat	1
Thermal conductivity	2	Chloride penetration	1
Slump flow	2	Rapid chloride permeability	1
Interface yield stress	2	Water absorption	1
Autogenous shrinkage	2	GWP	1
Rebound number	1	Fracture energy	1
Embodied carbon	1		

Optimised

Properties	**Qt.**	**Properties**	**Qt.**
Cost	13	Tensile strength	1

Compressive strength	6	Cracking risk coefficient	1
CO_2 emission	5	Energy consumption	1
Embodied CO_2	2	Resource consumption	1
Slump	2	GHG emission	1
Flexural strength	1	Diffusion coefficient of CO_2	1

4.4. Studies based on objective/output type

The following category breaks down the papers into four main classes. Two classes are associated with the predictive research line, characterised as single and multi-output. While the other two are linked to the optimisation research line, defined as single and multi-objective. The distribution of the studies addressing each of these classes is illustrated in Figure 6. As is seen from Figure 6, a prediction with a single output occupies the most significant proportion of studies, with 70% of the studies examined. Another interesting point is that the researchers focus more on single-objective optimisation studies than on multi-objective optimisation, representing only 7% of the documents.

A possible reason for many studies in the prediction area is exploring models more effectively for each type of concrete property. Regarding the optimisation of concrete, the studies related to single optimisation are still more significant than those with multiple objectives. It can also be explained by the effort in optimising different properties simultaneously, where the optimisation of one objective function affects the other's results. Both lines of research indicate the need to be further explored. The first one requires more reliable models for different types of concrete, including all possible constituents that affect the properties of concrete. And the second line should be better explored because the design of concrete mixtures is a trade-off process with few studies published in the literature so far.

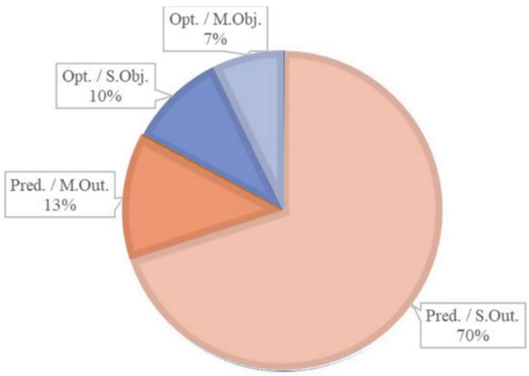

Figure II- 12. Number of studies in concrete mix design according to output or objective function type considered

Figure 7 reveals the distribution of these four classes over 2012-2021, which shows the papers according to the output or objective function type. Besides the significant growth in the prediction area of a single property, there is a slight development in single and multi-objective optimisations of concrete properties. Therefore, an increasing trend in the concrete mix optimisation research line is spotted.

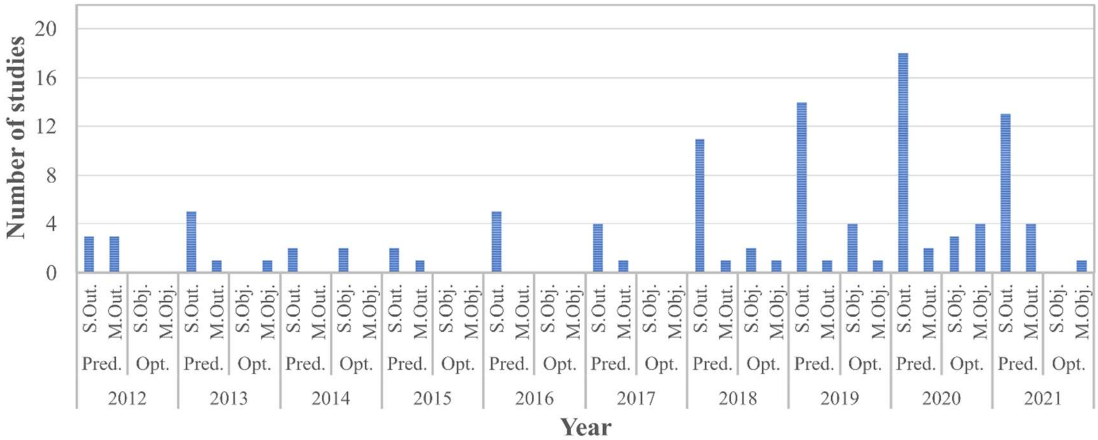

Figure II- 13. Studies distribution according to output or objective function type

4.5. Studies based on computer aid systems adopted

The last classification category takes into account the computer aid system adopted. Based on the selected papers, the works are categorised into three main groups: (1) machine learning techniques, (2) mathematical programming models, and (3) hybrid systems, which englobes the two methods mentioned previously. As seen previously in Figure 4, machine learning techniques dominate concrete mix design issues accounting for 45% of the studies reviewed. Regarding the mathematical models, the contribution decays to 26% of the papers, and the hybrid system takes 29% of the total concrete mix design studies examined. In order to comprehend the methods used, Figure 8 breaks down the three groups according to each research line from 2012 to 2021. From 2018 onwards, there has been significant growth in publications in general. Besides the spread of machine learning as the only tool for predicting properties, there is a notable development in adopting the hybrid system to predict properties and optimise the concrete mixture.

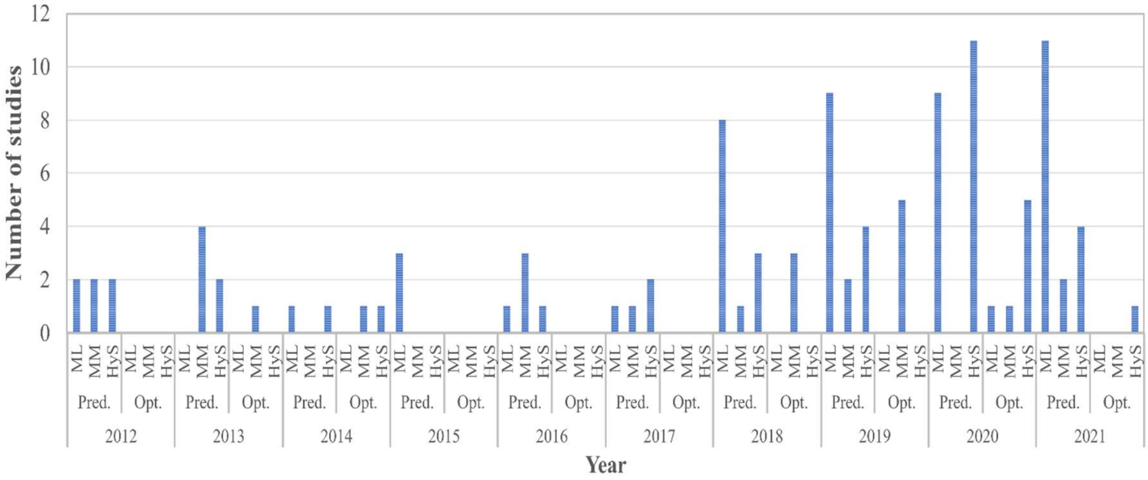

Figure II- 14. Studies distribution considering the techniques adopted
(ML – Machine learning, MM – Mathematical model, HyS – Hybrid system)

Further explaining the type of each technique of machine learning and mathematical models, Table 5 brings together the techniques of each method grouped by similar tools and algorithms used in the selected papers. Since the adopted approach to designing a concrete mixture depends on the research line and the properties chosen to be predicted or optimised, Table 4 is divided into two parts. The first part quantifies machine learning and mathematical models used to predict concrete properties, while the second part shows the techniques applied to optimise the concrete mixture.

Considering the prediction research line and the application of machine learning techniques, the first part of Table 5 shows that ANN covers most studies related to the properties prediction. They are responsible for 30% of the total works in this research line. Then, the two machine learning techniques most applied to determine concrete properties are DT and SVM with 14% and 13%, respectively. Among these three main groups, the most common techniques adopted are multilayer perceptron, feed-forward neural network, backpropagation neural network, random forest, and support vector machine. In order to evaluate the technique most adequate that generates the most reliable response, many works use a comparative methodology between the techniques. Therefore, more than one approach is used to predict the concrete properties, which allows the authors to select the technique with a minor error.

Regarding the mathematical models in the prediction line, metaheuristic methods have been gaining prominence and appear in 24% of the works, followed by exact methods with 8% and only 2% of heuristic methods. Among the main techniques, GA is the one that appears most frequently in the evaluated studies. In fewer studies, other techniques implemented are PSO, GEP, linear models, and non-linear models. Part of these mathematical algorithms in this research line is employed to tune the machine learning technique's hyperparameters and improve the property prediction.

The second part of Table 5 indicates a reduction in machine learning techniques compared to other lines of research. Techniques of machine learning groups account for only 7% for ANN, 7% for DT, and 5% for SVM of studies related to optimisation. On the other hand, exact and heuristic methods significantly increase, representing 18% and 18% of the total techniques used in this optimisation line. The approaches adopted are the polynomial model, GA, FA, GEP, and PSO. It is worth remembering that some techniques used to predict properties are used to formulate the objective functions to optimise concrete mixtures.

Table 5. Main techniques applied in prediction and optimisation problems related to concrete mix design

Predicted				
Machine Learning				
	Technique	**Qt.**	**Technique**	**Qt.**
Neural Network	Multilayer perceptron (MLP)	16	Adaptive neuro-fuzzy inference system (ANFIS)	2
	Feed-forward backpropagation (FBPNN)	15	Self-organising feature map (SOFM)	1
	Backpropagation (BPNN)	11	Convolutional neural network (CNN)	1
	Feed-forward (FNN)	9	Extreme learning machine (ELM)	1
	Radial basis function neural network (RBFNN)	3		
Support Vector Machine	Support vector machine (SVM)	14	Evolutionary support vector machine inference model (ESIM)	1
	Support vector regression (SVR)	6	Weighted support vector machine (wSVM)	1
	Least squares support vector machine (LSSVM)	2	Multi-output least squares support vector regression (MOLSSVR)	1
Decision Tree	Random forest (RF)	8	Adaptive boosting algorithm (AdaBoost)	2
	M5P model tree	5	XGboost	1
	Gradient boosting tree	4	Boosting Smooth Transition regression trees (BooST)	1
	Regression tree	3	Chi-squared automatic interaction detector (CHAID)	1
	Decision tree	2		
Linear Regression	Multivariate adaptive regression splines (MARS)	5	Ridge regression	2
	Linear regression (LR)	4	Lasso regression	1
Non-linear Regression	Gaussian process regression (GPR)	4	Minimax probability machine regression	1
	K-nearest neighbour (KNN)	2	Non-linear regression	1

Logistic Regression	Multilogistic regression	1		

Mathematical Optimisation

Metaheuristic Method	Genetic algorithm (GA)	12	Soccer league competition algorithm (SLCA)	1
	Particle swarm optimisation (PSO)	4	Group method of data handling (GMDH)	1
	Gene expression programming (GEP)	4	Ant lion optimisation (ALO)	1
	Firefly algorithm	3	Deep learning	1
	Beetle antennae search	3	Grasshopper optimisation algorithm (GOA)	1
	Genetic programming	3	Grey system theory	1
	Biogeography-based optimisation (BBO)	2	Direct search optimisation algorithm	1
	L-SHADE	1	Artificial bee colony algorithm (ABC)	1
	Imperialist competitive algorithm (ICA)	1	Grey Wolves Optimisation	1
	Multi-expression programming (MEP)	1	Mixed regression model	1
	Salp swarm optimisation	1	Harmony search	1
	Water cycle algorithm (WCA)	1		
Exact Method	Linear model/regression	5	Semi-empirical model based on Fick's law	1
	Non-linear model	3	Model-based on Abram's law	1
	Life cycle assessment	2	Densified mixture design algorithm	1
	Hydration model	1	Principal component analysis (PCA)	1
	Carbonation reaction model	1		
Heuristic Method	Five-layer fuzzy inference system	1	Fuzzy logic	1
	Fuzzy TSK	1		

Optimised Machine Learning

	Technique	Qt.	Technique	Qt.
Neural Network	Backpropagation (BPNN)	3	-	
Decision Tree	Random forest	2	Gradient boosting tree	1
Support Vector Machine (SVM)	Support vector regression (SVR)	1	Evolutionary support vector machine inference model (ESIM)	1

Mathematical Optimisation

Exact Method	Polynomial model	15	Sequential quadratic programming	1
	Life cycle assessment (LCA)	1	Adaptive surrogate model	1
Metaheuristic Method	Genetic algorithm (GA)	8	Beetle antennae search (BAS)	1
	Firefly algorithm (FA)	2	Biogeography-based optimisation (BBO)	1
	Gene expression programming (GEP)	2	Water cycle algorithm (WCA)	1
	Particle swarm optimisation (PSO)	2	Evolutionary algorithm (EA)	1

The proposed classification framework intends to guide researchers through the published works on concrete mix design and recognize patterns in each research line. Consequently, it is possible to clearly and quickly understand the main points covered in each work through the five categories adopted.

It is possible to visualise specific patterns of the prediction models if we evaluate the most common outputs in this systematic review, such as compressive strength, tensile strength, elastic modulus, and flexural strength. For example, regarding the compressive strength, although there is an additional feature related to the curing time, it is used in the same models. The most used models are artificial neural networks, random forests, support vector machines, and some metaheuristic methods. Considering the tensile strength, the elastic modulus, and flexural strength as the output, most works employ an artificial neural network model to predict these properties.

Another possible category not included in the classification framework is the input variables due to their dependency on the output property and the type of concrete. For instance, considering the compressive strength, the model could consider as input variables cement, blast-furnace slag, fly ash, water, superplasticizer, coarse aggregate, and fine aggregate [20]. While Zhang et al. [60] also considered the maximum size of coarse aggregate and curing age. Moon et al. [61] used different input variables: cement, coarse aggregate, sand, water, fineness modulus, and 3-day concrete strength. If we consider another property, such as thermal conductivity or permeability coefficient, the input variables would be quite distinct from the ones presented above. For the first property, the input can be, and for the second one, the water-to-cement ratio, aggregate-to-cement ratio, and aggregate size [50]. Due to this great diversity, this work designed the framework with only five categories and did not include the input variables.

Besides that, it is possible to broadly examine the evolution of research in concrete mix design. Table 6 summarises each of the documents selected in this review and briefly details the five categories of this framework.

Table 6. Description of the papers examined

N	Year	Description	Ref
1	2021	Developed a prior information-based NN model to obtain a 28-day concrete strength prediction model of conventional concrete.	[61]
2	2021	Proposed a hybrid model based on SVR and L-SHADE to predict the plastic viscosity of conventional concrete.	[62]
3	2021	Provided a multi-objective optimisation model using BPNN and beetle antennae search algorithm to maximise compressive strength and minimise cost and embodied CO_2 of silica fume concrete.	[60]
4	2021	Applied a predictive model based on BPNN to determine the 28-day concrete strength of geopolymer concrete.	[63]
5	2021	Determined the design relationship between various concrete mix composites of rubberised concrete and their multiple mechanical properties (7-d compressive strength, 28-d compressive strength, flexural strength, tensile strength, and elastic modulus) simultaneously.	[55]
6	2021	Determined a multilayer perceptron model with the highest prediction accuracy to predict the thermal conductivity of concrete containing fly ash, slag, and recycled aggregates.	[30]

7	2021	Compared self-organising feature map (SOFM) and multilayer perceptron to predict the compressive strength of slag concrete in early ages and more extended periods, with only three parameters as input variables.	[43]
8	2021	Established a prediction model based on the extreme learning machine to simultaneously predict rubberised concrete's 28-day compressive strength, splitting tensile, and flexural strength.	[42]
9	2021	Proposed a hybrid model using radial basis function neural network (RBFNN) and firefly optimisation algorithm (FOA) to predict the compressive strength of self-compacting concrete.	[28]
10	2021	Applied two machine learning methods (BPNN and SVM) to estimate the compressive strength of recycled aggregate concrete.	[47]
11	2021	Presented an RF model to predict two properties of fly ash-based alkali-activated concrete – slump flow and compressive strength.	[51]
12	2021	FBPNN was applied to predict a multi-output response (compressive strength, rebound hammer number, and ultrasonic pulse velocity) of sustainable concrete containing various amounts of fly ash, silica fume, and blast furnace slag.	[64]
13	2021	Different technical approaches (Linear, non-linear regressions, multi-logistic regression, M5P-tree, and FNN) were used and compared to predict the compressive strength of high-volume fly ash concrete.	[13]
14	2021	It was researched to develop an optimum machine learning algorithm for predicting steel fibre-reinforced concrete's compressive and flexural strengths. The linear regressor, lasso regressor, ridge regressor, K-nearest neighbour (KNN) regressor, decision tree regressor, random forest regressor, AdaBoost regressor, gradient boost regressor, and XGBoost regressor were used for machine-learning models.	[65]
15	2021	Imperialist Competitive Algorithm (ICA) was employed to develop new formulas to predict the tensile, compressive, and flexural strengths of recycled coarse aggregate concrete based on water, cement, RCA, natural coarse aggregates, and natural fine aggregates content.	[66]
16	2021	Development of Multi-Expression Programming (MEP) model to predict the split tensile strength and modulus of elasticity of concrete containing waste foundry sand.	[57]
17	2021	Implemented a hybrid ensemble model to predict the compressive strength of concrete containing fly ash and blast furnace slag. The predicted outputs of four conventional machine learning models (FNN, Linear and Non-Linear Multivariate Adaptive Regression Splines (MARS-L and MARS-C), Gaussian Process Regression (GPR), and Minimax Probability Machine Regression (MPMR)) were combined and trained using ANN to construct the hybrid model.	[67]
18	2020	A hybrid predicted model comprising an SVR model and FA was applied to predict compressive strength and flexural strength for steel fibre reinforced concrete. The FA-SVR model was then used as the objective function for a developed multi-objective optimisation to find the optimal concrete mixture proportion.	[68]
19	2020	Compared some hybrid models (BPNN-FA. SVM-FA, RF-FA) to predict compressive strength and tensile strength, then selected the best model to introduce it in a multi-objective optimisation with compressive strength, tensile strength, cost, and CO_2 emission.	[8]
20	2020	Adopted Boosting Smooth Transition regression trees (BooST) to investigate the predictive performance of Concrete Compressive Strength for high-performance compared with other contemporary methods for higher predictive accuracy.	[14]
21	2020	An optimum design method of high-strength concrete was developed for improving crack resistance based on ANN and GA.	[24]
22	2020	A particle swarm optimisation coupled with gradient boosting regression trees model was developed to optimise the mixture design of recycled aggregate concrete for various compressive strengths.	[29]
23	2020	A four-layer multi-layer-perceptron (MLP) model was developed for forecasting the compressive strength of ultra-high-performance concrete (UHPC) for a given mixture design.	[41]
24	2020	Proposed a multi-objective optimisation method integrating BPNN, RF, and PSO to optimise concrete mixture proportions of plastic concrete and high-performance concrete	[20]
25	2020	Some machine learning techniques (FNN, GPR, SVM, DT, LR) were compared to identify the high-strength concrete mixes while minimising the embodied carbon value of that mix composition.	[48]

26	2020	Predicted permeability and compressive strength of pervious concrete by applying multi-output least squares support vector regression (MOLSSVR) coupled with beetle- antennae search (BAS) algorithm.	[50]
27	2020	Estimated the compressive strength of the concretes containing blast furnace slag using a hybrid system composed of FBPNN and a multi-objective salp swarm algorithm.	[26]
28	2020	A hybrid artificial intelligence model based on RF and beetle antennae search (BAS) was proposed to predict the compressive strength of oil palm shell concrete.	[59]
29	2020	The water cycle algorithm, soccer league competition algorithm, GA, FNN, and SVR were applied in order to predict the compressive strength. Then, the most efficient sustainable objective function was used to estimate the mixture design of sustainable concrete.	[6]
30	2020	Presented an ensemble machine learning model (RF and SVM) to predict the modulus of elasticity (MOE) of concrete formulated using recycled concrete aggregate.	[52]
31	2020	A hybrid method using LSSVM and PSO was proposed for predicting interface yield stress and the plastic viscosity of conventional concrete mixes.	[31]
32	2020	A genetic algorithm was used to find the optimal mixture design of fly ash and slag ternary blended concrete, considering the costs of CO_2 emissions and material as objective functions.	[69]
33	2020	The AdaBoost algorithm is adopted to predict the compressive strength of concrete containing fly ash and slag.	[54]
34	2020	Estimated and compared the compressive strength of plastic concrete using four computational intelligence, including SVM, GMDH, and multi-gene genetic programming (MGGP).	[70]
35	2020	The strength characteristics of geopolymer self-compacting concrete with GEP were modelled to perform a single output response and the artificial neural networks FBPNN to predict the properties simultaneously.	[25]
36	2020	Four analytical models were built based on FBPNN, to predict the 1-day, 7-day, and 28-day compressive strengths and slump flow of ultra-high-performance concrete.	[71]
37	2020	Investigated an ensemble approach composed of MARS and GBM to predict high-performance concrete's concrete strength.	[72]
38	2020	ANFIS models were developed to predict the compressive strength of rubberised concrete.	[73]
39	2019	A slag concrete with low CO_2 emissions was designed using GA.	[56]
40	2019	A concrete mix design for low-CO_2 fly ash concrete considering climate change, carbonation, and CO_2 uptake was presented.	[74]
41	2019	It was investigated the efficiency of ALO for fine-tuning an MLP neural network in the field of concrete slump prediction and compared it with other benchmark models (BBO and GOA).	[22]
42	2019	Used a GA to search for the optimal mixture of high-strength concrete, considering the sum of the material cost and the cost of CO2 emissions as the optimisation objective of the GA.	[75]
43	2019	Predicted the relative viscosity, yield stress, and slump flow of self-consolidating concrete employing an FBPNN to predict the rheological properties under hot weather conditions and prolonged mixing.	[76]
44	2019	Linear regression and random forest were applied to predict the coefficient of thermal expansion of conventional concrete.	[77]
45	2019	Three soft computing methods of FBPNN, MARS, and M5Tree were employed to predict plastic concrete's slump, compressive strength, and elastic modulus.	[78]
46	2019	FNN, M5P model tree and RF techniques were used to predict the mechanical performance of recycled aggregate concrete containing mineral admixtures.	[79]
47	2019	Used a mathematical approach, namely Grey System Theory (GST), to examine the parametric sensitivity of the mechanical properties of recycled aggregate concrete. Then, a Multiple Nonlinear Regression (MNR) and BPNN simulated the mechanical properties.	[80]
48	2019	Four data mining models (Bayesian Ridge model, Gaussian Process, Regression Tree, and Gradient Boosting Regression Tree) were compared to predict hydraulic concrete's mechanical properties.	[81]
49	2019	A hybrid system composed of three neural networks and GA was proposed to predict the transport tortuosity of the pore system of fly ash concrete.	[82]

50	2019	A method based on the inverse heat transfer problem through a direct search optimization algorithm was developed to predict the heat of cement hydration, thermal conductivity, and specific heat.	[27]
51	2019	A multi-island GA method was combined with an adaptive surrogate modelling approach to achieve an optimal concrete mixture of fly ash and phosphorous slag concrete by optimising CO2 emissions and cost.	[83]
52	2019	Examination of an FNN model combined with an artificial bee colony (ABC) algorithm in predicting chloride penetration in self-consolidating concretes.	[23]
53	2019	Applied FNN to predict the compressive strength of waste rubber concrete.	[84]
54	2019	Applied MLP to predict the carbonation of fly-ash concrete, taking into account the most influential parameters, including mixture proportions and exposure conditions.	[40]
55	2019	Used BBP to predict the compressive strength of silica fume concrete, while constrained biogeography-based optimization (CBBO) was used to estimate its optimal mix design minimising cost.	[85]
56	2019	An MLP model was applied to investigate the modulus of elasticity of recycled aggregate concrete.	[86]
57	2019	Compared an MLP technique using four different training algorithms in steel fiber reinforced concrete to predict water absorption, compressive strength, flexural strength, split tensile strength, and slump.	[87]
58	2018	MARS, M5P Tree, LS-SVM, MLP, and Multiple Linear Regression (MLR) were employed to compare the prediction of compressive strength and ultrasonic pulse velocity of fiber reinforced concrete.	[88]
59	2018	A multi-objective Grey Wolves Optimisation (GWO) was applied to find a simple FNN model with acceptable error to predict the compressive strength of silica fume concrete.	[89]
60	2018	BPNN was used to analyse the significance of each aggregate characteristic and determine the best combinations of factors that would affect the compressive strength and elastic modulus of recycled aggregate concrete.	[90]
61	2018	An MLP model was designed to investigate the effect of concrete mix compositions, weathering effect, and exposure time on carbonation depth in concrete.	[91]
62	2018	Non-linear and mixed regression analyses were applied to model the compressive strength of limestone and fly ash concrete.	[15]
63	2018	Sequential quadratic programming was employed to solve the optimal mix design problem of self-compacting concrete in order to minimise the concrete cost.	[16]
64	2018	Mathematical models were developed to perform optimization of GHG emissions for four groups of concrete containing fly ash, ground granulated blast furnace slag, natural pozzolans, and limestone.	[18]
65	2018	BPNN, fuzzy TSK, SVR, and RBFNN compared the elastic modulus of recycled aggregate concrete.	[92]
66	2018	BPNN was employed to establish the prediction models of compressive strength and the slump of the pozzolanic concrete.	[93]
67	2018	Used FBPNN and SVM to estimate the compressive strength and flexural strength of carbon-fiber-reinforced lightweight concrete exposed to high temperatures.	[94]
68	2018	Applied gene expression models to predict slump flow and compressive strength of normal weight concrete containing granule blast furnace slag.	[95]
69	2018	A system based on an FNN model integrated with a modified firefly algorithm (MFA) was developed to predict high-performance concrete's compressive and tensile strength.	[96]
70	2018	Predicted the compressive strength of recycled concrete using a Convolutional Neural Network (CNN).	[97]
71	2018	A Multivariate Adaptive Regression Spline (MARS) model was designed to predict the compressive strength of fly ash-based geopolymer concrete.	[98]
72	2018	The recycled aggregate concrete's compressive strength was predicted by using FBPNN.	[99]
73	2017	Two BPNN were developed to predict the ingredients of self-compacting concrete, one to foresee all of them in one step and the other to predict one component per step.	[100]
74	2017	A GA was applied to optimise the connection weights for each MLP model neuron to predict the compressive strength of pozzolanic concrete.	[101]

75	2017	Three different models of MLR, BPNN, and ANFIS were compared in order to predict the compressive strength of concrete.	[102]
76	2017	Gene expression programming (GEP) was employed to predict the compressive strength, elastic modulus, flexural strength, and splitting tensile strength of recycled aggregate concrete.	[103]
77	2016	Proposed a model for predicting compressive strength and global warming potential (GWP) for concretes containing fly ash and blast furnace slag.	[17]
78	2016	Integrated PSO and MLP to predict compressive strength, splitting tensile strength, fracture energy, and flexural strength of self-compacting concrete.	[7]
79	2016	An SVM model was proposed for the autogenous shrinkage prediction of concrete mixtures.	[49]
80	2016	A numerical procedure consisting of a blended hydration model and a carbonation reaction model was presented to evaluate slag concrete's compressive strength and carbonation depth.	[104]
81	2016	A new meta-model was developed to calculate the carbonation front depth of reinforced concrete based on the analytic solution of Fick's first law.	[105]
82	2015	Applied FBPNN to predict the concrete mix composition for steel fiber-reinforced concrete.	[106]
83	2015	Employed a M5' model tree algorithm to predict the elastic modulus of recycled aggregate concrete.	[107]
84	2015	A prediction model through FBPNN was developed to predict concrete shrinkage containing silica fume and fly ash.	[108]
85	2014	Evolutionary Support Vector Machine Inference Model (ESIM) predicted the compressive strength of high-performance concrete and integrated ESIM and GA to optimise concrete mixtures.	[109]
86	2014	A numerical technique was developed to obtain optimum concrete mix proportions through a genetic algorithm (GA) for reinforced concrete structures under carbonation.	[110]
87	2014	Individual and ensemble learning classifiers with MLP, SVM, CART, and LR were constructed to predict the compressive strength of high-performance concrete.	[111]
88	2013	A five-layer fuzzy inference system (FIS) was designed to estimate the quantities of concrete ingredients based on the American Concrete Institute (ACI) method of concrete mixture design.	[112]
89	2013	The densified mixture design algorithm (DMDA) was applied in the concrete mix design of self-consolidating concrete to evaluate the effect of the paste amount on the slump, slump flow, compressive strength, and ultrasonic pulse velocity.	[113]
90	2013	Compared linear regression, non-linear regression, and FBPNN models to predict rapid chloride permeability of self-consolidating concretes based on their mixture proportions.	[114]
91	2013	Employed an evolutionary algorithm (EA) to minimise CO2 emission and cost of concrete containing fly ash and slag.	[115]
92	2013	Prediction of compressive strength of concretes containing silica fume and styrene-butadiene rubber with a mathematical model based on Abrams' law.	[116]
93	2013	The compressive strength of high performance concrete was predicted using genetic programming.	[117]
94	2013	The compressive strength of high performance concrete was predicted using individual and ensemble models composed of SVM, MLP, CART, CHAID, LR, and GENLIN.	[5]
95	2012	The HS algorithm was applied to estimate the mix proportion of high-performance concrete.	[21]
96	2012	Proposed an FBPNN model for predicting the early-age autogenous shrinkage of concrete containing silica fume and fly ash.	[118]
97	2012	Used RBFNN to predict yield stress and plastic viscosity and optimise the mix proportioning for high-performance concrete.	[119]
98	2012	Combined an MLP model with Principal Component Analysis (PCA) and developed six neural networks models for predicting slump and compressive strength in concrete with mineral additives such as blast furnace slag, fly ash, and silica fume.	[120]
99	2012	A hybrid system to predict high-performance concrete compressive strength was proposed integrating Fuzzy Logic (FL), weighted Support Vector Machines (wSVM) and fast messy genetic algorithms (fmGA).	[121]

100	2012	Utilized gene expression programming (GEP) to derive a new model for predicting the compressive strength of high performance concrete mixes.	[122]

5. Discussion

In this discussion section, the papers analysed through the classification framework developed in Section 3 answer the main research questions addressed in this review.

1. What are the main machine learning algorithms that have been used in order to come up with the concrete mix design?

After analysing the selected papers, six main groups are observed when it comes to machine learning techniques applied to concrete mix design: artificial neural networks, decision trees, support vector machines, linear regression, non-linear regression, and logistic regression. A predominance of the first three groups can be observed in the papers reviewed. In most documents, artificial neural networks have been used in several algorithms (feed-forward neural network, backpropagation neural network, multilayer perceptron) for property predictions and even used as objective functions for optimisation problems. However, despite the widespread use of this technique for predicting compressive strength at different concrete ages, it is also applied to predict other properties such as tensile strength, elastic modulus, slump, and carbonation depth.

Support Vector Machines (SVM) and Decisions Trees (DT) have also gained prominence in predicting concrete mixture properties, especially support vector regression, random forest, and M5P model trees. However, some studies reveal that in multi-output prediction, a model can have higher accuracy than others, depending on the property to be predicted. For example, the backpropagation neural network algorithm can achieve superior performance in predicting the compressive strength of concrete. At the same time, RF can provide greater accuracy in predicting properties such as slump [20]. Some comparative studies between the three types of machine learning show that ANN performed considerably better in predicting mechanical properties, such as compressive strength, followed by RF and M5P model tree [79].

Although ANN, SVM, and DT present excellent results in predicting properties, comparative studies between machine learning techniques indicate the superiority of other less explored ones. For example, the Multivariate Adaptive Regression Splines (MARS) technique was compared to the multilayer perceptron and least square support vector machine and performed better compared to these other techniques in predicting compressive strength and ultrasonic pulse velocity [88]. However, notwithstanding the promising results, this technique appears in only five studies.

Besides that, some studies focus on building a computational ensemble model consisting of a set of machine learning algorithms. Some papers compare the individual machine learning techniques with this ensemble approach, revealing a higher accuracy when employing multiple techniques. Thus, new models can be built by integrating different algorithms, allowing vast possibilities to find more accurate models.

2. What are the mathematical optimisation algorithms that have been used in order to help concrete mix design?

As mentioned in the preceding section, mathematical models are often used in concrete mix design to obtain the best concrete composition that meets some objective functions. According to each study, these functions are specified and aim at minimising or maximising some characteristics of the concrete. Different properties are used as objective functions; compressive strength and cost are the most evaluated.

Among the algorithms used in the 2012-2021 period, the Genetic Algorithm (GA) is the optimisation algorithm most widely adopted since the beginning of the period evaluated. However, all the articles selected for this review employed single-objective optimisations to find the best composition varying the objective functions. For instance, Kwon et al. [110] obtained optimum concrete mix proportions through GA using the diffusion coefficient of CO_2 as an objective function for reinforced concrete structures.

Other algorithms also appear with a particular frequency in concrete proportioning studies. For example, the Firefly Algorithm (FA) has been frequently used in multiobjective optimisations with compressive strength, cost, and CO2 emissions as objective functions [8]. Another example is Gene Expression Programming (GEP) is applied in single-objective optimization minimising cost [75] and multiobjective optimisation, evaluating compressive strength and slump [96]. Particle Swarm Optimisation (PSO) is also used in both types of optimisation, which is analysed the cost for the single case [29] and compressive strength, cost, and slump for the multiobjective analysis [20].

Another application of mathematical models is optimizing other techniques' hyperparameters and increasing the model's accuracy [82]. These four algorithms also appear frequently in optimising hyperparameters; however, GA is the most relevant in this algorithm optimisation application.

3. How is the literature linking both aspects, mathematical optimisation and machine learning, in order to design a concrete mix?

Based on the reviewed papers, a classification framework was proposed, noting the existence of two well-defined lines of research for the concrete mix design project.

The first line of research is linked to the optimisation process of concrete mixtures. Through mathematical models, objective functions that must be maximised or minimised are represented to obtain the best composition according to the project objective. The studies analysed showed mathematical models optimising the concrete mix design.

The second line is related to the prediction of concrete properties. In this line, most of the researchers found in the literature are concentrated, and about 83% of the works in this area use machine learning to predict concrete mixtures' characteristics.

Many studies have also adopted hybrid models in optimisation and prediction research lines. The former uses machine learning techniques to represent some properties as an objective function. For example, the application of BPNN represents the compressive strength and slump as objective functions. Then, a particle swarm algorithm is applied to perform a multi-objective optimisation [20]. The latter uses mathematical models to improve the machine learning techniques' parameters and thus optimise the model [50].

6. Conclusion

Operational research techniques for concrete mix design have been used to improve the results achieved by the conventional concrete mix proportioning methods based on many experimental tests. Mathematical programming and machine learning methods applied in concrete mix design have turned attractive for several reasons. To begin with, the different composition combinations due to the diversity of raw materials make the concrete design very challenging to perform all laboratory experiments. In addition, laboratory methods are massive, time-consuming, and do not guarantee an ideal concrete mixing solution. Furthermore, mathematical models and machine learning techniques innovate the concrete design method, ensuring that concrete prediction and optimisation models are carried out quickly and accurately.

This review gathered the most relevant papers on concrete mix design from 2012 to 2021. And then, it was proposed a classification framework that could group the works into different categories. Within the analysis of 100 studies, the framework was broken down into five main categories: type of research line, type of concrete, concrete properties, type of output or objective, and type of computer aid system adopted.

In order to focus on the most important and relevant papers, this systematic review was subjected to some constraints to select them. The main limitations of this work were databases and the boundary definition, which took into account only papers published from a specific timeline, written in English, and did not include any conference paper. Another limitation was regarding the design of the classificational framework, which could have included a variety of input variables. However, the inclusion of this category would be challenging to group in a single category since the input variables depend on the output to be predicted. For those reasons, the classification framework focused on these five categories. According to the reviewed papers, two research lines were spotted. Around 83% are mainly aimed at predicting the properties of concrete, where 70% developed a prediction model with a single output response, and 13% presented a model with multi-property prediction. The remaining documents, around 17%, look for optimising the concrete mixture, with 10% showing a single objective optimisation and 7% minimising or maximising multi-objective functions.

Among the types of concrete investigated, there is a growing trend of studies in both research lines with concrete containing recycled aggregate and supplementary cementitious material such as fly ash, slag, and silica fume. It can be justified by encouraging the reuse of wasted materials and developing more sustainable materials. Regarding the properties predicted or used as objective functions, mechanical properties are the most evaluated, especially compressive strength. Another objective function often used is cost, which is always considered in optimisation studies. These two properties are presented in 60% and 37% of the examined research lines.

Furthermore, the framework was broken down into three classes to evaluate the computer aid system: mathematical models, machine learning, and hybrid system. Most works applied machine learning to predict some concrete property when it comes to predicting research. About 45% of all documents use at least one machine learning technique. Mathematical models and hybrid systems are used in both research lines, the former accounting for 26% and the latter for 29% of the articles. Although several techniques have been used in both areas, a relevant domain of artificial neural networks, support vector machines, decision trees, and genetic algorithms was noted.

Although this subject presents many relevant works for predicting properties and optimising concrete mixtures, few focus on multi-objective optimisation. Since several objective functions can show a trade-off relationship, simultaneously optimising them is pretty challenging. Therefore, in order to simplify the mathematical optimisation model, most studies use only a single objective function. However, multi-objective optimisation can adequately represent the concrete mix problem, making it more reliable. Accordingly, future works could focus on multi-objective optimisations, analysing different properties so that the results encompass several concrete characteristics, for example, compressive strength, cost, CO2 emission, and thermal conductivity. Thus, it can achieve a more resistant, less costly, and more sustainable concrete material.

References

[1] M. A. DeRousseau, J. R. Kasprzyk, and W. V. Srubar, “Computational design optimization of concrete mixtures: A review,” Cem Concr Res, vol. 109, no. April, pp. 42–53, 2018, doi: 10.1016/j.cemconres.2018.04.007.

[2] A. M. Neville and J. J. Brooks, Concrete technology, 2nd editio. 2nd editio.Harlow: Pearson Education Limited, 2010.

[3] M. L. Gambhir, Concrete Technology, 5th Editio. 5th Editio.New Delhi: McGraw Hill Education, 2013.

[4] M. A. DeRousseau, “Concrete mixture design using machine learning, life cycle assessment, and multi-objective optimization,” University of Colorado Boulder, 2020.

[5] J. Chou and A.-Duc. Pham, “Enhanced artificial intelligence for ensemble approach to predicting high performance concrete compressive strength,” Constr Build Mater, vol. 49, pp. 554–563, 2013.

[6] H. Naseri, H. Jahanbakhsh, P. Hosseini, and F. Moghadas Nejad, “Designing sustainable concrete mixture by developing a new machine learning technique,” J Clean Prod, vol. 258, p. 120578, 2020, doi: 10.1016/j.jclepro.2020.120578.

[7] H. Mashhadban, S. S. Kutanaei, and M. A. Sayarinejad, “Prediction and modeling of mechanical properties in fiber reinforced self-compacting concrete using particle swarm optimization algorithm and artificial neural network,” Constr Build Mater, vol. 119, pp. 277–287, 2016, doi: 10.1016/j.conbuildmat.2016.05.034.

[8] J. Zhang, Y. Huang, F. Aslani, G. Ma, and B. Nener, “A hybrid intelligent system for designing optimal proportions of recycled aggregate concrete,” J Clean Prod, vol. 273, p. 122922, 2020, doi: 10.1016/j.jclepro.2020.122922.

[9] H.-S. Lee, S.-M. Lim, and X.-Y. Wang, “Optimal Mixture Design of Low-CO2 High-Volume Slag Concrete Considering Climate Change and CO2 Uptake,” Int J Concr Struct Mater, vol. 13, no. 1, 2019, doi: 10.1186/s40069-019-0359-7.

[10] W. ben Chaabene, M. Flah, and M. L. Nehdi, “Machine learning prediction of mechanical properties of concrete: Critical review,” Constr Build Mater, vol. 260, p. 119889, Nov. 2020, doi: 10.1016/j.conbuildmat.2020.119889.

[11] I. Nunez, A. Marani, M. Flah, and M. L. Nehdi, “Estimating compressive strength of modern concrete mixtures using computational intelligence: A systematic review,” Constr Build Mater, vol. 310, p. 125279, Dec. 2021, doi: 10.1016/j.conbuildmat.2021.125279.

[12] Y. Song, X. Wang, H. Li, Y. He, Z. Zhang, and J. Huang, “Mixture Optimization of Cementitious Materials Using Machine Learning and Metaheuristic Algorithms: State of the Art and Future Prospects,” Materials, vol. 15, no. 21, p. 7830, Nov. 2022, doi: 10.3390/ma15217830.

[13] A. Mohammed, S. Rafiq, P. Sihag, R. Kurda, and W. Mahmood, “Soft computing techniques: Systematic multiscale models to predict the compressive strength of HVFA concrete based on mix proportions and curing times,” Journal of Building Engineering, vol. 33, no. September 2020, p. 101851, 2021, doi: 10.1016/j.jobe.2020.101851.

[14] U. Anyaoha, A. Zaji, and Z. Liu, “Soft computing in estimating the compressive strength for high-performance concrete via concrete composition appraisal,” Constr Build Mater, vol. 257, p. 119472, 2020, doi: 10.1016/j.conbuildmat.2020.119472.

[15] R. Jin, Q. Chen, and A. B. O. Soboyejo, “Non-linear and mixed regression models in predicting sustainable concrete strength,” Constr Build Mater, vol. 170, pp. 142–152, 2018, doi: 10.1016/j.conbuildmat.2018.03.063.

[16] A. Habibi and J. Ghomashi, “Development of an optimum mix design method for self-compacting concrete based on experimental results,” Constr Build Mater, vol. 168, pp. 113–123, 2018, doi: 10.1016/j.conbuildmat.2018.02.113.

[17] S. A. Miller, P. J. M. Monteiro, C. P. Ostertag, and A. Horvath, “Concrete mixture proportioning for desired strength and reduced global warming potential,” Constr Build Mater, vol. 128, pp. 410–421, 2016, doi: 10.1016/j.conbuildmat.2016.10.081.

[18] C. Fan and S. A. Miller, “Reducing greenhouse gas emissions for prescribed concrete compressive strength,” Constr Build Mater, vol. 167, pp. 918–928, 2018, doi: 10.1016/j.conbuildmat.2018.02.092.

[19] K.-Q. Li, D.-Q. Li, P.-T. Li, and Y. Liu, "Meso-mechanical investigations on the overall elastic properties of multi-phase construction materials using finite element method," Constr Build Mater, vol. 228, p. 116727, 2019, doi: 10.1016/j.conbuildmat.2019.116727.

[20] J. Zhang, Y. Huang, Y. Wang, and G. Ma, "Multi-objective optimization of concrete mixture proportions using machine learning and metaheuristic algorithms," Constr Build Mater, vol. 253, p. 119208, 2020, doi: 10.1016/j.conbuildmat.2020.119208.

[21] J. H. Lee, Y. S. Yoon, and J. H. Kim, "A new heuristic algorithm for mix design of high-performance concrete," KSCE Journal of Civil Engineering, vol. 16, no. 6, pp. 974–979, 2012, doi: 10.1007/s12205-012-1011-0.

[22] H. Moayedi, B. Kalantar, L. K. Foong, D. T. Bui, and A. Motevalli, "Application of three metaheuristic techniques in simulation of concrete slump," Applied Sciences (Switzerland), vol. 9, no. 20, 2019, doi: 10.3390/app9204340.

[23] M. Najimi, N. Ghafoori, and M. Nikoo, "Modeling chloride penetration in self-consolidating concrete using artificial neural network combined with artificial bee colony algorithm," Journal of Building Engineering, vol. 22, no. July 2018, pp. 216–226, 2019, doi: 10.1016/j.jobe.2018.12.013.

[24] L. D. of H.-S. C. M. P. for C. R. U. A. N. N. and G. A. Yue, L. Hongwen, L. Yinuo, and J. Caiyun, "Optimum Design of High-Strength Concrete Mix Proportion for Crack Resistance Using Artificial Neural Networks and Genetic Algorithm," Front Mater, vol. 7, no. October, pp. 1–12, 2020, doi: 10.3389/fmats.2020.590661.

[25] P. O. Awoyera, M. S. Kirgiz, A. Viloria, and D. Ovallos-Gazabon, "Estimating strength properties of geopolymer self-compacting concrete using machine learning techniques," Journal of Materials Research and Technology, vol. 9, no. 4, pp. 9016–9028, 2020, doi: 10.1016/j.jmrt.2020.06.008.

[26] A. Kandiri, E. Mohammadi Golafshani, and A. Behnood, "Estimation of the compressive strength of concretes containing ground granulated blast furnace slag using hybridized multi-objective ANN and salp swarm algorithm," Constr Build Mater, vol. 248, p. 118676, 2020, doi: 10.1016/j.conbuildmat.2020.118676.

[27] G. Knor, R. Jaskulski, M. A. Glinicki, and J. Holnicki-Szulc, "Numerical identification of the thermal properties of early age concrete using inverse heat transfer problem," Heat and Mass Transfer/Waerme- und Stoffuebertragung, vol. 55, no. 4, pp. 1215–1227, 2019, doi: 10.1007/s00231-018-2504-2.

[28] G. Pazouki, E. M. Golafshani, and A. Behnood, "Predicting the compressive strength of self-compacting concrete containing Class F fly ash using metaheuristic radial basis function neural network," Structural Concrete, no. February, 2021, doi: 10.1002/suco.202000047.

[29] I. Nunez, A. Marani, and M. L. Nehdi, “Mixture optimization of recycled aggregate concrete using hybrid machine learning model,” Materials, vol. 13, no. 19, pp. 1–24, 2020, doi: 10.3390/ma13194331.

[30] Y. Sargam, K. Wang, and I. H. Cho, “Machine learning based prediction model for thermal conductivity of concrete,” Journal of Building Engineering, vol. 34, no. November 2020, p. 101956, 2021, doi: 10.1016/j.jobe.2020.101956.

[31] T.-D. Nguyen, T.-H. Tran, and N.-D. Hoang, “Prediction of interface yield stress and plastic viscosity of fresh concrete using a hybrid machine learning approach,” Advanced Engineering Informatics, vol. 44, no. February, p. 101057, 2020, doi: 10.1016/j.aei.2020.101057.

[32] Y. Sharifi and M. Hosseinpour, “Compressive strength assessment of concrete containing metakaolin using ANN,” Journal of Rehabilitation in Civil Engineering, vol. 8, no. 4, pp. 15–27, 2020, doi: 10.22075/JRCE.2020.19043.1358.

[33] I.-C. Yeh, “Modeling Concrete Strength with Augment-Neuron Networks,” Journal of Materials in Civil Engineering, vol. 10, no. 4, pp. 263–268, Nov. 1998, doi: 10.1061/(ASCE)0899-1561(1998)10:4(263).

[34] I.-C. Yeh, “Modeling of strength of high-performance concrete using artificial neural networks,” Cem Concr Res, vol. 28, no. 12, pp. 1797–1808, Dec. 1998, doi: 10.1016/S0008-8846(98)00165-3.

[35] I.-C. Yeh, “Analysis of Strength of Concrete Using Design of Experiments and Neural Networks,” Journal of Materials in Civil Engineering, vol. 18, no. 4, pp. 597–604, Aug. 2006, doi: 10.1061/(ASCE)0899-1561(2006)18:4(597).

[36] I.-C. Yeh, “Modeling slump flow of concrete using second-order regressions and artificial neural networks,” Cem Concr Compos, vol. 29, no. 6, pp. 474–480, Jul. 2007, doi: 10.1016/j.cemconcomp.2007.02.001.

[37] H. Yaprak, A. Karaci, and I. Demir, “Prediction of the effect of varying cure conditions and w/c ratio on the compressive strength of concrete using artificial neural networks,” Neural Comput Appl, vol. 22, no. 1, pp. 133–141, 2013, doi: 10.1007/s00521-011-0671-x.

[38] A. A. Shahmansouri, M. Yazdani, S. Ghanbari, H. Akbarzadeh Bengar, A. Jafari, and H. Farrokh Ghatte, “Artificial neural network model to predict the compressive strength of eco-friendly geopolymer concrete incorporating silica fume and natural zeolite,” J Clean Prod, vol. 279, p. 123697, 2021, doi: 10.1016/j.jclepro.2020.123697.

[39] A. Verm and I. Verma, “Use of artificial neural network in design of fly ash blended cement concrete mixes,” International Journal of Recent Technology and Engineering, vol. 8, no. 3, pp. 4222–4233, 2019, doi: 10.35940/ijrte.C5146.098319.

[40] Y. Kellouche, B. Boukhatem, M. Ghrici, and A. Tagnit-Hamou, "Exploring the major factors affecting fly-ash concrete carbonation using artificial neural network," Neural Comput Appl, vol. 31, no. s2, pp. 969–988, 2019, doi: 10.1007/s00521-017-3052-2.

[41] J. Abellán-García, "Four-layer perceptron approach for strength prediction of UHPC," Constr Build Mater, vol. 256, 2020, doi: 10.1016/j.conbuildmat.2020.119465.

[42] J. Zhang, J. Xu, C. Liu, and J. Zheng, "Prediction of Rubber Fiber Concrete Strength Using Extreme Learning Machine," Front Mater, vol. 7, no. January, pp. 1–12, 2021, doi: 10.3389/fmats.2020.582635.

[43] S. Czarnecki, M. Shariq, M. Nikoo, and Ł. Sadowski, "An intelligent model for the prediction of the compressive strength of cementitious composites with ground granulated blast furnace slag based on ultrasonic pulse velocity measurements," Measurement, vol. 172, no. 2021, pp. 1–10, 2021, doi: 10.1016/j.measurement.2020.108951.

[44] T. Al-Mughanam, T. H. H. Aldhyani, B. Alsubari, and M. Al-Yaari, "Modeling of compressive strength of sustainable self-compacting concrete incorporating treated palm oil fuel ash using artificial neural network," Sustainability (Switzerland), vol. 12, no. 22, pp. 1–13, 2020, doi: 10.3390/su12229322.

[45] J. A. K. Suykens and J. Vandewalle, "Least Squares Support Vector Machine Classifier," Neural Process Lett, vol. 9, pp. 293–300, 1999.

[46] A. Géron, Hands-on Machine Learning with Scikit-Learn, Keras, and TensorFlow: Concepts, Tools, and Techniques to Build Intelligent Systems, 2nd ed., . 2nd ed., O'Reilly Media, Inc., 2019.

[47] S. R. Salimbahrami and R. Shakeri, "Experimental investigation and comparative machine-learning prediction of compressive strength of recycled aggregate concrete," Soft comput, vol. 25, no. 2, pp. 919–932, 2021, doi: 10.1007/s00500-021-05571-1.

[48] P. S. M. Thilakarathna, S. Seo, K. S. K. Baduge, H. Lee, P. Mendis, and G. Foliente, "Embodied carbon analysis and benchmarking emissions of high and ultra-high strength concrete using machine learning algorithms," J Clean Prod, vol. 262, no. February 2019, p. 121281, 2020, doi: 10.1016/j.jclepro.2020.121281.

[49] J. Liu, K. Z. Yan, X. Zhao, and Y. Hu, "Prediction of autogenous shrinkage of concretes by support vector machine," International Journal of Pavement Research and Technology, vol. 9, no. 3, pp. 169–177, 2016, doi: 10.1016/j.ijprt.2016.06.003.

[50] J. Zhang, Y. Huang, G. Ma, J. Sun, and B. Nener, "A metaheuristic-optimized multi-output model for predicting multiple properties of pervious concrete," Constr Build Mater, vol. 249, p. 118803, 2020, doi: 10.1016/j.conbuildmat.2020.118803.

[51] E. Gomaa, T. Han, M. ElGawady, J. Huang, and A. Kumar, "Machine learning to predict properties of fresh and hardened alkali-activated concrete," Cem Concr Compos, vol. 115, no. November 2020, p. 103863, 2021, doi: 10.1016/j.cemconcomp.2020.103863.

[52] T. Han, A. Siddique, K. Khayat, J. Huang, and A. Kumar, "An ensemble machine learning approach for prediction and optimization of modulus of elasticity of recycled aggregate concrete," Constr Build Mater, vol. 244, p. 118271, 2020, doi: 10.1016/j.conbuildmat.2020.118271.

[53] A. Behnood and E. M. Golafshani, "Machine learning study of the mechanical properties of concretes containing waste foundry sand," Constr Build Mater, vol. 243, p. 118152, 2020, doi: 10.1016/j.conbuildmat.2020.118152.

[54] D. C. Feng et al., "Machine learning-based compressive strength prediction for concrete: An adaptive boosting approach," Constr Build Mater, vol. 230, p. 117000, 2020, doi: 10.1016/j.conbuildmat.2019.117000.

[55] X. Huang, J. Zhang, J. Sresakoolchai, and S. Kaewunruen, "Machine learning aided design and prediction of environmentally friendly rubberised concrete," Sustainability (Switzerland), vol. 13, no. 4, pp. 1–27, 2021, doi: 10.3390/su13041691.

[56] X.-Y. Wang and H.-S. Lee, "Effect of global warming on the proportional design of low CO2 slag-blended concrete," Constr Build Mater, vol. 225, pp. 1140–1151, 2019, doi: 10.1016/j.conbuildmat.2019.07.134.

[57] M. F. Iqbal et al., "Sustainable utilization of foundry waste: Forecasting mechanical properties of foundry sand based concrete using multi-expression programming," Science of the Total Environment, vol. 780, p. 146524, 2021, doi: 10.1016/j.scitotenv.2021.146524.

[58] H.-W. Chung and M. Tia, "Effects of Minimum Cementitious Paste Volume and Blended Aggregates on Compressive Strength and Surface Resistivity of Portland Limestone Cement Concrete," Journal of Materials in Civil Engineering, vol. 33, no. 5, p. 04021080, 2021, doi: 10.1061/(asce)mt.1943-5533.0003706.

[59] J. Zhang, D. Li, and Y. Wang, "Predicting uniaxial compressive strength of oil palm shell concrete using a hybrid artificial intelligence model," Journal of Building Engineering, vol. 30, no. February, p. 101282, 2020, doi: 10.1016/j.jobe.2020.101282.

[60] J. Zhang, Y. Huang, G. Ma, and B. Nener, "Mixture optimization for environmental, economical and mechanical objectives in silica fume concrete: A novel frame-work based on machine learning and a new meta-heuristic algorithm," Resour Conserv Recycl, vol. 167, no. April 2020, 2021, doi: 10.1016/j.resconrec.2021.105395.

[61] S. Moon and A. Munira Chowdhury, "Utilization of Prior Information in Neural Network Training for Improving 28-Day Concrete Strength Prediction," J Constr Eng Manag, vol. 147, no. 5, p. 04021028, 2021, doi: 10.1061/(asce)co.1943-7862.0002047.

[62] T. D. Nguyen, T. H. Tran, H. Nguyen, and H. Nhat-Duc, "A success history-based adaptive differential evolution optimized support vector regression for estimating plastic viscosity of fresh concrete," Eng Comput, vol. 37, no. 2, pp. 1485–1498, 2021, doi: 10.1007/s00366-019-00899-7.

[63] C. Gunasekara, P. Atzarakis, W. Lokuge, D. W. Law, and S. Setunge, "Novel analytical method for mix design and performance prediction of high calcium fly ash geopolymer concrete," Polymers (Basel), vol. 13, no. 6, pp. 1–21, 2021, doi: 10.3390/polym13060900.

[64] A. M. Tahwia, A. Heniegal, M. S. Elgamal, and B. A. Tayeh, "The prediction of compressive strength and non-destructive tests of sustainable concrete by using artificial neural networks," Computers and Concrete, vol. 27, no. 1, pp. 21–28, 2021, doi: 10.12989/cac.2021.27.1.021.

[65] M. C. Kang, D. Y. Yoo, and R. Gupta, "Machine learning-based prediction for compressive and flexural strengths of steel fiber-reinforced concrete," Constr Build Mater, vol. 266, p. 121117, 2021, doi: 10.1016/j.conbuildmat.2020.121117.

[66] M. Rezaiee-Pajand, J. Mohebi Najm Abad, A. Karimipour, and A. Rezaiee-Pajand, "Propose new implement models to determine the compressive, tensile and flexural strengths of recycled coarse aggregate concrete via imperialist competitive algorithm," Journal of Building Engineering, vol. 40, no. February, p. 102337, 2021, doi: 10.1016/j.jobe.2021.102337.

[67] P. G. Asteris, A. D. Skentou, A. Bardhan, P. Samui, and K. Pilakoutas, "Predicting concrete compressive strength using hybrid ensembling of surrogate machine learning models," Cem Concr Res, vol. 145, no. October 2020, p. 106449, 2021, doi: 10.1016/j.cemconres.2021.106449.

[68] Y. Huang, J. Zhang, F. Tze Ann, and G. Ma, "Intelligent mixture design of steel fibre reinforced concrete using a support vector regression and firefly algorithm based multi-objective optimization model," Constr Build Mater, vol. 260, p. 120457, 2020, doi: 10.1016/j.conbuildmat.2020.120457.

[69] X.-Y. Wang, "Impacts of climate change on optimal mixture design of blended concrete considering carbonation and chloride ingress," Frontiers of Structural and Civil Engineering, vol. 14, no. 2, pp. 473–486, 2020, doi: 10.1007/s11709-020-0608-5.

[70] A. T. Amlashi, P. Alidoust, A. R. Ghanizadeh, S. Khabiri, M. Pazhouhi, and M. S. Monabati, "Application of computational intelligence and statistical approaches for auto-estimating the compressive strength of plastic concrete," European Journal of

Environmental and Civil Engineering, vol. 0, no. 0, pp. 1–32, 2020, doi: 10.1080/19648189.2020.1803144.

[71] J. Abellán García, J. Fernández Gómez, and N. Torres Castellanos, "Properties prediction of environmentally friendly ultra-high-performance concrete using artificial neural networks," European Journal of Environmental and Civil Engineering, vol. 0, no. 0, pp. 1–25, 2020, doi: 10.1080/19648189.2020.1762749.

[72] M. R. Kaloop, D. Kumar, P. Samui, J. W. Hu, and D. Kim, "Compressive strength prediction of high-performance concrete using gradient tree boosting machine," Constr Build Mater, vol. 264, p. 120198, 2020, doi: 10.1016/j.conbuildmat.2020.120198.

[73] M. Jalal, Z. Grasley, N. Nassir, and H. Jalal, "Strength and dynamic elasticity modulus of rubberized concrete designed with ANFIS modeling and ultrasonic technique," Constr Build Mater, vol. 240, p. 117920, 2020, doi: 10.1016/j.conbuildmat.2019.117920.

[74] X.-Y. Wang, "Simulation for optimal mixture design of low-CO2 high-volume fly ash concrete considering climate change and CO2 uptake," Cem Concr Compos, vol. 104, no. September, p. 103408, 2019, doi: 10.1016/j.cemconcomp.2019.103408.

[75] X.-Y. Wang, "Effect of carbon pricing on optimal mix design of sustainable high-strength concrete," Sustainability (Switzerland), vol. 11, no. 20, 2019, doi: 10.3390/su11205827.

[76] M. I. Al-Khatib and S. Al-Martini, "Predicting the rheology of self-consolidating concrete under hot weather," Proceedings of Institution of Civil Engineers: Construction Materials, vol. 172, no. 5, pp. 235–245, 2019, doi: 10.1680/jcoma.16.00055.

[77] V. Nilsen, L. T. Pham, M. Hibbard, A. Klager, S. M. Cramer, and D. Morgan, "Prediction of concrete coefficient of thermal expansion and other properties using machine learning," Constr Build Mater, vol. 220, pp. 587–595, 2019, doi: 10.1016/j.conbuildmat.2019.05.006.

[78] A. T. Amlashi, S. M. Abdollahi, S. Goodarzi, and A. R. Ghanizadeh, "Soft computing based formulations for slump, compressive strength, and elastic modulus of bentonite plastic concrete," J Clean Prod, vol. 230, pp. 1197–1216, 2019, doi: 10.1016/j.jclepro.2019.05.168.

[79] S. Arora, B. Singh, and B. Bhardwaj, "Strength performance of recycled aggregate concretes containing mineral admixtures and their performance prediction through various modeling techniques," Journal of Building Engineering, vol. 24, no. March, p. 100741, 2019, doi: 10.1016/j.jobe.2019.100741.

[80] J. Xu, X. Zhao, Y. Yu, T. Xie, G. Yang, and J. Xue, "Parametric sensitivity analysis and modelling of mechanical properties of normal- and high-strength recycled

aggregate concrete using grey theory, multiple nonlinear regression and artificial neural networks," Constr Build Mater, vol. 211, pp. 479–491, 2019, doi: 10.1016/j.conbuildmat.2019.03.234.

[81] M. Zhang, M. Li, Y. Shen, Q. Ren, and J. Zhang, "Multiple mechanical properties prediction of hydraulic concrete in the form of combined damming by experimental data mining," Constr Build Mater, vol. 207, pp. 661–671, 2019, doi: 10.1016/j.conbuildmat.2019.02.169.

[82] B. Boukhatem, R. Rebouh, A. Zidol, M. Chekired, and A. Tagnit-Hamou, "An intelligent hybrid system for predicting the tortuosity of the pore system of fly ash concrete," Constr Build Mater, vol. 205, pp. 274–284, 2019, doi: 10.1016/j.conbuildmat.2019.02.005.

[83] X. Cen, Q. Wang, X. Shi, Y. Su, and J. Qiu, "Optimization of concrete mixture design using adaptive surrogate model," Sustainability (Switzerland), vol. 11, no. 7, 2019, doi: 10.3390/su11071991.

[84] M. Hadzima-Nyarko, E. K. Nyarko, N. Ademović, I. Miličević, and T. K. Šipoš, "Modelling the influence of waste rubber on compressive strength of concrete by artificial neural networks," Materials, vol. 12, no. 2, 2019, doi: 10.3390/ma12040561.

[85] E. M. Golafshani and A. Behnood, "Estimating the optimal mix design of silica fume concrete using biogeography-based programming," Cem Concr Compos, vol. 96, no. July 2018, pp. 95–105, 2019, doi: 10.1016/j.cemconcomp.2018.11.005.

[86] S. Sadati, L. E. Brito da Silva, D. C. Wunsch, and K. H. Khayat, "Artificial Intelligence to Investigate Modulus of Elasticity of Recycled Aggregate Concrete," ACI Mater J, vol. 116, no. 1, pp. 51–62, 2019, doi: 10.14359/51706948.

[87] T. F. Awolusi, O. L. Oke, O. O. Akinkurolere, A. O. Sojobi, and O. G. Aluko, "Performance comparison of neural network training algorithms in the modeling properties of steel fiber reinforced concrete," Heliyon, vol. 5, no. 1, p. e01115, 2019, doi: 10.1016/j.heliyon.2018.e01115.

[88] A. Ashrafian, M. J. Taheri Amiri, M. Rezaie-Balf, T. Ozbakkaloglu, and O. Lotfi-Omran, "Prediction of compressive strength and ultrasonic pulse velocity of fiber reinforced concrete incorporating nano silica using heuristic regression methods," Constr Build Mater, vol. 190, pp. 479–494, 2018, doi: 10.1016/j.conbuildmat.2018.09.047.

[89] A. Behnood and E. M. Golafshani, "Predicting the compressive strength of silica fume concrete using hybrid artificial neural network with multi-objective grey wolves," J Clean Prod, vol. 202, pp. 54–64, 2018.

[90] Z. Duan, S. Hou, C. S. Poon, J. Xiao, and Y. Liu, "Using neural networks to determine the significance of aggregate characteristics affecting the mechanical

properties of recycled aggregate concrete," Applied Sciences (Switzerland), vol. 8, no. 11, 2018, doi: 10.3390/app8112171.

[91] S. C. Paul, B. Panda, Y. Huang, A. Garg, and X. Peng, "An empirical model design for evaluation and estimation of carbonation depth in concrete," Measurement (Lond), vol. 124, no. March, pp. 205–210, 2018, doi: 10.1016/j.measurement.2018.04.033.

[92] E. M. Golafshani and A. Behnood, "Application of soft computing methods for predicting the elastic modulus of recycled aggregate concrete," J Clean Prod, vol. 176, pp. 1163–1176, 2018, doi: 10.1016/j.jclepro.2017.11.186.

[93] C. Y. Kao, C. H. Shen, J. C. Jan, and S. L. Hung, "A computer-aided approach to pozzolanic concrete mix design," Advances in Civil Engineering, vol. 2018, 2018, doi: 10.1155/2018/4398017.

[94] H. Tanyildizi, "Prediction of the strength properties of carbon fiber-reinforced lightweight concrete exposed to the high temperature using artificial neural network and support vector machine," Advances in Civil Engineering, vol. 2018, 2018, doi: 10.1155/2018/5140610.

[95] B. Şimşek, "Modeling and optimization of standard concrete containing granule blast furnace slag: A gene expression modeling based multi-response weighted non-linear programming application," International Journal of Industrial Engineering : Theory Applications and Practice, vol. 25, no. 4, pp. 490–506, 2018.

[96] D. K. Bui, T. Nguyen, J. S. Chou, H. Nguyen-Xuan, and T. D. Ngo, "A modified firefly algorithm-artificial neural network expert system for predicting compressive and tensile strength of high-performance concrete," Constr Build Mater, vol. 180, pp. 320–333, 2018, doi: 10.1016/j.conbuildmat.2018.05.201.

[97] F. Deng, Y. He, S. Zhou, Y. Yu, H. Cheng, and X. Wu, "Compressive strength prediction of recycled concrete based on deep learning," Constr Build Mater, vol. 175, pp. 562–569, 2018, doi: 10.1016/j.conbuildmat.2018.04.169.

[98] W. Lokuge, A. Wilson, C. Gunasekara, D. W. Law, and S. Setunge, "Design of fly ash geopolymer concrete mix proportions using Multivariate Adaptive Regression Spline model," Constr Build Mater, vol. 166, pp. 472–481, 2018, doi: 10.1016/j.conbuildmat.2018.01.175.

[99] H. Naderpour, A. H. Rafiean, and P. Fakharian, "Compressive strength prediction of environmentally friendly concrete using artificial neural networks," Journal of Building Engineering, vol. 16, no. October 2017, pp. 213–219, 2018, doi: 10.1016/j.jobe.2018.01.007.

[100] M. A. Yaman, M. A. Elaty, and M. Taman, "Predicting the ingredients of self compacting concrete using artificial neural network," Alexandria Engineering Journal, vol. 56, no. 4, pp. 523–532, 2017, doi: 10.1016/j.aej.2017.04.007.

[101] R. Rebouh, B. Boukhatem, M. Ghrici, and A. Tagnit-Hamou, "A practical hybrid NNGA system for predicting the compressive strength of concrete containing natural pozzolan using an evolutionary structure," Constr Build Mater, vol. 149, pp. 778–789, 2017, doi: 10.1016/j.conbuildmat.2017.05.165.

[102] F. Khademi, M. Akbari, S. M. Jamal, and M. Nikoo, "Multiple linear regression, artificial neural network, and fuzzy logic prediction of 28 days compressive strength of concrete," Frontiers of Structural and Civil Engineering, vol. 11, no. 1, pp. 90–99, 2017, doi: 10.1007/s11709-016-0363-9.

[103] A. Gholampour, A. H. Gandomi, and T. Ozbakkaloglu, "New formulations for mechanical properties of recycled aggregate concrete using gene expression programming," Constr Build Mater, vol. 130, pp. 122–145, 2017, doi: 10.1016/j.conbuildmat.2016.10.114.

[104] L. Han-Seung and X. Y. Wang, "Evaluation of compressive strength development and carbonation depth of high volume slag-blended concrete," Constr Build Mater, vol. 124, pp. 45–54, 2016, doi: 10.1016/j.conbuildmat.2016.07.070.

[105] V. L. Ta, S. Bonnet, T. Senga Kiesse, and A. Ventura, "A new meta-model to calculate carbonation front depth within concrete structures," Constr Build Mater, vol. 129, pp. 172–181, 2016, doi: 10.1016/j.conbuildmat.2016.10.103.

[106] M. Açikgenç, M. Ulaş, and K. E. Alyamaç, "Using an Artificial Neural Network to Predict Mix Compositions of Steel Fiber-Reinforced Concrete," Arab J Sci Eng, vol. 40, no. 2, pp. 407–419, 2015, doi: 10.1007/s13369-014-1549-x.

[107] A. Behnood, J. Olek, and M. A. Glinicki, "Predicting modulus elasticity of recycled aggregate concrete using M5′ model tree algorithm," Constr Build Mater, vol. 94, pp. 137–147, 2015, doi: 10.1016/j.conbuildmat.2015.06.055.

[108] K. Mermerdaş and M. M. Arbili, "Explicit formulation of drying and autogenous shrinkage of concretes with binary and ternary blends of silica fume and fly ash," Constr Build Mater, vol. 94, pp. 371–379, 2015, doi: 10.1016/j.conbuildmat.2015.07.074.

[109] M.-Y. Cheng, D. Prayogo, and Y.-W. Wu, "Novel Genetic Algorithm-Based Evolutionary Support Vector Machine for Optimizing High-Performance Concrete Mixture," Journal of Computing in Civil Engineering, vol. 28, no. 4, p. 06014003, 2014, doi: 10.1061/(asce)cp.1943-5487.0000347.

[110] S. J. Kwon, B. J. Lee, and Y. Y. Kim, "Concrete mix design for service life of rc structures under carbonation using genetic algorithm," Advances in Materials Science and Engineering, vol. 2014, pp. 6–8, 2014, doi: 10.1155/2014/653753.

[111] J. S. Chou, C. F. Tsai, A. D. Pham, and Y. H. Lu, "Machine learning in concrete strength simulations: Multi-nation data analytics," Constr Build Mater, vol. 73, pp. 771–780, 2014, doi: 10.1016/j.conbuildmat.2014.09.054.

[112] S. Y. Kute and R. S. Kale, "Five-layer fuzzy inference system to design a concrete mixture, based on ACI method," ACI Mater J, vol. 110, no. 6, pp. 629–639, 2013, doi: 10.14359/51686330.

[113] Y. Y. Chen, B. L. A. Tuan, and C. L. Hwang, "Effect of paste amount on the properties of self-consolidating concrete containing fly ash and slag," Constr Build Mater, vol. 47, pp. 340–346, 2013, doi: 10.1016/j.conbuildmat.2013.05.050.

[114] N. Ghafoori, M. Najimi, J. Sobhani, and M. A. Aqel, "Predicting rapid chloride permeability of self-consolidating concrete: A comparative study on statistical and neural network models," Constr Build Mater, vol. 44, pp. 381–390, 2013, doi: 10.1016/j.conbuildmat.2013.03.039.

[115] T. Kim, S. Tae, and S. Roh, "Assessment of the CO2 emission and cost reduction performance of a low-carbon-emission concrete mix design using an optimal mix design system," Renewable and Sustainable Energy Reviews, vol. 25, pp. 729–741, 2013, doi: 10.1016/j.rser.2013.05.013.

[116] M. Shafieyzadeh, "Prediction of Compressive Strength of Concretes Containing Silica Fume and Styrene-Butadiene Rubber (SBR) with a Mathematical Model," Int J Concr Struct Mater, vol. 7, no. 4, pp. 295–301, 2013, doi: 10.1007/s40069-013-0055-y.

[117] M. Castelli, L. Vanneschi, and S. Silva, "Prediction of high performance concrete strength using Genetic Programming with geometric semantic genetic operators," Expert Syst Appl, vol. 40, no. 17, pp. 6856–6862, 2013, doi: 10.1016/j.eswa.2013.06.037.

[118] M. L. Nehdi and A. M. Soliman, "Artificial intelligence model for early-age autogenous shrinkage of concrete," ACI Mater J, vol. 109, no. 3, pp. 353–362, 2012, doi: 10.14359/51683826.

[119] M. I. Khan, "Mix proportions for HPC incorporating multi-cementitious composites using artificial neural networks," Constr Build Mater, vol. 28, no. 1, pp. 14–20, 2012, doi: 10.1016/j.conbuildmat.2011.08.021.

[120] B. Boukhatem, S. Kenai, A. T. Hamou, D. Ziou, and M. Ghrici, "Predicting concrete properties using Neural Networks (NN) with Principal Component Analysis (PCA) technique," Computers and Concrete, vol. 10, no. 6, pp. 557–573, 2012, doi: 10.12989/cac.2012.10.6.557.

[121] M. Y. Cheng, J. S. Chou, A. F. V. Roy, and Y. W. Wu, "High-performance Concrete Compressive Strength Prediction using Time-Weighted Evolutionary Fuzzy Support

Vector Machines Inference Model," Autom Constr, vol. 28, pp. 106–115, 2012, doi: 10.1016/j.autcon.2012.07.004.

[122] S. M. Mousavi, P. Aminian, A. H. Gandomi, A. H. Alavi, and H. Bolandi, "A new predictive model for compressive strength of HPC using gene expression programming," Advances in Engineering Software, vol. 45, no. 1, pp. 105–114, 2012, doi: 10.1016/j.advengsoft.2011.09.014.

CHAPTER III

Methodology approach for prediction of the thermal conductivity of concrete by using neural networks

Ana Carolina Rosa [a,b], Youssef Elomari [b], Alejandro Calderón [c], Carles Mateu [d], Assed Haddad [a], Dieter Boer [b] *

[a] Programa de Engenharia Ambiental, UFRJ, Universidade Federal do Rio de Janeiro, Rio de Janeiro, 21941-901, Brazil

[b] Department of Mechanical Engineering, University Rovira i Virgili, Av. Països Catalans, 26, 43007 Tarragona, Spain

[c] Department of Materials Science and Physical Chemistry, Universitat de Barcelona, Martí i Franquès 1-11, Barcelona 08028, Spain

[d] Department of Computer Science and Industrial Engineering, University of Lleida, Lleida, Spain

KEYWORDS: Artificial neural networks, MLP, GAN, Concrete, Thermal conductivity.

NOMENCLATURE

Abbreviations	
Ad	Admixture content
ANN	Artificial neural network
C	Cement content
CAg	Coarse aggregate content
Ce	Ceramic content
CTGAN	Conditional Tabular GAN
DNN	Deep neural network
F	Fiber content
FA	Fly ash content
FAg	Fine aggregate content
FV	Foam volume content
GAN	Generative adversarial network
LAg	Lightweight aggregate content
ML	Machine learning
MLP	Multilayer perceptron
NAg	Natural aggregate content
RAg	Recycled aggregate content
S	Slag content

SCM	Supplemental cementitious material
SF	Silica fume content
Sp	Superplasticizer content
W	Water content
W/C	Water-cement ratio
WAg	Waste aggregate content
Symbols	
Coefficient of determination	R^2
Density - D	ρ [kg/m³]
Experimental data found in the literature	y
Mean of the experimental data	y_m
Mean of the normalized value	μ
Normalized value	z
Number of data points	n
Number to be normalized	x
Predicted data	y'
Root mean squared error	$RMSE$
Standard deviation	σ
Thermal conductivity - TC	k [W/m.K]

3.1. Introduction

Driven by a growing sustainability awareness, energy efficiency is a topic that researchers are addressing more and more. As in other areas, concerns about energy efficiency solutions for buildings have frequently increased. For this reason, the demand for the development of materials to improve thermal performance in buildings has increased. An appropriate design and material selection can lead to more environmentally responsible and economically sustainable construction practices. Using materials for better thermal efficiency can preserve indoor thermal comfort despite fluctuations in outdoor environment conditions and reduce energy consumption [1].

In the building design phase, vast possibilities of methods and materials exist to guarantee an adequate combination to achieve an energy-efficient building. Two methods can be used to improve the energy efficiency of a building: active methods and passive methods. The first method incorporates energy-efficient technologies to use or produce energy in the building, such as heat pumps and solar photovoltaic panels. The second method uses some countermeasures to reduce the thermal transmittance of the structure and enhance this efficiency using materials with low thermal conductivity. For instance, concrete can be used as part of the building envelope to improve energy efficiency [2]. It is the most used and stands out the most among the different building materials due to its unique features. Besides its resistance, durability, and low cost, its inherent thermal mass attracts decision-makers. It allows heat to be absorbed, stored, and gradually released, stabilizing indoor temperatures and reducing unnecessary heating or cooling [3]. However, as concrete is a composite material of cement, water, fine aggregate, and coarse aggregate, its composition presents a wide variety, leading its properties to undergo significant variations [4]. Adjusting these constituents and their quantities allows the properties of both fresh and hardened states to be tailored to achieve the required design specifications. Thus, a specific property can be increased or decreased by changing the mass composition or including different constituents in a concrete mixture.

Thermal conductivity is one fundamental property affecting the material's energy efficiency, which measures how well a material conducts heat. Regarding the effect of heat transfer in a building, materials with low thermal conductivity value are used as isolation. They reduce the heat transfer between the indoors and outdoors of the building, which maintains comfortable temperatures and reduces the energy consumption of heating and cooling systems. Factors influencing concrete's thermal conductivity include constituents' type and weight, such as fine and coarse aggregates, supplemental cementitious materials (SCMs), and fibers [5]. Since a concrete composition presents 50 - 70 % of aggregates, their mineralogy and volume on the mixture have most effect on the thermal conductivity of concrete [6].

Accurately predicting the thermal conductivity of concrete becomes vital to improving buildings' energy efficiency. A prediction model can help avoid many time-consuming experiments in the design and material selection phases. In recent years, Machine Learning (ML) models have been used to predict some properties of concrete, such as compressive strength [7], tensile strength [8], modulus of elasticity [9], flexural strength [10], slump [11], chloride penetration [12], carbonation depth [13] and surface chloride concentration [14]. Although many researchers used ML, there is a limitation in studies developing prediction models of the thermal properties of concrete [15]. Among the machine learning-based models, the Artificial Neural Network (ANN) is one of the most employed ones to solve complex problems and has applications in many fields [16]–[18]. The importance of ANNs lies in their ability to learn and make decisions based on data, which makes them highly valuable in different areas. ANN is composed of a set of networks of interconnected nodes, which work together to learn complex relationships between a group of inputs and outputs. This makes them well-suited for predicting some values based on a variety of parameters. In this way, ANNs can be used to solve problems that conventional or other computational methods have difficulties [19]. ANNs provide an alternative method for predicting concrete properties that is faster, cheaper, and more accurate than traditional methods.

One of the drawbacks typically encountered when developing an ANN is a limited data set, and effectively training an ANN requires a massive amount of data. When this problem arises, an alternative that has been emerging is the use of the Generative Adversarial Network (GAN). GANs were introduced by Goodfellow et al. [20] and became a revolutionary development in the world of generative modeling. They have different applications, such as image generation, super-resolution imaging, style transfer, and data augmentation. Regarding tabular data, it can create synthetic data to expand real datasets and prevent over-fitting in such data-limited situations, helping to improve the training of an ML model. Although GAN started with image generation, some authors have already used this technique to create tabular data and obtain satisfactory results. Since the introduction of GAN, several algorithms to model tabular data have been used, such as Conditional Tabular GAN (CTGAN) [21], TabGAN, and CopulaGAN [22]. Although many researchers have attempted to use ANN to predict different properties of concrete, only few works are progressing on models to determine the thermal properties. Additionally, the papers do not have a generalist model focusing on different types of concrete, including many constituents.

This work intends to fill this gap, and the novelty is the development of a methodology using an ANN model integrated with a data augmentation model to predict the thermal properties of concrete containing different types of materials such as slag, lightweight and recycled aggregates, fibers, and others. For this methodology, a Multilayer perceptron (MLP) model to predict thermal conductivity and a Copula GAN model to improve the tabular data from published papers will be developed based on the constituents' composition and the concrete's density.

The rest of the paper is organized as follows. Section 2 brings an overview of ANN used to predict the properties of concrete. Next, section 3 describes the methodology adopted for this work. Section 4 evaluates the results using the case study and discusses the main findings. Finally, the final section contains concluding remarks on the general findings.

3.2. Overview of neural networks

Over the years, ANN has been applied to predict different properties of concrete due to its ability to model complex non-linear relationships. Regarding the property predictions, the mechanical properties are the most evaluated, with compressive strength being the most investigated in different machine learning models. Yeh was one of the pioneers applying ANN to predict the properties of concrete and developed some models to predict compressive strength and slump [23], [24]. Kandiri et al. estimated the compressive strength using a hybrid model of ANN and salp swarm algorithm [25]. Abellán-García trained an MLP model to forecast the compressive strength for a given ultra-high-performance concrete mixture design [26]. Besides that, another work also developed an ANN with a feedforward backpropagation algorithm to predict the slump flow and compressive strength while incorporating silica fume, limestone powder, recycled glass powder, and fluid catalytic cracking residue [27].

Although many studies are developing new models to predict the properties of concrete, only a few papers investigated thermal conductivity or other thermal properties. Fidan et al. [28] trained different structures of an ANN model to predict thermal conductivity through five parameters – density, compressive strength, tensile strength, porosity, and ultrasonic pulse velocity. The best solution performance was achieved with a neural network with three layers and the following neuron sequence of 5, 25, and 1 in each layer. Sargam et al. [6] evaluated nine machine learning models, and MLP showed the highest prediction accuracy using the maxout activation function and three hidden layers, each containing 100 neurons. Kurpińska et al. [29] also investigated the influence of varying the number of neurons in the hidden layer to forecast the thermal conductivity of lightweight concrete. The model presented a sigmoid function and a structure with four layers: an input layer with two neurons, the first hidden layer varying from 2 to 12, the second hidden layer ranging from 2 to 17, and the output with one neuron. Kursuncu et al. [30] used ANN and ANOVA to investigate the effect of partial replacement of waste marble powder and rice husk ash instead of fine aggregate and cement into foam concrete. The results indicated ANN as the most adequate to estimate the thermal conductivity. Gence et al. [31] compared two neural networks, the radial basis neural network and MLP, to predict the thermal conductivity of concrete with vermiculite and concluded that the former had greater accuracy.

Different types of ANNs have been successfully applied to predict the thermal conductivity of concrete. Table 1 compares some studies using distinct architectures of ANN, which indicates that most models employ the backpropagation algorithm and do not have an extensive database.

Table 1. Comparative studies of different ANN-based methods to predict thermal conductivity.

Reference	ML method	Concrete	Number of inputs	Number of hidden layers	Neurons of hidden layers	Number of outputs	Number of datasets	Activation function	Evaluation criteria
Sargam et al. [6]	MLP	Concrete containing modern constituent materials	3, 5, 6, 8, 9, 10,13	3	100 100 100	1	213	Maxout	MAE
Fidan et al. [28]	ANN	Concrete	5	1	5,10,15, 20,25	1	132	Tangent sigmoid	MAE, MAPE, RMSE, R^2
Kurpińska et al. [29]	Backpropagation NN	Lightweight concrete	2	2	2-12 2-17	1	15	Sigmoid	MSE
Kursuncu et al. [30]	ANN	Foam concrete	-	-	-	1	18	Sigmoid	R
Gencel et al. [31]	Radial basis NN / MLP	Concrete with vermiculite	5	1	3	1	20	Non-linear	RMSE, MSE, R
Lee et al. [32]	Backpropagation NN	Concrete	11	2	20 20	1	152	Sigmoid / linear	MSE, R
Ozel and Topsakal [33]	Backpropagation NN	Construction materials	2	1	1	1	110	-	RMSE, R^2
Ipek et al. [34]	Backpropagation NN	Rubberized concrete	5	1	2 - 20	1	127	Tangent hyperbolic	R2, MAPE, MSE,

Besides this brief literature overview, a general bibliometric analysis was performed to identify the most relevant and influential literature in predicting the properties of concrete using machine learning. This approach enables the identification of critical research gaps and areas for future investigation. To proceed with this analysis, VOSviewer software [35] was used. It visually represents the research and supports researchers in finding insights into the research domain. Consequently, it can lead to more effective development of prediction models of thermal properties.

In order to evaluate the development of the research regarding property predictions using machine learning, a bibliometric study in the Web of Science database was performed. The following query was used: ("concrete") AND ("properties" OR "thermal conductivity") AND ("machine learning" OR "deep learning" OR "artificial neural network" OR "ANN"). The search criteria used in this bibliometric study summarized 1673 articles and review articles published between 2014 and 2023. Then, the overall result was imported into VOSviewer to analyze the keywords in each paper's title or abstract. This generated the overlay cluster representing the development of the research over the years (Figure 1).

Figure 1 depicts the relationship between machine learning models and the properties of concrete. Although the documents' timeline is longer, the figure shows a shorter and more current timeline due to the many published papers. It is possible to observe several machine learning techniques being used to predict the properties of concrete, in addition to other mathematical models such as genetic algorithms, particle swarm, and adaptive neuro-fuzzy inference systems. The models include random forest, support vector machine, extreme gradient boosting, and neural networks. Furthermore, different properties such as slump, compressive strength, porosity, and thermal conductivity are observed, and the different types of concrete studied include geopolymer, ultra-high performance concrete, self-compacting concrete, and recycled aggregate concrete. Hence, machine learning models are a very current research line with great expectations of development in this area. The efficacy of neural networks, which remain widely employed, is particularly notable in the present context. Hence, machine learning models are a very current research line with great expectations of development in this area. The effectiveness of neural networks is particularly notable in the present context.

Therefore, after the literature and the bibliographic review, we noticed a gap in prediction models regarding thermal properties using machine learning models. This study aims to fill this gap and develop a method to predict the thermal conductivity of concrete. Additionally, we intend to build a model for data augmentation that has been used in many fields.

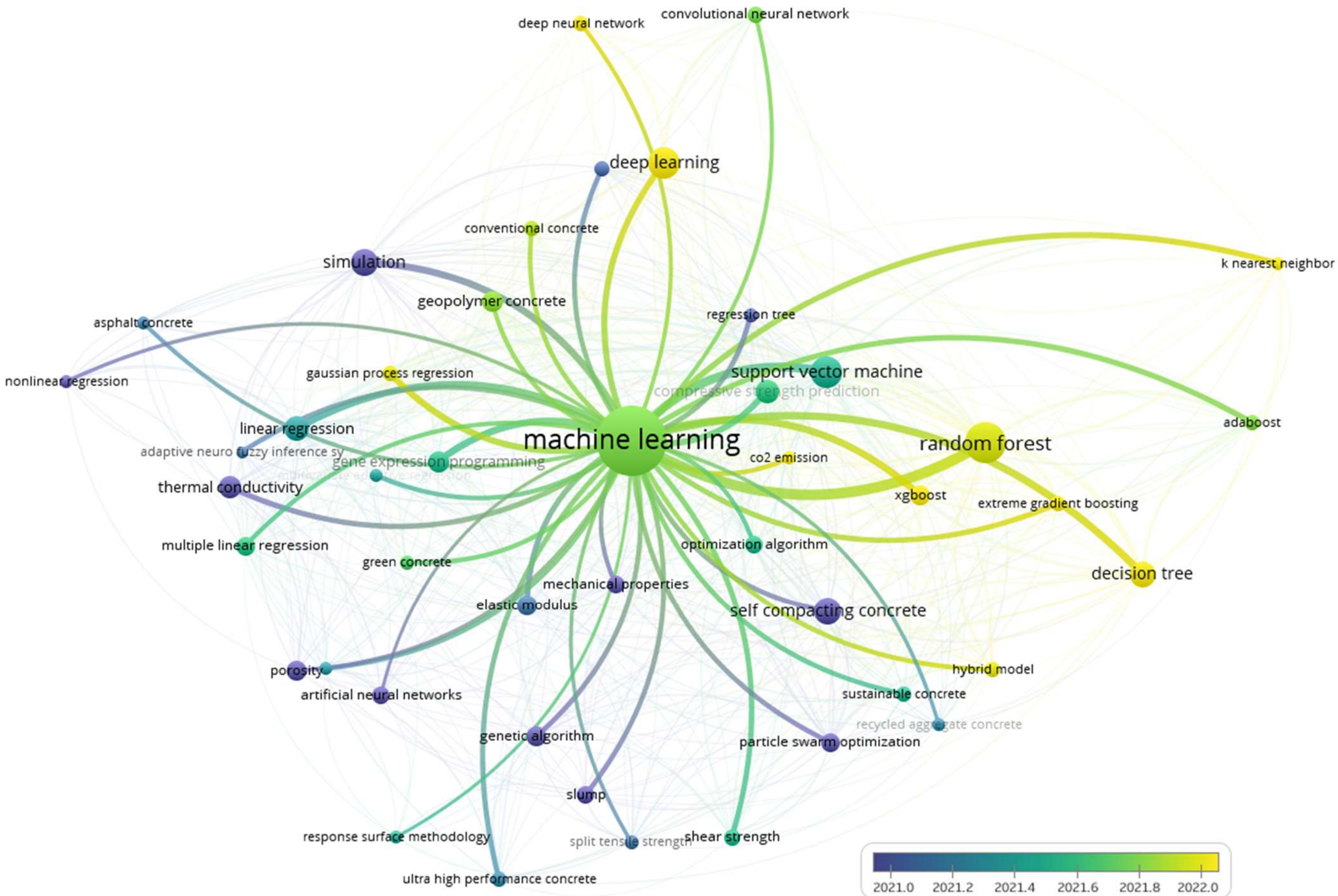

Figure 1. Cluster based on keywords (Network visualization).

3.3. Methodology

The proposed methodology encompasses two primary components: the development of an MLP for the material property prediction based on its features and the utilization of a GAN for data augmentation to enhance the dataset and improve the accuracy of the first model. This section outlines the proposed methodology applied to this work to achieve these objectives, broken down into three steps: data collection, model development, and the model's application in a case study (Figure 2).

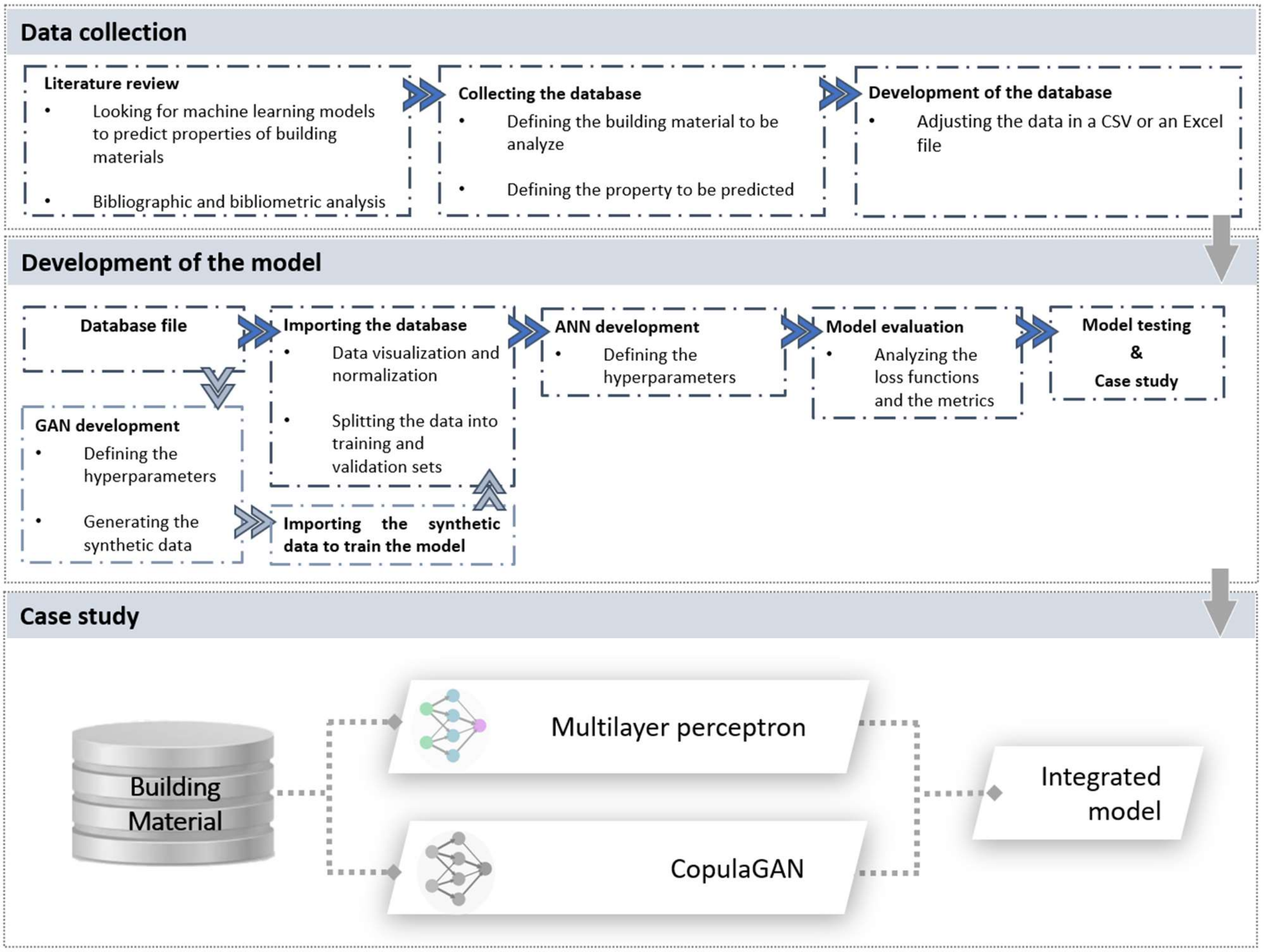

Figure 2. Methodology framework

3.3.1. Data collection

The first step corresponds to a literature review and gathering the database from diverse sources, including industry databases, research publications, and laboratory experiments. Collecting the database is the primordial step in building the predictive model, as the ANN will learn from this data. If it presents an inadequate representation of the problem, the model cannot predict the property effectively, thus reducing the model's reliability. Furthermore, for the model to be representative, there must be a sufficiently large amount of data to ensure diversity. In this step, it is necessary to recognize which relevant input features potentially impact the model's output response. After this process, all the information must be inserted into a CSV or Excel file.

3.3.2. Development of the model

The second one is related to the general process of building the prediction model, i.e., the data normalization, the selection of an appropriate neural network architecture for the prediction task, defining the learning rate of the neural network, the number of hidden layers and neurons in each layer, and the metrics to find the best model for the dataset.

3.3.2.1. Importing the database

Normalizing the dataset is vital in any deep-learning model. Since the database may present the parameters in a different and wide range of values, it is necessary to normalize them by putting all features on a standard scale to avoid the data values affecting the weights and biases during the training model. Generally, the normalization process transforms the inputs and output parameters, taking them to a standardized scale, i.e., transforming them to have a mean of zero and a standard deviation of one. This step guarantees that each parameter will contribute equally during the learning process and prevent any parameter from dominating this process due to its more extensive scale. Another advantage is that it enhances the generalization capabilities of the models, improving the prediction of new data. The normalization known as z-score is commonly used to bring the numerical features to the same scale. The mathematical equation of this process is presented in Equation 1.

$$z = (x - \mu)/\sigma \qquad \text{(Eq.1)}$$

Where z is the normalized value, x is the number we would like to normalize, μ is the mean, and σ is the standard deviation of this parameter. Therefore, in order to achieve the normalized values for each numeric column, the mean and the standard deviation are calculated. Each value in the numeric column is subtracted by the column's mean and then divided by the column's standard deviation.

3.2.2 Performance metrics

Assessing the model's performance is critical to data analysis and predictive modeling. Since each metric has a unique perspective on different aspects of the model, it is crucial to evaluate more than one metric to achieve a comprehensive and well-rounded assessment of the model's performance, which will give a complete picture of the problem and allow us to gain insights into the strengths and weaknesses of the model.

To quantify the predictive capabilities of the proposed model and enable meaningful comparisons among the different architectures and previous models published in the literature, it is advisable to use more than one performance metric. Researchers commonly employ complementary performance metrics, such as the coefficient of determination (R^2) and root mean squared error (RMSE). The former measures the proportion of the variance in the output that is predictable from the input variables. Its value varies from 0 to 1, and the closer this value to 1, the better the model's capability to explain the variability of the data. However, the latter measures the average error between the predicted and actual output values. Due to its squared term, it is more sensitive to outliers and more significant errors. The lower this metric, the more accurate the predictions are. These metrics are valuable indicators of how well the models capture the data's variance and their predictions' accuracy. The mathematical formulae for R^2 and RMSE are provided in Equations (2) and (3).

$$R^2 = \frac{\sum_{k=1}^{n}(y-y')^2}{\sum_{k=1}^{n}(y-y_m)^2} \qquad \text{Eq (2)}$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}\ (y-y')^2}{n}} \qquad \text{Eq (3)}$$

Where n is the number of data points, y is the experimental data found in the literature, y' is the predicted data, and y_m is the mean of the experimental data.

3.2.3. Multilayer Perceptron model

The Multilayer Perceptron (MLP) is an artificial neural network that can be applied to solve many problems, including classification, regression, and pattern recognition. A typical MLP model comprises three blocks of interconnected neurons: the input layer in which each neuron represents a data feature; the hidden layers that could have one or more layers depending on the complexity of the problem; and the output layer representing the response output (Figure 3). Each neuron in the network is bonded to other neurons through connection weights. Each neuron in a neural network receives an input (X_i) that is multiplied by a weight (W_i) and is summed with each other and added to the bias value (b). Then, the result is transferred to the activation function, which adjusts the final output (Figure 4). Each layer of neurons gathers input from the previous layers, and the outputs of neurons within one layer become the inputs to neurons in the following layer. Finally, the last layer produces the predictions of the model.

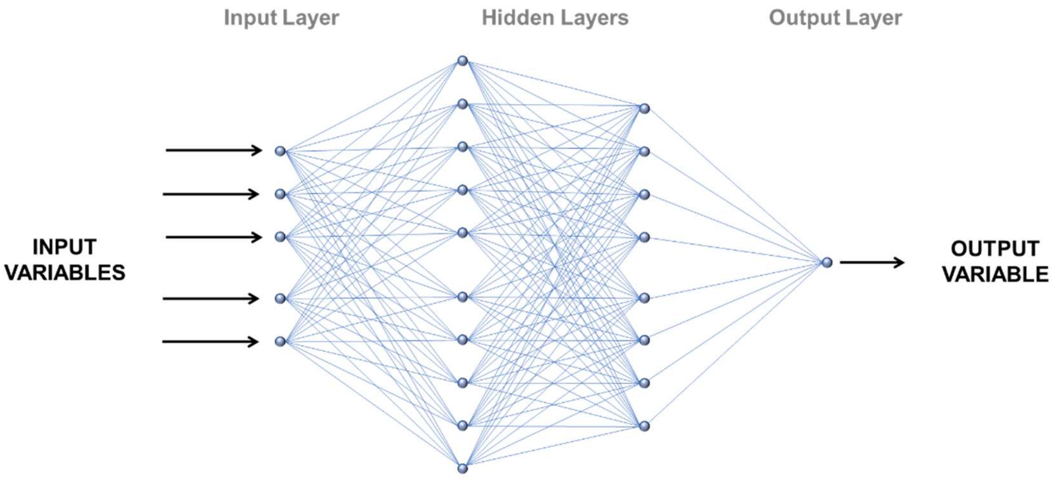

Figure 3. Example of the MLP model

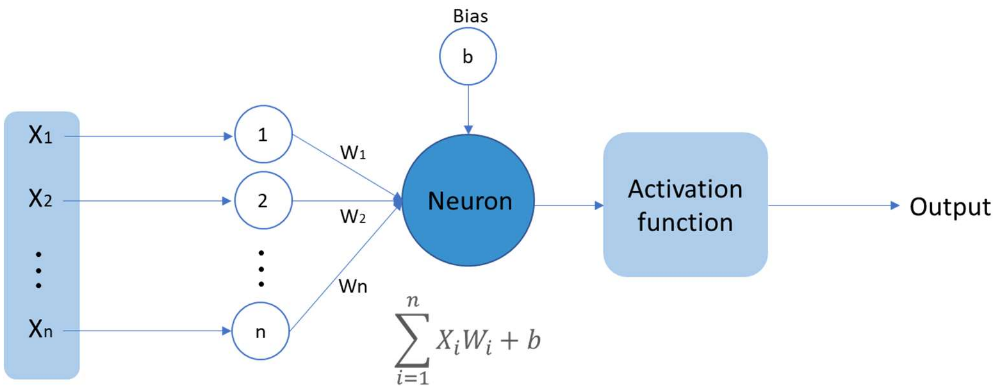

Figure 4. Graphical representation of a neuron

3.2.4. Generative Adversarial Network model

A Generative Adversarial Network (GAN) model consists of two distinct neural networks, the generator and the discriminator, which are trained simultaneously. The GAN captures the distribution pattern of a given dataset and generates data that resembles this dataset. First, the generator initiates the process, which takes a random noise as an input, creating samples similar to the original dataset. The discriminator then tries to distinguish the data and detect whether a sample is from the original dataset or the generator model distribution. During the training, the generator loss and the discriminator loss are evaluated, allowing the generator to get better and better at producing new artificial data and the discriminator to find if they are real or synthetic (Figure 5). This methodology framework used the Copula GAN synthesizer, a CTGAN variation. It uses the cumulative distribution function-based transformation that Gaussian copulas apply, making the learning process easier.

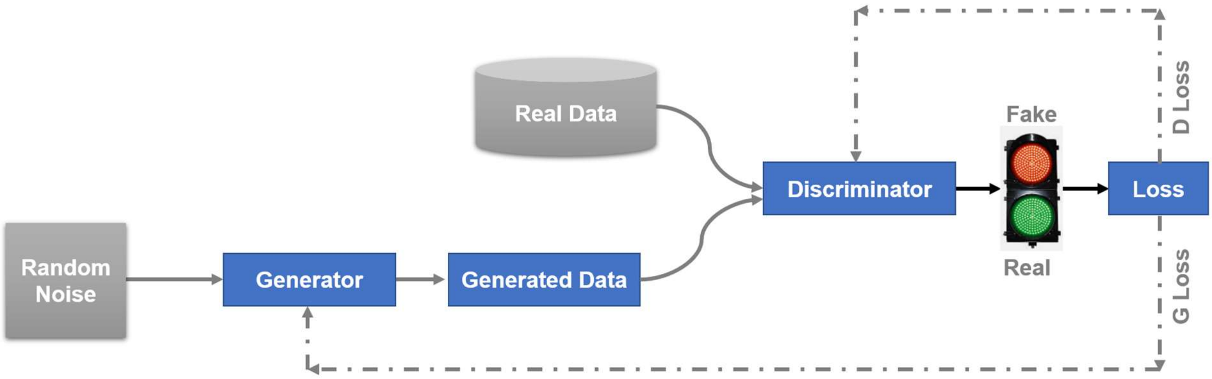

Figure 5. Flowchart of a Generative Adversarial Network (GAN)

3.3. Case study

The proposed model is applied to conduct a broad case study on concrete mixtures. As previously highlighted in the introduction, precise thermal conductivity predictions for building materials can significantly enhance building thermal performance. Given that the thermal conductivity of concrete can be influenced by the type and quantity of each constituent, this study seeks to predict this critical property based on the mass composition of these constituents.

A comprehensive literature review is conducted to create a database of 200 points compiled from various relevant studies published in the literature. Each selected paper presents a different type of concrete, such as recycled aggregate concrete, lightweight concrete, fiber-reinforced foamed concrete, moderate strength concrete, and concrete with vermiculite. This variation allows the development of a generalist model. The dataset compiled in this context consists of concrete from 12 sources (Table 2) with modern constituent materials: ceramic, slag, silica fume, fly ash, steel fiber, propylene fiber, recycled aggregate, lightweight aggregate, waste aggregate, and foam. The model's inputs are defined according to the available data and the influence of each feature on the thermal conductivity.

Here, a total of 18 input parameters that affect the output, thermal conductivity, are considered. The selected input parameters for this case study are density (D), water-to-cement ratio (W/C), water content (W), cement content (C), ceramic powder content (Ce), fine aggregate content (FAg), coarse aggregate content (CAg), natural aggregate content (NAg), recycled aggregate content (RAg), light aggregate content (LAg), waste aggregate content (WAg), fly ash content (FA), silica fume content (SF), slag content (S), fiber content (F), other admixture content (Ad), superplasticizer content (Sp), foam volume content (FV). The values of TC lie in the range of 0.37 – 2.70 W/m.K with a standard deviation of 0.5 W/m.K.

In order to train the model, the compiled dataset is randomly divided into training (80%) and validation sub-dataset (20%). Then, the reliability and reproducibility of the ANN model are evaluated on a new testing dataset. The predictive performance of the final model is thoroughly assessed using the metrics mentioned above. This assessment allows us to check the model’s accuracy and efficiency in capturing the underlying patterns within the database.

Table 2. The summary of the dataset used in the model

		Zhu et al. [36]	**Kim et al.** [37]	**Lee et al.** [32]	**Vejmelková et al.** [38]	**Demirboga** [39]	**Sargam et al.** [40]
N° of data		15	14	82	5	9	21
D	[kg/m³]	1617 - 2195	1540 - 2297	1144 - 3075	2017 - 2199	2290 - 2355	1626 - 2380
W/C	-	0.40 - 0.50	0.25 -0.57	0.22 – 0.62	0.32 - 0.80	0.35	0.35 - 0.55

W	[kg/m³]	130 - 150	140 - 558	133 - 524	124	167.5	120 - 190
C	[kg/m³]	260 - 375	0 - 1792	242 - 1762	154 - 385	245 – 324	90 - 110
Ce	[kg/m³]	-	-	-	0 - 231	-	-
FAg	[kg/m³]	0 - 658	0 - 702	0 - 1167	775	740	280 - 365
CAg	[kg/m³]	0 - 605	0 - 1103	0 - 1850	175	458	-
NAg	[kg/m³]	0 - 987	-	-	770	577	0 - 440
RAg	[kg/m³]	0 - 1710	-	-	-	-	0 - 460
LAg	[kg/m³]	-	-	-	-	-	0 - 320
WAg	[kg/m³]	0 - 1645	-	-	-	-	-
FA	[kg/m³]	-	0 - 973	0 - 104	392	26.25 - 105	0 - 30
SF	[kg/m³]	-	-	0 - 89	-	26.25 - 52.5	-
S	[kg/m³]	-	0 - 1282	-	-	26.25 - 105	-
F	[kg/m³]	-	-	-	-	-	0 - 20
Ad	[kg/m³]	-	0 - 8.81	-	-	-	-
Sp	[kg/m³]	-	-	-	3.4 - 4	1.8	-
FV	[kg/m³]	-	-	-	-	-	-
TC	[W/m.K]	0.74 – 1.40	0.66 – 1.69	0.30 – 2.50	1.21 – 1.55	0.95 – 1.17	0.70 – 1.20

		Sargam et al. [41]	**Sargam et al.** [42]	**Cavalline et al.** [43]	**Kurpinska et al.** [44]	**Gencel et al.** [31]	**Bayraktar et al.** [45]
N° of data		4	1	17	4	4	20
D	[kg/m³]	2281 - 2367	2389	1434 - 2219	1234 - 2244	1194 - 1370	1418 - 1988
W/C	-	0.45	0.43	0.29 - 0.54	0.40	0.63 - 1.37	0.75 – 1.88
W	[kg/m³]	157	151	155 - 218	220 - 310	471 - 580	338
C	[kg/m³]	349	281	268 - 454	500 - 700	422 - 750	180 - 450
Ce	[kg/m³]	-	-	-	-	-	-
FAg	[kg/m³]	772 - 821	890	0 - 1559	0 - 1300	-	0 - 1091
CAg	[kg/m³]	-	-	-	-	149 - 192	-
NAg	[kg/m³]	0 - 1171	900	0 - 1068	-	-	-
RAg	[kg/m³]	0 - 1036	-	-	-	-	-
LAg	[kg/m³]	-	-	0 - 1101	-	-	-
WAg	[kg/m³]	-	-	-	0 - 490	-	0 - 1091
FA	[kg/m³]	-	71	0 - 102	-	-	-
SF	[kg/m³]	-	-	-	-	-	-
S	[kg/m³]	-	-	-	-	-	0 - 270
F	[kg/m³]	-	-	-	-	-	0 - 9
Ad	[kg/m³]	-	-	-	-	-	-
Sp	[kg/m³]	-	-	-	3 - 4.9	-	-
FV	[kg/m³]	-	-	-	-	-	50 - 100
TC	[W/m.K]	0.89 - 1.14	2.70	0.69 - 1.78	0.37 - 1.19	0.33 - 0.52	0.50 - 0.75

4. Results and discussion

This section presents the model specifications, performance metrics, and outcomes derived from the implementation analysis of deep neural networks. The results are organized into three subsections: firstly, the implementation of the Multilayer Perceptron (MLP) model in Python; secondly, an assessment of the feasibility of generating synthetic data using CopulaGAN; and lastly, the training and evaluation of the MLP model with two distinct datasets.

4.1 Development of the multilayer perceptron model

The primary objective of this work is to develop an MLP model for predicting the thermal conductivity of concrete based on the constituents' mass composition and density of concrete. The hyperparameters of the MLP model are initially predefined and subsequently fine-tuned to enhance the prediction accuracy. Table 3 provides an overview of the model's key features, encompassing the hyperparameters, evaluation metrics, activation functions, optimization methods, and loss functions.

Table 3. Main features of the MLP model

Hyperparameters	
	Hidden layers = [1, 2, 3, 4]
	Batch size = 64
	Learning rate = 0.001
	Epochs = 100
Metrics	
	RMSE
	R^2
Activation function	
	ReLU
Optimization function	
	Stochastic Gradient Descent (SGD)
Loss function	
	MSE

To establish the optimal architecture of the MLP model, the number of hidden layers and neurons is systematically varied, and a combination of the metrics RMSE and R^2 is employed to measure the model's overall performance. A diligent trial-and-error approach is undertaken to determine the most effective configuration for the number of hidden layers and neurons in each layer. Table 4 displays the loss functions and performance metrics of multiple neural networks assessed in this study. Notably, the most proficient model features a two-hidden-layer structure with 200 and 100 neurons, yielding an RMSE of 0.111 W/m.K and an R^2 of 0.984 for the training dataset. As for the validation dataset, the model demonstrates metrics values of 0.183 and 0.960, respectively.

Table 4. Performance indices and metrics of the MLP model

Layers	ANN Architecture	Train Loss	Valid Loss	RMSE	R^2
4	200 – 100	0.083	0.012	0.111	0.984
5	200 – 100 – 50	0.107	0.024	0.155	0.968
5	200 – 100 – 40	0.099	0.037	0.192	0.950
5	200 – 100 – 30	0.087	0.028	0.167	0.960
5	200 – 100 – 20	0.012	0.037	0.191	0.949
6	200 – 100 – 50 – 25	0.102	0.064	0.253	0.918

Regarding the best-performing model, an analysis of the loss functions and the performance metrics versus the number of epochs is conducted (Figure 6). As the number of epochs progresses, a significant pattern appears in the training and validation processes. Both training and validation losses exhibit an exponential decrease as iterations continue. In parallel, the performance metrics on the validation dataset align with the behavior of the loss function. While RMSE steadily declines, R^2 shows a corresponding increase, achieving its minimum and maximum value at approximately 100 epochs. This observed trend reflects the model's capacity to optimize its predictive accuracy as it refines its learning over successive epochs.

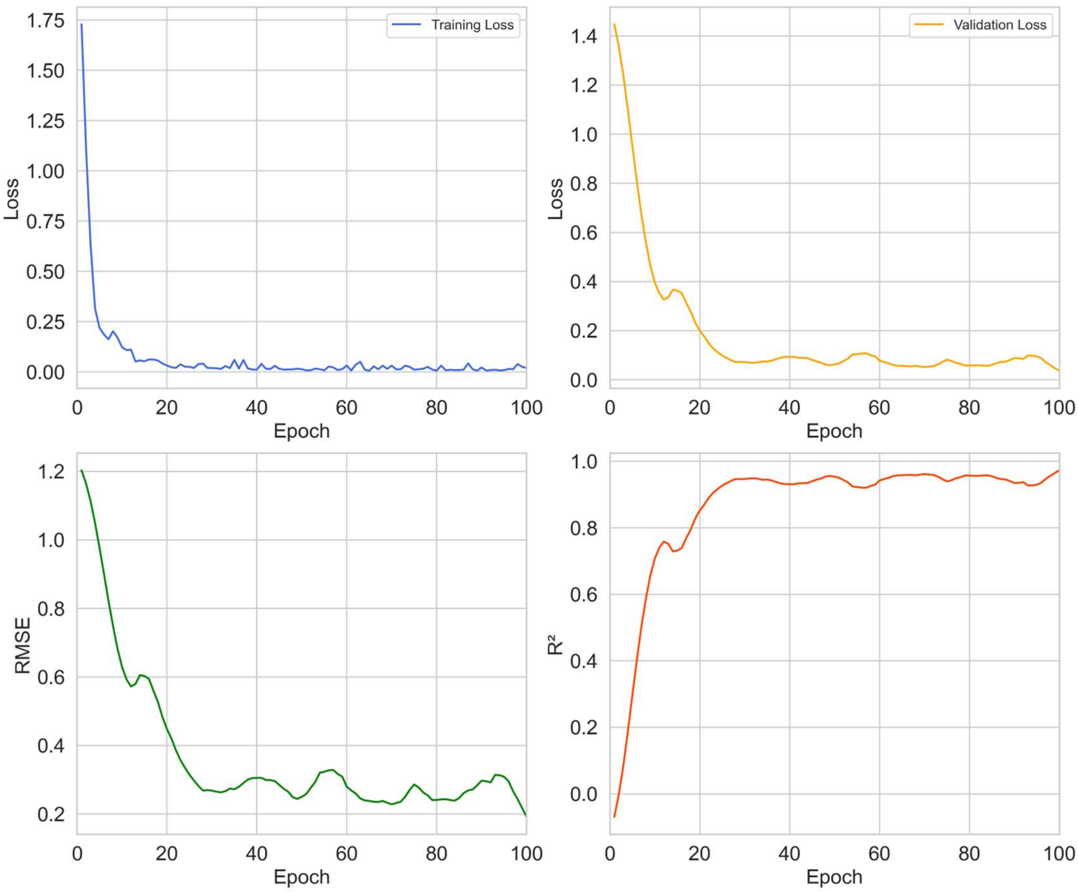

Figure 6. Loss function and performance measures during the training and validation

Comparing the findings of the MLP model to previous studies, they demonstrate a remarkable level of accuracy for the training and validation. A study by Fidan et al. [28], using a dataset with 132 entries, developed an ANN based on the mechanical properties and the density of concrete, which performed an overall R^2 performance of 0.996. However, the ANN architecture used only one hidden layer with 5 to 25 neurons. Sargam et al. [6] took a step further by implementing a neural network with a larger dataset, with an RMSE of 0.117 and R^2 of 0.964 for the training set and an RMSE of 0.215 and R^2 of 0.894 for the validation set. Nonetheless, large residuals were obtained. Another work with a high-performance ANN model was that developed by Gencel et al. [31], in which he investigated the influence of concrete composition and temperature on thermal conductivity. Despite the results of R^2 of 0.998 and RMSE of 0.005 during neural network training, only 12 inputs were used. Although these works showed substantial success in predicting thermal conductivity, our model upholds higher accuracy and introduces a broader array of concrete constituents into the modeling process.

4.2. Development of the Generative Adversarial Network model

The second part of the study considers the implementation of the CopulaGAN for data augmentation. As can be observed in Table 4, increasing the number of layers does not improve the model's learning capacity or performance. If we increase the hidden layers of the model with a small dataset, the model can overfit. For this reason, the loss functions exhibit higher values. In order to solve this issue, we intend to develop a GAN model to increase the dataset, creating new synthetic data based on the real data used to develop the MLP model. It will avoid the overfitting process and make the model more robust with extensive dataset training.

After generating the synthetic data, two aspects of the original data are evaluated: the column shapes and pair trends. Each aspect contributed differently to the overall score, which reached 78.14%, indicating that the synthetic data has acceptable quality. In Figure 7, we present the distribution of the augmented dataset. While it's important to note some variations, the overall trend is consistent with the distribution of normalized thermal conductivity in the real dataset. Suggesting that, despite minor differences, the augmented data closely approximates the distribution characteristics of the original dataset.

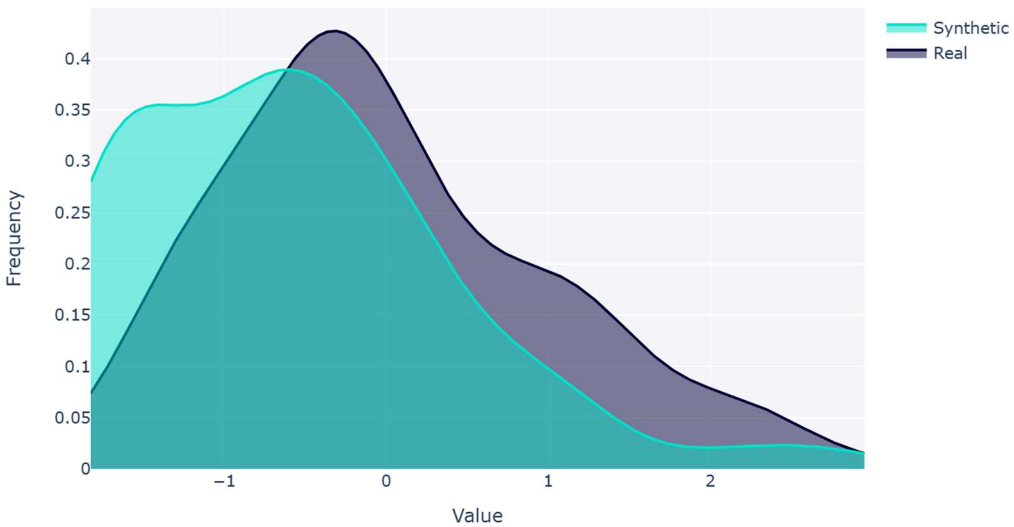

Figure 7. Distribution of the synthetic generated by GAN and the real data

After the synthetic data generation, we aim to incorporate it into the MLP model for joint training with real and artificial datasets. To assess the model's response to this diverse training approach, two distinct sets of synthetic data were created, one comprising 200 data points and the other consisting of 1000 data points. The seamless integration of these two neural network models and the subsequent validation of the model using a new test set are elaborated upon in the forthcoming section.

4.3. Integration of both models to predict thermal conductivity.

The third phase of this study is dedicated to examining the impact of synthetic data on the training of a neural network model. This evaluation is conducted to ascertain whether a model trained with augmented data can produce results comparable to the performance achieved through training with real data and possibly enhance it. In order to assess this impact, this study investigates three distinct scenarios: (a) training the MLP model exclusively with the real dataset containing 200 entries, (b) training the model with the synthetic data consisting of 200 entries first and subsequently incorporating the real dataset, (c) training the model with 1000 data points of the synthetic data and then integrating the real dataset.

The same hyperparameters utilized in developing the final model are maintained in this integration process. The key distinction lies in the dataset used for the training and validation phases. Our approach initially involves training the model first with synthetic data. Subsequently, the same model architecture uses the real dataset to enhance the model's performance. The results for the previous MLP model (a) and scenarios (b) and (c) are summarized in Table 5, indicating the outcomes of the final model.

Table 5. Performance metrics of the MLP model trained with both datasets

Scenario	Dataset	Overall	
		RMSE	R^2
(a)	Training	0.0443	0.9923
	Validation	0.3233	0.9228
	Test	0.0604	0.9846
(b)	Training	0.0221	0.9980
	Validation	0.1329	0.9870
	Test	0.0439	0.9919
(c)	Training	0.0219	0.9981

Validation	0.0735	0.9960
Test	0.0244	0.9975

The predictive capabilities of the developed model are thoroughly examined using a new dataset comprising 30 data points to assess its generalization performance. Despite the limited size of the dataset in scenario (a), the training and validation phases yield commendable results, with an RMSE of 0.0604 and an R^2 of 0.9846. In scenarios (b) and (c), where synthetic data plays a vital role, both demonstrate exemplary performance during training, achieving R^2 values of 0.9995 and 0.9962. This improvement in the results also extends to the validation dataset, with R^2 values of 0.9938 and 0.9754. Post-incorporation of real data into the training process, the metrics exhibit comparable values during training and improvements in validation, 0.9870 and 0.9960 for scenarios (b) and (c), respectively. Upon rigorous evaluation of the independent test set, scenario (c) emerges as the most robust and indicates the most favorable combination of performance metrics.

The predictive performance of the developed MLP model for the three scenarios is evaluated on an independent test dataset that can be graphically seen in Figure 8 and Figure 9. Comparing the actual and predicted thermal conductivity results for each scenario, a distinct pattern emerges, underscoring the model's significant performance enhancement when integrated with synthetic data. In Figure 8(a), we present the predictions for the first case, wherein a considerable portion of the points align with the ideal x=y line. However, some deviations are noticeable, with errors exceeding 0.12. Figures 8(b) and 8(c) illustrate scenarios where the model is trained using synthetic data. In both instances, incorporating synthetic data leads to a discernible improvement in prediction accuracy and a notable error reduction. Scenario (c) stands out, boasting a larger dataset and delivering the most compelling outcomes. This scenario achieves the optimal combination of performance metrics, represented by a best-fit line with an impressive R^2 value of 0.9975 and a minimal RMSE of 0.0244.

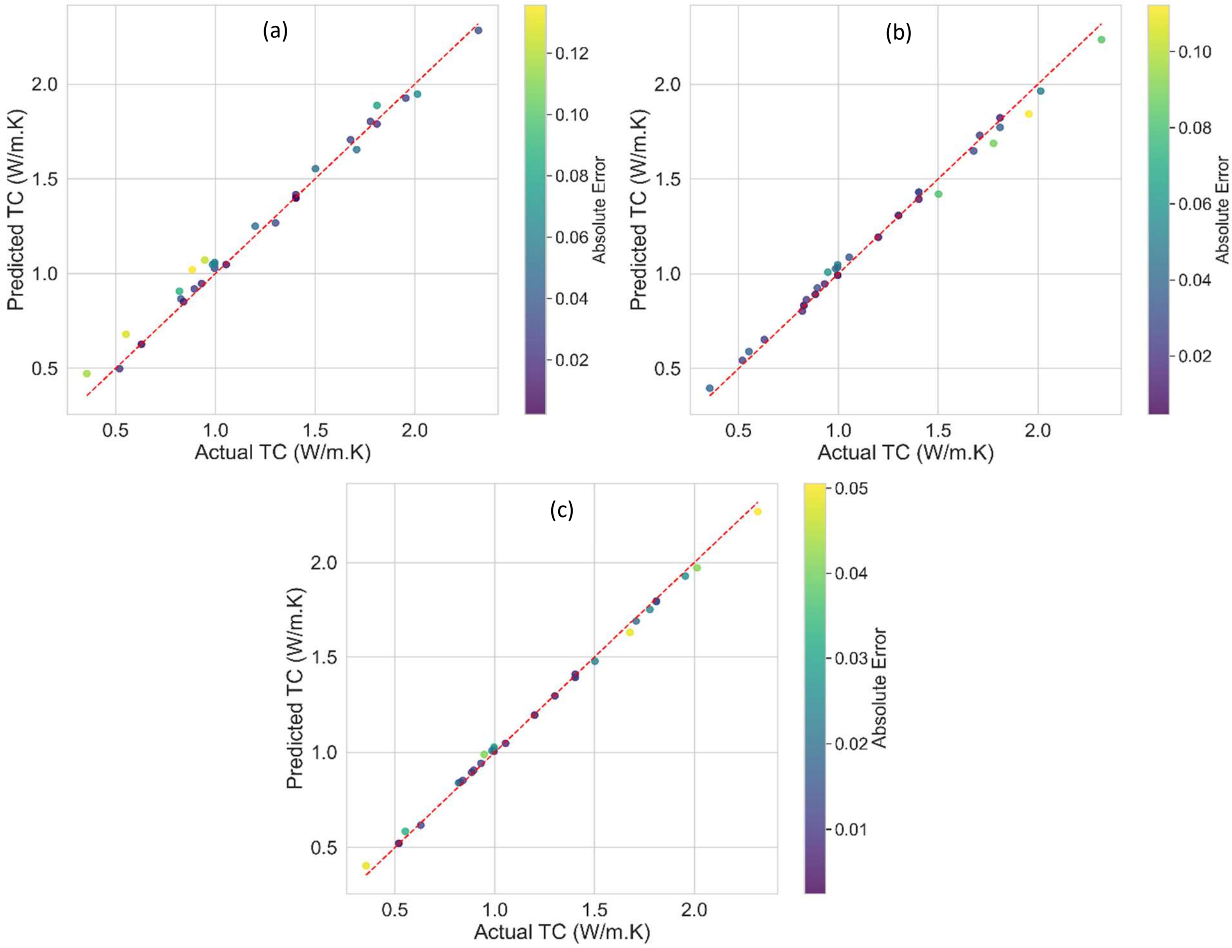

Figure 8. Performance and predictions of the thermal conductivity of the test dataset for the three scenarios: (a) training the model with real data only, (b) training the model with 200 synthetic data and real data, and (c) training the model with 1000 synthetic + real data.

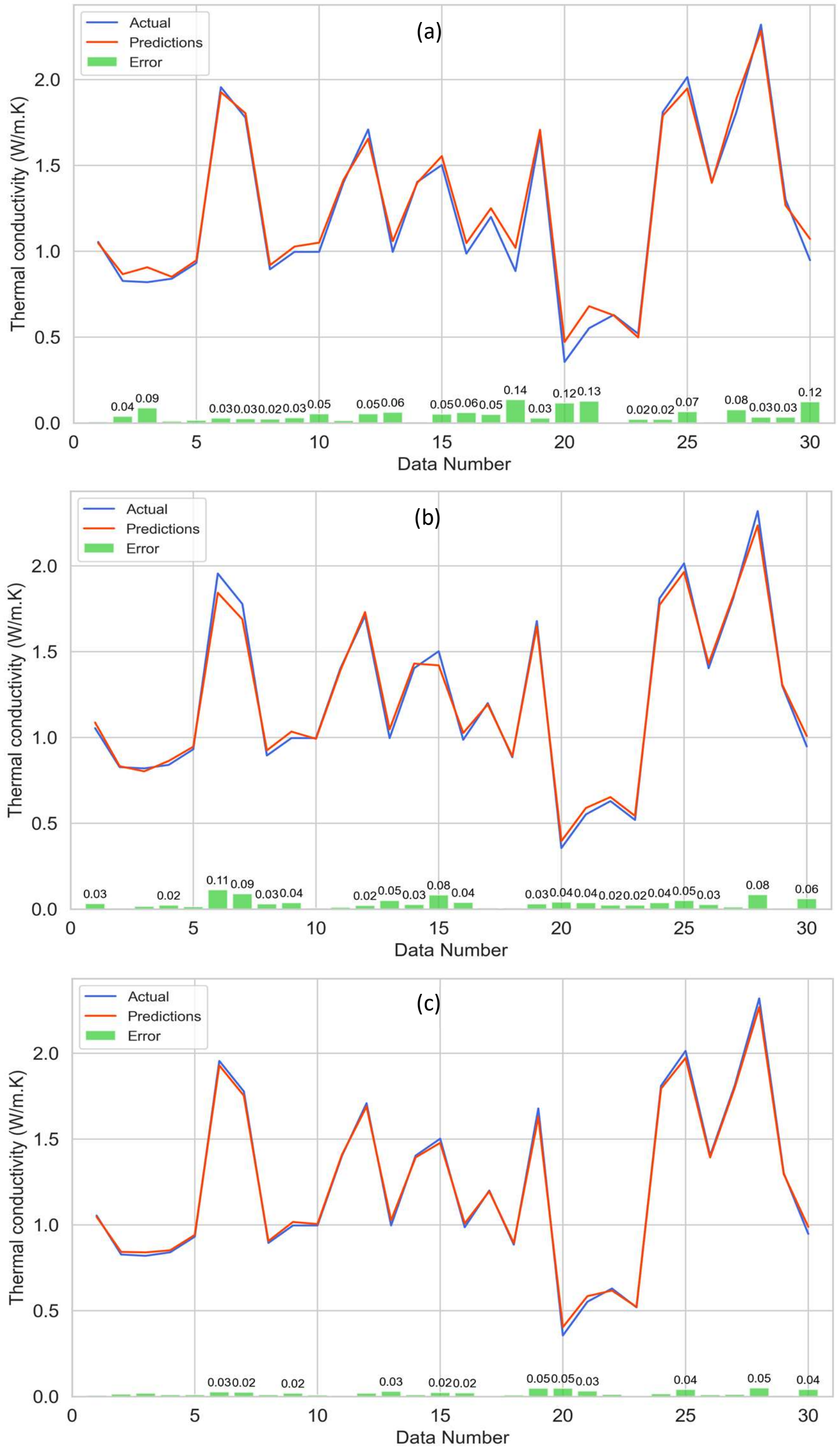

Figure 9. Performance and predictions of the thermal conductivity of the test dataset for the three scenarios: (a) training the model with real data only, (b)

training the model with 200 synthetic data and real data, and (c) training the model with 1000 synthetic + real data.

Figure 9 shows the relationship between the actual values derived from the test dataset from published papers and the corresponding predictions generated by the Multilayer Perceptron (MLP) model. This visual representation allows us to assess the model's accuracy in forecasting thermal conductivity, providing valuable insights into its performance.

In Figure 9(a), we depict the model utilizing solely the real dataset, while Figures 9(b) and 9(c) display the model trained with both the 200-point and 1000-point datasets, respectively. Although Figure 9(a) reveals a close alignment between predicted and real points, some discrepancies across various data points with errors as high as 0.13 and 0.14 W/m.K are observed. Moving to Figure 9(b), a significant reduction in these discrepancies becomes evident, especially within the low thermal conductivity values spanning from 0.5 to 1.0 W/m.K range. This evidence signifies a marked improvement in prediction accuracy, resulting in a substantial decrease in both the number of points with errors and the magnitude of those errors, reducing the maximum error to 0.11 W/m.K. Figure 9(c) unveils an even more prominent enhancement in accuracy, with predictions exhibiting a higher level of precision and a further reduction in errors. In this case, the most significant error recorded is 0.05 W/m.K.

In summary, an impressive alignment between the actual and predicted values is performed across all three cases. The first case exhibits a comparatively lower performance than the other two, potentially attributed to the limited size of the dataset. The presence of a smaller database may result in a more significant deviation between actual and predicted values, particularly within the thermal conductivity range of 0.5-1.0 W/m.K. This scarcity of data becomes apparent when synthetic data is introduced for training, ultimately enhancing the model's accuracy in predicting the thermal conductivity property.

The methodology employed in this study endeavors to show the efficacy of neural networks in predicting the thermal properties of concrete based on the mass composition of its constituents. Furthermore, it seeks to investigate the utilization of synthetic data generation for training neural networks, addressing the challenge of data scarcity, a significant impediment to the accuracy of models employing machine or deep learning techniques. While existing literature demonstrates notable success in utilizing neural networks for thermal conductivity prediction, our model achieves higher accuracy and incorporates a more extensive array of concrete constituents into the modeling process. This diversity enhances the model's versatility, expanding its applicability to a broader range of materials.

3.4. Conclusion

This study proposed a comprehensive methodology using deep neural networks. In the initial phase, various architectures of multilayer perceptron were explored, and the most effective one was selected for an in-depth case study. Subsequently, we introduced a generative adversarial network to create synthetic datasets, facilitating the training of the multilayer perceptron with both real and synthetic data. The application of this methodology was specifically demonstrated in the context of concrete mixtures. The dataset was meticulously gathered from published papers.

For a comprehensive evaluation of the case study, we compared and assessed three distinct scenarios: (a) training the MLP exclusively with the real database, (b) training the MLP with 200 synthetic data followed by real data, and (c) training the MLP with 1000 synthetic data followed by real data. The outcomes across all three scenarios demonstrated satisfactory and highly accurate predictions of concrete thermal conductivity. Notably, including synthetic data in the training process led to a remarkable enhancement in accuracy and a significant reduction in errors within the test dataset, which underscores the effectiveness of the CopulaGAN synthesizer in generating synthetic data and the efficiency of MLP training with a combined dataset. The evaluation of these scenarios utilized the performance metrics RMSE and R^2, yielding the following values for each scenario: scenario (a) – 0.0581 and 0.9858, scenario (b) – 0.0394 and 0.9935, and scenario (c) – 0.0314 and 0.9959, respectively.

This study aimed to develop a methodology for predicting the thermal properties of concrete using artificial neural networks. Our specific focus was to demonstrate the feasibility of forecasting concrete's thermal conductivity based on its constituents' mass composition. Additionally, we sought to show the benefits of integrating a generative adversarial network for data augmentation. The implications of this study can offer insights that can be applied to the development and enhancement of prediction models using deep learning for other materials. One notable impact lies in advancing the field of energy efficiency in buildings, providing decision-makers and manufacturers with advanced knowledge of a building material's thermal properties. While this work has made valuable contributions, it is essential to acknowledge certain limitations. Notably, the study could be enhanced by including other influential input variables, such as temperature, aggregate types, and mineralogy, and by expanding the real dataset to improve predictions across a broader range of thermal conductivity values. Future research studies could explore the application of this methodology to other composite materials and consider the inclusion of additional information, such as mechanical properties. An additional option is to explore the specific heat as an output variable, a fundamental property in transient conditions, complementing the importance of thermal conductivity in stationary scenarios. This way, we can understand how different materials behave and make the predictions even more accurate. This work underscores the importance of training models with sufficiently diverse datasets to ensure accuracy, and it highlights the potential of synthetic data to augment training when faced with a scarcity of real data.

REFERENCES

[1] S. H. Park, J. Chung, M. Souk Yeo, and K. W. Kim, "Evaluation of the thermal performance of a Thermally Activated Building System (TABS) according to the thermal load in a residential building," Energy Build, vol. 73, pp. 69–82, 2014, doi: 10.1016/j.enbuild.2014.01.008.

[2] D. Zhang, Z. Li, J. Zhou, and K. Wu, "Development of thermal energy storage concrete," Cem Concr Res, vol. 34, no. 6, pp. 927–934, Jun. 2004, doi: 10.1016/j.cemconres.2003.10.022.

[3] M. Wu, M. Li, C. Xu, Y. He, and W. Tao, "The impact of concrete structure on the thermal performance of the dual-media thermocline thermal storage tank using concrete as the solid medium," Appl Energy, vol. 113, pp. 1363–1371, 2014, doi: 10.1016/j.apenergy.2013.08.044.

[4] M. A. De Rousseau, J. R. Kasprzyk, and W. V. Srubar, "Computational design optimization of concrete mixtures: A review," Cem Concr Res, vol. 109, no. April, pp. 42–53, 2018, doi: 10.1016/j.cemconres.2018.04.007.

[5] S. Wang et al., “Thermal energy storage in concrete: Review, testing, and simulation of thermal properties at relevant ranges of elevated temperature,” Cem Concr Res, vol. 166, Apr. 2023, doi: 10.1016/j.cemconres.2023.107096.

[6] Y. Sargam, K. Wang, and I. H. Cho, “Machine learning based prediction model for thermal conductivity of concrete,” Journal of Building Engineering, vol. 34, Feb. 2021, doi: 10.1016/j.jobe.2020.101956.

[7] H. S. Ullah, R. A. Khushnood, J. Ahmad, and F. Farooq, “Predictive modelling of sustainable lightweight foamed concrete using machine learning novel approach,” Journal of Building Engineering, vol. 56, Sep. 2022, doi: 10.1016/j.jobe.2022.104746.

[8] A. Behnood and E. M. Golafshani, “Machine learning study of the mechanical properties of concretes containing waste foundry sand,” Constr Build Mater, vol. 243, p. 118152, 2020, doi: 10.1016/j.conbuildmat.2020.118152.

[9] T. Han, A. Siddique, K. Khayat, J. Huang, and A. Kumar, “An ensemble machine learning approach for prediction and optimization of modulus of elasticity of recycled aggregate concrete,” Constr Build Mater, vol. 244, p. 118271, 2020, doi: 10.1016/j.conbuildmat.2020.118271.

[10] V. Mehta, “Machine learning approach for predicting concrete compressive, splitting tensile, and flexural strength with waste foundry sand,” Journal of Building Engineering, vol. 70, Jul. 2023, doi: 10.1016/j.jobe.2023.106363.

[11] T. Ji, T. Lin, and X. Lin, “A concrete mix proportion design algorithm based on artificial neural networks,” Cem Concr Res, vol. 36, no. 7, pp. 1399–1408, Jul. 2006, doi: 10.1016/j.cemconres.2006.01.009.

[12] H. W. Song and S. J. Kwon, “Evaluation of chloride penetration in high performance concrete using neural network algorithm and micro pore structure,” Cem Concr Res, vol. 39, no. 9, pp. 814–824, Sep. 2009, doi: 10.1016/j.cemconres.2009.05.013.

[13] Z. Huo, L. Wang, and Y. Huang, “Predicting carbonation depth of concrete using a hybrid ensemble model,” Journal of Building Engineering, vol. 76, Oct. 2023, doi: 10.1016/j.jobe.2023.107320.

[14] R. Cai et al., “Prediction of surface chloride concentration of marine concrete using ensemble machine learning,” Cem Concr Res, vol. 136, Oct. 2020, doi: 10.1016/j.cemconres.2020.106164.

[15] M. M. Moein et al., “Predictive models for concrete properties using machine learning and deep learning approavhes: A review,” Journal of Building Engineering, vol. 63, 2023, doi: 10.1016/j.jobe.2022.105444.

[16] H. Naderpour, A. H. Rafiean, and P. Fakharian, “Compressive strength prediction of environmentally friendly concrete using artificial neural networks,” Journal of Building Engineering, vol. 16, no. October 2017, pp. 213–219, 2018, doi: 10.1016/j.jobe.2018.01.007.

[17] X. Li, S. Liu, L. Zhao, X. Meng, and Y. Fang, “An integrated building energy performance evaluation method: From parametric modeling to GA-NN based energy consumption prediction modeling,” Journal of Building Engineering, vol. 45, 2022, doi: 10.1016/j.jobe.2021.103571.

[18] A. Al-shawafi, H. Zhu, S. I. Haruna, Z. Bo, S. A. Laqsum, and S. M. borito, “Experimental study and machine learning algorithms for evaluating the performance of U-shaped ultra-high performance reinforced fiber concrete under static and impact loads,” Journal of Building Engineering, vol. 70, Jul. 2023, doi: 10.1016/j.jobe.2023.106389.

[19] Y. Sharifi and M. Hosseinpour, “Compressive strength assessment of concrete containing metakaolin using ANN,” Journal of Rehabilitation in Civil Engineering, vol. 8, no. 4, pp. 15–27, 2020, doi: 10.22075/JRCE.2020.19043.1358.

[20] I. J. Goodfellow et al., “Generative Adversarial Nets,” in Advances in neural information processing systems, Z. Ghahramani, M. Welling, C. Cortes, N. Lawrence, and K. Q. Weinberger, Eds., Neural Information Processing Systems Foundation, Inc. (NeurIPS), 2014, pp. 2672–2680. [Online]. Available: http://www.github.com/goodfeli/adversarial

[21] L. Xu, M. Skoularidou, A. Cuesta-Infante, and K. Veeramachaneni, “Modeling Tabular Data using Conditional GAN,” in Advances in Neural Information Processing Systems 32, 2019. [Online]. Available: https://github.com/DAI-Lab/CTGAN

[22] M. Abedi, L. Hempel, S. Sadeghi, and T. Kirsten, “GAN-Based Approaches for Generating Structured Data in the Medical Domain,” Applied Sciences (Switzerland), vol. 12, no. 14, Jul. 2022, doi: 10.3390/app12147075.

[23] I.-C. Yeh, "Modeling of strength of high-performance concrete using artificial neural networks," Cem Concr Res, vol. 28, no. 12, pp. 1797–1808, Dec. 1998, doi: 10.1016/S0008-8846(98)00165-3.

[24] I.-C. Yeh, "Modeling slump flow of concrete using second-order regressions and artificial neural networks," Cem Concr Compos, vol. 29, no. 6, pp. 474–480, Jul. 2007, doi: 10.1016/j.cemconcomp.2007.02.001.

[25] A. Kandiri, E. Mohammadi Golafshani, and A. Behnood, "Estimation of the compressive strength of concretes containing ground granulated blast furnace slag using hybridized multi-objective ANN and salp swarm algorithm," Constr Build Mater, vol. 248, p. 118676, 2020, doi: 10.1016/j.conbuildmat.2020.118676.

[26] J. Abellán-García, "Four-layer perceptron approach for strength prediction of UHPC," Constr Build Mater, vol. 256, 2020, doi: 10.1016/j.conbuildmat.2020.119465.

[27] J. Abellán García, J. Fernández Gómez, and N. Torres Castellanos, "Properties prediction of environmentally friendly ultra-high-performance concrete using artificial neural networks," European Journal of Environmental and Civil Engineering, vol. 0, no. 0, pp. 1–25, 2020, doi: 10.1080/19648189.2020.1762749.

[28] S. Fidan, H. Oktay, S. Polat, and S. Ozturk, "An Artificial Neural Network Model to Predict the Thermal Properties of Concrete Using Different Neurons and Activation Functions," Advances in Materials Science and Engineering, vol. 2019, 2019, doi: 10.1155/2019/3831813.

[29] M. Kurpińska, L. Kułak, T. Miruszewski, and M. Byczuk, "Application of artificial neural networks to predict insulation properties of lightweight concrete," Applied Sciences (Switzerland), vol. 11, no. 22, Nov. 2021, doi: 10.3390/app112210544.

[30] B. Kursuncu, O. Gencel, O. Y. Bayraktar, J. Shi, M. Nematzadeh, and G. Kaplan, "Optimization of foam concrete characteristics using response surface methodology and artificial neural networks," Constr Build Mater, vol. 337, Jun. 2022, doi: 10.1016/j.conbuildmat.2022.127575.

[31] O. Gencel, F. Koksal, M. Sahin, M. Y. Durgun, H. E. Hagg Lobland, and W. Brostow, "Modeling of thermal conductivity of concrete with vermiculite by using artificial neural Networks approaches," Experimental heat transfer, vol. 26, no. 4, pp. 360–383, 2013, doi: 10.1080/08916152.2012.669810.

[32] J. H. Lee, J. J. Lee, and B. S. Cho, "Effective Prediction of Thermal Conductivity of Concrete Using Neural Network Method," Int J Concr Struct Mater, vol. 6, no. 3, pp. 177–186, Sep. 2012, doi: 10.1007/s40069-012-0016-x.

[33] C. Ozel and A. Topsakal, "Comparison between ANFIS and ANN for estimation of the thermal conductivity coefficients of construction materials," 2015. [Online]. Available: www.scientiairanica.com

[34] S. İpek, G. Özer Yaman, and C. Kılınç, "Investigation of thermal conductivity of rubberized concrete as an energy-efficient building material and modeling by artificial intelligence," Archives of Civil and Mechanical Engineering, vol. 23, no. 3, Aug. 2023, doi: 10.1007/s43452-023-00701-y.

[35] H. Xu et al., "Application of Artificial Neural Networks in Construction Management: A Scientometric Review," Buildings, vol. 12, no. 7. MDPI, Jul. 01, 2022. doi: 10.3390/buildings12070952.

[36] L. Zhu, J. Dai, G. Bai, and F. Zhang, "Study on thermal properties of recycled aggregate concrete and recycled concrete blocks," Constr Build Mater, vol. 94, pp. 620–628, Jul. 2015, doi: 10.1016/j.conbuildmat.2015.07.058.

[37] K.-H. Kim, S.-E. Jeon, J.-K. Kim, and S. Yang, "An experimental study on thermal conductivity of concrete," Cem Concr Res, vol. 33, no. 3, pp. 363–371, 2003, doi: 10.1016/S0008-8846(02)00965-1.

[38] E. Vejmelková, D. Koňáková, T. Kulovaná, A. Hubáček, and R. Černý, "Mechanical and thermal properties of moderate-strength concrete with ceramic powder used as supplementary cementitious material," Adv Mat Res, vol. 1054, pp. 194–198, 2014, doi: 10.4028/www.scientific.net/AMR.1054.194.

[39] R. Demirboga, "Thermal conductivity and compressive strength of concrete incorporation with mineral admixtures," Build Environ, vol. 42, pp. 2467–2471, 2007.

[40] Y. Sargam, K. Wang, and J. E. Alleman, "Effects of Modern Concrete Materials on Thermal Conductivity," Journal of materials in civil engineering, vol. 32, no. 4, 2020, doi: 10.1061/(ASCE)MT.1943-5533.0003026.

[41] Y. Sargam, B. M. Shankaramurthy, and K. Wang, "Characterization of RCAs and their concrete using simple test methods," J Sustain Cem Based Mater, vol. 9, no. 2, pp. 61–77, 2019, doi: 10.1080/21650373.2019.1692093.

[42] Y. Sargam, M. Faytarouni, K. Riding, K. Wang, C. Jahren, and J. Shen, "Predicting thermal performance of a mass concrete foundation – a field

monitoring case study," Case studies in construction materials, vol. 11, 2019, doi: 10.1016/j.cscm.2019.e00289.

[43] T. L. Cavalline, R. W. Castrodale, C. Freeman, and J. Wall, "Impact of lightweight aggregate on concrete thermal properties," ACI Mater J, vol. 114, no. 6, pp. 945–956, Nov. 2017, doi: 10.14359/51701003.

[44] M. Kurpińska, J. Karwacki, A. Maurin, and M. Kin, "Measurements of thermal conductivity of LWC cement composites using simplified laboratory scale method," Materials, vol. 14, no. 6, Mar. 2021, doi: 10.3390/ma14061351.

[45] O. Yavuz Bayraktar, G. Kaplan, O. Gencel, A. Benli, and M. Sutcu, "Physico-mechanical, durability and thermal properties of basalt fiber reinforced foamed concrete containing waste marble powder and slag," Constr Build Mater, vol. 288, Jun. 2021, doi: 10.1016/j.conbuildmat.2021.123128.

CHAPTER IV

Data-Augmented Deep Learning Models for Assessing Thermal Performance in Sustainable Building Materials

Ana Carolina Rosa [1,2], Carles Mateu [3], Assed Haddad [1], Dieter Boer [2]

[1] Programa de Engenharia Ambiental, Universidade Federal do Rio de Janeiro, Rio de Janeiro, Brazil

[2] Department of Mechanical Engineering, University Rovira i Virgili, Tarragona, Spain.

[3] Department of Computer Science and Industrial Engineering, University of Lleida, Lleida, Spain

Keywords: Energy Efficiency, Phase Change Materials, Deep Neural Networks, Multilayer Perceptron, Data Augmentation, Thermal conductivity.

INTRODUCTION

Energy efficiency in buildings plays a fundamental role in the rational and effective use of energy consumption to meet human needs without compromising the environment. One of the major measures used to reduce energy consumption is the use of more energy-efficient materials, thus preserving thermal comfort despite temperature variations in the outside environment [1].

Concrete is one of the most used materials in construction due to its unique features, such as high compressive strength, heat capacity, and low cost. Depending on its composition, a more sustainable concrete mix with better thermal performance can be achieved. Therefore, any variations in the constituents and the properties of the concrete can undergo significant variations. A vast number of construction materials have been added to concrete to make it more sustainable. One of the materials that is gaining prominence is phase change material (PCM), which can store and release thermal energy as it changes from one phase to another. This characteristic makes them extremely useful in various applications, especially thermal management and energy storage systems.

Research on advanced energy storage materials is on the rise, as evidenced by the growing number of publications, indicating a heightened focus in this area [2]. The integration of PCMs into building materials like cement or concrete enhances their thermal properties, leading to reduced energy consumption [3]. For instance, experimental findings by Rajesh et al. [4] demonstrated a notable 20% decrease in energy consumption when testing PCM mortar specimens. The utilization of PCM for improving the thermal energy storage of building envelopes is widespread [5]. Qu et al. [6] utilized EnergyPlus software to analyze the incorporation of PCMs in building envelopes, aiming to achieve thermal comfort and energy savings. Additionally, Guo and Zhang [7] explored the performance of PCM-packed building envelopes against seasonal fluctuations. Haridass et al. [8] conducted both experimental and numerical investigations comparing the thermal behavior of buildings with bare and PCM-integrated RCC roofs, revealing a significant reduction in heat transfer into the building with PCM integration.

Due to the significance of this issue and the pressing need for energy-efficient building materials, numerous studies have turned their attention to the analysis of thermal properties in construction materials. Consequently, the development of a predictive model capable of accurately estimating their thermal behavior emerges as a valuable tool for optimizing their application. While the analysis of thermal properties, like thermal conductivity, is crucial, another fundamental property, compressive strength, profoundly impacts construction materials' performance. Compressive strength is vital to assess a material's ability to withstand axial loads and to ensure structural integrity in construction applications. To be effective as a thermal energy storage material and durable, it is crucial to have an ideal composition to ensure that its thermal and mechanical properties meet the required design specifications [9].
Machine learning (ML) and deep learning (DL) models can be used to develop such predictive models by learning from large datasets and finding patterns to make predictions. By incorporating these models, the predictive models can adapt to varying conditions and continuously improve their accuracy, making them a valuable tool in optimizing the performance of building materials [10]. In recent years, ML and DL models have been used to predict some properties of concrete, such as compressive strength [11], tensile strength [12], modulus of elasticity [13], flexural strength [14], slump [15], chloride penetration [16], carbonation depth [17] and surface chloride concentration [18]. Although many researchers used ML, there is a limitation in studies developing prediction models of the thermal properties of concrete [19].

Among the models, the Artificial Neural Network (ANN) is one of the most employed ones to solve complex problems and has various applications in several fields [20,21]. The importance of ANNs lies in their ability to learn and make decisions based on data, which makes them highly valuable. In this way, ANNs can be used to solve problems that conventional or other computational methods have difficulties [22]. One of the drawbacks typically encountered when developing an ANN is a limited data set, and effectively training an ANN requires a massive amount of data. When this problem arises, an alternative that has been emerging is the use of the Generative Adversarial Network (GAN).
GANs were introduced by Goodfellow et al. [23] and became a revolutionary development in the world of generative modeling. They have different applications, such as image generation, super-resolution imaging, style transfer, and data augmentation. Regarding tabular data, it can create synthetic data to expand real datasets and prevent over-fitting in such data-limited situations, helping to improve the training of an ML model. Although GAN started with image generation, some authors have already used this technique to create tabular data and obtain satisfactory results. Since the introduction of GAN, several algorithms to model tabular data have been used, such as Conditional Tabular GAN (CTGAN) [24], TabGAN, and CopulaGAN [25]. Although many researchers have attempted to use ANN to predict different properties of concrete, only a few works are progressing on models to determine the thermal properties and mechanical properties.

Before collecting the data for this study, a literature review was carried out to find gaps and trends in the development of property prediction. This analysis revealed a lack of data in the literature and the opportunity to develop deep neural network models with the integration of the data augmentation technique. This work intends to fill this gap and the novelty is the development of a predictive model encompassing augmented data to train the model. For this methodology, a Multilayer perceptron (MLP) model to predict thermal conductivity and compressive strength, and a Copula GAN model to improve the tabular data from published papers will be developed. This research seeks to advance the field of energy-efficient building materials while ensuring the structural resilience of construction materials, thus contributing significantly to sustainable construction practices.

METHODS

The proposed method intends to extend the method developed in a previous work and involves two fundamental steps: firstly, the development of an MLP model tailored to predict material properties based on their distinctive features; and secondly, the utilization of a GAN for data augmentation. This augmentation process enriches the dataset, enhancing the accuracy of the initial model. This section elucidates the method adopted in this study to accomplish these objectives. Additionally, to validate the effectiveness of the model, a case study is conducted and evaluated (Figure 1).

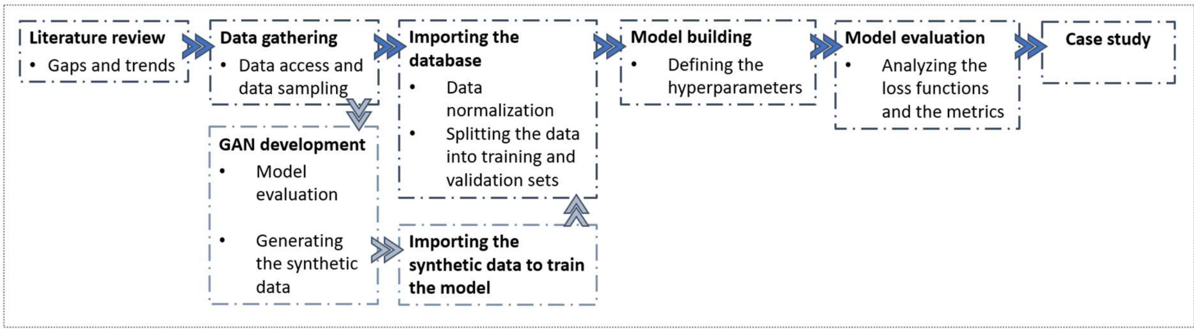

Figure 1. The method flowchart.

After the synthetic data generation, we aim to incorporate it into the MLP model for joint training with real and artificial datasets. To assess the model's response to this diverse training approach, a distinct set of synthetic data is created, comprising 1000 data points. The integration of these two neural network models and the subsequent validation of the model using a new test set are elaborated upon in the forthcoming section.

CASE STUDY

The proposed model is applied to a case study on concrete mixtures enriched with PCM and nano-silica aerogel. As previously highlighted in the introduction, thermal conductivity, and compressive strength are important features and can significantly enhance the thermal performance and resistance of a building. Given that these properties can be influenced by the type and quantity of each constituent, this study seeks to predict them based on the mass composition of the constituents.

To simulate a lack of data, this study gathered 33 experimental data with different compositions of PCM and nano-silica aerogel. Here, a total of 7 input parameters that affect thermal conductivity (TC) and compressive strength (CS) are considered (Figure 2). The model's inputs are density (D), cement content (C), sand content (S), aerogel aggregate content (AAg), PCM aggregate content (PAg), water content (W), and superplasticizer content (Sp). The values of TC lie in the range of 0.19 – 1.80 W/m.K with a standard deviation of 0.4 W/m.K and the values of CS lie in the range of 1.4 – 42.0 W/m.K with a standard deviation of 11.5 W/m.K.

In order to train the model, the compiled dataset is randomly divided into training (80%) and validation sub-dataset (20%). Then, the reliability and reproducibility of the ANN model are evaluated on a new testing dataset. The predictive performance of the final model is thoroughly assessed using the metrics RMSE and R^2. This assessment allows us to check the model's accuracy and efficiency in capturing the underlying patterns within the database.

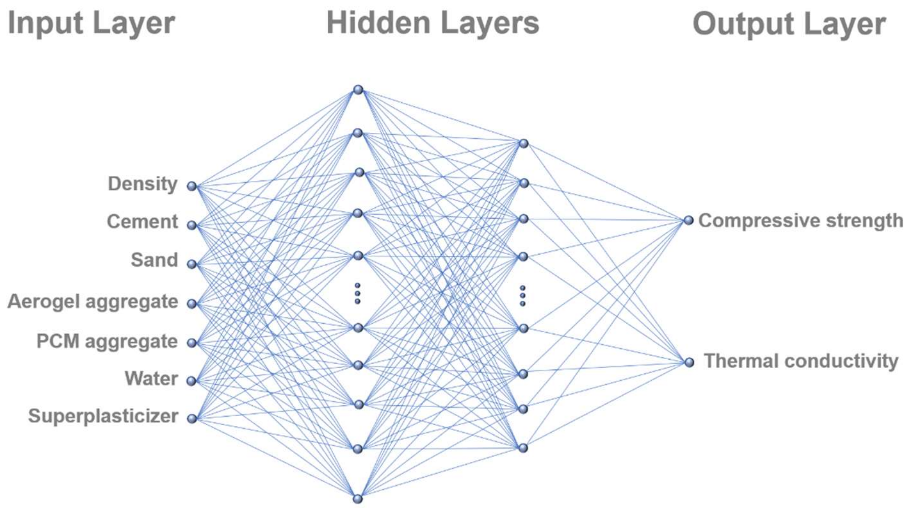

Figure 2. The multilayer perceptron model

RESULTS AND DISCUSSION

The first step of this work is to develop an MLP model for predicting the compressive strength and the thermal conductivity of concrete based on the constituents' mass composition and density of concrete. The hyperparameters of the MLP model are depicted in Table 1, which provides an overview of the model's key features, encompassing the hyperparameters, evaluation metrics, activation functions, optimization methods, and loss functions.

Table 1. The main features of the MLP model

Hyperparameters	
	Hidden layers = 2
	Neurons = [200,100]
	Batch size = 64
	Learning rate = 0.001
	Epochs = 100
Metrics	
	RMSE
	R^2
Activation function	
	ReLU
Optimization function	
	Stochastic Gradient Descent (SGD)
Loss function	
	MSE

The second step of this study is dedicated to evaluating the impact of synthetic data on the training of the MLP model. This analysis is conducted to determine whether a model trained with augmented data can produce results comparable to the performance achieved through training with real data and possibly enhance it. This study investigates two distinct scenarios: (a) training the MLP model exclusively with the real dataset containing 33 entries, and (b) training the model with the synthetic data consisting of 1000 entries first and then the real dataset.

The same hyperparameters are used in both scenarios. The only distinction lies in the dataset used for the training and validation phases. Our approach in scenario (b) involves training the model first with synthetic data. Subsequently, the same model architecture uses the real dataset to enhance the model's performance. In the training phase, the overall metrics achieved an RMSE of 0.0700 and an R^2 of 0.9940 for the first scenario, and an RMSE of 0.0530 and an R^2 of 0.9971 for the second scenario. The performance of the final model on the validation and test datasets for both scenarios (a and b) is summarized in Table 2, presenting the key outcomes.

Table 2. Comparison between the performance metrics of the MLP model trained with both datasets

Scenario	Dataset	CS		TC	
		RMSE	R^2	RMSE	R^2
(a)	Validation	0.6329	0.9970	0.0226	0.9964
	Test	1.2728	0.9922	0.0381	0.9921
(b)	Validation	0.5669	0.9975	0.0203	0.9971
	Test	0.7239	0.9975	0.0149	0.9988

Although scenario (a) had a limited dataset size, the training, validation, and test datasets still produced acceptable results for compressive strength and thermal conductivity. However, in scenario (b), where synthetic data plays a vital role, both predictions exhibited superior performance. This enhancement in results is consistently observed across the validation and test datasets as well.

The MLP model's predictive performance for the two scenarios is assessed using an independent test dataset consisting of 8 data points. Figure 3(a) e Figure 3(b) demonstrate the performance and predictions of the compressive strength for the test dataset, while Figure 4(a) and Figure 4(b) display the performance of the thermal conductivity. In Figure 3(a), the compressive strength predictions and actual values reveal some alignment with the ideal x=y line, alongside noticeable deviations with errors exceeding 0.18. Figure 3(b) illustrates the scenario where the model is trained with synthetic data. Here, the integration of synthetic data results in a notable improvement in prediction accuracy and a significant reduction in error. This scenario achieves an R^2 value of 0.9975 and an RMSE of 0.7239.

Comparing Figure 4(a) and Figure 4(b) for thermal conductivity prediction, a similar trend to that observed in the compressive strength prediction becomes apparent. However, noteworthy differences exist in the achieved errors between the two scenarios. In scenario (a), errors reach up to 0.05, whereas scenario (b) demonstrates a remarkable 50% reduction in error values, with a maximum of 0.025. Scenario (b) also boasts an R^2 of 0.9921 and an RMSE of 0.0381.

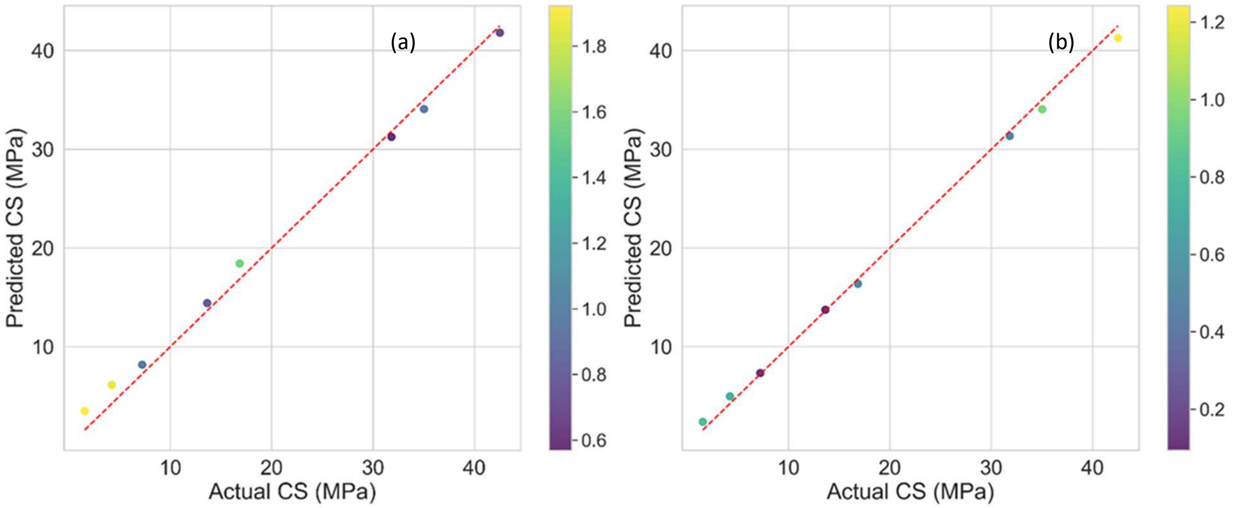

Figure 3. Performance and predictions of the compressive strength of the test dataset for the two scenarios: (a) training the model with real data only, (b) training the model with 1000 synthetic data and real data

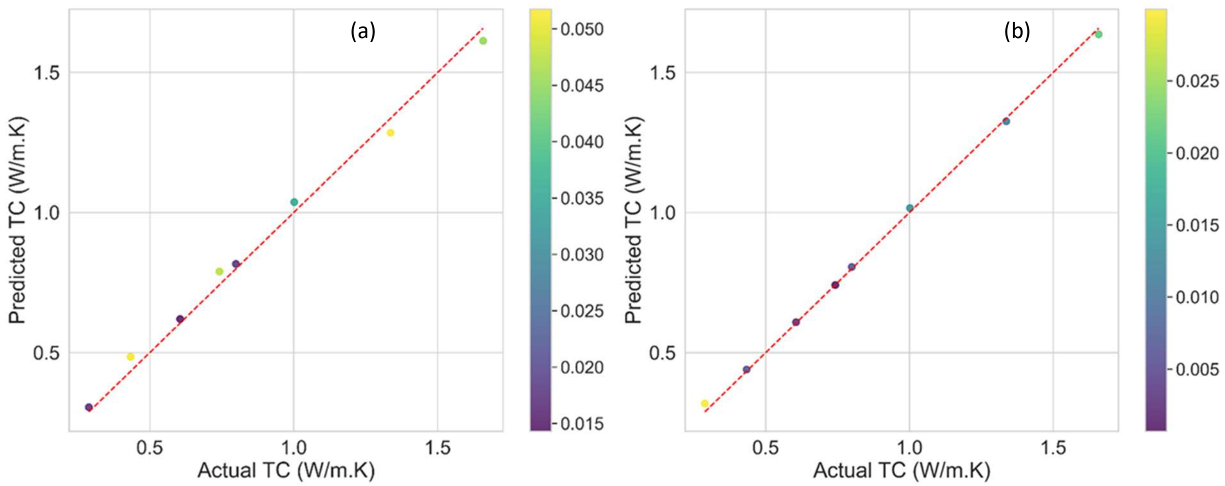

Figure 4. Performance and predictions of the thermal conductivity of the test dataset for the two scenarios: (a) training the model with real data only, (b) training the model with 1000 synthetic data and real data

Figure 5 and Figure 6 show the relationship between the actual values derived from the test dataset and the corresponding predictions generated by the Multilayer Perceptron (MLP) model. This graphical representation allows us to assess the model's accuracy in forecasting compressive strength and thermal conductivity, respectively, providing valuable insights into its performance.

In Figure 5(a) and Figure 6(a), we depict the model utilizing solely the real dataset, while Figure 5(b) and Figure 6(b) display the model trained with the 1000-point dataset and the real dataset, respectively.

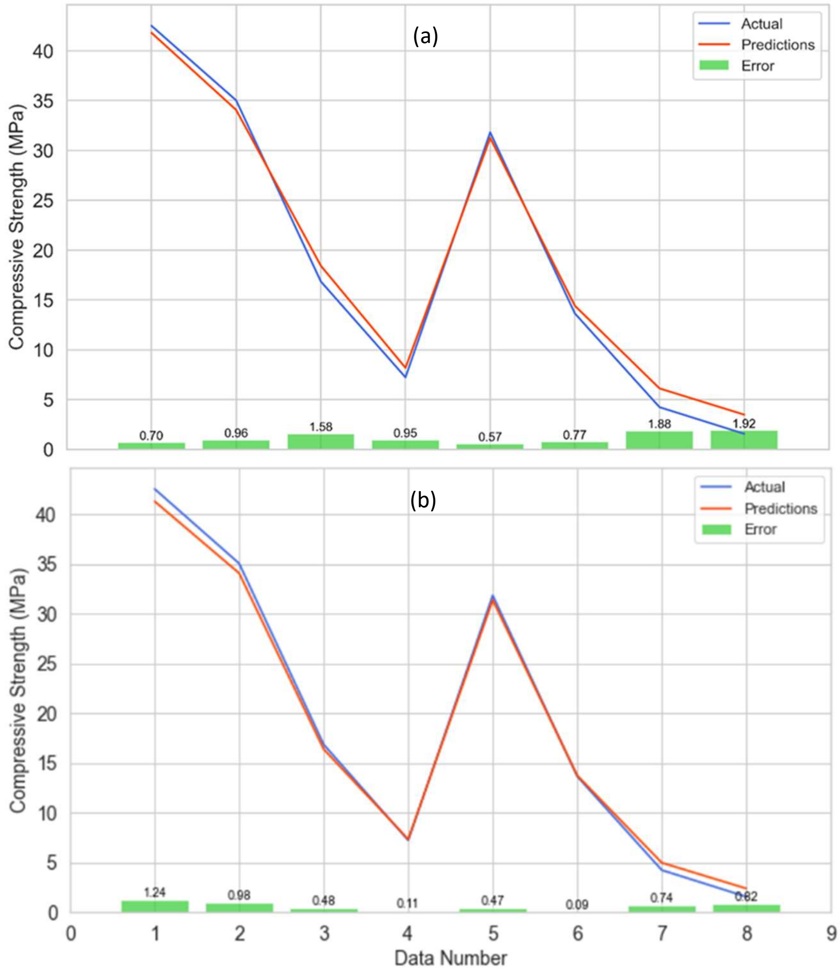

Figure 5. Predictions of the compressive strength of the test dataset for the three scenarios: (a) training the model with real data only, (b) training the model with 1000 synthetic data and real data

Regarding the compressive strength results, Figure 5(a) indicates a close alignment between predicted and real points. However, some discrepancies across various data points with errors varying between 0.57 and 1.92 MPa are observed. Moving to Figure 5(b), a significant reduction in these discrepancies becomes evident, especially in values ranging from 7MPa to 31MPa range.

The thermal conductivity results show even more promising outcomes. Figure 6(a) illustrates the proximity of predicted values to real values, with only half of the points exhibiting errors of 0.05 W/m.K. Meanwhile, Figure 6(b) underscores the high accuracy of the synthetic data utilized in training the network, with just two points with errors of 0.02 and 0.03 W/m.K.

A comparison of actual and predicted values reveals a significant enhancement in model performance when synthetic data is integrated. The model's predictions demonstrate a heightened level of precision and a further reduction in errors, affirming the model's accuracy and the benefits of incorporating synthetic data into the training process.

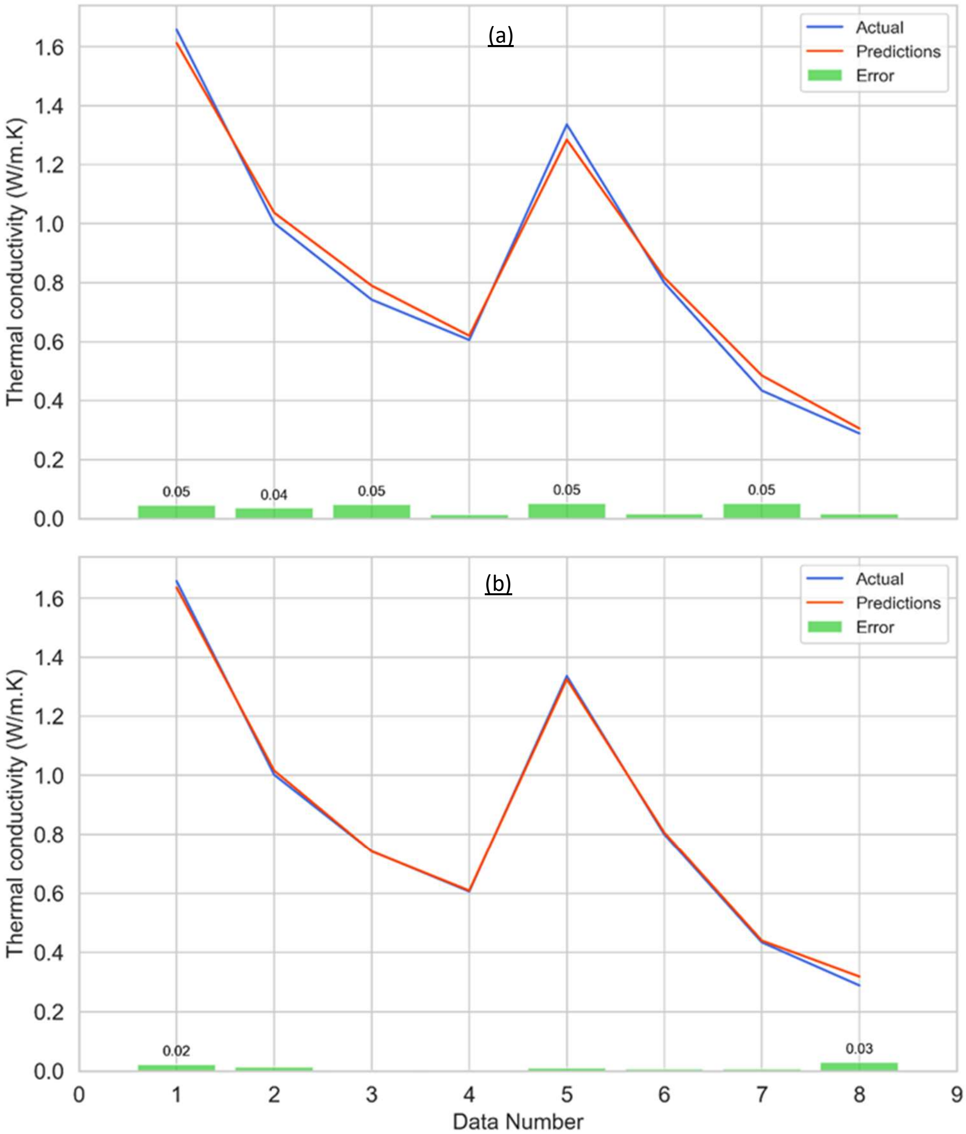

Figure 6. Predictions of the thermal conductivity of the test dataset for the three scenarios: (a) training the model with real data only, (b) training the model with 1000 synthetic data and real data

CONCLUSION

This study proposed a methodology using data augmentation to train a multilayer perceptron model. We introduced a generative adversarial network to create synthetic datasets, helping the training of the MLP with both real and synthetic data. The application of this methodology was specifically demonstrated in the context of a small dataset of concrete mixtures using PCM and nano-silica aerogel.

We compared and assessed two distinct scenarios: (a) training the MLP exclusively with the real database, and (b) training the MLP with 1000 synthetic data followed by real data. The outcomes of these two scenarios demonstrated satisfactory and highly accurate predictions of compressive strength and thermal conductivity. Notably, including synthetic data in the training process led to an enhancement in accuracy and a significant reduction in errors within the test dataset, which underscores the effectiveness of the CopulaGAN synthesizer in generating synthetic data and the efficiency of MLP training with a combined dataset. The evaluation of these scenarios relied on the performance metrics RMSE and R^2, resulting in the following values for each scenario: In scenario (a), the metrics for compressive strength were 1.2728 (RMSE) and 0.9922 (R^2), while for thermal conductivity, they were 0.0381 (RMSE) and 0.9921 (R^2). In scenario (b), the metrics for compressive strength were 0.7239 (RMSE) and 0.9975 (R^2), and for thermal conductivity, they were 0.0149 (RMSE) and 0.9988 (R^2).

This study aimed to expand a methodology for predicting the thermal properties of concrete using artificial neural networks. Our specific focus was to demonstrate the feasibility of data augmentation to forecast concrete's compressive strength and thermal conductivity based on the mass composition of its constituents. As previously mentioned, this model builds upon an earlier iteration developed by the authors. Through this study, we have demonstrated the remarkable flexibility and generalization of the MLP model. It seamlessly adapts to variations in the dataset, accommodating changes in inputs and incorporating compressive strength as an additional output. Furthermore, our model surpasses alternative models, as evidenced by the significant reduction in errors for both outputs. An additional advantage of the MLP model lies in its compatibility with synthetic data, addressing a common limitation associated with many models, which is the requirement for extensive datasets. This feature enhances the model's scalability and applicability, making it an attractive choice for diverse research and practical applications. The implications of this study can offer insights that can be applied to the development and enhancement of prediction models using deep learning for other energy-efficient materials.

While this work has made valuable contributions, it is essential to acknowledge certain limitations. Notably, the study could be enhanced by including other influential input variables, such as temperature, aggregate types, and mineralogy, and by expanding the real dataset to improve predictions across a broader range of compressive strength and thermal conductivity values. Future research studies could explore the application of this methodology to other composite materials and consider the inclusion of additional information as inputs and outputs. This study pointed out the significance of training models with diverse datasets and highlighted the potential of synthetic data to enhance training, particularly in situations where real data may be limited.

REFERENCES

[1] Park SH, Chung J, Souk Yeo M, Kim KW. Evaluation of the thermal performance of a Thermally Activated Building System (TABS) according to the thermal load in a residential building. Energy Build 2014;73:69–82. https://doi.org/10.1016/j.enbuild.2014.01.008.

[2] Hassan F, Jamil F, Hussain A, Ali HM, Janjua MM, Khushnood S, et al. Recent advancements in latent heat phase change materials and their applications for thermal energy storage and buildings: A state of the art review. Sustainable Energy Technologies and Assessments 2022;49. https://doi.org/10.1016/j.seta.2021.101646.

[3] Mustafa J, Alqaed S, Sharifpur M. Loading phase change material in a concrete based wall to enhance concrete thermal properties. Journal of Building Engineering 2022;56. https://doi.org/10.1016/j.jobe.2022.104765.

[4] Rajesh B, Mekala C. Experimental investigation of influence of phase change materials in energy consumption of air-conditioning units and prediction performance evaluation of modified deep neural network model. Constr Build Mater 2023;389. https://doi.org/10.1016/j.conbuildmat.2023.131582.

[5] Patel B, Rathore PKS, Gupta NK, Sikarwar BS, Sharma RK, Kumar R, et al. Location optimization of phase change material for thermal energy storage in concrete block for development of energy efficient buildings. Renew Energy 2023;218. https://doi.org/10.1016/j.renene.2023.119306.

[6] Qu Y, Zhou D, Xue F, Cui L. Multi-factor analysis on thermal comfort and energy saving potential for PCM-integrated buildings in summer. Energy Build 2021;241. https://doi.org/10.1016/j.enbuild.2021.110966.

[7] Guo J, Zhang G. Investigating the performance of the PCM-integrated building envelope on a seasonal basis. J Taiwan Inst Chem Eng 2021;124:91–7. https://doi.org/10.1016/j.jtice.2021.04.066.

[8] Haridass K, N. N, Rajkumar S. Experimental and numerical investigation on phase change material filled reinforced cement concrete roof slab for mitigating the heat transfer. Energy Sources, Part A: Recovery, Utilization, and Environmental Effects 2022;44:3850–65. https://doi.org/10.1080/15567036.2022.2068700.

[9] Nwe Htun N, Sukchai S, Hemavibool S. Properties of concrete material for thermal energy storage. In: Htun NN, Suckchai S, Hemavibool S, editors. International Conference and Utility Exhibition 2014 on Green Energy for Sustainable Development (ICUE 2014), Pattaya, Thailand: IEEE; 2014, p. 1–5.

[10] He Z, Guo W, Zhang P. Performance prediction, optimal design and operational control of thermal energy storage using artificial intelligence methods. Renewable and Sustainable Energy Reviews 2022;156:111977. https://doi.org/10.1016/j.rser.2021.111977.

[11] Ullah HS, Khushnood RA, Ahmad J, Farooq F. Predictive modelling of sustainable lightweight foamed concrete using machine learning novel approach. Journal of Building Engineering 2022;56. https://doi.org/10.1016/j.jobe.2022.104746.

[12] Behnood A, Golafshani EM. Machine learning study of the mechanical properties of concretes containing waste foundry sand. Constr Build Mater 2020;243:118152. https://doi.org/10.1016/j.conbuildmat.2020.118152.

[13] Han T, Siddique A, Khayat K, Huang J, Kumar A. An ensemble machine learning approach for prediction and optimization of modulus of elasticity of recycled aggregate concrete. Constr Build Mater 2020;244:118271. https://doi.org/10.1016/j.conbuildmat.2020.118271.

[14] Mehta V. Machine learning approach for predicting concrete compressive, splitting tensile, and flexural strength with waste foundry sand. Journal of Building Engineering 2023;70. https://doi.org/10.1016/j.jobe.2023.106363.

[15] Ji T, Lin T, Lin X. A concrete mix proportion design algorithm based on artificial neural networks. Cem Concr Res 2006;36:1399–408. https://doi.org/10.1016/j.cemconres.2006.01.009.

[16] Song HW, Kwon SJ. Evaluation of chloride penetration in high performance concrete using neural network algorithm and micro pore structure. Cem Concr Res 2009;39:814–24. https://doi.org/10.1016/j.cemconres.2009.05.013.

[17] Huo Z, Wang L, Huang Y. Predicting carbonation depth of concrete using a hybrid ensemble model. Journal of Building Engineering 2023;76. https://doi.org/10.1016/j.jobe.2023.107320.

[18] Cai R, Han T, Liao W, Huang J, Li D, Kumar A, et al. Prediction of surface chloride concentration of marine concrete using ensemble machine learning. Cem Concr Res 2020;136. https://doi.org/10.1016/j.cemconres.2020.106164.

[19] Moein MM, Saradar A, Rahmati K, Mousavinejad SHG, Bristow J, Aramali V, et al. Predictive models for concrete properties using machine learning and deep learning approavhes: A review. Journal of Building Engineering 2023;63. https://doi.org/10.1016/j.jobe.2022.105444.

[20] Deng F, He Y, Zhou S, Yu Y, Cheng H, Wu X. Compressive strength prediction of recycled concrete based on deep learning. Constr Build Mater 2018;175:562–9. https://doi.org/10.1016/j.conbuildmat.2018.04.169.

[21] Naderpour H, Rafiean AH, Fakharian P. Compressive strength prediction of environmentally friendly concrete using artificial neural networks. Journal of Building Engineering 2018;16:213–9. https://doi.org/10.1016/j.jobe.2018.01.007.

[22] Sharifi Y, Hosseinpour M. Compressive strength assessment of concrete containing metakaolin using ANN. Journal of Rehabilitation in Civil Engineering 2020;8:15–27. https://doi.org/10.22075/JRCE.2020.19043.1358.

[23] Goodfellow IJ, Pouget-Abadie J, Mirza M, Xu B, Warde-Farley D, Ozair S, et al. Generative Adversarial Nets. In: Ghahramani Z, Welling M, Cortes C, Lawrence N, Weinberger KQ, editors. Adv Neural Inf Process Syst, vol. 27, Neural Information Processing Systems Foundation, Inc. (NeurIPS); 2014, p. 2672–80.

[24] Xu L, Skoularidou M, Cuesta-Infante A, Veeramachaneni K. Modeling Tabular Data using Conditional GAN. Advances in Neural Information Processing Systems 32, 2019.

[25] Abedi M, Hempel L, Sadeghi S, Kirsten T. GAN-Based Approaches for Generating Structured Data in the Medical Domain. Applied Sciences (Switzerland) 2022;12. https://doi.org/10.3390/app12147075.

CHAPTER V

Enhanced Deep Learning Models with Data Augmentation for Ceramic Brick Property Assessment

Ana Carolina Rosa [1,2], Assed Haddad [1], Dieter Boer [2]

[1] Programa de Engenharia Ambiental, Universidade Federal do Rio de Janeiro, Rio de Janeiro, Brazil

[2] Department of Mechanical Engineering, University Rovira i Virgili, Tarragona, Spain.

Keywords: xxxx

INTRODUCTION

V. APPENDIX

5.1. List of publications

5.1.1. Research articles

5.1.2. Oral communications

5.2. Supplementary material